*Journal of Experimental and Theoretical Physics, Vol. 99, No. 4, 2004, pp. 669–689.*
*Translated from Zhurnal Éksperimental'noĭ i Teoreticheskoĭ Fiziki, Vol. 126, No. 4, 2004, pp. 771–798.*


**ATOMS, SPECTRA, RADIATION**

# Free-Space Quantum Cryptography Using Multiphoton States: Secure Key Distribution to Satellites

**S. N. Molotkov**

*Institute of Solid State Physics, Russian Academy of Sciences, Chernogolovka, Moscow oblast, 142432 Russia*
*Moscow State University, Moscow, 119991 Russia*
*e-mail: molotkov@issp.ac.ru*

Received March 3, 2004

**Abstract**—Whereas quantum cryptography ensures security by virtue of complete indistinguishability of nonorthogonal quantum states, attenuation in quantum communication channels and the unavailability of single-photon sources present major problems. In view of these difficulties, the security of quantum cryptography can change from unconditional to conditional. Since the restrictions imposed by nonrelativistic quantum mechanics and used to formulate key distribution protocols have been largely exhausted, new principles are required. The fundamental relativistic causality principle in quantum cryptography can be used to propose a new approach to ensuring unconditional security of quantum cryptosystems that eliminates the aforementioned difficulties. Quantum cryptosystems of this kind should obviously be called relativistic. It is shown that relativistic quantum cryptosystems remain unconditionally secure: first, attenuation in a quantum communication channel can only reduce the key generation rate, but not the security of the key; second, the source may not generate pure single-photon states, and a nonzero single-photon probability will suffice. The scheme remains secure even if the contribution of a single-photon component is arbitrarily small. This formally implies that a state may be characterized by an arbitrarily large mean photon number. The single-photon probability affects only the key generation rate, but not security. © *2004 MAIK "Nauka/Interperiodica".*

## 1. INTRODUCTION

Quantum cryptography, i.e., secure key distribution over public quantum communication channels, may be the only method for implementing absolutely secure "one-time pad" cryptography systems [1–3]. To date, several different fiber-based prototype quantum cryptosystems have been constructed. A Japanese group reported in [4] that the maximum range of secure key distribution in a quantum cryptosystem reached with self-compensation by means of Faraday mirrors is 100 km. The previous record, 67 km, was set by a Swiss group [5]. Recently, MagiQ presented a commercial quantum cryptography system with a fiber length of 120 km. Most existing prototype quantum cryptosystems rely on the following coding principles:

(i) in polarization coding, information about the key is encoded into polarization states [6];

(ii) in phase coding, an unbalanced Mach–Zehnder interferometer is employed and information is encoded into the phase difference between the input and output arms of the interferometer [7, 8];

(iii) in frequency coding, frequency modulation is applied to the carrier frequency [9];

(iv) in quantum cryptography using coherent states, homodyne detection is performed at the receiver end [10].

The best results have been achieved in the development of fiber-based cryptosystems with phase coding and self-compensation by means of Faraday mirrors [4, 5, 11].

Recently, the first local quantum cryptography network for key exchange between users within a range of 10 km was tested in Boston (the project is supported by DARPA) [12].

Several prototype quantum cryptosystems for secure key distribution over free space have been created [13, 14]. The longest range (documented in publications) is 23.4 km under both daylight and nighttime conditions. The purpose of all free-space quantum cryptosystems is secure key exchange between ground-based stations and low-orbit satellites (at altitudes of up to 1000 km) or between ground-based stations via satellites. According to estimates made by specialists from QinetiQ, a company developing surface-to-satellite communication systems, this can be done in the very near future by analyzing quantum cryptography schemes, because a sufficiently high level of technology development has been reached and the projected costs are reasonable [15]. Moreover, key distribution via satellites potentially has a much wider scope. Experiments in this line of research are being planned for the next year.

All schemes mentioned above, except for those based on homodyne detection, rely on single-photon detection. However, since strictly single-photon sources are not available to this day (some progress in construction of their laboratory prototypes was

reported in [16]), the pseudo-single-photon sources currently employed make use of laser pulses with mean photon numbers reduced to $\mu = 0.1$–$0.3$ photons per pulse. In this scheme, a communication channel may carry more than one photon, because the coherent state corresponding to a laser pulse is characterized only by the mean photon number. The actual attenuated laser radiation in a communication channel corresponds to a mixture of states with different photon numbers. This implies that a quantum cryptosystem may be insecure against attacks of certain kinds (e.g., see [17, 18]). Note also that the key generation rate substantially decreases when the mean photon number is low (when it is as low as 0.1, about 90% of pulses are empty).

Attenuation in a quantum communication channel presents an essential in terms of security. While attenuation obviously reduces the key transmission rate and some photons may not reach the receiver, the basic problem is that security cannot be guaranteed if the attenuation exceeds a certain value (see [18] and discussion below). Attenuation in a fiber-optic cable depends on its length. The critical length within which the system guarantees security is not known to this day, and its estimated values vary from tens of kilometers to 150 km [18].

An analysis of the basic quantum cryptographic protocols (BB84 and B92 and their derivatives) shows that the proofs of their security for a channel with attenuation essentially rely (explicitly or implicitly) on some prior information about the quantum bit error rate (QBER). For example, if the attenuation in the channel varies during the execution time for a key distribution protocol, then so does the error rate (even in the absence of an eavesdropper). In this situation, no security can be guaranteed if a constant QBER is assumed in the protocol. Whereas the attenuation can be treated as constant for fiber-based quantum cryptosystems (0.17–0.25 dB/km for a 1550-nm single-mode fiber), a similar assumption is not valid in the case of a free-space link, because atmospheric conditions cannot be controlled. Therefore, key-distribution protocols should be stable and guarantee security for a channel with attenuation that either varies during the execution time for a key distribution protocol or is known a priori. This serious problem must be solved in order to rule out any doubt concerning unconditional security in quantum cryptography (i.e., security that is guaranteed by fundamental limitations dictated by the laws of quantum mechanics rather than by the eavesdropper's technical restrictions).

The difficulties outlined above are due to the fact that the security of a protocol relies only on the geometric properties of quantum state vectors in a Hilbert space $\mathcal{H}$. More precisely, it is based on the no-cloning theorem for an unknown quantum state [19] and on Bennett's theorem on complete indistinguishability of nonorthogonal quantum states [20]. Roughly speaking, protocols are formulated in a Hilbert space $\mathcal{H}$, and the fact that quantum-state measurements and propagation take place in spacetime has never been used explicitly. Since attenuation takes place in spacetime rather than in a Hilbert space, additional fundamental limitations due to properties of propagating quantum states are required to rule out any insecurity due to attenuation, and spacetime information about them must be available. Limitations on quantum cryptographic protocols dictated only by geometric properties of quantum states in Hilbert spaces seem to be exhausted. Additional fundamental limitations should be found in a special theory of relativity. Note also that, since photons are truly relativistic massless particles (field quanta) propagating with the highest possible speed, it would be unreasonable not to use additional opportunities offered by nature.

Thus, quantum cryptosystems may not be unconditionally secure because of the lack of single-photon sources and high attenuation.

In what follows, several free-space quantum cryptosystems are proposed in which limitations on measurability dictated both by quantum mechanics and by special theory of relativity are employed. Since the quantum cryptosystems discussed below allow for quantum-state (key) propagation in spacetime, knowledge of the channel length is required. Relativistic quantum cryptosystems remain secure irrespective of attenuation in the communication channel: attenuation reduces the key transmission rate without affecting security. Moreover, security of a key is guaranteed even for multiphoton states. Formally, the scheme is effective for a quantum state characterized by an arbitrary mean photon number. The highest efficiency is reached for $\mu$ between 1 and 3, when empty pulses are almost absent (the contribution of vacuum to the coherent state is low). This means that the key generation rate is at least an order of magnitude higher than that in schemes relying only on geometric properties of quantum states and requiring faint laser pulses with $\mu = 0.1$–$0.3$. The rate is additionally increased because even orthogonal states can be used under the limitations imposed by special relativity theory, in which case no bases-reconciliation procedure (indispensable in the BB84 protocol) is required. Furthermore, since both legitimate users and the eavesdropper act in spacetime, collective attacks performed by the eavesdropper are not advantageous as compared to individual measurements of each pulse when orthogonal states are used. Finally, the system guarantees security even if the error rate in the adopted binary string is higher than 40% (when $\mu \approx 1$). For example, it should be recalled that security of the BB84 protocol is guaranteed only if the error rate is lower than 11% [21, 22].

The only additional requirement in relativistic quantum cryptography, as compared to nonrelativistic quantum cryptosystems using nonorthogonal states, is knowledge of the channel length. This appears to be a modest price to pay for the advantages gained in the relativistic case.

Several realizations of quantum cryptosystems are discussed below.

## 2. PHYSICS OF QUANTUM CRYPTOGRAPHY

In quantum cryptography, security is based on fundamental limitations imposed by the laws of quantum mechanics, rather than on any assumption about the eavesdropper's technical or computational resources (in contrast to classical coding schemes, e.g., RSA [23, 24]). In a quantum cryptosystem, any eavesdropping attempt is detected by virtue of the following interrelated fundamental limitations in quantum mechanics.

I. The process

$$\begin{aligned} |\varphi_0\rangle \otimes |A\rangle &\longmapsto |\varphi_0\rangle \otimes |\varphi_0\rangle \otimes |A_0\rangle, \\ |\varphi_1\rangle \otimes |A\rangle &\longmapsto |\varphi_1\rangle \otimes |\varphi_1\rangle \otimes |A_1\rangle, \\ &\text{if } \langle\varphi_0|\varphi_1\rangle \neq 0 \end{aligned} \tag{1}$$

(copying of unknown quantum states) is forbidden by the no-cloning theorem.

II. No information can be extracted about a nonorthogonal state without perturbing it; i.e., the following process is forbidden:

$$\begin{aligned} |\varphi_0\rangle \otimes |A\rangle &\longmapsto U(|\varphi_0\rangle \otimes |A\rangle) = |\varphi_0\rangle \otimes |A_0\rangle, \\ |\varphi_1\rangle \otimes |A\rangle &\longmapsto U(|\varphi_1\rangle \otimes |A\rangle) = |\varphi_1\rangle \otimes |A_1\rangle, \\ &\text{if } |A_0\rangle \neq |A_1\rangle, \end{aligned} \tag{2}$$

where $|A\rangle$ is a detector state and $U$ is a unitary operator describing simultaneous evolution of the detected and detector states. In essence, these limitations follow from Heisenberg's uncertainty principle, which rules out simultaneous measurement of observables associated with noncommuting operators.

For orthogonal states, nonperturbative copying or extraction of information is not forbidden. In nonrelativistic quantum mechanics, the observables

$$\rho_0 = |\varphi_0\rangle\langle\varphi_0|, \quad \rho_1 = |\varphi_1\rangle\langle\varphi_1|$$

are associated with commuting measurement operators defined as the orthogonal projectors

$$\mathcal{P}_{0,1} = |\varphi_{0,1}\rangle\langle\varphi_{0,1}|, \quad [\mathcal{P}_0, \mathcal{P}_1] = 0.$$

Limitations (1) and (2) follow from geometric properties of the state vectors $|\varphi_{0,1}\rangle$ in the Hilbert space corresponding to a quantum system. Orthogonal quantum states cannot be used in quantum cryptography without imposing additional fundamental limitations on measurability, such as those dictated by special theory of relativity. The photons used as quantum information carriers are inherently relativistic objects (quanta of a massless field). Therefore, novel secure quantum cryptosystems can be created by using additional limitations on measurability dictated by special theory of relativity.

Unavailability of single-photon sources and attenuation in communication channels are the main impediments to practical applicability of the quantum cryptography schemes whose security is guaranteed only by complete indistinguishability of nonorthogonal states. To demonstrate how a cryptosystem becomes insecure, consider a protocol using two nonorthogonal states (so-called B92 protocol). When attenuation is sufficiently high, the eavesdropper can use the intercept–resend strategy to extract complete information about the transmitted key without revealing the intervention. This far-from-optimal strategy can be described as follows. In each pulse, the eavesdropper makes use of measurements represented by the partition of the identity operator

$$\begin{aligned} I &= \mathcal{P}_0^\perp + \mathcal{P}_1^\perp + \mathcal{P}_?, \quad \mathcal{P}_0^\perp = a(I - |\varphi_0\rangle\langle\varphi_0|), \\ &\mathcal{P}_1^\perp = a(I - |\varphi_1\rangle\langle\varphi_1|), \\ \mathcal{P}_?^\perp &= I - \mathcal{P}_0^\perp - \mathcal{P}_1^\perp, \quad a = \frac{1}{1 + |\langle\varphi_0|\varphi_1\rangle|}. \end{aligned} \tag{3}$$

Consider the outcome space $\Omega$ consisting of three outcomes: $\Omega = \{0, 1, ?\}$. The probability of the outcome 0,

$$p_0 = \langle\varphi_0|\mathcal{P}_1^\perp|\varphi_0\rangle < 1,$$

is nonvanishing only for the input state $|\varphi_0\rangle$ and zero for the other input state, $|\varphi_1\rangle$. Similarly, the outcome 1 will take place with the probability

$$p_0 = \langle\varphi_1|\mathcal{P}_0^\perp|\varphi_1\rangle < 1,$$

which does not vanish only for the input state $|\varphi_1\rangle$. Since the states are nonorthogonal (i.e., completely indistinguishable), the outcome ? (inconclusive result) will occur with the probability

$$p_? = \langle\varphi_0|\mathcal{P}_?|\varphi_0\rangle = \langle\varphi_1|\mathcal{P}_?|\varphi_1\rangle = |\langle\varphi_0|\varphi_1\rangle| \neq 0.$$

Therefore, the eavesdropper can identify the state in the channel if the outcome 0 or 1 has occurred. After the outcome ?, the eavesdropper cannot tell which state has been transmitted. A pulse sent into a channel with attenuation may not reach the receiver. In the case of the outcome ?, the eavesdropper blocks the link and does not resend anything. When attenuation exceeds a certain critical level, the intervention cannot be detected, while the eavesdropper has complete information about the key and remains unnoticed. An analogous strategy can be used in the BB84 protocol.

Another essential security requirement is the availability of a single-photon source. Formally, it follows from (1) and (2) that any pair of nonorthogonal (even multiphoton) states can be used. These states must be distinguishable by means of the measurement procedure used at the receiver end; i.e., the procedure must implement projection onto multiphoton state vectors. The required source is a laser, and the corresponding coherent state is characterized by the mean photon number $\langle n\rangle = \mu$ only. However, since no phase reference

is available, the eavesdropper can determine a density matrix rather than a pure coherent state,

$$\rho = \int \frac{d\theta}{2\pi} |\mu e^{i\theta}\rangle\langle \mu e^{-i\theta}| = \sum_n p(\mu, n)|n\rangle\langle n|, \quad p(\mu, n) = e^{-\mu}\frac{\mu^n}{n!}. \tag{4}$$

The photons in a statistical mixture are described by the Poisson distribution. The actual experiments are performed on weak coherent radiation with $\mu \approx 0.1$–$0.3$, in which Fock states with photon numbers $n = 0, 1, \ldots$ can be detected with probabilities $p(\mu, n)$.

Thus, no photons are detected with probability $p(0, \mu)$; one photon, with probability $p(1, \mu)$; and so on. In the case of a multiphoton source, the eavesdropper can perform the so-called photon-number splitting (PNS) attack [17, 18]. A nondemolition measurement of the photon number is represented by a partition of the identity operator:

$$I = \bigoplus_{n=0}^{\infty} \mathcal{P}_n = \bigoplus_{n=0}^{\infty} |n\rangle\langle n|. \tag{5}$$

Formally, a nondemolition measurement that determines the photon number has two outcomes:

$$I = \mathcal{P}_{n=1} + \mathcal{P}_{n\neq 1}, \quad \mathcal{P}_{n\neq 1} = I - \mathcal{P}_{n=1}. \tag{6}$$

Therefore, if the outcome corresponds to $\mathcal{P}_{n=1}$ (one photon is detected), then the link is blocked. In the alternative case of $\mathcal{P}_{n\neq 1}$ (multiple photons are detected), the eavesdropper can split the photons into two parts, perform measurements on one part, and resend the remaining photons over a channel with lower attenuation. Since some photons do not reach the receiver, an attack of this kind cannot be revealed [18]. It should be noted here that the use of single-mode (monochromatic) photon states

$$|n\rangle = (a^+_{k_0})^n|0\rangle$$

(where $k_0$ is frequency and $\hbar = c = 1$) is a highly idealized model, because monochromatic waves are not localized in space and time. The measurement formally described by a projector $|n\rangle\langle n|$ requires access to a state as a whole, i.e., to the entire unbounded domain in space where the plane-wave amplitude of the state is not zero. Thus, the measurement requires an infinite time since the maximum speed is finite. If the distribution protocol makes use only of geometric properties of states in a Hilbert space, then it is formally beyond time (the discussion above is time-independent). For such protocols, the fact that a formally infinite time is required to measure the photon number in a communication channel is unimportant. Actually, any state occupies a bounded domain in space (associated with a finite spectral bandwidth). Accordingly, a finite time is required to formulate a real-time protocol for measuring the photon number (or another quantity), which leads to a delay in the arrival of measurement results at the receiver end. This makes it possible to reveal any eavesdropping attempt.

Thus, if the attenuation is such that the eavesdropper performing a nondemolition measurement of the photon number in a communication channel can block all signals when the channel carries a single-photon Fock state, then the protocol becomes insecure.

Reliable nonperturbative identification of orthogonal states is not forbidden by Bennett's theorem [20]. In a frequently articulated interpretation of the theorem, an orthogonal state is said to be "coupled" to an auxiliary system and change its state $|A\rangle$. This is inconsistent with the fact that the theorem has a purely geometric meaning: the state vector $|A\rangle$ of the auxiliary system can be rotated by a unitary transformation depending on the input vector $|\varphi_{0,1}\rangle$ into a new state $|A_0\rangle$ or $|A_1\rangle$ without changing the input vector. The theorem relies on the implicit assumption that the input vector $|\varphi_{0,1}\rangle$ is accessible as a whole. In other words, a unitary transformation $U$ can be performed only if the entire state space $\mathcal{H}_{\varphi_{0,1}}$ can be accessed. Otherwise, the transformation is not unitary. The fact that the proof deals with $|\varphi_{0,1}\rangle$ as a whole, disregarding its "internal" coordinates, implies that the "entire" state vector subject to the unitary transformation is available "instantly."

The Hilbert space $\mathcal{H}_{\varphi_{0,1}}$ representing a real physical system is tied to the Minkowski spacetime, where each state is characterized by a wavefunction amplitude. Access to the Hilbert space of states implies access to the spacetime domain of nonvanishing state amplitude (wavefunction). If only the domain in space where the state amplitudes do not vanish is accessible, then even orthogonal states cannot be reliably copied or distinguished. This is more or less obvious, because no manipulation (including copying and identification) cannot have an outcome with a probability higher than the partial norm corresponding to the states lying in the accessible spacetime domain, i.e., in the accessible part of the Hilbert space. An orthogonal state to be reliably copied or distinguished must be available instantly and as a whole.

If a state has nonzero amplitude within a bounded spacetime domain, then its accessibility implies that the domain is accessible. Since there is no limit on the maximum speed in nonrelativistic quantum mechanics, any bounded domain can be accessed instantly. In quantum field theory, the maximum speed is limited. Accordingly, a state with finite support can be accessed only after it has been mapped by a unitary transformation into a state having nonzero amplitude within an arbitrarily small domain in space. Then, the theorem from [20] can be applied. By the relativistic causality principle [25], a unitary transformation of a state defined on a bounded spacetime domain into a state localized within an arbitrarily small spatial domain can

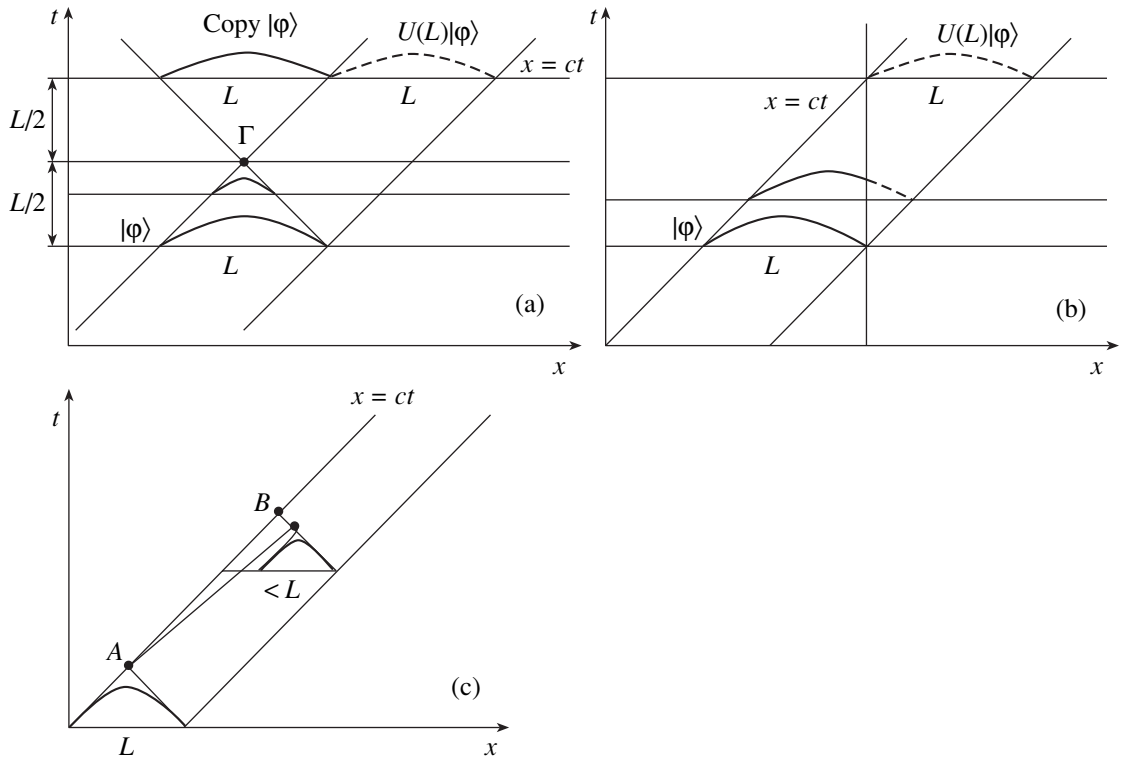



**Fig. 1.**

be performed only in a finite time. The minimal time is determined by the condition that the past light cone contains the original spatial domain of nonzero state amplitude (see Fig. 1). The cone's apex is the arbitrarily small domain (point) to which the original state amplitude is mapped by the unitary transformation. Each of two mutually orthogonal states mapped by a unitary transformation to the arbitrarily small domain can then be reliably copied or distinguished. Since these are states of a massless quantum field (photons) propagating with the maximum possible speed, a unitary transformation followed by copying would shift (delay) the propagation of a state in spacetime relative to the free evolution (propagation) of the state. This effect makes it possible to reveal any eavesdropping attempt. Note that relativistic limitations on measurements were examined by Landau and Peierls in a pioneering study [26] continued by Bohr and Rosenfeld [27].

In summary, a no-cloning theorem for orthogonal states of a massless quantum field can be stated as follows. Orthogonal states can be copied with a probability arbitrarily close to unity. The copied states have similar amplitudes shifted in spacetime. Thus, admissible manipulations are subject to a restriction weaker than nonrelativistic condition (1):

$$\begin{aligned} |\varphi_0\rangle &\longmapsto (U_L|\varphi_0\rangle) \otimes (U_L|\varphi_0\rangle), \\ |\varphi_1\rangle &\longmapsto (U_L|\varphi_1\rangle) \otimes (U_L|\varphi_1\rangle), \end{aligned} \tag{7}$$

where $U_L$ is the operator of translation along a light-cone branch over the interval $L = \Delta(x - t)$ (diameter of the domain of nonzero state amplitude). (For simplicity, it is assumed here that both states have nonzero amplitudes within the same spacetime domain, but their shapes $\varphi_{0,1}(x - t)$ are different.)

Bennett's theorem on identification of orthogonal states [20] is modified in a similar manner: admissible manipulations are subject to a restriction weaker than nonrelativistic condition (2):

$$\begin{aligned} |\varphi_0\rangle|A\rangle &\longmapsto (U_L|\varphi_0\rangle) \otimes |A_0\rangle, \\ |\varphi_1\rangle \otimes |A\rangle &\longmapsto (U_L|\varphi_1|) \otimes |A_1\rangle, \quad |A_0\rangle \neq |A_1\rangle. \end{aligned} \tag{8}$$

Figures 1a and 1b illustrate the above discussion.

Since the amplitudes of massless quantum-field states propagating in a certain direction along the $x$ axis depend only on $x - t$, it will suffice to consider the case in which time is fixed and coordinate is treated as variable (or vice versa). Consider one of the orthogonal states propagating at the speed of light, with amplitude $\varphi(x - t)$ ($c = 1$, and the state index 0 or 1 is omitted for brevity). Suppose that the state is localized within a domain $L$ in the sense that

$$\int_L |\varphi(x - t_0)|^2 dx \approx 1,$$

where $\varphi_{0,1}(x - t_0)$ is the amplitude at $t_0$.

To determine the values of the state amplitude at $t_0$ at all $x$ in the localization domain simultaneously, a unitary transformation must be applied to the entire state:

$$U\varphi_{0,1}(x - t_0) = \tilde{\varphi}_{0,1}(x' - t), \quad t > t_0.$$

The new state $\tilde{\varphi}(x' - t)$ has nonzero amplitude within a smaller spatial domain. The minimum domain size in $x'$ at time $t$ is dictated by the relativistic causality principle, which was formulated in its final form by Bogolyubov [25]. The unitary operator has nonzero matrix elements only if the points $(x, t_0)$ and $(x', t)$ lie in the past light cone with apex at the point $\Gamma$, which contains the domain of nonzero state amplitude at $t_0$. By any moment that precedes $L$, the original state amplitude can be mapped by a unitary transformation to a state amplitude localized in an arbitrarily small neighborhood of $\Gamma$. It is essential that this state is different from $\varphi(x - t_0)$. By the instant $\Gamma$, the values of state amplitude at all $x$ can be accessed instantly. Now, a measurement result can be obtained instantly, and complete information about the state is available (with probability one). If the states in the original pair are orthogonal, then a unitary transformation can be executed to obtain another pair of orthogonal states at the instant $\Gamma$, and, therefore, one state can be reliably distinguished from the other (by the theorem on complete distinguishability of orthogonal states [20], which is now applicable). It should be emphasized that these orthogonal states differ from the original ones. A state can be recovered or copied by executing the inverse unitary transformation "directed" forward in time. A state amplitude with shape identical to the original one cannot be obtained earlier than dictated by the relativistic causality principle. The state amplitude with an amplitude identical to the original one is localized in the future light cone with apex at $\Gamma$. Moreover, the resulting state also differs from the original one in the sense that it is delayed in time relative to the original state, which would have traveled the distance $L$ forward along the $x$ axis by the moment $L$ if no copying or information-extraction operations were attempted (see Fig. 1a). These considerations apply to extraction of information about states in a channel with probability one. A similar reasoning is true for extraction of information with probability less than unity, in which case the corresponding delay is shorter than $L$ (see Figs. 1a and 1b).

A similar analysis can be developed in the nonrelativistic case. If the limitations imposed by special relativity theory are ignored, the part of the analysis concerning the light cone should be removed. Formally, unitary transformations can then be executed instantly, and even the coordinate can be left out of analysis, while it should be kept in mind that any state subject to a unitary transformation can be instantly accessed as a whole (and so is the corresponding spatial domain).

A similar analysis can also be developed for a state mapped by a unitary transformation to a state of an auxiliary system. One example of unitary transformation of this kind is the so-called light trapping [28]. This transformation maps a state of the photon field to the vacuum state (since the field is massless and propagation speed cannot be zero), while an atomic state is mapped to a new one. Since the transformation is unitary, access to all values of the photon-packet amplitude is required at the point of localization of the atomic system. They are accessed naturally as the packet propagates at the speed of light and reaches ("enters" as a whole) the localized atomic system. If a result valid with unit is sought, a time interval equal to $L$ is required (the entire single-photon packet must "enter" the atomic system). The resulting photon field is in the vacuum state, while the auxiliary system is in a new state depending on the input photon state. By the moment $L$, the latter state can be identified with probability one, and its copy can be prepared with delay $L$ relative to the freely propagating original packet (see Fig. 1b).

Thus, extraction of any information about one of the orthogonal states modifies them, inducing translation (delay) in spacetime.

It is also important for further analysis that no evolution of a massless quantum field interacting with its environment (other quantum and classical degrees of freedom in a channel) cannot lead to "contraction" of a state in the sense that the partial norm of the state corresponds to a smaller spatial domain extending beyond the light cone as compared to free propagation (see Fig. 1c). Normally, the interaction gives rise to a mixed state, but the support of the density matrix in spacetime cannot be "contracted" and "pushed" out of the light cone (see Fig. 1c). Otherwise, information could be transmitted by using quantum states faster than the speed of light. Indeed, consider one of a pair of orthogonal states (see Fig. 1c). Alice cannot extract classical information from a quantum state before the instant defined by the condition that the state amplitude is covered by the past light cone. After that, classical information can be transmitted to Bob, but not faster than the speed of light (the partners are connected by a branch of the light cone in Fig. 1c). (Alice and Bob are the conventional names of the sender and receiver, respectively, and $A$ and $B$ denote their locations.) If a quantum state evolving in a channel could be "contracted" so that the cone's apex would lie in the spacelike region relative to the cone with apex at point $A$ and with a branch passing through point $B$ when the past light cone contains the state, then Bob could extract classical information from the quantum state before it would have been transmitted at the speed of light by Alice, because the apex of the light cone containing the "contracted" quantum state extends into the spacelike region.

With regard to cryptography, the analysis developed above implies that the eavesdropper can neither copy a state nor extract information about it from a noisy channel before the instants depicted in Figs. 1a and 1b. (If

the delay test is passed, then Eve's error probability cannot be smaller than those given by expressions (31) and (41) below.) This observation is of key importance for security of relativistic quantum cryptography, which remains secure irrespective of arbitrary attenuation. Since the security in question relies on the relativistic causality principle applied to evolution of quantum states, the limitations dictated by relativistic causality cannot be eliminated by any attenuation.

## 3. INFORMATION STATES AND MEASUREMENTS

Consider a scheme using multiphoton coherent states. The information states corresponding to 0 and 1 are defined, respectively, as orthogonal states $\rho_0$ and $\rho_1$ with nonoverlapping frequency bands of equal bandwidth. Let us show that there exist states with a given bandwidth characterized by the highest degree of localization of spatial amplitude (smallest spacetime extent). The use of a prescribed bandwidth facilitates analysis and does not affect the results. Polarization degrees of freedom are ignored as unimportant here. It is essential for the analysis developed below that photons are massless.

Since the relative phase of the states is not specified in each pulse sent into the communication channel, the channel contains a state described by the density matrix

$$\rho_{0,1} = \bigoplus_{n=0}^{\infty} p(n,\mu)|\varphi_{0,1}^{(n)}\rangle\langle\varphi_{0,1}^{(n)}| = \bigoplus_{n=0}^{\infty}\rho_{0,1}, \qquad \langle\varphi_i^{(n)}|\varphi_j^{(n)}\rangle = \delta_{ij}, \tag{9}$$

where

$$|\varphi_j^{(n)}\rangle = \int\!\dots\!\int d\hat{k}_1\dots d\hat{k}_n\varphi_{0,1}(\hat{k}_1)\dots\varphi_{0,1}(\hat{k}_n) \times \varphi^+(\hat{k}_1)\dots\varphi^+(\hat{k}_n)|0\rangle, \tag{10}$$

$$\varphi^+(\hat{k}) = \delta(\hat{k}^2)\theta(k_0)a^+(\hat{k}), \quad \hat{k} = (k_0, k). \tag{11}$$

Henceforth, states propagating in the same direction are considered, since such states are used in data transmission. By virtue of (10) and (11),

$$|\varphi_j^{(n)}\rangle = \int_0^{\infty}\!\dots\!\int_0^{\infty}\frac{dk_1\dots dk_n}{\sqrt{k_1\dots k_n}} \times\frac{\varphi_{0,1}(k_1)\dots\varphi_{0,1}(k_n)}{\sqrt{k_1\dots k_n}}|k_1,\dots,k_n\rangle, \qquad \varphi(k)\equiv\varphi(k,k_0=|k|)\equiv\varphi(k,k_0=k), \tag{12}$$

where

$$|k_1,\dots,k_n\rangle = a^+(\hat{k}_1)\dots a^+(\hat{k}_n)|0\rangle = \sqrt{\frac{k_1\dots k_n}{n!}}\sum_{\{i\}}\delta(k_1-q_i)\dots\delta(k_n-q_{i_n}) \tag{13}$$

denotes symmetrized basis vectors and $\{i\}$ means all possible indicial permutations. This expression automatically reflects the fact that photons are identical bosons. The orthogonality condition for generalized basis vectors is written as

$$\langle k_1\dots k_n|q_1\dots q_m\rangle = \delta_{nm}k_1\dots k_n \times\sum_{\{i\}}\delta(k_1-q_{i_1})\dots\delta(k_n-q_{i_n}). \tag{14}$$

The amplitudes of the orthogonal states $\varphi_{0,1}(k)$ are defined in frequency bands $\Delta k_{0,1}$ of equal bandwidth. It is assumed that

$$\operatorname{supp}\varphi_{0,1}(k)\in\Delta k_{0,1}, \quad \operatorname{supp}\varphi_0(k)\cap\operatorname{supp}\varphi_1(k) = \varnothing.$$

The analysis developed below makes use only of the bandwidths, but not the absolute frequencies. The state amplitudes are normalized to unity:

$$\int_0^{\infty}|\varphi(k)|^2dk = 1. \tag{15}$$

For single-mode (monochromatic) states,

$$|\varphi(k)|^2\longrightarrow\delta(k-k_0), \quad |k_1\dots k_n\rangle\equiv|n\rangle, \qquad |n\rangle = (a^+(\hat{k}))^n|0\rangle.$$

In a coordinate–time representation, the amplitudes of states propagating in the same direction depend only on the difference $\tau = x - t$ (since photons are massless). It holds that

$$|\varphi_{0,1}^{(n)}\rangle = \int_{-\infty}^{\infty}\!\dots\!\int_{-\infty}^{\infty}\frac{d\tau_1\dots d\tau_n}{(2\pi)^n}\varphi(\tau_1)\dots\varphi(\tau_n) \times|\tau_1\dots\tau_n\rangle, \tag{16}$$

$$\varphi(\tau) = \int_0^{\infty}\frac{dk}{\sqrt{k}}e^{ik\tau}\varphi(k), \qquad |\tau_1\dots\tau_n\rangle = \int_0^{\infty}\!\dots\!\int_0^{\infty}\frac{dk_1\dots dk_n}{\sqrt{k_1\dots k_n}} \times\exp[-i(k_1\tau_1+\dots+k_n\tau_n)]|k_1\dots k_n\rangle. \tag{17}$$

The fact that the amplitudes depend only on $\tau = x - t$ implies that, if a measurement result can be obtained at $t$ in the neighborhood of $x$, then it can be obtained with equal probability at $t' + (x' - x)$ in the neighborhood of $x'$, because the state amplitudes propagate at the speed of light. Since state amplitude depends on $x - t$, any measurement result can be obtained with a nonzero

probability only if the spacetime domain of nonzero state amplitude is accessible. The probability of any measurement result cannot be higher than the partial norm of the states corresponding to this domain. Since the state amplitude of a massless field depends only on $x - t$, the actual method used to access the domain of nonzero amplitude is formally irrelevant. Access can be gained if the coordinate is held constant (a conditionally local instrument is used). In this case, the state is recovered in the finite time required to reach the neighborhood of a certain point, which depends on the spatial extent of the nonzero-amplitude domain. This condition can be implemented by means of a unitary transformation that is local in space and nonlocal in time so that the photon state is mapped to a quantum state localized in space. Being massless, the photon field is mapped by the transformation to the vacuum state, while a localized (e.g., atomic) system is mapped to a new state (as in experiments on light trapping).

Formally, a measurement can be performed with a nonlocal instrument at a certain instant, since the amplitude depends only on $x - t$. (The instrument is nonlocal in the sense that it starts to interact with the state at a certain instant in the domain of nonzero state amplitude.) In this case, speed limitation and relativistic causality dictate that the measurement is nonlocal in space and requires a finite time to be performed, which cannot be shorter than that required for the instantaneous domain of nonzero state amplitude to fall into the past light cone. Moreover, the state belonging to the domain at a certain instant must be mapped to another localized state at a later instant, because photons are massless.

In any case, even to obtain a result with probability less than one, access is required to a subdomain where the amplitude is not zero. Formally, the delay due to the corresponding transformation of one state to another can be arbitrarily long (because the field is massless). A shorter delay implies a smaller accessible domain and a smaller partial norm corresponding to the domain. The next problem discussed here is the optimal (for the eavesdropper) relations between extraction of information about the transmitted state and the ensuing delay. For brevity, measurements that are feasible in the sense explained above are henceforth called measurements in the spacetime window $T$.

## 4. RELATIONS BETWEEN STATE DELAY AND INFORMATION EXTRACTED BY THE EAVESDROPPER: DIRECT ATTACK ON THE KEY

In this section probabilities for distinguishing between states are obtained as functions of delay, based on a qualitative analysis.

Any measurement on a quantum system can be described in terms of a set of operator-valued measures constituting a partition of unity. The identity operator is the direct sum of identity operators in symmetrized $n$-particle subspaces. Since a pair of states defined on frequency bands $\Delta k_0$ and $\Delta k_1$, respectively, are used (the speed of light is $c = 1$), it suffices to consider partition of the identity operator in basis vectors defined on these bands:

$$\begin{aligned} I &= \bigoplus_{n=0,\, i=0,\,1}^{\infty} I^{(n)}(\Delta k_i) \\ &= \bigoplus_{n=0}^{\infty} \int_{\Delta k_i} \dots \int_{\Delta k_i} \frac{dk_1 \dots dk_n}{\sqrt{k_1 \dots k_n}} |k_1 \dots k_n\rangle\langle k_1 \dots k_n| \\ &= \bigoplus_{n=0,\, i=0,\,1}^{\infty} \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} \mathcal{M}_i^{(n)}(d\tau_1 \dots d\tau_n), \end{aligned} \tag{18}$$

where operator-valued measure is defined as

$$\begin{aligned} \mathcal{M}_i^{(n)}(d\tau_1 \dots d\tau_n) &= |\tau_1 \dots \tau_n\rangle\langle \tau_1 \dots \tau_n| \\ &= \frac{d\tau_1 \dots d\tau_n}{(2\pi)^n}, \end{aligned} \tag{19}$$

$$\begin{aligned} |\tau_1 \dots \tau_n\rangle &= \int_{\Delta k_i} \dots \int_{\Delta k_i} \frac{dk_1 \dots dk_n}{\sqrt{k_1 \dots k_n}} \\ &\times \exp[-i(k_1\tau_1 + \dots + k_n\tau_n)] |k_1 \dots k_n\rangle. \end{aligned} \tag{20}$$

The operator-valued measure $\mathcal{M}_i^{(n)}(d\tau_1 \dots d\tau_n)$ is the probability of measurement results within the spacetime window $(\tau_1, \tau_1 + d\tau_1 \dots \tau_1, \tau_n + d\tau_n)$ for the input density matrix. The $n$-photon component of the density matrix is

$$\begin{aligned} \Pr^{(n)}(-\infty, \infty) &= \mathrm{Tr}\{\rho_{0,1}^{(n)} I(\Delta k_{0,1})^{(n)}\} \\ &= \mathrm{Tr}\{\rho_{0,1}^{(n)} (\mathcal{M}_{0,1}^{(n)}(T_1 \dots T_n) + \mathcal{M}_{0,1}^{(n)}(\bar{T}_1 \dots \bar{T}_n))\} \\ &= \left\{\left(\int_{T_1} + \int_{\bar{T}_1}\right) |\varphi_{0,1}(\tau_1)|^2 d\tau_1\right\} \dots \left\{\left(\int_{T_n} + \int_{\bar{T}_n}\right) |\varphi_{0,1}(\tau_n)|^2 d\tau_n\right\} \\ &= (p_{0,1}(T_1) + p_{0,1}(\bar{T}_1)) \dots (p_{0,1}(T_n) + p_{0,1}(\bar{T}_n)). \end{aligned} \tag{21}$$

In the case of $n = 1$ (single-photon packet corresponding to state 0 or 1), the probability of obtaining a measurement result defined on the entire space is

$$p_{0,1}(T_1) + p_{0,1}(\bar{T}_1) = 1.$$

Accordingly, if only the spacetime window $T_1$ is accessible for measurement, then the probability to obtain a corresponding result is

$$p_{0,1}(T_1) < 1.$$

If an outcome is obtained, than orthogonal states are reliably distinguished with this probability. However, no outcome obtained in the time window $T_1$ with the

probability $p_{0,1}(\bar{T}_1) < 1$, since the outcome takes place in the spacetime region $\bar{T}_1 = (-\infty, \infty) - T_1$, which is not accessible to the observer. For the observer, the absence of an outcome is an inconclusive result. In this case, the probability of correct identification of one of two states is 1/2, which is equivalent to that of a correct guess.

As noted above, in the case of (massless) photon field, access to the spacetime domain of nonzero state amplitude induces delay (translation) in spacetime. This implies that a result can be obtained with probability $p_{0,1}(T_1) < 1$ in the spacetime window $T_1$ if access to a domain of certain size is gained, which leads to delay.

In the case of measurement on a multiphoton input state represented by (9), the indistinguishability of photons implies that, in view of (21), the probability of detection in a time window $T$ satisfies the relation

$$1 = \sum_{k=0}^{n} C_n^k p^k(T) p^{n-k}(\bar{T}), \qquad p^k(T) + p^{n-k}(\bar{T}) = 1, \tag{22}$$

where $C_n^k$ is a binomial coefficient. The state indices 0 and 1 are omitted here for brevity. This relation means that $k$ photons can be detected in a time window, while the remaining $n - k$ photons cannot (formally, they are detectable outside the accessible window $T$), and, since photons are identical, the probability of such an event is

$$C_n^k p^k(T) p^{n-k}(\bar{T}).$$

Once a frequency band is chosen for state amplitudes, the next step is to determine the minimal time window $T$ required to detect a state with a probability arbitrarily close to unity, i.e., the minimal delay caused by the eavesdropper attempting to access the entire state in a given frequency band.

For a given finite frequency band $c\Delta k$, the maximum partial norm corresponding to a given time window $T$ is obtained for a state with support contained in $\Delta k$ and with a spacetime amplitude $\varphi(k)$ that solves the unconditional extremum value problem for the functional

$$\mathcal{F} = \frac{\frac{1}{2\pi}\int_{-T}^{T} |\varphi(\tau)|^2 d\tau}{\int_0^\infty |\varphi(k)|^2 dk}. \tag{23}$$

One-particle amplitudes are normalized to unity:

$$\int_0^\infty |\varphi(k)|^2 dk = 1. \tag{24}$$

The highest degree of localization within the time window $T$ is reached for the states whose amplitudes are the eigenfunctions of the integral equation obtained by varying the functional

$$\lambda_m \varphi_m(k) = \frac{1}{\pi}\int_{\Delta k} \frac{\sin(k-k')T}{k-k'} \varphi_m(k') dk', \tag{25}$$
$$\lambda_m(\zeta) \sim 1 - \frac{4\sqrt{\pi}8^m}{m!}\zeta^{m+1/2} e^{-2\zeta}, \quad \zeta = \Delta k T.$$

The largest eigenvalue is the probability of localization (partial norm within the time window $T$), and the corresponding eigenfunction determines the optimal form of the state amplitude. This equation was analyzed in [29, 30]. The eigenvalues are positive functions of $\Delta kT$ making up a decreasing sequence: $1 > \lambda_0 > \lambda_1 > \ldots > 0$ ($m = 0, 1, \ldots, \infty$). The first few eigenvalues were found for several values of $\Delta kT$ in [30], and an asymptotic expression valid at $\Delta kT \gg 1$ for any $m$ was derived in [29]. Equation (25) implies that the eigenvalues exponentially approach unity at $\zeta > 1$. Thus, no measurement result can be obtained within the spacetime window $T$ with a probability greater than $\lambda_0(\Delta kT)$ for a state with support in a frequency band $\Delta k$. The probability of detection outside the window can be reduced to an arbitrarily small value by adjusting its width $T$.

In quantum cryptography, an appropriate choice of the window guarantees that the eavesdropper will have only an arbitrarily small amount of information about the transmitted state outside the window. Thus, the minimal time window $T$ is required to detect of one of the orthogonal states with a probability arbitrarily close to unity (see Eq. (25)).

Because the state amplitude is defined on the mass surface ($\varphi(k, k_0)$ is a function of $k_0 = k$ rather than of independent $k$ and $k_0$), it has a nonzero value at every point outside any compactum [31–33]. The nonlocalizability of states in quantum field theory is a well-established fact (e.g., see [34] for discussion of the underlying physics). In the present context, nonlocalizability can be demonstrated as a corollary to the Wiener–Paley theorem [35]. If a function $\varphi(k)$ normalized by the relation

$$\int_0^\infty dk |\varphi(k)|^2 = 1,$$

vanishes at $k \le 0$, but not everywhere, then the admissible decay of its spacetime preimage at infinity is guaranteed by convergence of the integral

$$\int_{-\infty}^{\infty} \frac{\ln|\varphi(\tau)|}{1+\tau^2} d\tau < \infty.$$

Hence, the amplitude $\varphi(\tau)$ cannot be even exponentially decreasing (and much less vanishing outside a compactum), because if

$$|\varphi(\tau)| \propto \exp(-\alpha|\tau|),$$

then the integral is divergent. However, the decrease in

amplitude may be arbitrarily close to exponential with any $\alpha > 0$:

$$|\varphi(\tau)| \propto \exp(-\alpha|\tau|/\ln(\ln\ldots|\tau|)).$$

A similar degree of localization of photon field is feasible in the 3D case [36], even though it has been believed for a long time (after Newton and Wigner [37]) that the steepest possible decay in space is described by the inverse power 7/2.

The fact that amplitude is nonlocalizable (nonvanishing outside any compactum) is essentially due to relativistic causality. In particular, it was shown in [33] that free evolution of a state amplitude localized in a finite spatial region at an initial moment $t_0$ results in nonzero state amplitude in a region separated from the initial localization domain by an arbitrarily large spacelike interval at any $t > t_0$. This behavior is inconsistent with relativistic causality, since it implies that information can be transmitted faster than the speed of light even if the probability of a measurement outcome obtained in the domain separated from the initial localization domain by a spacelike interval is less than unity.

Suppose that the eavesdropper tries to identify one of the two orthogonal states described by (9), using a preset time window (controlled delay) $\delta T$. The index 0 or 1 is omitted hereinafter, because identification probabilities are equal for equal bandwidths corresponding to states 0 and 1.

Since the states are orthogonal, they can be distinguished by detecting at least one photon within the time window $\delta T$ chosen by the eavesdropper. Note that the vacuum components contained in (21) and (22) carry no information about the transmitted state (0 or 1). The total probability minus the vacuum-state probability must be normalized to unity, because only the probability of detection of at least one photon carries information about the transmitted state. Since the vacuum-state probability is

$$p(0, \mu) = 1 - e^{-\mu},$$

the normalization condition is

$$\frac{1}{1-e^{-\mu}}\sum_{n=1}^{\infty} p(n, \mu) = 1. \tag{26}$$

A simple combinatorial analysis using expression (26) leads to the following expression for the probability that the states 0 and 1 are distinguished by detecting at least one photon:

$$\begin{aligned}&\Pr_E(n \geq 1 \text{ detected})\\ &= \frac{1}{1-e^{-\mu}}\sum_{n=1}^{\infty} p(n, \mu)\Pr^{(n)}(n \text{ detected})\\ &= \frac{1}{1-e^{-\mu}}\sum_{n=1}^{\infty} p(n, \mu)[1-(1-p(\delta T))^n] = \frac{1-e^{-\mu p(\delta T)}}{1-e^{-\mu}}.\end{aligned} \tag{27}$$

This probability corresponds to a conclusive result and approaches unity as $\delta T \longrightarrow \infty$. Therefore, this probability exponentially approaches unity as $\Delta kT$ only if a time window $T$ where the state is localized almost entirely is accessible.

Accordingly, the no-detection probability (inconclusive result) corresponding to delay $\delta T$ is

$$\begin{aligned}&\Pr_E(\text{not detected}) = 1 - \Pr_E(n \geq 1 \text{ detected})\\ &= \frac{e^{-\mu p(\delta T)} - e^{-\mu}}{1-e^{-\mu}}.\end{aligned} \tag{28}$$

To extend these results to the case of a classical field, note that the probability that two orthogonal states of electromagnetic field with nonoverlapping frequency bands are distinguished (depending on $\delta T$) approaches the classical limit as $\mu \longrightarrow \infty$. In this limit, an arbitrarily small time window ($\delta T \longrightarrow 0$) is sufficient for distinguishing the states almost completely; i.e., it suffices to have access to an arbitrarily small spacetime domain where classical signal does not vanish. In the opposite limit ($\mu \ll 1$), the time window (delay) $\delta T$ required to distinguish the states almost completely must correspond to a domain where the partial norm of the squared state amplitude is close to unity.

Now, consider the measurements that must be performed at the receiver end to reveal the presence of the eavesdropper.

To ensure security of the key, it is essential that Bob accepts only pulses with outcomes in the time window defined by the relation

$$t_A + L_{ch}/c, t_A + L_{ch}/c + T, \quad c = 1.$$

Here, $t_A$ is the instant at which the state enters the communication channel, $L_{ch}$ is the channel length, and $T$ is the time window where the states are localized almost entirely. The time $(t_A + L_{ch} + T) - t_A$ is required for a state extending over $T$ and propagating with the speed of light to reach Bob completely. Henceforth, it is assumed that $t_A = 0$ for brevity.

Since $T$ is uniquely related to a frequency band, the measurements to be performed by Bob should be restricted to $\Delta k_{0,1}$, say, by filtering the signal before it reaches a light detector. Formally, a filtered measurement of this kind is described by a partition of the identity operator into operator-valued measures with supports in preset nonoverlapping frequency bands $\Delta k_{0,1}$:

$$\begin{aligned}I &= \bigoplus_{n=0, i=0,1}^{\infty} I^{(n)}(\Delta k_i)\\ &= \bigoplus_{n=0}^{\infty}\int_{\Delta k_i}\ldots\int_{\Delta k_i}\frac{dk_1\ldots dk_n}{k_1\ldots k_n}|k_1\ldots k_n\rangle\langle k_1\ldots k_n|\\ &= \bigoplus_{n=0, i=0,1}^{\infty}\Bigg(\underbrace{\int\ldots\int \mathcal{M}_i^{(n)}(d\tau_1\ldots d\tau_n)}_{\forall\tau_i \in T_{ch}}\\ &\quad + \underbrace{\int\ldots\int \mathcal{M}_i^{(n)}(d\tau_1\ldots d\tau_n)}_{\exists\tau_i \in \bar{T}_{ch}}\Bigg),\end{aligned} \tag{29}$$

where $T_{ch}$ denotes the time window ($L_{ch}$, $L_{ch} + T$) and $\bar{T}_{ch} = (-\infty, \infty)/T_{ch}$ (complementary to $T_{ch}$). In this scheme, the set of measurement outcomes can be divided into two subsets: those in which all counts are in the time window $T_{ch}$ controlled by Alice and represented by the first summand in (30) below and those in which at least one count falls outside $T_{ch}$. The pulses resulting in outcomes of the latter type are discarded by Bob. In this measurement scheme, filters are required to ensure that the eavesdropper cannot use states occupying shorter intervals, i.e., those with bandwidths wider than $\Delta k_{0,1}$. Moreover, the light-detector response time $\tau_d$ must satisfy the condition $\tau_d \ll T$ to ensure representative statistics within the state localization window. The probability that the eavesdropper identifies a bit sent by Alice and passes the delay test performed by Bob does not exceed

$$\begin{aligned}&\Pr(\mathrm{bit}_E = \mathrm{bit}_A \wedge \mathrm{test}(\forall\tau_i \in T_{ch}) = \mathrm{OK})\\ &\quad = \Pr_E(\text{not detected}) \times \frac{1}{2} \times 1\\ &\quad + \Pr_E(n \geq 1 \text{ detected}) \times 1 \times (1 - p(\delta T))\\ &\quad = \frac{1}{2}\frac{e^{-\mu p(\delta T)} - e^{-\mu}}{1 - e^{-\mu}} + (1 - p(\delta T))\frac{1 - e^{-\mu p(\delta T)}}{1 - e^{-\mu}}.\end{aligned} \tag{30}$$

The first summand in (30) is the product of three multipliers, which correspond to the absence of any count within the eavesdropper's time window $\delta T$, a correct guess in the absence of any count (with probability 1/2), and a passed delay test, respectively. The last probability is unity, because no counts occur within $\delta T$.

The second summand in (30) is the product of the probability that at least one photon is detected in the frequency band $\Delta k_0$ or $\Delta k_1$ within the time window $\delta T$, the unit probability of identification when at least one photon is counted, and the probability of passing the delay test (i.e., to be counted by Bob within $T_{ch}$). Since

$$(1 - p(\delta T))^n < 1 - p(\delta T),$$

the last probability does not exceed $1 - p(\delta T)$; i.e., the highest probability that the presence of the eavesdropper will not be revealed is attained if only single-photon states are resent to Bob, rather than multiphoton ones represented by density matrix (9). Physically, this means that the probability of counting outside $T_{ch}$ is higher for delayed multiphoton states. In particular, the probability of passing the delay test tends to zero in the limit of classical signal even for in the case of infinitesimal delay, because an arbitrarily short "tail" that does not fit into $T_{ch}$ will fail the test. (Note that delay is guaranteed by the existence of a maximum speed and by the "tuning" of the time window to the "shortest" states associated with the frequency band in question.) Denote the highest value of probability (30) by

$$\begin{aligned}&\delta_E^{OK}\\ &= \max_{\varphi(k),\,\delta T}\{\Pr(\mathrm{bit}_E = \mathrm{bit}_A \wedge \mathrm{test}(\forall\tau_i \in T_{ch}) = \mathrm{OK})\}.\end{aligned} \tag{31}$$

This value corresponds to the "shortest" states in (23)–(25); i.e., the function $p(\delta T)$ is known. Therefore, the eavesdropper can determine the corresponding $\delta T$; i.e., $p(\delta T)$ can henceforth be used as an independent variable. (Note that it cannot exceed (31) by virtue of relativistic causality.) Moreover, this scheme does not require collective measurement, because only orthogonal states are used and the protocol is implemented in real time, in contrast to schemes using only geometric properties of nonorthogonal states. In the present scheme, orthogonal states cannot be reliably distinguished because the eavesdropper does not actually have enough time to access the entire state. Whereas collective measurement may provide more information about nonorthogonal states in nonrelativistic cryptography (blocks, rather than individual states, are measured), no advantage over individual measurement on each pulse is gained by the eavesdropper in relativistic cryptography.

If the input states 0 and 1 are pure single-photon ones, i.e.,

$$\begin{aligned}|\varphi_{0,1}^{(1)}\rangle &= \int_0^\infty \frac{dk}{\sqrt{k}}\frac{\varphi_{0,1}(k)}{\sqrt{k}}|k\rangle,\\ \varphi(k) &\equiv \varphi(k, k_0 = |k|) \equiv \varphi(k, k_0 = k),\end{aligned} \tag{32}$$

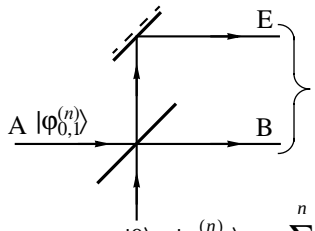
then the probability that the eavesdropper identifies the transmitted state and pass the delay test at the receiver end is

$$\begin{aligned}&\Pr(\mathrm{bit}_E = \mathrm{bit}_A \wedge \mathrm{test}(\tau \in T_{ch}) = \mathrm{OK})\\ &\quad = (1 - p(\delta T)) \times \frac{1}{2} \times 1\\ &\quad + p(\delta T) \times 1 \times (1 - p(\delta T)) \leq \frac{9}{16}.\end{aligned} \tag{33}$$

Its maximum is attained when $p(\delta T) = 1/4$. The optimal time window $\delta T$ for the eavesdropper can be found by solving the equation

$$\frac{9}{16} = \int_{-\delta T}^{\delta T} |\varphi(\tau)|^2 d\tau. \tag{34}$$

If $\mu = 1$, then the value of $\delta_E$ is not very close to 9/16 = 0.5625, which would be the case if the input states were single-photon Fock states with amplitudes $\varphi_{0,1}$ defined within nonoverlapping frequency bands of bandwidth $\Delta k$. It should be recalled here that a finite bandwidth is actually required to set a lower limit on the spacetime support of state amplitude. Moreover, the protocol does not dictate that Alice must prepare the "shortest" states




$$|\varphi_{\mathrm{E,B}}^{(n)}\rangle = \sum_{k=0}^{n} (\sqrt{\eta})^k (\sqrt{1-\eta})^{n-k} \times \sqrt{C_n^k}|\varphi^{(k)}\rangle_{\mathrm{E}} \otimes |\varphi^{(n-k)}\rangle_{\mathrm{B}}$$

**Fig. 2.**

within the preset frequency band, but Bob must accept only measurements resulting in counts fitting in the time window (measured from the leading front) corresponding to the "shortest" states.

## 5. PHOTON-NUMBER SPLITTING ATTACK

In the direct attack on the key analyzed above, the eavesdropper directly measures the transmitted state. In a photon-number splitting (PNS) attack dealing with a multiphoton state, the eavesdropper "splits" the photons into two parts by means of a beamsplitter (see Fig. 2) and measures the states of the "diverted" photons. The beamsplitter is characterized by the splitting ratio $\eta/(1-\eta)$.

The output of a beamsplitter is

$$|\varphi_{\mathrm{E,B}}^{(n)}\rangle = \sum_{k=0}^{n} (\sqrt{\eta})^k (\sqrt{1-\eta})^{n-k} \times \sqrt{C_n^k}|\varphi^{(k)}\rangle_{\mathrm{E}} \otimes |\varphi^{(n-k)}\rangle_{\mathrm{B}}, \quad (35)$$

$$|\varphi^{(k)}\rangle_{\mathrm{E}} = \left( \underbrace{\int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} \frac{d\tau_1 \dots d\tau_k}{(2\pi)^k} \varphi(\tau_1) \dots \varphi(\tau_k) |\tau_1 \dots \tau_k\rangle}_{k} \right)_{\mathrm{E}}, \quad (36)$$

$$|\varphi^{(n-k)}\rangle_{B} = \left( \underbrace{\int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} \frac{d\tau_1 \dots d\tau_{n-k}}{(2\pi)^{n-k}} \varphi(\tau_1) \dots \varphi(\tau_{n-k}) |\tau_1 \dots \tau_{n-k}\rangle}_{n-k} \right)_{\mathrm{B}}. \quad (37)$$

It is essential for further analysis that the beamsplitter output states measured by Eve (conventional name of the eavesdropper) and Bob are entangled. By using (35)–(37), the density matrix of an entangled state can be expressed as

$$\rho_{\mathrm{E,B}} = \sum_{n=0}^{\infty} p(\mu, n) |\varphi_{\mathrm{E,B}}^{(n)}\rangle\langle\varphi_{\mathrm{E,B}}^{(n)}|. \quad (38)$$

Because of entanglement, the measurements performed by Eve and Bob are correlated, as in the classical Einstein–Podolsky–Rosen paradox [38]. For the present analysis, it is important that the photon numbers in Eve's and Bob's channels are entangled. Since Bob accepts only measurements resulting in photon counts, entanglement implies that Eve gains reliable information about the transmitted state when she records a simultaneous count. The probability of simultaneous counts recorded by Bob and Eve is

$$\Pr(\text{E detected } n \geq 1 \wedge \text{B detected } n \geq 1) = \sum_{n=1}^{\infty} p(\mu, n) \mathrm{Tr}_{\mathrm{E,B}} \left\{ \left( \bigoplus_{n=1, i=0,1}^{\infty} I^{(n)}(\Delta k_i) \right)_{\mathrm{B}} \times \left( \bigoplus_{n'=1, i=0,1}^{\infty} I^{(n')}(\Delta k_i) \right)_{\mathrm{E}} \left( \sum_{k=0}^{n} (\sqrt{\eta})^k (\sqrt{1-\eta})^{n-k} \right. \right. \quad (39)$$

$$\left. \left. \times \sqrt{C_n^k}|\varphi^{(k)}\rangle_{\mathrm{E}} \otimes |\varphi^{(n-k)}\rangle_{\mathrm{B}} \right) \left( \sum_{k'=0}^{n} (\sqrt{\eta})^{k'} (\sqrt{1-\eta})^{n-k'} \times \sqrt{C_n^{k'}}\langle\varphi^{(k')}| \otimes {}_{\mathrm{B}}\langle\varphi^{n-k'}| \right) \right\},$$

where the operator-valued measure containing only summands with $n, n' \geq 1$ corresponds to counting at least one photon by Eve and Bob simultaneously. The eavesdropper identifies the transmitted states and remains unnoticed with probability (39). Note that photon-splitting measurements do not induce any delay at the receiver end. The operator-valued measures in (29) associated with measurements in a finite time window can be replaced with identity operators, because state splitting does not induce any delay and only exponentially small "tails" of states fall outside the window $T$.

Each $n$-photon component of the density matrix corresponds to $n$ photons counted with probability $\eta^n$ in Eve's beamsplitter channel. Since these outcomes do not result in any counts at Bob's end, this component does not contribute to the total probability. Similarly, each $n$-photon component of the density matrix corresponds to $n$ photons counted with probability $(1-\eta)^n$ in Bob's beamsplitter channel. Since these outcomes are accepted by Bob, but Eve records no counts in this case, the corresponding probability that the transmitted state is identified by Eve is the probability 1/2 of a correct guess.

The fact that the vacuum components of states does not contribute to counts recorded at Eve's or Bob's end should be taken into account in calculating the total probability, as done in the preceding section.

Finally, the probability that the eavesdropper performing a PNS attack identifies the transmitted states

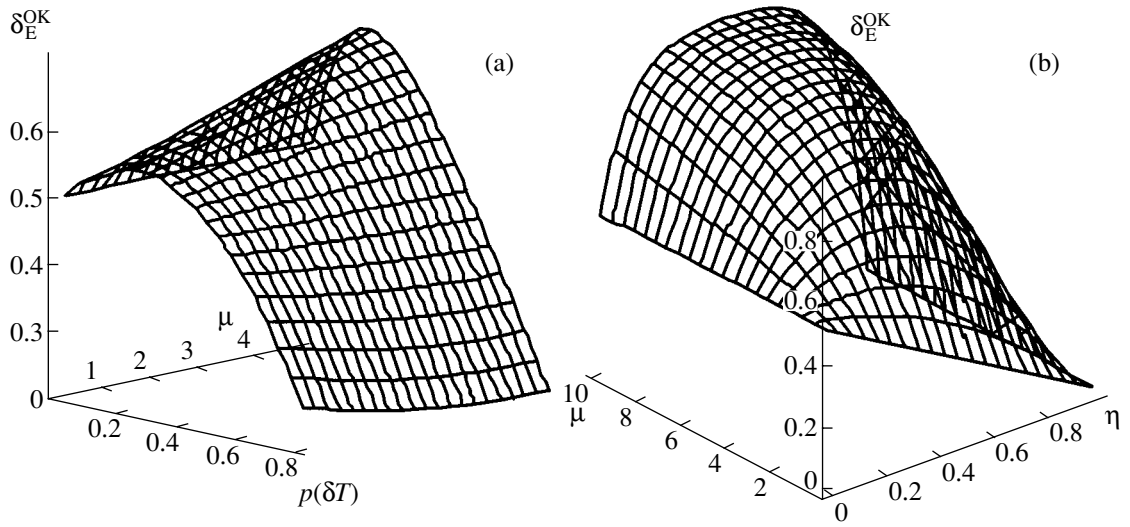



**Fig. 3.**

and remains unnoticed is expressed as

$$\Pr(\mathrm{bit}_E = \mathrm{bit}_{A,B} \wedge n \geq 1 \text{ detected by B})$$
$$= \frac{1}{2}\frac{e^{-\mu\eta} - e^{-\mu}}{1 - e^{-\mu}} + \left(1 - \frac{e^{-\mu\eta} + e^{-\mu(1-\eta)} - 2e^{-\mu}}{1 - e^{-\mu}}\right). \quad (40)$$

The first summand in (40) corresponds to the case when all of the $n$ photons in each $n$-photon component of the density matrix are counted by Bob. The probability that Eve distinguishes between states is 1/2 (the multiplier in the first summand). The second summand represents simultaneous photon counting by Eve and Bob, in which case the probability of identification is unity.

In this attack, the maximum value

$$\delta_E^{OK} = \max_\eta \{\Pr(\mathrm{bit}_E = \mathrm{bit}_{A,B} \wedge n \geq 1 \text{ detected by B})\} \quad (41)$$

for a preset photon number $\mu$ depends on the beamsplitter's splitting ratio $\eta$.

## 6. MAXIMUM $\delta_E^{OK}$

The maximum probability $\delta_E^{OK}$ that the eavesdropper measures each bit transmitted by Alice and passes the delay test is plotted in Fig. 3a as a function of $p(\delta T)$ and $\mu$. According to argumentation presented above (see Fig. 1c), attenuation in a communication channel cannot increase the probability of successful detection by the eavesdropper within a preset time window.

If the photon number is as small as $\mu = 0.1$, then $\delta_E^{OK} = 0.565$ (see Table 1 and Fig. 3a), which exceeds the probability of a correct guess only by 0.065. For convenience, the analysis presented below deals with the eavesdropper's error probability $\delta_E = 1 - \delta_E^{OK}$, which is interpreted as follows. Since Bob accepts only pulses received within the preset time window, $\delta_E$ is the probability that the eavesdropper fails to identify the bits that pass the test and are accepted by Bob. If $\mu = 5$, then $\delta_E = 30\%$. This means that the eavesdropper correctly detects about 70% of the bits that are accepted by the legitimate users. If the length $N$ of the string of bits adopted by Alice and Bob is sufficiently large, then $0.3N$ encrypted bits can be extracted (see below).

When the mean photon number is large, $\delta_E^{OK}$ approaches unity; i.e., $\delta_E \longrightarrow 0$ and the delay also tends to zero. Even though the partial norm $p(\delta T)$ approaches zero ($p(\delta T) = 0.04$ if $\mu = 100$), its small value is compensated for by the large number of photons that can be counted in the time window $\delta T$. When $\mu \gg 1$, the classical limit is approached and two states of electromagnetic field with nonoverlapping frequency bands ("colors") can be reliably identified almost instantly.

**Table 1**

| $\mu$ | $p(\delta T)$ | $\delta_E^{OK}$ |
|---|---|---|
| 0.1 | 0.25 | 0.565 |
| 0.3 | 0.25 | 0.569 |
| 1.0 | 0.25 | 0.587 |
| 3.0 | 0.23 | 0.642 |
| 5.0 | 0.2 | 0.691 |
| 10.0 | 0.14 | 0.771 |
| 20.0 | 0.11 | 0.847 |
| 50.0 | 0.06 | 0.918 |
| 100.0 | 0.04 | 0.952 |

**Table 2**

| $\mu$ | $\eta$ | $\delta_E^{OK}$ |
|---|---|---|
| 0.1 | 0.0001 | 0.502 |
| 1 | 0.20 | 0.515 |
| 3.0 | 0.40 | 0.746 |
| 5.0 | 0.45 | 0.893 |
| 10.0 | 0.5 | 0.997 |

Since coherent states with $\mu = 1$–3 (associated with relatively large $\delta_E$) can be practically implemented, faint laser pulses with $\mu = 0.1$–0.3 (the fraction of empty pulses is ≈90%) are not required. When $\mu = 3$, empty pulses are scarce (their fraction is 5%) and the key generation rate is higher by a factor of 10.

Since the probability $\delta_E$ depends only on the dimensionless parameter $\Delta kT$, any bandwidth can be combined with a time window sufficiently wide to "contain" the entire photon state. The value of $\delta_E$ depends only on the structure of a state, reaching a maximum for a pure single-photon state propagating with the speed of light. A single-photon packet corresponds to the quantum relativistic limit, in which case $\delta_E$ can be interpreted as a "universal" constant, because it is derived by using only the normalization of state amplitude and the fact of propagation at the speed of light.

The probability $\delta_E^{OK}$ associated with PNS attack is plotted in Fig. 3b as a function of the beamsplitter parameter $\eta$ and the mean photon number $\mu$. Table 2 lists the maximum $\delta_E^{OK}$ corresponding to the optimal $\eta$.

When $\mu = 0.1$, the probability that the eavesdropper identifies a state and remains unnoticed is close to 1/2 ($\delta_E^{OK} = 0.502$). This result is qualitatively clear. Indeed, the state corresponding to $\mu = 0.1$ (minus the vacuum component) is typically a single photon, and the probabilities of two or more photons are much lower. Therefore, a single-photon state can be detected either in Eve's or in Bob's beamsplitter output channel, but can never be detected in both channels simultaneously.

When $\mu > 10$, the classical limit is approached, and the probability of simultaneous counts recorded by Eve and Bob is close to unity for $\eta = 1/2$.

## 7. KEY SECURITY

The principal difficulty in proving the unconditional security of nonrelativistic quantum cryptography (based on complete indistinguishability of nonorthogonal states) arises from the fact that the aforementioned protocols are formulated as "exchange" protocols in a Hilbert space. When a protocol is formulated by using only the properties of the state space, it is necessary to prove its stability with respect to collective measurement, and this is the most difficult task. To date, several proofs of security have been proposed for various protocols [21, 22, 39, 40].

Transmission of information always involves preparation of the quantum system to be employed as an information carrier, transfer of its copy to a remote user over a communication channel, and subsequent measurement of its state. In relativistic quantum cryptography, there is no need to take into account collective measurement. Moreover, the error (delay) induced by the eavesdropper can be separated from those due to channel noise (e.g., dark counts). In the nonrelativistic case, even partial discrimination between these errors is impossible.

A criterion for key security can loosely be stated as follows: Alice and Bob must use identical copies of a key, and it must be known only to them. To state it formally, suppose that both Alice and Bob have a string of $m$ bits obtained by implementing a protocol and adopt it as a key. Then, the key must satisfy the following requirements,

**1. Identity.** The probability that the $i$th bits $b_A(i)$ and $b_B(i)$ in the $m$-bit strings adopted as keys by Alice and Bob, respectively, are different must be exponentially small for some preset values of security parameters $M$, $\varepsilon_1$:

$$\forall \varepsilon_1 > 0 \quad \exists M, \quad \Pr\{b_A(i) \neq b_B(i)\} \leq e^{-M} \leq \varepsilon_1. \quad (42)$$

The mutual information between Alice's and Bob's $m$-bit strings is

$$I(A; B) \geq m - 2^{-M'}. \quad (43)$$

**2. Security.** The probability that Eve knows every bit in the key exceeds by an exponentially small quantity the probability $2^{-m}$ of a correct guess of the $m$-bit string adopted by Alice. (The lowest probability 1/2 corresponds to a correct guess of a bit.) Equivalently, the probability that Eve knows any particular bit in the string adopted as a key exceeds 1/2 by an exponentially small quantity:

$$\forall \varepsilon_2 > 0 \quad \exists \eta_2, \zeta,$$
$$\Pr\{b_A(i) = b_E(i)\} \leq \frac{1}{2} + e^{-\eta_2}, \quad e^{-\eta_2} \leq \varepsilon_2. \quad (44)$$

This implies that Eve's information about $b_A(m)$ and $b_B(m)$ adopted by the legitimate users is exponentially small:

$$I(A; E) \leq e^{-\eta_2} \leq \varepsilon_2, \quad I(B; E) \leq e^{-\eta_2} \leq \varepsilon_2, \quad (45)$$

where the pairs $\varepsilon_1$, $\eta_1$ and $\varepsilon_2$, $\eta_2$ may be mutually independent. The string of bits adopted as a key should not be interpreted as the bits in the pulse sent by Alice.

Each bit in the key is a function of the set of bits in Alice's pulse.

Now, let us formulate a protocol for secure key generation. (Recall that the length of the quantum communication channel must be known.) Note that a protocol may or may not require synchronization of Alice's and Bob's clocks.

1. Alice and Bob publicly choose states to be employed. A bandwidth $\Delta k$ is set (which corresponds to a certain time window $T$ since the channel length is known). A mean photon number $\mu$ is chosen, which automatically defines $\delta_E$. An $N \gg 1$ is chosen. (The protocol contains $2N$ bits.)

2. Alice generates a random string $b_i$ of bits of length $i = 1, \ldots, 2N$.

3. Bob performs measurements described by partition of unity (21) and publicly announces the fact of recording a count.

4. After all states prepared by Alice have been sent, Bob publishes the indices of the pulses for which measurement outcomes were obtained in the time windows $(t_i + L_{ch}, t_i + L_{ch} + T)$, i.e., without delay. (Here, $t_i$ is the $i$th instant at which the pulse was sent.) The pulses with outcomes outside the time windows are discarded. Suppose that the number of outcomes that pass the delay test is $2n$.

5. The partners randomly choose substrings of $n$ outcomes out of the $2n$ retained ones and publish the corresponding bit values.

6. Using a public channel, Alice and Bob compare their respective published substrings bit by bit and evaluate the error rate. Denote the number of equal bit values by $n_{OK}$ and the number of unequal bit values by $n_{\overline{OK}}$ (these numbers are equal). The error rate is evaluated as $\delta_{AB} = n_{\overline{OK}}/n$. When $n$ is sufficiently large, the error rate for the unpublished substrings is exponentially close to $\delta_{AB}$.

7. If $\delta_{AB} > \delta_E$, then the protocol is aborted, because the key cannot be extracted.

8. Alice and Bob use the classical error correction code $[n, k, d]$ to correct $t = \delta_{AB} n$ errors in the unpublished $n$-bit substrings. In this code, the number of code words is $2^k$, and the minimal distance between the code words is either $d > 2\delta_{AB} n + 1$ (for linear code) or $d > 2t + 1$ (for random linear code). To do this, Alice announces $\nu_i$ test strings of the code ($i = 1, \ldots, r$, $r = n - k$). Alice also publishes $r$ test parity bits,

$$\text{parity}_i = \nu_i n_A$$

($n_A$ and $n_B$ denote Alice's and Bob's unpublished substrings, respectively; they contain about $\delta_{AB} n$ unequal bits with probability close to unity).

9. Knowing the correct substring parity, Bob corrects errors in his string. As a result, Alice's and Bob's have identical strings with probability close to one if $n$ is sufficiently large.

10. The condition $\delta_{AB} < \delta_E$ guarantees that Eve has an exponentially small amount of information about the key after the correction of errors is performed by Alice and Bob by using a public channel (see below). Thus, the application of the error correction code ensures the identity and security of the key.

If $2n$ is sufficiently large (see step 4), then the evaluation of the error rate for the randomly chosen $n$-bit substring of the total string of length $n$ guarantees an arbitrarily close-to-unity value of the probability that the number of errors in the $n$ bits that have not been published by Alice and Bob is

$$t = n\delta_{AB}. \tag{46}$$

This makes it possible to use the linear error correction code $[n, k, d]$ with $d \geq 2t + 1$. The length of the final key obtained by using this code does not exceed

$$k < n(1 - H(\delta_{AB})),$$

$$H(x) = -x\log x - (1 - x)\log(1 - x),$$

where $H(x)$ is the binary entropy function. This is a necessary condition for correcting errors characterized by the probability $\delta_{AB}$ in a binary symmetrical setup. The code can be used to correct $(d - 1)/2$ bit values with a probability arbitrarily close to unity [41–43]. This result corresponds to the maximum Shannon information [41] for random linear codes.

For the best error-correction codes $[n, r, d]$ with $d/n \geq \delta_{AB}$, the value of the "code speed" $R = k/n$ does not fall below the Varshamov–Hilbert boundary [43]:

$$k \geq n\left(1 - H\left(\frac{d}{n}\right)\right) \tag{47}$$

(Shannon limit). However, the estimate that follows from the Varshamov–Hilbert inequality is more constructive. There exist linear regular (nonrandom) codes for which this boundary is achieved, in contrast to the Shannon limit. The latter can be achieved only for random codes and is essentially a theorem of existence, rather than a constructive limit: no regular codes are known for which this boundary can be achieved [41, 43].

The protocol is executed as long as the condition $\delta_{AB} < \delta_E$ is satisfied. The code used by Alice and Bob must be optimal in that it must correct all error strings with probabilities lower than $\delta_{AB}$, but fail to correct $\delta_E n$ errors.

If $\delta_{AB} < \delta_E$, then there exists a random error-correction code such that

$$d/n \geq \delta_{AB}, \text{ but } d/n < \delta_E, \tag{48}$$

which will correct Alice's and Bob's errors with a probability arbitrarily close to unity (provided that $n$ is sufficiently large to ensure reliability), but will not correct Eve's errors. As a result, Alice and Bob will retain a

nearly equal number $nR$ of identical bits if $n \gg 1$ such that

$$nR < nC(\delta_{AB}), \quad C(\delta_{AB}) = 1 - H(\delta_{AB}), \tag{49}$$

where $C(x)$ is the capacity of the classical symmetric binary communication channel and $R$ is the "speed" of the error-correction code.

After Alice and Bob have corrected their errors by using a public communication channel, the probability

$$\Pr\{\{b_A(i)\} \neq \{b_B(i)\}$$

that the strings

$$\{b_{A,B}(i)\} = (b_{A,B}(1), \ldots, b_{A,B}(n))$$

are different is

$$\begin{aligned} &\Pr\{\{b_A(i)\} = \{b_B(i)\}\} \\ &> 1 - 4\exp[-nE(\delta_{AB}, z)], \\ &E(\delta_{AB}, z) = \max_{0 \le z \le 1}\{zR - E_0(\delta_{AB}, z)\}, \end{aligned} \tag{50}$$

where

$$\begin{aligned} E_0(z, \delta_{AB}) &= z\ln 2 - (1+z) \\ &\times \ln[\delta_{AB}^{1/(1+z)} + (1-\delta_{AB})^{1/(1+z)}]. \end{aligned} \tag{51}$$

Thus, identity conditions (42) and (43) for the keys possessed by the legitimate partners are satisfied. It remains to show that the key remains secure, i.e., the probability that Eve knows the corrected string exponentially decreases with increasing $n$.

When the condition $\delta_{AB} < \delta_E$ is satisfied, Eve deals with a binary symmetrical setup [42], for which the data rate (measured in bits per pulse) exceeds the capacity $C(\delta_E)$ of the channel connecting her with Alice. When $\delta_{AB} < \delta_E$, she perceives error correction as data transmission with a data rate higher than the capacity of the channel connecting her with Alice, whereas the capacity $C(\delta_{AB})$ of the channel connecting the legitimate partners exceeds $C(\delta_E)$. Error correction performed by using codes is perceived as data transmission with a data rate $R < C(\delta_{AB})$, which exceeds the capacity of the channel connecting Eve with the legitimate partners ($R > C(\delta_E)$).

When $R > C(\delta_E)$, the following theorem on the average probability of error per bit is valid [44].

**Theorem.** *If a discrete memoryless channel has a capacity $C$ and $R > C$ for any $(n, R)$-code, then the error probability $P_e$ is such that*

$$P_e \ge 1 - \frac{4A}{n(R-C)^2} - \exp\left(-\frac{n(R-C)}{2}\right), \tag{52}$$

*where $A$ is a positive constant depending on the channel characteristics and independent of $R$ and $n$.*

This bound can be improved [42]: for any memoryless channel,

$$P_e \ge 1 - 2\exp(-n\alpha(R)), \tag{53}$$

where

$$\begin{aligned} &\alpha(R) \\ &= \min\left\{\frac{R-C}{2}, \lim_{s \ge 0}\left[s\left(C + \frac{R-C}{2} - \ln g(s)\right)\right]\right\}, \end{aligned} \tag{54}$$

and

$$g(s) = \sum_{j,k} Q(k)P(j|k)\exp\left\{s\ln\frac{P(j|k)}{\sum_i Q(i)P(j|i)}\right\}. \tag{55}$$

For a discrete binary symmetric channel,

$$\begin{aligned} &Q(i) = \frac{1}{2}, \quad P(j|j) = 1 - \delta_E, \\ &P(j|k) = 1 - \delta_E, \quad j \neq k, \quad j, i, k = 0, 1. \end{aligned} \tag{56}$$

$C = C(\delta_E)$ and $R \approx C(\delta_{AB})$ in (52)–(55). Thus,

$$\begin{aligned} R \approx C(\delta_{AB}) &= 1 + (1-\delta_{AB})\log(1-\delta_{AB}) \\ &+ \delta_{AB}\log\delta_{AB}, \quad R < C, \end{aligned} \tag{57}$$

$$\begin{aligned} C &= C(\delta_E) \\ &= 1 + (1-\delta_E)\log(1-\delta_E) + \delta_E\log\delta_E. \end{aligned} \tag{58}$$

The function $g(s)$ is expressed as

$$\ln g(s) = (s-1)\ln 2 + \ln[\delta_E^{s+1} + (1-\delta_E)^{s+1}], \tag{59}$$

and the exponent in (53) is on the order of $n(R - C)$.

As a result, security criterion (44), (45) is satisfied: Eve has an exponentially small amount of information about the key:

$$\begin{aligned} &\Pr\{\{b_{A,B}(i)\} = \{b_E(i)\}\} \\ &< 2\exp(-n\alpha(R)) \approx 2\exp(-n(C(\delta_{AB}) - C(\delta_E))). \end{aligned} \tag{60}$$

Note that the BB84 protocol [45] guarantees key security in the Shannon limit only if the error rate is lower than 11% [20, 21]. This limit is due to the fact that phase errors arising in measurements performed in different bases must be corrected, as well as bit flips [21]. Accordingly, the limit is determined by solving the equation

$$1 = 2H(\delta_{AB}). \tag{61}$$

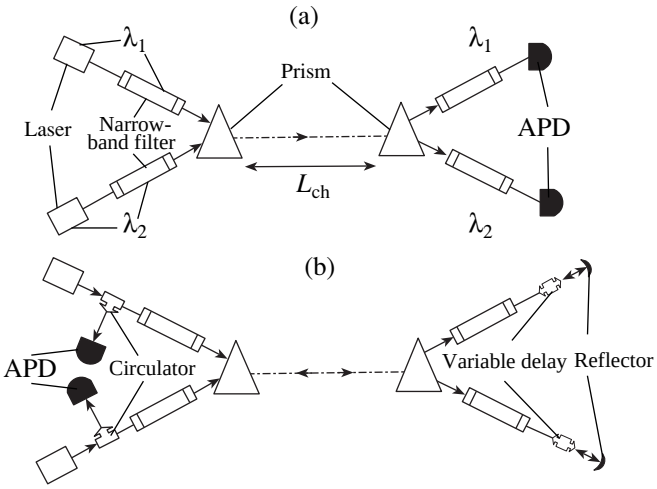



**Fig. 4.**

Even though Eve's information about the bits retained by Alice and Bob after performing the error correction has been performed tends to zero for $n \gg 1$, a privacy amplification procedure can be used to enhance security.

## 8. QUANTUM CRYPTOSYSTEM WITH CLOCK SYNCHRONIZATION AT THE SENDER AND RECEIVER ENDS

The cryptosystem described in this section requires synchronization of the clocks used at the sender and receiver ends, i.e., precise timing of each pulse.

The other scheme discussed below does not require synchronization. It suffices that two identical clocks (pulse generators) are started independently at the sender and receiver ends and the communication channel length is not required either.

Suppose that the communication channel length $L_{ch}$ is known, and the clocks used by the partners are synchronized (say, by using a public classical channel).

The system is schematized in Fig. 4a. The time evolution of states is illustrated by Fig. 5a, where pulses are depicted after they have been filtered. The spectra of the laser pulses emitted at the sender end are centered at different wavelengths $\lambda_1$ and $\lambda_2$ and do not overlap. The execution of the protocol is started at a certain instant by randomly triggering one of the lasers. The laser emits a pulse with a wide bandwidth $\Delta\omega_1$ and a short duration $\Delta t \sim 1/\Delta\omega_1$. The triggering instant is subsequently announced by using a public classical commu-

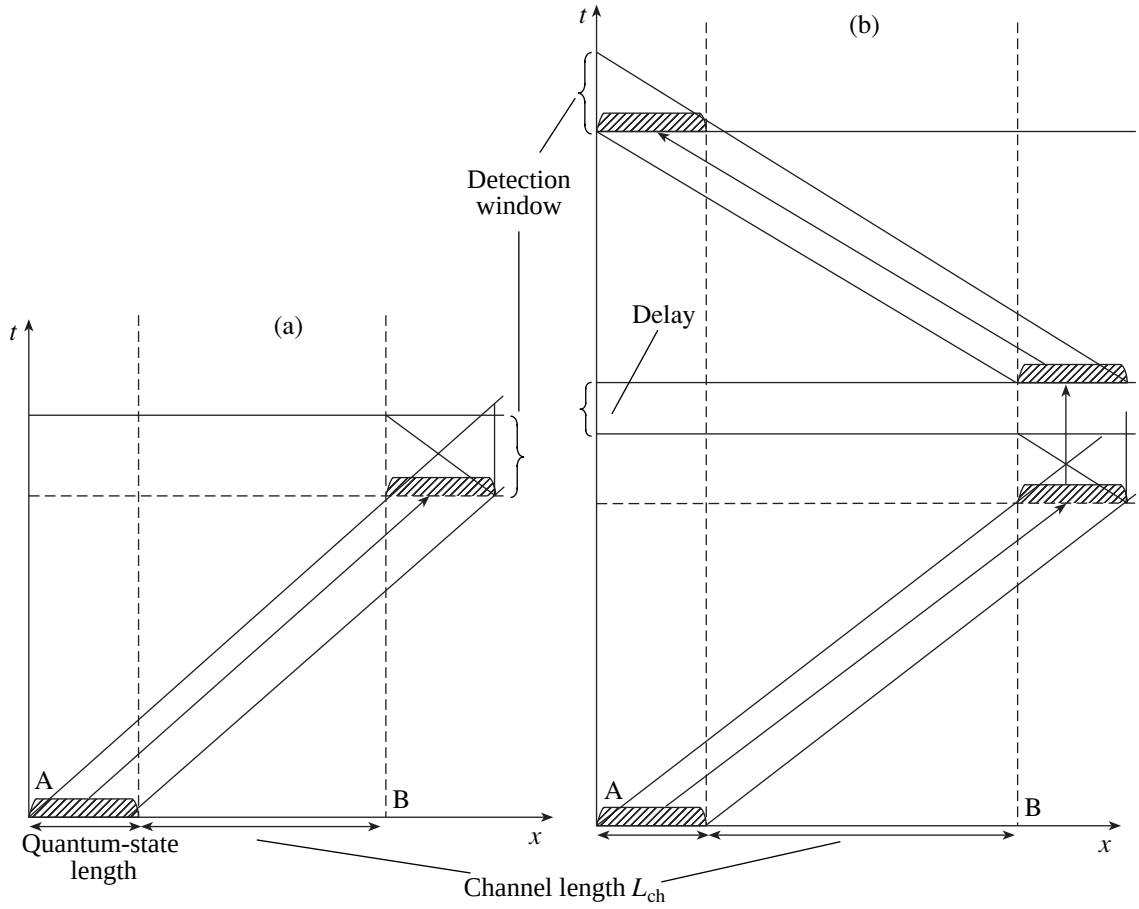



**Fig. 5.**

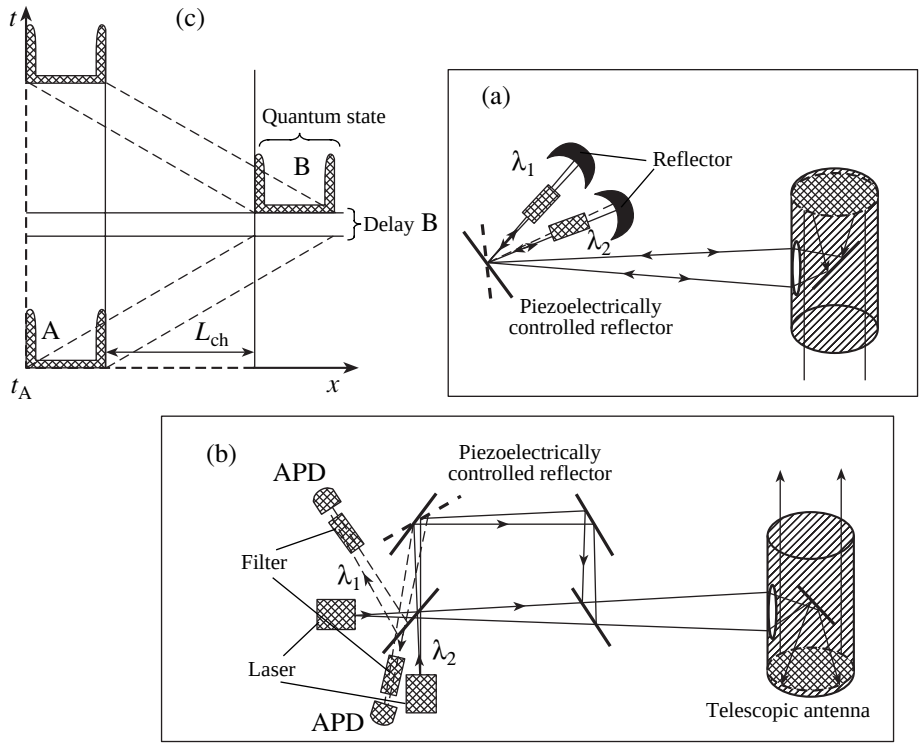



**Fig. 6.**

nication channel. The clocks are synchronized within $\Delta t$. Each filter (see Fig. 4a) reduces the broad laser spectrum to a narrow frequency band of width $\Delta k$. If the laser power corresponding to the starting spectral width $\Delta\omega$ is known and controlled, then the power corresponding to $\Delta k$ can be evaluated to find the mean number $\mu$ of filtered photons injected into the communication channel. The prism shown in Fig. 4 is required to deflect pulses centered at different wavelengths.

The narrow frequency band $\Delta k$ is associated with the spacetime interval $\Delta T \sim 1/\Delta k$ required for a pulse of length $c\Delta T$ to enter the communication channel entirely. The time when the leading front enters the channel is known within $\Delta t$. Accordingly, the duration of a pulse with a narrow bandwidth $\Delta k$ must be much greater than error $\Delta t$ of synchronization of the clocks used at the sender and receiver ends. The channel length, pulse length $\Delta T$, and the instant of leading-front entry can be used to calculate the time required for a pulse to reach the receiver end. The signals received by the avalanche photodiodes (APDs) at the receiver end pass through filters identical to those used at the sender end to ensure that pulses shorter than those corresponding to the bandwidth $\Delta k$ do not reach the APDs. The APDs operate in the so-called Geiger mode. The APD recovery time $\tau_D$ must be such that $\tau_D \ll \Delta T$. Bob accepts only pulses that trigger the APDs in the time window $[L_{ch}, L_{ch} + \Delta T]$ (see Fig. 5a). The emission time for the leading front is $t_A = 0$ (see Fig. 5a). Filtering at the receiver end is required to rule out the possibility of sending pulses with shorter duration (wider bandwidth) by the eavesdropper, which would compensate for the delay induced by measuring a pulse of duration $T$.

## 9. DOUBLE-PASS QUANTUM CRYPTOSYSTEM WITHOUT CLOCK SYNCHRONIZATION AT THE SENDER AND RECEIVER ENDS

The double-pass scheme shown in Fig. 4b differs from that described above in that it does not require synchronization. It suffices to know the length of the communication channel. States are prepared as in the other scheme. The channel length and the moment of pulse emission can be used to calculate the moment when the pulse returns to Alice. Thus, Alice can use her knowledge of the time window for the returning pulse to reveal Eve's presence, as illustrated by Fig. 5b.

Operating at the reflector end, Bob randomly opens one of the reflector channels, keeping the other channel shut. Alice uses the open channel to announce only the pulses that have returned to her and detected in the respective correct time windows. If the channel opened by Bob does not match the pulse sent by Alice, then the pulse is discarded, because it will not return to Alice. If Bob opens the right channel and then Alice announces that the pulse has been detected in the correct time window, then both partners know the transmitted bit.

If the leading front of the pulse enters the channel at a moment $t_A$, then Alice must accept only counts that fit in the time window ($t_A + 2L_{ch} + t_{B\,delay}$, $t_A + 2L_{ch} +$

$t_{\text{B delay}} + T$), where $t_{\text{B delay}}$ is a randomly varied delay introduced by Bob and subsequently specified for each pulse by using an open channel (see Fig. 5b). The use of a random delay is essential for security of the protocol.

Finally, let us discuss a prototype system for key distribution to an orbiting object (see Fig. 6, where tracking and positioning schemes are not shown to simplify presentation).

Technically, it is easier to construct "extended" states from short-duration pulses than prepare states with narrow bandwidths. This is done by combining an arm of an unbalanced interferometer with two beamsplitters. Figure 6a shows a double-pass scheme without clock synchronization at the two ends. "Short" states with pulse durations on the order of 1 ns are created by using two lasers with different wavelengths that match atmospheric windows, such as 850 and 1000 nm, which are characterized by transmittance of 74 to 75%. The triggering instant for one of the lasers is measured up to the pulse duration (say, 1 ns). The emitted pulse enters an arm of an unbalanced interferometer with a path difference greater than the length of the input pulse (the pulse does not interfere with itself). The output is the superposition of two "short" states with separation (output length) greater than the error in measuring the distance between the ground-based and orbiting objects (see Fig. 6c). After the pulse is focused by means of a telescopic system, it is sent to the orbiting object. The pulse is received by the telescopic system at the receiver end (see Fig. 6b) and directed to a piezoelectrically controlled mirror and reflectors. One of the two mirror positions is randomly set and is independent of the sender parameters. In this scheme, a random delay is also required at Bob's end and the delay time is announced by Bob by using a public channel after a string of bits has been transmitted. If the emitted state characterized by $\lambda_1$ matches the reflector position (set to reflect a state characterized by $\lambda_1$), then the state returns to the sender end. Otherwise, the pulse is empty and the state does not return to the sender end. The position of the controlled reflector used at the sender end depends on the transmitted state. After the state with $\lambda_1$ has returned, the reflector position is such that the pulse is directed to an APD tuned to $\lambda_1$. The pulse consists of three "spikes" (not shown in Fig. 6b) separated by the path difference between the interferometer arms, the amplitude of the central spike being twice as large as those of the leading and trailing "spikes." The APD randomly fires when triggered by one of the "spikes," and the triggering instant is recorded.

Procedures analogous to those described above are used to extract a key from the transmitted string and detect possible eavesdropping attempts.

An "extended" state consisting of two "halves" is advantageous as compared to a narrow-bandwidth state, because it is easier to implement in practice and to detect with an APD, which must be kept under reverse bias only during the short intervals when the "halves" reach the APD, so that the dark count rate is substantially reduced.

Let us perform some numerical estimates. For a satellite orbit radius $L_s \approx 1000$ km $= 10^8$ cm, a satellite speed $v_s \approx 5$ km/s $= 5 \times 10^5$ cm/s, and a speed of light $c = 3 \times 10^{10}$ cm/s, the error in measuring the distance between the ground-based object and the satellite by means of a classical signal is estimated as the shift in the satellite's position over the time interval required for the probe signal to reach the satellite:

$$\frac{L_s v_s}{c} = \frac{10^8 \times 5 \times 10^5}{3 \times 10^{10}} \approx 1.7 \times 10^3 \text{ cm} = 17 \text{ m}.$$

Therefore, the length of the "extended" state consisting of two "halves" must be greater than 17 m. Suppose that $L = 300$ m $= 3 \times 10^4$ cm. The shift is the satellite's position required for both "halves" to be reflected by Bob's reflectors is

$$\frac{L v_s}{c} = \frac{3 \times 10^4 \times 5 \times 10^5}{3 \times 10^{10}} = 5 \times 10^{-1} \text{ cm} = 5 \text{ mm}.$$

Therefore, no additional adjustment of the reflector is required during the reflection time if the diameter of the telescopic antenna is about 10 cm. The distance between the orbiting and ground-based objects can be measured by a third party with the use of GPS or the Russian GLONASS system.

Since this cryptosystem makes use of orthogonal states nonoverlapping spectra, the scheme can be generalized by introducing frequency multiplexing, in which case the key generation rate is proportional to the number of channels in the atmospheric window.

## 10. CONCLUSIONS

Attenuation in quantum communication channels presents the principal difficulty for security of nonrelativistic quantum cryptography schemes, which relies on complete indistinguishability of nonorthogonal states. Nonrelativistic key distribution protocols cannot guarantee unconditional security (dictated only by the laws of quantum mechanics) if attenuation exceeds some critical level (which cannot be accurately evaluated).

An analysis of quantum cryptographic protocols and the proofs of their security shows that security problems arise when the eavesdropper can "replace" the available attenuating channel with a channel characterized by lower attenuation. It is important that this can be done in both fiber-optic and free-space systems without installing, say, a new cable. The eavesdropper can follow a simple strategy of channel "replacement" equivalent to laying out a new fiber, which requires tapping into the channel at two points located near the sender and receiver ends, where the detector and transmitter are installed, respectively. In this setting, attenuation in the links between the eavesdropper and the legitimate partners can be neglected. Near the sender,

the eavesdropper performs measurements analogous to those described in Section 2. If a conclusive result (0 or 1) is obtained, then the eavesdropper sends this information to tapping location near the receiver over a public classical channel, which can always be treated as ideal. Another eavesdropper uses the information received over the classical channel to prepare a quantum state ($|\varphi_0\rangle$ or $|\varphi_1\rangle$) and sends it to the legitimate receiver over the fiber. If an inconclusive result "?" is obtained (see (3)), then no quantum state is sent to the legitimate receiver from the nearest tapping location. Since the quantum channel is attenuating, the presence of the eavesdropper cannot be revealed if the channel length exceeds a certain critical value. It is obvious that a similar strategy can be used in free-space cryptography without improving any properties of the transmitting medium (atmosphere).

Another principal difficulty is the unavailability of single-photon sources. Even though pure single-photon states are not required, the measurement procedure used at the receiver end must identify a pair of multiphoton states as a whole. Currently, the only suitable method is homodyne detection, which can be represented as projection onto a coherent state. However, this method is very complicated, and the cryptographic stability of key distribution protocols using multiphoton states is poorly studied. When a standard avalanche photodiode is employed as a detector, the use of a single-photon source is essential. Even if ideal single-photon sources were available (some progress in construction of their laboratory prototypes was reported in [16]), the problems due to attenuation would not be eliminated.

It is argued here that the attenuation problem in quantum cryptosystems relying only on geometric properties (complete indistinguishability) of nonorthogonal quantum states cannot be solved by using any tricks or complicated protocols. Additional fundamental laws must be invoked to ensure security for any attenuation. The relativistic causality principle provides the desired solution. The limitations for measurability and data rate imposed by special relativity theory remain valid at any level of attenuation. Otherwise, it would be possible to exchange classical information between remote observers faster than at the speed of light.

Additional fundamental physical principles are invoked to formulate a new approach and ensure unconditional security of quantum cryptosystems by eliminating the aforementioned difficulties.

Thus, relativistic quantum cryptosystems remain unconditionally secure: first, any attenuation in a communication channel only reduces the key generation rate without affecting its security; second, the source may not generate pure single-photon states and a nonzero single-photon probability will suffice. The scheme remains secure even if the contribution of a single-photon component is arbitrarily small. This formally implies that a state may be characterized by an arbitrarily large mean photon number. The single-photon probability affects only the key generation rate, but not security.

This work was supported by Academy of Cryptography and the Russian Foundation for Basic Research, project no. 02-02-16289.

## REFERENCES

1. V. A. Kotel'nikov, A Report (Moscow, 1941).
2. C. E. Shannon, Bell Syst. Tech. J. **28**, 658 (1949).
3. G. S. Vernam, J. Am. Inst. Elect. Eng. **55**, 109 (1926).
4. H. Kosaka, A. Tomita, Y. Nambu, *et al.*, quant-ph/0306066.
5. D. Stucki, N. Gisin, O. Guinnard, *et al.*, quant-ph/0203118.
6. A. Muller, J. Breguet, and N. Gisin, Europhys. Lett. **23**, 383 (1993); A. Muller, H. Zbinden, and N. Gisin, Nature **378**, 449 (1995); A. Muller, H. Zbinden, and N. Gisin, Europhys. Lett. **33**, 335 (1996).
7. Ch. Marand and P. D. Tounsend, Opt. Lett. **20**, 1695 (1995); P. D. Tounsend, Nature **385**, 47 (1997); IEEE Photonics Technol. Lett. **10**, 1048 (1998).
8. R. Hughes, G. G. Luther, G. L. Morgan, and C. Simmons, Lect. Notes Comput. Sci. **1109**, 329 (1996); R. Hughes, G. Morgan, and C. Peterson, J. Mod. Opt. **47**, 533 (2000).
9. P. C. Sun, Y. Mazurenko, and Y. Fainman, Opt. Lett. **20**, 1062 (1995); Y. Mazurenko, R. Giust, and J. P. Goedgebuer, Opt. Commun. **133**, 87 (1997).
10. F. Grosshans, G. Van Assche, J. Wenger, *et al.*, Nature **421**, 238 (2003).
11. M. Martinelli, Opt. Commun. **72**, 341 (1989); J. Mod. Opt. **39**, 451 (1992).
12. C. Elliot, D. Pearson, and G. Troxel, quant-ph/0307049.
13. C. Kurtsiefera, P. Zarda, M. Halder, *et al.*, Preprint (Los Alamos, 2002).
14. R. J. Hughes, J. E. Nordholt, D. Derkacs, and C. G. Peterson, New J. Phys. **4**, 43.1 (2002).
15. J. G. Rarity, P. R. Tapster, P. M. Gorman, and P. Knight, New J. Phys. **4**, 82.1 (2002).
16. A. Beveratos, R. Brouri, T. Gacoin, *et al.*, quant-ph/0206136.
17. N. Lütkenhaus, Phys. Rev. A **61**, 052304 (2000); G. Brassard, N. Lütkenhaus, T. Mor, and B. C. Sanders, Phys. Rev. Lett. **85**, 1330 (2000).
18. A. Acin, N. Gisin, and V. Scarani, quant-ph/0302037.
19. W. K. Wootters and W. H. Zurek, Nature **299**, 802 (1982).
20. C. H. Bennett, Phys. Rev. Lett. **68**, 3121 (1992); C. H. Bennett, G. Brassard, and N. D. Mermin, Phys. Rev. Lett. **68**, 557 (1992).
21. D. Mayers and A. Yao, quant-ph/9802025.
22. P. W. Shor and J. Preskill, quant-ph/0003004.
23. W. Diffi and M. E. Hellman, IEEE Trans. Inf. Theory **22**, 644 (1976).
24. R. L. Rivest, A. Shamir, and L. Adleman, Commun. ACM **21**, 120 (1978).

25. N. N. Bogolyubov and D. V. Shirkov, *Introduction to the Theory of Quantized Fields*, 3rd ed. (Nauka, Moscow, 1973; Wiley, New York, 1980).
26. L. D. Landau and R. Peierls, Z. Phys. **69**, 56 (1931); L. D. Landau, in *Collected Works* (Nauka, Moscow, 1969), Vol. 1, p. 56; L. D. Landau and R. Peierls, Z. Phys. **62**, 188 (1930); L. D. Landau, in *Collected Works* (Nauka, Moscow, 1969), Vol. 1, p. 33.
27. N. Bohr and L. Rosenfeld, Mat. Fys. Medd. K. Dan. Vidensk. Selsk. **12**, 3 (1933); N. Bohr, in *Collected Works* (Nauka, Moscow, 1969), Vol. 1, p. 39.
28. M. Fleischhauer and M. D. Lukin, Phys. Rev. Lett. **84**, 5094 (2000).
29. W. H. Fuchs, J. Math. Anal. Appl. **9**, 317 (1964).
30. D. Slepian and H. O. Pollak, Bell Syst. Tech. J. **40**, 40 (1961).
31. N. N. Bogolyubov, A. A. Logunov, A. I. Oksak, and I. T. Todorov, *General Principles of Quantum Field Theory* (Nauka, Moscow, 1987; Kluwer Academic, Dordrecht, 1990).
32. A. M. Jaffe, Phys. Rev. **158**, 1454 (1967).
33. G. C. Hegerfeldt, Phys. Rev. D **10**, 3320 (1974); G. C. Hegerfeldt and S. N. M. Ruijsenaar, Phys. Rev. D **22**, 377 (1980).
34. D. A. Kirzhnits, Usp. Fiz. Nauk **90**, 129 (1966) [Sov. Phys. Usp. **9**, 692 (1966)].
35. N. Wiener and R. Paley, *Fourier Transform in the Complex Domain* (Am. Math. Soc., New York, 1934; Nauka, Moscow, 1964).
36. I. Bialynicki-Birula, Phys. Rev. Lett. **80**, 5247 (1998).
37. T. D. Newton and E. P. Wigner, Rev. Mod. Phys. **21**, 400 (1949).
38. A. Einstein, B. Podolsky, and N. Rosen, Phys. Rev. **47**, 777 (1935).
39. E. Biham, M. Boyer, P. O. Boykin, *et al.*, quant-ph/9912053.
40. Hoi-Kwong Lo and H. F. Chau, quant-ph/9803006.
41. C. E. Shannon, Bell Syst. Tech. J. **27**, 397 (1948); Bell Syst. Tech. J. **27**, 623 (1948).
42. R. Gallagher, *Information Theory and Reliable Communication* (Wiley, New York, 1968; Sovetskoe Radio, Moscow, 1974).
43. E. J. Mac Williams and N. J. A. Sloane, *The Theory of Error-Correcting Codes* (North-Holland, Amsterdam, 1977).
44. J. Wolfowitz, Ill. J. Math. **1**, 591 (1957).
45. C. H. Bennett and G. Brassard, in *Proceedings of IEEE International Conference on Computer Systems and Signal Processes* (Bangalore, India, 1984), p. 175.

*Translated by A. Betev*

*Journal of Experimental and Theoretical Physics, Vol. 99, No. 4, 2004, pp. 690–707.*
*Translated from Zhurnal Éksperimental'noĭ i Teoreticheskoĭ Fiziki, Vol. 126, No. 4, 2004, pp. 799–818.*


**ATOMS, SPECTRA, RADIATION**

# Cascade Processes of Energy Loss by Emission of Hard Phonons

**M. Kh. Khokonov**

*Kabardino-Balkar State University, Nalchik, 360004 Russia*
*e-mail: khokon6@mail.ru*

Received March 26, 2004

**Abstract**—The method of cascade equations is used for studying the dynamics of the energy loss process (struggling) during passage of particles through a substance. Special attention is paid to multiphoton processes of emission of relativistic electrons in amorphous media and in oriented crystals. New analytic solutions to cascade equations are obtained and a method is developed for solving these equations, which makes it possible to express the results in terms of rapidly converging integrals. It is shown that the relatively simple Landau equation for the energy loss distribution function can be used for analyzing processes of energy loss commensurate with the initial energy of particles. The experimental data for 150-GeV electrons in oriented crystals are thoroughly analyzed. © *2004 MAIK "Nauka/Interperiodica".*

## 1. INTRODUCTION

An ultrarelativistic electron in a medium or in an external field can emit photons with an energy $\omega$ commensurate with the initial electron energy $E_0$. In this case, such quantities as mean radiation loss $\langle \Delta E \rangle$ or even mean square fluctuation $\langle (\Delta E)^2 \rangle$ fail to provide a correct pattern of energy loss due to a strong electron energy spread (struggling) at the output. Under these conditions, the energy loss distribution function $p(y, t)$ is required, where $y$ is the energy loss and $t$ is the time. This problem was formulated in such a form long ago by Bethe and Heitler in their pioneering work [1], in which simple but rather approximate expressions for $p(y, t)$ were derived. Another important case concerns ionization energy loss in relatively thin targets, when the energy loss is much smaller than the beam energy. Landau [2] proved that distribution $p(y, t)$ in this case can be determined from an integral equation leading to a standard distribution function. The problem is formulated as follows: knowing the differential probability of interaction over an infinitely short interval of time (path length) $\nu(y)dt$, find the probability $p(y, t)$ corresponding to a finite time interval. The Landau solution corresponds to the Rutherford differential cross section of energy loss per unit length, when $\nu(y) \sim 1/y^2$. In this case, the limitation on small energy loss (for $y \longrightarrow 0$) was introduced.

In applications to radiation problems, $\nu(y)$ is the emission probability per unit time. In the case of bremsstrahlung in an amorphous medium, this quantity is $\nu = \sigma_{BH}Nc$, where $\sigma_{BH}$ is the Bethe–Heitler cross section [1], $N$ is the number of atoms of the media per unit volume, and $c$ is the velocity of light. However, the value of $\nu$ for radiation processes can not always be defined in this way. The classical formulas, which are spectral–angular characteristics of radiation, are functionals of the electron trajectory and determine the characteristics of radiation over the entire time of interaction rather than per unit path length. Schwinger [3] derived the emission probabilities per unit time; however, the corresponding expressions for each instant depend on the past history and future of an electron; consequently, in the context being discussed, such an approach is applicable only in the constant field approximation (CFA), in which the formulas describing emission can be reduced to familiar expressions for synchrotron radiation (or their semiclassical generalizations). In this connection, the approach developed in [4] appears promising, since it makes it possible to express the characteristics of radiation in a nonuniform external field in terms of local interaction parameters (i.e., the field potential and its derivatives) rather than in terms of the trajectory. Nevertheless, we will confine our analysis of multiphoton processes of energy loss by radiation in oriented crystals to the CFA, assuming that the mean free pass $l \approx v/\nu$ between consecutive emission events exceeds the radiation formation length ($v$ is the electron velocity that approximately equals the velocity of light).

Lindhard and Nielsen [5, 6] thoroughly analyzed the main properties of integral equations for distribution function $p(y, t)$ and obtained analytic solutions for a number of important cases. In particular, they considered certain types of exponential cross sections with screening. Lindhard [7] developed an approach for obtaining an approximate solution to the Landau–Vavilov equations [8] with an error smaller than 2% in the second approximation. We present here a number of

new analytic solutions for $p(y, t)$, consider their applications to the problems of emission of electrons with energies 150–240 GeV in oriented crystals, and compare the results with the experimental data [9–12]. It is shown that a relatively simple Landau equation can be useful in the case when it seems to be inapplicable (i.e., when $y \sim E_0$).

Another aspect of the problem is the study of the multiphoton nature of radiation. Since all photons emitted by an electron arrive at a detector almost simultaneously and are indistinguishable, information on the initial one-photon radiation cross sections cannot be obtained from direct measurements. Nevertheless, we will show that the probabilities of processes and distribution functions for measurable quantities can be calculated.

Baier and Katkov [13] have recently studied in detail the multiphoton nature of radiation and its influence on the spectra being measured; in particular, an important case of bremsstrahlung in an amorphous medium was considered. Multiphoton processes accompanying the passage of electrons through oriented crystals were studied in detail by numerical simulation methods [14–19]. Some aspects of struggling (energy loss) were considered in [20]. The effect of distortion of the measured emission spectra in the case of channeling was studied in [21] for photons with energies much smaller than the electron energy. However, we propose an approach based on cascade-type equations. The results are applicable for studying the emission of hard photons as well as ionization energy loss. Some important probabilistic and mathematical aspects of the problems considered below are described in [22]. Here, we give new convenient representations of some special functions that play an important role in problems of this type.

## 2. POISSON-TYPE ENERGY LOSS PROCESS

We will apply the term Poisson-type processes to processes in which the statistics of individual events of interaction accompanied by energy loss for an electron passing through a medium can be described by the Poisson distribution. Our analysis is based on the solution of integral equations of motion determining the time evolution of energy loss distribution function $p(y, t)$, where spatial variable $y$ characterizes the energy loss. This is not necessarily the energy lost in the process, but a quantity that is in one-to-one correspondence with this energy. This problem is formulated in general form in [5, 6], where analytic solutions are obtained for a number of important cases. It will be shown below that an effective procedure for obtaining a wide class of solutions to the integral equations of motion can be constructed and new analytic solutions will be given for a practically important case of the exponential type of differential interaction cross sections with screening.

### *2.1. Basic Equation of Motion*

The integral equation describing the evolution of the energy loss distribution function was introduced by Landau [2] in the form

$$\frac{\partial p(y, t)}{\partial t} = \int_0^\infty [p(y-\mu, t) - p(y, t)]\nu(\mu)d\mu, \qquad (2.1)$$

where $p(y-\mu, t) = 0$ for $\mu > y$ and the kernel $\nu(\mu)d\mu$ is independent of $y$ and $t$ and is equal to the cross section of interaction per unit time, which leads to a decrease in the particle energy. This means that quantity $\nu(y)dy\,dt$ can be treated as the differential probability of interaction with a preset change in parameter $y$, referred to an infinitely small time interval $dt$, while distribution function $p(y, t)$ is the probability density corresponding to a finite time interval. In Eq. (2.1), we assume that the energy loss is much smaller than the initial energy of the particle. However, this circumstance is immaterial for further analysis. Distribution $p(y, t)$ is normalized,

$$\int_0^\infty p(y, t)dy = 1, \qquad (2.2)$$

and satisfies the initial condition $p(y, t = 0) = \delta(y)$, where

$$\nu(y) = \begin{cases} \nu(y), & y > 0 \\ 0, & y < 0. \end{cases} \qquad (2.3)$$

The differential cross sections of the equations describing energy degradation satisfy this condition in contrast to equations describing, for example, processes of the type of multiple scattering.

It will be shown below that relatively simple equation (2.1) permits generalizations that make it possible to analyze a wide class of much more complicated equations. We will distinguish between two types of differential cross sections $\nu(y)$ with finite and diverging values of the total cross section. The first type of cross sections mainly corresponds to processes of energy loss by emission of hard photons, while the second type corresponds to ionization losses in a substance.

Let the total interaction cross section be finite,

$$\int_0^\infty \nu(y)dy = \nu_0. \qquad (2.4)$$

In this case, we can introduce a function $g(y)$, such that $\nu(y) = \nu_0 g(y)$ and normalized to unity,

$$\int_0^\infty g(y)dy = 1. \qquad (2.5)$$

Then we can write Eq. (2.1) in the form

$$\frac{\partial p(y,\xi)}{\partial \xi} = -p(y,\xi) + \int_0^\infty p(y-\mu,\xi)g(\mu)d\mu, \qquad (2.6)$$

where the time variable $\xi = \nu_0 t$ has been introduced.

In the case considered here, total cross section $\nu_0$ for an individual interaction is a time-independent constant. Consequently, we assume that the distribution over individual events of interaction is a Poisson distribution; i.e., we assume that the probability that exactly $m$ interactions occur during a finite time interval $t$ is determined by the Poisson distribution $P_m(\xi)$. In this case, quantity $\xi = \nu_0 t$ has the meaning of the mean number of interaction events during time $t$, which lead to discrete events of energy loss. After the occurrence of $m$ events, the value of variable $y$ becomes $y = y_1 + y_2 + \ldots + y_m$. We introduce the probability density $g^{(m)}(y)$ for the fact that, after $m$ interactions, a finite value of $y$ belongs to the interval $(y, y + dy)$. Obviously, $g^{(1)}(y) = g(y)$.

It can be verified by direct substitution that the solution to Eq. (2.6) under the above initial condition has the form

$$p(y,\xi) = \sum_{m=0}^\infty P_m(\xi)g^{(m)}(y), \qquad (2.7)$$

where $P_m(\xi)$ is the Poisson distribution,

$$P_m(\xi) = \frac{\xi^m}{m!}\exp(-\xi). \qquad (2.8)$$

Thus, for Poisson-type processes, the temporal (depending on $\xi$) and spectral (depending on $y$) parts in Eq. (2.7) can be described independently.

Functions $g^{(m)}(y)$ in Eq. (2.7) are normalized to unity and satisfy the recurrence relations

$$g^{(m)}(y) = \int_0^y g^{(m-n)}(y-\mu)g^{(n)}(\mu)d\mu, \quad n \le m, \qquad (2.9)$$

where $g^{(0)}(y) = \delta(y)$ and $g^{(1)}(y) = \delta(y)$.

Partial probability distributions of type (2.9) can be calculated using the Fourier transformation,

$$g^{(m)}(y) = (2\pi)^{-1}\int_{-\infty}^\infty (g_k)^m \exp(iky)dk, \qquad (2.10)$$

where

$$g_k = \int_0^\infty g(\mu)\exp(-ik\mu)d\mu. \qquad (2.11)$$

Substituting expressions (2.10) and (2.8) into Eq. (2.7), after summation we arrive at the known representation of the solution to Eqs. (2.1) and (2.6) in the form [5]

$$p(y,\xi) = (2\pi)^{-1}\int_{-\infty}^\infty \exp(iky - \xi\lambda(k))dk, \qquad (2.12)$$

where

$$\lambda(k) = 2i\int_0^\infty \sin\frac{k\mu}{2}\exp\left(-\frac{ik\mu}{2}\right)g(\mu)d\mu. \qquad (2.13)$$

Consequently, we have verified that relation (2.7) indeed satisfies Eq. (2.1) and the known expressions (2.12) and (2.13) correspond to Poisson-type processes, which is not seen directly from expressions (2.12) and (2.13). In contrast to expression (2.7), a solution in the form (2.12) and (2.13) is also valid in the case when integral (2.4) diverges and Poisson distribution (2.8) is meaningless. Nevertheless, processes satisfying Eq. (2.1) will also be treated as Poisson-type processes.

Equation (2.1) describes energy-loss processes (Landau case) only for asymmetric differential cross sections (2.3). If a differential cross section is symmetric (i.e., if $\nu(-y) = \nu(y)$), Eq. (2.1) has the form of a Mollier equation [23] describing the processes of multiple scattering of fast charged particles in a substance. In this case, $p(y, t)$ is the angular distribution at instant $t$. It should be noted that formulas resembling Eqs. (2.7)–(2.9) also take place for multiple scattering [24]. These formulas are of practical importance in studying multiple scattering through relatively large angles, when the number of individual scattering events does not exceed 5–10. A detailed analysis for the case of symmetric nuclei is given in [5].

### *2.2. Finite Cross Section of Power Type with Screening*

A wide class of differential cross sections, which lead to analytic solutions and are important for practical application, can be described by a power function with an exponential cutoff parameter (screening),

$$g(y) = Cy^{\alpha-1}\exp(-\lambda y), \qquad (2.14)$$

where the normalization constant $C = \lambda^\alpha/\Gamma(\alpha)$, $\alpha > 0$, and $\lambda > 0$. In this case, from expressions (2.10) and (2.11), we obtain

$$g^{(m)}(y) = \frac{\lambda^{\alpha m}}{\Gamma(\alpha m)}y^{\alpha m-1}\exp(-\lambda y). \qquad (2.15)$$

Then the solution to Eq. (2.1) has the form

$$p(y, \xi) = y^{-1} \sum_{m=1}^{\infty} \frac{x^m}{m!\Gamma(\alpha m)} \exp(-\lambda y - \xi), \quad (2.16)$$

where $x = (\lambda y)^{\alpha}\xi$.

In particular, for $\alpha = 1$ (exponential differential cross section), expression (2.16) leads to the known result (see [22, Section 10])

$$p(y, \xi) = y^{-1}(\xi\lambda y)^{1/2} I_1[2(\xi\lambda y)^{1/2}] \times \exp(-\lambda y - \xi), \quad (2.17)$$

where $I_1$ is a modified Bessel function.

Formulas (2.15) and (2.16) lead to the following useful relations for the mean value of quantity $\exp(-ky)$:

$$\int_0^{\infty} \exp(-ky) g^{(m)}(y) dy = \left(\frac{\lambda}{\lambda + k}\right)^{m\alpha}. \quad (2.18)$$

This leads to the following relation for the value of $\exp(-ky)$ averaged over distribution function (2.16):

$$\langle \exp(-ky) \rangle = \exp\left[-\xi + \xi\left(\frac{\lambda}{\lambda + k}\right)^{\alpha}\right]. \quad (2.19)$$

### *2.3. Saddle-Point Method*

We will apply the saddle-point method not in the conventional sense as a method for approximate calculation or analysis of asymptotic behavior of quantities, but as a method of exact calculation, which makes it possible to avoid the emergence of rapidly oscillating integrals in final expressions. Such a method was developed in [25] for problems of calculating the spectral–angular characteristics of radiation by relativistic electrons. We will demonstrate the effectiveness of an analogous approach in solving integral equations of type (2.1).

Formulas (2.12) and (2.13) contain rapidly oscillating expressions in the integrands, which makes these formulas inapplicable for practical calculations. In some important cases, however, this difficulty can be overcome by choosing an appropriate integration contour. For this purpose, we write solutions (2.12), (2.13) to Eq. (2.1) in the form of a Laplace transform (see also [2]),

$$p(y, \xi) = \int_{\sigma - i\infty}^{\sigma + i\infty} \exp[zy - \xi\lambda(z)] \frac{dz}{2\pi i}, \quad (2.20)$$

where

$$\lambda(z) = \int_0^{\infty} g(\mu)[1 - \exp(-z\mu)] d\mu. \quad (2.21)$$

Here, $z = r + iw$ is a complex variable.

It is important for further analysis that the argument of the exponential in expression (2.20),

$$f(z) = zy - \xi \int_0^{\infty} g(\mu)[1 - \exp(-z\mu)] d\mu, \quad (2.22)$$

has a saddle point lying on the real axis of the complex variable $z$ for an arbitrary type of differential cross section $g(\mu)$. Indeed, saddle point $z_0$ is defined by the condition $f'(z_0) = 0$. This leads to the following transcendental equation for saddle point $z_0$:

$$y = \xi \int_0^{\infty} \mu g(\mu) \exp(-z_0\mu) d\mu. \quad (2.23)$$

Since $g(\mu)$ is a real-valued function, all quantities in Eq. (2.23) are real and the saddle point also lies on the real axis, $z_0 = r_0$.

Instead of the integration contour traditional for Laplace transformations, it is expedient to carry out integration in Eq. (2.20) along the steepest descent line (SDL). Let us represent function (2.22) as the sum of the real and imaginary parts, $f = u + iv$. Since the saddle point is real-valued, the integration path along the SDL in Eq. (2.20) is defined by the condition $v(r, w) = \mathrm{Im}\, f(z) = 0$, which gives the following equation defining the SDL for the integral in Eq. (2.20):

$$w = \frac{\xi}{y} \int_0^{\infty} g(\mu) \exp(-r\mu) \sin(\mu w) d\mu. \quad (2.24)$$

It can be seen that the SDL (2.24) passes via saddle point $z_0$ defined by formula (2.23).

By way of an example, let us consider differential cross section (2.14). Formulas (2.20) and (2.21) in this case give

$$p(y, \xi) = \lambda \exp(-\xi - \lambda y) \int_C \exp[f(z)] \frac{dz}{2\pi i}, \quad (2.25)$$

$$f(z) = \lambda y z + \xi z^{-\alpha}. \quad (2.26)$$

In accordance with expression (2.23), the saddle point on the real axis is given by

$$r_0 = \left(\frac{\alpha\xi}{\lambda y}\right)^{1/(\alpha+1)}. \quad (2.27)$$

Then the equation for the SDL has the form

$$\rho(\theta) = \left[\frac{\xi \sin(\alpha\theta)}{\lambda y \sin\theta}\right]^{1/(\alpha+1)}, \quad (2.28)$$

where we have introduced polar coordinates $\rho$ and $\theta$ in the complex plane.

The solution to Eq. (2.1) with differential cross section (2.14) can now be expressed in terms of the rapidly converging integral

$$p(y,\xi) = (\pi y)^{-1}\exp(-\xi-\lambda y)\left(\frac{\alpha}{\alpha+1}\right) \times\int_0^\pi u(\theta)\exp[u(\theta)]d\theta, \quad (2.29)$$

where

$$u(\theta) = \xi\rho^{-\alpha}\frac{\sin[\theta(1+\alpha)]}{\sin\theta}. \quad (2.30)$$

Here, $\rho = \rho(\theta)$ can be determined from Eq. (2.28) and quantity $u(\theta)$ in Eq. (2.30) is the real part of function (2.26); i.e., $u = \mathrm{Re}[f(z)]$.

This, a wide class of solutions to problems with a finite total cross section can be represented in two equivalent ways—in the form of infinite series (2.7) or in the form of a rapidly converging integral of type (2.25). Solutions of type (2.7) are convenient when the process is characterized by a relatively small number of interactions $k \le 10$–$20$ (e.g., the case of electrons with energies above 100 GeV in oriented crystals). If, however, the number of interaction events is large (e.g., in the case of ionization energy loss), it is much more convenient to use the integral representation of solutions of type (2.25) or (2.29). In addition, the integral representation makes it possible to describe processes with a diverging total cross section (bremsstrahlung) as well as the cases permitting only numerical integration with realistic expressions for differential cross sections. In the latter case, the initial rapidly oscillating expressions (2.12) and (2.13) are virtually invalid and a correct result can be obtained by integrating along the SDL.

### *2.4. Diverging Power-Type Cross Section with Screening*

If a differential cross section leads to a diverging expression for the total cross section, the solution to Eq. (2.1) cannot be represented in the form (2.6) and formulas (2.7)–(2.11) are meaningless.

Let us apply the saddle-point method for solving equation of motion (2.1) with a diverging cross section of the form

$$\nu(y) = Cy^{-1-\beta}\exp(-\lambda y), \quad 0<\beta<1. \quad (2.31)$$

The result has a form similar to expressions (2.29) and (2.30), in which we must carry out the substitution $\alpha \longrightarrow -\beta$, $\xi \longrightarrow -\kappa$ with

$$\kappa = \frac{\pi\xi\lambda^\beta}{\beta\sin(\pi\beta)\Gamma(\beta)}, \quad (2.32)$$

where we now have $\xi = Ct$.

The solution to Eq. (2.1) assumes the form

$$p(y,\xi) = (\pi y)^{-1}\exp(\kappa-\lambda y)\frac{\beta}{\beta-1} \times\int_0^\pi u(\theta)\exp[u(\theta)]d\theta, \quad (2.33)$$

$$u(\theta) = -\kappa\frac{\sin[\theta(1-\beta)]}{\sin\theta}\left[\frac{\kappa\sin(\beta\theta)}{\lambda y\sin\theta}\right]^{\beta/(1-\beta)}. \quad (2.34)$$

In contrast to solution (2.29), solution (2.33) with diverging cross section possesses the following property. Suppose that we know solution $p_0(y, \xi)$ for a power-type differential cross section without screening (i.e., defined by formula (2.31) with $\lambda = 0$). In this case, the solution for cross section (2.31) can be found by the formula

$$p(y,\xi) = \exp(\kappa-\lambda y)p_0(y,\xi). \quad (2.35)$$

Note that $u(\theta)$ in expression (2.34) is independent of $\lambda$.

Let us consider a few relatively simple analytic solutions for various values of parameter $\beta$.

1. $\beta = 1/3$, $\kappa = 3\xi\lambda^{1/3}\Gamma(3/2)$. In this case, the solution has the form

$$p(y,\xi) = y^{-1}\exp(\kappa-\lambda y)\frac{\kappa}{(3\lambda y)^{1/3}}\mathrm{Ai}\left[\frac{\kappa}{(3\lambda y)^{1/3}}\right], \quad (2.36)$$

where $\mathrm{Ai}(x)$ is the Airy function. In particular, for $\lambda = 0$, we arrive at the result obtained in [5] for a power-type potential with $\beta = 1/3$.

2. $\beta = 1/2$, $\kappa = 2\xi(\pi\lambda)^{1/2}$. In this case, expressions (2.33) and (2.34) can be reduces to the result obtained in [5]:

$$p(y,\xi) = \xi y^{-3/2}\exp\left[-\frac{(\pi^{1/2}\xi-\lambda^{1/2}y)^2}{y}\right]. \quad (2.37)$$

3. $\beta = 2/3$, $\kappa = (3/2)\xi\lambda^{2/3}\Gamma(1/3)$. The solution can be expressed in terms of the Airy function and its derivatives. We will write here only the solution for a power-type differential cross section $p_0(y, \xi)$ (i.e., for $\lambda = 0$) with $\beta = 2/3$:

$$p_0(y,\xi) = 2y^{-1}a^{1/2}\exp\left(-\frac{2}{3}a^{3/2}\right) \times[a^{1/2}\mathrm{Ai}(a) - \mathrm{Ai}'(a)], \quad (2.38)$$

where $a = 3^{1/3}(\xi/2)y^{-2/3}\Gamma(1/3)$. The solution in the case with screening $p(y, \xi)$ can be obtained using formula (2.35).

4. $\beta = 0$. In the limit $\beta \longrightarrow 0$ (see Appendix 1), we obtain from formula (2.33) the well-known distribution (gamma-density) [22]:

$$p(y,\xi) = \lambda\frac{\exp(-\lambda y)}{\Gamma(\xi)}(\lambda y)^{\xi-1}. \quad (2.39)$$

The mean value of quantity $\exp(-ky)$ averaged over distribution (2.39) is given by

$$\langle \exp(-ky) \rangle = \left(\frac{\lambda}{k+\lambda}\right)^{\xi}. \quad (2.40)$$

In the particular case $\lambda = 1$, gamma distribution (2.39) is transformed into the Bethe–Heitler distribution [1]

$$p(y, \xi) = \frac{\exp(-y)}{\Gamma(\xi)} y^{\xi-1}. \quad (2.41)$$

It will be shown below that distribution (2.39) has a number of considerable advantages as compared to (2.41) when applied to problems of radiation emitted by relativistic electrons in an amorphous medium.

It should be noted that solutions $p(y, t)$ to Eq. (2.1) satisfy the condition

$$p(y, t) = \int_0^y p(y-\mu, t_1) p(\mu, t_2) d\mu, \quad (2.42)$$

where $t = t_1 + t_2$.

## 3. EMISSION OF HARD PHONONS

The cascade process of successive emission of a large number of photons by an electron is completely determined by defining the differential probability of emission of a photon with an energy from the interval $(\omega, \omega + d\omega)$ by an electron with energy $E$ during an infinitely short time interval, $\nu_\omega(E)d\omega dt$. This quantity will be referred to as the differential radiation cross section per unit time. Then the electron distribution by instant $t$ will be characterized by function $W_\omega(E_0, t)d\omega$, which gives the probability density of the fact that the energy loss for an electron with initial energy $E_0$ after the passage of path $z = vt$ lies in the interval $(\omega, \omega + d\omega)$ (in the case considered here, the electron velocity can be assumed to be equal to the velocity of light, $v \approx c$).

Distribution function $W_\omega(E_0, t)$ is normalized,

$$\int_0^{E_0} W_\omega(E_0, t) d\omega = 1, \quad (3.1)$$

and satisfies the initial condition $W_\omega(E_0, 0) = \delta(\omega)$. In experiments, the radiation intensity

$$I_\omega(E_0, t) = \omega W_\omega(E_0, t) \quad (3.2)$$

is directly measured (see, for example, [9, 10]).

The total energy loss by radiation is given in this case by

$$\Delta E(t) = \int_0^{E_0} I_\omega(E_0, t) d\omega. \quad (3.3)$$

### 3.1. Cascade Equation for a Multiphoton Process

The time evolution of energy loss distribution function $W_\omega(E_0, t)$ is described by the cascade-type equation

$$\frac{\partial W_\omega(E_0, t)}{\partial t} = \int_0^\omega W_{\omega-\mu}(E_0, t)\nu_\mu(E_0 - \omega + \mu, t)d\mu - \int_0^{E_0-\omega} W_\omega(E_0, t)\nu_\mu(E_0 - \omega, t)d\mu, \quad (3.4)$$

and we consider the general case of the time-dependent differential cross section $\nu_\omega(E, t)$.

The solution to Eq. (3.3) can be expressed in terms of contributions from individual emission events. We assume that the total radiation cross section has a finite value,

$$\nu(E, t) = \int_0^E \nu_\omega(E, t) d\omega. \quad (3.5)$$

Let $W_\omega^{(k)}(E_0, t)d\omega$ be the probability of the fact that the total energy of $k$ emitted photons lines between $\omega$ and $\omega + d\omega$; i.e., $\omega$ in this case is the sum, $\omega = \omega_1 + \omega_2 + \ldots + \omega_k$. Then the energy loss distribution is determined by the sum of partial contributions,

$$W_\omega(E_0, t) = \sum_{k=0}^\infty W_\omega^{(k)}(E_0, t), \quad (3.6)$$

where $W_\omega^{(0)}(E_0, t)$ has the meaning of the probability that not a single emission event occurs within the time interval $(0, t)$.

We will find the values of $W_\omega^{(k)}(E_0, t)$ from the following considerations. The probability that an electron emit precisely $k$ photons during the time interval $(0, t + dt)$ is the sum of two terms. The first term corresponds to the process in which exactly $k$ emission events occur during the time interval $(0, t)$, and not a single emission event occurs during an infinitely short time interval $dt$. The second term corresponds to the case when $k - 1$ photons are emitted during time interval $(0, t)$ and one photon is emitted during time $dt$:

$$W_\omega^{(k)}(t + dt) = W_\omega^{(k)}(t)[1 - \nu(E_0 - \omega, t)dt] + \int_0^\omega W_{\omega-\mu}^{k-1}(t)\nu_\mu(E_0 - \omega + \mu, t)d\mu dt. \quad (3.7)$$

Here, $\nu(E, t)$ is the total radiation cross section per unit time, which is defined by formula (3.5).

Solving Eq. (3.7), we obtain the recurrence relation for the partial probabilities in expression (3.6),

$$W_\omega^{(k)}(E_0, t) = \exp\left[-\int_0^t \nu(E_0-\omega, \tau)d\tau\right] \times \int_0^t d\tau \exp\left[\int_0^\tau \nu(E_0-\omega, \tau')d\tau'\right] \times \int_0^\omega W_{\omega-\mu}^{k-1}(E_0, \tau)\nu_\mu(E_0-\omega+\mu, \tau)d\mu, \quad (3.8)$$

where

$$W_\omega^{(0)}(E_0, t) = \delta(\omega)\exp\left[-\int_0^t \nu(E_0-\omega, \tau)d\tau\right]. \quad (3.9)$$

Expressions (3.6)–(3.9) define the solution to the equation of motion of general form (3.4).

Since the total radiation cross section in these expressions is a function of time and energy, $\nu = \nu(E, t)$, the statistics of individual emission events in expressions (3.6) and (3.8) cannot be presented any longer by a simple relation of type (2.8). In particular, in the case when the total cross section is independent of $E$ (i.e., $\nu = \nu(t)$), the probability $P_k(t)$ of obtaining exactly $k$ photons during time interval $(0, t)$ is defined by the formula

$$P_k(t) = \exp\left[-\int_0^t \nu(\tau)d\tau\right] \times \int_0^t d\tau \exp\left[-\int_0^\tau \nu(\tau')d\tau'\right]\nu(\tau)P_{k-1}(\tau), \quad (3.10)$$

where

$$P_0(t) = \exp\left[-\int_0^t \nu(\tau)d\tau\right]. \quad (3.11)$$

Formulas (3.7)–(3.11) define a generalized Poisson process. If the radiation cross section is constant ($\nu = \nu_0$), expression (3.10) is transformed into Poisson distribution (2.8) with $\xi = \nu_0 t$. The motion of an electron in the field of atomic chains and planes of a crystal is an example of a time-dependent differential cross section [26].

### *3.2. Multiphoton Process of the Poisson Type*

The results of the previous section can be considerably simplified in the case of a time-independent differential cross section $\nu_\omega(E)$. We will also assume that the total cross section $\nu_0$ does not depend on the electron energy $E$. In this case, equation of motion (3.4) assumes the form

$$\frac{\partial W_\omega(E_0, t)}{\partial t} = \int_0^\omega W_{\omega-\mu}(E_0, t)\nu_\mu(E_0-\omega+\mu)d\mu - \int_0^{E_0-\omega} W_\omega(E_0, t)\nu_\mu(E_0-\omega)d\mu. \quad (3.12)$$

Let $g_\omega^{(1)}(E)$ be the differential probability density of the fact that the energy of an emitted photon is $\omega$:

$$g_\omega^{(1)}(E) = \nu_\omega(E)/\nu_0. \quad (3.13)$$

We introduce the probabilities $g_\omega^{(k)}(E_0)\,d\omega$ that the total energy of $k$ photons emitted by an electron with initial energy $E_0$ lies between $\omega$ and $\omega + d\omega$. These quantities are normalized to unity,

$$\int_0^E g_\omega^{(k)}(E)d\omega = 1, \quad (3.14)$$

and satisfy the recurrence relations

$$g_\omega^{(k)}(E_0) = \int_0^\omega g_{\omega-\mu}^{(k-m)}(E_0-\mu)g_\mu^{(m)}(E_0)d\mu, \quad m < k. \quad (3.15)$$

The probability $P_k(t)$ that exactly $k$ photons are emitted during time interval $[0, t]$ is now given by Poisson distribution (2.8) with $\xi_0 = \nu_0 t$. Functions $W_\omega^{(k)}(E_0, t)$ defined by formula (3.6) can now be simply reduced to products of the type $P_k(t)g_\omega^{(k)}(E_0)$. For the energy loss distribution, we now have

$$W_\omega(E_0, t) = \sum_{k=0}^\infty P_k(t)g_\omega^{(k)}(E_0). \quad (3.16)$$

This expression, which is a solution to Eq. (3.12) with a time-independent differential cross section and an energy-independent total cross section, satisfies the condition

$$W_\omega(E_0, t_1+t_2) = \int_0^\omega W_{\omega-\mu}(E_0-\mu, t_1)W_\mu(E_0, t_2)d\mu, \quad (3.17)$$

where $t = t_1 + t_2$.

If the total radiation cross section depends on the electron energy, but does not depend of time, formulas (3.6)–(3.9) give a solution of a more general form than (3.16), which will not be considered here.

### 3.3. Multiplicity of Radiation

Some quantities characterizing the multiphoton nature of radiation can be expressed in terms of distribution functions $W_\omega^{(k)}(E_0, t)$. For example, the average number of emitted photons as a function of their total energy $\omega$ is given by

$$\langle k(\omega, t)\rangle = \sum_{k=1}^{\infty} k W_\omega^{(k)}(E_0, t), \tag{3.18}$$

where $\omega$ is the total energy of $k$ emitted photons. In other words, if an experimenter sees that the total energy of all photons emitted by an electron is $\omega$, quantity (3.18) indicates their mean number.

The probability of emission of exactly $k$ photons during time $t$ is given by

$$W^{(k)}(t) = \int_0^{E_0} W_\omega^{(k)}(E_0, t)d\omega, \tag{3.19}$$

and the average number of emitted photons is

$$\langle k(t)\rangle = \int_0^{E_0} d\omega \sum_{k=1}^{\infty} k W_\omega^{(k)}(E_0, t). \tag{3.20}$$

The values of quantities (3.18)–(3.20) can be directly measured in experiments (see, for example, [9]). An important quantity (which, however, cannot be directly measured) is the number of single photons $n_\omega(E_0, t)d\omega$ corresponding to the given interval $(\omega, \omega + d\omega)$ of their energies and emitted by an electron with initial energy $E_0$ by instant $t$. The knowledge of such a radiation spectrum that is not distorted by multiphoton nature of emission is important for practical applications and provides the most complete experimental information on the type of radiation. If a process is of the Poisson type and evolves in accordance with Eq. (3.12), we obtain for the one-photon spectrum the expression

$$n_\omega(E_0, t) = \sum_{k=0}^{\infty} P_k(t) \sum_{m=0}^{k} n_\omega^{(m)}(E_0), \tag{3.21}$$

where $P_k(t)$ is the Poisson distribution. Partial contributions to spectrum (3.21) are defined by a recurrence relation of the form differing from (3.15):

$$n_\omega^{(m)}(E_0) = \int_0^{E_0-\omega} n_\omega^{(m-1)}(E_0 - \mu) n_\mu^{(1)}(E_0)d\mu. \tag{3.22}$$

Each partial quantity $n_\omega^{(m)}(E_0)$ defines the probability that the energy of the $m$th emitted photon lies in the interval $(\omega, \omega + d\omega)$. Here, quantity $n_\omega^{(1)}(E) = g_\omega^{(1)}(E)$ is defined by Eq. (3.13), $n_\omega^{(0)}(E) = \delta(\omega)$.

Functions (3.22) satisfy the normalization conditions,

$$\int_0^{E_0} n_\omega^{(m)}(E_0)d\omega = 1.$$

The integral of $n_\omega(E_0, t)$ gives the same quantity as (3.20); i.e., the average number of photons emitted by instant $t$ is

$$\langle k(t)\rangle = \int_0^{E_0} d\omega n_\omega(E_0, t). \tag{3.23}$$

For a Poisson-type process expressed by formulas (3.21) and (3.22), we have $\langle k(t)\rangle = \nu_0 t$. Thus, while quantity $\omega$ in expressions (3.21)–(3.23) is the energy of an individual photon, quantity $\omega$ in formulas (3.1)–(3.20) is the total energy of all emitted photons. One-photon spectra of type (3.21) were studied in [27] using computer simulation.

We assume that the passage of electrons through a substance is characterized by two independent mechanisms of emission. Then, intensity (3.2), as well as the distribution function $W_\omega(E_0, t)$ for a given $\omega$, is not an additive sum of contributions from the corresponding types of radiation, while the total loss by radiation (3.3), the total number of emitted photons (3.18), and the spectrum of single photons (3.21) are additive quantities. For example, the passage of electrons through an oriented crystal is accompanied by incoherent radiation emitted by individual atoms, $\nu_\omega^{BH}$ [1], and coherent radiation at continuous potential of atomic chains and planes, $\nu_\omega^{C}$ [28]. Then $I_\omega \neq I_\omega^{BH} + I_\omega^{C}$, while $n_\omega = n_\omega^{BH} + n_\omega^{C}$.

### 3.4. Analytic Solutions

Let us consider a particular case when electron energy $E$ and energy $\omega$ of an emitted photon appear in the differential emission probability in the form of a certain combination so that

$$\nu_\omega(E)d\omega = \nu(y)dy = \nu(y)\frac{d\omega}{E-\omega}, \tag{3.24}$$

where

$$y = \ln\frac{E}{E-\omega}; \tag{3.25}$$

i.e., the denominator in the argument of the logarithm in relation (3.25) contains the electron energy after the emission.

Let two photons with energies $\omega_1$ and $\omega_2$ be emitted in succession. These energies correspond to $y_1 = \ln(E/E_1)$ and $y_2 = \ln(E_1/E_2)$, where $E_1 = E - \omega_1$ and $E_2 = E_1 - \omega_2$. Then the total value of $y = y_1 + y_2$ is trans-

formed in accordance with the same law (3.25) as for components $y_1$ and $y_2$; i.e., $y = \ln(E/E_2)$, $E_2 = E - \omega$, where $\omega = \omega_1 + \omega_2$ is the total energy loss in two emission events. Thus, an important property of transformation (3.25) is that the sum of $y_k$ corresponding to $k$ photons emitted in succession satisfies a transformation of the same type as (3.25) for each quantity $y_k$ taken separately. Transformation (3.25) makes it possible in this case to reduce Eq. (3.12) to a much simpler equation of the type (2.1) or (2.6) if the corresponding differential cross section has the form (3.24). Indeed, let us pass in the first integral in relation (3.12) from variables $\omega$ and $\mu$ to variables $y$ and $\eta$ using the formulas $y = \ln(E_0/E')$ and $\eta = \ln(E''/E')$, where $E' = E_0 - \omega$ and $E'' = E' + \mu$. Then the distribution function $p(y, t)$ in the new variable is connected with the initial function by the relation

$$p(y, t) = W_\omega(E_0, t)E_0\exp(-y).$$

Quantity $\omega - \mu$ can be expressed in terms of the difference $y - \eta$, and the first integral in relation (3.12) can be expressed in terms of an integral of product $p(y - \eta, t)\nu(\eta)$ with respect to $\eta$ from 0 to $y$. Proceeding analogously with the second integral, we reduce Eq. (3.12) to an equation of the type (2.1). All these arguments correspond to the case when the differential cross section is independent of time and the total cross section (if the corresponding integral converges) is independent of the electron energy (i.e., if a Poisson-type process occurs). Thus, the results obtained in Section 2 for Eq. (2.1) can be applied to Eq. (3.12); a large class of solutions to the latter equation for distribution function $W_\omega(E_0, t)$ can be obtained by substituting expression (3.25) into the formulas derived for Eq. (2.1).

The simplest case corresponds to the uniform distribution

$$g_\omega^{(1)}(E) = E^{-1}, \quad \omega \le E. \tag{3.26}$$

Then relation (3.15) leads to the following expression for the $k$-photon probability:

$$g_\omega^{(k)}(E) = \frac{1}{(k-1)!E}\left(\ln\frac{E}{E-\omega}\right)^{k-1}. \tag{3.27}$$

Substitution of this expression into Eq. (3.16) leads to the solution

$$W_\omega(E_0, t) = \frac{1}{E_0}\exp(-\xi)\xi^{1/2}\left(\ln\frac{E_0}{E_0-\omega}\right)^{1/2} \times I_1\left[2\xi^{1/2}\left(\ln\frac{E_0}{E_0-\omega}\right)^{1/2}\right], \tag{3.28}$$

where $\xi = \nu_0 t$, which corresponds to formula (2.17) for $\lambda = 1$ with $y$ defined by formula (3.25).

To analyze the spread in the energy losses of relativistic electrons in a medium, Bethe and Heitler [1] introduced the following approximate expression for the differential cross section per unit time:

$$\nu_\omega(E)d\omega = \frac{\nu_0 d\omega}{E\ln\dfrac{E}{E-\omega}}; \tag{3.29}$$

here, $\nu_0 \approx (4/3)c/L_0$ and $L_0$ is the radiation length in an amorphous substance. Formula (3.29) corresponds to the expression

$$\nu(y)dy = \nu_0\exp(-y)dy/y. \tag{3.30}$$

The Bethe–Heitler solution [1] for cross section (3.29) is given by

$$W_\omega(E_0, t)d\omega = \frac{1}{E_0\Gamma(\xi)}\left(\ln\frac{E_0}{E_0-\omega}\right)^{\xi-1}d\omega, \tag{3.31}$$

where $\xi = \nu_0 t$ and $\Gamma(\xi)$ is the gamma function.

Result (3.31) is inconvenient for large energy losses, when $\omega \sim E_0$; i.e., for large values of $y$ (for $\xi > 1$, we have divergence for $\omega = E_0$). This drawback can be eliminated by replacing relation (3.30) by the expression

$$\nu(y)dy = C\exp(-\lambda y)dy/y, \tag{3.32}$$

where $\lambda > 1$.

Cross section (3.32) has the form (2.14) with $\alpha = 0$; consequently, the solution to the corresponding equation of motion has the form (2.39). This leads to the Lindhard distribution [29]

$$W_\omega(E_0, t)d\omega = \frac{\lambda}{E_0\Gamma(\xi)}\left(\frac{E_0-\omega}{E_0}\right)^{\lambda-1} \times\left(\lambda\ln\frac{E_0}{E_0-\omega}\right)^{\xi-1}d\omega, \tag{3.33}$$

where $\xi = Ct$.

A wide class of one-photon differential radiation cross sections with a regular asymptotic behavior for small frequencies $\omega \longrightarrow 0$ can be described by a less general expression as compared to those considered above,

$$g_\omega^{(1)}(E_0) = \frac{\lambda^\alpha}{E_0\Gamma(\alpha)}\left(\frac{E_0-\omega}{E_0}\right)^{\lambda-1}\left(\ln\frac{E_0}{E_0-\omega}\right)^{\alpha-1}, \tag{3.34}$$

where $\alpha > 0$. This expression corresponds to formula (2.14).

The energy loss distribution is connected with the energy distribution of particles through the relation $W_\omega(E_0, t) = F(E = E_0 - \omega, t)$, where function $F(E, t)$ satisfies the initial condition $F(E, t = 0) = \delta(E - E_0)$.

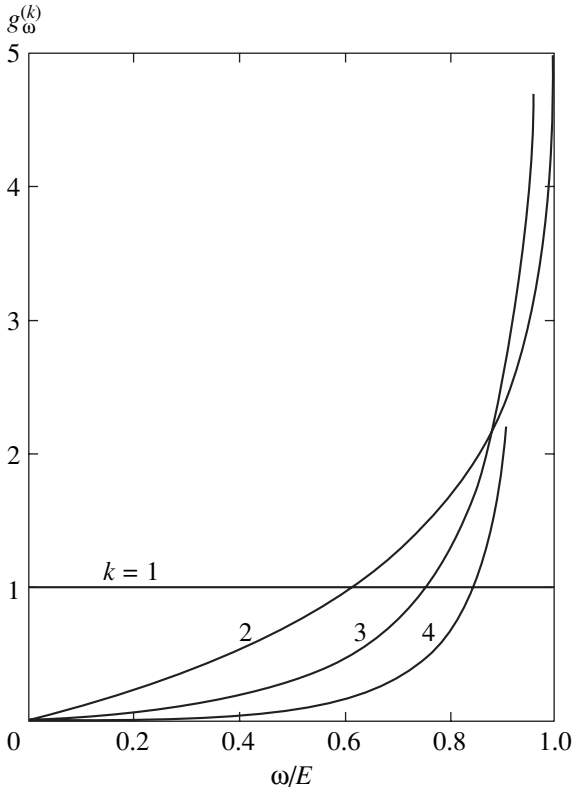



**Fig. 1.** Partial contributions $g_\omega^{(k)}(E)$ (3.27) for the uniform initial distribution (3.26) for different values of $k$.

In accordance with Eq. (3.16), the electron energy distribution function corresponding to differential cross section (3.34) has the form

$$F(E, t) = \sum_{k=0}^{\infty} P_k(\xi) \frac{\lambda^{k\alpha}}{E_0 \Gamma(k\alpha)} \left(\frac{E}{E_0}\right)^{\lambda-1} \left(\ln \frac{E_0}{E}\right)^{k\alpha-1}. \quad (3.35)$$

The mean value of quantity $(E/E_0)^m$ is given by

$$\left\langle \left(\frac{E}{E_0}\right)^m \right\rangle = \int_0^{E_0} \left(\frac{E}{E_0}\right)^m F(E, t) dE. \quad (3.36)$$

For distribution (3.35), we obtain, in accordance with Eq. (2.19),

$$\left\langle \left(\frac{E}{E_0}\right)^m \right\rangle = \exp\left[-\xi + \xi\left(\frac{\lambda}{\lambda + m}\right)^\alpha\right]. \quad (3.37)$$

In particular, for the Bethe–Heitler case (3.31), we have $\langle E/E_0 \rangle = 2^{-\xi}$.

Figure 1 shows partial contributions $g_\omega^{(k)}(E)$ (3.27) for the uniform initial distribution (3.26). It can be seen that, with increasing number $k$ of emitted photons, the partial distributions as functions of $\omega$ are rapidly dis-

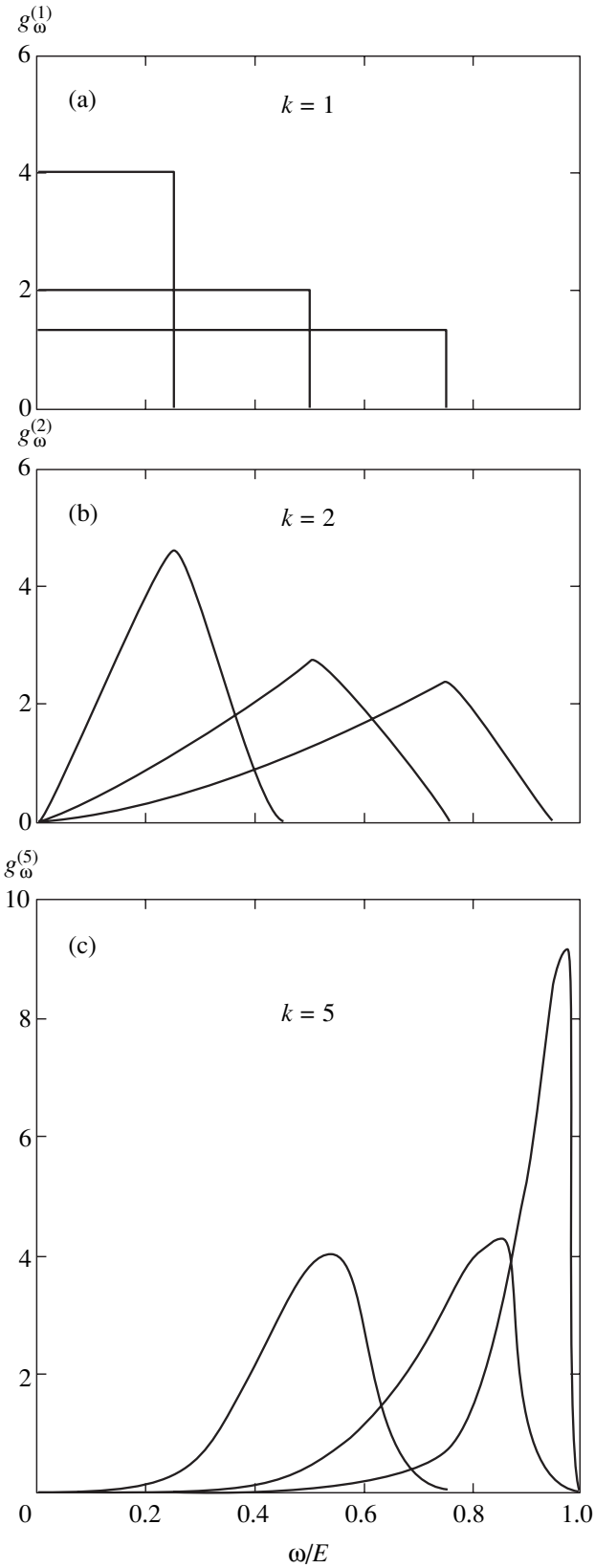



**Fig. 2.** (a) Stepwise one-photon distribution $g_\omega^{(1)}$ (3.38) for $\alpha = 0.25$, 0.5, and 0.75; (b, c) corresponding two-photon $g_\omega^{(2)}$ and five-photon $g_\omega^{(5)}$ distributions.

placed towards harder frequencies, exhibiting a logarithmic divergence for $\omega \sim E$. The step distribution

$$g_\omega^{(1)}(E) = \begin{cases} (\alpha E)^{-1}, & 0 < \omega < \alpha E \\ 0, & \alpha E < \omega < E, \end{cases} \quad (3.38)$$

where $0 < \alpha < 1$, possesses much more realistic properties. This distribution also has an analytic solution in the form of a combination of piecewise-continuous functions of a simple form, which are not given here in view of their cumbersome nature. Figure 2 illustrates the two-photon $g_\omega^{(2)}$ (Fig. 2b) and five-photon $g_\omega^{(5)}$ (Fig. 2c) distributions for various initial step distributions (Fig. 2a) defined by formula (3.38). In this case, all values of $g_\omega^{(k)}$ vanish for $\omega \longrightarrow E$. It can be seen from Fig. 2 that the shape of distributions $g_\omega^{(k)}$ is determined to a considerable extent by the initial conditions.

## 4. RADIATION IN ORIENTED CRYSTALS

The trajectory of a relativistic electron passing through a crystal at a certain small angle $\theta \sim \theta_L$ to principal crystallographic directions is determined by the correlated action on the electron by atoms forming this direction [26]. Here, $\theta_L = (4Ze^2/dE)^{1/2}$ is the Lindhard critical angle, $Z$ is the atomic number of the substance, and $d$ is the distance between neighboring atoms along the given direction. Such a correlated action of atoms in a chain is characterized by a continuous potential $U(r)$ [26] acting on the electron and depending only on the distance to the chain. The force $F = |\nabla U(r)|$ acting on the electron is on the order of $F \sim Ze^2/(da_F) \sim 10^2$–$10^3$ eV/Å and leads to intensive emission (the Kumakhov effect) [28, 30]. Here, $a_F$ is the Thomas–Fermi screening parameter. For electron energies above 50–100 GeV, we can assume that the field acting on the electron does not change over the length of radiation formation; consequently, the characteristics of radiation can be described in the constant field approximation [31]. In this case, the corresponding cross sections can be reduced to the familiar formulas for synchrotron radiation taking into account spin and quantum recoil effects during emission [32, 33]. From the standpoint of cascade equations of type (3.4), it is important that the radiation cross sections in the CFA are determined only by local characteristics of the field (namely, by the force acting at a given point) and, hence, can be used directly as the differential cross sections $\nu_\omega(E)$ appearing in formulas (3.4) and (3.12). A more consistent theory, which takes into account the field nonuniformity over the length of radiation formation and at the same time is local by nature, has been developed recently in [4]. Nevertheless, we will confine our analysis to the CFA since it is sufficient for studying typical properties of cascade processes of multiple emission of photons, which accompany the passage of electrons with energies exceeding 100 GeV through oriented crystals.

### 4.1. Constant Field Approximation

The differential cross section of emission of a photon with energy $\omega$ by an electron with $E$ per unit time is defined in the CFA by the quantum-mechanical synchrotron formula [32]

$$\nu_\omega(E) = \frac{\alpha}{\pi\sqrt{3}\hbar\gamma^2} \times \left[ 2K_{2/3}(\xi) - \int_\xi^\infty K_{1/3}(\eta)d\eta + \frac{u^2}{1-u}K_{2/3}(\xi) \right] \quad (4.1)$$

$$= \frac{\alpha}{\pi\hbar\gamma^2}\int_0^\infty \left[ \frac{9 + 36x^2 + 16x^4}{9 + 12x^2} + \frac{u^2}{1-u}\left(1 + \frac{2x^2}{3}\right) \right] \times \exp R\frac{dx}{Y(x)}, \quad (4.2)$$

where the argument of the exponential function is

$$R(x) = -\xi\left(1 + \frac{4x^2}{3}\right)Y(x). \quad (4.3)$$

Here, $\alpha = 1/137$, $u = \omega/E$, $\gamma = E/mc^2$ is the Lorentz factor,

$$\xi = \frac{2}{3}\frac{u}{1-u}\frac{1}{\chi},$$

$\chi = \hbar F\gamma/m^2c^3$ is the Lorentz-invariant field parameter, $F$ is the force acting on the electron, and $m$ is the electron rest mass. Function $Y(x) = (1 + x^2/3)^{1/2}$ is the SDL for the Macdonald functions $K_{1/3}(\xi)$ and $K_{2/3}(\xi)$ (see Appendix B). Our representation (4.2) is much more convenient for specific applications of standard formula (4.1) since it does not contain special functions and has the form of a rapidly converging integral, in which argument $\xi$ appears only in the exponential factor. Formula (4.2) was derived by the saddle point method [25] (see Appendix B). Formulas (4.1)–(4.3) are transformed into the classical formulas for synchrotron radiation for small values of field parameter $\chi < 0.1$.

Figure 3 shows the dependence of the total radiation cross section in the CFA per unit length (in $\mu m^{-1}$) on force $F$ acting on the electron (in keV/Å), calculated on the basis of formulas (4.1)–(4.3) for various values of energy $E$. In other words, the integral quantity (3.5) laid along the ordinate axis (divided by the velocity of light) can be treated as the number of photons emitted per unit length. The values of force $F$ typical of oriented crystals are indicated. It can be seen that the total radiation cross section in the CFA slowly decreases with increasing $E$. For example, as the electron energy

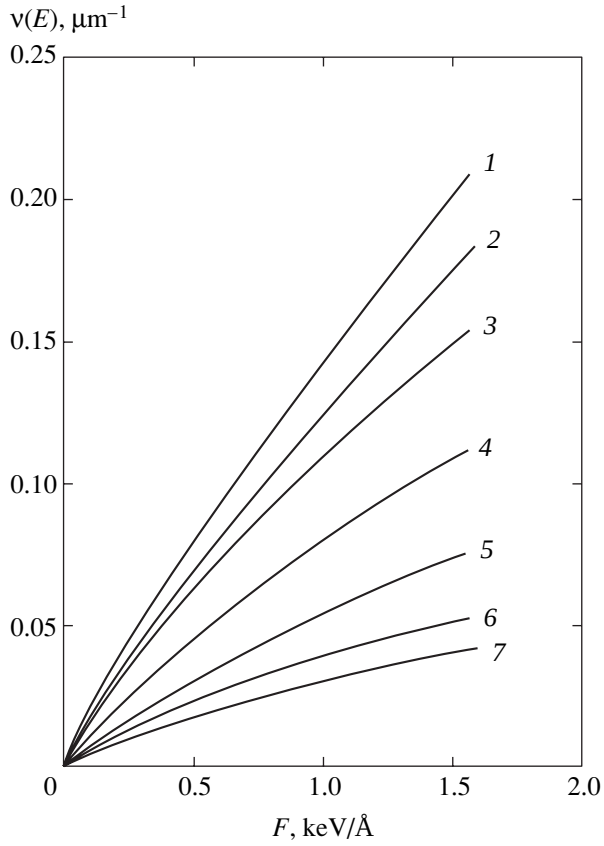



**Fig. 3.** Photon emission probability per unit length as a function of force $F$ calculated in the CFA using formulas (4.1)–(4.3) for various electron energies (in TeV): 0.075 (*1*), 0.150 (*2*), 0.3 (*3*), 1 (*4*), 4 (*5*), 10 (*6*), and 20 (*7*).

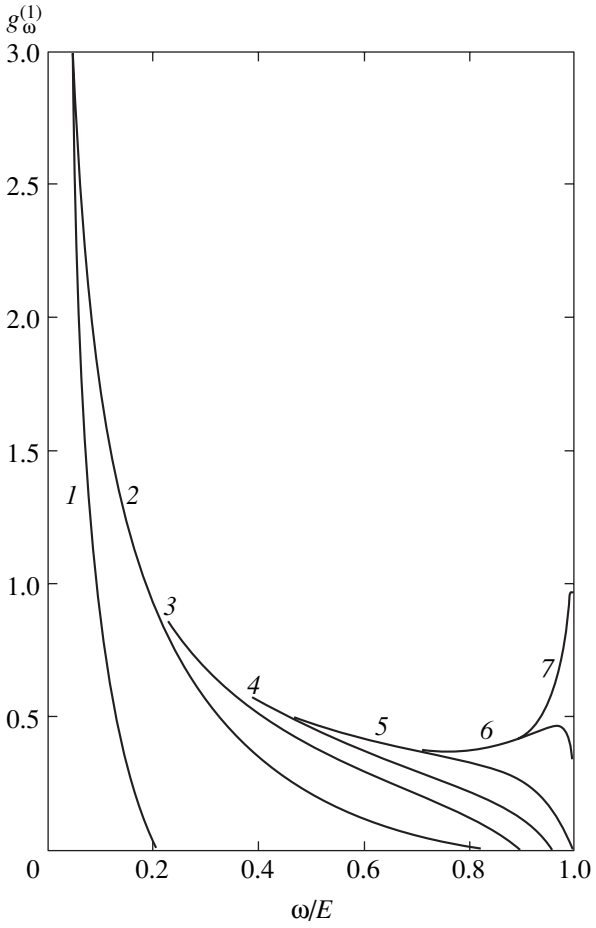



**Fig. 4.** One-photon emission probabilities $g_\omega^{(1)}$ (3.13) calculated in the CFA (4.1)–(4.3) for various values of field parameter $\chi$ = 0.05 (*1*), 0.5 (*2*), 1.45 (*3*), 3 (*4*), 9 (*5*), 45 (*6*), and 208 (*7*).

decreases by half from 150 GeV for $F \approx 500$ eV/Å, the total radiation cross section increases only by 10%. Thus, in the case considered here, the total radiation cross section can be approximately assumed to be independent of the electron energy. In the CFA, the differential cross section (4.1)–(4.3) has an integrable singularity $\nu_\omega \sim \omega^{-2/3}$ for $\omega \longrightarrow 0$. This leads to the presence of a large number of photons with relatively low energies in the radiation spectrum. For example, the fraction of photons with energies smaller that 0.01 of the electron energy ranges from 50% (for $E \sim 100$ GeV) to 25% (for $E \sim 20$ TeV) of all emitted photons. It should be noted, however, that such a form of the dependence for $\omega \longrightarrow 0$ is typical precisely for synchrotron-type spectra. More accurate calculations (as compared to those made in the CFA) taking into account the field nonuniformity indicate strong suppression of radiation in the soft part of the spectrum [4].

Figure 4 shows one-photon emission probabilities (spectra) (3.13) normalized to unity and calculated in the CFA for various values of field parameter $\chi$. For large values of $\chi > 40$–50, the photon spectrum acquires a peak at $\omega/E \sim \chi/(1 + \chi) \sim 1$. The emergence of this peak is associated with an increase in the contribution of the spin term to the spectrum with increasing $\chi$. The spin contribution to radiation is determined by the terms containing $u^2/(1 - u)$ in expressions (4.1), (4.2). For $E \approx 150$–300 GeV, the values of $\chi$ range from 1 to 10 in different crystals. The effect of spin on radiation has been studied recently in experiments [34]. The mean energy $\langle\omega\rangle$ of an emitted photon, i.e., the value of $\omega$ averaged over one-photon distributions $g_\omega^{(1)}$ (3.13), is of interest for our analysis. For $\chi \approx 1$–2, the value of this quantity is $\langle\omega\rangle \approx 0.1$–0.12 and tends to the limit $\langle\omega\rangle \longrightarrow 0.25$ as $\chi \longrightarrow \infty$.

Partial probabilities $g_\omega^{(k)}$ (3.15) for 150-GeV electrons are shown in Fig. 5 for two values of the force. The initial distributions $g_\omega^{(1)}$ corresponding to the CFA (4.1)–(4.3) are shown by the dashed curves. The numerical method for calculating $g_\omega^{(k)}$ for an arbitrary initial distribution is given in Appendix C. In Fig. 5, the total energy of $k$ emitted photons is laid along the abscissa axis. Figure 5 illustrates the strong dependence of the

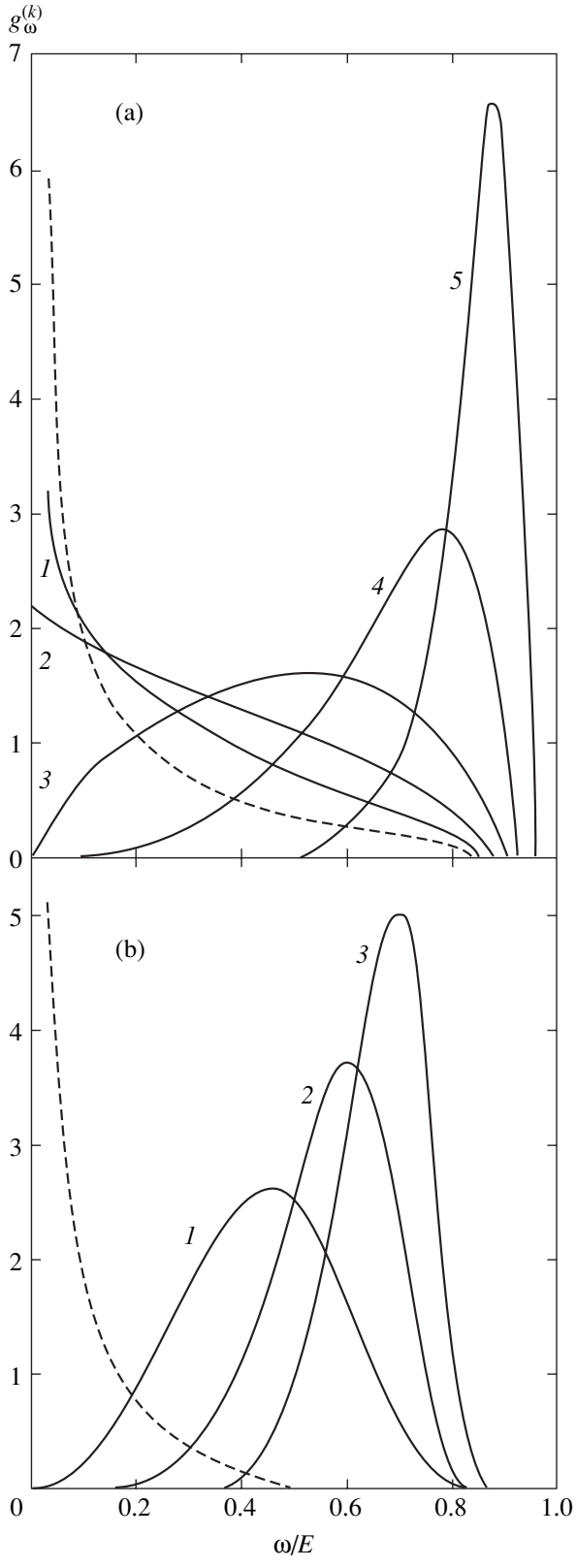



**Fig. 5.** Partial probabilities $g_\omega^{(k)}$ (3.15) for 150-GeV electrons in the CFA for $F = 500$ (a) and 100 (b) eV/Å. The initial distributions are shown by dashed curves. (a) $k = 2$ (*1*), 3 (*2*), 5 (*3*), 10 (*4*), and 20 (*5*); (b) $k = 10$ (*1*), 20 (*2*), and 30 (*3*).

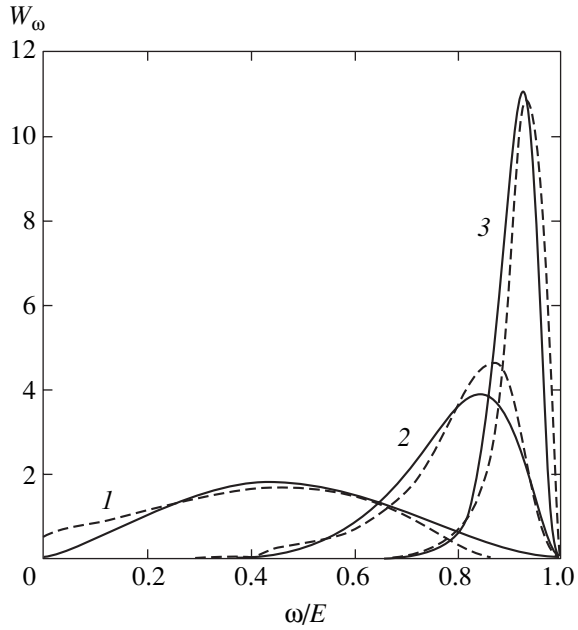



**Fig. 6.** Solutions to cascade equation (3.16) for 150-GeV electrons in a 185-μm-thick target. Solid curves are calculated in the CFA and dashed curves correspond to Lindhard distribution (3.33): *1*—$F = 200$ eV/Å, $\chi = 0.44$, $\langle k\rangle = 3.4$, $\lambda = 3$, $\xi = 2.3$; *2*—$F = 500$ eV/Å, $\chi = 1.1$, $\langle k\rangle = 14$, $\lambda = 5$, $\xi = 8.6$; *3*—$F = 1000$ eV/Å, $\chi = 2.22$, $\langle k\rangle = 24$, $\lambda = 12$, $\xi = 30.3$. Here, $\langle k\rangle$ is the mean number of emitted photons (3.20); $\lambda$ and $\xi$ are the parameters of the Lindhard distribution.

shape of the $k$-photon distribution of probabilities $g_\omega^{(k)}$ on the initial distribution for the same value of $k$ (cf. curves *5* and *2* in Figs. 5a and 5b for $k = 20$).

Figure 6 shows the solutions to the cascade equation (3.16) for 150-GeV electrons for three values of force in a 185-μm-thick target. The solid curves are calculated in the CFA, while the dashed curves describe the Lindhard distribution (3.33). It can be seen that, with appropriately chosen parameters $\xi$ and $\lambda$, the relatively simple Lindhard distribution can successfully approximate the results of more accurate calculations. Since the process studied here is of the Poisson type, the average number of emitted photons in this case is $\langle k\rangle = \nu_0 z$ (the values of $\nu_0$ can be found in Fig. 3).

### *4.2. Analysis of Experimental Results*

We will apply the results to analyze radiation emitted by electrons with energies of hundreds GeV during their motion at small angles $\theta \sim \theta_L$ to an atomic chain in a crystal (axial channeling and quasi-channeling [35–37]). The regular motion in the continuous potential $U(r)$ ($r$ is the distance to the atomic chain) is perturbed by incoherent multiple scattering from individual atoms of the crystal, which leads to a stochastic increase in the transverse electron energy $\varepsilon$.

In our calculations, we disregard multiple scattering and the competing effect of the decrease in the transverse energy due to emission of photons (transverse energy damping). In this case, the motion of an electron in an axially symmetric potential $U(r)$ is characterized by two integrals of motion, viz., transverse energy $\varepsilon$ and electron angular momentum $\mu$ relative to the atomic chains, while the pattern of axial channeling can be described as follows. The electron incident on the crystal at angle $\theta_{in}$ to the atomic chain at a distance $r_{in}$ from it acquires a transverse energy

$$\varepsilon = \frac{E\theta_{in}^2}{2} + U(r_{in})$$

and an angular momentum

$$\mu = \frac{E}{c}\theta_{in} r_{in} \cos\varphi_{in},$$

where $\varphi_{in}$ is the azimuthal angle of incidence relative to the chain. The electron moves in a trajectory with constant $\varepsilon$ and $\mu$ and emits photons whose spectrum can be calculated using formulas (3.13)–(3.16). The resultant spectrum can be obtained by averaging over points $r_{in}$ and angles $\theta_{in}$ of incidence, which is equivalent to averaging over the electron distribution functions over transverse energies and angular momenta in the crystal [35, 36]. In this way, the initial experimental conditions (angular divergence of the electron beam at the inlet, etc.) are taken into account.

To use formulas (3.13)–(3.16) for calculations, we must specify the initial one-photon distribution function $g_\omega^{(1)}$. In the CFA, this quantity depends on the current distance $r$ to the atomic chain at each instant. For this reason, for the one-photon function in formula (3.16), we must take distribution $g_\omega^{(1)}$ calculated in the CFA (4.1)–(4.3) and averaged over the equilibrium distribution function for electrons with preset values of $\varepsilon$ and $\mu$ over the transverse coordinate. For electrons getting in a state with a finite transverse motion (i.e., with $\varepsilon < 0$), the equilibrium distribution over the transverse coordinate has the form

$$dw(\varepsilon, \mu, r) = \frac{dr}{T(\varepsilon, \mu)}\left(\frac{2m\gamma}{\varepsilon - U(r) - \mu^2/2m\gamma r^2}\right)^{1/2}, \quad (4.4)$$

where $T(\varepsilon, \mu)$ is the period of transverse radial oscillations of an electron in a channel, $r_{min} < r < r_{max}$, and the values of $r_{min}$ and $r_{max}$ are determined from the condition that the radicand in formula (4.4) is equal to zero.

Quasi-channeled electrons with infinite transverse motion ($\varepsilon > 0$) are uniformly distributed in the transverse plane. The transverse region accessible to electrons is independent of $\varepsilon$ (and is equal to the cross-sectional area per chain, $S_0 = 1/Nd$, where $N$ is the number of atoms in the crystal per unit volume); consequently, the radiation cross section for channeled electrons in the CFA is independent of their transverse energy. For channeled electrons with $\varepsilon < 0$, the radiation cross sections strongly depend on the transverse energy because the transverse region $S(\varepsilon)$ accessible to electrons strongly depends on $\varepsilon$ and electrons with modulo large transverse energies emit more strongly since they move at smaller distances from the atomic chain.

We calculated the continuous atomic chain potential $U(r)$ on the basis of the Doyle–Turner atomic potential [38] with averaging over thermal vibrations of atoms in the crystal. We assumed that the potential is equal to zero ($U(r) = 0$) in region $r_1 < r < r_0$, where $r_1$ is the shortest distance between atomic chains along the preset direction and $r_0 = (\pi Nd)^{-1/2}$ is the channel radius. In other words, the maximal fraction of electrons that can be trapped in a channel is $N_{ch} = (r_1/r_0)^2$ (for $\theta_{in} = 0$). For example, for a $\langle 110 \rangle$ germanium crystal, this fraction amounts to about 30% of the beam.

The results of calculation of the radiation intensity spectrum (3.2) are shown in Figs. 7a and 7b, which present the results for 150-GeV electrons incident at angles in the intervals 0–7 (Fig. 7a) and 15–21 (Fig. 7b) μrad relative to the $\langle 110 \rangle$ direction of a germanium crystal of thickness 200 μm (in this example, $\theta_L$ = 55 μrad). The experimental data are borrowed from [9]. The spectra are given in units of intensity for an amorphous medium: $Ez/L_{rad}$, where $z$ is the crystal thickness and $L_{rad}$ is the radiation length ($L_{rad}$ = 2.36 cm for germanium). Thus, the values laid along the ordinate axis in Figs. 7a and 7b show that the radiation intensity in an oriented crystal exceeds that in an amorphous target of the same thickness. The solid curve in Fig. 7a corresponds to calculations performed in the cascade theory in accordance with formulas (3.2) and (3.16). Theoretical values in Fig. 7a and in the remaining figures were calculated at individual points connected by straight lines. It can be seen that, under the conditions represented in Figs. 7a and 7b, the theory developed here reproduces the experimental results to a high degree of accuracy.

The fraction of electrons trapped in the channeling mode under the condition represented in Fig. 7a is $N_{ch}$ = 0.25 (for angles of incidence of 0–7 μrad). The dashed curves show the calculated contributions to the radiation spectrum from electrons with different transverse energies: $\varepsilon > 0$ (*1*); $\varepsilon \in [-10, 0]$ eV, (0.113) (*2*); $\varepsilon \in [-20, -10]$ eV, (0.044) (*3*); and $\varepsilon \in [-U_m, -20]$ eV, (0.094) (*4*). The values in parentheses show the fraction of electrons with corresponding transverse energies. The depth of the potential well in this case is $U_m$ = 202 eV (the crystal temperature is 293 K, the mean amplitude of thermal vibrations was taken equal to $u_\perp$ = 0.12 Å). It follows from Fig. 7a that the main contribution to the experimentally observed peak at $\omega \approx 0.8E$

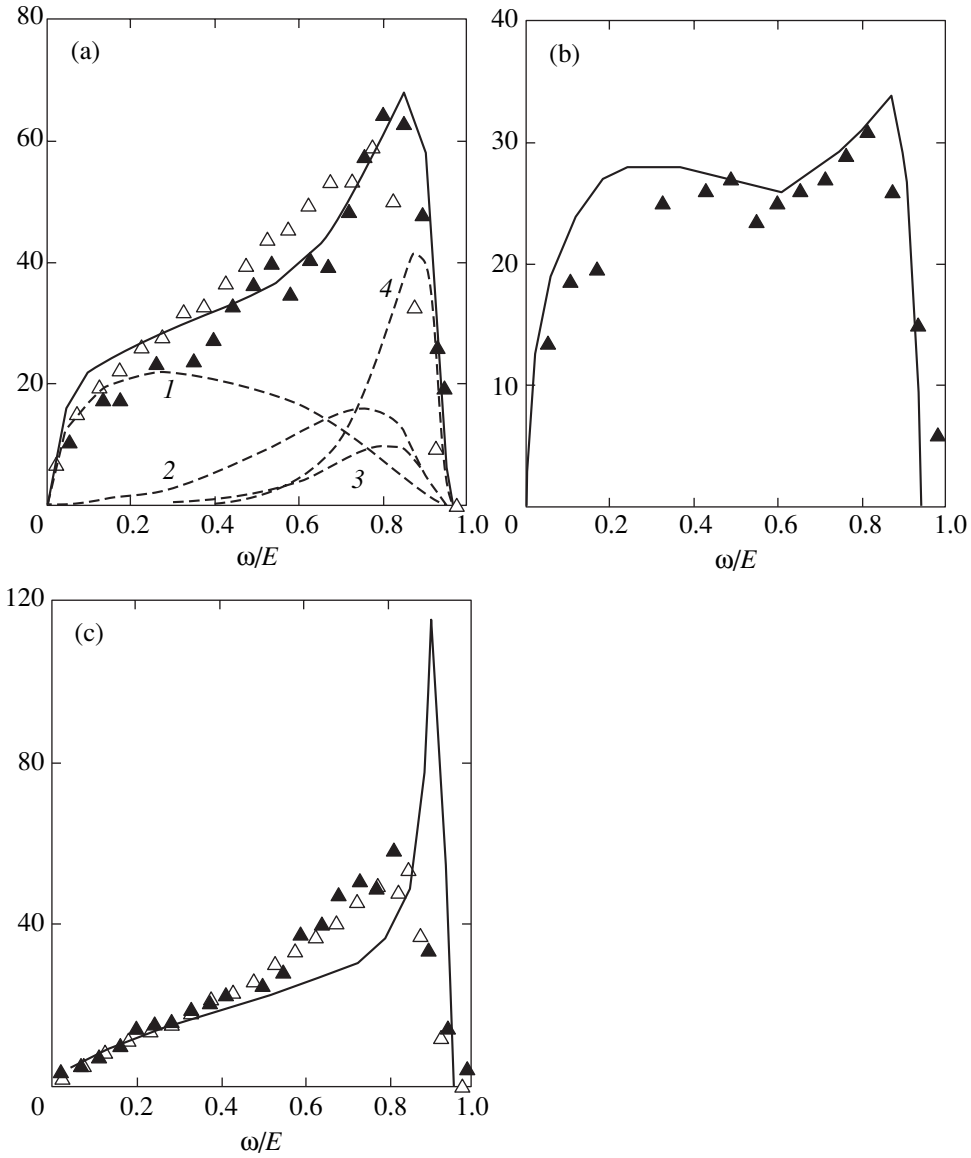



**Fig. 7.** Radiation intensity spectra for 150-GeV electrons in a germanium crystals of thickness 200 (a, b) and 400 (c) μm for incidence angles of 0–7 (a, c) and 15–21 (b) μrad relative to the ⟨110⟩ axis; ▲—experiment [9], △—numerical simulation; solid curves are calculated in the cascade theory and dashed curves describe the contributions from electrons with different transverse energies (see text).

comes from a relatively small number of electrons in deeply bound states of transverse motion (in our example, the number of such electrons does not exceed 10%). It is found that the main contribution to the peak comes from partial functions $g_\omega^{(k)}$ with $k \sim 10$. The results of simulation are shown in Fig. 7a by light symbols. Multiple incoherent scattering of impinging electrons from individual atoms of the crystal was taken into account as well as the damping of transverse energy due to radiation, $\delta\varepsilon = -(\omega/E)[\varepsilon - U(r)]$. The numerical simulation method is described in [19]. In contrast to [19], however, we separately simulated the point at which an emission event takes place (see also [16]).

Figures 7a and 7b illustrate the strong orientation dependence of the spectral characteristics of radiation (the fraction of electrons in a channel is 0.09 under the conditions depicted in Fig. 7b), which is successfully accounted for in the cascade theory. It should be noted that the theory developed here does not contain fitting parameters and leads to correct quantitative results.

However, the correspondence between the cascade theory and experiment is worse for thicker crystals (see Fig. 7c). The disregard of multiple scattering in the cascade theory leads to a sharp peak predicted by the theory but not observed for thick crystals. Nevertheless, in this case also, the theory provides values of the intensity in the soft part of the spectrum, which agree with

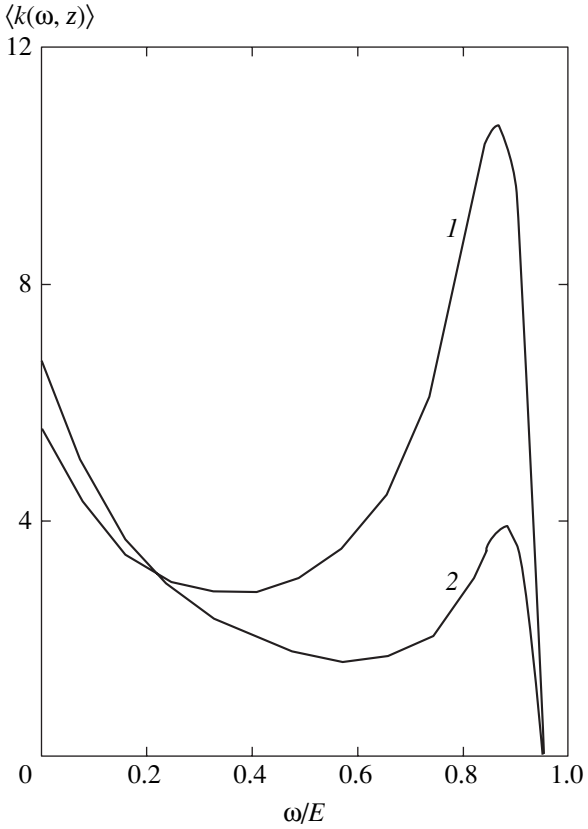



**Fig. 8.** Dependence of the mean number of emitted photons $\langle k(\omega, z)\rangle$ on their total energy $\omega$ for the conditions depicted in Fig. 7a. The angles of incidence are 0–7 (*1*) and 21–27 (*2*) μrad.

the experiment, and correctly predicts the total energy loss by radiation. The total energy loss by radiation is proportional to the area under the curves in Fig. 7. For the conditions depicted in Figs. 7a and 7c, the total energy losses amount to 27 and 44%, respectively, of the initial electron energy, which is close to the experimental results.

The quantity that can be measured in experiments is the average number of emitted photons as a function of their total energy $\langle k(\omega, z)\rangle$, defined by formula (3.18). This quantity for 150-GeV electrons is shown in Fig. 8 for various conditions of their entrance in a 200-μm-thick germanium crystal relative to the ⟨110⟩ direction. The area under the curves in Fig. 8 is equal to the total average number $\langle k(z)\rangle$ of emitted photons over a given thickness $z$. Figure 8 illustrates the strong orientation dependence of the number of emitted photons if the angles of incidence vary within the critical Lindhard angle $\theta_L$. The presence of a peak on the $\langle k(\omega, z)\rangle$ spectra for $\omega \longrightarrow 0$ in Fig. 8 is associated with the behavior of one-photon probabilities in the CFA, $g_\omega^{(1)} \sim \omega^{-2/3}$, which considerably exaggerates the number of soft photons in the radiation spectrum as compared to the results of a more accurate analysis [4].

## 5. CONCLUSIONS

Processes leading to a spread in energy losses can be divided into two groups, viz., the processes in which energy losses in a single interaction event are relatively small and the final distribution is determined by a large number of interactions exceeding $10^2$–$10^3$ (ionization losses) and the processes in which the energy lost in one interaction event is commensurate with the particle energy, while the resultant distribution is determined by the contribution of a relatively small number of interaction events <10–20 (bremsstrahlung and radiation emitted by electrons with energies above 100 GeV in oriented crystals). In the latter case, an adequate description of the energy loss spread corresponding to specific experimental conditions can be obtained as the sum of partial contributions from the probabilities of emission of different numbers of photons (3.6). The expressions for the energy loss distribution function are especially simple when the total interaction cross section is independent of time and energy (which corresponds to a Poisson-type process). In both cases, a large class of phenomena corresponding to power-type differential cross sections of interactions with screening can be described by exact solutions of a relatively simple form, which can be expressed in terms of rapidly converging integrals of elementary functions.

## ACKNOWLEDGMENTS

I am grateful to the late Prof. J. Lindhard, who attracted my attention to useful properties of distribution (3.33), for the opportunity to work at the Århus University in the first half of the 1990 and for numerous fruitful discussions. Thanks are also due to J.U. Anderson, Yu.V. Kononets, H. Nitta, E. Uggerhoj, and A.Kh. Khokonov for discussions clarifying important aspects of this study.

This work was supported by the program "Integration" (grant no. E0203).

## *APPENDIX A*

Formulas (2.33) and (2.34) lead to the gamma distribution (2.39) if we use the relation

$$\Gamma^{-1}(\xi) = \pi^{-1}\xi^{1-\xi}\int_0^\pi \left(\frac{\sin\theta}{\theta}\right)^\xi \exp\left(\frac{\xi\theta}{\tan\theta}\right)d\theta. \qquad \text{(A.1)}$$

This formula follows from the standard representation (see, for example, [39, 8.315])

$$\Gamma^{-1}(\xi) = i(2\pi)^{-1}\int_C (-z)^{-\xi}\exp(-z)dz, \qquad \text{(A.2)}$$

where the integration contour $C$ passes from $+\infty$ along the real axis, bypasses the origin counterclockwise, and returns to $+\infty$. Formula (A.1) can be obtained from for-

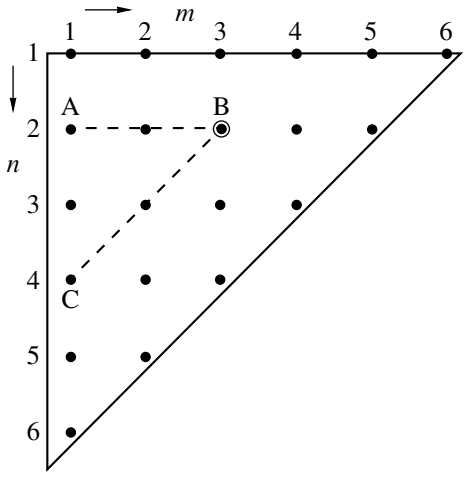



**Fig. 9.** Mesh in variables $u = \omega/E_0$ and $e = E/E_0$ for numerical calculation of partial probabilities (3.15).

mula (A.2) by integration along the steepest descent line (see text).

In contrast to standard representations [39–41], formula (A.1) is convenient in view of its rapidly converging integrand, which makes it very helpful on account of the fact that factor $\Gamma^{-1}(\xi)$ is often encountered in the distribution functions considered here.

## *APPENDIX B*

Standard representations for the Macdonald functions in formula (4.1) have the form (see [39, 8.433])

$$K_{1/3}(\xi) = \sqrt{3}\int_0^\infty \cos\left[\frac{3}{2}\xi\left(x + \frac{x^3}{3}\right)\right]dx, \qquad \text{(B.1)}$$

$$K_{2/3}(\xi) = \sqrt{3}\int_0^\infty x\sin\left[\frac{3}{2}\xi\left(x + \frac{x^3}{3}\right)\right]dx. \qquad \text{(B.2)}$$

Replacing the integration variable in formulas (B.1) and (B.2) by the complex variable $z = x + iy$, we find that the argument of the trigonometric functions in expressions (B.1) and (B.2) has a saddle point $z_0 = i$ lying on the imaginary axis, and the equation for the SDL has the form $Y(x) = (1 + x^2/3)^{1/2}$. Integrating along this line, we obtain

$$K_{1/3}(\xi) = \sqrt{3}\int_0^\infty \exp\left[-\xi\left(1 + \frac{4x^2}{3}\right)Y(x)\right]dx, \qquad \text{(B.3)}$$

$$K_{2/3}(\xi) = \frac{1}{\sqrt{3}}\int_0^\infty (3 + 2x^2) \times \exp\left[-\xi\left(1 + \frac{4x^2}{3}\right)Y(x)\right]\frac{dx}{Y(x)}. \qquad \text{(B.4)}$$

Our representation (B.3) and (B.4) is not encountered in the literature. It is convenient since parameter $\xi$ appears in the argument of the exponential of the rapidly converging integrand, which is beneficial in operations of differentiation and integration with respect to argument as well as in numerical calculations. The Airy function and its derivative in expressions (2.36) and (2.38) are in one-to-one correspondence with $K_{1/3}$ and $K_{2/3}$ and, hence can also be represented by formulas of type (B.3) and (B.4) [40].

## *APPENDIX C*

The problem lies in numerical calculation of integrals of the type

$$F(e, u) = \int_0^u f(e - x, u - x)w(e, x)dx, \qquad \text{(C.1)}$$

where $e$ and $u$ are dimensionless variables varying from zero to unity and corresponding to the electron energy and frequency (energy) of emitted photons. We assume that the integrands in expression (C.1) are defined for all $e > u$.

Integral (C.1) was evaluated on a triangular mesh (see Fig. 9) with the same number of partitions and identical steps $\delta$ in both variables,

$$\delta \equiv \Delta e = \Delta u = \frac{1 - 2u_{\min}}{N - 1}.$$

Here, $N$ is the maximal number of points at which $e$ and $u$ are calculated ($N = 6$ in Fig. 9), $u_{\min} = 1 - u_{\max}$, where $u_{\min}$ and $u_{\max}$ are the minimal and the maximal photon frequencies for which integrals (C.1) are evaluated. In our case, $u_{\min} = 0.0051529$. Such a choice of quantity $u_{\min}$ is dictated by the algorithm of numerical integration of functions with a singularity of the type $u^{-2/3}$ in the left integration limit. The values of frequency variable $m$ increase along the abscissa axis, while variable $n$ increases in the negative direction of the ordinate axis.

The functions appearing in the integrand of formula (C.1) are juxtaposed to the mesh functions $f_{nm}$ and $w_{nm}$, the first index corresponding to electron energy $e_n$ and the second, to photon frequency $u_m$. Then the $n$th value of energy on the mesh corresponds to $N - n + 1$ values of frequency variable $m$. By definition, $e_1 = 1$, $e_N = 1 - 2u_{\min}$, $u_1 = u_{\min}$, and $u_N = 1 - u_{\min}$.

On this mesh, integral (C.1) can be evaluated by the formula (method of trapezoids)

$$F_{nm} = \delta\sum_{k=1}^{m}\left(1 - \frac{1}{2}\delta_{1k} - \frac{1}{2}\delta_{mk}\right) \times f_{n+k-1, m-k+1}w_{nk}, \qquad \text{(C.2)}$$

where $\delta_{ij}$ is the Kronecker delta.

In accordance with this formula, the values of $f_{ij}$ in the indicated sum on a mesh of the type shown in Fig. 9 for given $e_n$ and $u_m$ (encircled point in the figure) are taken on the diagonal connecting points $(n, m)$ and $(n + m - 1, 1)$ (dashed line BC in Fig. 9, $n = 2$ and $m = 3$). The values of $w_{ij}$ in Eq. (C.2) are taken along the line $n = \text{const}$ (line AB). The values of integral (C.1) between zero and $u = u_1 = u_{\min}$ were calculated analytically taking into account the smallness of $u_{\min} \ll 1$; i.e., integral (C.1) was evaluated using formula (C.2) in the limits $[u_{\min}, u]$. At each stage of computations, the accuracy was corrected with the help of the normalization condition and the results were verified by comparing with analytic solutions.

## REFERENCES


1. H. Bethe and W. Heitler, Proc. R. Soc. London, Ser. A **146**, 83 (1934).
2. L. Landau, J. Phys. (Moscow) **8**, 201 (1944); in *Collected Works* (Nauka, Moscow, 1969), Vol. 1, p. 482 [in Russian].
3. J. Schwinger, Phys. Rev. **75**, 1912 (1949).
4. M. Kh. Khokonov and H. Nitta, Phys. Rev. Lett. **89**, 094801 (2002).
5. J. Lindhard and V. Nielsen, K. Dan. Vidensk. Selsk. Mat. Fys. Medd. **38** (9) (1971).
6. J. Lindhard and V. Nielsen, K. Dan. Vidensk. Selsk. Mat. Fys. Medd. **39** (1) (1974).
7. J. Lindhard, Phys. Scr. **32**, 72 (1985).
8. P. V. Vavilov, Zh. Éksp. Teor. Fiz. **32**, 920 (1957) [Sov. Phys. JETP **5**, 749 (1957)].
9. R. R. Medenwaldt, S. P. Moller, S. Tang-Peterson, *et al.*, Phys. Lett. B **242**, 517 (1990); R. Medenwaldt, S. P. Moller, A. H. Sorensen, *et al.*, CERN/SPSC 90-31 SPSC/P234 Add. 3 (1990).
10. K. Kirsebom, R. Medenwaldt, U. Mikkelsen, *et al.*, Nucl. Instrum. Methods Phys. Res. B **119**, 79 (1996).
11. A. Baurichter, K. Kirsebom, Yu. V. Kononets, *et al.*, Phys. Rev. Lett. **79**, 3415 (1997).
12. K. Kirsebom, U. Mikkelsen, E. Uggerhoj, *et al.*, Nucl. Instrum. Methods Phys. Res. B **174**, 274 (2001).
13. V. N. Baier and V. M. Katkov, Phys. Rev. D **59**, 056003 (1999).
14. X. Artru, Phys. Lett. A **128**, 302 (1998); Nucl. Instrum. Methods Phys. Res. B **48**, 278 (1990).
15. Yu. V. Kononets and V. A. Ryabov, Nucl. Instrum. Methods Phys. Res. B **48**, 269 (1990); Nucl. Instrum. Methods Phys. Res. B **48**, 274 (1990).
16. Yu. V. Kononets, J. Mosc. Phys. Soc. **2**, 71 (1992).
17. V. V. Beloshitsky, M. A. Kumakhov, and A. Kh. Khokonov, Nucl. Instrum. Methods Phys. Res. B **62**, 207 (1991).
18. M. Kh. Khokonov, Pis'ma Zh. Éksp. Teor. Fiz. **56**, 349 (1992) [JETP Lett. **56**, 333 (1992)].
19. M. Kh. Khokonov, Zh. Éksp. Teor. Fiz. **103**, 1723 (1993) [JETP **76**, 849 (1993)].
20. A. I. Akhiezer and N. F. Shul'ga, *High-Energy Electrodynamics in Matter* (Nauka, Moscow, 1993; Gordon and Breach, Luxembourg, 1996), Chap. 9.
21. G. V. Dedkov and R. I. Tegaev, Radiat. Eff. **91**, 265 (1986).
22. W. Feller, *An Introduction to Probability Theory and Its Applications*, 3rd ed. (Wiley, New York, 1967; Mir, Moscow, 1984), Vol. 2.
23. H. Bethe, Phys. Rev. **89**, 1256 (1953).
24. L. Meyer, Phys. Status Solidi B **44**, 253 (1971).
25. M. Kh. Khokonov, Phys. Scr. **55**, 513 (1997).
26. J. Lindhard, K. Dan. Vidensk. Selsk. Mat. Fys. Medd. **34** (14) (1964).
27. A. Kh. Khokonov, M. Kh. Khokonov, and A. A. Kizdermishov, Zh. Tekh. Fiz. **72** (3), 69 (2002) [Tech. Phys. **47**, 350 (2002)].
28. M. A. Kumakhov, Phys. Lett. A **57A**, 17 (1976).
29. J. Lindhard, private communication (1990).
30. V. V. Beloshitsky and F. F. Komarov, Phys. Rep. **93**, 117 (1982).
31. I. C. Kimball and N. Cue, Phys. Rep. **125**, 68 (1985).
32. N. P. Klepikov, Zh. Éksp. Teor. Fiz. **26**, 19 (1954).
33. V. N. Baĭer, V. M. Katkov, and V. M. Strakhovenko, Usp. Fiz. Nauk **159**, 455 (1989) [Sov. Phys. Usp. **32**, 972 (1989)].
34. K. Kirsebom, U. Mikkelsen, E. Uggerhoj, *et al.*, Phys. Rev. Lett. **87**, 054801 (2001).
35. V. V. Beloshitskiĭ and M. A. Kumakhov, Zh. Éksp. Teor. Fiz. **82**, 462 (1982) [Sov. Phys. JETP **55**, 265 (1982)].
36. V. I. Telegin and M. Kh. Khokonov, Zh. Éksp. Teor. Fiz. **83**, 260 (1982) [Sov. Phys. JETP **56**, 142 (1982)].
37. F. F. Komarov, A. F. Komarov, and M. Kh. Khokonov, Zh. Éksp. Teor. Fiz. **93**, 41 (1987) [Sov. Phys. JETP **66**, 23 (1987)].
38. R. A. Doyle and R. S. Turner, Acta Crystallogr. A **24**, 390 (1968).
39. I. S. Gradshteĭn and I. M. Ryzhik, *Table of Integrals, Series, and Products* (Fizmatgiz, Moscow, 1963; Academic, New York, 1980).
40. *Handbook of Mathematical Functions*, Ed. by M. Abramowitz and I. A. Stegun (Dover, New York, 1971; Nauka, Moscow, 1979).
41. E. Jahnke, F. Emde, and F. Losch, *Tables of Higher Functions*, 6th ed. (McGraw-Hill, New York, 1960; Nauka, Moscow, 1968).


*Translated by N. Wadhwa*

*Journal of Experimental and Theoretical Physics, Vol. 99, No. 4, 2004, pp. 708–718.*
*Translated from Zhurnal Éksperimental'noĭ i Teoreticheskoĭ Fiziki, Vol. 126, No. 4, 2004, pp. 819–832.*


**PLASMA, GASES**

# Plasma Satellites of X-ray Lines of Ions in a Picosecond Laser Plasma

**V. S. Belyaev[a], V. I. Vinogradov[a], A. S. Kurilov[a], A. P. Matafonov[a], V. S. Lisitsa[b], V. P. Gavrilenko[c,d], A. Ya. Faenov[e], T. A. Pikuz[e], I. Yu. Skobelev[e], A. I. Magunov[d,e], and S. A. Pikuz, Jr.[f]**

[a]*Central Research Institute of Machine Building, ul. Pionerskaya 4, Korolev, Moscow oblast, 141070 Russia*
[b]*Institute of Nuclear Fusion, Russian Research Center Kurchatov Institute, pl. Akademika Kurchatova 1, Moscow, 123182 Russia*
[c]*FGUP "Research Center for Surface and Vacuum Properties," State Standard of Russia, Moscow, 119421 Russia*
[d]*Prokhorov Institute of General Physics, Russian Academy of Sciences, Moscow, 119991 Russia*
[e]*Multicharged Ion Spectra Data Center, VNIIFTRI, Mendeleevo, Moscow oblast, 141570 Russia*
[f]*Department of Physics, Moscow State University, Vorob'evy gory, Moscow, 119992 Russia*
*e-mail: VadimBelyaev@mtu-net.ru*
Received April 5, 2004

**Abstract**—We present the results of our measurements of the spectra for multicharged ions in a plasma produced by moderately intense (about $10^{17}$ W cm$^{-2}$) picosecond laser pulses. They suggest the existence of intense plasma oscillations with a frequency appreciably lower than the frequency of the laser radiation. The observed spectrum for the plasma satellites of the Lyman $Ly_\alpha$ doublet of the hydrogenic F IX ion in a dense plasma was modeled theoretically. The resulting doublet profile was shown to have a complex structure that depends nontrivially both on the plasma density and on the frequency and amplitude of the plasma oscillations. The positions of the satellites and their separations allowed them to be associated with intense electrostatic oscillations with an amplitude of $(4–6) \times 10^8$ V cm$^{-1}$ and a frequency near $(0.7–1) \times 10^{15}$ s$^{-1}$. Assuming the oscillation frequency to be determined by the strength of the magnetic field $B$ generated in the plasma, we obtained an estimate of $B$ that is in reasonable agreement with other measurements and estimates of this quantity. Our theoretical analysis allowed explanation of the emission spectra observed when flat fluoroplastic targets were heated by intense picosecond laser pulses. © *2004 MAIK "Nauka/Interperiodica".*

## 1. INTRODUCTION

The X-ray spectra of a high-temperature plasma are shaped mainly by such atomic processes as electron–ion collisions and radiative or autoionization ionic level decays. The presence of slowly varying (quasi-static) electric and magnetic fields in a plasma also affects its radiative characteristics, causing the profiles of the emitted spectral lines to change mainly through the Zeeman and Stark effects.

In addition, strong oscillating electromagnetic fields attributable both to external exposure of the plasma (laser or microwave emissions used for plasma heating) and to the growth of strong plasma instabilities and the corresponding plasma oscillations can exist in the plasma.

The peculiarities of a plasma produced by intense sub- and picosecond laser pulses stem from the fact that the electron velocity distribution is highly anisotropic. This anisotropy gives rise to intense electromagnetic oscillations attributable to the growth of instabilities in such a plasma, primarily the so-called Weibel instability. Its growth is known to lead to the generation of strong quasi-static magnetic fields with the strength determined by the laser radiation intensity and reaching several hundred megagauss at intensities above $10^{19}$ W cm$^{-2}$. Recent experiments [1] have revealed such superstrong quasi-static magnetic fields generated in a laser plasma. Their presence, irrespective of the generation mechanism, is fundamentally important, because it radically changes the physical properties of the laboratory plasma. The experimentally detected magnetic fields [1] are ~500 MG, and they affect the propagation of ordinary and extraordinary waves in the visible spectral range. Naturally, the presence of magnetic fields also affects the pattern and frequency distribution of the plasma oscillations.

The frequencies of the electromagnetic oscillations in a laser plasma typically do not exceed $\hbar\omega \sim 1–10$ eV. Such oscillations may have the strongest effect on the emission spectra in the visible spectral range, where the atomic transition energy is comparable to the oscillation energy. A characteristic spectroscopic manifestation of the oscillating electromagnetic fields in a plasma

is the emergence of additional spectral lines (satellites) with the frequencies

$$\omega_s = \omega_a \pm l\omega,$$

where $\omega_a$ is the atomic transition frequency; and $l = 1, 3, 5, \ldots$ and $l = 2, 4, 6, \ldots$ are for parity-forbidden and parity-permitted atomic transitions, respectively. Depending on whether the oscillating field is an external laser field or generated by plasma oscillations, these satellites are called laser (laser satellites were detected, e.g., in [2–7]) or plasma (plasma satellites were detected, e.g., in [8, 9]). If the atomic transition frequency lies in the X-ray spectral range (multicharged ions with $\hbar\omega_a > 1$ keV), then, in general, the emerging satellites will no longer be isolated lines, but will form the profiles of the observed X-ray transitions. Note that their detection requires using high-resolution ($\lambda/\Delta\lambda \sim 10^3$–$10^4$) X-ray spectrographs. However, such X-ray instruments currently exist and are widely used in X-ray spectral studies of high-temperature plasmas.

Detection of the spectroscopic effects attributable to plasma oscillations is undoubtedly of interest, first, as an independent confirmation of the emergence of these oscillations and, second, as a method for measuring their parameters (amplitude, frequency, polarization). Of particular interest is the case where the frequency of these oscillations is close to the electron cyclotron frequency, which corresponds to the growth of Bernstein-mode-type plasma oscillations in a magnetic field. Observations of such oscillations could serve as an independent method for measuring the magnetic fields generated in a plasma.

The observed spectral characteristics must be compared with calculations that take into account the simultaneous presence of oscillating and quasi-static fields in a plasma and that are performed for actual atomic systems whose energy level structure can be fairly complex.

In this paper, we present the results of our measurements of the spectra for multicharged ions in a plasma produced by moderately intense (about $10^{17}$ W cm$^{-2}$) picosecond laser pulses, suggesting the existence of intense plasma oscillations with a frequency appreciably lower than the frequency of the laser radiation.

The observed spectrum for the plasma satellites of the Lyman $2p_{1/2,\,3/2}$–$1s_{1/2}$ doublet of the hydrogenic fluorine ion in a dense plasma was modeled theoretically. The resulting doublet profile was shown to have a complex structure that depends nontrivially both on the plasma density and on the frequency and amplitude of the plasma oscillations. In this paper, we assume that the oscillation frequency is determined by the strength of the magnetic field $B$ generated in a plasma. Under this assumption, we were able to obtain an estimate for $B$ that is in reasonable agreement with other measurements and estimates of this quantity. Our theoretical analysis allowed us to explain the emission spectra observed when flat fluoroplastic targets were heated by intense picosecond laser pulses.

## 2. EXPERIMENTAL RESULTS

We carried out our experiments with the Neodim terawatt laser facility [10]. This laser facility provided the following laser pulse parameters: an energy as high as 1.7 J, a wavelength of 1.055 μm, and a duration of 1.5 ps. A laser beam 60 mm in diameter was focused on the target by a 14.5-mm-thick aspherical lens with a focal length of 140 mm that concentrated 50% of the beam energy into a circle 15 μm in diameter. As a result, the intensity of the beam when focused on the target reached $3 \times 10^{17}$ W cm$^{-2}$. Flat 200-μm-thick fluoroplastic plates were used as the targets. The residual gas pressure in the vacuum chamber did not exceed $10^{-3}$ Torr.

The X-ray radiation from a plasma produced by the interaction of a laser pulse with the target (see Fig. 1) was recorded by two FSPR spectrographs [11] with spherically bent quartz or mica crystals (the radius of curvature of the crystal surface was 150 mm). In all experiments, the angle of observation was 5° and 85° to the normal to the target surface for spectrographs *1* and *2*, respectively. A Kodak-2492 X-ray photographic film protected from visible light by two 1-μm-thick polypropylene layers with evaporated Al 0.2 μm in total thickness was used as the detector of the radiation reflected from the crystal.

The plasma emission spectra were investigated in the spectral range 1.49–1.51 nm. Under experimental conditions, the spectrographs provided a spectral resolution $\lambda/\Delta\lambda$ of at least 5000. Figure 1 shows typical F IX $Ly_\alpha$ line spectrograms emitted by the plasma in different directions: almost along the normal to the target and parallel to its surface. Both spectrograms were taken with a spatial resolution, but the spatial resolution was along the $x$ axis (i.e., in the target plane) in the former case and along the $z$ axis (i.e., in the direction of the preferential plasma expansion) in the latter case.

These spectrograms show that the size of the plasma emission region is large, about 800 μm along both the $x$ and $z$ axes. (Recall that the diameter of the laser focusing spot was about 15 μm.) We see from the spectrogram shown in Fig. 1b that the profile of the $Ly_\alpha$ line recorded along the $x$ axis is symmetric. Since its full width at half maximum (FWHM) is no larger than 0.04 Å, this implies that the plasma expansion velocity component $v_x$ does not exceed $4 \times 10^7$ cm s$^{-1}$. We see from the spectrogram shown in Fig. 1a that in the region $x \approx$ –100–400 μm, the $Ly_\alpha$ line profile is asymmetric and have deep dips and local peaks in the long-wavelength wing. Since $v_x < 4 \times 10^7$ cm s$^{-1}$, the emission from this spatial region occurs $t \sim 0.25$–1 ns after the onset of plasma expansion, i.e., much later than the termination of the heating laser picosecond pulse. Thus,

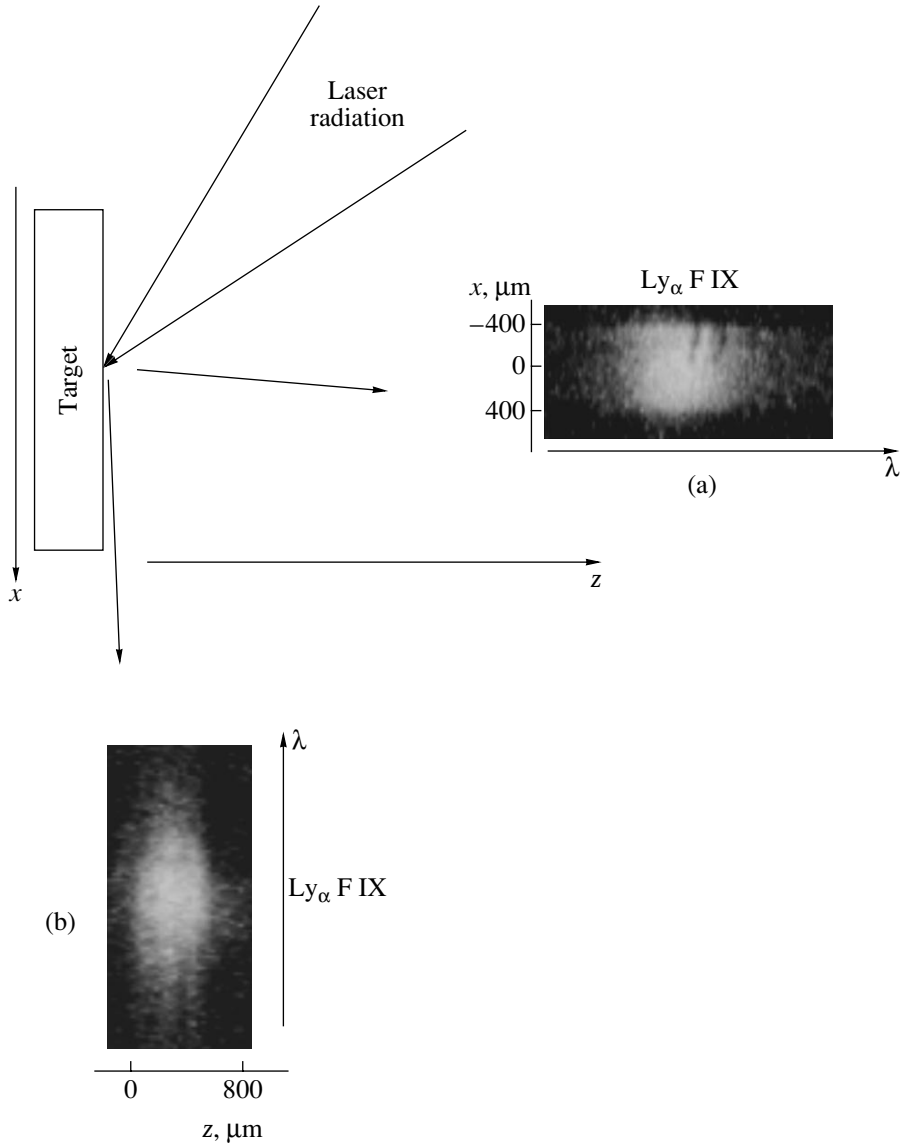



**Fig. 1** Experimental setup for observing the $Ly_\alpha$ line profiles of F IX in a laser plasma and spectrograms taken with a spatial resolution along (a) and perpendicular (b) to the surface of a fluoroplastic target.

the spectral features emitted from the plasma regions with $x \approx -100$–$400$ μm cannot be laser satellites.

In general, the dips in the $Ly_\alpha$ line profile could be associated with self-absorption. Whereas the dips (due to self-absorption) for a homogeneous plasma emerge only at the line center, the dips for an inhomogeneous plasma, in which the absorbing and emitting regions have different velocities, can emerge due to the Doppler shift in both the short- and long-wavelength wings, depending on whether the absorbing region moves toward the observer faster or more slowly than the emitting region (see, e.g., [12]). However, two dips (absorption on the $Ly_1$ and $Ly_2$ components) whose separation must have been exactly equal to the fine splitting (5.4 mÅ for F IX) must have emerged in any case. Since the separation between the dips in our experiments, first, changed with laser flux density from 8 to 11 mÅ and, second, differed significantly from the fine splitting, the observed profiles cannot be explained in terms of self-absorption.

Figure 2 shows the $Ly_\alpha$ line profiles observed from various spatial plasma regions for two experiments where the laser flux density was $q_{las} = 2 \times 10^{17}$ (a) and $3 \times 10^{17}$ W cm$^{-2}$ (b). We see that the observed profiles depend most strongly on $q_{las}$. However, even at the same intensity, different spatial plasma regions yield significantly differing structures of the emission spectrum.

Note also that the spectra in our experiments were recorded without a time resolution; i.e., the experimental spectra are an integral of the plasma emission over its lifetime. Since the density and temperature of the

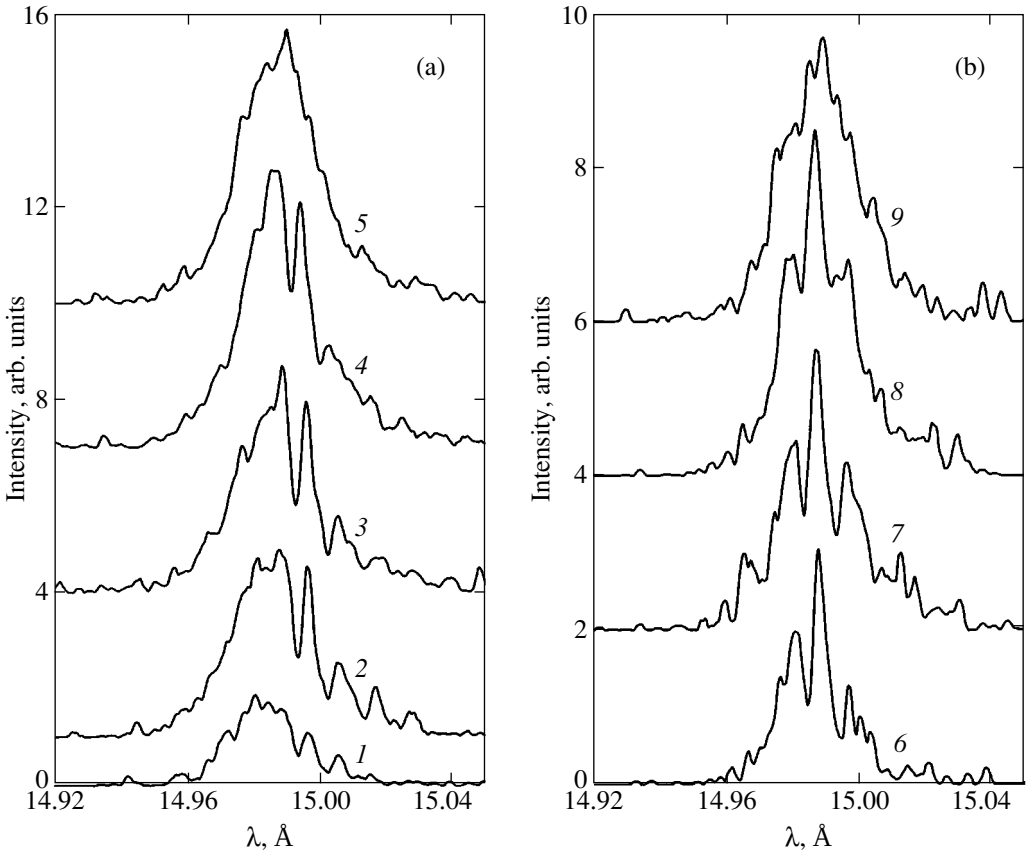



**Fig. 2.** $Ly_\alpha$ line profiles of F IX emitted by various spatial plasma regions: $x$ = –400 (*1*, *6*), –315 (*2*), –230 (*3*, *7*), –60 (*4*, *8*), and 0 μm (*5*, *9*). The laser flux density was $q_{las} = 2 \times 10^{17}$ and $3 \times 10^{17}$ W cm$^{-2}$ in experiments (a) and (b), respectively. For clarity, the profiles were displaced along the vertical axis.

plasma during its evolution change and since the parameters of the oscillating field can also change, describing the observed spectrum by only one set of parameters $N_e$, $T$, ω, and $E_0$ is a rough approximation. However, as we show below, this approximation yields excellent qualitative (and even quantitative) agreement with experimental data.

## 3. CALCULATIONS OF THE $Ly_\alpha$ EMISSION LINE SPECTRUM FOR F IX IN A DENSE PLASMA

Calculations of the emission spectra under the simultaneous influence of the quasi-static fields of ions and the variable fields of plasma oscillations underlie any theoretical modeling. We assumed that the quasi-static field generated by plasma ions obeyed the Holtsmark distribution with an ion density of about $10^{20}$ cm$^{-3}$. The electric field strength of the plasma oscillations has the largest uncertainty. This strength is determined by the plasma turbulence level, which is characterized by the ratio of the oscillation energy density to the plasma thermal energy density. In our calculations, it was varied over a wide range corresponding to various turbulence levels from $10^{-1}$ to $10^{-3}$. As regards the plasma oscillation frequency, it was also varied to obtain the best fit to the experimental data. This frequency proves to be appreciably lower than the frequency of the incident laser radiation. The frequency scale of such oscillations may be related to the electron cyclotron frequency of the oscillations in a magnetic field with the characteristic scale determined by the Weibel instability.

A characteristic feature of the experimental $Ly_\alpha$ line profiles for F IX ions is the presence of many peaks and dips in the profiles (see Fig. 2). Although the Stark effect attributable to the influence of quasi-static electric plasma ion microfields **F** on hydrogenic ions plays an important role in shaping the $Ly_\alpha$ line profiles, this effect alone cannot explain the severe raggedness of the observed $Ly_\alpha$ line profiles for F IX ions. The additional influence of a quasi-monochromatic electric field of the following form is the principal mechanism that can give rise to features (peaks and dips) in the smooth quasi-static line profiles of the hydrogenic ions formed by the low-frequency ion electric microfields **F**:

$$\mathbf{E}(t) = \sum_k \mathbf{E}_{0k} \cos(\omega_k t + \alpha_k)$$

(see, e.g., [13–15]). Here, we assume that the characteristic spectral width of the oscillating field $\mathbf{E}(t)$ is much smaller than the frequencies of the individual harmonics $\omega_k$. The specific form of these features depends on the parameters of the electric fields $\mathbf{F}$ and $\mathbf{E}(t)$ (on the characteristic strengths of these fields, the characteristic frequency of the field $\mathbf{E}(t)$, and the pattern of $\mathbf{E}(t)$ in a plasma), the mutual arrangement of the atomic energy levels, and the matrix elements of the dipole moment between these atomic levels.

The $n = 2$ energy level of a hydrogenic ion consists of three sublevels: $2P_{3/2}$, $2S_{1/2}$, and $2P_{1/2}$. In the absence of electric fields, the $Ly_\alpha$ spectral line is produced by two radiative transitions: $2P_{3/2} \longrightarrow 1S_{1/2}$ and $2P_{1/2} \longrightarrow 1S_{1/2}$. For F IX, the $2P_{3/2} \longrightarrow 1S_{1/2}$ and $2P_{1/2} \longrightarrow 1S_{1/2}$ spectral components are spaced approximately $\delta\lambda = 5.4 \times 10^{-4}$ nm apart. Note that the value of the fine structure (i.e., the separation between the $2P_{3/2}$ and $2P_{1/2}$ levels) on the frequency scale is $\delta\omega = 4.54 \times 10^{14}$ s$^{-1}$. Stark state mixing of the $2P_{3/2}$, $2S_{1/2}$, and $2P_{1/2}$ levels takes place under the influence of plasma electric microfields. This effect, together with the Doppler effect, causes a rearrangement of the $Ly_\alpha$ line profile: the individual $2P_{3/2} \longrightarrow 1S_{1/2}$ and $2P_{1/2} \longrightarrow 1S_{1/2}$ components disappear, and one broad resulting profile appears in place of them.

Here, to theoretically analyze the experimental $Ly_\alpha$ line profiles for F IX ions, we used a model in which each F IX ion was assumed to experience two electric fields: a quasi-static field $\mathbf{F}$ (generated by plasma ions) and a linearly polarized harmonic field $\mathbf{E}(t) = \mathbf{E}_0\cos\omega t$. Since there was no a priori information about the possible amplitude–angular distribution of the oscillating field $\mathbf{E}(t)$ in a plasma, we considered the simplest case where $\mathbf{F} \parallel \mathbf{E}(t)$ in terms of our model. The Holtsmark function [16] was used as the distribution function of the ion microfield strength $W(F)$. Although the distribution function of the plasma ion microfields $W(F)$ slightly differs from the Holtsmark function due to the cross correlation between the ions and the Debye screening of the ion-produced electric fields for the plasma parameters under consideration (the electron temperature $T_e$ is more than or ~100 eV, and the plasma density $N_e$ is more than or ~$10^{20}$ cm$^{-3}$), we disregarded this difference, because our prime objective was to describe the features observed in the experimental $Ly_\alpha$ line profiles for F IX ions.

When theoretically analyzing the modification of the $Ly_\alpha$ emission spectrum for F IX ions under the influence of two electric fields, the ion microfield $\mathbf{F}$ and the oscillating field $\mathbf{E}(t) = \mathbf{E}_0\cos\omega t$, we chose the quantization $z$ axis of the Cartesian coordinate system along the vector $\mathbf{F}$. (Here, the coordinate origin is assumed to coincide with the F IX nucleus.) Let us denote the wavefunctions $\varphi_j$ ($j = 1, 2, \ldots, 8$) of the states belonging to the level with the principal quantum number $n = 2$ for F IX as follows:

$$\varphi_{1,4} \equiv |2S_{1/2}, m_j\rangle, \quad \varphi_{2,5,7,8} \equiv |2P_{3/2}, m_j\rangle,$$

$$\varphi_{3,6} \equiv |2P_{1/2}, m_j\rangle,$$

$m_j = 1/2$ for $\varphi_1$, $\varphi_2$, and $\varphi_3$; $m_j = -1/2$ for $\varphi_4$, $\varphi_5$, and $\varphi_6$; $m_j = 3/2$ for $\varphi_7$; and $m_j = -3/2$ for $\varphi_8$. In our model ($\mathbf{F} \parallel \mathbf{E}(t) \parallel z$), the combined electric field $\boldsymbol{\varepsilon}(t) = \mathbf{F} + \mathbf{E}_0\cos\omega t$ interacts with each of the two individual three-level systems:

$$\varphi_1 \rightleftharpoons \varphi_2 \rightleftharpoons \varphi_3, \quad \varphi_4 \rightleftharpoons \varphi_5 \rightleftharpoons \varphi_6.$$

In this case, to calculate the $Ly_\alpha$ emission spectrum, it will suffice to consider the radiative transitions from the three level system $\varphi_1 \rightleftharpoons \varphi_2 \rightleftharpoons \varphi_3$ perturbed by the field $\boldsymbol{\varepsilon}(t)$ to the lower $n = 1$ level and to take into account the spectral component corresponding to the radiative transition from the $\varphi_7$ state to the $n = 1$ state. The radiative transitions to the $n = 1$ level from the states with negative magnetic quantum numbers $m_j$ ($m_j = -1/2, -3/2$) belonging to the upper $n = 2$ level yield an emission spectrum that matches the emission spectrum for the transitions from the states with positive $m_j$ ($m_j = 1/2, 3/2$).

To determine the evolution of the F IX states in the combined electric field $\boldsymbol{\varepsilon}(t)$, we solved the Schrödinger equation

$$i\frac{\partial\Psi}{\partial t} = [\hat{H}_0 + \hat{V}_1 + \hat{V}_2(t)]\Psi. \qquad (1)$$

In Eq. (1) and below, unless specified otherwise, we use atomic units: $\hbar = e = m_e = 1$. In Eq. (1), $\hat{H}_0$ is the unperturbed Hamiltonian for F IX with the energy eigenvalues $\varepsilon_j^{(0)}$ and the corresponding wave eigenfunctions $\varphi_j$, $\hat{V}_1 = zF$ is the operator for the dipole interaction of F IX with the field $\mathbf{F}$, and $\hat{V}_2(t) = zE_0\cos\omega t$ is the operator for the dipole interaction of F IX with the oscillating field $\mathbf{E}_0\cos\omega t$. Analyzing the behavior of the three-level system $\varphi_1 \rightleftharpoons \varphi_2 \rightleftharpoons \varphi_3$ in the field $\boldsymbol{\varepsilon}(t) = \mathbf{F} + \mathbf{E}_0\cos\omega t$, we represent the solution of the Schrödinger equation (1), according to the Floquet theorem (see, e.g., [17]), as the wavefunction of a quasi-energy state

$$\Psi(t) = \exp(-i\mu t)\sum_{j=1}^{3}\sum_{s=-\infty}^{+\infty} A_{js}\exp(-is\omega t)\varphi_j. \qquad (2)$$

In expression (2), the quasi-energy $\mu$ and the time-independent coefficients $A_{js}$ are unknown quantities. Substituting (2) into (1), we obtain a system with an infinite

number of algebraic equations with constant coefficients after simple transformations:

$$(\varepsilon_k^{(0)} - \mu - s\omega)A_{k\omega} + \sum_{j=1}^{3}[A_{js}\langle\varphi_k|\hat{V}_1|\varphi_j\rangle + (A_{j,s-1} + A_{j,s+1})\langle\varphi_k|\hat{V}_2|\varphi_j\rangle] = 0, \quad (3)$$

where $k = 1, 2, 3$; $s = 0, \pm1, \pm2, \ldots$; and $\hat{V}_2 = zE_0/2$. To obtain a system with a finite number of algebraic equations from (3), we will retain only the terms containing $A_{j0}$, $A_{j1}$, $A_{j,-1}$, $A_{j2}$, and $A_{j,-2}$ in (3). This imposes an upper limit on the amplitude of the oscillating field in our calculations:

$$E_0\ [\mathrm{V\ cm^{-1}}] < 11.5 \times 10^{-7}\omega\ [\mathrm{s^{-1}}].$$

As a result, we obtain a system of 15 linear algebraic equations with constant coefficients in place of (3). The system of equations derived in this way describes the behavior of an effective 15-level quantum-mechanical system under the influence of a constant (time-independent) perturbation $\hat{U}$. The unperturbed energy levels for this effective system are

$$E_{ks}^{(0)} = \varepsilon_k^{(0)} + s\omega, \quad (4)$$

where $k = 1, 2, 3$ and $s = 0, \pm1, \pm2$. The matrix elements of the operator $\hat{U}$ calculated using the wavefunctions $|ks\rangle$ that correspond to the unperturbed energy levels $E_{ks}^{(0)}$ in (4) can be easily deduced from system (3). They are a linear combination of the matrix elements for the operators $\hat{V}_1$ and $\hat{V}_2$ calculated using the wavefunctions $\varphi_k$ ($k = 1, 2, 3$) of the initial three-level system $\varphi_1 \rightleftarrows \varphi_2 \rightleftarrows \varphi_3$. Thus, we reduced the problem of the dynamic Stark effect for the three-level system $\varphi_1 \rightleftarrows \varphi_2 \rightleftarrows \varphi_3$ in the electric field $\boldsymbol{\varepsilon}(t) = \mathbf{F} + \mathbf{E}_0\cos\omega t$ to the simpler problem of the evolution of an effective 15-level system that undergoes a static perturbation $\hat{U}$. We solved the latter problem numerically by diagonalizing the corresponding $15 \times 15$ energy matrix. This allowed us to numerically calculate the three quasi-energies ($\mu_1$, $\mu_2$, and $\mu_3$) and the corresponding three wavefunctions $\Psi_p(t)$ (see (2)) for the three-level system $\varphi_1 \rightleftarrows \varphi_2 \rightleftarrows \varphi_3$:

$$\Psi_p(t) = \exp(-i\mu_p t)\sum_{j=1}^{3}\sum_{s=-2}^{2} A_{js}^{(p)}\exp(-is\omega t)\varphi_j, \quad p = 1, 2, 3. \quad (5)$$

Using expression (5), we can represent the emission spectrum for the radiative transition from the three-level system $\varphi_1 \rightleftarrows \varphi_2 \rightleftarrows \varphi_3$ to the lower $n = 1$ level as

$$I(\Delta\omega, F) = \sum_{p=1}^{3}\sum_{j=1}^{3}\sum_{s=-2}^{2}\left|A_{js}^{(p)}\right|^2|\langle\varphi_j|z|\varphi_0\rangle|^2 \times \delta(\Delta\omega - \mu_p - s\omega), \quad (6)$$

where $\varphi_0$ is the wavefunction of the $n = 1$ state and all energies are assumed to be measured from the energy of the $n = 1$ level. Note that the arguments of the $\delta$ function in (6) specify the positions of the spectral components:

$$\Delta\omega_{p,s} = \mu_p + s\omega, \quad (7)$$

where $p = 1, 2, 3$ and $s = 0, \pm1, \pm2$.

Based on formula (6), our computer code numerically computed the "Doppler" spectrum $I_D(\Delta\omega)$ that was obtained by substituting Gaussian profiles with the FWHM determined by the F IX ion temperature for the $\delta$ functions in (6). To find the resulting $\mathrm{Ly}_\alpha$ line profile, our code then numerically averaged the Doppler spectrum $I_D(\Delta\omega)$ over the ion microfield distribution $W(F)$ and added the Doppler profile of the line corresponding to the radiative transition from the $\varphi_7$ state to the $n = 1$ state to the spectrum obtained in this way.

Thus, in our model, the computed line profile depended on four parameters: the plasma temperature $T$ and density $N$ as well as the frequency $\omega$ and amplitude $E_0$ of the oscillating field. The results of our calculations for various values of these parameters are presented in Figs. 3–6.

Figure 3 shows how the emission spectrum in the wavelength range 14.94–15.04 Å depends on the amplitude of the oscillating electric field $E_0$ at its frequencies $\omega = 5 \times 10^{14}$ (a) and $10^{15}$ $\mathrm{s^{-1}}$ (b). We see that a change in $E_0$ leads not only to an intensity redistribution between the emerging spectral components, but also to a change in their positions. The latter stems from the fact that the positions of the spectral components are determined (at fixed ion microfield strength $F$) by $\Delta\omega_{p,s} = \mu_p + s\omega$ (see (7)), and the quasi-energy $\mu_p$ depends on the amplitude $E_0$. Note that $\mu_p$ also depend on the ion microfield strength $F$. Since different F IX ions in a plasma experience different fields $F$, the above spectral components broaden in accordance with the plasma ion microfield distribution function. When the frequency of the oscillating field $\omega$ exceeds the fine splitting of the $n = 2$ level for F IX and the characteristic FWHM of the Stark $\mathrm{Ly}_\alpha$ line profile for F IX in the plasma ion microfields, some of these spectral components can be roughly treated as Blochinzew's satellites [18] (see also [19]). As an example, note that profile *9* in Fig. 3b exhibits the $\pm\lambda_0^2\omega/2\pi c$ satellites (denoted by $S_{+1}$ and $S_{-1}$), while profile *5* exhibits the $\pm\lambda_0^2\omega/2\pi c$ and $\pm\lambda_0^2\omega/\pi c$ satellites (denoted by $S_{+1}$, $S_{-1}$

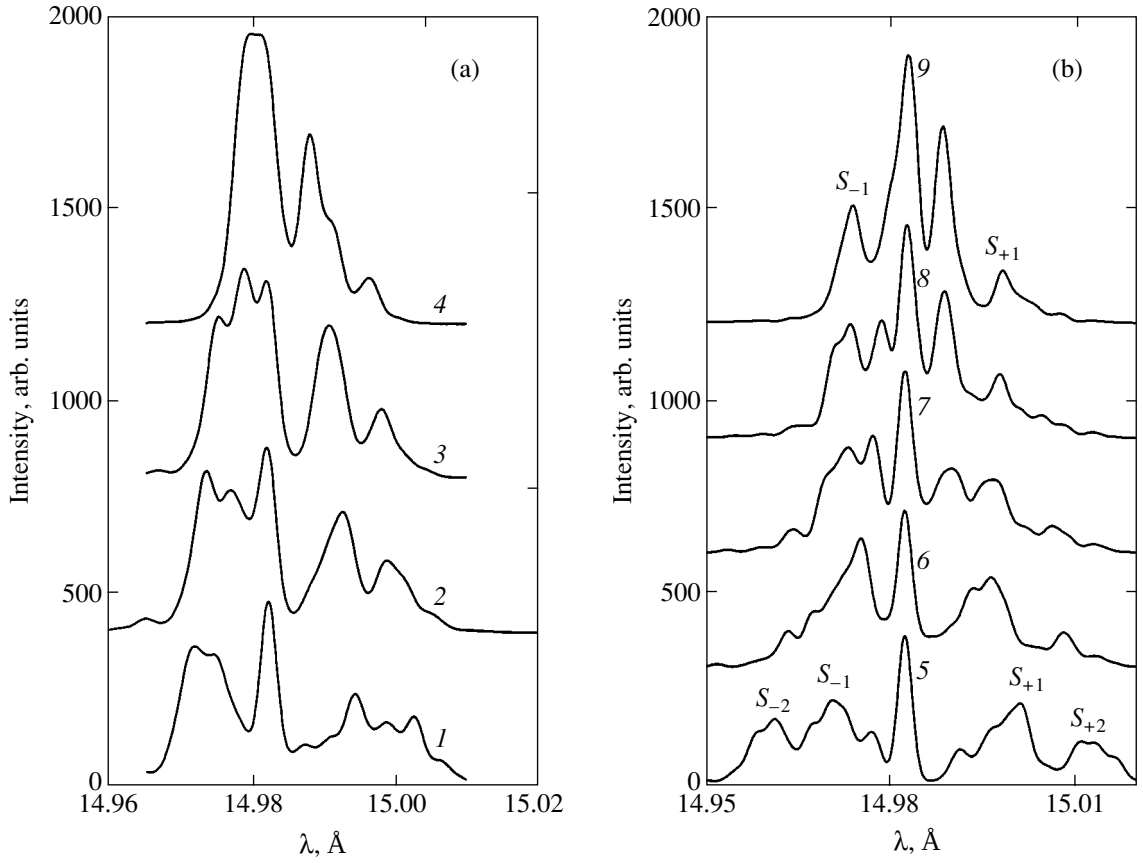



**Fig. 3.** Computed $Ly_\alpha$ line profiles for F IX in a plasma with (a) $T_i = 100$ eV, $N_e = 10^{20}$ cm$^{-3}$, $\omega = 5 \times 10^{14}$ s$^{-1}$ and (b) $T_i = 100$ eV, $N_e = 2 \times 10^{20}$ cm$^{-3}$, $\omega = 10^{15}$ s$^{-1}$ for various oscillating electric field strengths: $E_0 = 6 \times 10^8$ (*1*, *7*), $5 \times 10^8$ (*2*, *8*), $4 \times 10^8$ (*3*, *9*), $2.5 \times 10^8$ (*4*), $10^9$ (*5*), and $7 \times 10^8$ V cm$^{-1}$ (*6*). For clarity, the computed profiles were displaced along the vertical axis.

and $S_{+2}$, $S_{-2}$, respectively). Here, $\lambda_0$ is the wavelength of the $Ly_\alpha$ transition for F IX ions.

Figure 4 shows how the spectrum depends on the frequency $\omega$ at fixed field amplitudes $E_0$. We see that a change in $\omega$ affects both the positions of the spectral components and their intensities.

It is clearly from the above figures that the emitted spectrum has a complex structure that depends nontrivially on both the amplitude and frequency of the oscillating field, and the positions of the local maxima and minima emerging in the $Ly_\alpha$ line profile are determined by the pair of parameters $E_0$ and $\omega$, not by one of them.

Figure 5 shows how the spectrum depends on the ion temperature. Here, the situation is simpler: an increase in the plasma temperature causes the spectral components to broaden, and the line profile is smoothed; i.e., the maxima and minima emerging in the profile become less distinct, although their positions remain unchanged.

The change in plasma density (see Fig. 6) from $N_e = 10^{19}$ cm$^{-3}$ to $N_e = 6 \times 10^{20}$ cm$^{-3}$ has a more complex effect on the emitted spectrum. We see that an increase in $N_e$ results both in an intensity redistribution between the spectral components and in their shift.

Thus, it follows from Figs. 3–6 that, depending on $N_e$, $T$, $E_0$, and $\omega$, the Lyman doublet can have a completely different profile that does not resemble the classical doublet structure even remotely. If the X-ray spectrogram recording the plasma radiation has a sufficiently high spectral resolution, $\lambda/\Delta\lambda > 10^3$, and if electric oscillations of sufficient amplitude exist in the plasma, then such complex profiles must be observed experimentally. We recorded these profiles in the radiation from a plasma heated by intense picosecond pulses at laser flux densities $q_{las} \approx (2\text{–}3) \times 10^{17}$ W cm$^{-2}$.

## 4. COMPARISON WITH CALCULATIONS

We see from Fig. 7 that the observed line profiles can be well described by calculations that take into account the presence of oscillating electric fields. For example, the profile observed at $q_{las} = 2 \times 10^{17}$ W cm$^{-2}$ is satisfactorily reproduced by calculations for $T_i = 100$ eV, $N_e = 10^{20}$ cm$^{-3}$, $\omega = 7 \times 10^{14}$ s$^{-1}$, and $E_0 = 4 \times 10^8$ V cm$^{-1}$ (see Fig. 7a), while the profile observed at

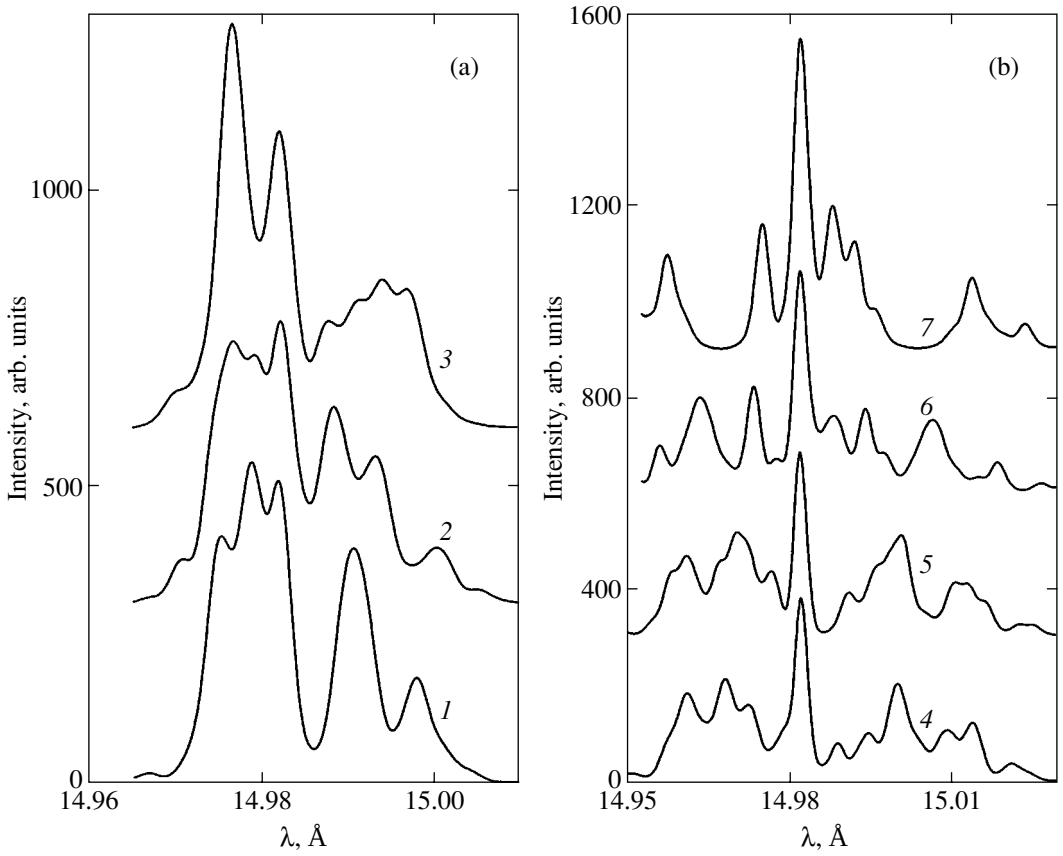



**Fig. 4.** Computed $Ly_\alpha$ line profiles for F IX in a plasma with (a) $T_i = 100$ eV, $N_e = 10^{20}$ cm$^{-3}$, $E_0 = 4 \times 10^8$ V cm$^{-1}$ and (b) $T_i = 100$ eV, $N_e = 2 \times 10^{20}$ cm$^{-3}$, $E_0 = 10^9$ V cm$^{-1}$ for various frequencies of the oscillating electric field: $\omega = 5 \times 10^{14}$ (*1*), $7 \times 10^{14}$ (*2*), $10^{15}$ (*3*, *5*), $8.88 \times 10^{14}$ (*4*), $1.78 \times 10^{15}$ (*6*), and $3.36 \times 10^{15}$ s$^{-1}$ (*7*). For clarity, the computed profiles were displaced along the vertical axis.

$q_{las} = 3 \times 10^{17}$ W cm$^{-2}$ is satisfactorily reproduced by calculations for $T_i = 100$ eV, $N_e = 2 \times 10^{20}$ cm$^{-3}$, $\omega = 10^{15}$ s$^{-1}$, and $E_0 = 6 \times 10^8$ V cm$^{-1}$ (see Fig. 7b). Note that the experimental line profiles slightly differ from the computed ones in the short-wavelength wings. This is because fast ions are present in the plasma (as we showed previously [20], the fraction of the fast F IX ions can reach 20% under our experimental condi-

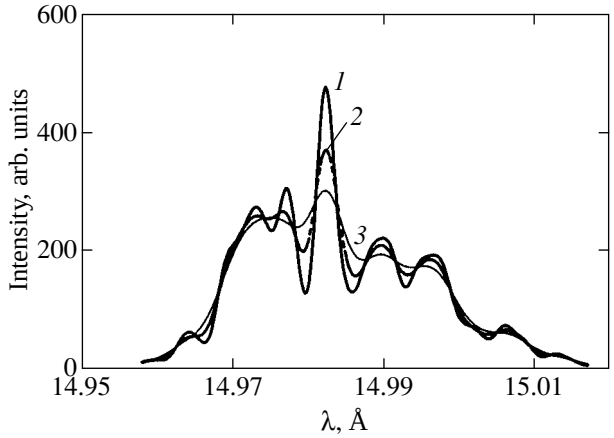



**Fig. 5.** Computed $Ly_\alpha$ line profiles for F IX in a plasma with $N_e = 2 \times 10^{20}$ cm$^{-3}$, $\omega = 10^{15}$ s$^{-1}$, $E_0 = 6 \times 10^8$ V cm$^{-1}$, and $T_i = 100$ (*1*), 200 (*2*), and 400 eV (*3*).

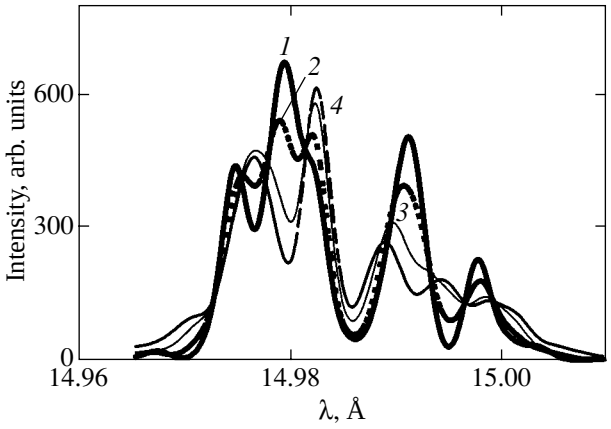



**Fig. 6.** Computed $Ly_\alpha$ line profiles for F IX in a plasma with $T_i = 100$ eV, $\omega = 5 \times 10^{14}$ s$^{-1}$, $E_0 = 4 \times 10^8$ V cm$^{-1}$, and $N_e = 10^{19}$ (*1*), $10^{20}$ (*2*), $3 \times 10^{20}$ (*3*), and $6 \times 10^{20}$ cm$^{-3}$ (*4*).

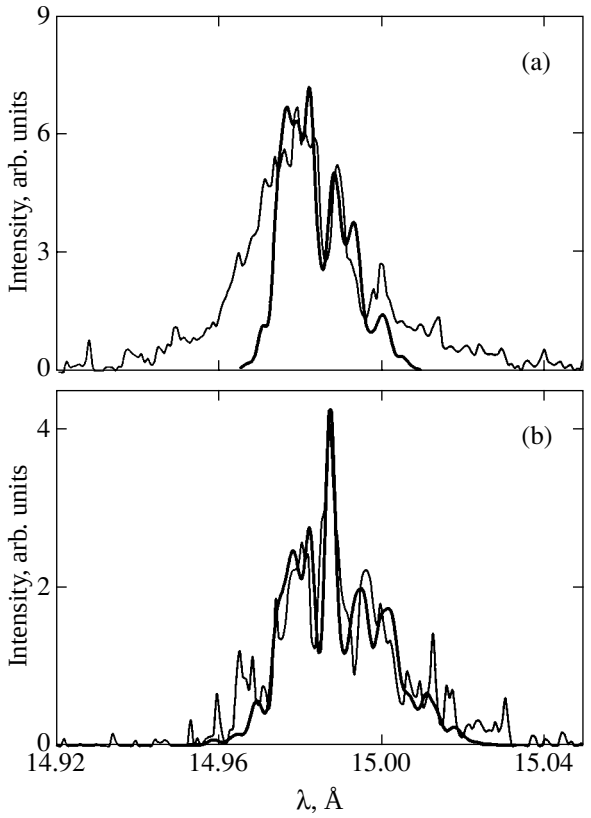



**Fig. 7.** Comparison of the experimental (thin curves) and theoretical (thick curves) $Ly_\alpha$ line profiles for F IX: (a) the experiment at $q_{las} = 2 \times 10^{17}$ W cm$^{-2}$, and the calculation for $T_i = 100$ eV, $N_e = 10^{20}$ cm$^{-3}$, $\omega = 7 \times 10^{14}$ s$^{-1}$, $E_0 = 4 \times 10^8$ V cm$^{-1}$; (b) the experiment at $q_{las} = 3 \times 10^{17}$ W cm$^{-2}$, and the calculation for $T_i = 100$ eV, $N_e = 2 \times 10^{20}$ cm$^{-3}$, $\omega = 10^{15}$ s$^{-1}$, $E_0 = 6 \times 10^8$ V cm$^{-1}$.

tions); their motion toward the spectrograph must cause the short-wavelength wings of the spectral lines to increase. We disregarded the presence of fast ions in our theoretical model. That is why the theoretical and experimental profiles in Fig. 7 differ.

Note also that we recorded the spectra without a time resolution; i.e., the experimental spectra are an integral of the plasma emission over its entire lifetime. Since the density and temperature of the plasma during its evolution change and since the parameters of the oscillating field can also change, describing the observed spectrum by only one set of parameters $N_e$, $T$, $\omega$, and $E_0$ is a rough approximation. However, this approximation yields excellent qualitative (and event quantitative) agreement with experimental data.

Thus, the assumption about the presence of an oscillating electric field in a plasma allows us to explain observed X-ray spectral results that cannot be interpreted in any other way. We consider the possible generation mechanisms for such fields in the next section.

## 5. POSSIBLE GENERATION MECHANISMS FOR AN OSCILLATING ELECTRIC FIELD IN A PICOSECOND LASER PLASMA

The above calculations were essentially based on phenomenological ideas of the existence of intense electrostatic oscillations in a plasma with a given frequency and amplitude. Here, we briefly analyze the possible generation mechanisms for such oscillations in a plasma produced by picosecond laser pulses.

A characteristic feature of the laser plasma is anisotropy in the electron velocity distribution that is attributable to the formation of a strong electric current and that gives rise to a number of plasma instabilities. Ion–acoustic oscillations are the best-known type of instability. However, these are low-frequency oscillations whose effect on an atom is quasi-static and can cause only an additional broadening of the profiles of spectral lines without characteristic satellite structures. High-frequency electron Langmuir oscillations with the electron plasma frequency may be considered as a candidate for the generation of an oscillating electric field that affects the atomic spectra. However, the dispersion law for Langmuir oscillations results in a wide smearing of the oscillation spectrum due to intense thermal electron motion, which also affects the spectral smearing of the atomic satellites attributable to such oscillations.

The mechanisms related to the generation of strong magnetic fields when the solid target surface is exposed to a laser pulse are preferred among other possible candidates for the generation mechanism of an oscillating electric field in a plasma. The recorded magnetic field strengths under these conditions reach $(3–5) \times 10^8$ G at a radiation flux density of $10^{19}$ W cm$^{-2}$ [1].

The relationship of these strengths to the laser radiation intensity is well described by the formula [21]

$$B\ [\mathrm{G}] \approx 10^{-1} J^{1/2},\ [\mathrm{W\ cm^{-2}}]^{1/2}, \qquad (8)$$

which can be derived by setting the energy density of the laser wave equal to the energy density of the generated magnetic field.

Using this similarity law under our experimental conditions, we estimated the magnetic field strength to be $B = 5 \times 10^7$ G. An electron cyclotron frequency of about $8 \times 10^{14}$ s$^{-1}$, which is in good agreement with the frequencies of the electrostatic oscillations used above to interpret the experimental data, corresponds to this field strength. Thus, it seems natural to associate the oscillating electromagnetic fields affecting an atom with the plasma oscillations in a magnetic field at the electron cyclotron frequency.

The Bernstein modes [22] propagating across the magnetic field are the most suitable candidate for such oscillations. This is because they, first, have frequencies close to the electron cyclotron ones, second, possess a dispersion law with (in contrast to, e.g., Langmuir oscillations) an isolated discrete spectrum (a kind of

Landau levels), and, finally, they correspond to the oscillations in a collisionless plasma, which most closely matches the conditions of short picosecond laser pulses.

The generation mechanisms for Bernstein oscillations have been extensively studied in the literature and are associated with the growth of electron cyclotron drift instability [23, 24]. The drift of electrons relative to ions attributable to the flow of a strong electric current across the magnetic field underlies the growth of this instability. In the laser plasma conditions under consideration, a quasi-static magnetic field is generated due to the growth of Weibel instability [25] that is related to the anisotropy in the electron velocity distribution function caused by the current of ionization of the medium by a laser pulse. The direction of this constant magnetic field coincides with the direction of the laser wave magnetic field; hence, it is perpendicular to the direction of the current. The emerging current is probably attributable to the reverse currents that appear when a beam of fast electrons is generated by the main laser pulse. These currents are responsible for the laser plasma pinching observed in recent experiments [26] that arises when thin wire targets are exposed to picosecond laser pulses. The growth rate of the electron cyclotron drift instability is described by the formula [23]

$$\frac{\gamma}{\omega_B} = \frac{\cos\theta}{\pi^{1/2}}\frac{v_d}{v_{e0}}\frac{T_e}{2T_i}\frac{1.5}{[1+(k\lambda_D)^2]^2}\frac{\nu_{ee}}{\omega_B}, \quad (9)$$

where $\omega_B$ is the electron cyclotron frequency, $\theta$ is the angle between the wave vector of the oscillations $k$ and the velocity vector of the current drift $v_d$, $v_{e0}$ is the thermal electron velocity, $T_e$ and $T_i$ are the electron and ion temperatures, $\lambda_D$ is the electron Debye length, and $\nu_{ee}$ is the electron–electron collision frequency.

The second term in (9) corresponds to the stabilization of the cyclotron instability by electron–electron collisions, and its contribution is generally negligible under laser plasma conditions. Indeed, the collision frequency estimated using the formula

$$\nu_{ee} \approx 10^{10}(N_e\ [\mathrm{cm}^{-3}]/10^{16})(10/T_e\ [\mathrm{eV}])^{3/2}\ \mathrm{s}^{-1}$$

yields $\nu_{ee} \approx 10^{11}$ s$^{-1}$ for $N_e = 10^{20}$ cm$^{-3}$ and $T_e = 1$ keV, which is more than three orders of magnitude lower than the electron cyclotron frequency.

The growth of this instability is effective at a high electron drift velocity $v_d$ determined by the current. The condition that determines this velocity is [23]

$$\frac{v_d}{v_{e0}} \geq \frac{\tilde{n}B}{(2\times 10^{-5}N_e)^{1/2}}, \quad (10)$$

where the magnetic field $B$ is in gausses, and the electron density $N_e$ is in cm$^{-3}$.

Substituting the typical magnetic field strength and electron density used above, we obtain a value of the right-hand side in formula (10) that is on the order of unity for the fundamental $\tilde{n} = 1$ mode. The corresponding growth rates estimated using formula (9) are $\gamma \approx 10^{13}$–$10^{14}$ s$^{-1}$ for these $v_d/v_{e0}$ ratios, whose reciprocal is smaller than the laser pulse duration. Thus, the conditions for intense generation of Bernstein oscillations due to the growth of electron cyclotron drift instabilities are realized in the laser plasma.

The intensity of the Bernstein plasma oscillations can be roughly estimated on the grounds that the effective electron–wave collision frequency $\nu_{E0}$ that limits the field strength $E_0$ must be on the order of the wave growth rate. In this case, the effective electron–wave collision frequency $\nu_{E0}$ in a thermally equilibrium plasma differs, according to [27], from the electron–electron collision frequency $\nu_{ee}$ only by a logarithmic factor of ~10. The wave energy density at thermodynamic equilibrium, $E_{cn}^2/8\pi$ (where $E_{cn}$ is the wave amplitude), is lower than the thermal energy density $N_eT_e$ by a factor of $N_e\rho_D \gg 1$. Setting the effective collision frequency equal to the growth rate of the oscillations under consideration using formula (9), we obtain an estimate for the amplitude of the generated waves:

$$(E_0^2/8\pi N_eT_e)(N_e\rho_D^3)\nu_{ee}/10 = \gamma. \quad (11)$$

Substituting the plasma parameters used above into (11) yields an estimate of $E_0 \approx (4\text{–}6)\times 10^8$ V cm$^{-1}$, in reasonable agreement with the strengths of the oscillating fields adopted above.

## 6. CONCLUSIONS

Our measurements have revealed intense plasma oscillations in a laser plasma produced by picosecond laser pulses. The separation between the observed satellites, which correlates with the estimates of the magnetic field emerging in a plasma, allows them to be associated with the growth of Bernstein-mode-type plasma oscillations. If this correlation will receive further experimental confirmation, then its presence will make it possible to directly measure the magnetic fields generated in a laser plasma.

Actually, the observed spectral pattern proves to be more complex than the computed one. This is probably attributable, first, to the presence of low-frequency ion–acoustic oscillations and, second, to the presence of additional satellite lines related to the possible sublevel intersection of the fine structure of the $n = 2$ level for F IX in a strong magnetic field (see [28] for more details). In the latter case, the influence of the variable electric field of plasma oscillations near these points of intersection gives rise to Blochinzew's standard structure of equidistant satellites [18] whose intensity depends on the ratio of the field strength and frequency. The observed spectrum probably corresponds to a superposition of these types of spectra. Unfortunately, the distribution of the magnetic field emerging from the

growth of Weibel instability is not known. Further experiments on observing satellites in the spectra of multicharged ions will allow one to determine the type of distribution of these fields and, thus, the peculiarities of the growth of Weibel instability as well as another type of oscillations in a plasma.

## ACKNOWLEDGMENTS

This work was supported by the International Scientific and Technical Center (project no. 2155) and INTAS (grant no. 01-0233). The work of V.P.G. was also supported in part by the Russian Foundation for Basic Research (project no. 03-02-17282). We are grateful to V.P. Krainov for valuable discussions

## REFERENCES

1. M. Tatarakis, A. Goral, I. Watts, *et al.*, Phys. Plasmas **9**, 2244 (2002).
2. I. Yu. Skobelev, A. Ya. Faenov, A. I. Magunov, *et al.*, in *Proceedings of 7th International Conference on Multiphoton Processes* (Garmish-Partenkirchen, Germany, 1996); Inst. Phys. Conf. Proc., Ed. by P. Lambropoulos and H. Walther, Vol. 154.
3. S. A. Pikuz, A. Maksimchuk, D. Umstadter, *et al.*, Pis'ma Zh. Éksp. Teor. Fiz. **66**, 454 (1997) [JETP Lett. **66**, 480 (1997)].
4. R. C. Elton, H. R. Griem, B. L. Welch, *et al.*, J. Quant. Spectrosc. Radiat. Transf. **58**, 559 (1997).
5. A. L. Osterheld, B. K. F. Young, J. Dunn, *et al.*, J. Quant. Spectrosc. Radiat. Transf. **58**, 827 (1997).
6. I. Yu. Skobelev, A. Ya. Faenov, A. I. Magunov, *et al.*, Phys. Scr. T **73**, 104 (1997).
7. O. Renner, O. Peyrusse, P. Sondhauss, and E. Forster, J. Phys. B: At. Mol. Opt. Phys. **33**, L151 (2000).
8. D. Riley and O. Willi, Phys. Rev. Lett. **75**, 4039 (1995).
9. V. S. Belyaev, V. I. Vinogradov, A. S. Kurilov, *et al.*, Pis'ma Zh. Éksp. Teor. Fiz. **78**, 1216 (2003) [JETP Lett. **78**, 703 (2003)].
10. V. S. Belyaev, V. I. Vinogradov, A. S. Kurilov, *et al.*, Kvantovaya Élektron. (Moscow) **30**, 229 (2000).
11. T. A. Pikuz, A. Ya. Faenov, S. A. Pikuz, *et al.*, J. X-ray Sci. Technol. **5**, 323 (1995).
12. V. A. Boĭko, A. V. Vinogradov, A. A. Ilyukhin, *et al.*, Kvantovaya Élektron. (Moscow) **8**, 28 (1981).
13. E. Oks, *Plasma Spectroscopy. The Influence of Microwave and Laser Fields* (Springer, Berlin, 1995).
14. L. A. Bureeva, V. P. Gavrilenko, and V. S. Lisitsa, in *An Encyclopedia of Low-Temperature Plasma*, Ed. by V. E. Fortov (Nauka, Moscow, 2000), Vol. 1, p. 351 [in Russian].
15. V. P. Gavrilenko, V. N. Ochkin, and S. N. Tskhai, Proc. SPIE **4460**, 207 (2002).
16. H. R. Griem, *Spectral Line Broadening by Plasmas* (Academic, New York, 1974; Mir, Moscow, 1978).
17. N. B. Delone and V. P. Kraĭnov, *Atom in a Strong Light Field* (Énergoatomizdat, Moscow, 1984) [in Russian].
18. D. I. Blochinzew, Phys. Z. Sowjetunion **4**, 501 (1933).
19. V. P. Gavrilenko and E. A. Oks, Kvantovaya Élektron. (Moscow) **10**, 1910 (1983).
20. V. S. Belyaev, V. I. Vinogradov, A. S. Kurilov, *et al.*, Zh. Éksp. Teor. Fiz. **125**, 1295 (2004) [JETP **98**, 1133 (2004)].
21. V. S. Belyaev, Quantum Electron. **34**, 41 (2002).
22. *Radiation Processes in Plasmas*, Ed. by G. Bekefi (Wiley, New York, 1966; Mir, Moscow, 1971).
23. D. Forslund, R. Morse, C. Nielson, and J. Fu, Phys. Fluids **15**, 1303 (1972).
24. A. A. Galeev, D. G. Lominadze, A. D. Pataraya, *et al.*, Pis'ma Zh. Éksp. Teor. Fiz. **15**, 417 (1972) [JETP Lett. **15**, 294 (1972)].
25. E. S. Weibel, Phys. Rev. Lett. **2**, 83 (1959).
26. F. N. Beg, E. L. Clark, and M. S. Wei, Phys. Rev. Lett. **92**, 095001 (2004).
27. B. I. Davydov, in *Plasma Physics and Problem of Controlled Thermonuclear Reactions*, Ed. by M. A. Leontovich (Akad. Nauk SSSR, Moscow, 1958) [in Russian].
28. L. A. Bureeva and V. S. Lisitsa, *Perturbed Atom* (IzdAt, Moscow, 2000) [in Russian].

*Translated by V. Astakhov*

*Journal of Experimental and Theoretical Physics, Vol. 99, No. 4, 2004, pp. 719–726.*
*Translated from Zhurnal Éksperimental'noĭ i Teoreticheskoĭ Fiziki, Vol. 126, No. 4, 2004, pp. 833–842.*


**PLASMA, GASES**

# Different Modes of Capacitively Coupled Radio-Frequency Discharge in Methane

**I. V. Schweigert**

*Institute of Theoretical and Applied Mechanics, Siberian Division, Russian Academy of Sciences, Novosibirsk, 630090 Russia*

*e-mail: ischweig@itam.nsc.ru*

Received December 5, 2003

**Abstract**—Two different modes of capacitively coupled radio-frequency (RF) discharge in methane at gas pressures between 0.01 and 1.00 Torr are examined by numerical simulation based on a combined approach. It is shown that transition between volume-dominated and active-sheath modes is caused by variation of discharge current or gas pressure. Hysteretic behavior is revealed as the discharge current density is varied along the growing- and falling-current branches of the current-density curve. A phase diagram representing the domains of different discharge modes is obtained in wide current and pressure ranges. © *2004 MAIK "Nauka/Interperiodica".*

## 1. INTRODUCTION

The existence of different discharge modes was examined in a pioneering experimental study by Levitskii [1]. Starting with that outstanding work, the change between different (high- and low-voltage) discharge modes were attributed to the so-called α–γ transition. The sharp increase in plasma density and decrease in electron temperature were explained by the effect of secondary electrons on the discharge dynamics. Depending on power input, discharge can be sustained either by electrons in the quasi-neutral region, which are heated as the electrode sheath moves (α mode), or by electrons emitted by the electrodes (γ mode). Transition between different modes of RF discharge characterized by high and low voltages in argon and helium was studied experimentally in [2]. It has been demonstrated that electron density is much higher in the γ mode, and the corresponding electron energy distribution function (EEDF) is Maxwellian owing to electron–electron collisions. Both in kinetic simulations [3] and in simulations using a two-fluid model for the electron gas [4], the α–γ transition was examined for an RF discharge in helium. Transition of different type was revealed experimentally for a low-pressure RF discharge in argon [5]. A sharp transition from low to high electron temperature associated with increase in gas pressure was observed. The increase in electron temperature was explained in [5] by a change in the electron-heating mechanism involving the Ramsauer effect. Transition between different modes of discharge in silane was observed experimentally in [6, 7] and analyzed numerically in [8]. The transition from volume glow to electrode-sheath glow observed in [6, 7] was attributed to secondary electrons. It should be noted that a higher rate of α-Si:H film deposition was measured for the volume-dominated discharge mode. Even though capacitively coupled RF discharge plasmas are widely applied in technological processes (e.g., in diamond-like film deposition), the dynamics of transition between different modes of discharge in molecular gases has never been investigated.

## 2. MODEL AND NUMERICAL TECHNIQUE

In this study, transition between different capacitively coupled RF discharge modes is analyzed by using the algorithm developed in [9], which combines fluid-dynamic and kinetic approaches. The purely fluid-dynamic approach is disadvantageous in that the inertia of both electrons and ions is neglected, which restricts the scope of the fluid model to high gas pressures. On the other hand, solution of kinetic equations by the PIC–MCC (Particle-In-Cell Monte Carlo Collision) method requires a very large particle number $N$, particularly at low gas pressures. It is well known that an increase in $N$ reduces electric-field fluctuations, which lead to spurious heating of electrons. Currently, hybrid models combining kinetic and fluid-dynamic approaches are widely used in computations of glow discharges (e.g., see [4, 10, 11]). In the present combined model [9], the kinetic equations for electrons and ions (three-dimensional in velocity space and one-dimensional in space), the continuity equations for the electron and ion densities and fluxes, and the Poisson equation for electric-field strength $E$ are solved self-consistently. Kinetic equations are solved to find the electron energy distribution function and the kinetic

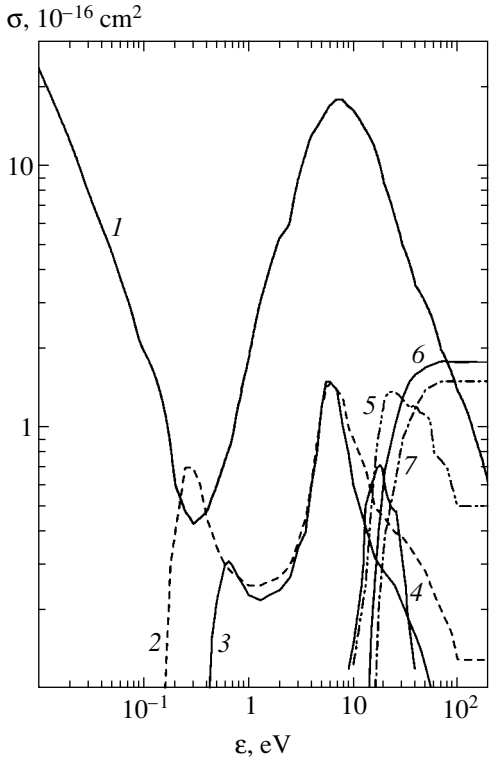



**Fig. 1.** Cross sections for electron scattering by $CH_4$ molecules versus electron energy: elastic scattering (*1*), excitation of vibrational levels (*2*, *3*), $CH_2$ or $CH_3$ radical production by dissociation (*4*, *5*), ionization (*6*), and dissociative ionization (*7*).

coefficients in transport equations. One advantage of the present combined PIC–MCC approach over the conventional PIC–MCC algorithm (see [12]) is the lower time complexity achieved by substantially reducing the number of particles. Even at low gas pressures, we use only $N = 5000$ for each plasma component, whereas the conventional PIC–MCC algorithm requires more than $N = 256000$ particles to ensure agreement with experiment (see [9]).

We computed a one-dimensional capacitively coupled RF discharge with sinusoidal current density $j$ of frequency 13.56 MHz. One electrode was grounded, and the potential at the other was computed self-consistently by assuming that a prescribed current was sustained. The interelectrode spacing $d$ was varied from 3 to 6 cm. Methane was treated as an electropositive gas [13]. To simplify analysis, we follow [14] and consider only $CH_5^+$ ions. We used a model of electron kinetics with electron–neutral [13, 14] and electron–electron collisions computed by the method proposed in [15]. Figure 1 shows cross sections for elastic collisions, excitation of vibrational levels, dissociation, ionization, and dissociative ionization. Since the methane decomposition rate was assumed to be low, only collisions between electrons and $CH_4$ molecules are taken into consideration. Plasma chemistry is not discussed in this paper.

## 3. STRUCTURE AND CHARACTERISTICS OF DIFFERENT RADIO-FREQUENCY DISCHARGE

Our discharge simulations revealed two different modes of capacitively coupled RF discharge. Figure 2 shows typical electron density distributions $n_e$ and electron energies $\varepsilon$ averaged over the RF discharge period. The volume-dominated discharge mode (VD mode) is characterized by low plasma density and high electron energy (solid curves in Fig. 2). The distribution of mean electron energy across the discharge gap is flat. This mode occurs at lower gas pressures and current densities. The active-sheath mode (AS mode) is associated with a much higher plasma density (dashed curves in Fig. 2). The distribution of mean electron energy has peaks in the sheaths and a deep minimum at the center of the discharge gap. Transition between the modes occurs under a certain critical conditions. In computations, variation of either current density or gas pressure causes gradual change in discharge characteristics until a critical point is reached, and then the discharge structure changes abruptly. Figure 2 demonstrates how a slight change in current density by 0.1 mA/cm$^2$ makes the system switch from VD to AS mode. The plasma density increases by a factor of 4 to 5, while the mean electron energy drops at the center of the discharge gap. Figure 3 shows the variation of mean electron energy at the center of the gap with discharge current density for several values of gas pressure. The drop in $\varepsilon$ spans a narrow current-density interval, and the lowest current density corresponds to $P = 0.03$ Torr.

It should be noted that the transition involves a drop in discharge power and a sharp increase in the total degree of ionization.

## 4. ROLE OF SECONDARY ELECTRONS IN DISCHARGE DYNAMICS

In previous studies, the change between the modes was attributed to the $\alpha$–$\gamma$ transition. For this reason, we performed an analysis of the effect of secondary electrons on discharge dynamics. The calculated secondary ion–electron emission coefficient $\gamma$ was varied between 0 and 0.5, while the initial secondary-electron temperature $T_e$ was set equal to 1 eV. The results of computations performed for several values of $\gamma$ revealed a weak effect of secondary electrons on discharge characteristics. Even at the highest value, $\gamma = 0.5$, the critical current density $j^*$ was only 25% lower than $j^*$ at $\gamma = 0$. Only the $\alpha$ mode was obtained at $P = 0.01$–$1.00$ Torr and $j = 0.45$–$2.2$ mA/cm$^2$. Therefore, the $\alpha$–$\gamma$ transition is not responsible for the change between different modes of capacitively coupled RF discharge in methane, in contrast to inert gases. The key difference in electron kinetics between molecular and inert gases is

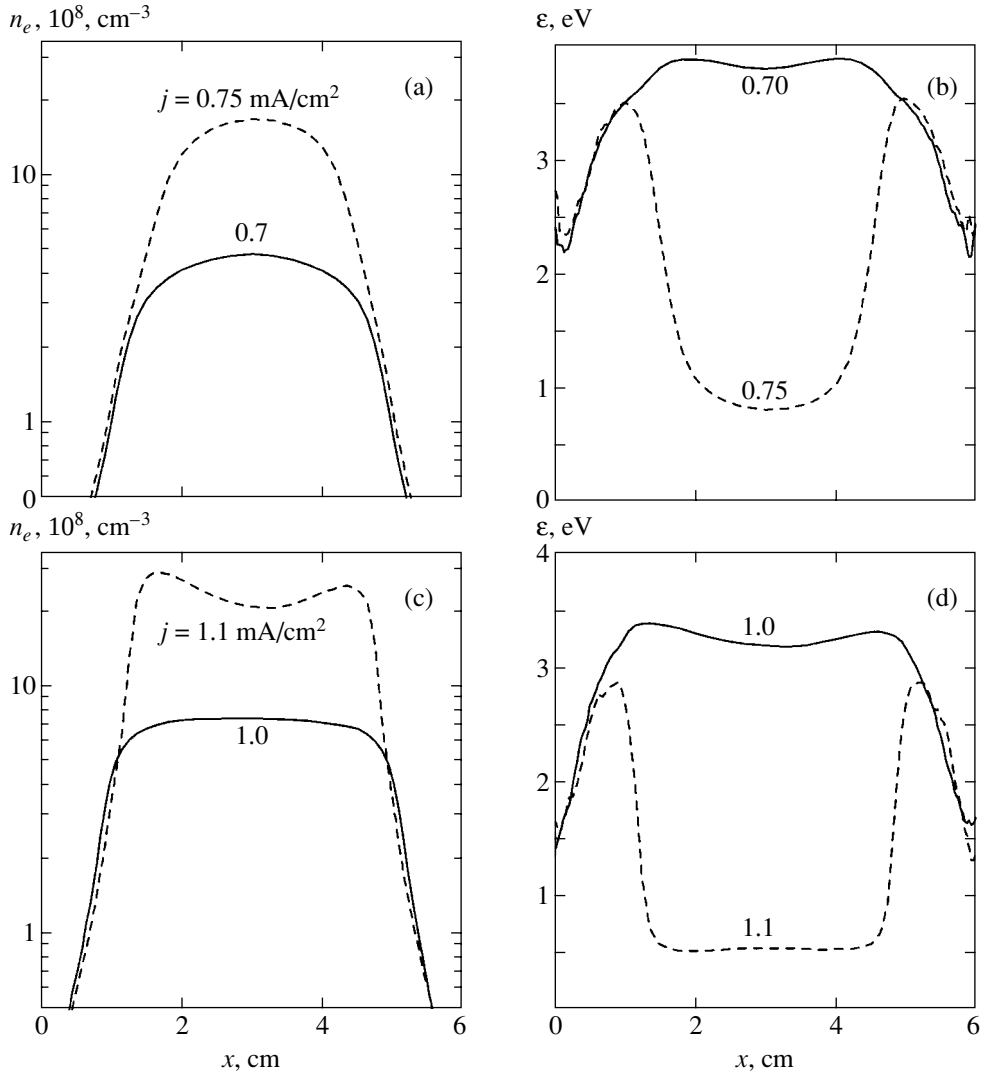



**Fig. 2.** Distributions of electron density $n_e$ and electron energy ε across the discharge gap averaged over the discharge period for $j$ = 0.70 and 0.75 mA/cm$^2$ at $P$ = 0.03 Torr (a, b) and for $j$ = 1.0 and 1.1 mA/cm$^2$ at $P$ = 0.075 Torr (c, d).

that the electron-energy relaxation length is much smaller in methane, as compared to inert gases, because vibrational levels with very low threshold energies (0.162 and 0.361 eV) are excited. Accordingly, the secondary electrons emitted by electrodes cannot give rise to a high-energy beam, as they do in the γ mode of discharge in an inert gas.

## 5. SCENARIO FOR TRANSITION BETWEEN DIFFERENT MODES

The electric-field strength in the electrode sheath increases with current density. At the critical point, the mean free path for ionization $l_{ion}$ becomes smaller than the sheath thickness $l_{sh}$, and the degree of ionization in the sheath sharply increases. In other words, the electron energy exceeds the ionization threshold energy inside the electrode sheath, which causes transition from VD to AS mode. To verify this hypothesis, we varied the ionization cross section $\sigma_i$ by δ so that the critical current density increased for $\delta > 0$ and decreased for $\delta < 0$. For example, $j^* = 0.725$ mA/cm$^2$ for $\delta = 0$, and the critical current density decreased to 0.70 mA/cm$^2$ as ionization threshold energy was reduced by using $\delta = -0.5$ eV.

Figure 4 shows the electron energy distribution functions in the electrode sheath and at the center of the discharge gap corresponding to $j$ = 1 mA/cm$^2$ (subcritical current density) and $j$ = 1.1 mA/cm$^2$ (supercritical current density). In the volume-dominated discharge mode, a low ionization rate in the sheaths is compensated for by ionization in the quasi-neutral region. The EEDFs shown in Fig. 4a demonstrate that the concentration of ionizing electrons in the VD mode is low both in the sheaths and at the center. In contrast, the EEDF

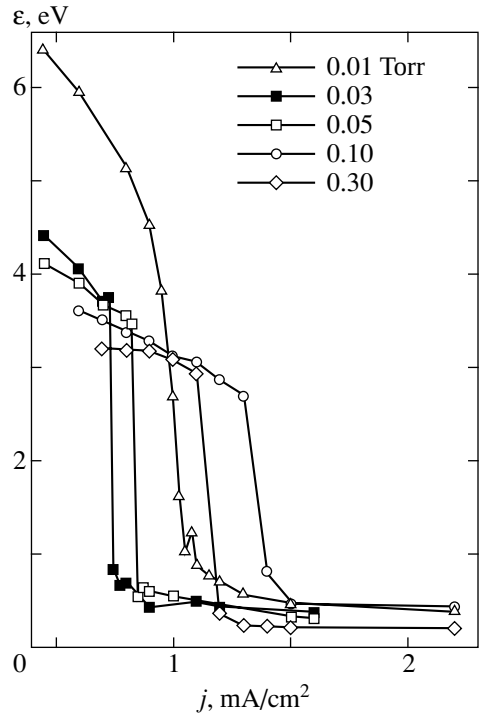



**Fig. 3.** Mean electron energy at the center of the discharge gap versus current density for several values of gas pressure.

corresponding to the active-sheath mode has a high-energy tail in the interval above the ionization threshold (12.6 or 14.3 eV) (curve *1* in Fig. 4b). These hot electrons can ensure a sufficiently high ionization rate in an electrode sheath to sustain a discharge with prescribed current density. At the center of the discharge gap, electrons have low energies ($\varepsilon < 0.5$ eV), compared to sheath electrons, and play a passive role (curve *2* in Fig. 4b).

The processes responsible for the EEDF shape in different modes are illustrated as follows. Figure 5 shows the rates of electron heating by electric field and electron-energy dissipation due to various inelastic processes in the VD and AS modes for $j = 1.0$ and 1.1 mA/cm$^2$ at $P = 0.075$ Torr. In the VD mode, electrons are heated in the quasi-neutral region (curve *1* in Fig. 5a), whereas electron energy is gained only in the electrode sheaths in the AS mode (curve *2* in Fig. 5a). Figure 5b demonstrates that the greater part of electron energy is transferred in the VD mode to excite vibrational levels in $CH_4$ molecules (curve *1*). However, note that a smaller interelectrode spacing width corresponds to a lower energy transfer to vibrational degrees of freedom and a higher power consumption due to dissociative processes. Note also that, even though the ionization rate is lower at $j = 1$ mA/cm$^2$ (solid curve in Fig. 5c) as compared to 1.1 mA/cm$^2$, a higher degree of methane dissociation is obtained for the VD mode of discharge at $j = 1$ mA/cm$^2$ (dashed curve in Fig. 5d). A similar trend was observed experimentally in [6] for discharge in silane. The electron kinetics are different in that electrons gain some additional energy in the central discharge region in the VD mode, whereas the mean

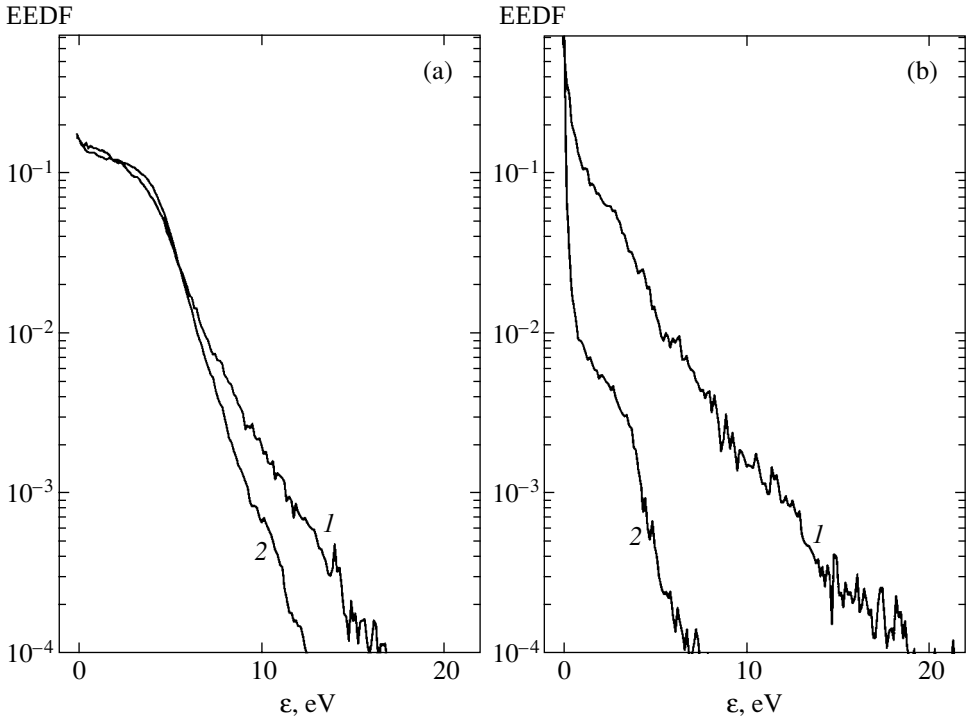



**Fig. 4.** Electron energy distribution function for $P = 0.075$ Torr and $d = 6$ cm: (a) $j = 1.1$ mA/cm$^2$, $x = 0.86$ (*1*) and 3.0 cm (*2*); (b) $j = 1.0$ mA/cm$^2$, $x = 1.36$ (*1*) and 3.0 cm (*2*).

electron energy is as low as tenths of an electronvolt in the AS mode because the electric field in the quasi-neutral region is weak. In the latter mode, thermalized electrons are trapped by a potential well and plasma density increases, as predicted in [16]. The heating of electrons in the AS mode is illustrated by Fig. 6, where the distributions of electric field strength $E$ and electron concentration are shown at several instants in the cathode part of the discharge period. It is clear that an electron-density wave approaches the electrode as the field strength decreases. The phase shift between the electron wave and $E$ is responsible for electron heating. Figure 6b shows the distribution of $E$ on a coarser scale. The electric field strength changes sign across the plasma–sheath interface, giving rise to a potential well analogous to negative glow in DC glow discharges.

## 6. PHASE DIAGRAMS FOR VOLUME-DOMINATED AND ACTIVE-SHEATH MODES

Determination of the critical parameters of the transition between the two RF discharge modes requires considerable computational resources. To find conditions for existence of the volume-dominated and active-sheath modes, we computed the dynamics of capacitively coupled RF discharge in methane over wide ranges of current density and gas pressure. In Fig. 7, the numerical results obtained for the growing-current branch are summarized in a phase diagram including VD- and AS-mode regions in the $j$–$P$ plane for $d$ = 3, 4, and 6 cm.

In the computations performed for the growing-current branch, we started from a low current density cor-

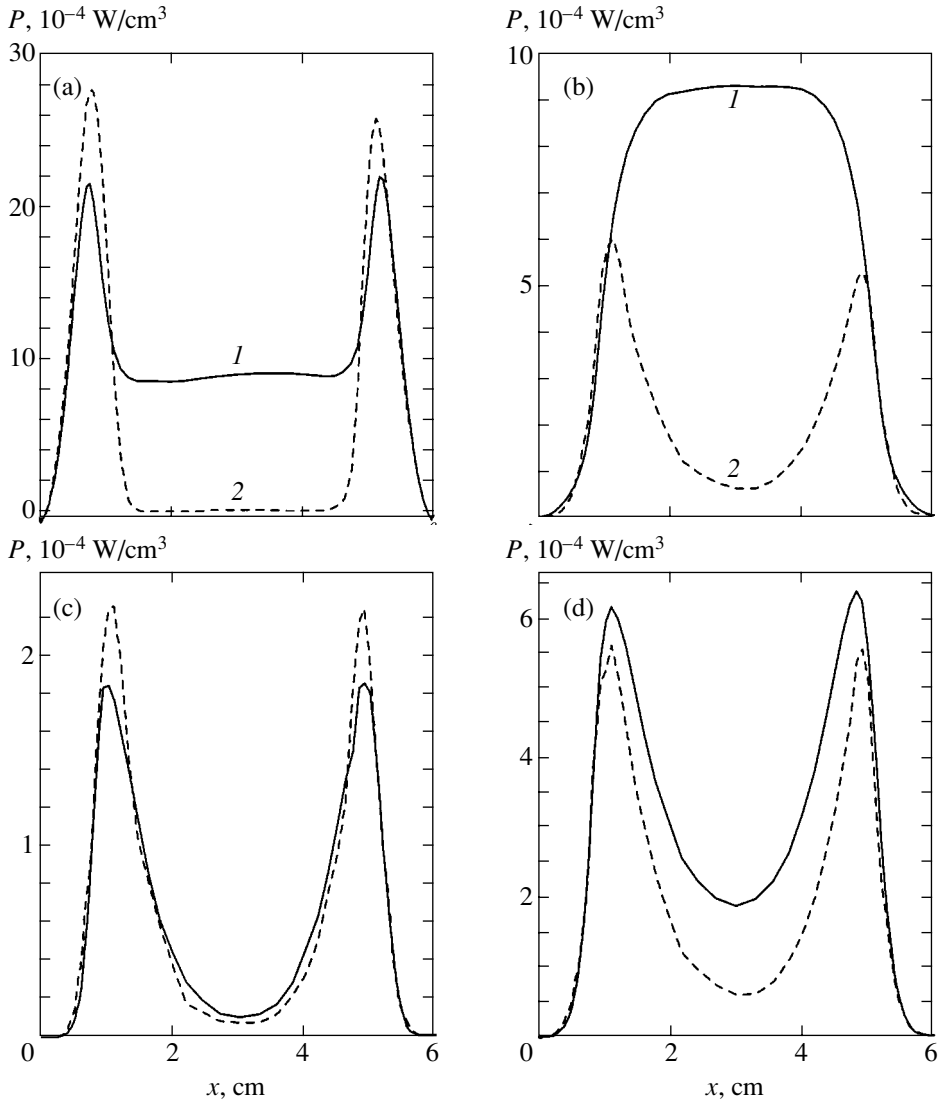



**Fig. 5.** Distributions of (a) time-averaged total rate of electron heating and power transferred to (b) excited vibrational levels, (c) ionized states, and (d) $CH_2$ or $CH_3$ radicals produced by dissociation for $d$ = 6 cm at $P$ = 0.075 Torr: $j$ = 1.0 mA/cm$^2$ (solid curves); $j$ = 1.1 mA/cm$^2$ (dashed curves).

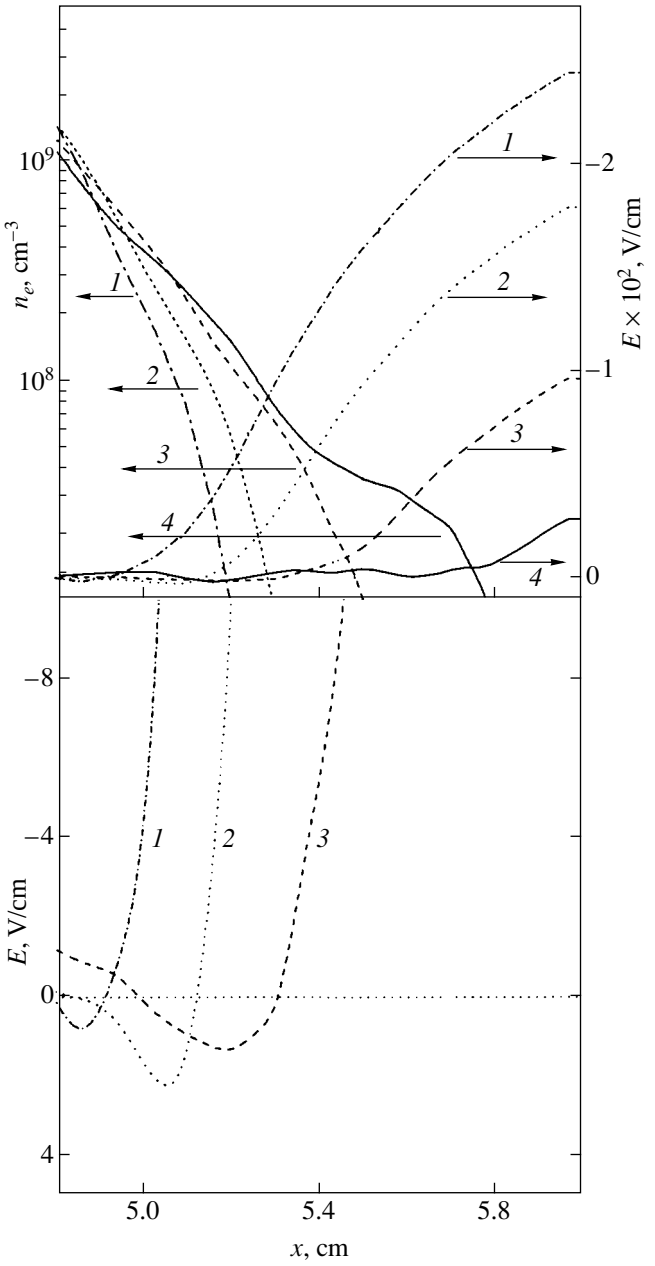



**Fig. 6.** Distributions of (a) $E$ and $n_e$ near an electrode and (b) $E$ on a coarser scale at $t = 0.2T$ (*1*), $0.3T$ (*2*), $0.4T$ (*3*), and $0.5T$ (*4*), where $T$ is discharge period, at $P = 0.075$ Torr for $j = 1.1$ mA/cm$^2$ and $d = 6$ cm.

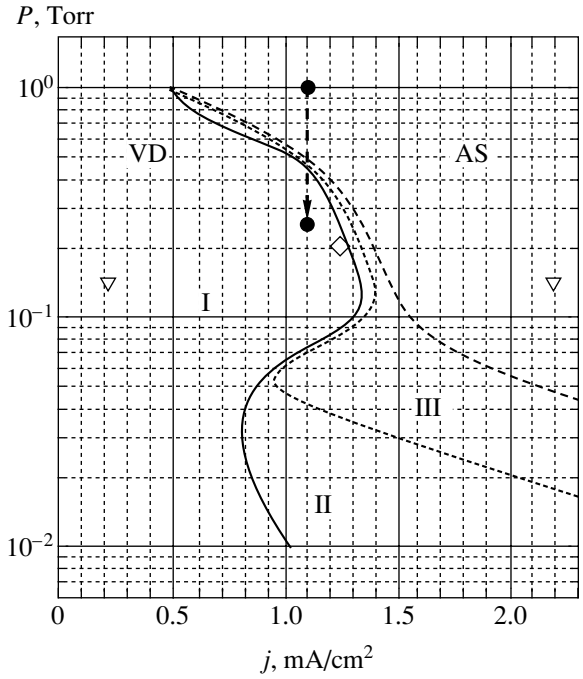



**Fig. 7.** Phase diagram for capacitively coupled RF discharge in methane. Solid, dotted, and dashed curves separate the VD- and AS-mode regions for $d = 6$, 4, and 3 cm, respectively. Symbols ▽, ◇, and ● represent data borrowed from [15], [12], and [16], respectively.

responding to the volume-dominated discharge mode. With increasing discharge current density, the discharge switched into the active-sheath mode at a critical point. Domain I in Fig. 7 corresponds to the VD mode of discharge with $d = 6$ cm. The solid curve separates the VD- and AS-mode regions. Note that the critical current density substantially varies with gas pressure. As the interelectrode spacing is reduced, the phase-diagram region representing the volume-dominated mode tends to expand. The VD-mode region corresponding to $d = 4$ cm consists of domains I and II. When the discharge gap is reduced to $d = 3$ cm, domain III joins the region. When $P < 0.05$ Torr and $j < 2.2$ mA/cm$^2$, the discharge is sustained in the volume-dominated mode. Figure 7 demonstrates that previous results concerning capacitively coupled RF discharge in methane are in good agreement with the present phase diagram. In [17], VD and AS modes (with $j = 0.2$ and 2.2 mA/cm$^2$, respectively) were obtained numerically for $P = 0.14$ Torr and $d = 3$ cm (triangles in Fig. 7). The discharge structure computed in [18] for $P = 0.1$–$0.35$ Torr, $j = 3$ mA/cm$^2$, and $d = 3.5$–$6.5$ cm corresponds to the active-sheath mode. The volume-dominated mode was also revealed in [14] for a capacitively coupled RF discharge at a voltage drop $U = 275$ V, $P = 0.2$ Torr, and $d = 3.5$ cm (diamond in Fig. 7). In [19], an RF discharge in methane was simulated by using a two-dimensional fluid model, and transition from a corner-dominated mode to a volume-dominated mode was observed at $U = 100$ V. The transition resulted in a drop in plasma density and an increase in electron energy in the entire discharge gap. The transition revealed in [19] is in excellent agreement with the present phase diagram (circles in Fig. 7). Note that a relatively high, smoothly distributed electron temperature at $P = 0.25$ Torr was obtained by using the fluid model employed in [19], whereas our kinetic simulations predict a pronounced minimum in electron energy in the central discharge region.

## 7. HYSTERESIS

The results presented above were obtained for the growing-current branch by approaching a preset current-density value "from below." An attempt to approach the same current-density value "from above" produced an unexpected result: we obtained a new stable solution. Figures 8a and 8b show the values of $n_e$ and ε, respectively, at the center of the discharge gap corresponding to the growing- and falling-current branches of the current-density curve. For $d = 6$ cm and $P = 0.075$ Torr, the VD mode is observed as the current density is increased up to $j^* = 1$ mA/cm$^2$ along the growing-current branch. As the discharge switches into the active-sheath mode, ε decreases and $n_e$ substantially increases. As the current density is decreased along the falling-current branch, the AS mode persists until a very low value of $j$ is reached. Hysteresis of this kind has been observed experimentally for discharge in silane [6]. For $d = 3$ cm and $P = 0.03$ Torr, discharge characteristics varied without hysteresis, because only the VD mode can be observed according to the phase diagram.

To elucidate the nature of the hysteretic behavior, we computed ionization dynamics for the growing- and falling-current branches. Figures 9a and 9b show total ionization as a function of the number of discharge periods obtained for $j = 0.70$ and 0.75 mA/cm$^2$ by increasing and decreasing the current density, respectively. Figure 9a demonstrates that $N_i$ is increased by a factor of about 5 by changing from $j = 0.70$ to $j = 0.75$ mA/cm$^2$. It is obvious that the transition is initiated by a rapidly increasing ionization rate in the electrode sheaths, because a much higher ionization rate is required to sustain the AS mode as compared to the VD mode. The change in total ionization $N_i$ due to the decrease in current density from 0.75 to 0.70 mA/cm$^2$ is insignificant (see Fig. 9b).

The hysteretic behavior is explained by the existence of two stable solutions (different modes) in a certain interval of current density. The actual mode is determined by the system's history. When the discharge

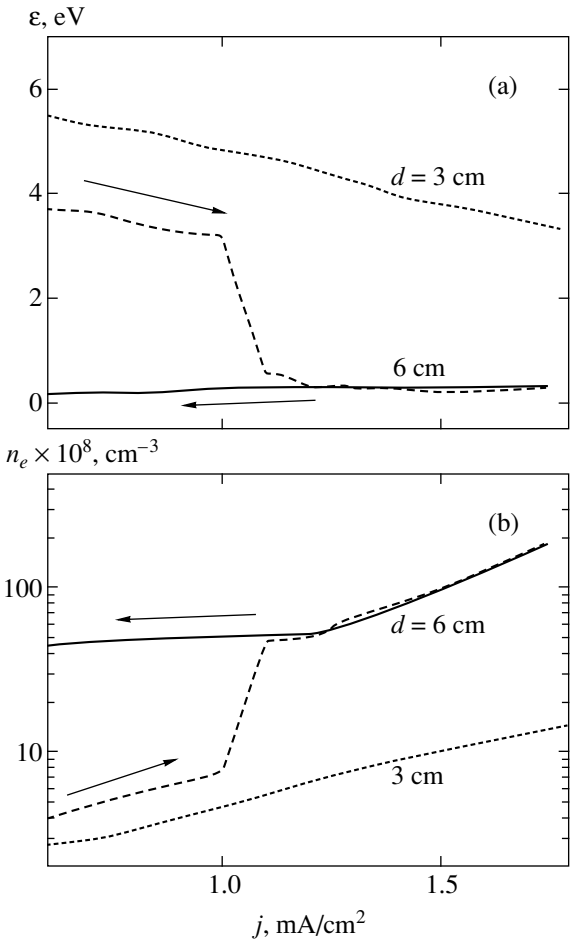



**Fig. 8.** Current-density dependence of $n_e$ (a) and ε (b) at the center of the discharge gap at $P = 0.075$ Torr: dashed and solid curves correspond to the growing- and falling-current branches at $d = 6$ cm; dotted curves, to $d = 3$ cm.

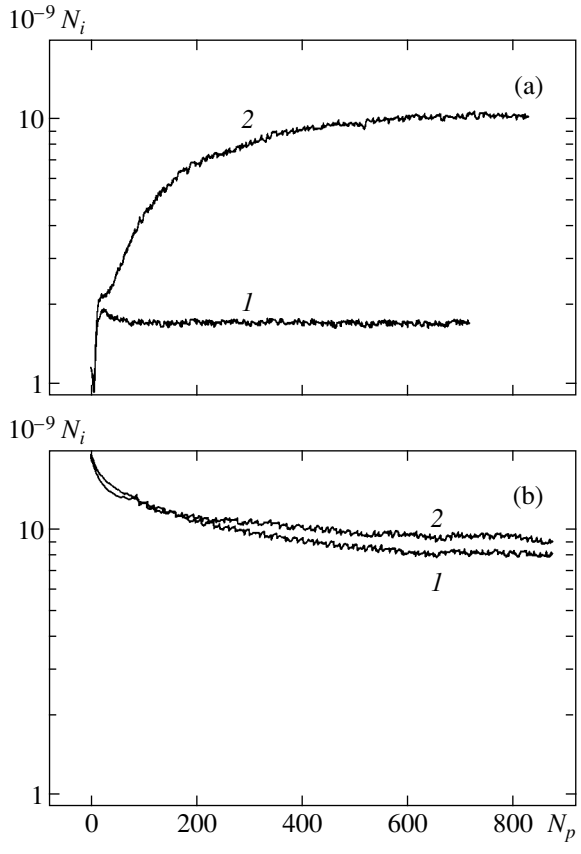



**Fig. 9.** Total ionization $N_i$ versus number of discharge periods for $j$ = (*1*) 0.70 and (*2*) 0.75 mA/cm$^2$ approached (a) "from below" and (b) "from above" for $P = 0.03$ Torr and $d = 6$ cm.

switches from VD into AS mode along the ascending current-density branch, the electron density $n_e^{VD}$ in the VD mode remains relatively low. Therefore, the electric field $E_{sh}$ in the electrode sheath must be sufficiently strong for the energy of a greater part of electrons to exceed the ionization threshold. When $j$ varies along the falling-current branch, $n_e^{AS} \gg n_e^{VD}$, and the high ionization rate required to sustain the AS mode can be reached for lower $E_{sh}$. Thus, the transition from AS to VD mode occurs at a lower $j$*, as compared to the VD–AS transition. This hysteretic behavior is not specific to discharge in methane. It is characteristic of any RF discharge that exhibits two modes.

## 8. CONCLUSIONS

Two modes of capacitively coupled RF discharge in methane are systematically analyzed by numerical simulation based on a combined PIC–MCC algorithm. A phase diagram representing the VD- and AS-mode regions was obtained for $j$ = 0.45–2.2 mA/cm$^2$ and $P$ = 0.01–1 Torr. Critical conditions for the VD–AS transition were computed for wide intervals of current density and gas pressure. It was shown that the transition is not due to the effect of secondary electrons, which is characteristic of discharges in inert gases. Hysteretic behavior of the discharge is analyzed. Different discharge modes were observed when the current density was increased and decreased to a certain value.

## ACKNOWLEDGMENTS

I thank A.S. Smirnov and A.L. Aleksandrov for helpful discussions of the problem and to T.V. Rakhimova for providing data on methane cross sections. This work was supported by grant no. 974354 under the NATO Science for Peace Program and by the Russian Foundation for Basic Research, project no. 04-02-16783.

## REFERENCES

1. S. M. Levitskiĭ, Zh. Tekh. Fiz. **27**, 970 (1957) [Sov. Phys. Tech. Phys. **2**, 887 (1957)].
2. V. A. Godyak, R. B. Piejak, and B. M. Alexandrovich, Phys. Rev. Lett. **68**, 40 (1992).
3. G. J. Parker, W. N. G. Hitchon, and J. E. Lawler, Phys. Fluids B **5**, 646 (1993).
4. Ph. Belenguer and J. P. Boeuf, Phys. Rev. A **41**, 4447 (1990).
5. V. A. Godyak and R. B. Piejak, Phys. Rev. Lett. **65**, 996 (1990).
6. C. Bohm and J. Perrin, J. Phys. D: Appl. Phys. **24**, 865 (1991).
7. J. L. Andujar, E. Bertran, A. Canillas, *et al.*, J. Vac. Sci. Technol. A **9**, 2216 (1991).
8. J. P. Boeuf and Ph. Belenguer, J. Appl. Phys. **71**, 4751 (1992).
9. I. V. Schweigert and V. A. Schweigert, Plasma Sources Sci. Technol. **13**, 315 (2004).
10. V. A. Shveĭgert and I. V. Shveĭgert, Prikl. Mekh. Tekh. Fiz. **29** (4), 16 (1988).
11. T. J. Sommerer and M. J. Kushner, J. Appl. Phys. **71**, 1654 (1992).
12. C. K. Birdsall and A. B. Langdon, *Plasma Physics via Computer Simulation* (McGraw-Hill, New York, 1985).
13. E. Gogolides, D. Mary, A. Rhallabi, and F. Turban, Jpn. J. Appl. Phys., Part 1 **34**, 261 (1995).
14. K. Nagayama, B. Farouk, and Y. H. Lee, IEEE Trans. Plasma Sci. **26**, 125 (1998).
15. W. M. Manheimer, M. Lampe, and G. Joyce, J. Comput. Phys. **138**, 563 (1997).
16. S. V. Berezhnoi, I. D. Kaganovich, and L. D. Tsendin, Plasma Phys. Rep. **24**, 556 (1998).
17. V. Ivanov, O. Proshina, T. Rakhimova, *et al.*, J. Appl. Phys. **91**, 6296 (2002).
18. K. Bera, B. Farouk, and Y. H. Lee, Plasma Sources Sci. Technol. **10**, 211 (2001).
19. K. Bera, B. Farouk, and Y. H. Lee, Plasma Sources Sci. Technol. **8**, 412 (1999).

*Translated by A. Betev*

*Journal of Experimental and Theoretical Physics, Vol. 99, No. 4, 2004, pp. 727–732.*
*Translated from Zhurnal Éksperimental'noĭ i Teoreticheskoĭ Fiziki, Vol. 126, No. 4, 2004, pp. 843–848.*


**PLASMA, GASES**

# Magnetic Field Generation in a Plasma Produced through Atomic Ionization by Circularly Polarized Radiation

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

*Lebedev Physical Institute, Russian Academy of Sciences, Leninskii pr. 53, Moscow, 119991 Russia*
*e-mail: silin@sci.lebedev.ru*

Received March 25, 2004

**Abstract**—We have established the dependences of the maximum Weibel instability growth rate and the corresponding wavenumber on the degree of anisotropy in the photoelectron distribution formed through tunnel atomic ionization in the field of a circularly polarized short laser pulse. We show how the relaxation of the initial distribution of photoelectrons due to their collisions with ions affects the pattern of generation of a quasi-static magnetic field. © *2004 MAIK "Nauka/Interperiodica".*

## 1. INTRODUCTION

When material is exposed to intense ultrashort laser pulses, a plasma with a highly nonequilibrium photoelectron velocity distribution is produced in a time comparable to or shorter than the reciprocal of the fundamental pulse frequency. In particular, if the laser radiation has such parameters that tunnel atomic ionization takes place, then the photoelectron distribution is highly anisotropic. The pattern of the anisotropy depends significantly on the degree of polarization of the ionizing radiation. In the case of linearly polarized radiation, an anisotropic bi-Maxwellian electron velocity distribution elongated in the direction of the field polarization is formed [1, 2]. When the radiation is circularly polarized, the photoelectron velocities are concentrated mainly in the plane of its polarization near the velocity of the electron oscillations in the field, $v_E$ [1]. A plasma with an anisotropic electron velocity distribution is unstable against the growth of Weibel instability, which generates a quasi-static magnetic field [3, 4]. Arefyev *et al.* [5] showed that the Weibel instability could grow via atomic ionization by linearly polarized radiation. Krainov [6] demonstrated a similar possibility for the conditions under which a plasma is produced in the field of circularly polarized radiation. This author found the maximum possible growth rate of the Weibel instability $\gamma_0 \approx \omega_L v_E/\sqrt{2}\,c$, where $\omega_L$ is the electron plasma frequency and $c$ is the speed of light, and showed that more favorable conditions for the generation of a magnetic field are created when atoms are ionized by circularly polarized radiation.

In this paper, we continue to study the Weibel instability due to atomic ionization by circularly polarized radiation. In contrast to [6], the emphasis is on studying the conditions under which the degree of anisotropy in the photoelectron distribution is not assumed to be anomalously large, which is of interest for wide ranges of frequencies and flux densities of the ionizing radiation. Special attention is given to studying the dependence of the maximum instability growth rate $\gamma_{\max}$ on the degree of anisotropy in the photoelectron distribution function. We show that the maximum value of $\gamma_0$ found in [6] is realized only for an anomalously large ratio $v_E/v_T \geq 100$, where $v_T$ is the characteristic photoelectron velocity in a direction orthogonal to the plane of polarization of the radiation; i.e., when the electron oscillation energy in the pump field is four orders of magnitude higher than $mv_T^2$. The value of $\gamma_{\max}$ decreases with decreasing ratio $v_E/v_T$ and is appreciably lower than $\gamma_0$ for $100 \gg v_E/v_T \gg 1$. Lower values of the wavenumbers $k_{\max}$, which determine the spatial structure of the generated magnetic field, correspond to lower values of $v_E/v_T$. As $v_E/v_T$ decreases from 100 to $\sqrt{10}$, the length scale of the exponentially growing magnetic field perturbations increases by a factor of 8. Collisions of photoelectrons with ions and between themselves can significantly affect the growth of Weibel instability, because these collisions cause the degree of anisotropy in the photoelectron distribution function to decrease. Below, we show how the initial distribution of photoelectrons changes due to their collisions with ions. This change is insignificant at the initial instability growth stage if $\gamma_{\max} \gg \nu v_E^2/v_T^2$, where $\nu$ is the electron–ion collision frequency in a strong laser field. In contrast, if $\nu \ll \gamma_{\max} \ll \nu v_E^2/v_T^2$, then the degree of anisotropy in the photoelectron distribution will decrease by a factor of $v_E\sqrt{\nu}/v_T\sqrt{\gamma_{\max}} \gg 1$ before the growth of instability begins. Because of the decrease in

1063-7761/04/9904-0727$26.00 

the degree of anisotropy, the most effectively growing magnetic field perturbations will have a larger length scale and a lower growth rate. The corresponding quantitative changes in $\gamma_{max}$ and $k_{max}$ are plotted as a function of the parameter $v_E/v_T$. For $\nu \ll \gamma_{max} \ll \nu v_E^2/v_T^2$, $v_E/v_T$ should be substituted with the smaller quantity $\sqrt{\gamma_{max}/\nu}$ in the plots.

## 2. THE PHOTOELECTRON VELOCITY DISTRIBUTION

Let us consider the interaction between an ionizing ultrashort laser pulse and material. The pulse duration is assumed to be longer than the short atomic ionization time, but shorter than the time of change in the nonequilibrium photoelectron velocity distribution. Suppose that the laser field is circularly polarized and is roughly described by a relation of the form

$$\mathbf{E} = -\mathbf{i}E\sin(\omega t) + \mathbf{j}E\cos(\omega t), \quad (1)$$

where $\mathbf{i} \perp \mathbf{j}$ are unit vectors, $\omega$ is the fundamental pulse frequency, and $E$ is the electric field strength. We assume that the frequency and strength of the electric field (1) satisfy the conditions

$$mv_E^2 \gg 2I \gg \frac{3}{2}\hbar\omega\sqrt{\frac{mv_E^2}{2I}}, \quad (2)$$

where $v_E = |eE/m\omega|$, $e$ is the electron charge, $m$ is the electron mass, and $I$ is the atomic ionization potential. Under these conditions, tunnel atomic ionization takes place in the electric field (1), and the photoelectron distribution in velocity $\mathbf{v}$ corresponds to the ionization probability $W(\mathbf{v})$ derived in [1]:

$$W(\mathbf{v}) \propto \exp\left\{-\frac{2}{3\hbar\omega v_E\sqrt{m}} \times [2I + mv_z^2 + m(v_\perp - v_E)^2]^{3/2}\right\}. \quad (3)$$

It follows from this relation and from inequalities (2) that the distribution function for the bulk of the photoelectrons with velocities

$$v_z^2 \ll \frac{2I}{m}, \quad (v_\perp - v_E)^2 \ll \frac{2I}{m}, \quad (4)$$

can be roughly represented as

$$f(\mathbf{v}) \approx \frac{n}{2\pi^2 v_T^2} \times \left[v_E + v_E\,\mathrm{erf}\left(\frac{v_E}{\sqrt{2}v_T}\right) + \sqrt{\frac{2}{\pi}}v_T\exp\left(-\frac{v_E^2}{2v_T^2}\right)\right]^{-1} \times \exp\left[-\frac{v_z^2}{2v_T^2} - \frac{1}{2v_T^2}(v_\perp - v_E)^2\right]. \quad (5)$$

Here,

$$\mathrm{erf}(x) = \frac{2}{\sqrt{\pi}}\int_0^x dt\exp(-t^2)$$

is the error function,

$$v_T = \sqrt{\frac{k_B T}{m}}, \quad k_B T = \frac{\hbar\omega}{2}\sqrt{\frac{mv_E^2}{2I}}, \quad (6)$$

$k_B$ is the Boltzmann constant, $T$ is the effective temperature, and $n$ is the photoelectron density. It follows from relations (6) and (2) that $v_E \gg v_T$. This inequality allows distribution (5) to be substituted with a simpler one:

$$f(\mathbf{v}) \approx \frac{n}{4\pi^2 v_E v_T^2}\exp\left[-\frac{v_z^2}{2v_T^2} - \frac{1}{2v_T^2}(v_\perp - v_E)^2\right]. \quad (7)$$

Relation (7) yields $\sqrt{\langle v_z^2\rangle} \propto v_T$ and $\langle v_\perp\rangle \propto v_E$ for the mean photoelectron velocities. Taking into account this estimate, by the degree of anisotropy in the photoelectron distribution we will mean the ratio $v_E/v_T$. This estimate is justified when the influence of electron–electron collisions is negligible. The highly nonequilibrium photoelectron velocity distribution (7) forms the basis for the subsequent analysis of the pattern of generation of a quasi-static magnetic field in a plasma.

## 3. THE GROWTH RATE OF THE WEIBEL INSTABILITY

A plasma with an anisotropic photoelectron distribution (7) is unstable against the growth of Weibel instability. Assuming, as usual, that the small electric and magnetic field perturbations are in the form

$$\delta\mathbf{E} \propto \delta\mathbf{B} \propto \exp(-i\omega t + i\mathbf{k}\cdot\mathbf{r}), \quad (8)$$

we find from the Maxwell equations that

$$\Lambda_{ij}(\omega, \mathbf{k})\delta E_j = 0, \tag{9}$$

$$\Lambda_{ij}(\omega, \mathbf{k}) = \delta_{ij}\left(1 - \frac{\omega_L^2}{\omega^2} - \frac{k^2c^2}{\omega^2}\right) + \frac{\omega_L^2}{n\omega^2} \times \int d\mathbf{v}\frac{v_i v_j}{\omega - \mathbf{k}\cdot\mathbf{v}}\left(\mathbf{k}\cdot\frac{\partial f}{\partial \mathbf{v}}\right), \tag{10}$$

where $\mathbf{k}$ is the wave vector of the perturbations and $\omega_L = \sqrt{4\pi e^2 n/m}$ is the photoelectron plasma frequency. According to [7], the instability growth rate is at a maximum for field perturbations with the wave vector $\mathbf{k} = (0, 0, k)$ and the fields $\delta\mathbf{E} = (\delta E, 0, 0)$, $\delta\mathbf{B} = (0, \delta B, 0)$ or $\delta\mathbf{E} = (0, \delta E, 0)$, $\delta\mathbf{B} = (\delta B, 0, 0)$. For this configuration of the field perturbations, we derive the following dispersion relation for the instability growth rate from Eqs. (9) and (10):

$$\Lambda_{xx}(\omega, \mathbf{k}) = \Lambda_{yy}(\omega, \mathbf{k}) = 1 - \frac{\omega_L^2}{\omega^2} - \frac{k^2c^2}{\omega^2} + \frac{\omega_L^2}{2\omega^2}\left(3 + \frac{v_E^2}{v_T^2}\right)\left[1 - J_+\left(\frac{\omega}{kv_T}\right)\right] = 0, \tag{11}$$

where the function $J_+(\beta)$ is [8]

$$J_+(\beta) = \beta\exp\left(-\frac{\beta^2}{2}\right)\int_{i\infty}^{\beta} dy \exp\left(\frac{y^2}{2}\right). \tag{12}$$

At a low perturbation frequency, $|\omega| \ll kv_T \ll kc$, taking into account the approximate relation

$$J_+(\beta) \approx -i\beta\sqrt{\pi/2}, \tag{13}$$

we find from Eq. (11) that

$$\left(\frac{3}{2} + \frac{v_E^2}{2v_T^2}\right)\left(1 + i\sqrt{\frac{\pi}{2}}\frac{\omega}{kv_T}\right) = 1 + \frac{k^2c^2}{\omega_L^2}. \tag{14}$$

Since the ratio $v_T/v_E$ is small, we obtain the growth rate of the Weibel instability from (14):

$$\gamma = \mathrm{Im}\,\omega = \sqrt{\frac{2}{\pi}}kv_T\left(1 - \frac{2v_T^2}{v_E^2}\frac{k^2c^2}{\omega_L^2}\right) \ll kv_T. \tag{15}$$

The instability growth rate (15) is comparatively small in absolute value for wavenumbers

$$k \approx \frac{\omega_L}{\sqrt{2}c}\frac{v_E}{v_T}, \tag{16}$$

its value is much smaller than $v_E\omega_L/c\sqrt{2}$, which is much lower than $\omega_L$.

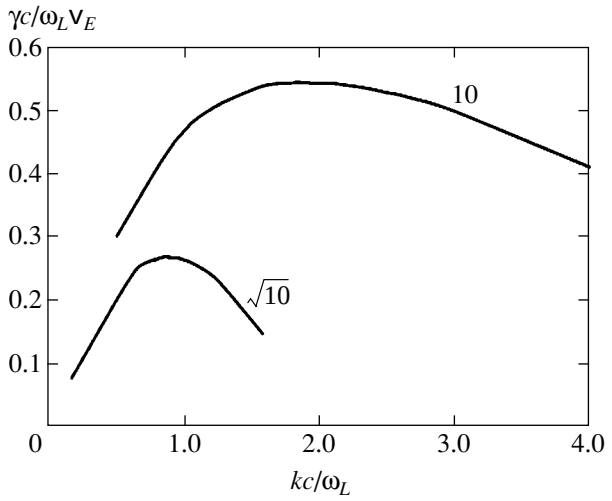



**Fig. 1.** Weibel instability growth rate $\gamma$ versus wavenumber $k$. The curves correspond to two values of $v_E/v_T = \sqrt{10}$ and 10.

In the opposite limiting case where $|\omega| \gg kv_T$, using the expansion

$$J_+(\beta) \approx 1 + \beta^{-2}, \quad |\beta| \gg 1, \tag{17}$$

we find from (11) that

$$\omega^4 - \omega^2(\omega_L^2 + k^2c^2) - \frac{1}{2}\omega_L^2k^2(v_E^2 + 3v_T^2) = 0. \tag{18}$$

The dispersion relation (18) has the approximate solution

$$\gamma \approx \frac{\omega_L}{\sqrt{2}\sqrt{\omega_L^2 + k^2c^2}}\sqrt{v_E^2 + 3v_T^2} \approx \frac{\omega_L k v_E}{\sqrt{2}\sqrt{\omega_L^2 + k^2c^2}} \gg kv_T, \tag{19}$$

where the inequality $\gamma \gg kv_T$ defines the range of wavenumbers in which solution (19) is applicable. It follows from this inequality that relation (19) has a wide range of applicability: $k \ll \omega_L v_E/cv_T\sqrt{2}$. At wavenumbers $\sim\omega_L v_E/cv_T\sqrt{2}$, (19) is joined with (15). The instability growth rate $\gamma$ is plotted against $k$ in Fig. 1. The curves in Fig. 1 describe the numerical solution of the equation

$$\left(\frac{kc}{\omega_L}\right)^2 + 1 = \left(1 + 3\frac{v_T^2}{v_E^2}\right)\frac{1}{\sqrt{8\pi}}\int_{-\infty}^{\infty} dt t^2 \times \exp\left(-\frac{t^2}{2}\right)\left[\left(\frac{\gamma}{kv_E}\right)^2 + \left(\frac{v_T}{v_E}t\right)^2\right]^{-1}, \tag{20}$$

that follows from (11) for $\gamma \ll \omega_L$. The curves in Fig. 1 correspond to two values of the parameter $v_E/v_T$, $\sqrt{10}$

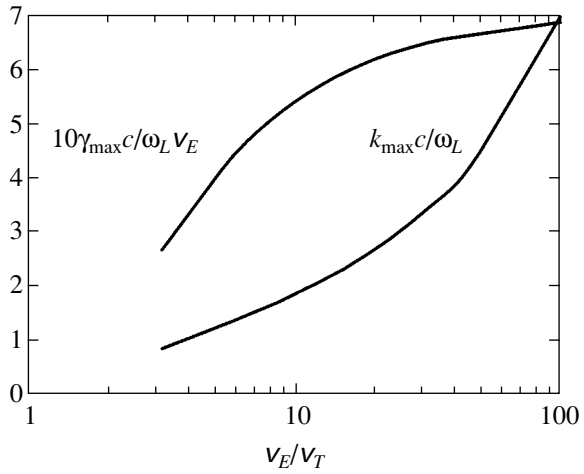



**Fig. 2.** Maximum Weibel instability growth rate $\gamma_{max}$ and wavenumber $k_{max}$, which determines the inhomogeneity length scale of the most rapidly growing perturbations, versus $v_E/v_T$.

and 10. Figure 1 shows that the growth rate of the Weibel instability has a distinct maximum. The larger the ratio $v_E/v_T$, which characterizes the degree of anisotropy in the photoelectron distribution function, the higher the maximum growth rate $\gamma_{max}$. The wavenumber $k_{max}$, which determines the length scale of the most rapidly growing field perturbations, increases with increasing $v_E/v_T$. The maximum instability growth rate and the corresponding wavenumber $k_{max}$ are plotted against $v_E/v_T$ in Fig. 2. Figure 2 shows that the growth rate $\gamma_{max} \approx 0.28\omega_L v_E/c$ at $v_E/v_T = \sqrt{10}$ first increases with $v_E/v_T$ and then reaches its maximum. We see from Fig. 2 that the maximum growth rate obtained previously [6] is defined by the formula

$$\gamma_0 \approx \frac{\omega_L v_E}{\sqrt{2}c} \approx 0.7\frac{\omega_L v_E}{c} \tag{21}$$

and is reached at an anomalously large ratio $v_E/v_T \geq 100$. This value is approximately twice $v_E/v_T \propto \sqrt{10}$. This change in the exponent that determines the change in the magnetic energy density,

$$\frac{\delta B^2(k_{max}, t)}{\delta B^2(k_{max}, t = 0)} \propto \exp(2\gamma_{max}t),$$

where $2\gamma_{max}t \gg 1$, is significant. The increase in $k_{max}$ with $v_E/v_T$ is larger than that in $\gamma_{max}$. As we see from Fig. 2, $k_{max}$ increases by a factor of 8 as $v_E/v_T$ changes from $\sqrt{10}$ to 100. This implies that the length scale of the generated quasi-static magnetic field decreases by the same factor. The curves in Fig. 2 indicate that the growth rate of the magnetic field and its spatial structure depend significantly on the degree of anisotropy in the photoelectron distribution function.

## 4. RELAXATION OF THE PHOTOELECTRON DISTRIBUTION

The growth time scale of the Weibel instability is $1/\gamma_{max} \geq \sqrt{2}\,c/\omega_L v_E$. It makes sense to talk about the instability growth and the related generation of a quasi-static magnetic field if the reciprocal of the instability growth rate is shorter than the time of relaxation of an anisotropic photoelectron distribution to an isotropic one. Collisions of photoelectrons with ions and between themselves cause their distribution to be isotropized. The influence of collisions on the degree of anisotropy in the photoelectron distribution when atoms are exposed to linearly polarized radiation was considered in [5] (see also [9] for numerical estimates and a discussion). Let us perform the corresponding analysis for circularly polarized radiation whose exposure time is shorter than the time scale of the change in the photoelectron distribution, but longer than the comparatively short tunnel atomic ionization time. If the degree of ionization of ions $Z$ is greater than one, then the isotropization is determined mainly by electron–ion collisions. In this case, the evolution of the initial photoelectron distribution (7) is described by an equation of the form

$$\frac{\partial f}{\partial t} = \frac{1}{2}\nu(v)\frac{\partial}{\partial \xi}(1-\xi^2)\frac{\partial f}{\partial \xi}, \quad -1 \leq \xi \leq 1, \tag{22}$$

where $f = f(v, \xi, t)$, $\xi = \cos\theta$, $\theta$ is the angle between the velocity vector $\mathbf{v}$ and the axis of symmetry of distribution (7),

$$\nu(v) = 4\pi Z e^4 n \Lambda m^{-2} v^{-3} \tag{23}$$

is the electron–ion collision frequency, and $\Lambda$ is the Coulomb logarithm. Since $v_E \gg v_T$ for the initial distribution, the bulk of the photoelectrons at the initial time is concentrated in the relatively narrow velocity range $v_E - v_T \leq v \leq v_E + v_T$ and in a range of angles close to $\pi/2$, when $\xi \leq v_T/v_E \ll 1$. This implies that to describe the relaxation of the initial photoelectron distribution on short time scales, Eq. (22) may be substituted with a simpler approximate equation:

$$\frac{\partial f}{\partial t} \approx \frac{1}{2}\nu\frac{\partial^2 f}{\partial \xi^2}, \tag{24}$$

where $\nu \approx \nu(v = v_E)$. This equation yields an estimate for the influence of collisions on the anisotropy in the photoelectron distribution function that is sufficient for our discussion. Assuming that $\nu t \ll 1$, the approximate solution of Eq. (24) can be represented as

$$f(v, \xi, t) = \int_{-\infty}^{\infty} \frac{d\xi'}{\sqrt{2\pi\nu t}} \times \exp\left[-\frac{(\xi - \xi')^2}{2\nu t}\right] f(v, \xi', t = 0), \tag{25}$$

where $\xi \le \nu t \ll 1$, and the initial distribution $f(v, \xi, t = 0)$ is described by Eq. (7) in which $v_z = v\xi$ and $v_\perp = v\sqrt{1-\xi^2}$. On short time scales when $\nu t \ll v_T^2/v_E^2$, distribution (25) is close to the initial photoelectron distribution. If, however,

$$v_T^2/v_E^2 \ll \nu t \ll 1,$$

then we derive an approximate relation from (25):

$$f(v, \xi, t) \approx \frac{n}{4\pi^2 v_E^2 v_T \sqrt{\nu t}} \times \exp\left[-\frac{\xi^2}{2\nu t} - \frac{(v - v_E)^2}{2v_T^2}\right]. \quad (26)$$

According to (25) and (26), electron–electron collisions cause the angular photoelectron distribution to broaden on short time scales. As we see from relation (26), the localization region of the bulk of the electrons in velocity space expands with time as $-\sqrt{\nu t} \le \xi \le \sqrt{\nu t}$. According to Eq. (26), the degree of anisotropy $v_E/\sqrt{\langle v_z^2 \rangle}$ in the electron distribution is given by

$$\frac{v_E}{v_T} \gg \frac{v_E}{\sqrt{\langle v_z^2 \rangle}} \propto \frac{1}{\sqrt{\nu t}} \gg 1.$$

At $\nu t \approx 1$, electron–ion collisions cause the photoelectron distribution to be isotropized. For $Z \gg 1$ on time scales $\nu t \le 1$, we may disregard the influence of electron–electron collisions on the relaxation of the initial photoelectron distribution. If, however, $Z = 1$, then electron–electron collisions speed up the isotropization of the photoelectron distribution by a factor of about 2. In addition, collisions between electrons result in relaxation of the electron energy, which is accompanied by the formation of a Maxwellian electron distribution also in a time ~$1/\nu$. It follows from the above discussion that the conservation time scale of the anisotropy in the photoelectron distribution is approximately $1/\nu$.

## 5. DISCUSSION

The pattern of relaxation of the initial photoelectron distribution described above allows us to specify the conditions under which the Weibel instability can grow. If the maximum instability growth rate $\gamma_{max}$ is larger than $\nu v_E^2/v_T^2$, then the exponential growth of a spontaneous magnetic field is characterized by the rate $\gamma_{max}$ dependent on $v_E/v_T$ that arises immediately after the tunnel ionization of atoms. The wavenumber $k_{max}$, which determines the length scale of the magnetic field, corresponds to the same value of $v_E/v_T$. If, however, the growth rate satisfies the inequalities $\nu \ll \gamma_{max} \ll \nu v_E^2/v_T^2$, then, before a spontaneous magnetic field begins to grow exponentially, electron–ion collisions will cause the degree of anisotropy in the initial velocity distribution to decrease. According to relation (26), because of electron–ion collisions at $\nu t \gg v_T^2/v_E^2$, the photoelectron distribution function along the axis of symmetry is characterized not by the velocity $v_T$, but by the larger quantity $v_E\sqrt{\nu t}$, which satisfies the inequalities $v_T \ll v_E\sqrt{\nu t} \ll v_E$. If we ignore the quantitative changes in $\gamma_{max}$ due to the decrease in the anisotropy of the photoelectron distribution (see Fig. 2), then we may assume that by the beginning of the exponential growth of the magnetic field, the anisotropy will be characterized not by the parameter $v_E/v_T$, but by the smaller, but still fairly large parameter $\sqrt{\gamma_{max}/\nu}$, $v_E/v_T \gg \sqrt{\gamma_{max}/\nu} \gg 1$. Under these conditions, the instability growth is also possible, but, as we see from Fig. 2, it is characterized by a slightly lower rate $\gamma_{max}$ and a smaller wavenumber $k_{max}$. It follows from the above discussion that the necessary condition for the growth of Weibel instability is $\gamma_{max} \propto \omega_L v_E/c \gg \nu$, where the electron–ion collision frequency decreases as $\nu_E^{-3}$ with increasing $v_E$. This condition is much weaker than that obtained previously [5] when considering the Weibel instability of a plasma with a photoelectron distribution formed through tunnel atomic ionization in the field of linearly polarized radiation. The degree of weakening of the inequality $\gamma_{max} \gg \nu$ is characterized by the large parameter $A^4 \gg 1$, where

$$A = \left[\frac{2(2I)^{3/2}}{3\hbar\omega\sqrt{m}v_E}\right]^{1/2} \gg 1. \quad (27)$$

The relation $A \gg 1$ follows from inequalities (2), which define the boundaries of the region where the regime of tunnel atomic ionization is realized. At $A \gg 1$, the parameter $A$ approximately characterizes the ratio of the kinetic energies of the photoelectrons produced in the fields of circularly and linearly polarized radiation. For linearly polarized radiation, the kinetic energy is appreciably lower. The relative increase in $\gamma_{max}$ due to the change in radiation polarization has $A \gg 1$. The weakening of the inequality $\gamma_{max} \gg \nu$ and the relative increase in $\gamma_{max}$ confirm the conclusion reached in [6] that in the regime of tunnel atomic ionization, circularly polarized radiation generates a quasi-static magnetic field much more efficiently than linearly polarized radiation. Another advantage of using circularly polarized radiation is that at the same flux density of the ionizing radiation, it generates a magnetic field that is a factor of $A \gg 1$ stronger than that generated by linearly polarized radiation. Following [5, 9], we can estimate the maxi-

mum strength of the generated magnetic field from the relation

$$\frac{\delta B^2}{8\pi} \approx 0.1\frac{1}{2}nmv_E^2, \quad (28)$$

where the coefficient 0.1 was taken from the numerical study by Wallace *et al.* [10] and means that ~10% of the particle kinetic energy transforms into magnetic energy. For our estimates, we assume that radiation with a frequency of $\omega = 2 \times 10^{15}$ s$^{-1}$ and a flux density of $q = 2 \times 10^{15}$ W cm$^{-2}$ ionizes helium atoms with an ionization potential of $I \approx 24.4$ eV. The photoelectron density is taken to be $n \approx 10^{21}$ cm$^{-3}$. For such plasma and radiation parameters, the tunnel ionization of helium atoms takes place, and the radiation itself penetrates deep into the plasma being produced. In this case, relation (28) yields an estimate for the magnetic field strength of $B \approx 1$ mG.

## ACKNOWLEDGMENTS

This work was supported by the Russian Foundation for Basic Research (project no. 02-02-16047), State Support for Leading Scientific Schools of Russia (NSh-1385.2003.2), and INTAS (grant no. 03-51-5037).

## REFERENCES

1. N. B. Delone and V. P. Krainov, J. Opt. Soc. Am. B **8**, 1207 (1991).
2. N. B. Delone and V. P. Kraĭnov, Usp. Fiz. Nauk **168**, 531 (1998) [Phys. Usp. **41**, 487 (1998)].
3. E. S. Weibel, Phys. Rev. Lett. **2**, 83 (1959).
4. A. F. Aleksandrov, L. S. Bogdankevich, and A. A. Rukhadze, *Principles of Plasma Electrodynamics*, 2nd ed. (Vysshaya Shkola, Moscow, 1988; Springer, Berlin, 1984).
5. V. I. Arefyev, V. P. Silin, and S. A. Uryupin, Phys. Lett. A **255**, 307 (1999).
6. V. P. Kraĭnov, Zh. Éksp. Teor. Fiz. **123**, 487 (2003) [JETP **96**, 430 (2003)].
7. A. Yu. Romanov, V. P. Silin, and S. A. Uryupin, Zh. Éksp. Teor. Fiz. **111**, 1245 (1997) [JETP **84**, 687 (1997)].
8. V. P. Silin and A. A. Rukhadze, *Electromagnetic Properties of Plasma and Plasma-Like Media* (Atomizdat, Moscow, 1961) [in Russian].
9. V. P. Krainov, J. Phys. B: At. Mol. Opt. Phys. **36**, 3187 (2003).
10. M. Wallace, J. U. Brackbill, C. W. Granfill, *et al.*, Phys. Fluids **30**, 1085 (1987).

*Translated by V. Astakhov*

*Journal of Experimental and Theoretical Physics, Vol. 99, No. 4, 2004, pp. 733–740.*
*Translated from Zhurnal Éksperimental'noĭ i Teoreticheskoĭ Fiziki, Vol. 126, No. 4, 2004, pp. 849–856.*


**PLASMA, GASES**

# Transport of Strongly Comptonized Radiation

**S. A. Wagner**

*Moscow Institute of Physics and Technology, Dolgoprudnyĭ, Moscow oblast, 141700 Russia*

*e-mail: s_wagner@mail.ru*

Received May 24, 2004

**Abstract**—For strongly Comptonized radiation in a nonrelativistic plasma, we calculate the bremsstrahlung source of Comptonized photons with an accurate allowance for free–free absorption and nonlinear stimulated Compton emission. We formulate radiation hydrodynamics equations that are valid in the limit of strong Compton energy exchange between plasma and radiation. We derive a formula for the energy dissipation rate under these conditions. For an optically thick region, we have found an equation that describes the spatial variation in the exponential fall-off factor of the radiation spectrum. © *2004 MAIK "Nauka/Interperiodica".*

## 1. INTRODUCTION

The issues of interaction between intense radiation and matter are of great practical and fundamental importance. The manifestations of inverse Compton scattering feature prominently among them. For example, on a laboratory scale, it can be essential for estimating the radiative energy losses when a body undergoes shock compression to a high density. Under astrophysical conditions, this effect is also noticeable for rarefied media due to their large scales. It has an effect on the radiation spectrum called Comptonization. The bremsstrahlung amplification and the spectral pattern under strong Comptonization were first reported by Kompaneets [1]. Subsequently, Comptonization was studied in a homogeneous case (see, e.g., the review [2]) and was widely used to interpret astrophysical data. A number of authors also generalized the Kompaneets kinetic equation to an inhomogeneous plasma (see, e.g., [3] and references therein) and considered various solutions of this equation (see, e.g., [4, 5] and references therein). Nevertheless, these solutions allow one neither to determine the emergent luminosity of a strongly Comptonized plasma region nor to describe the Compton energy exchange to find the plasma temperature in a self-consistent way.

Under strong Comptonization, the part of the spectrum that makes a major contribution to the radiation energy density has a Wien distribution and can be described by its photon number density. In a spatially inhomogeneous case, the parameters that completely characterize the radiation spectrum require an equation that describes their spatial variation. In particular, this is needed to calculate the radiative losses, because the emergent radiation is determined not only by the bremsstrahlung emission, but also by the optical depth of the emitting medium and by its temperature profile. A continuity equation for the photon number in the system of radiation hydrodynamics equations under the domination of Compton scattering was formulated in [6]. However, no accurate expression for the photon number source has been published as yet. Such an expression implies an allowance for both absorption and nonlinear stimulated Compton emission. The irreversible energy release in an inhomogeneous plasma with strongly Comptonized radiation has not been considered in general form at all.

The main goal of this study is to calculate the effective source of the photon number generated by bremsstrahlung processes in a nonrelativistic plasma. This quantity and the photon number density calculated from it are used in the formula for the total energy release in a Comptonized region that is obtained from the radiation hydrodynamics equations under strong Compton energy exchange.

## 2. THE KINETIC EQUATION

In this paper, we consider the continuum X-ray radiation produced by thermal bremsstrahlung emission in a plasma. The plasma is assumed to be in local equilibrium and have a temperature $T$, an electron density $n_e$, and a nonrelativistic flow velocity $\mathbf{u}$. This radiation can be locally described by the occupation number of unpolarized photons in momentum ($\mathbf{p}$) space at point $\mathbf{r}$ in the reference frame comoving with each small plasma part in local equilibrium. Several parameters that characterize the radiation transport impose further conditions on these quantities. They usually hold in astrophysical applications [7].

The angular distribution of the radiation depends on Thomson scattering with the opacity

$$k_T = \sigma_T n_e,$$

where $\sigma_T$ is the Thomson cross section. The coefficient

$k_T$ is also related to the optical depth $\tau$ by the expression

$$d\tau = k_T dl$$

that gives the fraction of the scattered radiation on the path $dl$ in the direction under consideration. Below, we assume that $\tau \gg 1$ for all significant sizes. The kinetic description of the radiation then reduces to a photon occupation number $n(t, \mathbf{r}, p)$ averaged over the $\mathbf{p}$ directions, which obeys the equation

$$\hat{L}n = \frac{1}{c}\left(\frac{\delta n}{\delta t}\right)^C + \frac{1}{c}\left(\frac{\delta n}{\delta t}\right)^B. \quad (1)$$

For a plasma moving with a nonrelativistic velocity $\mathbf{u}$, the left-hand side of Eq. (1) may be written as [8]

$$\hat{L}n = \frac{1}{c}\frac{\partial n}{\partial t} + \nabla\left\{\frac{\mathbf{u}}{c}n - \frac{1}{3k_T}\nabla n\right\} - \frac{1}{p^2}\frac{\partial}{\partial p}\left\{p^3\frac{\nabla \mathbf{u}}{3c}n\right\}. \quad (2)$$

Here, the terms related to the spatial variation describe the radiation advection by the plasma flow and the radiation diffusion. The term related to the variation in momentum space describes the contribution of the Doppler effect to the radiation spectrum. For clarity, it makes sense to emphasize that the differentiation operations $\nabla$ in the expression for $\hat{L}$ are performed at fixed $p$. In particular, when another variable $q = pc/T(\mathbf{r})$ is used below, they transform as

$$\nabla = \nabla|_q - \left(\frac{\nabla T}{T}\right)q\frac{\partial}{\partial q}. \quad (3)$$

The terms on the right-hand side of Eq. (1) are the collision integrals for Compton scattering and bremsstrahlung processes. For plasma temperatures $T \ll m_e c^2$, where $m_e$ is the electron mass and $c$ is the speed of light, the Compton collision integral takes the form [1]

$$\left(\frac{\delta n}{\delta t}\right)^C = \frac{1}{p^2}\frac{\partial}{\partial p}p^2 J^C$$

with the radial flux density in momentum space

$$J^C = ck_T p^2\left(\frac{n(1+n)}{m_e c} + \frac{T}{m_e c^2}\frac{\partial n}{\partial p}\right).$$

Here, the first and second terms are responsible for cooling and heating the radiation by plasma electrons, respectively. In astrophysical literature, the action of inverse Compton scattering for optical depths $\tau \gg 1$ is commonly described by the parameter

$$Y_C = \frac{4T}{m_e c^2}\tau^2.$$

This is the relative fraction of the energy gained by the radiation from plasma electrons when scattered from a region with an optical depth $\tau$. In order of magnitude, this parameter can also be interpreted as the ratio of the Compton heating term to the spatial diffusion term in Eq. (1).

The contributions of emission and absorption in the bremsstrahlung collision integral can be written using one quantity, the effective bremsstrahlung (free–free) opacity $k_B$:

$$\frac{1}{c}\left(\frac{\delta n}{\delta t}\right)^B = k_B n_P[(1+n) - ne^q] \equiv k_B[n_p - n].$$

The latter equality implies that the Plank occupation number $n_P = (e^q - 1)^{-1}$ is an equilibrium one. For photon energies $pc \lesssim T$ and plasma temperatures $T \gg I_\mathrm{H}$, where $I_\mathrm{H}$ is the atomic hydrogen ionization potential, the effective bremsstrahlung opacity [9] is

$$k_B = \kappa\frac{\phi(q)}{q^2}, \quad \phi(q) = \frac{1 - e^{-q}}{q}e^{q/2}K_0\left(\frac{q}{2}\right), \quad (4)$$

where $K_0(s)$ is the modified Bessel function of the second kind. For collisions between electrons and ions of each type $a$ with density $n_a$ and charge $z_a e$,

$$\kappa = \left(\sum_a z_a^2 n_a\right)n_e\sigma_T\frac{(hc)^3(TI_\mathrm{H})^{1/2}}{2\pi^{5/2}T^4}.$$

Below, we consider the problem of finding $n$ for Eq. (1) in the absence of extrinsic sources of high-energy photons, i.e., with the boundary condition

$$p^2\left(\frac{p}{3}(\nabla\cdot\mathbf{u})n + J^C\right) = o(1) \text{ for } p \longrightarrow \infty. \quad (5)$$

The boundary conditions for the spatial variable $\mathbf{r}$ are not used here. They are assumed to have no effect on the intrinsic spatial and energy scales of the $n(t, \mathbf{r}, p)$ variations in the spatial regions under consideration. The satisfaction of a similar condition is also implied for the time dependence of the plasma parameters $n_e$, $T$, and $\mathbf{u}$. In other words, we consider only the quasi-stationary limit for both Eq. (1) and the hydrodynamic plasma motion.

The choice of characteristic scales in Eq. (1) depends on the question under consideration. If plasma is accreted into an inner region and the emergent luminosity needs to be found, then the characteristic intrinsic spatial scales are determined by an order-of-magnitude equality of the diffusion and advection terms in Eq. (2). For flows with a moderately complex streamline topology, for example, with a spherically or axially symmetric topology, this equality is achieved on a cer-

tain surface. The geometric size of this surface $R$ for each direction of propagation of the radiation is then a characteristic scale. The corresponding optical depth is $\tau \sim k_T R$, and the plasma parameters must be taken at the above surface. If, however, the local radiation spectrum must be investigated, then the local sizes of the $n_e$, **u**, and $T$ profiles will be characteristic ones. For example, for spherically symmetric outflow and power-law radial dependences of these quantities, the radius (the distance to the center of symmetry) at a given point will be a characteristic scale. In this study, the conditions $\tau \gg 1$ and $Y_C \gg 1$ are assumed to be satisfied both for global characteristic scales of the problem and locally in a certain spatial region.

Therefore, it seems convenient to reduce Eq. (1) to dimensionless equations:

$$-q^4\left(\frac{\partial n}{\partial q} + n + n^2\right) = j, \tag{6}$$

$$\frac{\partial j}{\partial q} = -q^2\hat{\zeta}n + \xi(n_P - n)\phi(q). \tag{7}$$

Here, the operator

$$\hat{\zeta} = \frac{m_e c^2}{k_T T}\hat{L}$$

is on the order of $\zeta \sim 1/Y_C \ll 1$. The relative role of the bremsstrahlung processes in Eq. (7)is described by the parameter

$$\xi = \frac{\kappa m_e c^2}{k_T T}.$$

The inequality $\xi \ll 1$ is the last condition imposed on the plasma parameters in this paper.

## 3. THE EFFECTIVE PHOTON SOURCE

### *3.1. Preliminary Estimates*

For $q \ll 1$, the function $\phi$ in (4) is

$$\phi(q) = \ln\frac{4\eta_E}{q}[1 + O(q)], \quad \eta_E = 0.561459\ldots.$$

Therefore, we can conclude from our estimates of the derivatives $q\partial/\partial q \sim 1$ in (6) and (7) that the absorption is significant for $q \sim \xi^{1/2}$, with $n \sim \xi^{1/2}$ and $j \sim \xi$. Indeed, since the absorption dominates for $q \ll \xi^{1/2}$, the occupation number must approach its equilibrium Planck value $n_P \approx 1/q$ and remains the same in order of magnitude for $q \approx \xi^{1/2}$.

A further comparison of the terms in (7) in the range $\xi^{1/2} \ll q \ll (\xi/\zeta)^{1/3}$ shows that an increase in the Comptonizing flux $j$ can be caused by the bremsstrahlung source:

$$\frac{\partial j}{\partial q} \sim \xi\frac{1}{q}\ln\frac{4\eta_E}{q}.$$

Hence,

$$j \sim \xi\frac{1}{2}\ln^2\frac{4\eta_E}{\sqrt{\xi}}.$$

This flux acts as an effective source along with the bremsstrahlung proper. Being spent on the spatial escape of the radiation, it decreases for $q \gg (\xi/\zeta)^{1/3}$.

### *3.2. Radiation in a Strongly Comptonized Region*

More accurate relations arise if we construct the expansion in the lowest order of $\xi \ll 1$ and $\zeta \ll 1$ uniformly for $q$ and $n$. It is clear from (6) and (7) that on a scale $q \sim 1$ and for any $n$, this expansion must be identical to the expansion in powers of the right-hand side of (7). In other words,

$$j(q) = j_0 + j_1(q) + \ldots \tag{8}$$

with

$$j_1(q) = \int_q^{q_0} dq_1 q_1^2\hat{\zeta}n_0(q_1; j_0)$$

$$-\xi\int_q^{q_0} dq_1[n_P(q_1) - n_0(q_1; j_0)]\phi(q_1)$$

and

$$n = n_0(q; j_0) + n_1(q; A_1; j_0) + \ldots \tag{9}$$

with

$$n_1(q; A_1; j_0) = u(q; j_0)\left(A_1 + \int_q^{q_1} dq_2\frac{j_1(q_2)}{u(q_2; j_0)q_2^4}\right).$$

In these expansions, the zero-order term $n_0(q; j)$ denotes the solutions of Eq. (6) at fixed $j$. The function $u(q; j)$ in (9) denotes the solution of the homogeneous equation

$$\frac{\partial u}{\partial q} + u[1 + 2n_0(q; j)] = 0$$

that corresponds to the inhomogeneous equation for the first-order correction $n_1(q; A_1; j)$, which can be obtained from (6).

However, to use the boundary condition (5), we need to find out whether this expansion is valid for $q \gg 1$. At $j_0 \neq 0$, it is easy to establish the asymptotic behavior of the functions

$$n_0(q; j_0) = -j_0 q^{-4}[1 + O(1/q)],$$

$$u(q; j_0) = e^{-q}[1 + O(1/q)], \quad q \longrightarrow \infty.$$

It thus follows that, on the one hand, expansion (9) proves to be uniform for $q \longrightarrow \infty$, thereby allowing external boundary conditions to be imposed on it, but, on the other hand, it does not satisfy the specific condition (5). Therefore, we should choose $j_0 = 0$ and set $q_0 = \infty$ in expansion (8).

Solving Eq. (6) with $j = 0$, we can easily make sure that

$$n_0(q; j_0)\big|_{j_0 = 0} = (A_0^{-1}e^q - 1)^{-1},$$

$$u(q) = e^q(e^q - A_0)^{-2}.$$

To further study the uniformity, we must retain the leading (for $q \gg 1$) terms in Eq. (9). Since these terms contain $\hat{\zeta}$, we must turn to the explicit form (2) of the operator $\hat{L}$ with allowance made for (3). In this way, we obtain

$$n = A_0\left[1 + \frac{m_e c^2}{3k_T^2 T}\left(\frac{\nabla T}{T}\right)^2 q + O(\ln q)\right]e^{-q} + \ldots,$$

suggesting that the expansion is nonuniform because of the contribution from the spatial variations at $q \sim 1/\zeta$. Nevertheless, one might expect the matching with the uniform (in $\mathbf{r}$) expansion of $n$ for $q \gg 1$ that satisfies (5) to yield $q_1$. However, this quantity is not needed for the final result of this study. Therefore, the next section only briefly shows the existence of this expansion.

### *3.3. Exponential Asymptotics of the Spectrum*

To construct the expansion that satisfies (5), we must take into account the established exponential behavior of $n$ for $q \gg 1$. In other words, we must substitute $n = \exp(-g)$ into the original equation (1) and expand it in terms of $1/g \ll 1$ by assuming that $p\partial g/\partial p \sim g$ and $|R\nabla g| \sim g$. This procedure yields the equation

$$-\frac{1}{3k_T}(\nabla g)^2 = -\frac{p^2 k_T}{m_e c}\frac{\partial g}{\partial p}\left(1 - \frac{T}{c}\frac{\partial g}{\partial p}\right), \tag{10}$$

for the lowest-order approximation of $g$ that balances the terms of order $g^2$ and $p^2$. It has a solution of the special form

$$g = p\beta(\mathbf{r})c - \ln A, \quad A = \text{const}, \tag{11}$$

called the complete integral. Here, the function $\beta$ obeys the equation

$$\frac{1}{3k_T^2}(\nabla\beta)^2 = \beta^2\frac{(1/\beta) - T}{m_e c^2}. \tag{12}$$

Since the boundary condition (5) can be satisfied only for real $\beta > 0$, the more stringent requirement $1/\beta > T$ follows from (12). Different solutions of this equation differ, for example, by the distribution of $\beta$ over a closed surface. For the standard formulation of the problem where the emergent radiation spectrum observed far from sources needs to be calculated, a sphere with a large radius can be such a surface. In particular, for a spherically symmetric flow and a monotonically decreasing plasma temperature $T(r)$ with increasing distance $r$ to the center of symmetry, the solution $\beta(r)$ of Eq. (12) is uniquely determined by the condition $1/\beta \longrightarrow T$ as $r \longrightarrow 0$, yielding a specific value of $\beta(\infty)$ at large radii.

For $\zeta \ll 1$, the difference $1/\beta - T$ is small. Therefore, eliminating it in the expansion

$$\exp(-g) = A\exp\left(-\frac{pc}{T}\right)\left\{1 + \left(\frac{c}{T} - c\beta\right)p + \ldots\right\}$$

using (12), we can see it to match the last expansion of the previous section. Closer matching that could yield $q_1$ in (9) requires the next approximation. In this case, both $\ln A$ and the integration constants in (12) should be regarded as functions of $\ln p$, which let the substitution $n = \exp(-g)$ to balance the terms of order $g$ and $p$ in (1). However, this procedure is beyond the scope of this paper, which is limited to $\zeta \ll 1$.

### *3.4. Dependence of the Photon Source on the Bremsstrahlung Parameter*

To take into account the action of the bremsstrahlung source, it remains to consider the limit $q \ll 1$ in (9). This yields the expressions

$$j(q) = j_1 + O(q\zeta) - \xi\left[\frac{1}{2}\ln^2\frac{4\eta_E}{q} - \frac{\pi^2}{12} + O\left(q\ln^2\frac{4\eta_E}{q}\right) - f_1(A_0)\right] + \ldots, \tag{13}$$

$$n = \frac{A_0}{1 - A_0 + q} + \frac{A_1}{(1 - A_0 + q)^2} + O(q) + \frac{j(q) + O(\xi\ln(4\eta_E/q))}{3q^3} + \ldots. \tag{14}$$

Here,

$$f_1(A_0) = \int_0^\infty \frac{dq}{q}\frac{e^q-1}{A_0^{-1}e^q-1}e^{-q/2}K_0\left(\frac{q}{2}\right)$$

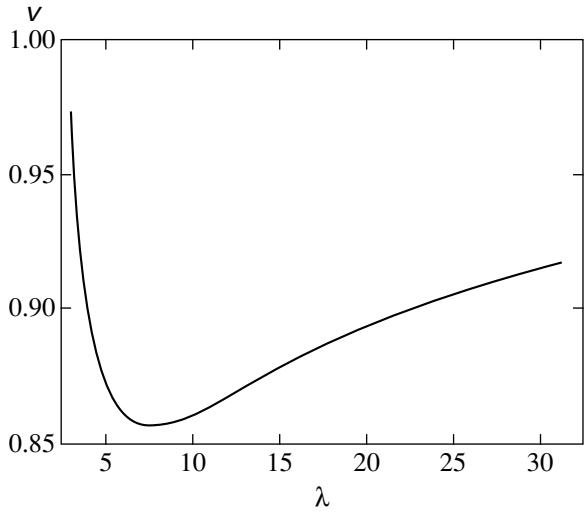
$$= \begin{cases} \frac{1}{2}[\ln^2(3+2\sqrt{2})]A_0[1+O(A_0)], & A_0 \ll 1 \\ \frac{1}{2}\ln^2\frac{4\eta_E}{1-A_0}+\frac{\pi^2}{12}+O(1-A_0), & 1-A_0 \ll 1, \end{cases}$$

$$j_1 = \int_0^\infty dq q^2 \hat{\zeta}\frac{1}{A_0^{-1}e^q-1}$$

is the sought effective source of the photon number. The corresponding dimensional quantity is

$$J_r = \frac{ck_TT}{m_ec^2}\frac{8\pi T^3}{(hc)^3}j_1$$
$$= 2\int_0^\infty \frac{4\pi p^2 dp}{h^3}c\hat{L}\frac{1}{A_0^{-1}\exp(pc/T)-1}. \quad (15)$$

Expression (14) shows that expansion (9) can again become invalid, but now when $q \longrightarrow 0$. To find $j_1$, we must reconcile (13) and (14) with the asymptotic behavior of this solution of (6) and (7) that would satisfy the equilibrium condition $n \longrightarrow n_P$ as $q \longrightarrow 0$ for $q \gg \xi^{1/3}$. In principle, this solution can be found numerically. The satisfaction of the above three conditions for the two constants $A_0$ and $j_1$ would entail an additional functional dependence of $j_1$ on $A_0$. However, the role of this dependence is determined by the relation between the scales of variations in $n$ with $q$. Expansion (14) indicates that the scales $1-A_0$ and $[j_1(1-A_0)/A_0]^{1/3}$ are possible. The estimates in Section 3.1 revealed the scale $\xi^{1/2}$ and the relation $j_1 \sim \xi$. Meanwhile, the relations $\xi^{1/2} \gtrsim [\xi(1-A_0)/A_0]^{1/3}$ or $\xi^{1/2} \gtrsim (1-A_0)$ strongly suggest that $1-A_0 \ll 1$, which corresponds to small deviations from the Planck occupation number for all $q$. In this case, however, it makes no sense to calculate the photon source, because all radiation parameters prove to be well known. If, alternatively, $\xi^{1/2} \ll [\xi(1-A_0)/A_0]^{1/3}$ (which also include nearly Planck values of the occupation number), then the significant integration range is $\xi^{1/2} \lesssim q \ll [\xi(1-A_0)/A_0]^{1/3}$. We may then substitute $j(q) = \xi u$, $q = \xi^{1/2}x$, and $n(q) = \xi^{-1/2}w$ into Eqs. (6) and (7) and pass to the lowest order of the expansion in terms of $\xi$ and $\zeta$. Thus, we obtain the system of equations




Residual dependence of the effective photon source on the bremsstrahlung parameter λ.

$$-x^4\left(\frac{dw}{dx}+w^2\right) = u(x),$$

$$\frac{du}{dx} = \left(\frac{1}{x}-w\right)\ln\frac{4\eta_E}{x\sqrt{\xi}}$$

for the final integration under the conditions

$$u(x) \longrightarrow f+\frac{\pi^2}{12}-\frac{1}{2}\ln^2\frac{4\eta_E}{x\sqrt{\xi}}, \quad w(x) \longrightarrow w_0$$

at $x \gg 1$ and

$$w \longrightarrow \frac{1}{x}$$

at $x \ll 1$.

In this case, the influence of

$$w_0 = \xi^{1/2}\frac{A_0}{1-A_0} \ll 1$$

on the values of $f$ and, hence, on the sought source $j_1 = \xi[f - f_1(A_0)]$ proves to be an insignificant small correction to the case of $w_0 = 0$.

The results can be represented as

$$f = 0.5\lambda^2 v(\lambda, w_0) - \frac{\pi^2}{12}, \quad \lambda \equiv \ln\frac{4\eta_E}{\xi^{1/2}},$$

where $\nu(\lambda, w_0) \approx \nu(\lambda, 0)$ is a comparatively slow function. It is shown in the figure. For comparison, it is worth noting that the calculation in [1] corresponds $f = 0.5\lambda^2$.

## 4. RADIATION HYDRODYNAMICS UNDER STRONG COMPTON ENERGY EXCHANGE

Problems associated with outflows of interstellar matter from compact objects and its accretion onto them figure prominently in astrophysics. The required accuracy of calculating the radiation that accompanies these processes and its interaction with matter can often be achieved by using only integrated radiation parameters: energy density $E$, pressure $P$, energy flux density $\mathbf{F}$, and Compton temperature $T_r$. In particular, in an optically thin case, $\tau \ll 1$, calculating the radiation spectrum reduces to adding up the actions of the sources. In this case, $\mathbf{F}$, $E$, $P$, and $ET_r$ can be determined in the same way and independently of spectral information. In general, based on the conservation law for the energy–momentum tensor, a system of equations similar to the hydrodynamic one can be formulated for these quantities (see, e.g., [10]). Under the domination of multiple scattering, $\tau \gg 1$, the radiation has a nearly isotropic angular distribution that yields the relation $P = E/3$. For $Y_C \ll 1$, when the inverse Compton effect may be ignored, this relation closes the system of equations for $\mathbf{F}$, $E$, and $P$. Finding other radiation parameters, in particular, the Compton temperature $T_r$, may still require turning to the kinetic equation [11].

The Compton energy exchange between plasma and radiation, which is commonly expressed in terms of $T_r$, can no longer be ignored for $Y_C \gtrsim 1$. The results of the previous section enable us to formulate the radiation hydrodynamics equations in the limit $Y_C \gg 1$. In particular, the photon occupation number at energies $pc \sim T$ was found in this section to be close to the Bose–Einstein distribution

$$n_0(p) = \left(A_0^{-1}\exp\frac{pc}{T} - 1\right)^{-1}.$$

Moreover, in the lowest order of the expansion in terms of $1/Y_C \ll 1$, this energy range makes a major contribution to the radiation energy density:

$$E \approx 2\int_0^\infty n_0(p)\,pc\,\frac{4\pi p^2 dp}{h^3} = \frac{8\pi T^4}{(hc)^3} \times \begin{cases} 6A_0[1 + O(A_0)], & A_0 \ll 1 \\ \dfrac{\pi^4}{15} + O(1 - A_0), & 1 - A_0 \ll 1. \end{cases} \quad (16)$$

The Comptonized photon number density may be more convenient to use in place of the quantity $A_0$, which yields the chemical potential $\mu_r = T\ln A_0$:

$$N \equiv 2\int_0^\infty n_0(p)\frac{4\pi p^2 dp}{h^3} = N_0 \begin{cases} A_0[1 + O(A_0)], & A_0 \ll 1 \\ 1.202\ldots + O(1 - A_0), & 1 - A_0 \ll 1, \end{cases} \quad (17)$$

where

$$N_0 \equiv N_0(T) = \frac{16\pi T^3}{(hc)^3}.$$

In this case,

$$E = T\begin{cases} 3N\left[1 + O\left(\dfrac{N}{N_0}\right)\right], & N \ll N_0 \\ \dfrac{\pi^4}{30}N_0 + O\left(1 - \dfrac{N}{N_0^*}\right), & 1 - \dfrac{N}{N_0^*} \ll 1 \end{cases}$$

and

$$\mu_r = T\begin{cases} \ln\dfrac{N}{N_0} + O\left(\dfrac{N}{N_0}\right), & N \ll N_0 \\ -0.731\ldots\left(1 - \dfrac{N}{N_0^*}\right), & 1 - \dfrac{N}{N_0^*} \ll 1, \end{cases}$$

where $N_0^* = 1.202\ldots N_0$.

The integration in (15) yields the photon production equation

$$\frac{\partial N}{\partial t} + \nabla\cdot\left\{\mathbf{u}N - \frac{c}{3k_T}\nabla N\right\} = J_r. \quad (18)$$

For $N \ll N_0$, the absorption term in $J_r$ may be disregarded, so the photon production rate takes the form

$$J_r \approx \frac{4}{\pi^{3/2}}\left(\sum_a z_a^2 n_a\right) n_e \sigma_T c\left(\frac{I_\mathrm{H}}{T}\right)^{1/2} f.$$

Integrating the original kinetic equation (1) over $p^3dp$ from 0 to $\infty$, we can make sure that the energy density (16) also obeys the equation for energy exchange with plasma electrons:

$$\frac{\partial E}{\partial t} + \nabla\cdot\left\{\mathbf{u}E - \frac{c}{3k_T}\nabla E\right\} + \frac{E}{3}(\nabla\cdot\mathbf{u}) = Q_{re}. \quad (19)$$

In accordance with the separation of the plasma–radiation interaction processes in the collision integral, the rate of energy exchange in this equation can also be separated into the bremsstrahlung and Compton ones:

$$Q_{re} = Q_{re}^B + Q_{re}^C.$$

Here, the Compton energy exchange rate can be expressed in terms of the Compton radiation temperature [12]:

$$Q_{re}^C = n_e \sigma_T c E \frac{4(T - T_r)}{m_e c^2}.$$

Since the condition $Y_C \gg 1$ implies that the Compton energy exchange dominates over the spatial transport of the energy density $E$, Eq. (19) in the approximation under consideration yields only the value of $T_r$ or, more precisely, its small deviation from $T$.

Equation (19) also implicitly contains the irreversible energy release. To find it requires introducing the radiation entropy density

$$S_r \equiv 2\int_0^\infty [(1+n_0)\ln(1+n_0) - n_0 \ln n_0] \frac{4\pi p^2 dp}{h^3}$$

$$= T\begin{cases} N\left[4 - \ln\frac{N}{N_0} + O\left(\frac{N}{N_0}\right)\right], & N \ll N_0 \\ \frac{2\pi^4}{45} N_0 \left[1 + O\left(1 - \frac{N}{N_0^*}\right)\right], & 1 - \frac{N}{N_0^*} \ll 1. \end{cases}$$

This quantity obeys the relations

$$T dS_r = dE - \mu_r dN, \quad \mu_r N = \frac{4}{3}E - TS_r.$$

Using them, we can rewrite (19) as an equation for entropy:

$$T\left\{\frac{\partial S_r}{\partial t} + \nabla\left[\left(\mathbf{u} - \frac{c}{3k_T}\nabla\right)S_r\right]\right\} = Q_{re} + Q^C - \mu_r J_r, \quad (20)$$

where

$$Q^C = \frac{c}{3k_T}\left[\frac{1}{T}(\nabla T)(\nabla E) + T\left(\nabla\frac{\mu_r}{T}\right)(\nabla N)\right].$$

Considering the differential

$$d\left(S + \frac{E}{T}\right) = \frac{E}{T^2}dT - \frac{\mu_r}{T}dN,$$

we can establish the Maxwell identity

$$\frac{1}{T^2}\left(\frac{\partial E}{\partial N}\right)_T = -\left(\frac{\partial(\mu_r/T)}{\partial T}\right)_N.$$

Therefore,

$$Q^C = \frac{c}{3k_T}\left[\frac{1}{T}\left(\frac{\partial E}{\partial T}\right)_N (\nabla T)^2 + T\left(\frac{\partial(\mu_r/T)}{\partial N}\right)_T (\nabla N)^2\right]$$

is positive and describes the irreversible energy dissipation.

The equation

$$T\left\{\frac{\partial}{\partial t}\sum_a S_a + \nabla\left[\mathbf{u}\sum_a S_a\right]\right\} = Q_{er} - \sum_a \mu_a J_a \quad (21)$$

for the entropy density $S_a(n_a, T)$ of all plasma components with the particle number density $n_a$, chemical potential $\mu_a(n_a, T)$, and particle production rate $J_a$ also holds for plasma. Adding Eqs. (20) and (21) and taking into account $Q_{er} + Q_{re} = 0$, we can obtain the equation

$$T\left\{\frac{\partial S}{\partial t} + \nabla\left[\mathbf{u}S - \frac{c}{3k_T}\nabla S_r\right]\right\} = Q^C - \mu_r J_r - \sum_a \mu_a J_a \quad (22)$$

for the production of the total entropy with the density

$$S = S_r + \sum_a S_a.$$

The radiation parameters and the plasma temperatures under strong Comptonization must be found simultaneously from Eqs. (18) and (22).

To conclude, it is worth noting that finding the remaining plasma parameters used in this section requires including the momentum balance equation for the plasma velocity **u** and the continuity (particle production/loss) equations for the number density $n_a$ of each component. These equations are not written out here, because they are unrelated to the Compton energy exchange.

## 5. CONCLUSIONS

We have derived the equations that describe the spatial transport of the photon number and determine the energy output in a region of strong Compton exchange with an inhomogeneous moving plasma. Together with the hydrodynamic equations, they determine the plasma and radiation parameters in a self-consistent way.

## ACKNOWLEDGMENTS

I thank S.S. Gershtein for his attention to the paper and Yu.M. Belousov for his stimulating support. This work was supported by the Ministry of Education of Russia and the CRDF grant no. REC-011.

## REFERENCES

1. A. S. Kompaneets, Zh. Éksp. Teor. Fiz. **31**, 876 (1956) [Sov. Phys. JETP **4**, 730 (1956)].
2. L. A. Pozdnyakov, I. M. Sobol', and R. A. Syunyaev, Itogi Nauki Tekh., Ser.: Astron., 100 (1982).
3. D. Psaltis and F. K. Lamb, Astrophys. J. **488**, 881 (1997).
4. M. Colpi, Astrophys. J. **326**, 223 (1988).
5. P. A. Becker, Mon. Not. R. Astron. Soc. **343**, 215 (2004).
6. V. S. Imshennik and Yu. I. Morozov, *Radiative Relativistic Gas Dynamics of High-Temperature Phenomena* (Atomizdat, Moscow, 1981) [in Russian].
7. G. B. Rybicki and A. P. Lightman, *Radiative Processes in Astrophysics* (Wiley, New York, 1979).
8. R. D. Blandford and D. G. Payne, Mon. Not. R. Astron. Soc. **194**, 1033 (1981).
9. J. Green, Astrophys. J. **130**, 693 (1959).
10. D. Michalas and B. Michalas, *Foundations of Radiation Hydrodynamics* (Oxford Univ. Press, Oxford, 1984).
11. S. A. Vagner, Zh. Éksp. Teor. Fiz. **107**, 151 (1995) [JETP **80**, 81 (1995)].
12. E. V. Levich and R. A. Syunyaev, Astron. Zh. **48**, 461 (1971) [Sov. Astron. **15**, 363 (1971)].

*Translated by V. Astakhov*

*Journal of Experimental and Theoretical Physics, Vol. 99, No. 4, 2004, pp. 741–748.*
*Translated from Zhurnal Éksperimental'noĭ i Teoreticheskoĭ Fiziki, Vol. 126, No. 4, 2004, pp. 857–865.*


**SOLIDS**
**Structure**

# Interaction of Nitrogen Dioxide Molecules with the Surface of Silicon Nanocrystals in Porous Silicon Layers

**E. A. Konstantinova*, L. A. Osminkina, K. S. Sharov, E. V. Kurepina, P. K. Kashkarov, and V. Yu. Timoshenko**

*Moscow State University, Vorob'evy gory, Moscow, 119992 Russia*
**e-mail: liz@vega.phys.msu.ru*
Received March 3, 2004

**Abstract**—The methods of infrared absorption spectroscopy and electron paramagnetic resonance are used for studying the effect of adsorption of $NO_2$ molecules, which are strong acceptors of electrons, on the electronic and optical properties of silicon nanocrystals in mesoporous silicon layers. It is found that the concentration of free charge carriers (holes) in silicon nanocrystals, which exhibits a nonmonotonic dependence on the $NO_2$ pressure, sharply increases in the presence of these molecules. At the same time, a monotonic increase in the concentration of dangling silicon bonds ($P_{b1}$ centers) is observed. A microscopic model proposed for explaining this effect presumes the formation of donor-acceptor pairs $P_{b1}^{+}$-$(NO_2)^{-}$ on the surface of nanocrystals, which ensure an increase in the hole concentration in nanocrystals, as well as $P_{b1}$ centers, which are hole-trapping centers. The proposed model successfully explains a substantial increase in photoconductivity (by two or three orders of magnitude) in the layers of porous silicon in the presence of $NO_2$ molecules; the increment in the concentration of free charge carriers is detected within an order of magnitude of this quantity. The results can be used in designing electronic and luminescence devices based on silicon nanocrystals. © *2004 MAIK "Nauka/Interperiodica".*

## 1. INTRODUCTION

Porous silicon (PS) obtained by electrochemical etching of monocrystalline silicon (*c*-Si) in a solution of hydrofluoric acid remains an object of intense studies (see, for example, reviews [1, 2]). This is due to the variety of physicochemical properties of this material, which is formed by an aggregate of silicon residues of nanometer size (nanocrystals). One of the interesting and most thoroughly studied properties of PS is its effective luminescence in the visible spectral range at room temperature. This property is successfully explained on the basis of the model of radiative recombination of excitons existing in small silicon nanocrystals (*nc*-Si), taking into account their interaction with the subsystem of free charge carriers capable of nonradiative recombination at surface defects [3]. At the same time, other model have been developed, in which the possibility of radiative recombination at surface centers is considered (see reviews [1, 2]). Both these approaches allow for the fact that silicon nanocrystals in PS have a huge surface (up to $10^3$ $m^2/g$ [2]), which is exposed to the action of various molecules from the surrounding medium. This leads to an exceptionally high sensitivity of physical properties of PS to the molecular coating of the nanocrystal surface [4]. Surface effects are apparently responsible for instability of luminescence characteristics of PS, which hampers the development of light emitting devices on its basis. On the other hand, the presence of the developed inner surface is a significant merit of PS in the study of fundamental aspects of adsorption processes in *nc*-Si.

From various adsorption effects modifying the properties of PS, the influence of various polar and nonpolar hydrocarbons has been studied best of all (see, for example, [3, 4] and the literature cited therein). The adsorption effect of active molecules forming negatively and positively charged complexes on the surface of a solid (i.e., exhibiting the properties of strong electron acceptors and donors, respectively) has been studied least. At the same time, analysis of fundamental mechanisms of interaction of such active molecules with a solid is very important in connection with the problem of *nc*-Si doping, which is required for optoelectronic applications of PS and other materials containing *nc*-Si [2]. In addition, such investigations are important for the development of a new generation of gaseous sensors. In this respect, nitrogen dioxide ($NO_2$) molecules are of considerable interest, since such molecules can be adsorbed at the surface of solids in the form of anion complexes [5]. Another important criterion dictating the choice of an adsorbate is the practical (ecological) aspect of the problem considered here. Indeed, the control of the $NO_2$ concentration in the atmosphere is an urgent problem since this substance is

very toxic and the search for new materials for developing high-sensitivity $NO_2$-based sensors is of considerable importance [6, 7].

It was established in [6, 8, 9] that the adsorption of $NO_2$ molecules leads to an increase in the concentration of free charge carriers (holes) and in the electrical conductivity in the layers of so-called mesoporous silicon (MPS). This type of PS is characterized by pores 4–100 nm in diameter and the characteristic size of *nc*-Si in it exceeds 5 nm [1]. The quantum size effect in *nc*-Si of this size is very weak and does not exceed the thermal energy at room temperature [1, 2]. Consequently, the electronic properties of MPS are close to those of *c*-Si except the contribution associated with the surface states of *nc*-Si.

Among the publications devoted to the effect of $NO_2$ adsorption on the electronic properties of PS, we can mention only a few studies in which the microscopic mechanism of the interaction of $NO_2$ molecules with the surface of *nc*-Si is considered [7, 9, 10]. In accordance with one of these models, $(NO_2)^-$ anion complexes formed on the surface of *nc*-Si activate boron atoms by their Coulomb field [9]. Thus, the hole concentration substantially increases as a result of $NO_2$ adsorption. However, using this approach, it is difficult to explain, for example, why the free carrier concentration varies within an order of magnitude in the course of adsorption of $NO_2$ molecules, while the conductivity increases by two or three orders of magnitude [6]. Analysis of the data presented in [6, 7] revealed that the PS samples subjected to the action of $NO_2$ molecules were initially characterized by different degrees of oxidation of the surface. The concentration of defects in the initial samples and those placed in the $NO_2$ atmosphere was not monitored. At the same time, defects in PS are known to be charge carrier trapping centers that can limit to a considerable extent their concentration and electric transport and, hence, can affect the electronic properties of *nc*-Si [11].

In this study, we employed IR and EPR spectroscopy to study the main physicochemical processes occurring as a result of adsorption of $NO_2$ molecules on the surface of *nc*-Si in MPS. This enabled us to monitor both the chemical composition of the surface coating of Si nanocrystals and their electronic properties (namely, the concentrations of equilibrium charge carriers and defects containing an unpaired electron spin).

## 2. EXPERIMENTAL TECHNIQUE

The PS samples were prepared using the standard technique of electrochemical etching of monocrystalline plates with the (100) surface orientation in a solution based on hydrofluoric acid [1]. The substrates were boron-doped *c*-Si plates with a resistivity of 10–20 mΩ cm (the equilibrium hole concentration was $\sim 5 \times 10^{18}$ cm$^{-3}$). The electrolyte was an aqueous solution of hydrofluoric acid mixed with ethyl alcohol in a proportion of HF(48%) : $C_2H_5OH$ = 1 : 1. The current density was 50 mA/cm$^2$. After the termination of pore formation, PS films were separated from the substrate by an abrupt short-term increase in the current density to 500 mA/cm$^2$. The thickness of the layers formed in this way was measured with the help of an optical microscope and was found to be 60 μm. The PS porosity was determined gravimetrically was found to be 50%.

Nitrogen dioxide was obtained from the chemical reaction

$$\mathrm{Cu(chip)} + 4\mathrm{HNO_3}$$

$$= 2\mathrm{NO_2(gas)} + \mathrm{Cu(NO_3)_2} + 2\mathrm{H_2O}.$$

Gaseous $NO_2$ was purified from water admixture by passing it through a flask containing dehydrator $P_2O_5$. Adsorption of $NO_2$ molecules was carried out from vacuum; all experiments were made in situ. We used vacuum equipment (membrane and turbomolecular pumps) manufactured at Varian.

The EPR measurements were made on the PS_100.X spectrometer (with a working frequency of 9.5 GHz and a sensitivity of $5 \times 10^{10}$ spin/G). To calculate the *g* factors and the defect concentrations, the following standards were used: MgO with $Mn^{++}$ ions and $CuCl_2 \cdot 2H_2O$, respectively. The IR transmission spectra of detached PS films were measured in a 400–6000 cm$^{-1}$ range at a resolution of 2 cm$^{-1}$ using a Perkin Elmer RX I instrument with the inverse Fourier transformation.

## 3. EXPERIMENTAL RESULTS

### *3.1. IR Absorption Spectroscopy*

Figure 1 shows the spectra of the absorption coefficient $\alpha(\nu)$ of PS layers in vacuum and after absorption of $NO_2$ molecules under various pressures. The $\alpha(\nu)$ dependences were calculated from the measured transmission spectra in accordance with the relation

$$\alpha(\nu) \approx -d^{-1}\ln[T(\nu)], \quad (1)$$

where $T(\nu)$ is the transmission coefficient and $d$ is the layer thickness.

In the IR spectrum of freshly prepared PS (Fig. 1a), the predominant absorption bands correspond to various local surface vibrations, namely, valence modes of Si–$H_x$ bonds ($x$ = 1, 2, 3) with frequencies of 2070–2170 cm$^{-1}$; scissors vibrations of Si–$H_2$ bonds at a frequency of ~910 cm$^{-1}$; and deformation vibrations of Si–$H_x$ bonds with the band maximum at a frequency of ~660 cm$^{-1}$ [2]. The presence of these bands indicates that the *nc*-Si surface is predominantly coated with hydrogen. These bands are observed against the back-

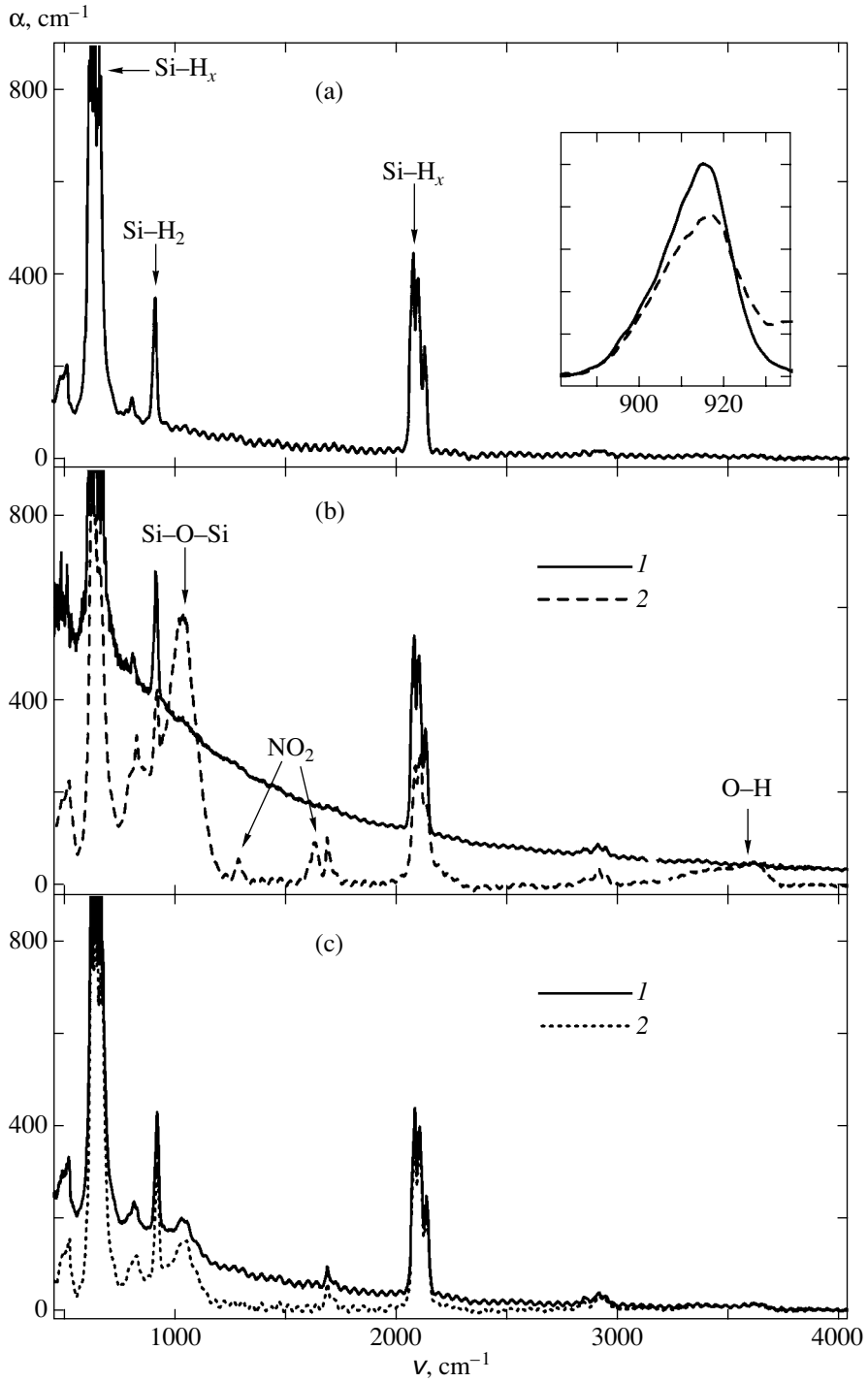



**Fig. 1.** Spectra of the absorption coefficient of PS layers in vacuum and under various $NO_2$ pressures: (a) in a vacuum of $10^{-6}$ Torr; (b) $P_{N_2O} = 0.1$ (*1*) and 10 Torr (*2*); (c) $P_{N_2O} = 1$ Torr (*1*) and subsequent evacuation to $10^{-6}$ Torr (*2*). The inset shows the absorption band at scissors vibrations of $Si–H_2$ bonds for freshly prepared samples (solid curve) and samples subjected to $NO_2$ adsorption at $P_{N_2O} = 10$ Torr (dashed curve).

ground of absorption at free charge carriers, which monotonically increases with decreasing wavenumber (increasing wavelength); this adsorption will be analyzed in detail at a later stage.

The adsorption of $NO_2$ molecules under a pressure $P_{NO_2} > 0.1$ Torr led to the emergence of new bands in the ranges 1050–1100 and 3100–3800 $cm^{-1}$, which are associated with adsorption under valence vibrations of

Si–O–Si and O–H bonds in water molecules adsorbed at the surface of *nc*-Si, respectively [2]. With increasing $NO_2$ pressure, the intensity of these lines increases, indicating stronger oxidation of the *nc*-Si surface. In addition to the above-mentioned bands, the PS samples subjected to the action of $NO_2$ molecules under high pressures exhibit adsorption at frequencies of 1290 $cm^{-1}$ (*trans*-dimers of covalent nitroso compounds $Si–N^+O^-=^-O^+N–Si$) and 1620–1680 $cm^{-1}$ (*cis*- and *trans*-isomers of covalent nitrites Si–O–N=O (spectrum *2* in Fig. 1b). "$N^+O^-$" indicates a semipolar bond, viz., a covalent pair of electrons connecting atoms that carry a total positive (N) and a total negative (O) charge [12].

Thus, the interaction of $NO_2$ molecules with the surface of silicon nanocrystals in PS layers exhibits all the features of chemisorption and can be represented by the following reaction (disregarding stoichiometric coefficients):

$$\begin{aligned} NO_2 + Si_3–Si–H &\longrightarrow Si–O–N=O \\ + Si–N^+O^-=^-O^+N–Si &+ SiO_y + H_2O, \end{aligned} \qquad (2)$$

where $1 \le y \le 2$ accounts for the possibility of partial of complete oxidation of surface silicon bonds. The second term on the left-hand side of relation (2) indicates the region of the *nc*-Si surface in the initial PS, which is characterized by a predominantly hydrogen coating (see above). It should be noted that the formation of $H_2O$ molecules, which is reflected on the right-hand side of relation (2), occurs due to partial consumption of the hydrogen coating of the *nc*-Si surface, which can be detected, in particular, from a decrease in adsorption at scissors vibrations of $Si–H_2$ bonds in a sample subjected to $NO_2$ adsorption (see the inset to Fig. 1a).

It is well known that the silicon surface is effectively oxidized during the interaction with nitric acid, the oxidation rate exceeding the analogous quantity for oxidation by gaseous $NO_2$ by several orders of magnitude [12]. Nitric acid can be formed as a reaction product during the interaction of $NO_2$ with water ($2NO_2 + H_2O \longrightarrow HNO_3 + HNO_2$). To clarify the role of these oxidation processes, we performed a control experiment in which the numbers of Si–O–Si groups formed as a result of adsorption of molecules of dehydrated $NO_2$ (used in these experiments) and $NO_2$ containing water vapor (in proportion 1 : 1) were compared using the IR spectra. It was found that no additional oxidation in the $NO_2$ atmosphere saturated with wave vapor takes place. Thus, under our experimental conditions, $HNO_3$ is apparently not formed and the oxidation of PS layers is due only to the interaction with $NO_2$ molecules. The experimental results match the theoretical estimates of the probability of nitric acid synthesis from $NO_2$ and $H_2O$, which are based on the value of the standard Gibbs energy of the corresponding chemical reaction [12].

It can be seen from Figs. 1a and 1b that the monotonic component of the $\alpha(\nu)$ spectrum associated with the absorption of IR radiation by free charge carriers increases in the $NO_2$ atmosphere. This indicates the increase in the hole concentration in *nc*-Si and can be formally written in the form

$$NO_2 + nc\text{-Si} \longrightarrow nc\text{-Si}\text{---}(NO_2)^- + h^+, \qquad (3)$$

where the first term on the right-hand side denotes a complex of the anion type on the *nc*-Si surface, the dashed line denotes a complex with charge transfer [5] (see Section 4 for details), and $h^+$ is a hole ejected to the bulk of a nanocrystal.

Adsorption at free charge carriers in MPS was observed earlier in [8, 9]. It was found that the type of adsorption for samples obtained under conditions similar to ours (the characteristic size of *nc*-Si exceeds 5 nm) is in good agreement with the predictions of the classical Drude model under weak scattering conditions. In this case, for frequencies on the order of or higher than the plasma frequency, the $\alpha(\nu)$ dependence can be expressed as [13]

$$\alpha(\nu) \propto N_p n^{-1} \tau \nu^{-2}, \qquad (4)$$

where $N_p$ is the concentration of free charge carriers (holes in our case), $n$ is the refractive index of the sample, and $\tau$ is the scattering time for the quasimomentum of holes. It should be noted that the concentration of free charge carriers in the bulk of *nc*-Si, which is defined as $N_p^* = N_p(1-p)^{-1}$, where $p$ is the PS porosity, is approximately twice as high as $N_p$.

To determine the value of $N_p$ from the $\alpha(\nu)$ spectra, we must know the value of $\tau$. It was shown in [8, 9] that the scattering time for holes in *nc*-Si with characteristic sizes far from the conditions of the quantum size effect are close to the values for the *c*-Si substrate used for obtaining PS. On account of this fact, we can derive from formula (4) the following relation for the concentration of free holes in PS:

$$N_p = N_{Si} \frac{n\alpha}{n_{Si}\alpha_{Si}}; \qquad (5)$$

here, $N_{Si}$ is the concentration of free holes in the silicon substrate and $n_{Si}$ and $\alpha_{Si}$ are the refractive index and the absorption coefficient of the substrate, respectively. The latter values were borrowed from [13].

Figure 2 shows the values of concentration of free holes in PS layers in vacuum and under various $NO_2$ pressures, calculated by formula (5). For freshly prepared samples, $N_p = 3.7 \times 10^{17}$ $cm^{-3}$, which is approximately an order of magnitude smaller than the doping level for the initial *c*-Si. Depletion of PS layers in equilibrium free charge carriers as compared to the mono-

crystalline substrate was studied in detail in [7, 11, 14], where several mechanisms responsible for this effect were proposed (namely, trapping at surface defects [9, 11] and deactivation of acceptor impurities as a result of their preferred localization at the *nc*-Si surface and an increase in the ionization energy [9, 14]).

The increase in the free hole concentration induced by $NO_2$ adsorption was a nonmonotonic function of the pressure produced by molecules (see Fig. 2). The maximal effect was observed for $P_{NO_2} = 10^{-1}$ Torr, and the value of $N_p$ was $2 \times 10^{18}$ cm$^{-3}$, which is an order of magnitude higher than for the initial freshly prepared sample. Moreover, the charge carrier concentration per unit volume of *nc*-Si, $N_p^* = 4 \times 10^{18}$ cm$^{-3}$, closely approaches the substrate doping level ($\sim 5 \times 10^{18}$ cm$^{-3}$). At first glance, the decrease in the value of $N_p$ upon a further increase in $P_{NO_2}$ was rather unexpected. Moreover, the value of $N_p$ for $P_{NO_2} = 10$ Torr was smaller by approximately a factor of 2.5 than for a freshly prepared sample. Another unexpected effect was that, after the removal of $NO_2$ molecules, the value of $N_p$ for samples in vacuum decreased below the initial level for freshly prepared layers (cf. curves in Fig. 1a and 1c).

The results indicate that, along with doping of *nc*-Si described by formula (3), another process occurs during the interaction of $NO_2$ molecules with the PS surface, which limits the increase in $N_p$ for small values of $P_{NO_2}$ and reduces the values of $N_p$ for large values of $P_{NO_2}$ as well as upon subsequent evacuation. Since the adsorption of $NO_2$ molecules is accompanied by oxidation of the PS surface (see reaction equation (2)), we can expect the emergence of surface defects [15]. We investigated such defects by the EPR method both in freshly prepared samples and in PS layers subjected to $NO_2$ adsorption in the entire range of pressures used in our experiments.

### *3.2. EPR Measurements*

The results of investigation by the EPR method are presented in Figs. 2 and 3. The EPR spectrum of the samples being studied (Fig. 3) exhibits the characteristic three-component structure. The measured orientation dependence of the EPR signal indicates the orthorhombic symmetry of the *g* tensor of spin centers being recorded. The values of the three main tensor components $g_1 = 2.0068 \pm 0.0005$, $g_2 = 2.0054 \pm 0.0005$, and $g_3 = 2.0027 \pm 0.0005$, as well the corresponding line width $\Delta H_1 = 7.6 \pm 0.8$ G, $\Delta H_2 = 3.2 \pm 0.8$ G, and $\Delta H_3 = 3.2 \pm 0.8$ G, indicate that the defects being detected are the so-called $P_{b1}$ centers, i.e., defects of a type of ruptured Si bond at the $Si/SiO_2$ interface [15, 16]. It should be noted that an ultimate atomic model of a $P_{b1}$ center

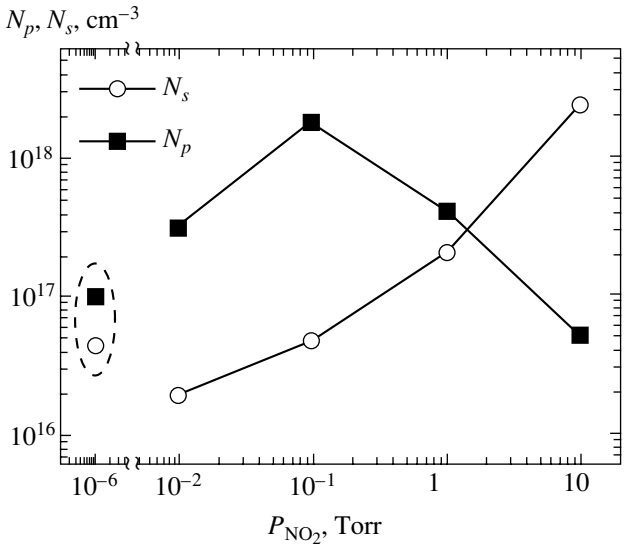



**Fig. 2.** Dependence of concentrations of free holes ($N_p$) and spin centers ($N_s$) in PS on the $NO_2$ gas pressure. The values for freshly prepared samples in a vacuum of $10^{-6}$ Torr (encircled by the dashed curve) are shown for comparison.

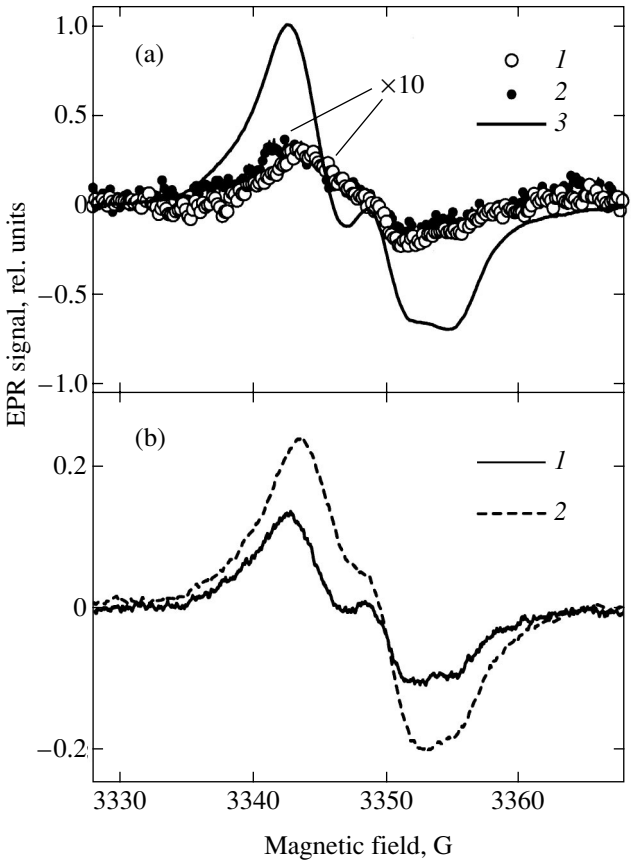



**Fig. 3.** EPR spectra of PS samples in vacuum and under various pressures of $NO_2$: (a) in a vacuum of $10^{-6}$ Torr (*1*), $P_{N_2O} = 0.1$ (*2*) and 10 Torr (*3*); (b) $P_{N_2O} = 1$ Torr (*1*) and subsequent evacuation to $10^{-6}$ Torr (*2*).

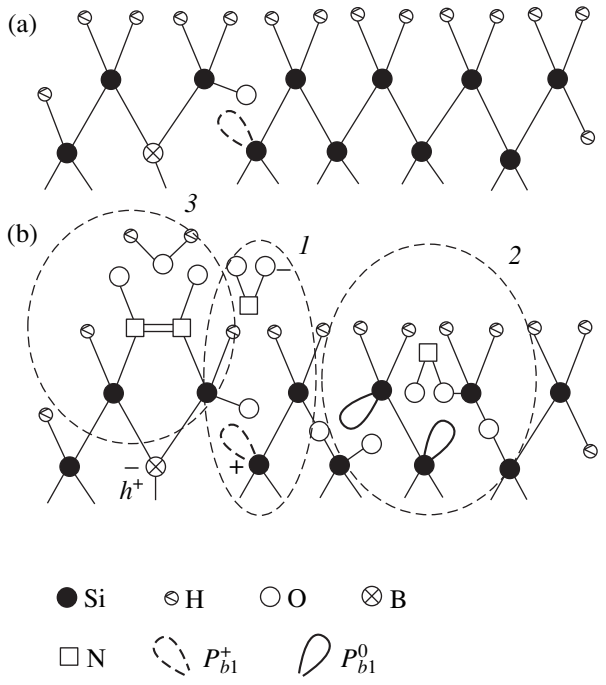



**Fig. 4.** Schematic representation of a surface region of the initial PS sample (a) before and (b) after the adsorption of $NO_2$ molecules and basic adsorption-induced processes: *1*—formation of donor-acceptor pairs $P^+_{b1}$–$(NO_2)^-$ and generation of a free hole $h^+$; *2*—formation of a covalent chemical bond (oxidation of silicon surface) and the emergence of $P_{b1}$ centers; *3*—formation of Si–$N^+O^-$=$^-O^+$N–Si *trans*-dimers and water molecules.

has not been developed as yet; it is known only that its nearest neighborhood contains one or two oxygen atoms [15]. Figure 4a schematically shows a region of the PS surface containing a $P_{b1}$ center. It should be noted that $P_{b1}$ centers exhibit amphoteric properties (i.e., such centers can trap both electrons and holes [17]).

The adsorption of $NO_2$ molecules under a pressure of $10^{-2}$ Torr led to a certain decrease in the EPR signal amplitude. An increase in the value of $P_{NO_2}$ up to $10^{-1}$ Torr virtually restored the signal level, while a further increase in $P_{NO_2}$ led to a substantial increase in the EPR signal amplitude (see Fig. 3a).

The dependence of concentration $N_s$ of $P_{b1}$ centers on $P_{NO_2}$, obtained from the measured EPR spectra, is shown in Fig. 2. The value of $N_s$ in vacuum (~$10^{-6}$ Torr) for a freshly prepared PS sample ($N_s = 4.4 \times 10^{16}$ cm$^{-3}$) is given for comparison. It can be seen that the number of defects is approximately halved (as compared to freshly prepared samples) after letting in $NO_2$ molecules under a pressure of $10^{-2}$ Torr and then monotonically increases in the pressure interval $10^{-1}$–10 Torr, attaining the maximal value $N_s = 2.3 \times 10^{18}$ cm$^{-3}$ at $P_{NO_2} = 10$ Torr. This value is equal in order of magnitude to the concentration of the boron impurity and free holes in the substrates used for obtaining PS. It should be noted that each next evacuation of PS after its being filled with $NO_2$ increased the intensity of the EPR signal by a factor of approximately 1.5–2 as compared to the value in the $NO_2$ atmosphere (see Fig. 3b). The variations of values of $N_s$ were reproduced in the next $NO_2$ admission-evacuation cycles. It should also be noted that holding of the samples at a given value of $P_{NO_2}$ (or in vacuum) for several hours hardly changed the value of $N_s$ at all.

## 4. DISCUSSION

Analysis of the data described in the previous section leads to the conclusion that the interaction of $NO_2$ molecules with the surface of silicon nanocrystals in PS is a complex adsorption process that modifies the molecular coating of nanocrystals and noticeably affects their electronic properties.

Let us consider the electronic properties of PS. We can expect that relatively small values of $N_s$ (~$5 \times 10^{16}$ cm$^{-3}$) and $N_p$ (~$10^{17}$ cm$^{-3}$) detected in a freshly prepared PS sample can be due to the fact that a large part of defects are transformed into the nonparamagnetic state after the capture of a hole: $P^0_{b1} + h^+ = P^+_{b1}$. In spite of the fact that most boron atoms are in the ionized state, the hole concentration is not high. In other words, the boron acceptor impurity in the initial PS is compensated by deeply lying donor states of $P^+_{b1}$ centers at the surface of *nc*-Si. This conclusion is in good agreement with our experimental data and with the results obtained in [9, 11].

Adsorption of $NO_2$ in PS for small values of $P_{NO_2}$ leads to a sharp increase in $N_p$, which can be formally explained by the formation of anion complexes $(NO_2)^-$ at the surface of *nc*-Si (see formula (3)). Such adsorption complexes with charge transfer can play the role of a doping impurity [5]. This must lead to the emergence of corresponding acceptor levels in the forbidden gap of *nc*-Si. These levels are apparently quite deep and cannot ensure by themselves the emergence of free holes in *nc*-Si at room temperature. However, adsorbed $NO_2$ molecules and $P_{b1}$ centers can experience a Coulomb interaction leading to the formation of donor-acceptor pairs $P^+_{b1}$–$(NO_2)^-$. The formation of such pairs increases the value of $N_p$ in view of passivation of $P_{b1}$ centers, which acquire a positive charge and cease to be the trapping centers for holes. Figure 4b illustrates schematically the doping mechanism for *nc*-Si considered here (process *1*). This model explains why the maximal value of $N_p$ detected after the adsorption of

$NO_2$ in our experiment, as well as in [6, 8, 9], does not exceed the doping level of the substrate.

The decrease in the charge carrier concentration in the interval of $P_{NO_2} = 10^{-1}$–10 Torr can be explained by trapping of a part of holes at new defects ($P_{b1}$) centers) formed as a result of chemisorption (see expression (2)). Indeed, it can be seen from Fig. 2 that a sharp increase in the value of $N_s$ takes place in the pressure range considered here. Oxidation of the PS surface in the atmosphere of $NO_2$ and the emergence of additional $P_{b1}$ centers is shown schematically in Fig. 4b (mechanism *2*). This process obviously limits the value of $N_p$ for small pressures $P_{NO_2}$ and causes its decrease at large pressures of the adsorbate. In addition, the $P_{b1}$ centers emerging as a result of chemisorption lead to an additional and irreversible decrease in the hole concentration upon the removal (desorption) of weakly bound $NO_2$ molecules from the surface of *nc*-Si (see Fig. 1c). The interaction of such weakly bound molecules with the *nc*-Si surface cannot be described by formula (2), but is apparently a weak form of chemisorption (with partial or complete charge transfer) [5].

The decrease in the value of $N_s$ for small values of $P_{NO_2} = 10^{-2}$ Torr is successfully explained by trapping of holes at defects (which thus become diamagnetic), i.e., by the formation of above-mentioned $P_{b1}^+$–$(NO_2)^-$ pairs. The increase in the number of spin centers being detected after desorption of weakly bound $NO_2$ molecules (see Fig. 3b) is apparently due to the transition of part of the defects from the $P_{b1}^+$ state to the $P_{b1}^0$ state.

Thus, we can now single out the main processes occurring at the surface of silicon nanocrystals during the adsorption of $NO_2$ molecules and affecting the electronic properties of PS. These processes are presented schematically in Fig. 4b, while Fig. 4a shows the same region of the PS surface prior to desorption.

1. The formation of donor–acceptor pairs $P_{b1}^+$–$(NO_2)^-$ at the surface of silicon nanocrystals, which is accompanied by an increase in the free hole concentration (process *1*).

2. The formation of a covalent chemical bond with an adsorbent in the form Si–O–Si and Si–O–N=O, which leads to oxidation of the silicon surface and the formation of $P_{b1}$ centers (process *2*).

3. The formation of Si–$N^+O^-$=$^-O^+$N–Si *trans*-dimers and water molecules (process *3*).

The hypothesis concerning the formation of donor–acceptor pairs $P_{b1}^+$–$(NO_2)^-$ at the *nc*-Si surface during the adsorption of $NO_2$ molecules makes it possible to explain the contradiction between the relatively small increase in the concentration of free charge carriers in PS (within an order of magnitude) and a considerable increase in the conductivity (by two or three orders of magnitude) [6]. Indeed, the total charge of the pairs formed in this case is equal to zero, while freshly prepared samples contain a large number of charged defects ($P_{b1}^+$) at which charge carriers undergo scattering [11]. Neutralization of charged spin centers in the course of adsorption of $NO_2$ molecules leads to a sharp increase in the mobility of holes due to a decrease in their scattering. Consequently, the substantial increase in the PS conductivity in the $NO_2$ atmosphere is mainly associated not with an increase in the concentration of holes, but rather with an increase in their mobility in ensembles of coupled *nc*-Si.

## 5. CONCLUSIONS

The interaction of $NO_2$ molecules with the *nc*-Si surface in PS layers is a complex physicochemical process including adsorption with charge transfer and the formation of donor-acceptor pairs $P_{b1}^+ - NO_2^-$, as well as chemisorption leading to the oxidation of the *nc*-Si surface. The emergence of adsorption-induced $P_{b1}^+ - NO_2^-$ complexes leads to an increase in the concentration of free charge carriers (holes) in *nc*-Si due to passivation of $P_{b1}$ centers. The oxidation of the PS surface occurring in this case is accompanied by the formation of new $P_{b1}$ centers, which limits to a considerable extent the increase in the hole concentration due to trapping of holes at the defects. Under $NO_2$ pressures on the order of 0.1 Torr and below, defect formation is virtually absent and the hole concentration in nanocrystals approaches the doping level of the initial monocrystalline substrate. Thus, the process of adsorption-aided doping of PS is most effective for low pressures of gaseous $NO_2$. The results can be used in designing sensors based on the $NO_2$ molecule.

## ACKNOWLEDGMENTS

One of the authors (E.A.K.) expresses thanks for financial support offered by the Presidential grant no. MK-2036.2003.02.

This study was carried out in the framework of the program of the Ministry of Industry and Science of the Russian Federation (contract no. 40.012.1.1.1153) on the equipment of Moscow State University.

## REFERENCES

1. A. G. Cullis, L. T. Canham, and P. D. J. Calcott, J. Appl. Phys. **82**, 909 (1997).
2. O. Bisi, S. Ossicini, and L. Pavesi, Surf. Sci. Rep. **38**, 1 (2000).
3. P. K. Kashkarov, B. V. Kamenev, E. A. Konstantinova, *et al.*, Usp. Fiz. Nauk **168**, 577 (1998) [Phys. Usp. **41**, 511 (1998)].

4. P. K. Kashkarov, E. A. Konstantinova, and V. Yu. Timoshenko, Fiz. Tekh. Poluprovodn. (St. Petersburg) **30**, 1479 (1996) [Semiconductors **30**, 778 (1996)]
5. F. F. Vol'kenshteĭn, *Electron Processes on the Semiconductor Surface under Chemisorption* (Nauka, Moscow, 1987) [in Russian].
6. L. Boarino, C. Baratto, F. Geobaldo, *et al.*, Mater. Sci. Eng. B **69**–**70**, 210 (2000).
7. L. Boarino, F. Geobaldo, S. Borini, *et al.*, Phys. Rev. B **64**, 205308 (2001).
8. V. Yu. Timoshenko, Th. Dittrich, and F. Koch, Phys. Status Solidi B **222**, R1 (2000).
9. V. Yu. Timoshenko, Th. Dittrich, V. Lysenko, *et al.*, Phys. Rev. B **64**, 085314 (2001).
10. F. Geobaldo, B. Onida, P. Rivolo, *et al.*, Chem. Commun., 2196 (2001).
11. V. Lehmann, F. Hofmann, F. Möller, *et al.*, Thin Solid Films **255**, 20 (1995).
12. *A Brief Handbook of Physicochemical Values* (Ivan Fedorov, St. Petersburg, 2002) [in Russian].
13. H. Hara and Y. Nishi, J. Phys. Soc. Jpn. **21**, 1222 (1966).
14. G. Polisski, D. Kovalev, G. G. Dollinger, *et al.*, Physica B (Amsterdam) **273**–**274**, 951 (1999).
15. E. H. Poindexter, P. J. Caplan, B. E. Deal, *et al.*, J. Appl. Phys. **52** (2), 879 (1981).
16. H. J. von Bardeleben, D. Stievenard, A. Grosman, *et al.*, Phys. Rev. B **47**, 10 899 (1993).
17. V. S. Vavilov, V. F. Kiselev, and B. N. Mukashev, *Defects in the Bulk and at the Surface of Silicon* (Nauka, Moscow, 1990) [in Russian].

*Translated by N. Wadhwa*

*Journal of Experimental and Theoretical Physics, Vol. 99, No. 4, 2004, pp. 749–755.*
*Translated from Zhurnal Éksperimental'noĭ i Teoreticheskoĭ Fiziki, Vol. 126, No. 4, 2004, pp. 866–873.*


**SOLIDS**
**Structure**

# Raman Scattering and Optical Normal Vibrations of the $Zn_{1-x}Cd_xSe$ Crystal Lattice

**E. A. Vinogradov[a], B. N. Mavrin[a], and L. K. Vodop'yanov[b]**

[a]*Institute of Spectroscopy, Russian Academy of Sciences, Troitsk, Moscow oblast, 142190 Russia*
[b]*Lebedev Institute of Physics, Russian Academy of Sciences, Leninskii pr. 53, Moscow, 119991 Russia*
*e-mail: mavrin@isan.troitsk.ru*

Received March 10, 2004

**Abstract**—The nonresonant spectra of the Raman scattering in $Zn_{1-x}Cd_xSe$ crystals ($0 \le x \le 1$) are investigated. The mode doublet was found to gradually shift with increasing $x$ from the longitudinal–transverse frequencies of ZnSe to the longitudinal–transverse frequencies of CdSe. Furthermore, an additional branch of weak modes was observed between the doublet components. The frequencies of this branch are shown to correspond to the frequencies of the cadmium impurity vibrations in the ZnSe lattice at $x = 0$ and the zinc impurity vibrations in the CdSe lattice at $x = 1$. The concentration dependences of the modes and atomic displacements are analyzed by using the isodisplacement model that includes the interaction between ZnSe- and CdSe-like vibrations. It is concluded that a one-mode behavior with the formation of an additional weak branch in the longitudinal–transverse splitting region is characteristic of this system. © *2004 MAIK "Nauka/Interperiodica".*

## 1. INTRODUCTION

The system of solid solutions $Zn_{1-x}Cd_xSe$ is of great interest in practical applications, in particular, in optoelectronics, owing to its unusual physical properties. Structures with quantum wells [1] and quantum dots [2] have been created and investigated on the basis of thin layers of these solid solutions. These are promising in creating sources of blue light. The system of solid solutions $Zn_{1-x}Cd_xSe$ is also interesting in that a structural phase transition from the cubic phase (sphalerite) to the hexagonal phase (wurtzite) occurs in it at $x = 0.3$–$0.4$ [3]. The Raman scattering spectroscopy (RSS) of solid solutions of tetrahedral type-II–VI semiconductors is an efficient method for investigating the dynamics of mixed crystal lattices.

The lattice dynamics for this system of solid solutions has been studied inadequately. The resonant RSS spectra were investigated in [4]. The Fermi resonance of a transverse optical phonon with two-phonon acoustic states was found in [5]. According to the current criterion for the rearrangement of the phonon spectrum as the composition of the solid solution changes [6], a two-mode type of optical phonon rearrangement must take place in the $Zn_{1-x}Cd_xSe$ system. However, having analyzed the RSS spectra, Brafman [7] concluded that this system has a one-mode behavior. Subsequently, while investigating the RSS of some of the solid solutions $Zn_{1-x}Cd_xSe$, Valakh *et al.* [8] obtained results consistent with the one-mode behavior of this system. Studies of epitaxial films of the solid solutions $Zn_{1-x}Cd_xSe$ by means of RSS [9] and infrared spectroscopy [10] also provided evidence for its one-mode behavior. The authors of [7–10] investigated only isolated ranges of alloy concentrations, which did not provide the full pattern of rearrangement of the phonon spectrum, while the accuracy and reliability of the measurements have not always corresponded to the present state of the art of the RSS technique.

In this paper, using a large set of compositions of high-quality $Zn_{1-x}Cd_xSe$ crystals and taking the spectra with an RSS Fourier spectrometer with infrared excitation of the RSS spectra, we analyzed the concentration rearrangement of the phonon spectrum in detail. In contrast to previous studies of the RSS spectra, we excited the spectra under nonresonant conditions. One of the complex problems in interpreting the spectra of solid solutions is to determine the positions of the local (impurity) modes. Here, we show that the frequencies of the impurity modes for the extreme compounds ($x \approx 0$ and $x \approx 1$) can be estimated by using the model of independent displacements of the ZnSe and CdSe atomic groups [11]. If the frequencies of the impurity modes are known, then the isodisplacement model can be used to analyze the concentration dependence of the vibration frequencies in a solid solution [12, 13]. We have experimentally and theoretically studied the concentration dependences of the lattice vibration frequencies, determined the type of mode rearrangement, calculated the composition dependences of the oscillator strengths, and found the amplitudes of the atomic displacement vectors for normal vibrations in the $Zn_{1-x}Cd_xSe$ system.

## 2. THE EXPERIMENTAL TECHNIQUE

The $Zn_{1-x}Cd_xSe$ crystals were grown by crystallization from the gas phase. The composition of the grown

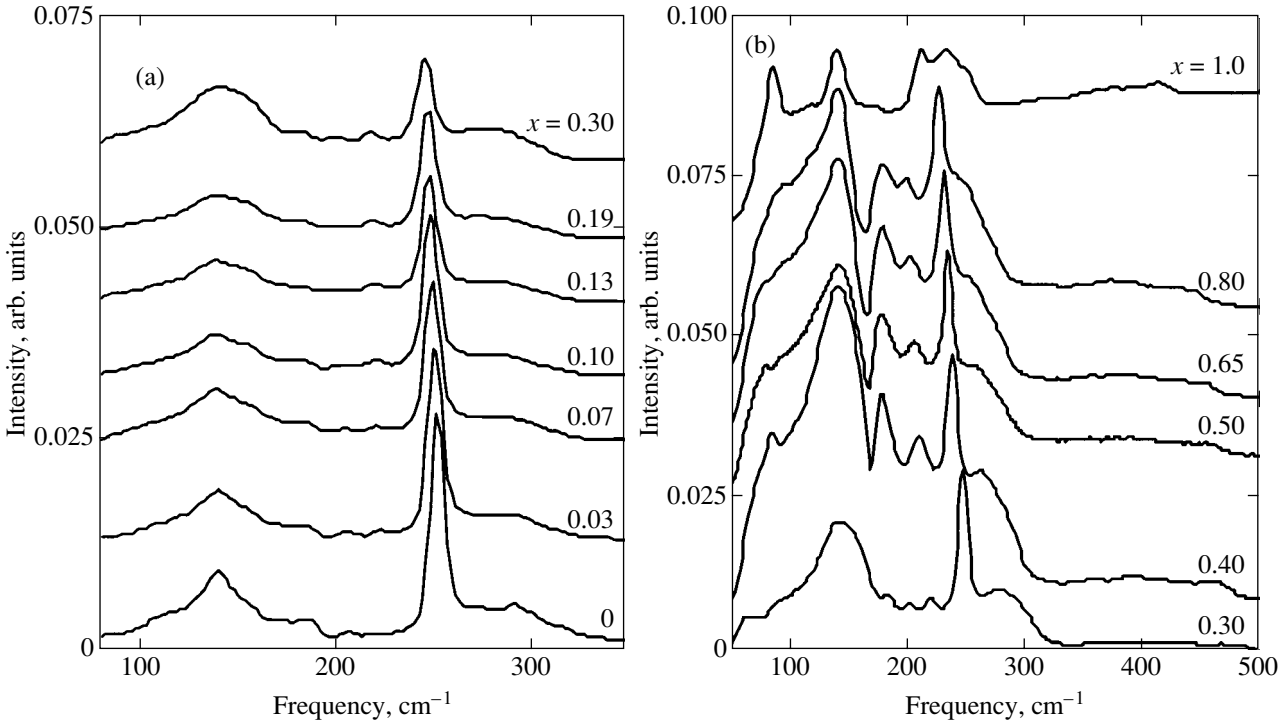



**Fig. 1.** RSS spectra of $Zn_{1-x}Cd_xSe$ crystals for (a) $x \le 0.30$ and (b) $x \ge 0.30$.

crystals was determined by measuring the lattice constants by the X-ray diffraction method and by analyzing the cathodoluminescence spectra.

The RSS spectra were excited by the 1.06-μm line of a continuous-wave YAG : Nd laser in a backscattering geometry at room temperature and recorded with an RFS-100 Bruker Fourier spectrometer. The drawbacks of the infrared excitation (the decrease in scattering cross section by almost a factor of 18 compared to the excitation by the 0.5145-μm line of an argon laser and the large detector noise in the near infrared) were offset by the following: the Fourier spectrometer had a large aperture (a gain by more than a factor of 10), the recording system was a multichannel one (a gain by a factor of $10^2$–$10^3$), and a liquid-nitrogen-cooled Ge detector was used. The exciting stray light in the spectrometer was suppressed by a notch filter that attenuated the signal at the frequency of the exciting line by more than a factor of $10^6$ and that allowed both the Stokes (from 50 to 3500 $cm^{-1}$) and anti-Stokes (from –1900 to –80 $cm^{-1}$) parts of the spectrum to be investigated.

## 3. EXPERIMENTAL RESULTS AND THEIR DISCUSSION

### *3.1. Results of Measurements*

The RSS spectra were measured for twelve compositions of the solid solutions $Zn_{1-x}Cd_xSe$ that covered the entire range from $x = 0$ to 1 almost uniformly (Fig. 1). All of the crystals under study were transparent to the exciting radiation. The RSS spectrum of the binary semiconductor compound ZnSe (sphalerite) with a cubic symmetry exhibits two bands of the longitudinal–transverse class-$F_2$ dipole vibration splitting: $\omega_{TO} = 207$ $cm^{-1}$ and $\omega_{LO} = 252.2$ $cm^{-1}$. At lower and higher frequencies, we see the second-order vibrational excitation bands that were studied in detail previously [14]. When a cadmium impurity concentration up to $x = 0.03$ is introduced, an additional band at $\omega_{ad} = 224$ $cm^{-1}$, whose nature is discussed below, appears in the region of the $\omega_{LO}$–$\omega_{TO}$ phonon splitting in ZnSe. As the cadmium atomic concentration increases further to $x = 0.3$, an appreciable decrease in the LO-phonon frequency, to 246.7 $cm^{-1}$, and a slight decrease in the frequencies $\omega_{TO}$ and $\omega_{ad}$ are observed. As regards the frequency $\omega_{TO}$ of the transverse optical mode, more complex processes related to the Fermi-resonance interaction of the TO mode with two-phonon acoustic states take place here [5]. The anharmonic interaction between them gives rise to an antiresonance dip in the RSS spectra near 160 $cm^{-1}$ at $x = 0.4$–$0.8$ (Fig. 1b). For $x > 0.3$, the position of the TO phonon is difficult to determine due to the Fermi resonance, and the rearrangement of the phonon spectrum in $Zn_{1-x}Cd_xSe$ as $x$ changes is described only by two vibration branches, $\omega_{LO}$ and $\omega_{ad}$.

The additional band $\omega_{ad}$ is of greatest interest. It shows up only in the spectra of solid solutions and disappears in the extreme binary compounds ZnSe and CdSe. The intensity of the RSS lines at the LO mode as

the composition of the solid solution changes remains almost constant, while the intensity of the additional mode $\omega_{ad}$ increases for the middle compositions and decreases at the edges. The attempts to explain the appearance of peaks in the RSS spectra at $\omega_{ad}$ in terms of two-phonon processes have failed, because combinations of two-phonon states that fall within this frequency range could not be found [14]. It is also difficult to explain its appearance in terms of the formation of macroscopic clusters caused by the deviation of the distribution of cadmium or zinc ions from the random one [15], because the studies of solid solutions by investigating the fine structure of the X-ray absorption spectra (EXAFS) and by other methods have not confirmed the existence of such clusters in alloys. When investigating the films of the cubic system $Zn_{1-x}Cd_xSe$, Alonso *et al.* [9] attributed the additional band $\omega_{ad}$ = 224 cm$^{-1}$ at $x = 0$ to the impurity vibration of cadmium atoms in ZnSe that smoothly transforms into the impurity vibrations of zinc atoms in CdSe at $x = 1$ as $x$ increases.[1] Since this attribution is not obvious, this assumption must be justified. Below, we provide estimates for the possible frequency of the impurity mode.

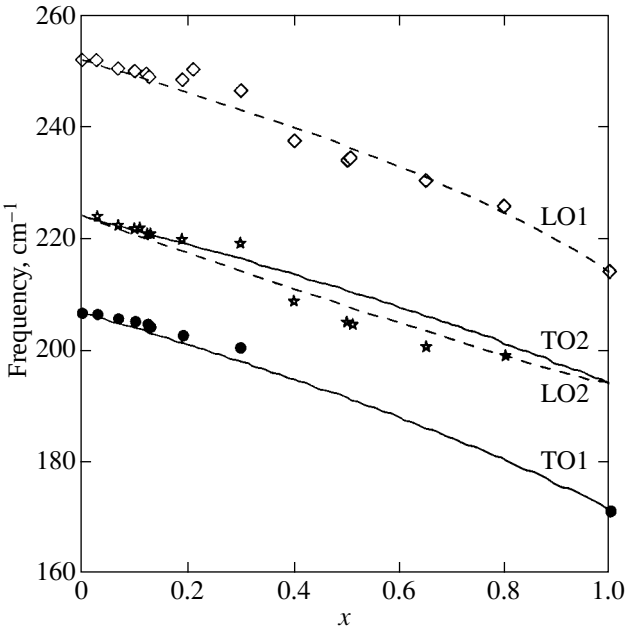



**Fig. 2.** Concentration dependences of the frequencies of optical atomic vibrations in the solid solution $Zn_{1-x}Cd_xSe$: the solid lines indicate the computed dependences for the transverse modes; the dashed lines indicate the computed dependences for the longitudinal modes; and the dots represent the experimental data.

### *3.2. The Frequencies of the Atomic Impurity Vibrations*

The frequencies of the impurity modes in the $Zn_{1-x}Cd_xSe$ system can be estimated if the equations of motion for the ZnSe (CdSe) groups are represented as [11]

$$\mu_i\ddot{\mathbf{u}}_i = -F_i(x)\mathbf{u}_i - \mu_i\Gamma_i\dot{\mathbf{u}}_i + e_i(x)\mathbf{E}_{\text{loc}}, \quad i = 1, 2, \quad (1)$$

where $\mu_i$, $\mathbf{u}_i$, $F_i(x)$, and $e_i(x)$ are, respectively, the reduced mass, the atomic displacement, the force constant, and the effective charge of the ZnSe- and CdSe-group ions. As follows from the analysis in [13], Eqs. (1) are acceptable only for the extreme compounds (near $x = 0$ and 1), where the interaction between ZnSe-like and CdSe-like vibrations may be disregarded. In this approximation, the dielectric constant $\varepsilon(\omega, x)$ of the $Zn_{1-x}Cd_xSe$ system has two poles at $x \approx 0$,

$$f_1^2 = \omega_{\text{TO},1}^2, \quad f_2^2 = \omega_{IM,2}^2, \quad (2)$$

and two poles at $x \approx 1$ [18],

$$f_1^2 = \omega_{IM,1}^2, \quad f_2^2 = \omega_{\text{TO},2}^2, \quad (3)$$

where

$$f_i^2 = \frac{1}{\mu}\left[F_i(x) - \frac{4\pi}{9V_i(x)}(\varepsilon_{\infty,i}(x) + 2)e_i^2(x)c_i(x)\right],$$

$$c_i(x) = \begin{cases} 1-x, & i = 1 \\ x, & i = 2, \end{cases}$$

$V_i(x)$ is the volume of the primitive cell, $\omega_{IM,1}$ and $\omega_{IM,2}$ are the impurity vibration frequencies of the zinc atoms in CdSe and the cadmium atoms in ZnSe, respectively, and $\varepsilon_\infty$ is the high-frequency dielectric constant.

Using the macroscopic parameters of the CdSe crystal ($\mu_i$ = 46.38 amu, $\varepsilon_{\infty,i}$ = 6.2, $e_i = 0.815e$, and $\omega_{TO,2}$ = 171 cm$^{-1}$), we obtain $F_2(1) = 1.1 \times 10^4$ g s$^{-2}$. According to (2), $F_2(0)$ must be known to determine the impurity vibration frequency of the cadmium atoms. In the solid solution $Zn_{1-x}Cd_xSe$, the bond length $r_{\text{CdSe}}$ decreases as $x$ changes from 1 to 0 [19]. Assuming the dependence of the force constant on the bond length in type-II–VI semiconductors to be [20]

$$f = f_0\exp(-1.138r), \quad (4)$$

we obtain $F_2(0) = 1.3 \times 10^4$ g s$^{-2}$. We then find from Eq. (2) that $\omega_{IM,2}$ = 217 cm$^{-1}$. The calculated impurity vibration frequency of the cadmium atoms in ZnSe is close to the frequency $\omega_{ad}$ of the additional band

[1] Similar results with the detection of impurity vibrations in the frequency range of the longitudinal–transverse optical phonon splitting were obtained in the RSS and infrared reflectivity spectra of $Cd_xZn_{1-x}Te$ crystals at small $x$ [16] and in the infrared reflectivity spectra of $Zn_{1-x}Cd_xS$ crystals at all $x \neq 0$ and $x \neq 1$ [17].

detected at $x = 0.03$. Given that Eq. (4) is approximate, the agreement between the calculations and the experimental data may be considered to be excellent, and the additional band $\omega_{ad}$ may be assumed to pertain to the impurity vibration of the cadmium atoms in ZnSe.

The impurity vibration frequency of the Zn atoms in CdSe can be estimated in the same way.

### *3.3. The Concentration Dependence of the Frequencies and Oscillator Strengths*

The normal vibrations in the solid solution $Zn_{1-x}Cd_xSe$ can be found in the isodisplacement model [12, 13]. The fundamental assumption of this model is that the anions and cations of the ZnSe (CdSe) groups vibrate in phase with the same amplitude, and the forces statistically averaged over all neighbors act on each ion. The fact that only the macroscopic parameters of the extreme compounds are required to describe the normal modes of the solid solution is among the advantages of this model. However, the frequencies of the impurity modes in the extreme compounds must be known for a proper description.

To determine the concentration dependences of the mode parameters in the solid solution, we used the potential energy density function for the TO and LO modes [13]. This function allows not only the mode dispersion, but also the atomic displacements for normal vibrations to be determined. In particular, in the system of $(\mathbf{w}_1, \mathbf{w}_2)$ coordinates, where $\mathbf{w}_1$ are the displacements of the zinc and cadmium atoms relative to the selenium atoms, and $\mathbf{w}_2$ are the relative displacements of the zinc and cadmium atoms, the TO-mode frequencies are the solution of the equation

$$\det(D_w - \omega^2 I) = 0, \tag{5}$$

where

$$D_w = \begin{vmatrix} D_{11} & D_{12} \\ D_{21} & D_{22} \end{vmatrix},$$

$$D_{11} = \frac{(1-x)F_{\mathrm{ZnSe}} + xF_{\mathrm{CdSe}}}{\mu_{\mathrm{VCA}}},$$

$$D_{12} = D_{21} = \frac{\sqrt{\mu_{\mathrm{DSL}}}}{\mu_{\mathrm{VCA}}}\left(\frac{F_{\mathrm{CdSe}}}{m_{\mathrm{Cd}}} - \frac{F_{\mathrm{ZnSe}}}{m_{\mathrm{Zn}}}\right), \tag{6}$$

$$D_{22} = \left[(1-x)\omega^2_{IM,\,\mathrm{Cd}} + x\omega^2_{IM,\,\mathrm{Zn}} + \mu_{\mathrm{DSL}}\left(\frac{1}{m_{\mathrm{Cd}}} - \frac{1}{m_{\mathrm{Zn}}}\right)\left(\frac{F_{\mathrm{CdSe}}}{m_{\mathrm{Cd}}} - \frac{F_{\mathrm{ZnSe}}}{m_{\mathrm{Zn}}}\right)\right],$$

and $\mu_{\mathrm{VCA}}$ and $\mu_{\mathrm{DSL}}$ are functions of the atomic masses $m_i$ and the composition $x$ [13]. For the LO modes, the terms that include the effective ion charges in the ZnSe and CdSe crystals are added to the matrix elements [13]. As in [13], we assumed a linear composition dependence of the force constants. In this case, we took into account the fact that the force constants $F_{\mathrm{CdSe}}$ and $F_{\mathrm{ZnSe}}$ must decrease with increasing $x$, because the bond lengths $r$(CdSe) and $r$(ZnSe) increase with $x$ [19].

The solution of Eq. (5) yields two frequencies for the TO modes and two frequencies for the LO modes. Figure 2 shows the computed dependences of the frequencies of optical phonons in the $Zn_{1-x}Cd_xSe$ crystals ($0 \le x \le 1$) together with the experimental data obtained from the RSS spectra. In general, there is satisfactory agreement between the computational and experimental data. Note that the TO1 branch gradually deviates from the computed curve due to the resonant interaction of the TO1 phonon with the two-phonon acoustic states (Fig. 1b). The deviations of the experimental data from the computational data are also clearly seen in the region of the structural phase transition ($x = 0.3$–$0.4$).

The concentration dependence of the TO- and LO-branch frequencies is indicative of the one-mode behavior of the vibrations in the solid solution $Zn_{1-x}Cd_xSe$. Indeed, the TO1 (LO1) mode of the ZnSe crystal smoothly passes into the TO1 (LO1) mode of the CdSe crystal as $x$ increases from 0 to 1. An additional pair of frequencies, TO2 and LO2, with an inverse sequence of frequencies ($\omega_{\mathrm{TO2}} > \omega_{\mathrm{LO2}}$) is formed between the TO1- and LO1-mode frequencies. The frequency inversion occurs when one mode falls into the TO–LO splitting region of the other mode [21]. The TO2–LO2 mode pertains to the quasi-local vibrations, because it has the impurity modes as its limit at the boundaries ($x = 0$ and $x = 1$).

The derived dispersions of the TO and LO modes in the solid solution $Zn_{1-x}Cd_xSe$ (Fig. 2, the computational data) make it possible to calculate the concentration dependence of the oscillator strengths of both the fundamental and quasi-local vibrations. The concentration dependences of the oscillator strengths allow the intensities of the modes in the RSS spectra to be qualitatively explained. We see from Fig. 3 that the oscillator strengths of the quasi-local TO and LO modes are smaller than those of the fundamental modes, and these are at a maximum for the middle compositions. This may suggest that the intensities of the quasi-local TO2 (LO2) modes are weaker than those of the fundamental TO1 (LO1) modes, because the RSS cross section for the dipole modes is proportional to the oscillator strength [22, 23]. However, when comparing the relative intensities of the TO and LO modes, we should take into account the fact that the RSS tensor components for the TO and LO modes are different, because the tensor component of the TO mode is determined only by the deformation contribution, while the tensor component of the LO mode also depends on the electrooptic contribu-

tion. In particular, a detailed analysis indicates that the intensity ratio $I_{TO}/I_{LO}$ in a diatomic cubic crystal is [23]

$$\frac{I_{TO}}{I_{LO}} = \frac{\varepsilon_\infty \varepsilon_0}{[\varepsilon_\infty - \gamma(\varepsilon_0 - \varepsilon_\infty)]^2},$$

where $\gamma$ is the ratio of the deformation and electrooptic contributions, and $\varepsilon_0$ is the static dielectric constant. According to the measurements [24], the parameter $\gamma$ is negative for the ZnSe crystal, which is why $I_{TO} \ll I_{LO}$. If the same intensity ratio is assumed to be also preserved for the quasi-local modes, then the RSS spectra of solid solutions must be dominated by the longitudinal quasi-local vibration mode, while the transverse mode must have a very low intensity.

As we see from Figs. 2 and 3, no phase transition shows up in the computed dependences of the phonon frequencies and the oscillator strengths on the alloy composition. This may be primarily because the computed model is approximate. In particular, this model disregards the possible change in the number of oscillators during the phase transition, because two atoms in the primitive cell are contained in the zinc blende structure and four atoms are contained in the wurtzite structure. As a result, the number of optically active modes must change. However, since the anisotropy is small, these new modes do not show up in the experiment. It is well known that the CdSe structure may be considered as a slightly distorted zinc blende structure with a very small anisotropy. As a result of the small anisotropy, only two bands pertaining to the TO and LO modes are observed in the RSS spectrum of hexagonal CdSe, as in the spectrum of cubic ZnSe. Therefore, we performed our calculations by assuming that all compositions of the solid solution $Zn_{1-x}Cd_xSe$ are characterized by the same structure and only two optical frequencies. In addition, in our computed model, we used the assumption employed previously [12, 13] about a linear composition dependence of the force constants. It disregards the small changes in the parameters of the vibrational system of the crystal during the phase transition. However, these changes manifest themselves in the experiment (see Figs. 1 and 2). Therefore, it should be noted that, although the spectra for low CdSe concentrations ($x \le 0.3$) (Fig. 1a) significantly differ in appearance from the spectra for CdSe-enriched compositions (Fig. 1b), this difference is largely attributable not to the phase transition, but, most probably, to a manifestation of the Fermi-resonance interaction between the TO mode and the two-phonon acoustic states [5] that distorts significantly the RSS spectrum.

### *3.4. Atomic Displacements for Normal Vibrations in the Solid Solution*

If the equations of motion for atoms in the solid solution were set up not in ($\mathbf{w}_1$, $\mathbf{w}_2$) coordinates, but in normal $\mathbf{Q}_i$ coordinates, then the matrix $D_Q$ would be

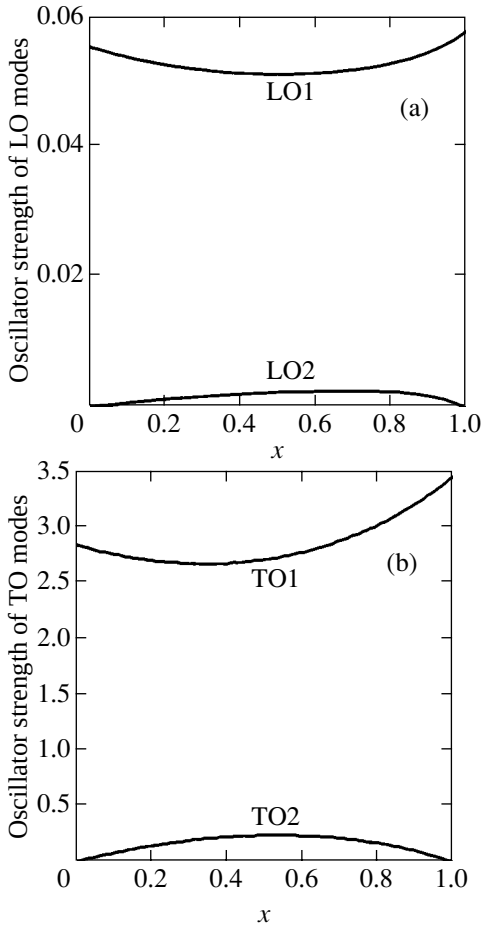



**Fig. 3.** Concentration dependences of the oscillator strengths for (a) the longitudinal and (b) transverse modes in the solid solution $Zn_{1-x}Cd_xSe$.

diagonal. The relationship between these matrices can be represented as

$$D_Q = RD_w,$$

i.e., the transformation matrix is

$$R = D_Q D_w^{-1}.$$

An expression for the diagonal elements of the matrix $D_Q$ for the TO modes follows from relation (6):

$$D_{Q,ii} = \frac{D_{11} + D_{22}}{2} \pm \sqrt{\left(\frac{D_{11} + D_{22}}{2}\right)^2 - (D_{12}^2 - D_{11}D_{22})}.$$

The normal coordinates are then defined by the transformation matrix $R$:

$$\mathbf{Q}_i = R_{ij}\mathbf{w}_j.$$

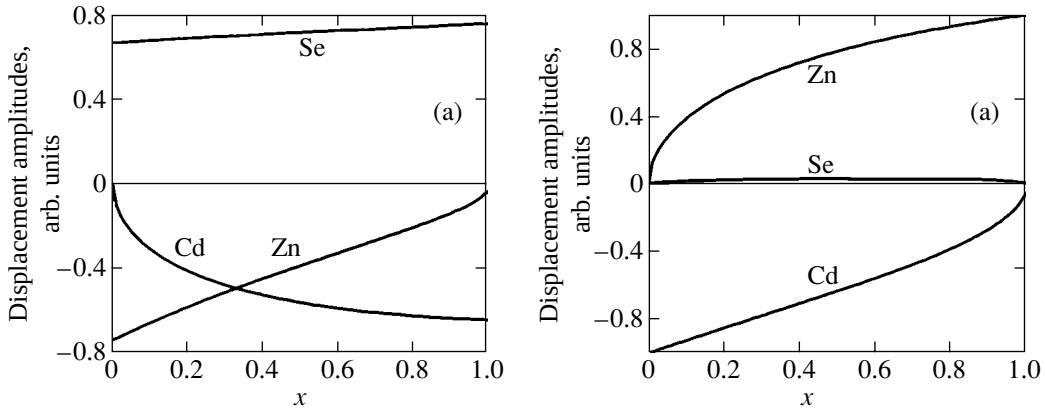



**Fig. 4.** Concentration dependences of the displacement amplitudes for atoms in the solid solution $Zn_{1-x}Cd_xSe$ during normal vibrations in the (a) fundamental and (b) local modes.

Let us now pass from the ($\mathbf{w}_1$, $\mathbf{w}_2$) coordinates to the ($\mathbf{u}_{Zn}$, $\mathbf{u}_{Cd}$, $\mathbf{u}_{Se}$) displacement coordinates, which are related by simple relations presented in [13]. We obtain

$$\mathbf{Q}_i = R'_{ik}\mathbf{u}_k, \quad (7)$$

where $R'_{ik}$ is a $2 \times 3$-rank matrix, and the index $k$ pertains to the Zn, Cd, and Se atoms.

Equation (7) allows the vibration amplitudes for each atom to be determined in normal coordinates $\mathbf{Q}_1$ (fundamental mode) and $\mathbf{Q}_2$ (quasi-local mode). Figure 4 shows the concentration dependences of the displacement amplitudes for each of the atom for the fundamental LO1 mode and the quasi-local LO2 mode.

At $x = 0$, the zinc and selenium atoms in the fundamental mode vibrate out of phase (Fig. 4a), which is well known for the optical mode in the binary ZnSe crystal. As $x$ increases, the cadmium atom that vibrates in phase with the zinc atom begins to be involved in the vibration; its vibrations amplitude increases, while the vibration amplitude of the zinc atoms decreases. The vibration amplitude of the selenium atom changes only slightly with concentration of the solid solution.

In the quasi-local mode, the cadmium and zinc atoms vibrate out of phase, while the displacement amplitudes of the selenium atoms in this mode are small and noticeable only for $x \neq 0$ and $x \neq 1$ (Fig. 4b). As $x$ increases, the displacement amplitude of the cadmium atoms decreases, but the displacement amplitude of the zinc atoms that vibrate out of phase with the cadmium atoms increases. The dipole moment that arises in these vibrations is determined by the difference between the effective charges of the zinc and cadmium ions. The effective ion charge (Szigeti charge) is $e_S = 0.75e$ and $0.81e$ for the ZnSe and CdSe crystals, respectively. This difference is small, and the dipole moment of the (Zn–Cd) vibrations is almost a factor of 10 smaller than that of the fundamental (Zn–Se or Cd–Se) vibrations; hence the small oscillator strength of the (Zn–Cd) vibrations and the low intensity of the RSS peaks at the impurity modes.

The composition dependence of the atomic displacements in normal vibrations for the TO modes is the same as that for the LO modes.

## 4. CONCLUSIONS

We have investigated the crystals of solid solutions $Zn_{1-x}Cd_xS$ over a wide range of concentrations ($0 \le x \le 1$) by means of Raman scattering spectroscopy for the nonresonant excitation of spectra by the 1.06-μm laser line. The frequencies of both the TO and LO phonons in the ZnSe crystal were found to smoothly decrease with increasing $x$. At $x \neq 0$ and $x \neq 1$, an additional branch of optical phonons within the longitudinal–transverse splitting region of the fundamental optical mode was observed in the RSS spectra. This additional mode was shown to pertain to the impurity vibration of cadmium atoms at $x \approx 0$ and to the impurity vibration of zinc atoms at $x \approx 1$. Using the model of isodisplacements, we computed the composition dependences of the mode dispersion and the oscillator strengths that agree with the experimental data. We conclude that the type of rearrangement of the phonon spectrum for the solid solution $Zn_{1-x}Cd_xS$ fits into the one-mode behavior with the formation of an additional branch of quasi-local modes. In [25], this type of rearrangement is called intermediate. We determined the displacement amplitudes for each of the atoms for normal vibrations in the fundamental and quasi-local modes over the entire range of concentrations.

## ACKNOWLEDGMENTS

We thank P.V. Shapkin who provided high-quality zinc–cadmium–selenium alloy crystals. This work was supported by the Russian Foundation for Basic Research (project no. 03-03-17110) and the "New

Materials an Structures" Program of the Russian Academy of Sciences.

## REFERENCES


1. P. Gingo, M. De Vittorio, R. Rinald, and R. Cingolani, Phys. Rev. B **54**, 16 934 (1996).
2. M. Strassberg, V. Kutzer, U. Pohl, *et al.*, Appl. Phys. Lett. **72**, 942 (1998).
3. K. V. Shalimova, A. F. Botev, V. A. Dmitriev, *et al.*, Kristallografiya **14**, 629 (1969) [Sov. Phys. Crystallogr. **14**, 531 (1969)].
4. M. Ya. Valakh, A. P. Litvinchuk, G. S. Pekar, and G. N. Polisskii, Phys. Status Solidi B **104**, 743 (1981).
5. M. Ya. Valakh, V. P. Lisitsa, G. S. Pekar, *et al.*, Phys. Status Solidi B **113**, 635 (1982).
6. J. Dow, W. Packard, H. Blackstead, and D. Jenkins, in *Dynamical Properties of Solids*, Ed. by G. K. Horton (North-Holland, Amsterdam, 1995), Vol. 7, p. 349.
7. O. Brafman, Solid State Commun. **11**, 447 (1974).
8. M. Ya. Valakh, V. P. Lisitsa, G. N. Polisskii, and V. I. Sidorenko, Phys. Lett. A **78A**, 115 (1980).
9. R. G. Alonso, E.-K. Suh, A. K. Ramdas, *et al.*, Phys. Rev. B **40**, 3720 (1989).
10. V. S. Vinogradov, L. K. Vodop'yanov, S. P. Kozyrev, and Yu. G. Sadof'ev, Fiz. Tverd. Tela (St. Petersburg) **43**, 1310 (2001) [Phys. Solid State **43**, 1365 (2001)].
11. A. Polian, R. Le Toullec, and M. Balkanski, Phys. Rev. B **13**, 3558 (1976).
12. D. L. Peterson, A. Petrou, W. Giriat, *et al.*, Phys. Rev. B **33**, 1160 (1986).
13. E. Yahne, Phys. Status Solidi B **74**, 275 (1976); Phys. Status Solidi B **75**, 221 (1976).
14. J. Ch. Irwin and J. La Combe, Can. J. Phys. **50**, 2596 (1972).
15. H. Verleur and A. Barker, Phys. Rev. **149**, 715 (1966).
16. L. K. Vodopyanov and E. A. Vinogradov, Cryst. Lattice Defects **5**, 125 (1974).
17. Yu. A. Mityagin, L. K. Vodop'yanov, and E. A. Vinogradov, Fiz. Tverd. Tela (Leningrad) **17**, 2054 (1975) [Sov. Phys. Solid State **17**, 1341 (1975)].
18. W. Lu, H. J. Ye, Z. Y. Yu, *et al.*, Phys. Status Solidi B **147**, 767 (1988).
19. M. H. Tsai, F. C. Peiris, S. Lee, and J. K. Furdyna, Phys. Rev. B **65**, 235202 (2002).
20. V. Kumar, J. Phys. Chem. Solids **61**, 91 (2000).
21. F. Gervais, Opt. Commun. **22**, 116 (1977).
22. M. V. Belousov, Fiz. Tverd. Tela (Leningrad) **15**, 1206 (1973) [Sov. Phys. Solid State **15**, 813 (1973)].
23. G. S. Barker and R. Loudon, Rev. Mod. Phys. **44**, 18 (1972).
24. J. F. Scott, T. C. Damen, and J. Shah, Opt. Commun. **3**, 384 (1971).
25. L. Genzel, T. P. Martin, and C. H. Perry, Phys. Status Solidi B **62**, 83 (1974).


*Translated by V. Astakhov*

*Journal of Experimental and Theoretical Physics, Vol. 99, No. 4, 2004, pp. 756–765.*
*Translated from Zhurnal Éksperimental'noĭ i Teoreticheskoĭ Fiziki, Vol. 126, No. 4, 2004, pp. 874–886.*


**SOLIDS**
**Structure**

# Magnetic Ordering in $La_{1-x}Sr_xMnO_{3-x/2}$ Anion-Deficient Manganites

**S. V. Trukhanov[a], M. V. Bushinsky[a], I. O. Troyanchuk[a], and H. Szymczak[b]**

[a]*Institute of Solid-State and Semiconductor Physics, National Academy of Sciences of Belarus, Minsk, 220072 Belarus*

*e-mail: truhanov@ifttp.bas-net.by*

[b]*Institute of Physics, Polish Academy of Sciences, Warsaw, 02-668 Poland*

Received March 16, 2004

**Abstract**—The structural, magnetic, and electrotransport properties of $La_{1-x}Sr_xMnO_{3-x/2}$ ($0 \le x \le 0.30$) manganites with perovskite structure are investigated experimentally as a function of oxygen deficiency. In the solid solutions $La_{1-x}Sr_xMnO_3$, a change in the type of symmetry of the unit cell is observed at $x = 0.125$. Samples with $x \le 0.125$ are characterized by an *O*'-orthorhombic unit cell, whereas samples with $x > 0.125$ are characterized by a rhombohedral unit cell. The structural properties of the anion-deficient solid solutions $La_{1-x}Sr_xMnO_{3-x/2}$ are analogous to those of the stoichiometric system. It is assumed that, as the oxygen content decreases, $La_{1-x}Sr_xMnO_{3-x/2}$ anion-deficient solid solutions experience a series of successive magnetic phase transformations in the ground state: from an *A*-type ($x = 0$) antiferromagnet to a cluster spin-glass-type inhomogeneous magnetic state ($0.175 < x \le 0.30$) through a two-phase (antiferromagnetic and ferromagnetic) state ($0 < x \le 0.175$). The anion-deficient solid solution with $x = 0.175$ has the maximal value of the ferromagnetic component. As the oxygen deficiency increases, the resistivity of $La_{1-x}Sr_xMnO_{3-x/2}$ samples first decreases (up to a value of $x = 0.175$), acquiring an activation character, and then increases (up to a value of $x = 0.30$). In this case, none of the anion-deficient solid solutions exhibits a metal–semiconductor transition in the whole range of concentrations considered. A peak of magnetoresistance at a temperature below the point of magnetic ordering is observed only in the sample with $x = 0.175$. The results of experiments carried out with a series of $La_{1-x}Sr_xMnO_{3-x/2}$ anion-deficient solid solutions are summarized in the concentration diagrams of the spontaneous magnetic moment and the critical temperature of magnetic phase transitions. Hypothetical magnetic phase states are pointed out. The experimental results obtained can be interpreted in terms of the phase-separation model and the competition between ferromagnetic and antiferromagnetic indirect superexchange interactions. It is assumed that $Mn^{3+}$–O–$Mn^{3+}$ indirect superexchange interactions in the orbitally disordered phase are positive in the case of octahedral coordination of manganese ions and are negative when the coordination of at least one $Mn^{3+}$ ion is pentahedral. © *2004 MAIK "Nauka/Interperiodica".*

## 1. INTRODUCTION

Substituted manganites of perovskite structure with the general chemical formula $Ln_{1-x}A_xMnO_3$ ($Ln^{3+}$ is a lanthanide ion and $A^{2+}$ is an alkaline-earth ion) have attracted the keen interest of specialists in the physics of magnetic phenomena for more than 50 years [1–12]. This interest has been aroused due to the large number of various phase states and phase transformations observed in this class of compounds [13–15].

Substituted manganites represent a good model object for studying the properties of strongly correlated electron systems and are of interest both from fundamental and practical points of view. These compounds attract interest due to the close connection between orbital, charge, spin, and lattice degrees of freedom and the associated diversity of physical properties. However, currently, the general state of the theory does not allow one to adequately describe the whole volume of experimental information available. Moreover, experimental information on the physical properties of magnets is being continuously supplemented, sometimes with contradictory data. The nature of magnetic and electric processes in manganites is not yet quite clear and has been a subject of extensive discussion [10–12]. Note that the physical properties of manganites are very sensitive to the method of production, the type of symmetry of a unit cell, size effects, the concentration of a substitute, the presence of ions of different valence at equivalent crystallographic sites, nonstoichiometry, etc. Moreover, as new experimental results are obtained, one feels more confident that even the fundamentals of the theory can be revised to take into account the effect of various defects and irregularities, including fluctuations. In spite of all the problems of fundamental character in understanding the physical properties of manganites, these compounds have already found wide application in the technology of magnetic recording as an active element of reading devices. Due to their high chemical stability, these compounds are

also employed as electrode materials in high-temperature fuel cells and in cathodes for $CO_2$ lasers [16].

Among the properties of substituted manganites, of special importance are magnetoresistive ones, which provide favorable conditions for the practical application of substituted manganites. There is a demand for materials with high magnetoresistance at room temperature in small magnetic fields [17]. However, the current interest in manganites is associated not only with possible engineering applications of the so-called colossal magnetoresistance but also with the fact that these compounds represent a suitable object for investigating the physics of strongly correlated systems. In particular, as we pointed out above, the close connection between orbital, charge, spin, and lattice degrees of freedom attracts great interest.

The original compound $LaMnO_3$ is an antiferromagnetic semiconductor with the magnetic structure of *A*-type. The weak ferromagnetic component is attributed to the Dzialoshinsky–Moriya antisymmetric exchange [18, 19]. The Néel temperature of this compound is about 140 K [20, 21]. The magnetic properties of manganites are related to the spins of manganese ions because their orbital magnetic moments are frozen into the crystalline field of anions, while the $La^{3+}$ and $O^{2-}$ ions are diamagnetic. The presence of Jahn–Teller $Mn^{3+}$ ions with the $t_{2g}^3e_g^1$ ($S = 2$) electron configuration in the stoichiometric $La^{3+}Mn^{3+}O_3^{2-}$ is responsible for the *O*'-orthorhombic symmetry of the unit cell of this compound. The *d* shell of $Mn^{3+}$ ions in the octahedral surrounding of oxygen is split into a doublet and a triplet. The threefold-degenerate $t_{2g}$ level is filled with three electrons, whereas the twofold-degenerate $e_g$ level contains only one electron. Such a high-spin configuration is associated with strong Hund's exchange coupling, which arranges the spins of all electrons in the same direction, while the $t_{2g}^3$ electrons form a local spin of $S = 3/2$.

The substitution of $A^{2+}$ ions for $La^{3+}$ in $La_{1-x}A_xMnO_3$ gives rise to $Mn^{4+}$ ions with the $t_{2g}^3$ ($S = 3/2$) electron configuration. When the concentration $x$ of substitute ions equals about 0.10, cooperative static Jahn–Teller distortions are removed [22]. To explain the magnetic and electrical properties of substituted manganites, one applies the so-called double exchange mechanism [23–25]. The double exchange provides an actual transition of an electron from the partially filled $e_g$ orbital of a $Mn^{3+}$ ion to a free $e_g$ orbital of a $Mn^{4+}$ ion. However, a great deal of the experimental results on the properties of substituted manganites can only be understood in terms of the model of phase separations [8, 9] and indirect superexchange interactions [26, 27].

The system of solid solutions $La_{1-x}Sr_xMnO_3$ represents the best example of a double exchange system. The samples of this system are characterized by the greatest width *W* of the one-electron band and, hence, are the least subject to Coulomb correlations [28]. An exception is the $La_{1-x}Sr_xMnO_3$ ($x \le 10$) samples, which are characterized by cooperative static Jahn–Teller distortions. The temperature of a Jahn–Teller transition is determined by the concentration of $Mn^{4+}$ ions. An electronic phase diagram for $La_{1-x}Sr_xMnO_3$ was first proposed in [1], and the modern diagram has been refined in [29–33].

A review of the literature shows that the effect of oxygen deficiency on the physical properties of the substituted manganites $La_{1-x}Sr_xMnO_3$ has been studied poorly. It is clear that oxygen vacancies will break the ordering of magnetic ions ($Mn^{3+}$ and $Mn^{4+}$) and change the angular and radial distributions of the electron density, which causes considerable changes in the magnetic and electrical properties of substituted manganites. Currently, there are only dispersed experimental data on the oxygen nonstoichiometry for separate manganite compounds [34–37]. Detailed systematic studies of this kind have not yet been performed. Therefore, the present paper is devoted to the systematic study of the crystalline structure and the magnetic and electric states of the anion-deficient manganites $La_{1-x}Sr_xMnO_{3-x/2}$. The solid solutions of this system contain only $Mn^{3+}$ ions; therefore, the magnetic properties of these solutions cannot be interpreted in terms of the double exchange model.

## 2. EXPERIMENT

Stoichiometric ceramic solid solutions of $La_{1-x}Sr_xMnO_3$ ($x = 0$; 0.05; 0.075; 0.10; 0.125; 0.175; 0.225; 0.30) were prepared by the conventional ceramic technology. $La_2O_3$ and $Mn_2O_3$ oxides and $SrCO_3$ carbonate (all materials of special purity) were weighted to a prescribed ratio of cations (La : Sr : Mn = $(1 - x)$ : $x$ : 1) and thoroughly mixed. The chemical materials were ground in an agate mortar with a small amount of ethyl alcohol. Since $La_2O_3$ is rather hygroscopic, it was annealed in air at 1000°C for 5 h before weighting to remove moisture and carbon dioxide. The mixtures of lanthanum and manganese oxides and a strontium carbonate thus obtained were compressed into pellets with a diameter of 2 cm and a height of 1.5 cm by a hydraulic press in a steel press mold under a pressure of about $10^8$ P and annealed in air at 1100°C for 2 h. Then the pellets were ground again, pressed, and finally synthesized in air at 1550°C for 2 h and slowly (100 deg/h) cooled to room temperature. During the synthesis, the samples were kept on a platinum substrate. The temperature in a furnace with chromide lanthanum heaters was controlled by a platinum–platinum-rhodium (10%) thermocouple. The reference junction of the thermocouple was placed in melting ice. The heating and cooling rates of samples in the furnace were controlled by an RIF-101 device.

The equation of the chemical reaction for obtaining stoichiometric samples of lanthanum manganite substituted by strontium ions can be represented as follows:

$$[(1-x)/2]La_2O_3 + xSrCO_3 + 0.50Mn_2O_3 \longrightarrow La_{1-x}Sr_xMnO_z + xCO_2\uparrow. \quad (1)$$

To obtain solid solutions with the content of oxygen close to stoichiometric, $La_{1-x}Sr_xMnO_3$ samples with $x < 0.20$ were tempered at 1000°C, and the samples with $x \geq 0.20$ were kept in air at 900°C for 100 h and then cooled to room temperature at a rate of 100 deg $h^{-1}$. It is well known that the concentration of oxygen in the substituted manganites obtained by such technology is close to that in stoichiometric samples [38].

An X-ray analysis of the synthesized samples was carried out on a DRON-3 diffractometer with the $K_\alpha$ radiation of Cr at room temperature in the interval of angles $30° \leq 2\theta \leq 100°$. A graphite monochromator was used to filter out the $K_\beta$ radiation. The content of oxygen was determined by a thermogravimetric analysis. According to our investigations, the samples synthesized in air were oxygen stoichiometric.

Anion-deficient solid solutions of $La_{1-x}Sr_xMnO_{3-x/2}$ ($x$ = 0; 0.05; 0.075; 0.10; 0.125; 0.175; 0.225; 0.30) were prepared by the method of topotactic reactions. Samples were placed in evacuated ($P \sim 10^{-4}$ Pa) quartz ampoules together with a certain amount of metallic tantalum, which was used as an oxygen absorber. The quartz ampoules were kept at 800°C for 24 h and then cooled to room temperature at a rate of 100 deg $h^{-1}$. The equation of the chemical reaction of reduction is expressed as follows:

$$La_{1-x}Sr_xMnO_3 + (x/5)Ta \longrightarrow La_{1-x}Sr_xMnO_{3-x/2} + (x/10)Ta_2O_5. \quad (2)$$

The content of oxygen in the anion-deficient solid solutions $La_{1-x}Sr_xMnO_{3-x/2}$ was determined by a change in the mass of samples by weighting them before and after reduction. To reduce the relative error in measuring the content of oxygen, a sample with a mass of about 3 g was usually placed in a quartz ampoule. In this case, the error was no greater than 0.3%. The anion-deficient samples were subjected to oxidation in air at 900°C for 5 h. This reaction can be described by the equation

$$La_{1-x}Sr_xMnO_{3-x/2} + (x/4)O_2 \longrightarrow La_{1-x}Sr_xMnO_3. \quad (3)$$

After oxidation, the samples were weighted again to control the content of oxygen. These investigations have shown that the absolute accuracy of determining the oxygen index is ±0.01, and the chemical formula of anion-deficient solid solutions can be expressed as $La_{1-x}Sr_xMnO_{3-x/2\pm0.01}$.

The magnetic properties were measured on an OI-3001 commercially available vibrational magnetometer in the range of temperatures from 4 to 300 K. The specific magnetic moment was measured as a function of temperature in weak fields (below 100 Oe) in the zero-field-cooling (ZFC) and field-cooling (FC) regimes, and also as a function of a field at a low temperature (5 K). The Néel temperature $T_N$ was determined from the temperature dependence of the inverse magnetic susceptibility. The dynamic magnetic susceptibility was measured by an inductance bridge in the temperature interval from 77 to 350 K. The field amplitude was equal to 200 A/m, and the frequency was 1200 Hz. The temperature of magnetic ordering ($T_{mo}$) was determined by the temperature dependence of the FC curve of the specific magnetic moment in a sufficiently weak magnetic field of 100 Oe as the temperature of the fastest decrease of the specific magnetic moment ($\min\{dM_{FC}/dT\}$). The freezing temperature $T_F$ of the magnetic moments of ferromagnetic clusters was determined as the temperature corresponding to the maximum of the ZFC curve of the specific magnetic moment. The electric resistivity was measured on well-sintered samples with a size of $10 \times 2 \times 2$ $mm^3$ and without obvious macroscopic cracks by the conventional four-probe technique in the range of temperatures from 77 to 330 K. Indium contacts were formed by ultrasonic deposition. The magnetoresistance was calculated by the formula

$$MR(\%) = \{[\rho(H) - \rho(0)]/\rho(0)\} \times 100\%, \quad (4)$$

where MR(%) is the negative isotropic magnetoresistance in percent, $\rho(H)$ is the resistivity in a magnetic field of 9 kOe, and $\rho(0)$ is the resistivity in zero magnetic field. The electric current was directed along the longest side of a sample. The magnetic field was applied parallel to the electric field in a sample.

## 3. EXPERIMENTAL RESULTS

To assess the crystalline structure of the stoichiometric $La_{1-x}Sr_xMnO_3$ and anion-deficient $La_{1-x}Sr_xMnO_{3-x/2}$ samples, we carried out a powder X-ray diffraction analysis. The X-ray diagrams of some of the solid solutions are shown in Fig. 1. We found that all the samples have the perovskite structure. The symmetry of a unit cell depends on the concentration of the substitute ion. For instance, stoichiometric samples of $La_{1-x}Sr_xMnO_3$ exhibit a change in the type of symmetry of the unit cell at $x = 0.125$. Samples with $x \leq 0.125$ have an *O'*-orthorhombic unit cell (the space group *Pbnm*, $Z = 4$), whereas samples with $x > 0.125$ have a rhombohedral unit cell (the space group $R\bar{3}c$, $Z = 2$). Earlier, we observed such a change in the symmetry of the unit cell in the system of barium-substituted solid solutions

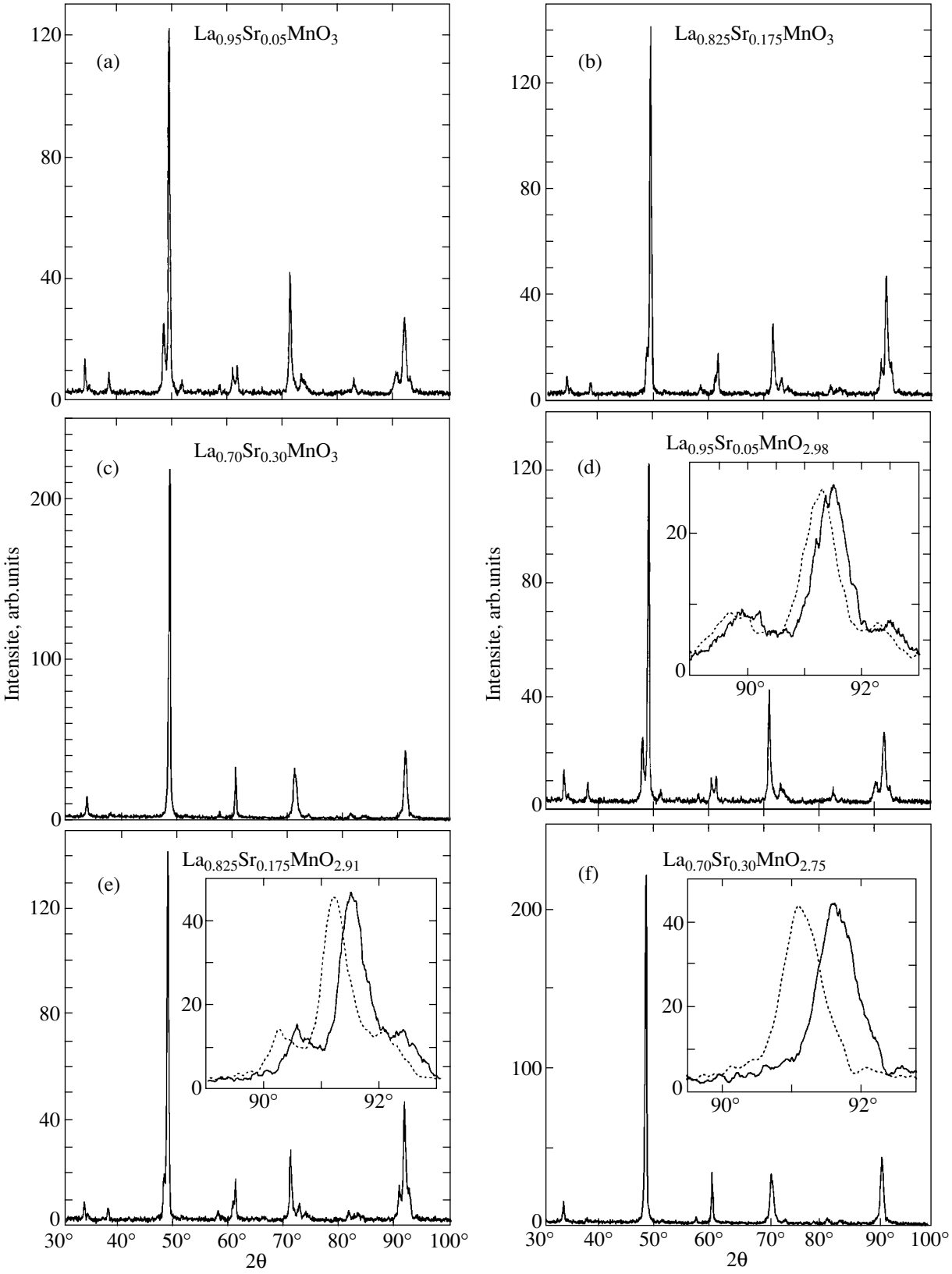



**Fig. 1.** Powder X-ray diffraction patterns obtained at room temperature for (a–c) stoichiometric solid solutions $La_{1-x}Sr_xMnO_3$ and (d–f) anion-deficient solid solutions $La_{1-x}Sr_xMnO_{3-x/2}$ for the following values of $x$: (a, d) 0.05, (b, e) 0.175, and (c, f) 0.30. The inset shows the X-ray multiplets (132 + 024 + 204 + 312) for samples with $x = 0.075$ and 0.175 and ($21\bar{1}$ + 211 + $\bar{2}11$) for $x = 0.30$. The solid lines represent data for stoichiometric samples, and the dashed lines, for anion-deficient samples.

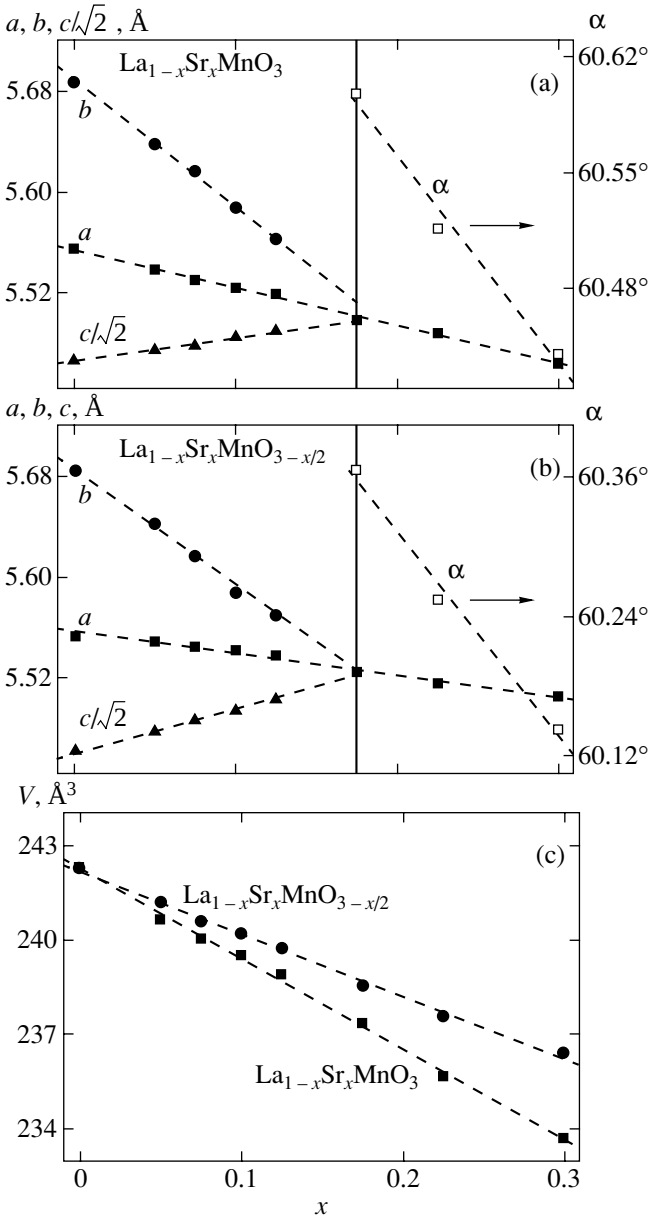



**Fig. 2.** Concentration dependence of the unit-cell parameters (filled squares) *a*, (filled circles) *b*, (filled triangles) *c*, and (open squares) α at room temperature for (a) stoichiometric solid solutions $La_{1-x}Sr_xMnO_3$ and (b) anion-deficient solid solutions $La_{1-x}Sr_xMnO_{3-x/2}$. Panel (c) shows the concentration dependence of comparable volumes of a unit cell for (filled squares) stoichiometric and (filled circles) anion-deficient samples.

$La_{1-x}Ba_xMnO_{3-x/2}$ [39]. The concentration dependence of structural distortions in anion-deficient samples of $La_{1-x}Sr_xMnO_{3-x/2}$ is analogous to that in the stoichiometric system. It is noteworthy that identical X-ray reflexes are displaced toward smaller Bragg angles in the case of anion-deficient samples. This fact indicates that the volume of the unit cell of anion-deficient samples is greater than that of the stoichiometric samples.

According to Goodenough [27], the *O*'-orthorhombic symmetry ($c/\sqrt{2} < a \le b$) is attributed to the cooperative static Jahn–Teller effect, which is responsible for the antiferrodistortion ordering of $d_{z^2}$ orbitals in $Mn^{3+}$ ions. This effect consists in reducing the free energy of a degenerate electron system by reducing its symmetry through the removal of the degeneracy of electron states. A necessary condition for the Jahn–Teller effect is the presence of a degenerate electron system. These may be $d^9d^7$ ions in a low-spin and $d^4$ ions in high-spin states. In the case of manganites, the degeneracy of $e_g$ levels of $Mn^{3+}$ in an octahedral surrounding of oxygen anions is removed. Earlier [40], it was established that the system $La_{1-x}Ca_xMnO_3$ with $x \sim 0.1$ experiences a transition from the *O*'-orthorhombic to the *O*-orthorhombic symmetry ($a < c/\sqrt{2} < b$). One may assume that the $O' \longrightarrow R$ transition in our case is associated with the fact that the size effect dominates the Jahn–Teller effect because a similar transition is observed in barium-substituted manganites. It is well known that, for a coordination number of 12, the effective radii of $Sr^{2+}$ ($r\{Sr^{2+}\}$ = 1.44 Å) and $Ba^{2+}$ ($r\{Ba^{2+}\}$ = 1.61 Å) ions are greater than those of $La^{3+}$ ($r\{La^{3+}\}$ = 1.36 Å) and $Ca^{2+}$ ($r\{Ca^{2+}\}$ = 1.34 Å) ions [41].

A variation in the concentration dependence of the parameters and the unit-cell volume of stoichiometric $La_{1-x}Sr_xMnO_3$ and anion-deficient $La_{1-x}Sr_xMnO_{3-x/2}$ samples is illustrated in Fig. 2. As the substitution of Sr ions increases, the parameters *a* and *b* monotonically decrease, while $c/\sqrt{2}$ increases. The change in the type of symmetry produces virtually no effect on the monotonicity of the variation in the parameter *a*. The angle α also varies monotonically (Figs. 2a and 2b).

As *x* increases, the unit-cell volume of the stoichiometric $La_{1-x}Sr_xMnO_3$ samples monotonically decreases (Fig. 2c). Such behavior is associated with the fact that $Mn^{4+}$ ions arise in the system whose radius ($r\{Mn^{4+}\}$ = 0.530 Å) in the octahedral surrounding of anions is less than the radius of $Mn^{3+}$ ions ($r\{Mn^{3+}\}$ = 0.645 Å) and that the size effect of increasing the *A* sublattice, $\langle \Delta r_A \rangle = +0.08$, is overbalanced by the decrease of the *B* sublattice, $\langle \Delta r_B \rangle = -0.115$. When oxygen vacancies arise, the coordination number decreases, and, hence, the effective radius of $Mn^{3+}$ ions also decreases. The effective radius of $Mn^{3+}(V)$ ions in the pentahedral coordination is equal to 0.580 Å [41]. In this case, the effect of the *B* sublattice is as small as $\langle \Delta r_B \rangle = -0.065$. However, due to the presence of actual vacancies of oxygen anions, the unit-cell volume of anion-deficient samples of $La_{1-x}Sr_xMnO_{3-x/2}$ also decreases as *x* increases, although this decrease is somewhat smaller (Fig. 2c).

The temperature and field dependence of the magnetic properties of certain stoichiometric $La_{1-x}Sr_xMnO_3$ samples are shown in Fig. 3. The dynamic magnetic susceptibility of the $La_{0.95}Sr_{0.05}MnO_3$ (Fig. 3a) and $La_{0.70}Sr_{0.30}MnO_3$ (Fig. 3b) samples obeys the Curie–Weiss law with $T_N \approx 141$ K and $T_C \approx 350$ K, respectively. As $x$ increases, the spontaneous atomic magnetic moment also increases. For $La_{0.70}Sr_{0.30}MnO_3$, it amounts to about $3.7\mu_B$ per manganese ion. These results are in good agreement with the results of other authors [1, 2, 20, 42].

The temperature dependence of the specific magnetic moment in the ZFC and FC regimes for the anion-deficient solid solutions $La_{1-x}Sr_xMnO_{3-x/2}$ is shown in Fig. 4. The samples investigated show substantially different behavior of the ZFC and FC curves. Most probably, this fact points to the absence of long-range ferromagnetic ordering in these samples. For $La_{0.925}Sr_{0.075}MnO_{2.95}$ (Fig. 4a), the FC curve starts to slowly increase below $T_{mo} \sim 110$ K. The ZFC curve has a peak at $T \sim 90$ K and is virtually independent of temperature below this point. The samples of $La_{0.775}Sr_{0.225}MnO_{2.89}$ (Fig. 4b) and $La_{0.70}Sr_{0.30}MnO_{2.85}$ (Fig. 4c) have a rather wide transition region to a paramagnetic state. The peaks of the ZFC curves are observed at about 45 and 40 K, respectively. These temperatures are taken as the freezing points $T_F$ of the magnetic moments of ferromagnetic clusters. The FC curves are virtually constant below $T_F$. Thus, anion-deficient samples with $x > 0.10$ show properties that are characteristic of cluster spin glasses. A slow decrease in $T_F$ with increasing $x$ points to a decrease in the size of ferromagnetic clusters.

The field dependence of the atomic magnetic moment of anion-deficient samples of the system $La_{1-x}Sr_xMnO_{3-x/2}$ (Fig. 5) substantially differs from that of the stoichiometric system $La_{1-x}Sr_xMnO_3$. First, the magnetic moment is not completely saturated in fields of up to 16 kOe, which slightly complicates the determination of the spontaneous atomic magnetic moment $M_s$. Second, under the assumption of complete ferromagnetic ordering of $Mn^{3+}$ ions, the theoretically possible value of $M_s$, equal to $4\mu_B$ per formula unit, is attained in none of the samples. Third, $M_s$ gradually increases with $x$ and then decreases, and the maximal value of the ferromagnetic component is observed in the $La_{0.825}Sr_{0.175}MnO_{2.91}$ sample.

All anion-deficient solid solutions of $La_{1-x}Sr_xMnO_{3-x/2}$ are characterized by the semiconductor type of electric conductivity. The temperature dependence of the resistivity and the magnetoresistance of $La_{0.825}Sr_{0.175}MnO_{2.91}$ is shown in Fig. 6. As $x$ increases, the room-temperature resistivity of anion-deficient samples first decreases ($\rho \approx 5\ \Omega$ cm) for the values of $x$ up to 0.175 and then increases (up to $x = 0.30$). In the whole range of concentrations, anion-deficient solid solutions do not show a metal–semiconductor transition. Note that the resistivity of samples with large oxygen deficiency ($x \geq 0.175$) is satisfactorily described by the equation $\ln\rho \propto T^{-1}$ (see also the inset to Fig. 6). In almost all the samples, the absence of anomalies in the magnetoresistance is attributed to the absence of anomalies in the electric resistance near the transition temperature to the magnetically ordered state. A magnetoresistance peak equal to about 17% was observed only in $La_{0.825}Sr_{0.175}MnO_{2.91}$ near the temperature $T \approx 90$ K.

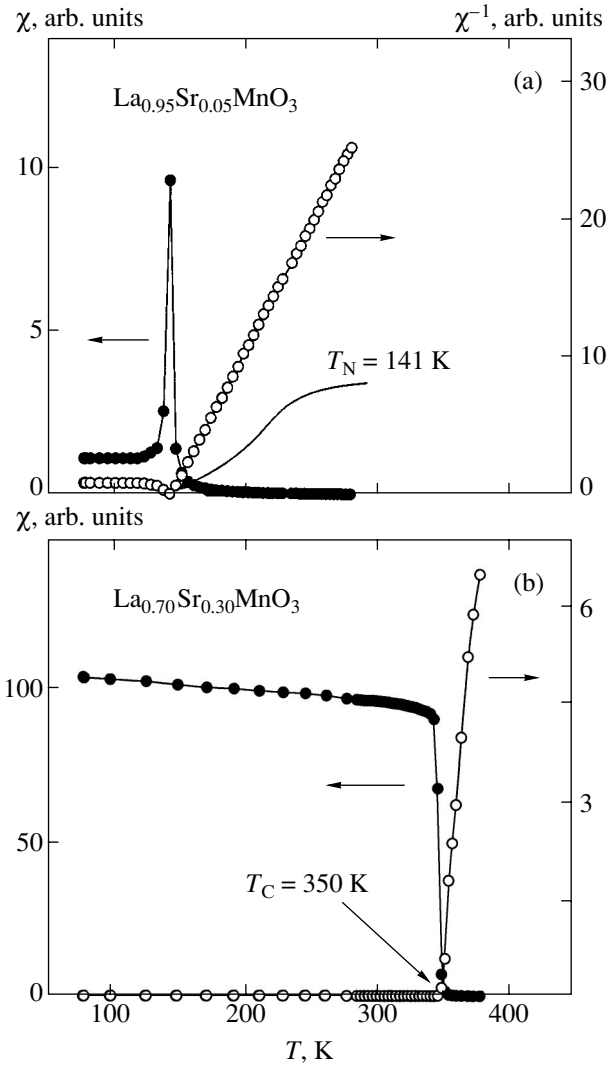



**Fig. 3.** Temperature dependence of (filled symbols) dynamic magnetic susceptibility and (open symbols) inverse dynamic magnetic susceptibility for the stoichiometric solid solutions (a) $La_{0.95}Sr_{0.05}MnO_3$ and (b) $La_{0.70}Sr_{0.30}MnO_3$.

The results of magnetic investigations of anion-deficient samples of $La_{1-x}Sr_xMnO_{3-x/2}$ are summarized in Figs. 7 and 8. As the concentration of the substitute ion increases, the spontaneous atomic magnetic moment

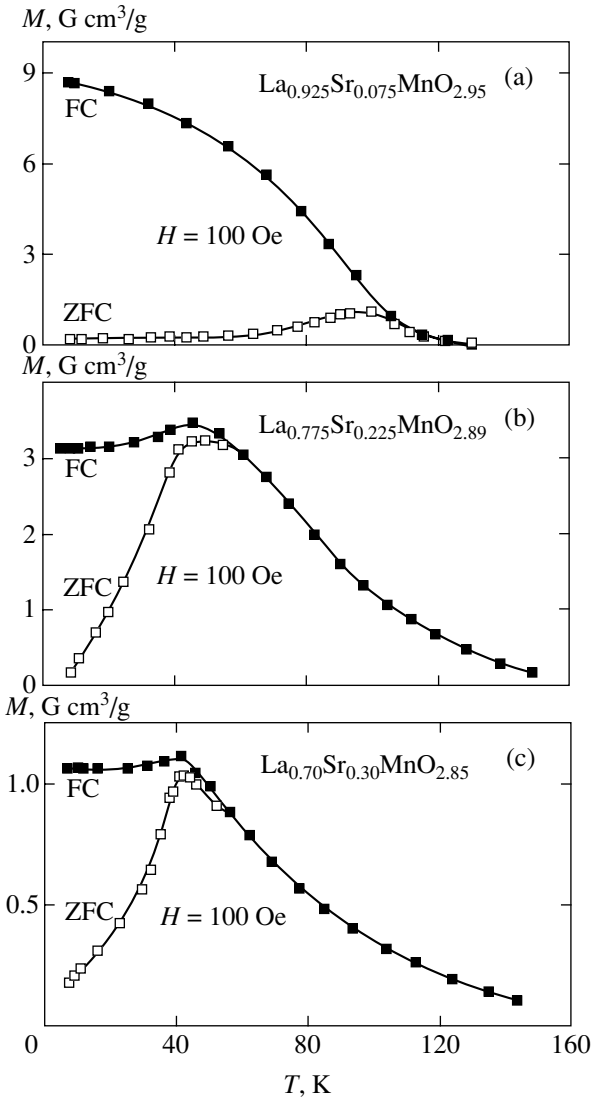



**Fig. 4.** Temperature dependence of the specific magnetic moment in an external magnetic moment of 100 Oe measured on warming after zero-field cooling and field cooling at 100 Oe for the anion-deficient solid solutions (a) $La_{0.925}Sr_{0.075}MnO_{2.95}$, (b) $La_{0.775}Sr_{0.225}MnO_{2.89}$, and (c) $La_{0.70}Sr_{0.30}MnO_{2.85}$.

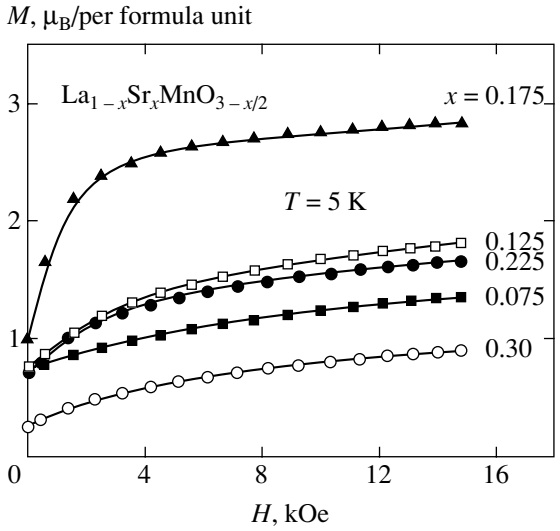



**Fig. 5.** Atomic magnetic moment versus external magnetic field at 5 K for the anion-deficient solid solutions $La_{1-x}Sr_xMnO_{3-x/2}$ with (filled squares) $x = 0.075$, (open squares) 0.125, (filled triangles) 0.175, (filled circles) 0.225, and (open circles) 0.30.

slowly increases (at $x = 0$, we have $M_s \approx 0.1\mu_B$ per formula unit) and then decreases (at $x = 0.30$, we have $M_s \approx 0.35\mu_B$ per formula unit). The point $x = 0.175$, which corresponds to the maximal value of the ferromagnetic component $M_s \approx 2.52\mu_B$ per formula unit, is a critical point. However, for any sample composition, $M_s$ does not reach its theoretically possible value calculated under the assumption of purely ferromagnetic ordering of $Mn^{3+}$ ions.

$LaMnO_3$ is an *A*-type ferromagnet with $T_N \approx 140$ K. Because of the weak ferromagnetic component associated with the Dzialoshinsky–Moriya antisymmetric exchange interaction [18, 19], this compound is often called a weak ferromagnet. As $x$ increases from 0 to 0.10, $T_N$ decreases from 140 to 100 K. From the viewpoint of magnetic structure, $La_{1-x}Sr_xMnO_{3-x/2}$ samples in this range of concentrations represent an antiferromagnetic charge-ordered phase with ferromagnetic cluster inclusions. In the interval $0.10 < x \le 0.175$, the ground state is the mixture of ferromagnetic and antiferromagnetic charge-disordered phases with $T_C$ increasing up to about 130 K. A transition to the paramagnetic state occurs through a mixed state with short-range ferromagnetic correlations below $T_F$. For $x > 0.175$, all anion-deficient samples are cluster spin glasses with $T_F \approx 40$ K. It should be noted that a temperature of 40 K is typical for the state of a cluster spin glass in manganites [43]. Information about the magnetic phase states of the stoichiometric solid solutions of $La_{1-x}Sr_xMnO_3$ can be found in [29–33].

## 4. DISCUSSION OF THE RESULTS

Substituted lanthanum manganites $La_{1-x}Sr_xMnO_{3-x/2}$ with oxygen deficiency are of interest for the experimental investigation of an indirect 180° exchange between manganese ions [44]. Note that $La^{3+}$, $Sr^{2+}$, and $O^{2-}$ ions are diamagnetic. This fact significantly simplifies the interpretation of the results of magnetic investigations.

The following arguments can be used to understand the nature of magnetic processes that take place in anion-deficient samples of $La_{1-x}Sr_xMnO_{3-x/2}$. As pointed out above, to explain the magnetic properties of substituted manganites, one often applies the theory of indirect 180° exchange interactions—the double exchange mechanism, which was proposed by Zener

[23, 24] and developed in detail by de Gennes [25]. The double exchange mechanism can be realized in a system with a mixed valence of manganese ions ($Mn^{3+}/Mn^{4+}$). The samples of the system considered here contain $Mn^{3+}$ ions alone; therefore, the double exchange mechanism cannot be applied to interpret the magnetic properties of these samples.

Goodenough formulated the basic principles of the superexchange theory—a specific type of the indirect 180° exchange interaction [26]. It is assumed that ferromagnetism can be attributed to the specific character of exchange interactions in the system of $Mn^{3+}$ Jahn–Teller ions. The orbital configuration of 3*d* electrons in case of removed cooperative static Jahn–Teller distortions is determined by the positions of the nuclei of manganese ions. In order that the $Mn^{3+}$–O–$Mn^{3+}$ interactions become ferromagnetic, a correlation should exist between the electron configuration and the vibrational modes of the nuclei. The $Mn^{3+}$–O–$Mn^{3+}$ superexchange interactions are anisotropic in the orbitally ordered phase (positive in the (001) plane and negative along the [001] direction) but isotropic in the orbitally disordered phase (positive along any direction) [27].

The compound $LaMnO_3$ ($x = 0$) exhibits an antiferrodistortion orbital ordering (the ordering of $d_{z^2}$ orbitals) caused by the Jahn–Teller effect. Therefore, the $Mn^{3+}$–O–$Mn^{3+}$ superexchange interactions are antiferromagnetic, and $LaMnO_3$ is an *A*-type antiferromagnet (Fig. 8).

The substitution of $Sr^{2+}$ ions for $La^{3+}$ leads to the dilution of the Jahn–Teller system and removes the orbital ordering. As a result, one gets ferromagnetic clusters in an antiferromagnetic charge-ordered matrix. In the interval $0 < x \le 0.10$, anion-deficient samples show an increase in the ferromagnetic component and a decrease in $T_N$. The elimination of the orbital ordering also gives rise to oxygen vacancies that change the angular and radial distributions of the electron density.

In the interval $0.10 < x \le 0.175$, the ferromagnetic component continues to grow and $T_C$ also increases. The value $x = 0.175$ at which the Jahn–Teller effect and, as a result, the orbital ordering are completely eliminated is a critical concentration. However, a purely ferromagnetic state cannot be realized because the appearance of oxygen vacancies changes the sign of $Mn^{3+}$–O–$Mn^{3+}$ exchange interactions. A $Mn^{3+}$–O–$Mn^{3+}$ superexchange interaction for a pentahedral coordination of at least one manganese ion is antiferromagnetic. The compound $Ca^{2+}Mn^{3+}O_{2.50}^{2-}$, which is antiferromagnetic, serves as confirmation of this fact [45]. The $Mn^{3+}$ ions in this compound are in a pentahedral coordination of oxygen anions. The samples in this region consist of antiferromagnetic (orbitally ordered and disordered) and ferromagnetic (orbitally disordered) phases. A transition to the paramagnetic state occurs through a mixed state with short-range ferromagnetic correlations below $T_F$.

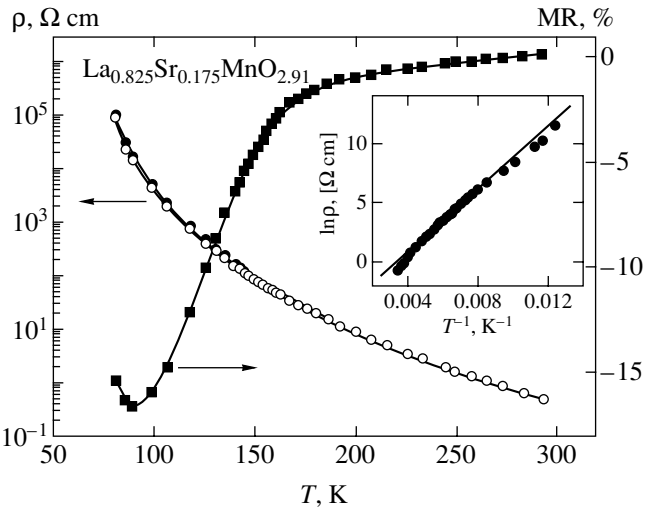



**Fig. 6.** Temperature dependence of the resistivity, measured (filled circles) in zero magnetic field and (open circles) in a magnetic field of 9 kOe, and (filled squares) the magnetoresistance for the anion-deficient sample $La_{0.825}Sr_{0.175}MnO_{2.91}$. The inset shows the natural logarithm of resistivity versus inverse temperature for the same sample.

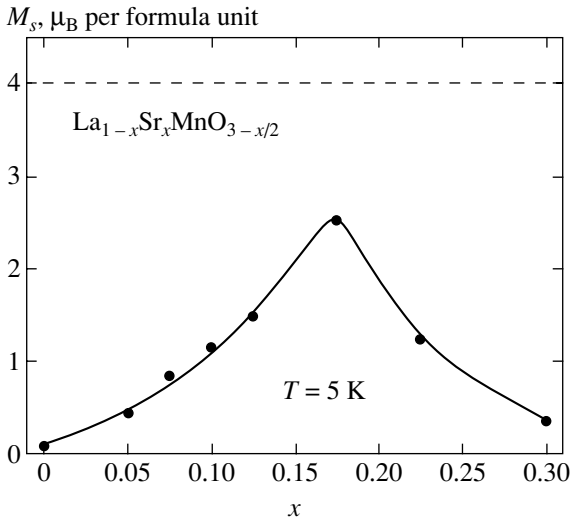



**Fig. 7.** Concentration dependence of experimental values (filled circles) of spontaneous atomic magnetic moment at 5 K for the anion-deficient solid solutions $La_{1-x}Sr_xMnO_{3-x/2}$. The dashed line represents expected theoretical values of the spontaneous atomic magnetic moment for the system of samples considered.

For $x > 0.175$, the volume of the orbitally disordered antiferromagnetic phase that is associated with the $Mn^{3+}$(V)–O–$Mn^{3+}$(V) interaction starts to rapidly increase, thus reducing the spontaneous magnetic moment. Competition between the antiferromagnetically and ferromagnetically ordered clusters is likely to

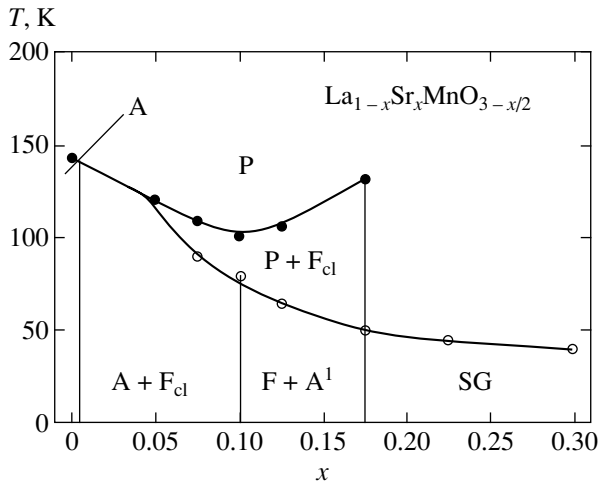



**Fig. 8.** Concentration dependence of the critical temperatures of magnetic phase transitions for the anion-deficient solid solutions $La_{1-x}Sr_xMnO_{3-x/2}$; A is an *A*-type charge-ordered ferromagnetic phase, A + $F_{cl}$ is a mixed magnetic state consisting of an antiferromagnetic phase and ferromagnetic clusters, F + $A^1$ is a mixed magnetic state consisting of ferromagnetic and antiferromagnetic charge-disordered phases, P + $F_{cl}$ is a mixed magnetic state consisting of a paramagnetic phase and ferromagnetic clusters, SG is a cluster spin glass, and P is a paramagnet. Filled symbols denote the transition temperature to a paramagnetic state, and open symbols denote the freezing temperature of the magnetic moments of ferromagnetic clusters.

give rise to a cluster spin-glass-type state with the slowly decreasing freezing temperature of the magnetic moments of ferromagnetic clusters. The orientation of the magnetic moments of spin-glass clusters below $T_F$ does not show any spatial periodicity. It randomly varies in space, similar to the configuration of atoms in ordinary glasses. In contrast to paramagnets, where the magnetic moment fluctuates with time, spin glasses are characterized by "frozen" magnetic moments, i.e., by nonzero time-average vector quantities. A cluster spin-glass-type state is frequently observed in inhomogeneous magnetic systems such as Co–Cu and Co–Ag granular films [46, 47]. In these films, ferromagnetic grains are embedded into a nonferromagnetic matrix. The realization of this possibility in the anion-deficient solid solutions $La_{1-x}Sr_xMnO_{3-x/2}$ is confirmed by a clear-cut peak on the temperature dependence of the ZFC curve. A slow decrease in $T_F$ points to a decrease in the size of ferromagnetically ordered clusters.

A similar behavior of magnetic properties is observed in the substituted anion-deficient manganites $La_{1-x}Ca(Ba)_xMnO_{3-x/2}$ [39, 40]. The specific features of the resistivity and the magnetoresistance of polycrystalline anion-deficient samples of $La_{1-x}Sr_xMnO_{3-x/2}$ can be understood based on the fact that they consist of two contributions: (1) intragranular and (2) intergranular domains (the surface of grains). As $x$ increases, the volume of ferromagnetic clusters increases. This increases the mobility of charge carriers and thereby reduces the resistivity for $x \le 0.175$. As the volume of the antiferromagnetic (orbitally disordered) phase increases for $x > 0.175$, the scattering of charge carriers by randomly distributed spins outside ferromagnetic clusters increases, thus increasing the resistivity.

It is well known that the diffusion coefficient of oxygen anions situated on the surface of grains is an order of magnitude greater than that of oxygen anions inside a grain [48]; therefore, the reduction reaction gives rise to oxygen-anion-poor microscopic regions on the surface of grains. Thus, the reduction leads to a domination of the intergranular contribution and increases the resistivity. A metal–semiconductor transition due to the intragranular contribution does not occur. The higher the concentration of oxygen vacancies, the higher the resistivity.

## 5. CONCLUSIONS

The phase composition and the structural, magnetic, and electrotransport properties of $La_{1-x}Sr_xMnO_{3-x/2}$ ($0 \le x \le 0.30$) manganites have been investigated experimentally as a function of the concentration of oxygen deficiency. In the stoichiometric $La_{1-x}Sr_xMnO_3$ solid solutions, a change in the type of symmetry of the unit cell has been observed for $x = 0.125$. Samples with $x \le 0.125$ are characterized by an *O*'-orthorhombic unit cell, whereas samples with $x > 0.125$ are characterized by a rhombohedral unit cell. The structural properties of the anion-deficient solid solutions $La_{1-x}Sr_xMnO_{3-x/2}$ are analogous to those of the stoichiometric system. It is assumed that, as the oxygen content decreases, $La_{1-x}Sr_xMnO_{3-x/2}$ anion-deficient solid solutions experience a series of successive magnetic phase transformations in the ground state: from an *A*-type ($x = 0$) antiferromagnet to a cluster spin-glass-type inhomogeneous magnetic state ($0.175 < x \le 0.30$) through a two-phase (antiferromagnetic and ferromagnetic) state ($0 < x \le 0.175$). The anion-deficient solid solution with $x = 0.175$ has the maximal value of the ferromagnetic component. As the oxygen deficiency increases, the resistivity of $La_{1-x}Sr_xMnO_{3-x/2}$ samples first decreases (up to a value of $x = 0.175$), and then increases (up to a value of $x = 0.30$). In this case, none of the anion-deficient solid solutions exhibits a metal–semiconductor transition in the whole range of concentrations considered. A peak of magnetoresistance below the temperature of magnetic ordering has been observed only in a sample with $x = 0.175$. The results of experiments have been summarized in the concentration diagrams of the spontaneous atomic magnetic moment and the critical temperature of magnetic phase transitions. Hypothetical magnetic phase states have been pointed out. The experimental results obtained can be interpreted in terms of the phase-separation model and the competition between ferromagnetic and antiferromagnetic indirect superexchange interactions. It is

assumed that $Mn^{3+}$–O–$Mn^{3+}$ indirect superexchange interactions are positive in the orbitally disordered phase in the case of octahedral coordination of manganese ions and are negative when the coordination of at least one $Mn^{3+}$ ion is pentahedral.

## ACKNOWLEDGMENTS

This work was supported in part by the GPOFI "Nanomaterials and Nanotechnology" (Task 3.3) and by the Belarussian Foundation for Basic Research (project no. F04R-076).

## REFERENCES

1. G. H. Jonker and J. H. van Santen, Physica (Utrecht) **16**, 337 (1950).
2. G. H. Jonker and J. H. van Santen, Physica (Utrecht) **16**, 599 (1950).
3. A. Ramirez, J. Phys.: Condens. Matter **9**, 8171 (1997).
4. M. Imada, A. Fujimori, and Y. Tokura, Rev. Mod. Phys. **70**, 1039 (1998).
5. J. Coey, M. Viret, and S. von Molnar, Adv. Phys. **48**, 167 (1999).
6. Y. Tokura and Y. Tomioka, J. Magn. Magn. Mater. **200**, 1 (1999).
7. V. M. Loktev and Yu. G. Pogorelov, Low Temp. Phys. **26**, 171 (2000).
8. E. Dagotto, T. Hotta, and A. Moreo, Phys. Rep. **344**, 1 (2001).
9. E. L. Nagaev, Phys. Rep. **346**, 387 (2001).
10. M. B. Salamon and M. Jaime, Rev. Mod. Phys. **73**, 583 (2001).
11. D. M. Edwards, Adv. Phys. **51**, 1259 (2002).
12. S. M. Dunaevskiĭ, Phys. Solid State **46**, 193 (2004).
13. R. von Helmholt, J. Wecker, B. Holzapfel, *et al.*, Phys. Rev. Lett. **71**, 2331 (1993).
14. K. Chabara, T. Ohno, M. Kasai, *et al.*, Appl. Phys. Lett. **63**, 1990 (1993).
15. S. Jin, T. H. Tiefel, M. McCormack, *et al.*, Science **264**, 413 (1994).
16. D. Hawe, Sens. Actuators **81**, 268 (2000).
17. K. M. H. Lenssen, D. J. Adelerhof, H. J. Gassen, *et al.*, Sens. Actuators **85**, 1 (2000).
18. I. Dzialoshinsky, J. Phys. Chem. Solids **4**, 241 (1958).
19. T. Moriya, Phys. Rev. **120**, 91 (1960).
20. G. Matsumoto, J. Phys. Soc. Jpn. **29**, 606 (1970).
21. G. Matsumoto, J. Phys. Soc. Jpn. **29**, 615 (1970).
22. I. O. Troyanchuk, Zh. Éksp. Teor. Fiz. **102**, 132 (1992) [Sov. Phys. JETP **75**, 132 (1992)].
23. C. Zener, Phys. Rev. **82**, 403 (1951).
24. C. Zener, Phys. Rev. **82**, 440 (1951).
25. P.-G. de Gennes, Phys. Rev. **118**, 141 (1960).
26. J. B. Goodenough, Phys. Rev. **100**, 564 (1955).
27. J. B. Goodenough, A. Wold, R. J. Arnott, and N. Menyuk, Phys. Rev. **124**, 373 (1961).
28. T. Chatterji, B. Ouladdiaf, P. Mandal, *et al.*, Phys. Rev. B **66**, 054403 (2002).
29. A. Urushibara, Y. Moritomo, T. Arima, *et al.*, Phys. Rev. B **51**, 14103 (1995).
30. H. Kawano, R. Kajimoto, M. Kubota, and H. Yoshizawa, Phys. Rev. B **53**, R14709 (1996).
31. D. Louca, T. Egami, E. L. Brosha, *et al.*, Phys. Rev. B **56**, R8475 (1997).
32. J. S. Zhou, J. B. Goodenough, A. Asamitsu, and Y. Tokura, Phys. Rev. Lett. **79**, 3234 (1997).
33. G. L. Liu, J. S. Zhou, and J. B. Goodenough, Phys. Rev. B **64**, 144414 (2001).
34. B. C. Tofield and W. R. Scott, J. Solid State Chem. **100**, 183 (1974).
35. H. L. Ju, J. Gopalakrishnan, J. L. Peng, *et al.*, Phys. Rev. B **51**, 6143 (1995).
36. M. Itoh, K. Nishi, J. D. Yu, and Y. Inaguma, Phys. Rev. B **55**, 14408 (1997).
37. A. M. De Leon-Guevara, P. Berthet, J. Berthon, *et al.*, Phys. Rev. B **56**, 6031 (1997).
38. P. Schiffer, A. P. Ramirez, W. Bao, and S.-W. Cheong, Phys. Rev. Lett. **75**, 3336 (1995).
39. S. V. Trukhanov, I. O. Troyanchuk, N. V. Pushkarev, and G. Shimchak, Zh. Éksp. Teor. Fiz. **123**, 128 (2003) [JETP **96**, 110 (2003)].
40. I. O. Troyanchuk, S. V. Trukhanov, G. Shimchak, *et al.*, Zh. Éksp. Teor. Fiz. **120**, 183 (2001) [JETP **93**, 161 (2001)].
41. R. D. Shannon, Acta Crystallogr. A **32**, 751 (1976).
42. H. Watanabe, J. Phys. Soc. Jpn. **16**, 433 (1961).
43. S. V. Trukhanov, I. O. Troyanchuk, M. Hervieu, *et al.*, Phys. Rev. B **66**, 184424 (2002).
44. E. E. Havinga, Philips Res. Rep. **21**, 432 (1966).
45. K. R. Poeppelmeier, M. E. Leonowicz, J. C. Scanlon, *et al.*, J. Solid State Chem. **45**, 71 (1982).
46. S. Nafis, J. A. Woollam, Z. S. Shan, and D. J. Sellmyer, J. Appl. Phys. **70**, 6050 (1991).
47. F. Conde, C. Gomez-Polo, and A. Hernando, J. Magn. Magn. Mater. **138**, 123 (1994).
48. K. A. Thomas, P. S. I. P. N. de Silva, L. F. Cohen, *et al.*, J. Appl. Phys. **84**, 3939 (1998).

*Translated by I. Nikitin*

*Journal of Experimental and Theoretical Physics, Vol. 99, No. 4, 2004, pp. 766–775.*
*Translated from Zhurnal Éksperimental'noĭ i Teoreticheskoĭ Fiziki, Vol. 126, No. 4, 2004, pp. 887–897.*


**SOLIDS**
**Structure**

# Antiferromagnetic Resonance and Phase Diagrams of Gadolinium Ferroborate $GdFe_3(BO_3)_4$

**A. I. Pankrats, G. A. Petrakovskiĭ, L. N. Bezmaternykh, and O. A. Bayukov**

*Kirenskiĭ Institute of Physics, Siberian Division, Russian Academy of Sciences, Akademgorodok, Krasnoyarsk, 660036 Russia*

*e-mail: pank@iph.krasn.ru*

Received March 29, 2004

**Abstract**—Antiferromagnetic resonance in single crystals of rhombohedral gadolinium ferroborate $GdFe_3(BO_3)_4$ was studied. The frequency–field dependences of antiferromagnetic resonance over the frequency range 26–70 GHz and the temperature dependences of resonance parameters for magnetic fields oriented along the crystal axis and in the basal plane were determined. It was found that the iron subsystem, which can be treated as a two-sublattice antiferromagnet with anisotropy of the easy-plane type, experienced ordering at $T$ = 38 K. At temperatures below 20 K, the gadolinium subsystem with the opposite anisotropy sign strongly influenced the anisotropic properties of the crystal. This resulted in a spontaneous spin-reorientation transition from the easy-plane to the easy-axis state at 10 K. Below 10 K, magnetic field-induced transitions between the states were observed. Experimental phase diagrams on the temperature–magnetic field plane were constructed for fields oriented along the crystal axis and in the basal plane. A simple model was used to calculate the critical transition fields. The results were in close agreement with the experimental values measured at $T$ = 4.2 K for both field orientations. © *2004 MAIK "Nauka/Interperiodica".*

## 1. INTRODUCTION

Crystals whose magnetic subsystems are formed by ions of different kinds have interesting magnetic properties. These properties are especially pronounced in rare-earth magnets in which rare-earth ions interact with 3*d* iron family ions. Most rare-earth metal ions are strongly anisotropic, and the magnetic anisotropy of such crystals is as a rule determined by the competition of the anisotropic interactions of these groups of ions. This results in the appearance of spin-reorientation transitions, both spontaneous that occur when temperature changes and field-induced. Transitions of this type and magnetic phase diagrams have been thoroughly studied for rare-earth metal ferrite–garnets and orthoferrites [1].

Another class of crystals of rare-earth metal compounds of the general formula $RM_3(BO_3)_4$, where R is a rare-earth metal and M = Fe, Cr, Al, Ga, and Sc [2–5], has received much less attention. These compounds are of interest not only from the point of view of their magnetic properties: they offer promise as materials for laser techniques and second optical harmonic generation. The crystals have hantite rhombohedral structures, space group *R*32 [6]. The magnetic properties of rare-earth crystals of this class remain virtually unstudied. The temperature dependences of magnetization measured for polycrystalline $RFe_3(BO_3)_4$ samples, where R = Y, La, Nd, Eu, and Ho [6, 7], made it possible to suggest antiferromagnetic ordering in these compounds.

More detailed and informative studies have recently been performed for $GdFe_3(BO_3)_4$ single crystals [8]. Temperature dependence anomalies were observed for magnetization at about 10 and 40 K. In addition, the field dependences of magnetization contained jumps below 10 K when the field was directed along the crystal axis. These jumps were interpreted as spin flop transitions. On the basis of these data, a model of the magnetic structure of $GdFe_3(BO_3)_4$ was suggested in [8]; its reliability, however, is questionable. It is very difficult to unambiguously determine the magnetic structure of a crystal solely on the basis of magnetic measurements. The purpose of this work was to study the magnetic structure and phase transitions of gadolinium ferroborate $GdFe_3(BO_3)_4$ by the antiferromagnetic resonance method, which is very sensitive to the magnetic structure of crystals.

We stress that the $GdFe_3(BO_3)_4$ crystal is the first representative of the family of crystals with the hantite structure for which the magnetic structure and magnetic phase diagram were studied in detail. This crystal offers much promise for studies of this kind, because both magnetically active ions, $Fe^{3+}$ and $Gd^{3+}$, are *S* ions. This circumstance is of special importance for resonance studies, because the corresponding resonance absorption should be fairly narrow.

## 2. EXPERIMENTAL DATA

We used $GdFe_3(BO_3)_4$ single crystals; the procedure for preparing them was described in [8]. Measurements

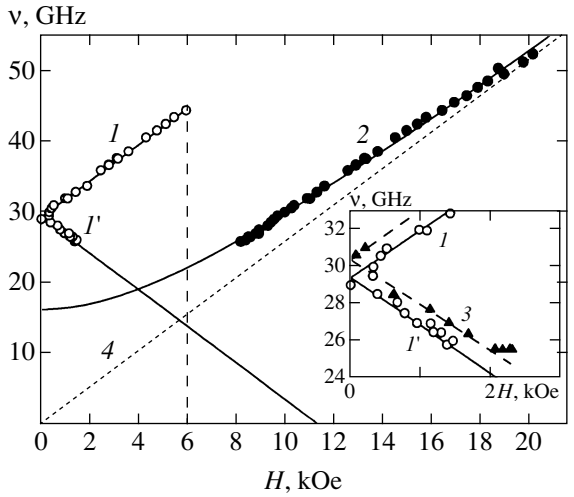



**Fig. 1.** Frequency–field dependence of antiferromagnetic resonance in $GdFe_3(BO_3)_4$ at $T = 4.2$ K and $\mathbf{H} \parallel c$. Line *4* corresponds to the paramagnetic dependence $\nu = \gamma H$. The initial portions of the dependences for (*1*, *1'*) $GdFe_3(BO_3)_4$ and (*3*) $GdFe_{3-x}Ga_x(BO_3)_4$ are shown in the inset.

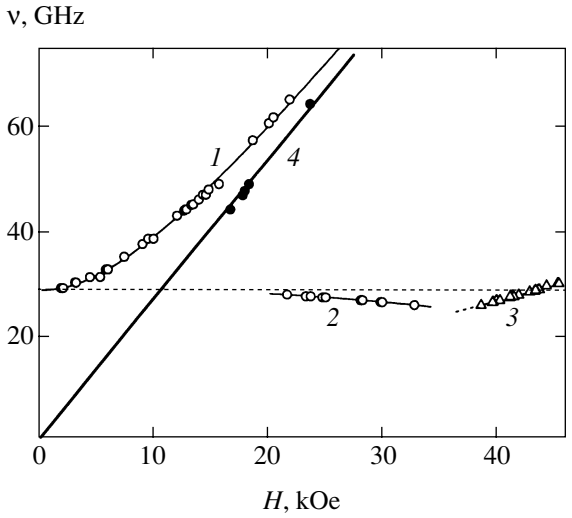



**Fig. 2.** Frequency–field dependences of antiferromagnetic resonance at $T = 4.2$ K, magnetic field is perpendicular to the crystal axis.

were performed for bulk single crystals with well-defined faceting of size up to 2 mm.

Magnetic resonance was studied over the frequency and temperature ranges 25–70 GHz and 4.2–60 K, respectively, using an automated magnetic resonance spectrometer with a pulsed magnetic field [9].

The frequency–field dependence of antiferromagnetic resonance measured at $T = 4.2$ K in a magnetic field directed along the principal crystal axis $c$ is shown in Fig. 1. Two dependence regions can easily be distinguished. At low magnetic fields, we observe two oscillation branches (*1* and *1'*) whose frequencies linearly depend on the field and the line width does not exceed 110 Oe at 37 GHz for the best samples. Both branches have equal initial splittings of about 29 GHz; the frequency of one of them linearly increases as the field grows, and that of the other decreases. In fields higher than $H_c \approx 6.0$ kOe, both magnetic resonance branches disappear (the highest resonance frequency of the upper branch is about 44.4 GHz). In their place, one oscillation branch (*2*) appears; its frequency increases as the field rises, and its splitting is smaller than that of the first two oscillation branches.

The frequency–field dependences of antiferromagnetic resonance measured at 4.2 K for the magnetic field orientation in the basal plane are shown in Fig. 2. We observe two oscillation branches with the same initial splitting as for $H \parallel c$ (Fig. 1). The frequency of one of them (*1*) nonlinearly increases as the field grows stronger. The other branch (*2*) is a decreasing function. Its linewidth measured in a magnetic-field sweep is 6–8 times larger than that for the first branch, and its frequency slowly decreases as the field increases. In strong fields, an additional line is observed, due to a narrow resonance. The corresponding frequency is a slowly increasing function of magnetic field intensity (see curve *3* and triangles in Fig. 2).

The temperature dependences of resonance fields for the magnetic field oriented along the crystal axis were measured at frequencies of 26.11 and 44.48 GHz (Fig. 3). At the lower frequency, we observed resonance absorption corresponding to both the low- (line *1*) and high-field (line *2*) regions. In the low-field region, the resonance field first decreased to zero and then increased as the temperature grew. This resonance disappeared above approximately 8.5 K. The resonance

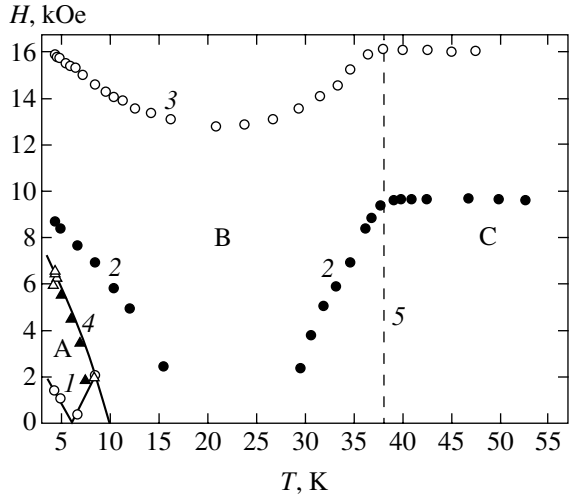



**Fig. 3.** Temperature dependences of resonance and critical fields at $\mathbf{H} \parallel c$ and frequencies of (*1*, *2*) 26.11 and (*3*) 44.48 GHz. State regions: A, easy-axis; B, easy-plane; and C, paramagnetic. Open and solid triangles are the antiferromagnetic resonance and magnetic measurement [8] data, respectively.

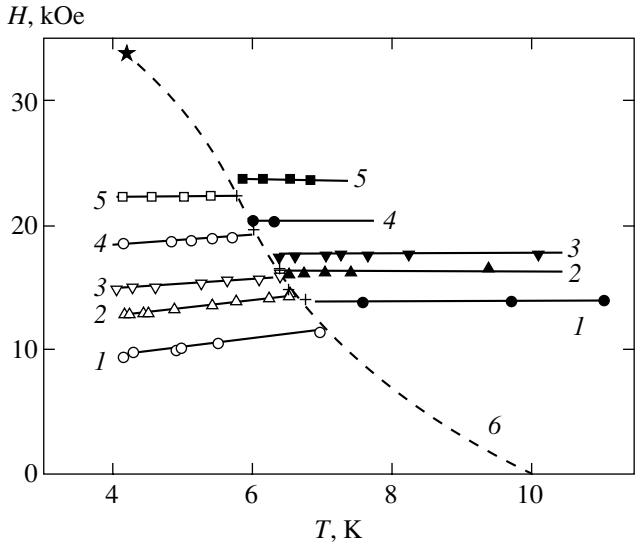



**Fig. 4.** Temperature dependences of resonance and critical fields at $\mathbf{H} \perp c$ and frequencies of (*1*) 38.63, (*2*) 44.32, (*3*) 47.04, (*4*) 57.55, and (*5*) 64.45 GHz. The dashed line *6* was drawn "by eye."

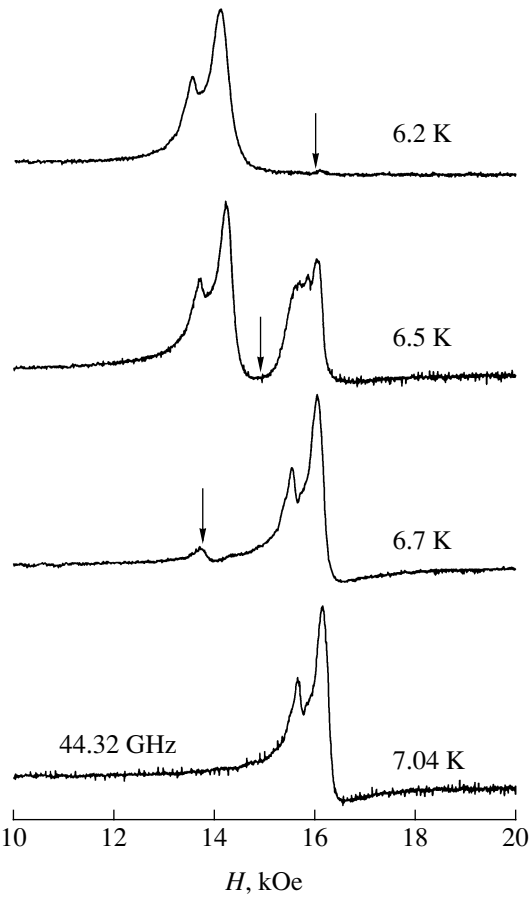



**Fig. 5.** Temperature-induced transformation of the resonance absorption spectrum at a 44.32 GHz frequency in the magnetic field perpendicular to the crystal axis.

field for the high-field resonance branch also sharply decreased to zero as the temperature increased to approximately 18 K. Over the temperature range 18–28 K, no resonance absorption was observed at this frequency. Starting with $T \approx 28$ K, resonance absorption reappeared, and the resonance field increased as the temperature grew and reached a plateau at $T \approx 38$ K. As the 44.48 GHz frequency was slightly higher than the highest frequency of the upper oscillation branch in the low-field resonance region, we only observed resonance absorption corresponding to the high-field branch over the whole temperature range. The temperature dependence of the resonance field was qualitatively similar to that characteristic of the lower frequency, namely, the resonance field first decreased and passed a minimum at about 20 K and then increased and reached a plateau close to 38 K.

The temperature dependence of the resonance parameters was different when the field was oriented in the basal plane of the crystal. The temperature dependences of the resonance fields measured at various frequencies are shown in Fig. 4. The fields first monotonically increase as the temperature of the crystal grows. Next, at a certain temperature, the resonance spectrum sharply changes, the temperature interval of these changes being less than 1 K. The transformation of the resonance spectrum measured at 44.32 GHz is shown in Fig. 5. The spectrum contains one inhomogeneously broadened line corresponding to branch *1* in Fig. 2 at temperatures up to about 6 K (the weak peak on the left wing of the resonance line is in our view an inhomogeneous peak). Additional resonance absorption appears at $T = 6.2$ K in higher magnetic fields above some critical value marked by an arrow in the figure. The critical field decreases as the temperature grows, and, at $T = 6.5$ K, the absorption spectrum contains two absorption lines. The low-field absorption line disappears as the temperature increases further, and only the high-field line remains.

The region of resonance spectrum transformation shifts to lower temperatures as the resonance frequency increases. The temperature dependence of the critical field is shown in Fig. 4 by crosses (line *6*). We also see that the resonance field of the high-field peak is independent of temperature to the right of line *6*. The frequency–field dependence for this crystal state is shown in Fig. 2 (line *4*); it can be approximated by a linear dependence starting at the origin.

## 3. RESULTS AND DISCUSSION

It is reasonable to assume that the magnetic properties of $GdFe_3(BO_3)_4$ are determined by the coexistence of two magnetic subsystems of iron and rare-earth metal ions interrelated by exchange interaction. An analysis of the resonance data obtained at $T = 4.2$ K leads us to conclude that the crystal at low temperatures in the ground state is an antiferromagnet with an easy anisotropy axis parallel to the principal crystal axis.

Indeed, if magnetic field $H_0$ is applied along the crystal axis and is lower than the critical spin flop transition field $H_{sf}$, the field dependences of the resonance frequencies have the form [10, 11]

$$\frac{\nu_{1,2}}{\gamma_\parallel} = \sqrt{(2H_E + H_A)H_A} \pm H_0\left(1 - \frac{\chi_\parallel}{2\chi_\perp}\right), \quad H < H_{sf} = \sqrt{(2H_E - H_A)H_A}. \tag{1}$$

For the spin flop phase, we have

$$\left(\frac{\nu_1}{\gamma_\parallel}\right)^2 = H_0^2\frac{2H_E(2H_E + H_A)}{(2H_E - H_A)^2} - 2H_E H_A, \quad \nu_2 = 0, \quad H > H_{sf}. \tag{2}$$

Here, $H_E$ and $H_A$ are the effective exchange and anisotropy fields, respectively, with respect to the crystal axis $c$ ($H_A > 0$); $\chi_\parallel$ and $\chi_\perp$ are the antiferromagnetic susceptibilities along the principal axis and in the basal plane, respectively; and $\gamma_\parallel$ is the gyromagnetic ratio for the axial magnetic field direction. The solid lines in Fig. 1 are the theoretical dependences constructed according to (1) for oscillation branches *1* and *1'*. The dependence parameters are as follows: $\nu_{\parallel c} = \gamma_\parallel\sqrt{(2H_E - H_A)H_A} = 29.4 \pm 0.2$ GHz and $\gamma_\parallel(1 - \chi_\parallel/2\chi_\perp) = 2.55 \pm 0.05$ MHz/Oe. Using the $\gamma_\parallel = 2.808$ MHz/Oe value obtained from EPR measurements at room temperature, we obtain the susceptibility ratio $\chi_\parallel/2\chi_\perp = 0.081$, which is close to the experimental value 0.083 measured at $T = 4.2$ K [8].

The magnetic field value $H_c = 6.0$ kOe at which the oscillation branches described by (1) for $H < H_{sf}$ disappear at $T = 4.2$ K is close to the characteristic field $H_c = 6.15$ kOe corresponding to the magnetization jump observed in [8] and interpreted as a spin flop transition. However, oscillation branch *2* observed at this magnetic field orientation in fields higher than 6.0 kOe cannot be assigned to a resonance in the spin flop phase of an easy-axis antiferromagnet, primarily because the critical field $H_c$ of the disappearance of the resonance of the easy-axis phase is much lower than the spin flop transition field $H_{sf} = 11.3$ kOe calculated by (1). In addition, the experimental frequency–field dependence in fields higher than 6.0 kOe cannot be approximated by Eq. (2) for the spin flop phase with any reasonable parameter values. At the same time, these data are well described by the dependence characteristic of antiferromagnets with easy-plane anisotropy [11],

$$\left(\frac{\nu_\parallel}{\gamma_\parallel}\right)^2 = 2H_E|H_A'| + H_0^2. \tag{3}$$

Here, $H_A' < 0$ is the anisotropy field in the induced easy-plane state. The solid line in Fig. 1 for oscillation branch *2* corresponds to this equation with an energy gap $\nu_{c2} = \gamma_\parallel\sqrt{2H_E|H_A'|} = 16.5 \pm 0.5$ GHz. It follows that the magnetization jump observed in [8] in the magnetic field $H_c = 6.15$ kOe aligned with the principal axis is the magnetic field-induced spin-reorientation transition from the easy-axis to the easy-plane state rather than the spin flop transition in the easy-axis $GdFe_3(BO_3)_4$ phase. This magnetization jump is then caused by the transition from parallel susceptibility $\chi_\parallel$ to perpendicular susceptibility $\chi_\perp$, which is much larger in magnitude. According to the antiferromagnetic resonance and magnetic [8] measurements, the temperature dependence of the critical field shown in Fig. 3 by triangles is the phase boundary separating these two states on the plane temperature–magnetic field along the $c$ axis.

The conclusion that the low-temperature state of the $GdFe_3(BO_3)_4$ crystal in the region of low fields is an easy-axis state is also substantiated by magnetic resonance in the field oriented in the basal plane. The frequency–field dependence at $H_A \ll H_E$ then takes the form [11]

$$\left(\frac{\nu_{\perp 1}}{\gamma_\perp}\right)^2 = (2H_E + H_A)H_A + \frac{2H_E - H_A}{2H_E + H_A}H_0^2 \approx (2H_E + H_A)H_A + H_0^2, \quad \left(\frac{\nu_{\perp 2}}{\gamma_\perp}\right)^2 = (2H_E + H_A)H_A - \frac{H_A}{2H_E + H_A}H_0^2. \tag{4}$$

The first equation describes oscillation branch *1* (Fig. 2), whose frequency increases as the field grows. Solid line *1* in Fig. 2 is the theoretical dependence of $\nu_{\perp 1}$ with the parameters $\nu_{\perp c} = \gamma_\perp\sqrt{(2H_E + H_A)H_A} = 29.0 \pm 0.2$ GHz and $\gamma_\perp = 2.66$ MHz/Oe; the $\nu_{\perp c}$ value coincides with $\nu_{\parallel c}$ for $H \parallel c$.

We found that a magnetic field applied in the basal plane of the crystal can also induce the transition to the easy-plane state. In this state, the frequency–field dependences of antiferromagnetic resonance for a magnetic field in the basal plane take the form [11]

$$\frac{\nu_{\perp 1}}{\gamma_\perp} = H_0\sqrt{1 + \frac{|H_A'|}{2H_E}} \approx H_0, \quad \left(\frac{\nu_{\perp 2}}{\gamma_\perp}\right)^2 = 2H_E|H_A'| - \frac{|H_A'|}{2H_E}H_0^2, \tag{5}$$

where magnetic anisotropy in this plane, which is weak according to the antiferromagnetic resonance data, is ignored.

The frequency–field dependences for the $\nu_{\perp 1}$ branches in the easy-plane and easy-axis states are

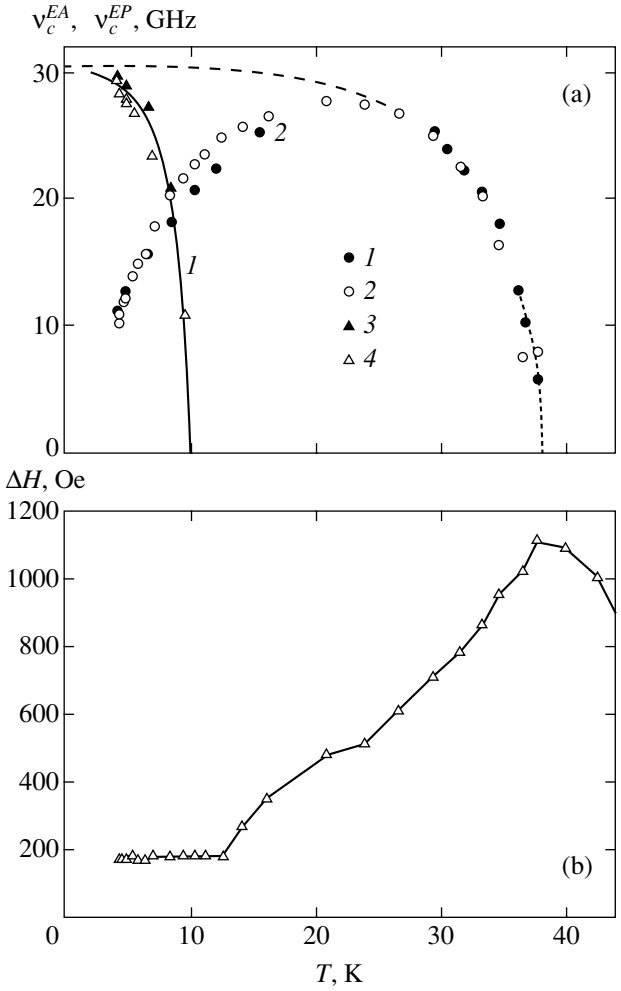



**Fig. 6.** (a) Temperature dependence of the energy gaps in the spectra of the (*1*) easy-axis and (*2*) easy-plane states. Frequencies and magnetic field orientations: (*1*) 26.11 GHz, **H** ∥ *c*; (*2*) 44.48 GHz, **H** ∥ *c*; (*3*) 26.11 GHz, **H** ∥ *c*; and (*4*) 38.63 GHz, **H** ⊥ *c*. (b) Temperature dependence of antiferromagnetic resonance line width at a 44.48 GHz and **H** ∥ c.

sharply different. Because of the presence of a gap in the spectrum of oscillations, the resonance field in the easy-axis state is always lower than in the state with an easy-plane anisotropy. The transformation of the resonance spectrum caused by heating the crystal (see Fig. 5) is just the transition between the $\nu_{\perp 1}$ oscillation branches of the easy-axis and easy-plane states.

It follows that the temperature dependence of the critical field shown in Fig. 4 is the phase boundary between the states with easy-axis and easy-plane anisotropies on the plane temperature–magnetic field in the basal plane. The temperature dependence of resonance fields for the $\nu_{\perp 1}$ branch in the easy-axis state of the crystal is caused by the temperature dependence of the energy gap in spectrum (4). In the easy-plane state, the resonance field for $\nu_{\perp 1}$ is temperature-independent according to (5). Resonance branch *2* in Fig. 2 corresponds to just this oscillation branch of the induced easy-plane phase.

Using (1) and (4), we can calculate the temperature dependences of the energy gaps $\nu_c^{EA}$ and $\nu_c^{EP}$ for the easy-axis and easy-plane states, respectively, from the temperature dependences of resonance fields (Fig. 3). The results are shown in Fig. 6a, which also contains the temperature dependence of $\nu_c^{EA}$ obtained from the temperature dependence of resonance absorption measured at 38.63 GHz for field $H \perp c$. These results closely agree with those obtained for $H \parallel c$. The energy gap $\nu_c^{EA}$ (curve *1*) tends to zero as the temperature approaches $T = 10$ K. It follows that the uniaxial anisotropy field $H_A$ does indeed change sign at this temperature, and a spontaneous spin-reorientation transition occurs. The anomaly of magnetic properties close to 10 K observed in [8] is related to just this phase transition.

The temperature dependence of the energy gap $\nu_c^{EP}$ (Fig. 6a, curve *2*) has an unusual form. The gap first increases as the temperature grows, reaches a plateau at about 20 K, and then decreases to zero as the temperature approaches $T = 38$ K. Apparently, this is the temperature of magnetic ordering of the subsystem of iron ions, because spontaneous ordering of the subsystem of rare-earth metal ions usually occurs at much lower temperatures [1, 12]. The boundary between the paramagnetic state of the crystal and the ordered state of the iron subsystem is shown by dashed line *5* in the phase diagram (Fig. 3). The ground state of the iron subsystem is antiferromagnetic, and its anisotropy is of the easy-plane type. An increase in the energy gap as the temperature decreases from $T_N = 38$ K is typical of antiferromagnets. It is caused by an increase in the effective exchange and anisotropy fields as the crystal is cooled. In the absence of a rare-earth metal subsystem, the temperature dependence of the energy gap would reach a plateau below approximately 15–20 K, and the gap would remain virtually unchanged during cooling to 0 K, as is shown by the dashed line in the figure. It is reasonable to suggest that a decrease in the energy gap starting with about 20 K is related to the influence of the gadolinium subsystem. It is likely that this subsystem has the opposite anisotropy sign, its contribution to the total anisotropy increases in magnitude as the temperature lowers, and spontaneous reorientation from the easy-plane to the easy-axis state occurs at $T = 10$ K.

As concerns the supposed state of the gadolinium subsystem in this temperature region, it cannot be entirely ruled out that this subsystem experiences ordering precisely at about $T = 20$ K. Apart from the strong line broadening as the iron subsystem approaches the Néel temperature $T_N = 38$ K, we observe an obvious broadening in the region of 20 K, see the temperature dependence of line width (Fig. 6b) measured at a frequency of 44.48 GHz for $H \parallel c$. This broad-

ening can be caused by spontaneous ordering in the gadolinium subsystem. However, no magnetic property anomalies were detected at this temperature [8]. In addition, an analysis of the crystallographic structure (see below) leads us to suggest that indirect exchange interaction in the gadolinium subsystem is fairly weak. Lastly, it should be mentioned that, according to magnetic measurements, spontaneous ordering of the gadolinium subsystem in the $GdGa_3(BO_3)_4$ crystal, where all iron ions are replaced by diamagnetic ions, does not occur even at $T = 4.2$ K [13]. For this reason, the gadolinium subsystem should rather be treated as polarized by a biasing field created as a result of $Fe^{3+}$–$Gd^{3+}$ exchange interaction over the whole temperature range of magnetic order studied, 4.2–38 K.

The suggestion of the polarization of the rare-earth metal subsystem and its influence on crystallographic anisotropy at low temperatures is substantiated by the results obtained in analyzing the gadolinium ferroborate crystal structure and exchange interactions. The crystal structure of rare-earth metal ferroborates is rhombohedral, space group $R32$ at room temperature [5, 6]. The lattice parameters are $a = 9.567(3)$ Å and $c = 7.578(2)$ Å [5], and the unit cell contains three $GdFe_3(BO_3)_4$ molecules. The $Fe^{3+}$ ions have an oxygen environment in the form of slightly distorted octahedra. Neighboring octahedra share edges to produce a helicoidal chain along the $c$ axis. The oxygen environment of $Gd^{3+}$ is a distorted triangular prism sharing vertex oxygen atoms with Fe–$O_6$ octahedra of three neighboring helicoidal chains. A fragment of the crystal structure of $GdFe_3(BO_3)_4$ is shown in Fig. 7a, where two chains of Fe–$O_6$ octahedra and two adjacent Gd–$O_6$ prisms are shown. Two neighboring iron ions in a chain are coupled by indirect Fe–O–Fe exchange interaction through two oxygen ions with bond angles of 101.10° and 103.43°. The estimates obtained using the simple model of exchange coupling [14] show that exchange interactions between iron ions in chains are antiferromagnetic and the exchange integral value is $J^{Fe\text{–}Fe} \approx -9$ K. Neighboring chains interact with each other through the Fe–O–Gd–O–Fe bonds and B–$O_3$ complexes shown as triangles in the figure. Judging from the absence of a broad maximum of the temperature dependence of magnetic susceptibility characteristic of low-dimensional magnetism [8], these interchain interactions are sufficiently strong for the establishment of three-dimensional magnetic order in the iron subsystem.

If the structure of gadolinium ferroborate is treated as planes containing iron and gadolinium ions that alternate along the $c$ axis, then each gadolinium ion is coupled by indirect Gd–O–Fe interactions with iron ions of two neighboring planes that belong to one sublattice and is not coupled with iron ions of its own plane that form another sublattice. According to estimates obtained by following [14], Gd–O–Fe exchange interactions that polarize the gadolinium subsystem at $T < T_N$ are also antiferromagnetic and are weaker than exchange interactions in the iron subsystem. These interactions are weak because the distance between the gadolinium and oxygen ions $R_{Gd\text{–}O} = 2.4$ Å in this chain is much larger than the Fe–O distances, $R_{Fe\text{–}O} = 1.99$–$2.04$ Å. It is likely, however, that polarizing exchange coupling between the iron and rare-earth metal ions in the hantite structure, that is, in $GdFe_3(BO_3)_4$, is stronger than similar coupling in rare-earth metal orthoferrites, where rare-earth metal ions interact with iron ions from different sublattices, which results in almost complete isotropic exchange balancing [1, 15, 16].

The gadolinium ions are in turn coupled with each other only through B–$O_3$ complexes. Both arms in the Gd–$BO_3$–Gd interaction chain are fairly long, $R_{Gd\text{–}O} = 2.4$ Å. It appears that indirect exchange coupling in the gadolinium subsystem is for this reason weaker than the polarizing action of the iron subsystem, which explains the absence of spontaneous magnetic ordering in the gadolinium subsystem of $GdGa_3(BO_3)_4$ even at $T = 4.2$ K.

A consideration of all these special features leads us to suggest that the $GdFe_3(BO_3)_4$ crystal at low temperatures has the magnetic structure shown in Fig. 7b. This structure consists of planes perpendicular to the $c$ axis and alternating along it. The planes contain ferromagnetically ordered iron and gadolinium ions. Neighboring planes are ordered antiferromagnetically. The size of the magnetic unit cell is doubled along the $c$ axis compared with the crystal lattice cell.

It follows that the magnetic anisotropic properties of gadolinium ferroborate at low temperatures are formed in the competition between anisotropic interactions of iron and gadolinium ion subsystems, which have different anisotropy signs. The iron subsystem is a collinear two-sublattice antiferromagnetic subsystem spontaneously ordered at $T < T_N = 38$ K. The gadolinium subsystem is polarized by antiferromagnetic interactions with the iron subsystem and can also be divided into two sublattices.

This view on the magnetic structure of gadolinium ferroborate allows us to write the energy of this crystal in the presence of a magnetic field as

$$U = J_{11}\mathbf{M}_1 \cdot \mathbf{M}_2 + J_{12}(\mathbf{M}_1 \cdot \mathbf{m}_1 + \mathbf{M}_2 \cdot \mathbf{m}_2) - \frac{K_1}{2M_0^2}(M_{1z}^2 + M_{2z}^2) - \frac{K_2}{2m_0^2}(m_{1z}^2 + m_{2z}^2) - \mathbf{H} \cdot (\mathbf{M}_1 + \mathbf{M}_2 + \mathbf{m}_1 + \mathbf{m}_2). \quad (6)$$

Here, $J_1 > 0$ and $J_{12} > 0$ are the exchange coupling parameters in the iron subsystem and between the iron and gadolinium subsystems, respectively; $\mathbf{M}_{1,2}$ and $\mathbf{m}_{1,2}$ are the magnetic moments of the iron and gadolinium subsystem sublattices, $|\mathbf{M}_1| = |\mathbf{M}_2| = M_0$ and $|\mathbf{m}_1| = |\mathbf{m}_2| = m_0$; and $K_1$ and $K_2$ are the uniaxial anisotropy

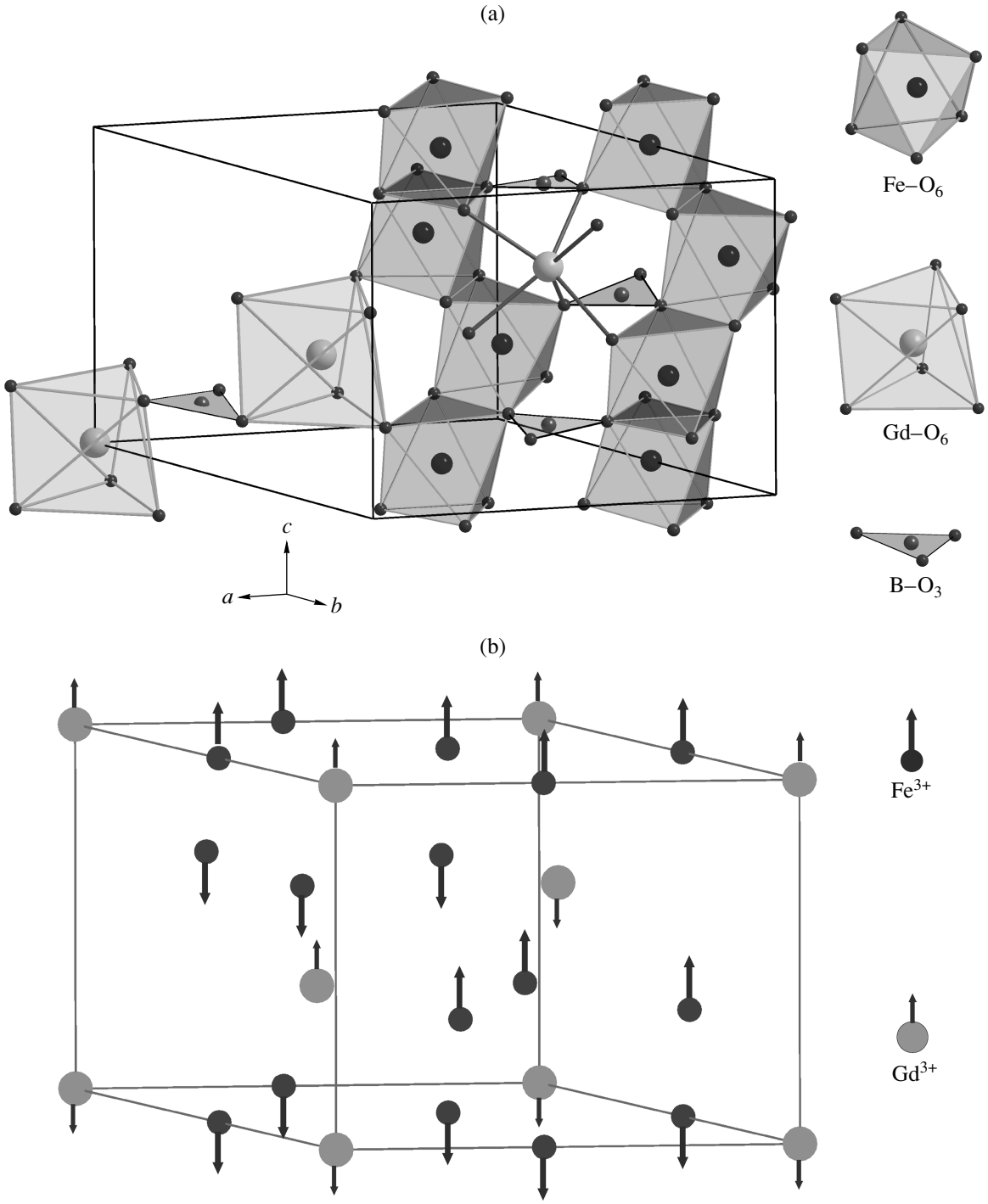



**Fig. 7.** (a) Fragment of the crystal structure of $GdFe_3(BO_3)_4$ and (b) supposed magnetic structure of $GdFe_3(BO_3)_4$ at $T < 10$ K.

constants of the iron and gadolinium subsystems. It follows from the experimental data that $K_1 < 0$ and $K_2 > 0$. Exchange coupling in the gadolinium subsystem and anisotropy in the basal plane are ignored. In addition, the second energy term is written on the assumption that the ions of each iron sublattice only interact with ions of one gadolinium sublattice. The minimization of energy (6) in the absence of a magnetic field gives the following solutions for polar angles $\theta_i$ and $\eta_i$ and sublattice magnetic moments $\mathbf{M}_i$ and $\mathbf{m}_i$:

1) $\theta_1 = 0$, $\theta_2 = \pi$, $\eta_1 = \pi$, $\eta_2 = 0$;

2) $\theta_1 = \pi/2$, $\theta_2 = -\pi/2$, $\eta_1 = -\pi/2$, $\eta_2 = \pi/2$;

3) $\theta_1 = \theta$, $\theta_2 = \pi + \theta$, $\eta_1 = \pi + \eta$, $\eta_2 = \eta$.

Generally, θ and η are not equal to each other and to 0 and π/2.

The last solution describes the angular phase in the narrow region $K_2 \approx |K_1|$, more exactly, in the region bounded by the inequalities

$$K_1' < K_2 < K_1'',$$

$$K_1' = |K_1|\frac{\cos^2\theta}{\cos^2\eta} + 2J_{12}mM\frac{1-\cos(\theta-\eta)}{\cos^2\eta}, \quad (7)$$

$$K_1'' = |K_1|\frac{\sin^2\theta}{\sin^2\eta} - 2J_{12}mM\frac{1-\cos(\theta-\eta)}{\sin^2\eta}.$$

Outside this region, we have the state with an easy-axis anisotropy if $K_2 > |K_1|$ [solution (1)] or the state with an easy-plane anisotropy [solution (2)].

It follows that the sequence of changes in the magnetic state of gadolinium ferroborate as the temperature lowers in the absence of a magnetic field can be explained as follows. The crystal is a two-sublattice antiferromagnet with an easy-plane anisotropy in the region of 20–38 K, where only the subsystem of iron ions is predominantly ordered. The degree of the polarization of the gadolinium subsystem with the opposite anisotropy sign increases as the temperature lowers; accordingly, the contribution of this subsystem to the total crystal anisotropy grows. This contribution becomes noticeable at temperatures lower than 20 K, and, at $T = 10$ K, the total anisotropy energy changes sign. The anisotropy fields $H_A$ and $H_A'$ present in equations (1)–(5) for antiferromagnetic resonance frequencies are resultant anisotropy fields at temperatures below 20 K; they are determined by the contributions of both magnetic subsystems. The magnetic structure of $GdFe_3(BO_3)_4$ shown in Fig. 7b corresponds to low temperatures.

This model can be used to estimate the anisotropy fields of both magnetic subsystems of gadolinium ferroborate at $T = 4.2$ K. Using the experimental exchange field value $H_E = 180$ kOe obtained from perpendicular susceptibility at $T = 4.2$ K [8] and the energy gap in the spectrum extrapolated to low temperatures $\nu_c \approx 30$ GHz for the iron subsystem (see Fig. 6a), we obtain the anisotropy field for the iron subsystem at low temperatures $H_A^{Fe} \approx -320$ Oe. The experimental energy gap 29.4 GHz value for the easy-axis crystal state at $T = 4.2$ K, which is determined by the resultant anisotropy field $H_A \approx 310$ Oe, can be used to estimate the anisotropy field for the gadolinium subsystem at the helium temperature, $H_A^{Gd} \approx 630$ Oe.

A magnetic field changes the orientations of the magnetic moments of the sublattices of both subsystems, primarily the gadolinium subsystem coupled by comparatively weak exchange interaction with the iron subsystem. Moreover, further calculations show that, in very strong magnetic fields, the gadolinium subsystem experiences the transition from the polarized antiferromagnetic to the "collapsed" spin-flip state. This changes the ratio between the contributions of the iron and gadolinium subsystems to the total anisotropy. As a result, the temperature of the orientation transition should also change as the magnetic field increases, and the very transition between the easy-axis and easy-plane states can occur under changes not only in temperature but also in magnetic field. The dependences of the equilibrium polar angles $\theta_1$ and $\eta_2$ on the magnetic field along axis $c$ obtained by numerically minimizing crystal energy (6) are shown in Fig. 8. The minimization was performed using the following potential parameter values:

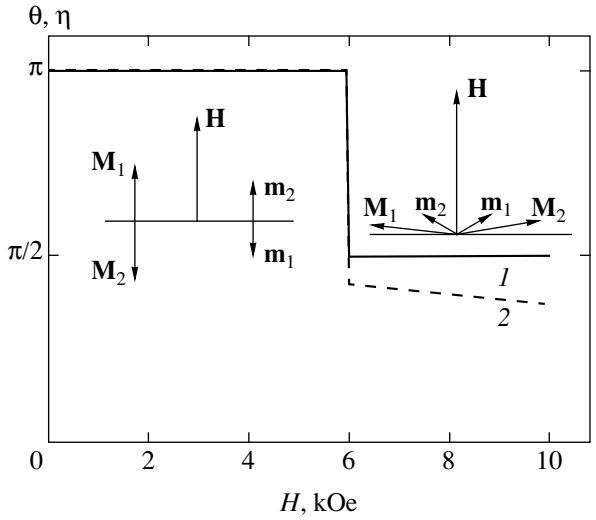



**Fig. 8.** Field dependence of equilibrium polar angles of the (*1*) iron and (*2*) gadolinium sublattice magnetic moments.

$$J_1 = 2400, \quad J_{12} = 320, \quad K_1 = -24000 \text{ erg/cm}^3,$$

$$K_2 = 62000 \text{ erg/cm}^3, \quad M_0 = 75 \text{ G cm}^3/\text{g},$$

$$m_0 = 25 \text{ G cm}^3/\text{g}.$$

The exchange parameter $J_1 = H_E/M_0$ and the anisotropy constant $K_1 = H_A^{Fe}M_0$ were calculated from the exchange field $H_E = 180$ kOe and the anisotropy field of the iron subsystem $H_A^{Fe} \approx -320$ Oe given above, and the $J_{12}$, $K_2$, and $m_0$ values played the role of adjustment parameters. The magnetization of the gadolinium subsystem sublattice was taken to be smaller than the saturation value $m_0 = 35$ G cm$^3$/g. Because of weak exchange interactions with the iron subsystem, the gadolinium subsystem was considered unsaturated at $T = 4.2$ K (this is the temperature at which the calculation results were compared with the experimental data). For

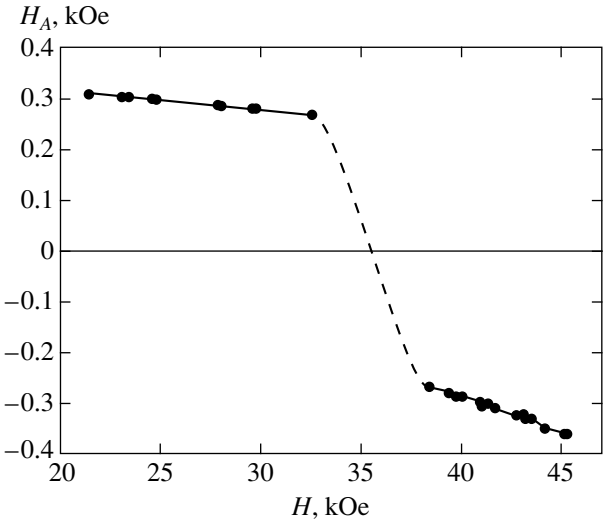



**Fig. 9.** Field dependence of the resultant anisotropy field for $H \perp c$ at $T = 4.2$ K.

the same reason, the magnetic moments of the sublattices in the gadolinium subsystem "collapse" in a magnetic field more rapidly than the iron sublattice moments, see Fig. 8.

Calculations show that, at this set of parameters, the transition from the easy-axis to the easy-plane state occurs in a magnetic field of $H_c = 5.92$ kOe. This value is close to the experimental critical field values obtained from antiferromagnetic resonance and magnetic measurements [8]. Also note that, according to our calculations, magnetic moment reorientation occurs in a jump, without the formation of an angular phase.

If the magnetic field is directed in the basal plane of the crystal, field-induced spin reorientation occurs via an angular phase at the same thermodynamic potential parameters. This phase exists in the interval of magnetic fields from 30.2 to 33.8 kOe (the larger value is labeled by an asterisk in Fig. 4). Within this interval, which is in agreement with the antiferromagnetic resonance data, the antiferromagnetic vectors of both magnetic subsystems rotate in the plane perpendicular to the magnetic field. According to the magnetic measurements [8, 13], this transition extends from 30.8 to 33.5 kOe at $T = 4.2$ K and has no field hysteresis, which also closely agrees with the results of our calculations.

The suggestion of the field dependence of the resultant anisotropy field of gadolinium ferroborate allows the unusual character of frequency–field dependences *2* and *3* (Fig. 2) to be explained. As they are situated above and below the critical spin reorientation field, it is reasonable to suggest that these dependences correspond to oscillations $\nu_{\perp 2}$ of the easy-axis and field-induced easy-plane states. A comparison of Eqs. (4) and (5) shows that the oscillation frequencies $\nu_{\perp 2}$ in both states virtually identically depend on the magnetic field because $|H_A| \ll H_E$. Using the exchange field $H_E \approx 180$ kOe and the experimental energy gap value, 29 GHz, we find that the $H_A/2H_E$ ratio, which determines the steepness of the frequency–field dependence for this branch, is about $2 \times 10^{-4}$. This value was used in Fig. 2 to plot the frequency–field dependence of $\nu_{\perp 2}$ [Eq. (4)] for the easy-axis state (the dashed line). On the scale of the figure, this branch can be considered magnetic-field–independent. As concerns the experimentally observed fairly strong magnetic field dependence of resonance frequencies, it probably originates from changes in the energy gaps in the spectrum caused by the field dependence of the resultant anisotropy and determined by the competing contributions of the iron and gadolinium subsystems. The field dependence of the resultant uniaxial anisotropy field calculated on the basis of these considerations is shown in Fig. 9.

Also note that the suggestion of competitive contributions of the iron and gadolinium subsystems to the uniaxial anisotropy of $GdFe_3(BO_3)_4$ is qualitatively substantiated by the magnetic resonance data on $GdFe_{3-x}Ga_x(BO_3)_4$ ($x \approx 1$) crystals. Partial substitution of iron ions by diamagnetic gallium ions decreases the contribution of the iron subsystem to anisotropy. For this reason, the total anisotropy and energy gap at 4.2 K increase compared with pure $GdFe_3(BO_3)_4$ (see inset to Fig. 1), and the critical field of the transition to the easy-plane state for $\mathbf{H} \parallel c$ increases from 6.0 to 9.3 kOe at $T = 4.2$ K.

## 4. CONCLUSIONS

To summarize, we thoroughly studied antiferromagnetic resonance and magnetic phase transitions in gadolinium ferroborate $GdFe_3(BO_3)_4$. This is the first such study for rare-earth metal magnets with hantite structures. An analysis of the experimental data led us to the following conclusions on the magnetic structure of this compound.

Ordering of the subsystem of iron ions, which is a two-sublattice antiferromagnet with easy-plane anisotropy, occurs at the Néel temperature $T_N = 38$ K. In our view, the influence of the anisotropy of the gadolinium subsystem polarized by exchange interaction with the iron subsystem becomes noticeable as the temperature decreases to 20 K. The gadolinium subsystem can also be represented in the form of a two-sublattice antiferromagnet with easy-axis anisotropy at temperatures below 20 K. The competition of the anisotropic contributions of the iron and gadolinium subsystems results in a spontaneous transition from the easy-plane to the easy-axis state as the temperature decreases to $T = 10$ K. An analysis of the crystal structure and exchange interactions at temperatures below 20 K led us to suggest the magnetic structure of $GdFe_3(BO_3)_4$ comprising planes that contain ferromagnetically ordered iron and gadolinium ions, are perpendicular to the $c$ axis, and alternate along this axis. The neighboring planes are

ordered antiferromagnetically. The size of the magnetic unit cell along the *c* axis equals two times the size of the crystallographic unit cell.

It follows from the resonance data that transitions between the easy-axis and easy-plane states occur not only depending on the temperature but also as the magnetic field changes. We constructed the experimental magnetic phase diagrams for magnetic fields oriented along the crystal axis and in the basal plane.

A simple model was suggested to describe spontaneous and induced phase transitions. The model takes into account antiferromagnetic exchange interactions within the iron subsystem and between the iron and gadolinium subsystems, the anisotropy energies of both subsystems, and the Zeeman energy. This model was used to find the conditions of the existence of the easy-axis, easy-plane, and angular phases. Calculations were performed to determine the critical fields of spin reorientation transitions in magnetic fields along the crystal axis and in the basal plane. The results were in close agreement with the experimental data obtained at $T = 4.2$ K.

## ACKNOWLEDGMENTS

The authors thank M.A. Popov for useful discussions.

## REFERENCES

1. K. P. Belov, A. K. Zvezdin, A. M. Kadomtseva, and R. Z. Levitin, *Reorientational Transitions in Rare-Earth Magnets* (Nauka, Moscow, 1979) [in Russian].
2. A. D. Mills, Inorg. Chem. **1**, 960 (1962).
3. G. Blasse and A. Bril, Phys. Status Solidi **20**, 551 (1967).
4. V. I. Chani, M. I. Timoshechkin, K. Inoue, *et al.*, Inorg. Mater. **30**, 1466 (1992).
5. N. I. Leonyuk and L. I. Leonyuk, Prog. Cryst. Growth Charact. Mater. **31**, 179 (1995).
6. J. A. Campa, C. Cascales, E. Guttierez-Puebla, *et al.*, Chem. Mater. **9**, 237 (1997).
7. Y. Hinatsu, Y. Doi, K. Ito, *et al.*, J. Solid State Chem. **172**, 438 (2003).
8. A. D. Balaev, L. N. Bezmaternykh, I. A. Gudim, *et al.*, J. Magn. Magn. Mater. **258–259**, 532 (2003).
9. V. I. Tugarinov, I. Ya. Makievskiĭ, and A. I. Pankrats, Prib. Tekh. Éksp., No. 4 (2004) [Instrum. Exp. Tech. **47**, 472 (2004)].
10. I. S. Jackobs, R. A. Beyerline, S. Foner, and J. P. Remeika, Int. J. Magn. **1**, 193 (1971).
11. A. G. Gurevich, *Magnetic Resonance in Ferrites and Antiferromagnets* (Nauka, Moscow, 1973) [in Russian].
12. A. K. Zvezdin, V. M. Matveev, A. A. Mukhin, and A. I. Popov, *Rare-Earth Ions in Magnetic-Ordered Crystals* (Nauka, Moscow, 1985) [in Russian].
13. A. D. Balaev, private communication.
14. O. A. Bayukov and A. F. Savitskiĭ, Fiz. Tverd. Tela (St. Petersburg) **36**, 1923 (1994) [Phys. Solid State **36**, 1049 (1994)].
15. D. V. Belov, A. K. Zvezdin, A. M. Kadomtseva, *et al.*, Fiz. Tverd. Tela (Leningrad) **23**, 2831 (1981) [Sov. Phys. Solid State **23**, 1654 (1981)].
16. J. D. Cashion, A. H. Cooke, D. M. Martin, and M. R. Wells, J. Phys. C: Solid State Phys. **3**, 1612 (1970).

*Translated by V. Sipachev*

*Journal of Experimental and Theoretical Physics, Vol. 99, No. 4, 2004, pp. 776–783.*
*From Zhurnal Éksperimental'noĭ i Teoreticheskoĭ Fiziki, Vol. 126, No. 4, 2004, pp. 898–906.*


**SOLIDS**
**Structure**

# CARS Study of Deuterium Clusters Stabilized in Solid Helium¶

**E. B. Gordon[a], T. Kumada[b], M. Ishiguro[b], and Ya. Aratono[b]**

[a]*Institute of Problems of Chemical Physics, Russian Academy of Sciences, Chernogolovka, Moscow oblast, 142432 Russia*
*e-mail: gordon@binep.ac.ru*

[b]*Advanced Science Research Center, Japan Atomic Energy Research Institute, 319-1195, Tokai, Ibaraki, Japan*

Received April 27, 2004

**Abstract**—Solid deuterium clusters that for the first time have been isolated in a matrix of solid helium have been investigated at $T = 1.3$ K and $P = 3$ MPa by the coherent anti-Stokes Raman spectroscopy (CARS) technique. The vibronic $Q_1(J = 0)$ and $Q_1(J = 1)$ line intensity, shape, and positions have been studied as functions of *ortho* and *para* content in the solid, as well as of the size of clusters. The strong effect of Raman scattering cross-section sensitivity to the molecular environment nuclear spin state has been found in CARS: the ratio of probabilities for the scattering by *para* ($J = 1$) and *ortho* ($J = 0$) deuterium, which is equal to 1 in a gas, is as high as 10000 in nearly pure *ortho* deuterium, whereas it is about 50 in spontaneous Raman scattering. This effect has been shown to occur starting from a cluster size corresponding to the onset of the phonon band.


## 1. INTRODUCTION

The most inert matrix ever applied in matrix isolation spectroscopy is *para* hydrogen [1]; only helium should be more inert than hydrogen in the symmetric $J = 0$ state ($J$ is the rotational number). Because helium cannot be solid at ambient pressure, the conventional technique of matrix isolation spectroscopy is inapplicable. Instead, spectroscopic studies of molecules and clusters surrounded by helium have been performed by the very interesting technique of their capture into cold liquid helium droplets. Although the molecule densities in the droplet experiments do not exceed those in the gas phase, the result of a photoinduced process could be detected with very high sensitivity by the depletion technique: the energy of visible and even infrared light photons released in a droplet induces evaporation of several thousand helium atoms, which is readily detectable if the droplet is sufficiently small. Although the finite size of a droplet, of course, causes some small distortion in the spectra observed, the approach makes it possible to obtain very interesting information, mainly on quantum helium liquid interaction with microinclusions [2].

Meanwhile, the argument underlying the principal suggestions in [2], in particular, the effect of $^4$He superfluidity on the rotation of molecules and clusters embedded in liquid helium has been restricted by the experimental impossibility of varying the matrix temperature and pressure in the framework of the droplet technique. The simplest solution would be experiments at high pressure in a helium crystal doped by the species under study; however, this has been widely accepted as impossible.

Nevertheless, a method of an impurity embedded in solid helium has been recently created [3, 4]. Among its first applications was the problem of matrix isolation spectroscopy of hydrogen; being the best material for isolation, hydrogen itself could not to this point have been placed in a matrix more inert than its own. Because of the high mobility of hydrogen molecules in solid helium [5], only molecular clusters can be stabilized in a helium crystal.

There is a very pronounced effect in the low-temperature vibrational $v = 0 \longrightarrow v = 1$ Raman spectra of condensed molecular hydrogen, which is especially expressed for deuterium [6]: the ratio of the Raman scattering cross sections corresponding to the $Q_1(J = 1)$ and $Q_1(J = 0)$ transitions, which is equal to 1 in the gas phase, has been found to strongly depend on the concentration of molecules in the $J = 1$ state in condensed phases; this ratio varies from 5 for nearly pure *para* ($J = 1$) deuterium ($p$-$D_2$) to 50 for nearly pure *ortho* deuterium ($o$-$D_2$) ($J = 0$). This unexpectedly strong cooperative effect has been reliably explained for solid hydrogen and deuterium in [6] to be caused by the delocalization of the $v = 1$ impurity states in a crystal. However, at least two questions remain unanswered. First, how does this effect manifest itself in stimulated Raman scattering (while in spontaneous Raman scattering the transition probability is proportional to the

¶ This article was submitted by authors in English.

square of the polarizability, the probability of coherent anti-Stokes Raman scattering (CARS) is proportional to the square of the third-order susceptibility). Second starting from what size of the microcrystal does the ratio of the cross sections become large?

Although the size effect in Raman scattering by solid hydrogen should obviously exist and should be strong, to our knowledge, nobody has discussed it, possibly due to the experimental impossibility of isolating small hydrogen clusters in a sufficiently inert matrix. Because the same effect has also been observed for liquid hydrogen and deuterium [7], it should not be too sensitive to crystal structure faults. Two cluster sizes may therefore be critical for the effect: first, when the phonon band appears (it corresponds to about ten layers of molecules in a crystallite and to the number of molecules in it about $N \approx 10^3$), and second, when the crystallite size becomes comparable to the scattered light wavelength, i.e., when $N \approx 10^9$ (an ideal infinite crystal was considered in the calculations in [6]).

## 2. TECHNIQUES

Unlike in the case of heavier rare gases commonly used for matrix isolation, the introduction of impurities into solid helium represents a very sophisticated problem. Both liquid and solid helium are self-purified species because the solubility in them of any other substances is negligibly small. Epitaxial growth of a doped crystal from the gas phase is also impossible because helium is the only substance possessing no triple point, nor even a coexistence of gas and solid phases in equilibrium. Moreover, due to the high plasticity of solid helium, a noticeable pressure gradient does not exist in it and the position of the liquid–solid interface in helium is governed only by the temperature profile; hence, the growth of a crystal from a liquid layer, which should be accompanied by the interface motion, is impossible. Only by using laser ablation from a target placed inside solid helium can the metallic clusters be stabilized, trapped in solid helium in the very vicinity of a target [8, 9].

The technique that we use allows the creation of large doped helium crystals with impurities uniformly distributed in the bulk. The permanent growth of the crystal from its upper edge has been carried out under continuous moving down of the crystal body at the expense of helium exhaust from the bottom of the crystal; due to a small amount of friction of helium solid on a cell wall, even a very small pressure gradient is already sufficient for this motion. As a result, helium gas from the gas supply system enters the cell to restore the former crystal surface position by condensation. Thus, although neither the position nor the shape of the crystal do not visually change in time, its contents permanently move down, with the velocity determined by the helium mass outflow.

As shown in Fig. 1, the experimental setup consists of an optical cell submerged into a liquid helium bath of an optical helium cryostat and a stainless steel gas-handling system. The body of the optical cell is a cylindrical sapphire tube (outer diameter 20 mm, inner diameter 17 mm, and length 80 mm) placed between indium-sealed brass flanges connected with the gas source at the top and with the outlet tube at the bottom. Sapphire was chosen because it simultaneously possesses the necessary strength for high-pressure operation, optical transparency, and a reasonable thermal conductivity at low temperatures. The gas source represents two coaxial thin-wall stainless tubes with outer diameters of 10 and 4 mm, and the inner diameter had an orifice of 0.25 mm in diameter at its bottom to form the gas jet entering the optical cell. For thermal insulation, the space between the tubes was pumped out and the lower end of the tubes was covered with a teflon cap with an orifice at the center to allow passage of the jet. The heater was wired and glued to the surface of the inner tube along its entire length, inserted into the cryostat to keep the temperature in the tube well above the dew point of dopant $D_2$; the temperature near the orifice was measured by a Cu-constantan thermocouple. Several layers of aluminum mesh were interposed at the bottom of the cell to provide uniformity of the outflow and to create a pressure gradient between solid helium in the optical cell and liquid helium in the outlet tube. To stabilize the position of the helium crystal's upper interface, it was useful to insert a constantan heater inside the outlet tube (heating prevents solidification of helium in the tube, violating the stability of the outflow). The cell as a whole can be moved up and down by Wilson sealing at the cap of the cryostat, allowing study of the whole height of the grown sample.

Deuterium with enhanced abundance of $o$-$D_2$ was prepared by mixing *normal* deuterium ($n$-$D_2$) with $o$-$D_2$ produced in the cryocooler (Nagase & Co. Ltd. Model UV204SC) filled with a ferric oxide (FeO(OH)) catalyst at a temperature close to the $D_2$ triple point. Deuterium was then premixed with helium at a ratio of 1 : 500 ~ 1000 in a stainless steel cylinder at a pressure of 6.0 MPa.

The procedure of sample preparation was as follows. During the pouring of liquid helium into the optical cryostat, the pure helium gas from the outer high-pressure cylinder was allowed to condensate in the cell at the pressure of about 1 MPa up to filling of the entire cell by liquid helium and subsequent inflow termination; the cylinder was kept connected to the cell after that. Then, the main helium bath of the cryostat was cooled by pumping with booster (Shinko SMB-C60) and rotary (Alcatel T2063SD) pumps down to 1.2 K. The pressure of pure helium in the gas supply system was then increased to 2.8 MPa, the crystal was seen to grow in a few minutes after pressure equalization (it was sometimes noticeable that a solid–solid transition occurred at the beginning, and the ascending interface

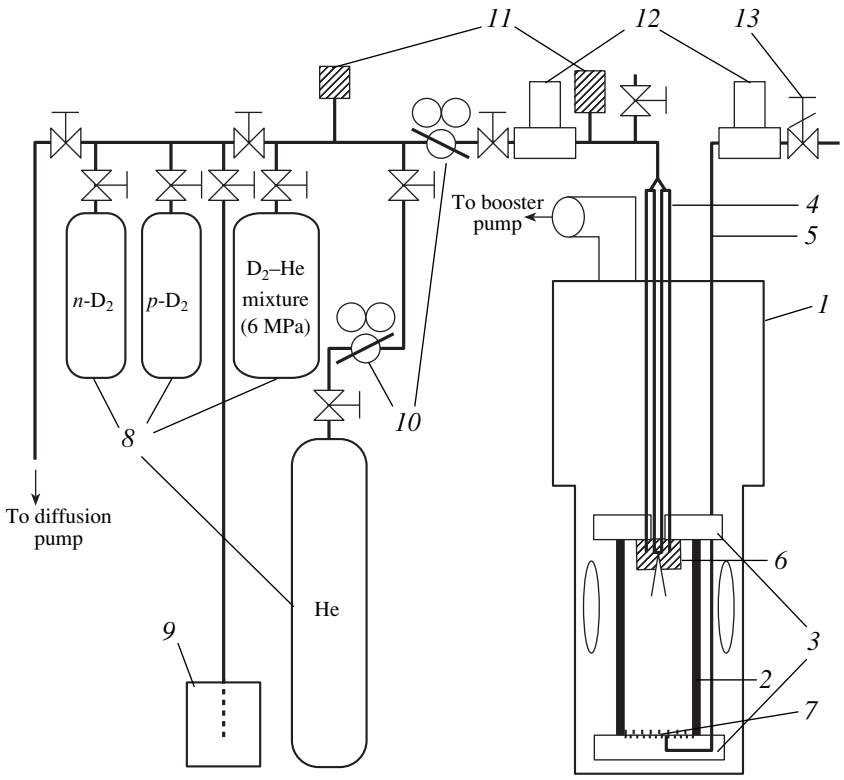



**Fig. 1.** Diagram of optical cell, He cryostat, and stainless gas-handling system: (*1*) cryostat, (*2*) sapphire cell, (*3*) brass flanges, (*4*) double-walled inlet tube (a constantan heater and a thermocouple are glued to the inner tube), (*5*) outlet tube (a constantan heater is inside), (*6*) teflon cap, (*7*) alumina mesh, (*8*) gas cylinders, (*9*) *para–ortho* converter, (*10*) pressure regulators, (*11*) pressure gauge, (*12*) mass flow meters, (*13*) needle valve.

was clearly seen; this proceeded at a rather small pressure while, as follows from the phase diagram [10], the bcc phase of solid helium formed earlier and was then converted to the hcp phase upon further cooling). Afterward, the needle valve at the end of the outlet tube was opened to start helium gas exhaust from the bottom of optical cell, the outflow usually kept at 200–400 sccm (it corresponds to $(0.9–1.8) \times 10^{20}$ atoms/s). In steady conditions, the arising inflow was equal to the outflow; however, periodical jumps of the flows occurred, especially when the pressure in the cell was sufficiently high. We believe that this happened when the pressure in the cell was more than 2.8 MPa (because of the impedance caused by the orifice, we could not accurately determine the pressure in flow conditions) and the film of liquid helium covering the crystal surface was not consequently superfluid anymore (both normal liquid and solid helium have rather low thermal diffusivity [10]). Indeed, the liquid–solid interface was concave in the case in which the main impedance for condensation heat removal is inside the cell. By decreasing the pressure, one can achieve conditions when the interface becomes flat (the main impedance is the heat transfer through a sapphire wall) and stable. The photo in Fig. 2 demonstrates such a pattern. For growing a doped crystal, the flow of pure helium from the cylinder was replaced by that of the He-impurity mixture preliminarily prepared in the storage vessel. The typical rate of $D_2$-doped helium crystal growth was 1 mm/min. The pumping rate was high enough to keep the temperature of the bath less than 1.7 K during the sample preparation procedure when the heat release due to both the inlet heating and the gas mixture condensation was about 3 W. In all cases, the doped crystals were nicely transparent. After sample preparation, the inflow of the He-impurity mixture was again replaced by the flow of pure helium for several minutes and then, keeping the pure helium supply system connected with the cell, the outflow was terminated and the inlet heater switched off. The sample thus prepared could be investigated for a long time without any noticeable changes.

The slow decrease of the pressure inside the cell by introducing a small outflow without any inflow led to frontal melting of the sample starting from its upper edge, as demonstrated in Fig. 3. The impurity from the melted part of the crystal then concentrated at the liquid–solid interface, forming a semitransparent sediment if the impurity content in the gas mixture was large. The interface and sediment motion can be terminated at a given position by stopping outflow and opening inflow to solidify the whole sample again. By vertical shift of the cell as a whole, the place with the

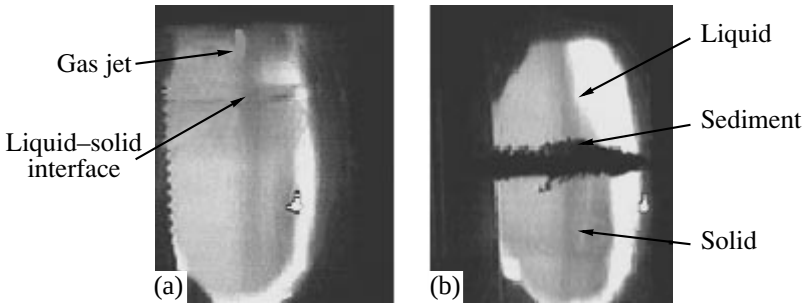



**Fig. 2.** Photos of a He crystal doped by 700 ppm of nitrogen: (a) the procedure of sample preparation, $\Phi = 200$ sccm, $P = 3.15$ MPa, $T = 1.5$ K, the temperature of orifice is around 250 K; (b) crystal frontal melting by pressure decreasing to 2.6 MPa.

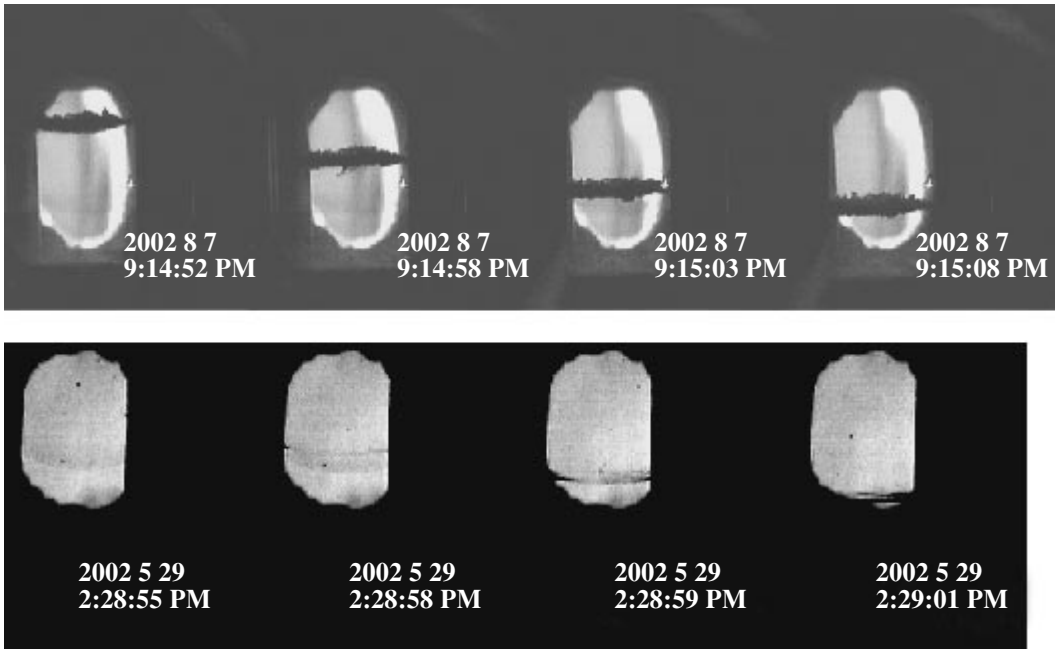



**Fig. 3.** Motion of the liquid–solid interface under diminishing helium pressure in real time for solid He doped by $N_2$ (upper) and by $D_2$ (lower), the window diameter is 3 cm.

enhanced impurity density could be placed into the observation zone.

The vibronic spectrum of coherent anti-Stokes Raman scattering of deuterium at our temperatures (around 1.5 K) consists of $Q_1(1)$ and $Q_1(0)$ transitions, the first of which belongs to $p$-$D_2$, and the second, to $o$-$D_2$. The scheme of CARS measurements is presented in Fig. 4. Two light beams with wavelengths of 566 (green) and 670–675 nm (red) from tunable pulsed dye lasers (Lambdaphysik Scanmate and Lumonics Hyperdye-300) pumped by the second harmonic of a pulsed Nd:YAG laser with a pulse duration of 5 ns and repetition rate of 10 Hz (Continuum Powerlite 7000) were focused in one spot, usually near the axis of the optical cell. The CARS signal generated at the focal point was detected by a photomultiplier (Hamamatsu R-4220). Laser light and parasite luminescence were attenuated by dichroic filters and a monochromator (Nikon P-250) tuned to the wavelength of CARS emission. After integration by Boxcar (Stanford SR250), the CARS signal was digitized by an AD converter (Interface PCI3133)

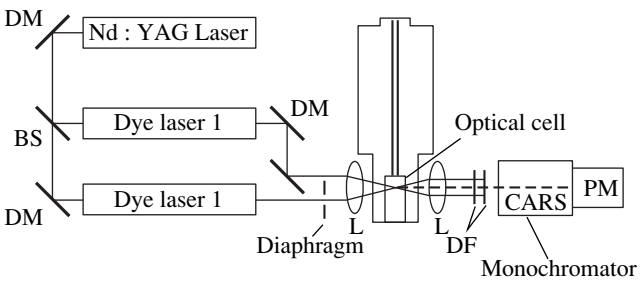



**Fig. 4.** Scheme of CARS detection: DM, dichroic mirror; BS, beam splitter; DF, dichroic filter; L, lenses; PM, photomultiplier.

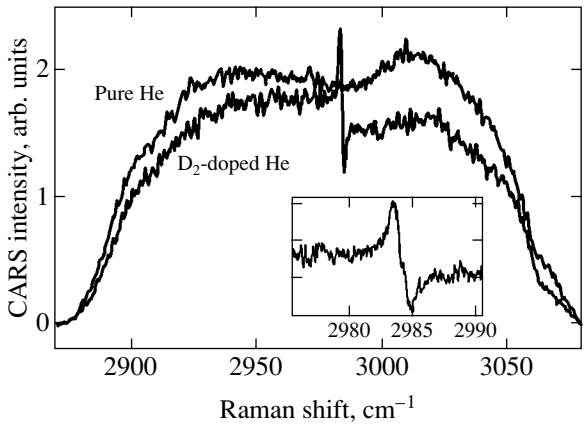



**Fig. 5.** CARS signal from pure He solid and from that doped by *n*-$D_2$ (2000 ppm). Apparent nonresonant signal shape is determined by the monochromator spectral width. The shape of $D_2$ line is shown in the inset.

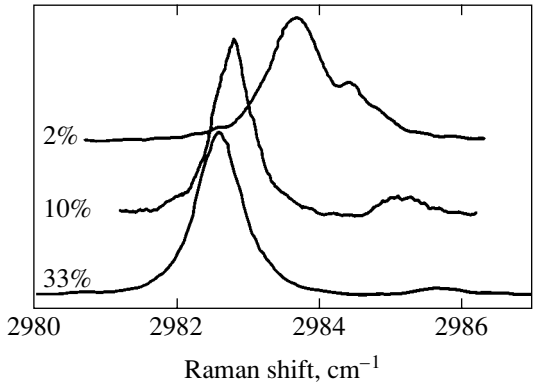



**Fig. 6.** CARS signal from $D_2$-doped solid He with the doped $D_2$ molecules (with different contents of *p*-$D_2$: 2, 10, and 33%) concentrated by the crystal temporal melting. The large left and small right peaks correspond to $Q_1(1)$ and $Q_1(0)$ transitions of $D_2$, respectively.

and was then accumulated in a personal computer. Using this technique, we were able to detect the $Q_1(1)$ and $Q_1(0)$ lines of gaseous deuterium even at a pressure of $10^{-4}$ MPa with a frequency resolution of about 0.5 cm$^{-1}$. As is known, the intensity of CARS signal $I_{\mathrm{CARS}}$ is

$$I_{\mathrm{CARS}} \propto I_g^2 I_r n^2, \tag{1}$$

where $I_g$ and $I_r$ are the respective intensities of green and red laser light and $n$ is the density of deuterium molecules at the focal point.

## 3. EXPERIMENTAL RESULTS

The presence of solid helium in the cell resulted in the appearance of a rather intense nonresonant blue signal, whose intensity was approximately proportional to $I_g^2$ and to $I_r$ and whose wavelength was $2\omega_g - \omega_r$. The signal shown in Fig. 5 corresponds to nonresonant four-wave scattering by helium. The probability of such a process is known to be proportional to the square of the nonresonant third-order susceptibility, which is ten times less for helium than for deuterium and is equal to $4 \times 10^{-39}$ cm$^6$/erg per molecule [11]. The presence of an impurity in the helium crystal grown by condensation of the gas mixture containing a rather small (less than 1000 ppm) amount of deuterium did not cause any significant change of this spectrum in either shape or intensity.

However, if part of the sample was temporarily melted, the intensive narrow resonant deuterium signal was observed under focusing the laser beams to the former border between melted and nonmelted solid helium. The signal was stable in time (see Fig. 6).

Under the condensation of the gas mixture with large deuterium content, the resonant signal at a wavelength close to the deuterium line position exists from the very beginning (without preliminary melting) on the background of the nonresonant CARS signal. Its shape, shown in Fig. 5, is typical of the resonance CARS signal superimposed on the intensive nonresonant background [12, 13]; the interference of these scatterings forms a so-called modulation dip distorting the line and "eating away" the background. The depth of the modulation dip shows that a significant part of nonresonant scattering originates from the focal point of laser beams, i.e., caused by solid helium; the rest may proceed from the sapphire tube, optical windows, etc. It is worth mentioning that the CARS signal intensity was found independently of the focus point moving up and down or along the cell radius, thereby proving the uniformity of the stabilized cluster distribution in solid helium.

The impurity concentrating near the bottom of the "melted" zone during crystal temporal melting evidences a large size of free inclusions formed by this procedure in condensed helium. Indeed, the height $h$ of the region of impurity clusters precipitation by gravity in liquid helium can be evaluated from the Boltzmann distribution

$$\frac{Nmgh}{kT}\frac{\rho_{D_2} - \rho_{He}}{\rho_{He}} \approx 1, \tag{2}$$

where $m$ is the mass of deuterium, $N$ is the number of molecules in a cluster, and $\rho_{D_2}$ and $\rho_{He}$ are the respective densities of solid deuterium and liquid helium. The height of the region with large cluster concentration estimated by the existence of an intensive CARS resonant signal is about 0.3 mm. Such an estimate gives $N = 10^8$, which corresponds to a cluster size of 100 nm. The results of X-ray analysis of clusters formed by helium–deuterium jet injection directly to superfluid helium gave a size of 3–6 nm [14]. However, this estimate, made from the diffraction peak shape, represents

a lower bound of size because the width of the diffraction peak is sensitive not to the crystallite size but to the length of the regularity of the crystal's structure. Most probably, temporal melting of solid helium causes mutual gluing of small clusters possessing rather regular structure, into large particles.

The above observations make the following scenario acceptable.

The condensation of mixtures containing a relatively small amount of deuterium (less than 1000 ppm) leads to the formation of rather small deuterium clusters distributed in bulk solid helium. As the content of the impurity in the gas mixture increases, the probability of molecule coalescence in a cooled gas jet just before it enters liquid helium and especially in liquid helium (during particle motion to the solid–liquid interface) significantly increases and the clusters that have stabilized in solid helium become larger.

Finally, under helium crystal temporal melting, the clusters merge together in large objects precipitating due to gravity. The cluster coalescence occurs in spite of the pressure being only about 1 bar less than that at the helium melting point, but the effective centripetal pressure caused by van der Waals interaction of the second layer of surrounding helium atoms with the deuterium core, being several bars, is therefore sufficient for keeping this layer solid even in a liquid [15].

It is reasonable to begin the analysis of CARS spectra of deuterium stabilized in solid helium from this sediment modeling a massive deuterium sample. In this case, the $Q_1(0)$ and $Q_1(1)$ line positions for different $o$-$D_2$ contents presented in Fig. 7 fit well the literature data available for solid deuterium with $o$-$D_2$ content more than 20% [6] and correctly reflect superlinear decrease of the $Q_1(0)$–$Q_1(1)$ splitting at small $o$-$D_2$ content observed recently [16] (the Fabri–Perrot resonator technique used in this work makes it possible to determine only splitting, not line positions).

It is immediately seen from Fig. 6 that, even at a $p$-$D_2$ content as low as 2%, the $Q_1(1)$ line is still much more intense than the $Q_1(0)$ one (in the CARS spectrum of that mixture in a gas, we did not observe the $Q_1(1)$ line at all because it should be more than three orders of magnitude less intense than the $Q_1(0)$ line). Taking into account that, according to (1), the CARS signal intensity is proportional to the square of the scatterer concentration, we calculated the ratio of the probabilities of the $Q_1(1)$ and $Q_1(0)$ transitions in stimulated Raman scattering for deuterium crystallites with different *ortho–para* ratios. CARS line profiles both in a gas and in a solid were found to be nicely approximated by a Gaussian width of 0.62 $cm^{-1}$, which probably represented the frequency resolution of our technique, and we simulated the line shape by this form in all cases. These probability ratios are compared in the table with those known for common Raman scattering. One can easily notice that the crystal-field effect under consideration is much more pronounced in CARS: in deuterium containing 2% of *para* modification, that ratio is as large as 10 000, although it is only about 50 in Raman scattering [6].

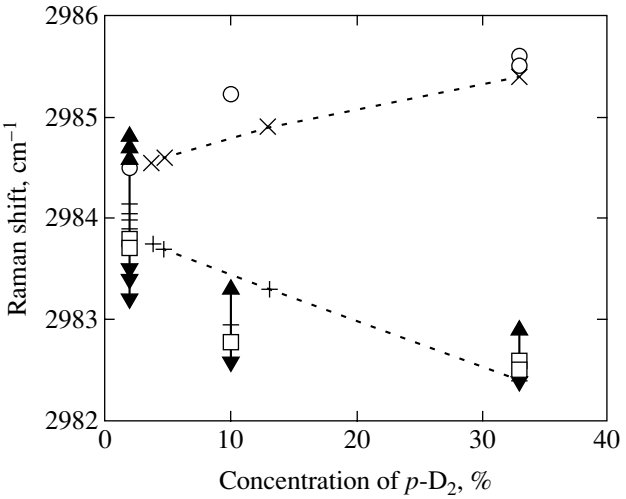



**Fig. 7.** Raman line for $D_2$ positions vs. *ortho–para* content in a mixture: distorted line for $D_2$ molecules trapped in solid He ((▼) negative peak, (▲) positive peak); (○, □) $o$- and $p$-$D_2$ molecules lines in remelted solid He; (×, +) the same for solid $D_2$ [16].

For deuterium clusters isolated in solid helium (see Fig. 5), line distortion by a coherent interaction with the background does not allow precise determination of the position of the resonant CARS line. However, the peak of the ungarbled line should be positioned around half the distance between the maximum and the minimum of the distorted line [17], and as seen from Fig. 7, the signals for clusters with 10 and 33% $p$-$D_2$ content consist mainly of $Q_1(1)$ lines. This means that, for these clusters, presumably containing $10^3$–$10^4$ molecules, the effect of $Q_1(1)$ line predominance already exists: assuming the equal probabilities of $Q_1(0)$ and $Q_1(1)$ transitions, as is the case for the gas phase, the $Q_1(0)$ line intensity should be 10 and 100 times larger for $n$-$D_2$ and $D_2$ containing 10% of *para* modification correspondingly. This effect seems to be weaker than in a massive sample: in deuterium clusters containing 2% of *para* modification, both the spectral width of the signal (much broader) and its position are evidence that it actually consists of both $Q_1(1)$ and $Q_1(0)$ lines with

The ratios of probabilities of the $Q_1(1)$ and $Q_1(0)$ transitions for different contents of *para*-$D_2$ in a solid; for spontaneous RS, data were taken from [6], the CARS data are our results

| | 100% | 33% | 10% | 2% |
|---|---|---|---|---|
| RS | 5 | 10 | 30 | 50 |
| CARS | ? | 150 | 1260 | 10000 |

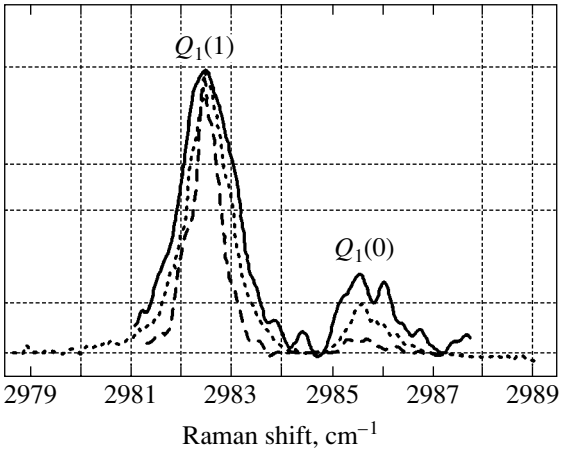



**Fig. 8.** $Q_1(1)$ and $Q_1(0)$ lines for He: *n*-$D_2$ crystal ($T$ = 1.3 K, $P$ = 3.15 MPa). Solid line, condensation of He gas containing 200 ppm of $D_2$. Dashed line, the same sample after fast annealing by decreasing the pressure to 2.6 MPa. Dotted line, solid obtained by condensation of He containing 1000 ppm of $D_2$.

comparable intensities, whereas in sediment, the intensity of the $Q_1(1)$ line is still much higher than that of $Q_1(0)$ (see Fig. 6). This conclusion has been proved directly by comparing the line profiles before and after solid helium sample temporary melting; for reliable line resolution, we then used *n*-$D_2$, where $Q_1(0)$ and $Q_1(1)$ splitting is rather large. The results of such experiments presented in Fig. 8 clearly demonstrate that, in smaller clusters, the lines are significantly more broadened and the ratio of their intensities is smaller.

## 4. DISCUSSION

In spite of the obvious presence of deuterium in solid helium bulk, no signal has been found for small deuterium content in a mixture when the smallest clusters are supposed to have been stabilized. The low deuterium densities should not be the only reason for failure to observe CARS. Indeed, we have reliably detected the CARS signal from gas deuterium at a density of $3 \times 10^{17}$ cm$^{-3}$. In experiments with solid helium, there was extra noise connected with the instability of nonresonant scattering by solid helium, caused as usual by fluctuation of laser energies from pulse to pulse. However, with the same linewidth, we should easily detect deuterium lines at a $D_2$ density of about $3 \times 10^{18}$ cm$^{-3}$ (100 ppm of deuterium in gas mixture). That means the deuterium lines from small clusters should be noticeably broadened, either by multiple faults of the crystal structure or due to the large contribution from molecules located at a cluster surface where the matrix shift should be about half that in the bulk; both reasons should be valid only for small molecular clusters consisting of a few layers. In principle, small clusters could be studied by using delayed CARS [18]: because the helium third-order susceptibility should be formed already in femtoseconds and hydrogen CARS characteristic times are about a nanosecond even in a nonideal crystal, the nonresonant background may be removed there.

The modulation dip we observed for larger deuterium clusters stabilized in solid helium can be used to measure the helium nonresonant third-order susceptibility $\gamma^{(3)}_{non}$ by the procedure proposed in [19] and applied there for argon $\gamma^{(3)}_{non}$ determination from analysis of the shape of the dip in modulation of the nitrogen CARS signal. Although the $\gamma^{(3)}_{non}$ value of helium is of general interest [12], the accuracy of its value in the literature is far from satisfactory. Owing to the small $\gamma^{(3)}_{non}$ value, it is difficult to measure it in a helium gas; this has been impossible unless by embedding small (less than $\lambda$) separated optically active impurities into condensed helium.

Surprisingly, although the $Q_1(0)$ and $Q_1(1)$ lines in condensed hydrogen and deuterium have been extensively studied by Raman scattering, the CARS technique has not been used for this. Our results for deuterium sediment in remelted solid helium should therefore be considered the first study of the large effect of the surrounding nuclear spin modification on the $Q_1(0)$ and $Q_1(1)$ transition probability ratio in solid deuterium by the CARS technique. From a general standpoint, such an effect should be much stronger in CARS than in Raman scattering. Indeed, in the harmonic oscillator approximation valid for simple molecules, $\gamma^{(3)}_{res} \propto \alpha^2$, where $\alpha$ is the polarizability and $\gamma^{(3)}_{res}$ is the third-order susceptibility at a resonance [20], and there is a simple relation between the ratios of the $Q_1(1)$ and $Q_1(0)$ line intensities in spontaneous RS and in CARS,

$$\frac{I_{CARS}(1)}{I_{CARS}(0)} = \left(\frac{I_{RS}(1)}{I_{RS}(0)}\right)^2. \qquad (3)$$

The data in the table show that this is nearly true. Nevertheless, one may notice that the effect is stronger than follows from (3): for example, in Raman scattering experiments [6] for deuterium crystals containing 2% of *para* modification, the intensities of $Q_1(1)$ and $Q_1(0)$ lines are nearly the same, whereas in CARS, as is seen from Fig. 6, the $Q_1(0)$ line is significantly weaker. In fact, this effect may be used for determination of completeness of the *para–ortho* conversion in deuterium with sensitivity better than $10^{-3}$%.

The impossibility of detecting the resonant CARS signal in small clusters does not allow us to fully trace the size dependence of the effect—it already exists when the resonant deuterium peak appears in our experiments. This is evidence that the effect of huge predominance of the CARS accompanied by the $Q_1(1)$

transition develops together with phonon band formation in a deuterium cluster. However, the transfer to a large probability ratio proceeds gradually: we managed to observe enhancement of the effect as the cluster size grows further (see Fig. 5), synchronously with the line narrowing.

## 5. CONCLUSIONS

The promises of the new technique of embedding impurities into solid helium have been demonstrated with the example of deuterium cluster stabilization in a matrix more inert than hydrogen. It was experimentally shown that the impurity clusters are distributed homogeneously in the helium crystal; this opens the possibility of using laser irradiation to evaporate individual molecules from the clusters preliminarily stabilized in solid helium, as has been done in [8, 9] for metallic atoms.

The extremely strong dependence of the third-order resonant susceptibility related to $Q_1(0)$ and $Q_1(1)$ transitions in solid deuterium on the ratio of *ortho–para* modifications has been found. Of course, this effect must exist in a massive deuterium crystal, as well as in hydrogen. It is significant that, in spontaneous Raman scattering, the effect of the $J = 1$ transition dominance consists of redistribution of probabilities of scattering by *ortho* and *para* states, with the total cross section being independent of their content [6]. This should not be true for stimulated Raman scattering, where the probabilities are proportional to square of the species concentration.

The phenomenon of the $J = 1$ Raman scattering dominance has been found to occur just as the size of deuterium crystallite becomes sufficient for phonon band formation.

Nonresonant CARS from solid helium has been observed for the first time; in a crystal doped by small clusters of Raman-active molecules, the modulation dip of their CARS line made it possible to calculate the absolute value of the nonresonant third-order susceptibility for helium.

## ACKNOWLEDGMENTS

This work was supported in part by the Grant in Aid of Scientific Research of the Ministry of Education, Culture, Sports, Science, and Technology of Japan. E.G. appreciates support from the JAERI Foreign Researcher Invitation Program and from the Russian Foundation for Basic Research (project nos. 00-15-97400 and 01-03-32247). The authors are grateful to T. Momose for useful discussions.

## REFERENCES

1. T. Oka, Annu. Rev. Phys. Chem. **44**, 299 (1993).
2. S. Grebenev, J. P. Toennies, and A. F. Vilesov, Science **279**, 2083 (1998).
3. E. B. Gordon, A. Usenko, and G. Frossati, J. Low Temp. Phys. **130** (1–2), 15 (2003).
4. E. B. Gordon, G. Frossati, A. Usenko, *et al.*, Physica B (Amsterdam) **329–333**, 404 (2003).
5. V. A. Mikheev, B. A. Maidanov, and N. P. Mikhin, Sov. J. Low Temp. Phys. **9**, 901 (1983).
6. J. Van Kronendonk, *Solid Hydrogen, Theory of the Properties of Solid $H_2$, HD, and $D_2$* (Plenum, New York, 1983).
7. S. S. Bhatnagar, E. J. Allin, and H. L. Welsh, Can. J. Phys. **40**, 9 (1962).
8. S. I. Kanorsky and A. Weis, Adv. At. Mol. Opt. Phys. **38**, 87 (1997).
9. K. Ishikawa, A. Hatakeyama, K. Gosyono-o, *et al.*, Phys. Rev. B **56**, 780 (1997).
10. J. Wilks, *Properties of Liquid and Solid Helium* (Oxford Univ. Press, Oxford, 1959).
11. T. Lundeen, S.-Y. Hou, and J. W. Nibler, J. Chem. Phys. **79**, 6301 (1983).
12. J. W. Hahn, S. N. Park, E. S. Lee, *et al.*, Appl. Spectrosc. **47**, 710 (1993).
13. M. D. Levenson, IEEE J. Quantum Electron. **10**, 110 (1974).
14. S. I. Kiselev, V. V. Khmelenko, D. M. Lee, *et al.*, Phys. Rev. B **65**, 024 517 (2002).
15. E. B. Gordon and A. F. Shestakov, Low Temp. Phys. **26**, 1 (2000).
16. J. L. Fredman, J. H. Eggert, J. D. Kinder, *et al.*, J. Low Temp. Phys. **115**, 181 (1999).
17. N. Bloembergen, *Nonlinear Optics*, 4th ed. (World Sci., River Edge, N.J., 1996; Mir, Moscow, 1966).
18. M. Karavitis, R. Zadoyan, and V. A. Apkarian, J. Chem. Phys. **114**, 4131 (2001).
19. K. Akihama, T. Asai, and S. Yamazaki, Appl. Opt. **32**, 7434 (1993).
20. S. A. Akhmanov and N. I. Koroteev, *Methods of Nonlinear Optics in Spectroscopy of Light Scattering* (Nauka, Moscow, 1981) [in Russian].

*Journal of Experimental and Theoretical Physics, Vol. 99, No. 4, 2004, pp. 784–791.*
*From Zhurnal Éksperimental'noĭ i Teoreticheskoĭ Fiziki, Vol. 126, No. 4, 2004, pp. 907–914.*




# Ultrarelativistic Nuclei in a Crystal Channel: Coulomb Scattering, Coherence, and Absorption¶

**V. R. Zoller**
*Institute for Theoretical and Experimental Physics, Moscow, 117218 Russia*
*e-mail: zoller@itep.ru*
Received May 12, 2004

**Abstract**—We incorporate the effect of lattice thermal vibrations into the Glauber-theory description of particle and nucleus–crystal Coulomb interactions at high energy. We show that taking the lattice thermal vibrations into account produces a strong absorption effect: the phase shift function of the multiple-diffraction scattering on a chain of $N$ identical atoms acquires a large imaginary part, and the radius of the absorption region in the impact parameter plane grows logarithmically with $N$. Consequences of this observation for the elastic and quasi-elastic Coulomb scattering are discussed. The practically interesting example of the coherent Coulomb excitation of ultrarelativistic particles and nuclei passing through a crystal is considered in detail. © *2004 MAIK "Nauka/Interperiodica".*

## 1. INTRODUCTION

In this paper, we develop the description of the absorption phenomenon in coherent particle and nucleus–crystal Coulomb interactions at high energy in the framework of the Glauber theory [1].

As is well known, multiloop corrections generate the imaginary part of the scattering amplitude even if the tree-level amplitude is purely real. For example, the purely real Born amplitude of the high-energy Coulomb scattering in a crystal acquires the imaginary part due to the multiple scattering (MS) effects. However, in the widely used static/frozen lattice approximation (SL approximation), taking rescatterings into account alters only the overall real phase of the full amplitude, thus producing no absorption effect. The latter is related to the creation and annihilation of excited intermediate states of the crystal and as such manifests itself only beyond the SL approximation (see [2] for the analysis of elastic scattering based on the SL approximation). Indeed, the amplitude of small-angle elastic scattering on a chain of $N$ identical atoms in the impact parameter representation is equal to

$$1 - \langle S(b)\rangle,$$

where the scattering matrix placed between the ground states of crystal is

$$\langle S(b)\rangle = \langle \exp[i\chi(b)]\rangle,$$

with the purely real phase shift function

$$\chi(b) = \sum_{j=1}^{N} \chi_j(b).$$

In the SL approximation,

$$\langle \exp[i\chi(b)]\rangle \approx \exp[i\chi(b)].$$

In general, therefore, the Coulomb phase shift function acquires a nonvanishing imaginary part, which is interpreted as an absorption effect, only with the lattice thermal vibrations taken into account. The imaginary part appears only as a second-order perturbation,

$$\sim \frac{i}{2}[\langle \chi^2\rangle - \langle \chi\rangle^2].$$

But the strength of the effect is proportional to $\beta^2 N$, where $\beta$ is the coupling constant. For the coherent scattering of relativistic nuclei (the electric charge $Z_1$) on the chain of $N$ atoms (the atomic number $Z_2$) in a crystal, the effective coupling

$$\beta = 2\alpha Z_1 Z_2$$

is strong and the absorption effect is also strong. The absorption is strong for the impact parameters $b$ smaller than some characteristic value

$$b_a \propto \log(\beta N)$$

and vanishes toward the region of larger $b$. This phenomenon provides a natural ultraviolet regulator of the theory and enables, in particular, consistent calculation of the coherent elastic scattering cross section. The latter is calculated and turns out to be equal to half the total cross section. As we see in what follows, the absorption effect is also of prime importance for quantitative understanding of the phenomenon of the coherent Coulomb excitation of relativistic particles and

¶ This article was submitted by the author in English.

nuclei passing through a crystal. A consistent description of this phenomenon is the goal of our paper.

The outline of the paper is as follows. We start with the well-known example of the coherent Coulomb elastic scattering of charged particle/nucleus by a linear chain of $N$ identical atoms in a crystal target (Section 2). In Section 3, we derive the scattering matrix with absorption and calculate the cross section $\sigma_{el}$ of the coherent elastic scattering and the cross section $\sigma_{Qel}$ of the incoherent excitation and breakup of the target. Then we find that in the large-$N$ limit,

$$\sigma_{el} \approx \sigma_{Qel} \approx \frac{1}{2}\sigma_{tot}.$$

In Section 4, we discuss the coherent Coulomb excitation of ultrarelativistic particles and nuclei passing through the crystal to the lowest order of perturbation theory. The higher-order effects are considered in Section 5, where the cross section of the process is calculated. We finally conclude with a brief summary in Section 6.

## 2. COHERENT ELASTIC SCATTERING AND ABSORPTION

The interatomic distances in a crystal, $a$, are large compared to the Thomas–Fermi screening radius $r_0$,

$$a \sim 3\text{–}5\ \text{Å} \gg r_0 = r_B Z_2^{-1/3} \sim 0.1\ \text{Å},$$

where $Z_2$ is the atomic number of the target atom and $r_B$ is the Bohr radius [3]. The relevant impact parameters $b$ satisfy the condition $b \ll a$, and the amplitudes of scattering by different atomic chains parallel to a given crystallographic axis are incoherent.

The amplitude of small-angle scattering of a charged particle (charge $Z_1$) by a linear chain of $N$ identical atoms in the eikonal approximation is given by [1]

$$F_{fi}(q) = \frac{ip}{2\pi}\int d^2\mathbf{b}\exp(i\mathbf{q}\cdot\mathbf{b}) \times \langle\Psi_f(\{\mathbf{r}_j\})|1 - S(\mathbf{b}, \mathbf{s_1}, \ldots, \mathbf{s_A})|\Psi_i(\{\mathbf{r}_j\})\rangle, \quad (1)$$

where $\Psi_i$ and $\Psi_f$ are the initial and final state wavefunctions of the crystal and $\mathbf{q}$ is the two-dimensional vector of the momentum transfer. The incident particle momentum $p$ is assumed to be large enough to satisfy the condition of applicability of the straight-paths approximation, $p/q^2 \gg aN$. The latter condition ensures the coherence of interactions with different atoms.

The elastic scattering corresponds to $i = f$ and the brackets $\langle\ \rangle$ signify that the average is to be taken over all configurations of atoms in the ground state,

$$\langle\Psi(\{\mathbf{r_j}\})|1 - S(\mathbf{b}, \mathbf{s}_1, \ldots, \mathbf{s}_N)|\Psi(\{\mathbf{r}_j\})\rangle = \int d^3\mathbf{r}_1 \ldots d^3\mathbf{r}_N |\Psi(\{\mathbf{r}_j\})|^2 \times \left[1 - \exp\left(i\sum_{j=1}^{N}\chi(\mu|\mathbf{b} - \mathbf{s}_j|)\right)\right]. \quad (2)$$

In (2), the total scattering phase is the sum of the phase shifts contributed by the individual atoms. The positions of the $N$ atoms that make up the target are defined by the three-dimensional vectors $\mathbf{r}_j$, $j = 1, \ldots, N$. The two-dimensional vectors $\mathbf{s}_j$ are the projections of these vectors on the impact parameter plane. We neglect all position correlations of the atoms and describe the ground state of the crystal by the wavefunction $|\Psi\rangle$ such that

$$|\Psi(\{\mathbf{r}_j\})|^2 = \prod_{j=1}^{N}|\psi(\mathbf{u}_j)|^2, \quad (3)$$

where the three-dimensional vectors $\mathbf{u}_j$ are defined by

$$\mathbf{r}_j = (j-1)\mathbf{a} + \mathbf{u}_j, \quad j = 1, \ldots, N, \quad \mathbf{a} = (0, 0, a)$$

and $\mathbf{u}_j = (\mathbf{s}_j, z_j)$.

From Eq. (2), it follows that

$$F_{ii}(q) = F(q) = ip\int bdbJ_0(qb)\{1 - \langle\exp[i\chi(\mu b)]\rangle^N\}. \quad (4)$$

Hereinafter, $J_{0,1}(x)$ and $K_{0,1}(x)$ are the Bessel functions and the screened Coulomb phase shift function is

$$\chi(\mu b) = -\beta K_0(\mu b) \quad (5)$$

with

$$\beta = 2\alpha Z_1 Z_2, \quad \mu = r_0^{-1}.$$

After integration over longitudinal variables $\{z_j\}$ followed by the azimuthal integration, the term $\langle\exp(i\chi)\rangle$ takes the form

$$\langle\exp(i\chi)\rangle = \int d^2\mathbf{s}\rho(s)\exp[i\chi(\mu|\mathbf{b} - s|)] = \exp(-\Omega^2 b^2)\int_0^\infty dx\exp(-x) \times I_0(2b\Omega\sqrt{x})\exp[-i\beta K_0(\mu\sqrt{x}/\Omega)]. \quad (6)$$

The two-dimensional vector $\mathbf{s}$, describes the position of the target atom in the impact parameter plane. The one-

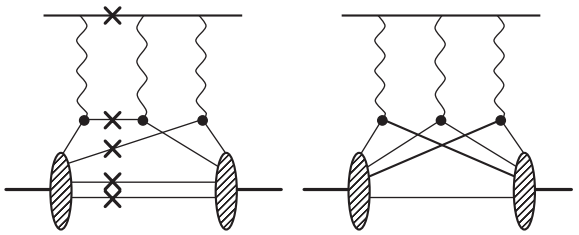



**Fig. 1.** Example of the relevant multiple scattering diagrams to the order $\beta^3$. The unitarity cut of the elastic amplitude (a) that contributes to the absorption is shown by crosses. Diagram (b) allows cuts only between the projectile-atom blocks and does not contribute to the absorption effect.

particle probability distribution $\rho(s)$ is given by

$$\rho(s) = \int dz|\psi(\mathbf{s}, z)|^2 = (\Omega^2/\pi)\exp(-\Omega^2\mathbf{s}^2). \quad (7)$$

For the most commonly studied elements at room temperature, the ratio $\mu/\Omega$ varies in a wide range, from $\mu/\Omega \sim 0.1$ to $\mu/\Omega \sim 1$ [3]. We first consider the region of small impact parameters including $b \lesssim 1/\Omega$.[1] For $b \lesssim 1/2\Omega$, only small $s$, such that $\mu s \lesssim 1$, contribute. We can then set

$$K_0(\mu s) \approx \log(1/\mu s)$$

and integrate over $s$,

$$\langle \exp(i\chi)\rangle \approx \left(\frac{\mu}{\Omega}\right)^{i\beta}\exp(-\Omega^2 b^2)$$

$$\times \int_0^\infty dx x^{i\beta/2}\exp(-x)I_0(2b\Omega\sqrt{x}) \quad (8)$$

$$= \left(\frac{\mu}{\Omega}\right)^{i\beta}\Gamma\left(1+\frac{i\beta}{2}\right)\Phi\left(-\frac{i\beta}{2}; 1; -\Omega^2 b^2\right),$$

where

$$\Phi(a; b; z) = 1 + \frac{a}{b}\frac{z}{1!} + \frac{a(a+1)}{b(b+1)}\frac{z^2}{2!} + \ldots \quad (9)$$

is the confluent hypergeometric function and

$$\Phi(a; b; z) = \exp(z)\Phi(b-a; b; -z).$$

From Eq. (8), it follows that

$$|\langle \exp(i\chi)\rangle|_{b=0} = \left[\frac{\pi\beta}{2\sinh(\pi\beta/2)}\right]^{1/2}, \quad (10)$$

where identity

$$|\Gamma(i\beta/2)|^2 = \frac{2\pi}{\beta\sinh(\pi\beta/2)} \quad (11)$$

was used. In the weak coupling regime $\beta \ll 1$,

$$|\langle \exp(i\chi)\rangle|_{b=0} \approx 1 - \frac{1}{2}(\langle\chi^2\rangle - \langle\chi\rangle^2)$$

and

$$\langle\chi^2\rangle - \langle\chi\rangle^2 = \frac{\pi^2\beta^2}{24}. \quad (12)$$

For $\beta \gtrsim 1$,

$$|\langle \exp(i\chi)\rangle|_{b=0} \approx \sqrt{\pi\beta}\exp(-\pi\beta/4). \quad (13)$$

Therefore, at small impact parameters, $b \lesssim 1/2\Omega$, the intensity of outgoing nuclear waves as a function of $N$ exhibits exponential attenuation. In terms of the unitarity cuts of the elastic scattering amplitude, the imaginary part of the phase shift function comes from the cuts through the multiphoton projectile-atom blocks, as shown in Fig. 1a. The account of diagrams like that in Fig. 1b, which allows cuts only between projectile-atom blocks, gives a scattering matrix of the form $\exp(iN\langle\chi\rangle)$ and affects only the overall real phase of the amplitude.

The absorption effect becomes weaker toward the region of large impact parameters $b \gtrsim 1/2\Omega$,

$$|\langle \exp(i\chi)\rangle|^N \approx |\langle \exp(i\chi)\rangle|_{b=0}^N \times\left[1 + \frac{N\beta^2}{16}(\Omega b)^4 + \ldots\right]. \quad (14)$$

For still larger $b$, $b \gg 1/2\Omega$, using the asymptotic form

$$I_0(z) \approx (2\pi z)^{-1/2}\exp(z),$$

we obtain

$$\langle \exp(i\chi)\rangle \approx 2\Omega\int \frac{sds}{\sqrt{\pi bs}}\exp[-\Omega^2(b-s)^2] \times \exp[i\chi(\mu s)]. \quad (15)$$

To evaluate the integral in (15), we expand $\chi(\mu s)$ in powers of $(s-b)$,

$$\chi(\mu s) \approx \chi(\mu b) + \omega(s-b).$$

If the frequency $\omega$,

$$\omega = \frac{d\chi}{db} = \mu\beta K_1(\mu b),$$

is small compared to $\Omega$,

$$\omega \ll \Omega, \quad (16)$$

[1] We note that the smallness of the ratio $r_A^2/u^2 \sim 10^{-6}$–$10^{-5}$ where $r_A$ is the nuclear radius and $u = 1/\Omega$ is the amplitude of lattice thermal vibrations, allows one to neglect the nuclear interactions of the projectile up to $N \sim 10^5$. As we see below, the absorption effect in which we are interested enters the game at much smaller $N$.

then

$$\langle \chi^2 \rangle - \langle \chi \rangle^2 = \frac{\omega^2}{2\Omega^2}$$

and

$$\langle \exp(i\chi) \rangle \approx \exp[i\chi - i\omega b] \times \Omega \int_0^\infty \frac{sds}{\sqrt{\pi bs}} \exp[-\Omega^2 (b-s)^2] \exp[i\omega s] \quad (17)$$

$$\approx \exp(i\chi) \exp[-\omega^2/4\Omega^2].$$

Condition (16) is satisfied if $b \gg \beta/\Omega$. For the impact parameters from the region

$$\beta/\Omega \ll b \ll 1/\mu,$$

we can write $\omega \approx \beta/b$; for larger $b$ such that $b \gg \mu^{-1}$,

$$\omega \approx \mu\beta \sqrt{\frac{\pi}{2\mu b}} \exp(-\mu b). \quad (18)$$

From the consideration presented above, it follows that the absorption effect in the elastic scattering amplitude is especially strong for impact parameters

$$b \lesssim b_a = \frac{1}{2\mu} \log \frac{\pi \mu^2 \beta^2 N}{4\Omega^2}. \quad (19)$$

For $b \ll b_a$, the atomic chain acts as an opaque "black" disc. Certainly, the value of this finding differs for different observables and for different processes proceeding at different impact parameters. The only thing that is worth noticing here is the representation of the scattering matrix in the form

$$\langle S(b) \rangle \approx \exp\left( iN\chi - \frac{N\omega^2}{4\Omega^2} \right), \quad b \gg \beta\Omega^{-1}. \quad (20)$$

Equation (20) supplemented with the observation that

$$\langle S(b) \rangle \approx (\pi\beta)^{N/2} \exp\left( -\frac{N\pi\beta}{4} \right), \quad b \lesssim \Omega^{-1}, \quad (21)$$

simplifies all further calculations greatly.

## 3. THE CROSS SECTIONS

### *3.1. The Elastic Cross Section*

Integrating by parts reduces $F(q)$ to a form convenient for evaluation of the total cross section,

$$F(q) = \frac{ip\mu N}{q} \int_0^\infty b db J_1(qb) \langle i\chi' \exp(i\chi) \rangle \times \langle \exp(i\chi) \rangle^{(N-1)}. \quad (22)$$

At small impact parameters $b \ll 1/2\mu$,

$$\langle \chi' \exp(i\chi) \rangle \approx \beta \left( \frac{\mu}{\Omega} \right)^{i\beta - 1} \exp(-\Omega^2 b^2) \times \Gamma\left( \frac{i\beta + 1}{2} \right) \Phi\left( \frac{i\beta + 1}{2}; 1; \Omega^2 b^2 \right). \quad (23)$$

Because of multiple scatterings, only large impact parameters $b$ may contribute to $F(q)$ at large $N$ and small $q$. Hence,

$$F(q) \approx \frac{ip\mu N}{q} \int_{1/\mu}^\infty b db J_1(qb) [i\chi' - \omega\omega'/2\Omega^2] \times \exp(iN\chi) \exp(-N\omega^2/4\Omega^2), \quad (24)$$

where the explicit form of $\langle \exp(i\chi) \rangle$ at large $b$, Eq. (17), is used. For large $b$,

$$\omega^2 \propto \exp(-2\mu b);$$

as $b$ grows, $\omega^2$ decreases much faster than the phase shift function $\chi(\mu b)$, which is proportional to $\exp(-\mu b)$. We see that the leading contribution to the elastic scattering amplitude (24) comes from

$$b \sim \mu^{-1}\xi \gg b_a,$$

where

$$\xi = \log(\beta N).$$

For large $N$, the second term in the square brackets in (24) is small compared to the first one. Then, for

$$q \lesssim q_0 = \mu/\xi$$

and $\xi \gg 1$, the steepest descent from the saddle point

$$b_0 = \mu^{-1}[\xi + i\pi/2] \quad (25)$$

in Eq. (24) yields

$$F(q) \approx \frac{ipb_0}{q} J_1(qb_0). \quad (26)$$

The effect of lattice thermal vibrations at small $q$ appears to be marginal and reduces to the factor $\exp(\mu^2/4\Omega^2 N)$ in (26), which is irrelevant at large $N$. The amplitude $F(q)$ in Eq. (26) coincides with the elastic scattering amplitude given by the SL approximation [2].

If $q \gtrsim q_0$, the stationary phase approximation gives elastic scattering amplitude of the form

$$F(q) \approx \frac{-ip\sqrt{\eta}}{\mu q} \exp\left( -\frac{iq\eta}{\mu} \right) \exp\left( -\frac{q^2}{4\Omega^2 N} \right), \quad (27)$$

where

$$\eta = \log(\mu\beta N/q) \gg 1.$$

Taking into account the lattice thermal vibrations

insures the convergence of the integral for the coherent elastic scattering cross section,

$$\sigma_{el} = \frac{\pi}{p^2}\int dq^2|F(q)|^2 \approx \frac{\pi\xi^2}{\mu^2}\int_0^{q_0^2}\frac{dq^2}{q^2}J_1^2\left(\frac{q\xi}{\mu}\right) + \frac{\pi}{\mu^2}\int_{q_0^2}^{\infty}\frac{dq^2}{q^2}\log\left(\frac{\mu\beta N}{q}\right)\exp\left(-\frac{q^2}{2\Omega^2 N}\right), \quad (28)$$

which for $\xi \gg 1$ is simply

$$\sigma_{el} \approx \frac{\pi}{\mu^2}\xi^2. \quad (29)$$

### 3.2. The Quasi-Elastic Cross Section

In this paper, we focus on the coherent nucleus–atom interactions. The incoherent process of ionization of the target atom is suppressed by a factor on the order of $Z_2^{-1}$. Then the inelastic process, which by unitarity gives rise to attenuation of the elastic amplitude, is the process of the quasi-elastic scattering (Fig. 1a). Its cross section is given by [1]

$$p^2\frac{d\sigma_{Qel}}{d^2\mathbf{q}} = \sum_f |F_{fi}(q)|^2 - |F_{ii}(q)|^2, \quad (30)$$

where the sum extends over all final states of the crystal in which no particle production occurs. The closure relation then yields

$$\frac{d\sigma_{Qel}}{d^2\mathbf{q}} = \frac{1}{4\pi^2}\int d^2\mathbf{b}d^2\mathbf{b}'\exp[i\mathbf{q}(\mathbf{b}-\mathbf{b}')] \times\{\langle\exp[i\chi(\mu b)-i\chi^*(\mu b')]\rangle^N - \langle\exp[i\chi(\mu b)]\rangle^N\langle\exp[-i\chi^*(\mu b')]\rangle^N\} \quad (31)$$

and

$$\sigma_{Qel} = \int d^2\mathbf{b}\{1-|\langle\exp[i\chi(\mu b)]\rangle|^{2N}\}. \quad (32)$$

In the SL approximation,

$$|\langle\exp[i\chi(\mu b)]\rangle| = 1$$

and

$$\sigma_{Qel} = 0.$$

From (20) and the discussion of the absorption radius $b_a$ presented above, it follows that for $\xi \gg 1$,

$$1-|\langle\exp[i\chi(\mu b)]\rangle|^{2N} \approx \theta(2b_a - b) \quad (33)$$

and

$$\sigma_{Qel} \approx \pi(2b_a)^2 \approx \frac{\pi}{\mu^2}\xi^2. \quad (34)$$

### 3.3. The Total Cross Section

From Eq. (26), by means of the optical theorem, we find the total cross section

$$\sigma_{tot} = \frac{4\pi}{p}\mathrm{Im}F(0) \approx \frac{2\pi}{\mu^2}\xi^2. \quad (35)$$

We therefore conclude that at high energy and in the large-$N$ limit,

$$\sigma_{el} \approx \sigma_{Qel} \approx \frac{1}{2}\sigma_{tot}. \quad (36)$$

## 4. COULOMB EXCITATION OF ULTRARELATIVISTIC PARTICLES AND NUCLEI IN A CRYSTAL CHANNEL: THE EXCITATION CROSS SECTION TO THE LOWEST ORDER AND THE BORN APPROXIMATION

We now consider the process of the coherent Coulomb excitation of ultrarelativistic particles and nuclei passing through the crystal. This method for the experimental study of rare processes was proposed in [4–10].

The ultrarelativistic projectile nucleus (the mass number $A$, the charge $Z_1$, and the 4-momentum $p$) moving along a crystal axis undergoes a correlated series of soft collisions that give rise to diagonal ($A \longrightarrow A$, $A^* \longrightarrow A^*$) and off-diagonal ($A \longrightarrow A^*$, $A^* \longrightarrow A$) transitions.

In [4, 5, 9], it was proposed that the electric dipole transition in $^{19}$F be studied, the excitation of the state $|J^\pi = 1/2^-\rangle$ from the ground state $|1/2^+\rangle$. The phenomenological matrix element of the transition $1/2^+ \longrightarrow 1/2^-$ is [11]

$$\mathcal{M} = \frac{1}{2}d\bar{u}(p')\gamma_5(\hat{q}\hat{\varepsilon}-\hat{\varepsilon}\hat{q})u(p), \quad (37)$$

where $u(p')$ and $u(p)$ are bispinors of the initial and final states of the projectile, $d$ is the transition dipole moment, and $\varepsilon$ is the photon polarization vector. The transverse and longitudinal components of the 4-vector $p - p'$ are denoted by $\mathbf{q}$ and $\kappa$, respectively. In what follows, $q = |\mathbf{q}|$. The only phenomenological parameter in the problem is the dipole moment $d$. The measured lifetime of the 110-keV level $^{19}$F($1/2^-$) is $\tau = (0.853 \pm 0.010)\times 10^{-9}$ s [12]; the dipole moment of the $1/2^+ \longrightarrow 1/2^-$ transition, determined from the width of the level $^{19}$F($1/2^-$), is $d \approx 5\times 10^{-8}$ keV$^{-1}$ [11]. Then, first, because of the large value of $\tau$, the decay of the excited state inside the target crystal can be safely neglected, and, second, due to the smallness of $d$, the excitation amplitude is much smaller than the elastic Coulomb amplitude for all $q$ up to $q \sim \sqrt{4\pi\alpha}\,Z_1/d$ and can be considered a perturbation. Thus, the multichannel problem reduces to a one-channel one.

The high-energy helicity-flip Born amplitude of the transition $1/2^+ \longrightarrow 1/2^-$ in collision of the projectile nucleus with $N$ bound atoms in the crystal is given by

$$F_{\rm ex}^B(\mathbf{q}) = S(\kappa)\frac{p}{2\pi}\frac{g(\boldsymbol{\sigma}\cdot\mathbf{q})}{q^2+\lambda^2}\exp\left(-\frac{q^2}{4\Omega^2}\right), \qquad (38)$$

where $\boldsymbol{\sigma} = (\sigma_1, \sigma_2, \sigma_3)$ is the Pauli spin vector, $\{\sigma_i, \sigma_j\} = 2\delta_{ij}$, and the amplitude we are constructing is to be regarded as an operator that transforms the initial helicity state of the projectile into its final state. In the denominator of Eq. (38), $\lambda^2 = \mu^2 + \kappa^2$. In the Glauber approximation, the longitudinal momentum transfer, which determines the coherency length $l_c \sim \kappa^{-1}$, is given by [13]

$$\kappa = \frac{M\Delta E}{p}, \qquad (39)$$

where $M$ is the mass of the projectile and $\Delta E$ is the excitation energy.[2]

In the first order in $g$, the structure factor of crystal is

$$S(\kappa) = \exp\left[-\frac{\kappa^2}{4\Omega^2}\right]\frac{\sin(\kappa Na/2)}{\sin(\kappa a/2)}. \qquad (40)$$

If the projectile momentum satisfies the resonance condition [4, 5, 7, 9]

$$\frac{M\Delta E}{p} = \frac{2\pi n}{a}, \quad n = 0, 1, 2, \ldots, \qquad (41)$$

then $S(\kappa) \sim N$. In the first order in $g$ and in the zeroth order in $\beta$ (the Born approximation), the cross section of the coherent excitation of the projectile in scattering on a chain of $N$ atoms in a crystal is

$$\sigma_{\rm ex}^B = \frac{\pi}{p^2}\int dq^2\left|F_{\rm ex}^B(\mathbf{q})\right|^2 \sim \frac{g^2N^2}{4\pi}\left[\log\left(1+\frac{2\Omega^2}{\lambda^2}\right)-\frac{2\Omega^2}{\lambda^2+2\Omega^2}\right], \qquad (42)$$

where

$$g = \sqrt{4\pi\alpha}dZ_2.$$

The central idea in [4–7, 9, 10], based on the Born approximation, is that the transition rate can be enhanced substantially due to coherency of interactions, which is assumed to be sustained over a long-distance scale. The law $\sigma_{\rm ex} \propto N^2$ is expected to hold up to the crystal thicknesses $N = L/a \sim 10^5$–$10^6$ in a tungsten target. In [10], the Born approximation for the coherent excitation of $\Sigma^+$ in high-energy proton–crystal interactions $p\gamma \longrightarrow \Sigma^+$ was assumed to be valid up to $N \sim 10^8$.

[2] The Fresnel corrections to the eikonal approximation, which are neglected here, become important at large $N$ or at large $q$; they diminish the coherency length and additionally suppress coherent processes [14].

However, taking into account the initial and final state Coulomb interactions dramatically changes the dependence of $\sigma_{\rm ex}$ on $N$. For instance, at $N = 2$, the excitation amplitude is of the form

$$F_{\rm ex}^{(2)}(\mathbf{q}) = \frac{p}{\pi}\int d^2\mathbf{b}\exp(i\mathbf{q}\cdot\mathbf{b})\langle f_{\rm ex}^B\exp(i\chi)\rangle \times\langle\exp(i\chi)\rangle. \qquad (43)$$

The first of the two bracketed factors in Eq. (43) corresponds to the nuclear excitation amplitude in scattering on a bound atom. It differs from the excitation amplitude of the Born approximation, $f_{\rm ex}^B(\mathbf{b})$, by the multiplicative phase factor that occurs due to the initial and final state multiple Coulomb scattering. At small impact parameters $b \lesssim 1/2\Omega$,

$$\langle f_{\rm ex}^B\exp(i\chi)\rangle \approx S(\kappa)\frac{g}{2\pi b}(\boldsymbol{\sigma}\cdot\mathbf{n}_b) \times\sinh\left(\frac{1}{2}\Omega^2b^2\right)\exp\left(-\frac{1}{2}\Omega^2b^2\right). \qquad (44)$$

For large $b$,

$$\langle f_{\rm ex}^B\exp(i\chi)\rangle = S(\kappa)\frac{g}{4\pi}\int d^2\mathbf{s}\rho(s)(\boldsymbol{\sigma}\cdot(\mathbf{n}_b-\mathbf{n}_s)) \times\lambda K_1(\lambda|\mathbf{n}_b-\mathbf{n}_s|)\exp[i\chi(\mu|\mathbf{n}_b-\mathbf{n}_s|)] \qquad (45)$$

$$\approx S(\kappa)\frac{g}{4\pi}(\boldsymbol{\sigma}\cdot\mathbf{n}_b)\lambda K_1(\lambda b)\exp(i\chi)\exp\left(-\frac{\omega^2}{4\Omega^2}\right),$$

where $\mathbf{n}_b = \mathbf{b}/|\mathbf{b}|$, $\mathbf{n}_s = \mathbf{s}/|\mathbf{s}|$, and $b \gtrsim \mu^{-1}$. Because

$$\left|\langle f_{\rm ex}^B\exp(i\chi)\rangle\right|^2 \propto \Omega^2b^2$$

for small $b$ and

$$\mu bK_1^2(\mu b) \propto \exp(-2\mu b)$$

for large impact parameters $b \gtrsim 1/\mu$, the cross section

$$\sigma_{\rm ex}^{(2)} = \int d^2\mathbf{b}\left|\langle f_{\rm ex}^B\exp(i\chi)\rangle\right|^2|\langle\exp(i\chi)\rangle|^2 \approx 4\sigma_{\rm ex}^{(1)}\left(1-\frac{\omega^2}{2\Omega^2}\right) \qquad (46)$$

is dominated by $b \sim 1/2\mu$. For the diamond crystal, $\mu/\Omega \approx 0.16$ [3]. Hence,

$$\frac{\omega^2}{2\Omega^2} \approx \frac{2\beta^2\mu^2}{\Omega^2} \sim \frac{1}{20}.$$

This estimate shows that even for the diamond crystal target, the Born approximation is invalid already at $N \gtrsim 10$.

## 5. MULTIPLE SCATTERING EFFECTS AND ABSORPTION IN THE COHERENT COULOMB EXCITATION PROCESSES

The transition amplitude on a chain of $N$ identical atoms including all the multiphoton $t$-channel exchanges is given by

$$F_{ex}(\mathbf{q}) = \frac{p}{\pi}\int d^2\mathbf{b}\exp(i\mathbf{q}\cdot\mathbf{b})\langle f_{ex}^B\exp(i\chi)\rangle \times \langle\exp(i\chi)\rangle^{N-1}. \quad (47)$$

Because of both the multiple scattering effect and absorption, only large impact parameters $b \gg \mu^{-1}$ may contribute to $F_{ex}(\mathbf{q})$. Evaluation of $F_{ex}(q)$ then gives

$$F_{ex}(\mathbf{q}) \approx \frac{gp}{2\pi}S(\kappa)(\boldsymbol{\sigma}\cdot\mathbf{n}_q)\int_{1/\mu}^{\infty} bdbJ_1(qb) \times \lambda K_1(\lambda b)\exp(iN\chi)\exp(-N\omega^2/4\Omega^2), \quad (48)$$

where $\mathbf{n}_q = \mathbf{q}/|\mathbf{q}|$. The contribution of the domain $q \lesssim q_0 = \mu/\xi$ to the excitation cross section can be neglected because $F_{ex} \propto q$ in this region. If $q \gg q_0$ and $\xi \gg 1$, the stationary phase approximation gives the coherent excitation amplitude of the form

$$F_{ex}(\mathbf{q}) \approx \frac{ipg(\boldsymbol{\sigma}\cdot\mathbf{n}_q)}{2\pi\beta}\frac{S(\kappa)}{N}\frac{\lambda}{\mu}\sqrt{\eta}\exp(-\delta\eta) \times \exp\left(-\frac{iq\eta}{\mu}\right)\exp\left(-\frac{q^2}{4\Omega^2 N}\right). \quad (49)$$

We see that the helicity-flip dynamics removes the factor $1/q$ from the elastic amplitude (27), thus making the UV regularization of the excitation cross section indispensable. This cross section is evaluated as

$$\sigma_{ex} = \frac{\pi}{p^2}\int dq^2|F_{ex}(\mathbf{q})|^2 \sim \frac{g^2N^{1-\delta}}{8\pi}C\log\left(\frac{N}{\delta\gamma}\right), \quad (50)$$

where

$$C = \gamma^{\Delta}\Delta^2\Gamma(\Delta), \quad \gamma = 2\Omega^2/\beta^2\mu^2,$$

$$\Delta = \lambda/\mu, \quad \delta = \Delta - 1 \sim \kappa^2/2\mu^2 \ll 1.$$

In (50), we simply set $S(\kappa) = N$. Thus, the account of multiple scatterings and absorption turns the Born approximation cross section

$$\sigma_{ex} \propto N^2$$

into

$$\sigma_{ex} \propto N^{1-\delta}\log N.$$

In the limit as $p \longrightarrow \infty$ and $\delta \longrightarrow 0$,

$$\sigma_{ex} \sim \frac{g^2N}{8\pi}\gamma\log\left(\frac{N}{\gamma}\right). \quad (51)$$

The dependence of $\sigma_{ex}$ on $N$ differs from that of the fully unitarized elastic cross section,

$$\sigma_{el} \propto \log^2 N.$$

The reason is that in $\sigma_{ex}$, we sum the eikonal diagrams to all orders in $\beta$ but only to the first order in $g$. Such a unitarization procedure is, of course, incomplete, but this is not important for practical purposes because the smallness of $d^2\Omega^2$ makes the next-to-leading-order terms negligibly small up to

$$N \sim \alpha Z_1^2/\delta\Omega^2 d^2 \sim 10^{12}.$$

## 6. SUMMARY

The main goal we pursued in this paper is a consistent description of the coherent Coulomb excitation of ultrarelativistic particles and nuclei passing through the aligned crystal. We started with a discussion of the elastic scattering and found that taking into account the lattice thermal vibrations within the Glauber multiple scattering theory gives rise to a strong absorption effect. The radius of the absorption region in the impact parameter space appeared to grow logarithmically as the crystal thickness grows. We derived a convenient representation for the scattering matrix with absorption and calculated the coherent elastic and the incoherent quasi-elastic cross sections. The suppression of scattering amplitudes in the absorption region was shown to serve as a natural UV regulator and enables consistent calculation of the cross section $\sigma_{ex}$ of the coherent nuclear excitation. The dependence of $\sigma_{ex}$ on the crystal thickness was found. The multiple scattering effects were shown to become numerically important already at $N \gtrsim 1$, thus leaving no room for the Born approximation widely used in early analyses of the problem.

## ACKNOWLEDGMENTS

The author thanks N.N. Nikolaev for useful comments.

## REFERENCES

1. R. J. Glauber, in *Lectures in Theoretical Physics*, Ed. by W. E. Brittin *et al.* (Interscience, New York, 1959), Vol. 1, p. 315.

2. N. P. Kalashnikov, E. A. Koptelov, and M. I. Ryazanov, Zh. Éksp. Teor. Fiz. **63**, 1107 (1972) [Sov. Phys. JETP **36**, 583 (1972)]; N. P. Kalashnikov and V. D. Mur, Yad. Fiz. **16**, 1117 (1972) [Sov. J. Nucl. Phys. **16**, 613 (1973)].
3. D. S. Gemmel, Rev. Mod. Phys. **46**, 1 (1974).
4. V. V. Okorokov, Yad. Fiz. **2**, 1009 (1965) [Sov. J. Nucl. Phys. **2**, 719 (1966)]; V. V. Okorokov, Yu. L. Pivovarov, A. A. Shirokov, and S. A. Vorobev, Preprint ITEP-90-49 (Inst. of Theoretical and Experimental Physics, Moscow, 1990).
5. V. V. Okorokov and S. V. Proshin, Preprint ITEP-13-1980 (Inst. of Theoretical and Experimental Physics, Moscow, 1980).
6. Yu. L. Pivovarov, H. Geissel, Yu. M. Filimonov, O. E. Krivosheev, and C. Scheidenberger, Preprint GSI-95-38 (Darmstadt, 1995).
7. Yu. L. Pivovarov and A. A. Shirokov, Yad. Fiz. **37**, 1101 (1983) [Sov. J. Nucl. Phys. **37**, 653 (1983)].
8. Yu. L. Pivovarov, A. A. Shirokov, and S. A. Vorobev, Nucl. Phys. A **509**, 800 (1990).
9. R. Fusina and J. C. Kimball, Nucl. Instrum. Methods Phys. Res. B **33**, 77 (1988).
10. A. Dubin, Yad. Fiz. **52**, 1243 (1990) [Sov. J. Nucl. Phys. **52**, 790 (1990)].
11. V. R. Zoller, Pis'ma Zh. Éksp. Teor. Fiz. **75**, 147 (2002) [JETP Lett. **75**, 119 (2002)].
12. F. Ajzenberg-Selove, Nucl. Phys. A **190**, 1 (1972).
13. V. N. Gribov, Zh. Éksp. Teor. Fiz. **57**, 1306 (1970) [Sov. Phys. JETP **30**, 709 (1970)].
14. V. R. Zoller, Phys. Lett. B **416**, 447 (1998).

*Journal of Experimental and Theoretical Physics, Vol. 99, No. 4, 2004, pp. 792–800.*
*Translated from Zhurnal Éksperimental'noĭ i Teoreticheskoĭ Fiziki, Vol. 126, No. 4, 2004, pp. 915–925.*


**SOLIDS**
**Electronic Properties**

# Classical and Quantum Percolation Phenomena and Controlling Them in Eutectic Semiconductor–Superconductor Compositions

**G. I. Isakov**

*Institute of Physics, Academy of Sciences of Azerbaijan, pr. Dzhavida 33, Baku, 370143 Azerbaijan*
*e-mail: gudrat@physics.ab.az*

Received December 22, 2003

**Abstract**—Experimental data on the conduction of heterogeneous systems have been traditionally interpreted in the context of the theory of percolation phenomena taking into account the relative threshold volume fraction ($\eta_c \approx 0.16$) of the high-conductivity phase. This work is concerned with the conduction of eutectic compositions semiconductor–normal metal at $T > T_c$ (the classical limit) and semiconductor–superconductor at $T < T_c$ (the quantum limit) obtained at various material growth rates; these materials contain metal particles as oriented whiskers in semiconducting matrices. The paper presents spatial and energy models of discrete, finite, and infinite clusters that well explain classical and quantum percolation conductivities. Depending on the growth rate of eutectic compositions, their classical and quantum conductivities can manifest themselves at arbitrary percolation thresholds $\eta_p$ ($0 < \eta_p \le \eta_c$). It is shown that the density of whiskers, the distances between them, their diameters, and the critical supercurrent density per whisker can be controlled by varying the rate of composition growth. © *2004 MAIK "Nauka/Interperiodica".*

## 1. INTRODUCTION

The determination of the mechanisms of classical and quantum phenomena and controlling them in dielectric (I) and semiconducting (Sm) materials, normal metals (N), superconductors (S), and composite structures $I_1$–$I_2$, I–Sm, I–N, I–S, $Sm_1$–$Sm_2$, Sm–N, Sm–S, $N_1$–$N_2$, N–S, and $S_1$–$S_2$ are fundamental problems of modern experimental, theoretical, and applied physics.

In [1, 2], we determined the mechanisms and possibilities of controlling such classical kinetic phenomena as heat conductivity determined by changes in phonon processes and the tensometric effect determined by changes in electronic processes in semiconductor–normal metal eutectic compositions. We also showed that semiconductor–normal metal eutectic compositions were good models of inhomogeneous compensated semiconductors [3]. The special features of the large-scale relief in these materials depend on the morphology of the eutectic compositions, the degree of doping, and the degree of compensation of their semiconducting matrices. We found that electrical conduction, Hall effect, thermal electromotive force, heat conductivity, and the other classical physical effects exhibited strong anisotropy in semiconductor–normal metal eutectic compositions [4–9]. In these compositions at various mutual electric current, heat flux, and magnetic field directions and directions of metallic cylinders, ellipsoids, plates and rods, and other forms, the anisotropy of classical phenomena can be controlled within wide ranges depending on changes in the electronic and phonon subsystems.

In recent years, much attention has been given to quantum phenomena in symmetrical high-$T_c$ superconducting bicrystalline Josephson junctions and the dependence of their electrophysical properties on the grain-boundary angle [10–12].

Among numerous Josephson junctions, superconductor–semiconductor–superconductor (S–Sm–S) junctions hold a special position, primarily because of the possibility to control semiconducting interlayer parameters by doping and applying external electric and magnetic fields [13–18].

In [19], we reported the observation of quantum superconductivity and Josephson effect phenomena for a eutectic semiconductor–superconductor composition. In [20], we showed that the superconductivity of a semiconductor–superconductor eutectic could be controlled by varying the angle between the electric current and superconducting filamentary crystal (whisker) directions.

This work is concerned with GaSb–$V_2Ga_5$ eutectic compositions obtained at various growth rates in the normal and superconducting states. The work presents original results on the mechanism and control of classical and quantum percolation phenomena in eutectic

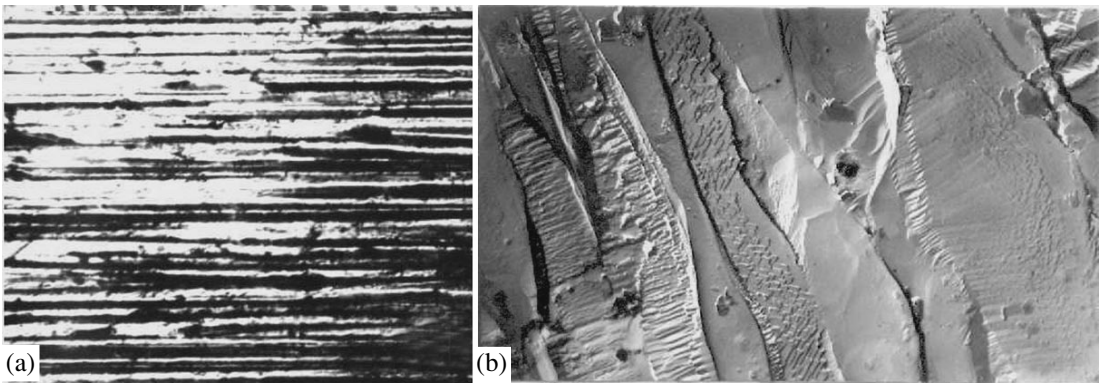



**Fig. 1.** Photomicrographs of the surface of the eutectic composition GaSb–$V_2Ga_5$ obtained at a $v$ = 7 cm/h growth rate parallel to whiskers: (a) magnification ×200 and (b) magnification ×3000.

compositions above and below the superconducting transition temperature.

## 2. EXPERIMENTAL RESULTS

GaSb–$V_2Ga_5$ eutectic compositions were prepared by the Bridgman method at various crystallization front rates ($v$ = 0.85, 2.8, 7, 22, 70, and 200 cm/h). The preparation of eutectics at various growth rates is of interest because changes in the growth rate can influence the morphology of the products and geometric dimensions of metallic whiskers. The dependence of the distance between whiskers $R$ on the rate of eutectic crystallization is given by the equation [21]

$$R^2 v = \text{const}. \quad (1)$$

The data on the density of whiskers, their size, and distances between them at various crystallization front rates are listed in the table.

The morphology of the samples was studied using a MIM-8M microscope and an electron microscope. Photomicrography (magnification ×200) of the surface of the GaSb–$V_2Ga_5$ eutectic composition parallel to the growth direction ($v$ = 7 cm/h) obtained using the MIM-8M microscope is shown in Fig. 1a. According to this figure, the metallic $V_2Ga_5$ phase in the GaSb matrix is formed in the form of whiskers largely oriented along the direction of crystallization front motion.

We see in Fig. 1a certain irregularities in the density of metallic whiskers; there are also some voids in the material. Irregularities in the growth of the GaSb–$V_2Ga_5$ eutectic composition are also quite obvious in electron microscopic images. An electron microscopic image of the surface of the eutectic composition along the direction of growth of metallic whiskers at $v$ = 7 cm/h is shown in Fig. 1b. There are three whiskers in this figure, one at the left bottom corner, another closer to the center, and the third at the top right corner of the figure. The figure shows that whiskers grow irregularly. They bend, and their diameters decrease. On the right of the whisker in the center, there is an empty bent channel that originally contained a metallic whisker that was fully etched off when the surface was treated with an etching agent. It is clearly seen that the diameter of the empty channel is not the same everywhere. Photomicrograph 1b was obtained with a magnification of ×3000. In the region of narrowing, the diameter of metal whiskers is approximately 0.5 μm. Eutectic growth irregularities and the appearance of regions with whisker narrowings are responsible for the Josephson effect [19] in this composition. Note that the

Growth rate ($v$) of eutectic GaSb–$V_2Ga_5$ compositions, whisker densities ($n$) and diameters ($d$), microbridge diameters ($d'$), distances between whiskers ($R$), and critical current densities ($I$) per infinite whisker

| Composition no. | $v$, cm/h | $n$, $10^4$ mm$^{-2}$ | $d$, μm | $d'$, μm | $R$, μm | $I$, A/cm$^2$ |
|---|---|---|---|---|---|---|
| 1 | 0.85 | 0.4 | 6 | 1–1.5 | 8 | $10^5$ |
| 2 | 2.8 | 1.2 | 3 | 0.7–1 | 6 | $10^4$ |
| 3 | 7 | 2 | 2 | 0.5–0.6 | 4–5 | $10^3$ |
| 4 | 22 | 6 | 1 | 0.3–0.4 | 3–4 | $(3–4)\times10^2$ |
| 5 | 70 | 10–20 | 0.3–0.5 | 0.1–0.2 | 1–3 | $(1–2)\times10^2$ |
| 6 | 200 | – | – | – | – | – |

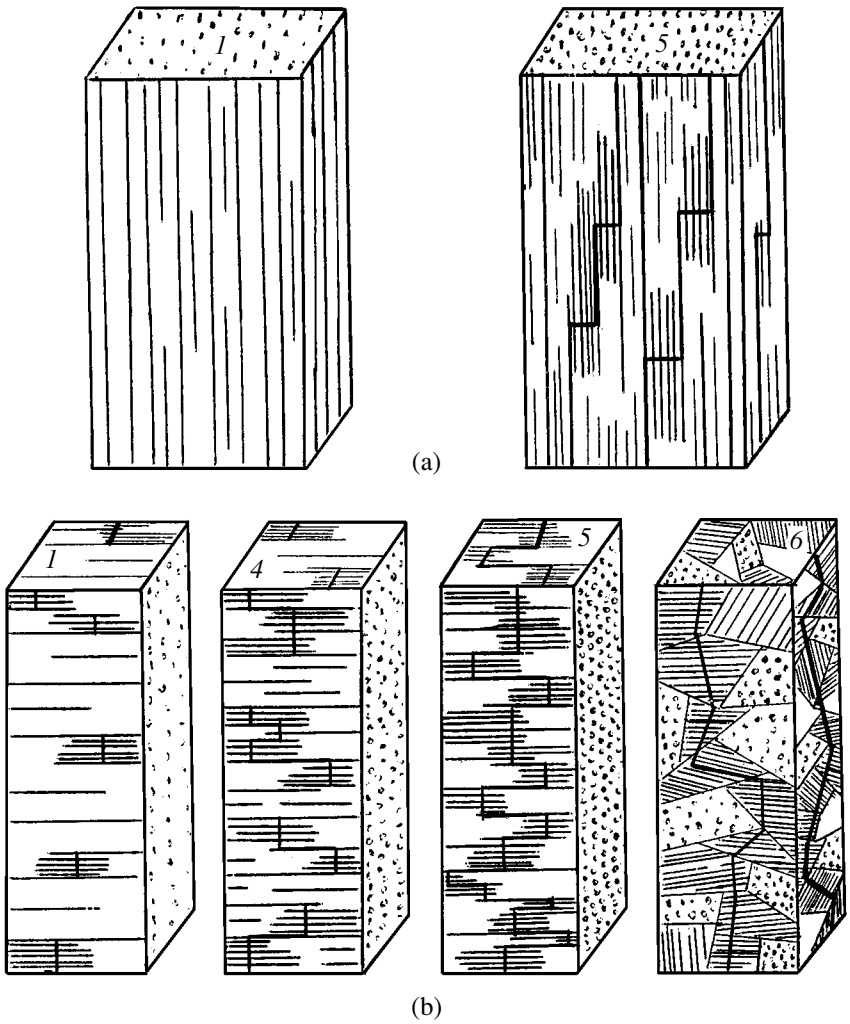



**Fig. 2.** Schematic models of eutectic compositions obtained at various growth rates $v$: (a) $I \parallel X$, $v$ = (*1*) 0.85 and (*5*) 70 cm/h and (b) $I \parallel X$, $v$ = (*1*) 0.85, (*4*) 22, (*5*) 70, and (*6*) 200 cm/h. Composition or curve numbers in all figures correspond to those in the table.

volume of the $V_2Ga_5$ superconducting phase was 4.4% of that of the GaSb–$V_2Ga_5$ eutectic composition.

Two kinds of samples were cut from eutectic compositions obtained at various growth rates. The samples had the shape of rectangular parallelepipeds (Fig. 2). Height $h$ of the parallelepipeds was either parallel to the direction of composition growth $X$ (the orientation of whiskers), $h \parallel X$ (samples a), or perpendicular to this direction, $h \perp X$ (samples b). The specific resistances $\rho$ of the samples were measured over the temperature range 2–300 K. In $h \parallel X$ samples, the electric current $I$ direction through a sample was parallel to whiskers $X$ ($I \parallel X$). For $h \perp X$ samples, we had $I \perp X$.

The temperature dependences of the specific resistances $\rho_\parallel$ of the samples obtained at various growth rates $v$ measured with $I \parallel X$ are shown in Fig. 3. According to this figure, the specific resistances $\rho_\parallel$ of all samples except sample 6 were close to each other at $T$ = 300 K. The $\rho_\parallel(T)$ curves diverged as the temperature decreased, and, at $T \le 200$ K, the specific resistances of the samples were sharply different. The temperature dependences of the specific resistances of the samples, $\rho_\parallel(T)$, were metallic in character over the whole temperature range. However, the slope of curves *1–5* decreased as the growth rate increased. Curve *6* was weakly metallic in character over the whole temperature range.

According to Fig. 3, all the samples under study underwent superconducting transition at $T \le 5.5$ K. It follows that, at $I \parallel X$, we can write the following inequalities for $v$ and $\rho_\parallel$ (at equal temperatures):

$$v_1 < v_2 < \ldots < v_6, \tag{2}$$

$$\rho_{\parallel 1} < \rho_{\parallel 2} < \ldots < \rho_{\parallel 5} \ll \rho_{\parallel 6}. \tag{3}$$

The temperature dependences of the specific resistances $\rho_\perp$ ($I \perp X$) of samples obtained at various growth rates $v$ are shown in Fig. 4. According to this figure, the

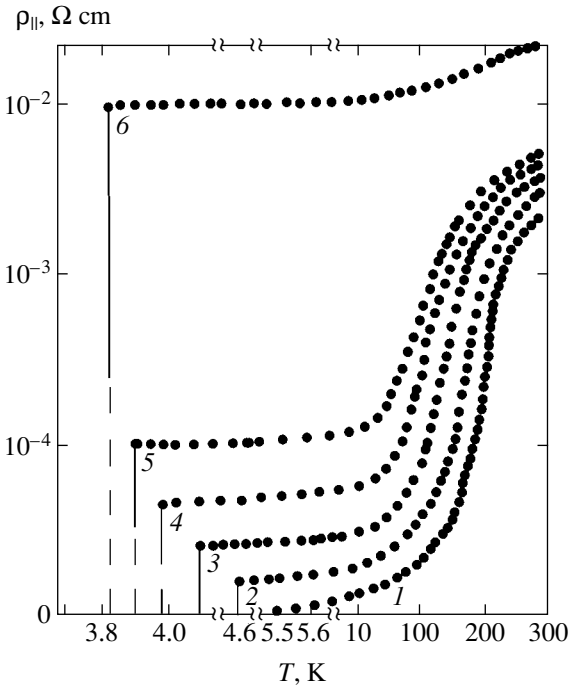



**Fig. 3.** Temperature dependences of specific resistances $\rho_{\parallel}$ of eutectic compositions grown at different rates $v$ ($I \parallel X$).

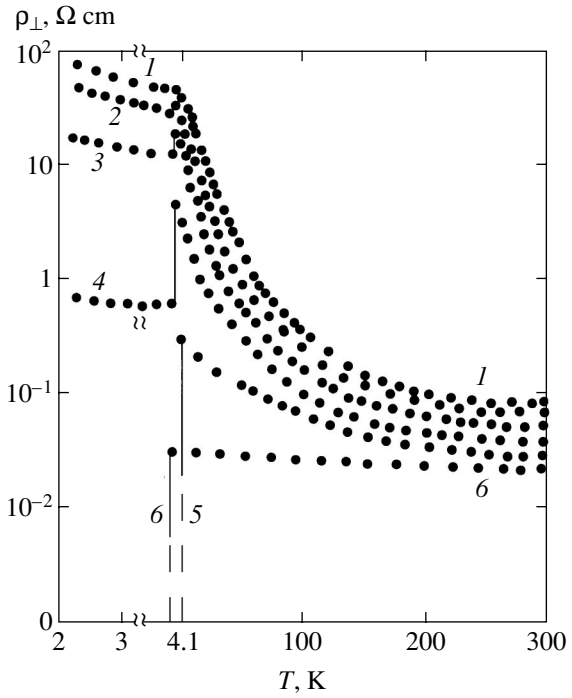



**Fig. 4.** Temperature dependences of specific resistances $\rho_{\perp}$ of eutectic compositions grown at different rates $v$ ($I \perp X$).

specific resistances $\rho_{\perp}(T)$ of the samples are again close to each other at $T = 300$ K. The $\rho_{\perp}$ curves diverge as the temperature decreases. The temperature dependences of the specific resistances $\rho_{\perp}$ of the samples, except sample 6, are semiconducting in character. Semiconducting dependences, however, become weaker as the growth rate increases. For sample 6, $\rho_{\perp}$ almost does not change over the temperature range 300–3.8 K. It follows that, for $I \perp X$, we can write the following inequalities for $v$ and $\rho_{\perp}$ (at equal temperatures):

$$v_1 < v_2 < \ldots < v_6,$$
$$\rho_{\perp 1} > \rho_{\perp 2} > \ldots > \rho_{\perp 6}. \quad (4)$$

We also see that, below $T = 4.1$K, the shape of the $\rho_{\perp}(T)$ curves for samples 1 and 2 strongly changes (Fig. 4). The specific resistances of samples 3 and 4 sharply decrease close to 4.1 K; this is followed by a change in the shape of the $\rho_{\perp}(T)$ dependences. Samples 5 and 6 undergo the superconducting transition at $T = 4.1$ K.

## 3. DISCUSSION

Note that the diameters of whiskers are not everywhere equal (Fig. 1b). In the region of whisker narrowings, their statistical mean diameter $d'$ is on the order of the coherence length ($d' \approx 10^{-4}$–$10^{-5}$ cm). For this reason, each whisker in a eutectic composition represents variable-thickness microbridges connected in series. We showed in [19] that, when electric current was parallel to superconducting whiskers $X$ ($I \parallel X$), the Josephson effect was observed. These variable-thickness microbridges are therefore weak Josephson bonds. Each whisker in a eutectic sample is an S–S'–S–S'–S… chain (Fig. 2a, *1*), where S is a superconducting whisker (or a superconducting bank) and S' is a Josephson variable-thickness microbridge (or a Josephson weak bond).

A metallographic study of the surface of the samples shows that the number of whiskers penetrating them and the lengths of many other whiskers decrease as the growth rate $v$ of the eutectic composition increases. However, new whiskers grow close to broken ones. These whiskers are weakly bound with each other by Josephson junctions of the type superconductor–semiconductor–superconductor (S–Sm–S). For this reason, two types of chains penetrating the sample, namely, S–S'–S–S'–S… and S–S'–S–Sm–S–S'–S…, are responsible for superconductivity in the $I \parallel X$ direction (Fig. 2a, *5*). These chains will be called infinite clusters in what follows. When $I$ is parallel to $X$, infinite clusters of both types are connected in parallel. At low sample growth rates, Josephson infinite clusters of the type S–S'–S–S'–S… predominate in superconductivity. The number of S–S'–S–S'–S… chains decreases as the rate of sample growth increases, and the contribution of superconducting clusters of the type S–S'–S–Sm–S–S'–S… begins to grow.

Note that infinite clusters of the type S–S'–S–Sm–S–S'–S… are bound more weakly than infinite clusters

of the type S–S'–S–S'–S… because of the presence of Sm semiconducting interlayers. For this reason, when electric current $I$ is parallel to whiskers $X$ ($I \parallel X$), infinite clusters of the S–S'–S–S'–S… type predominate in all samples except those grown at a rate of $v = 200$ cm/h. A study of the surface of the composition grown at $v = 200$ cm/h showed that the orientation relations between the matrix and whiskers were strongly disturbed in this sample (Fig. 2b, *6*). For this reason, the specific resistance $\rho_\parallel$ of this composition was 2–3 orders of magnitude higher at low temperatures than those of the other samples.

At $I \parallel X$ above the superconducting transition temperature $T_c$, infinite S–S'–S–S'–S…. and S–S'–S–Sm–S–S'–S… clusters transform into N–N'–N–N'–N… and N–N'–N–Sm–N–N'–N… clusters, where N is the normal metal, N' is the normal metal narrowing, and Sm is the semiconducting interlayer between normal metals. At low composition growth rates, the number of infinite clusters comprising N–N'–N–N'–N… chains is much larger (Fig. 2a, *1*) than the number of infinite N–N'–N–Sm–N–N'–N… clusters. The number of infinite N–N'–N–N'–N… clusters decreases and the contribution of infinite N–N'–N–Sm–N–N'–N… clusters grows as the rate of composition growth increases (Fig. 2a, *5*). Note that the resistance of the semiconducting interlayer ($R_{Sm}$) is much larger than the resistance of normal metals ($R_N$), $R_{Sm} \gg R_N$. For this reason, the specific resistance $\rho_\parallel$ decreases as the growth rate becomes lower (Fig. 3). The temperature dependences of the specific resistances of the compositions obtained at low growth rates exhibit steeper metallic trends. Note that, at $I \parallel X$, normal electric current flows in infinite N–N'–N–N'–N… and N–N'–N–Sm–N–N'–N… clusters connected in parallel. A decrease in the growth rate also causes an increase in the diameter of the metallic whiskers and metallic bridges on them (see table).

It follows that the specific resistances of eutectic compositions $\rho_\parallel$ can be controlled by varying the rate of their growth.

According to Fig. 4, the specific resistances $\rho_\perp$ of eutectic compositions [electric current $I$ is perpendicular to whiskers $X$ ($I \perp X$)] can also be controlled by varying the rate of eutectic composition growth. The semiconducting character of the dependences of the specific resistances $\rho_\perp$ is caused by current passage through junctions normal metal (N)–semiconductor (Sm)–normal metal (N) (N–Sm–N). At a $v = 0.85$–70 cm/h rate of eutectic composition growth, samples largely contain N–Sm–N–Sm–N… chains (Fig. 2b, *1*–*5*). All compositions except sample 6 (Fig. 2b, *6*) exhibit semiconductor-type dependences (Fig. 4). A change in the temperature dependence of the resistance $\rho_\perp$ or a rapid decrease in $\rho_\perp$ at $T = 4.1$ in samples 1, 2, 3, and 4 is related to the transition to the superconducting state of metallic whiskers and the appearance of Josephson junctions comprising S–Sm–S–Sm–S….

The weakening of semiconductor-like behavior of $\rho_\perp$ in sample *5*, its complete elimination in sample *6*, and transition of both samples to the superconducting state manifests the formation of infinite clusters of S–Sm–S–Sm–S… ($v = 70$ cm/h) and S–S'–S–Sm–S–Sm–S–S'–S… ($v = 200$ cm/h) Josephson chains. At a $v = 70$ cm/h eutectic composition growth rate, statistical mean distances between superconducting whiskers are of 1–3 μm (see table), which is closer to weak binding conditions in semiconductors.

The inverse proportionality $\rho_\perp \propto f(1/v)$ is caused by a decrease in the thickness of the Sm semiconducting interlayer as a result of an increase in the rate of composition growth. For this reason, at low growth rates, discrete N–Sm–N–Sm–N… cascades appear in the samples.

The conductivities of strongly inhomogeneous media are explained by percolation phenomena in experimental and theoretical works [22, 23]. According to the percolation conductivity condition, the relative critical volume $\eta_c$ of the high-conductivity phase should exceed some threshold value. For instance, $\eta_c \approx 0.16$ for isometric grains and can be slightly smaller than 0.16 for differently shaped grains. Note that the threshold value $\eta_c$ for percolation conductivity was treated in [22, 23] as averaged over the whole material volume. In these works, arbitrary distributions over the volume of a material were ignored. For instance, when eutectic compositions of the semiconductor–metal type are obtained as a result of oriented crystallization, one infinite metallic whisker is quite sufficient for the appearance of percolation conductivity at $I \parallel X$ in both the quantum ($T < T_c$) and classical ($T > T_c$) limits. The threshold $\eta_c$ value is then closer to zero. Note that, in the GaSb–$V_2Ga_5$ compositions, the relative volume of $V_2Ga_5$ whisker crystals was $\eta_{c1} \approx 0.04$, and, at $I \perp X$, we had $\eta_{c1} \ll \eta_c \approx 0.16$. Depending on the rate of eutectic composition growth, the threshold $\eta_c$ value can therefore be attained in discrete, long finite, and "infinite" composition regions (Figs. 2–4). Theory has ignored such percolation conductivity features thus far.

Note that the tunneling mechanisms in S–Sm–S-type junctions were theoretically considered in [13–15] for three cases: (a) nondegenerate semiconductors (low charge carrier concentrations in the Sm semiconducting interlayer), (b) intermediate charge carrier concentrations, and (c) degenerate semiconductors.

When a semiconductor is nondegenerate, the chemical potential μ is far below the bottom of the conduction band and impurity levels $E_d$ are scattered close to the chemical potential μ. The spread of impurity level values can be large compared with the barrier height $V - \mu$. It is assumed that the gap value or order parameter Δ is zero in the semiconductor. All free electrons then pass from the semiconductor to superconductors. Such junctions behave like superconductor–insulator–superconductor (SIS) junctions. The thickness of the

insulator is then of several dozen angstrom units. In our samples, the thickness of the semiconducting layer was on the order of one micron.

At strong degeneracy, the major tunneling mechanism is the neighborhood effect, and the junction behaves like a superconductor–normal metal–superconductor (SNS) junction. The normal metal thickness in such junctions can be as large as several tens of thousands of angstrom units.

In our compositions, the thickness of the semiconducting layer in junctions (S–Sm–S), cascades of junctions (S–Sm–S–Sm–S…), and finite and infinite clusters (S–S'–Sm–S'–S…) was of the same order of magnitude as the normal metal thickness in SNS junctions. The concentration of carriers in GaSb ($n = 10^{16}$ cm$^{-3}$), however, corresponded to intermediate concentrations [14]. In such junctions, resonance tunneling of Cooper pairs occurs. Because of resonance tunneling, the thickness of the Sm semiconducting interlayer in the case of a nondegenerate semiconducting barrier can be larger than the thickness of the dielectric layer in SIS junctions by several orders of magnitude. Overlapping of the wavefunctions of superconductors is then unnecessary.

Resonance tunneling of electrons was also considered in [24]. It was shown that, if the energy of electrons was close to impurity levels, resonance passage of electrons in the junction over special impurity configurations was possible. It was found in [14] that the probability of formation of a trajectory with strictly periodic arrangement of impurities was zero. Resonance-percolation trajectories appeared as the concentration of the impurity increased up to the development of degeneracy, and Cooper pairs were tunneled along these trajectories.

We found that interphase boundaries in semiconductor–metal eutectic compositions were strongly inhomogeneous [3, 8], and many features of transport phenomena in these materials were determined by interphase boundaries. For this reason, interphase boundaries in S–Sm–S and N–Sm–N junctions, their cascades, and finite and infinite clusters S–S'–S–Sm–S–Sm–S–S'–S… and N–N'–N–Sm–N–Sm–N—N'–N… are also strongly inhomogeneous. These inhomogeneous regions can overlap because of various fluctuation effects during eutectic composition growth. The overlap of inhomogeneous regions between whiskers at $T < T_c$ is responsible for the Josephson effect. The top of the valence band and the bottom of the conduction band in inhomogeneous semiconductors are modulated. The formation of an impurity band and interphase boundary inhomogeneities then increase the probability of the appearance of resonance-percolation trajectories for both Cooper pairs (at $T < T_c$) and single electrons (at $T > T_c$). The bottom of the conduction band in S–Sm–S and N–Sm–S junctions, their cascades, and infinite clusters lowers with respect to the chemical potential μ level as the thickness of Sm semiconducting interlayers increases; it decreases on both sides of the semiconducting barrier and is modulated only in a narrow region of interphase boundaries. Two- or one-particle percolation currents on both sides of the barrier then flow over the modulated bottom of the conduction band. In the middle of a semiconductor, these currents can flow along resonance-percolation trajectories. In thinner semiconducting interlayers Sm, the modulated bottom of the conduction band decreases with respect to the chemical potential μ level over the whole semiconductor thickness. Both two- and one-particle currents then flow only over the modulated bottom of the conduction band of the semiconductor. A schematic model of variable-thickness microbridges in an infinite S–S'–S–S'–S… whisker cluster and a scheme of the band structure of a cascade of S–Sm–S–Sm–S… Josephson junctions are shown in Fig. 5.

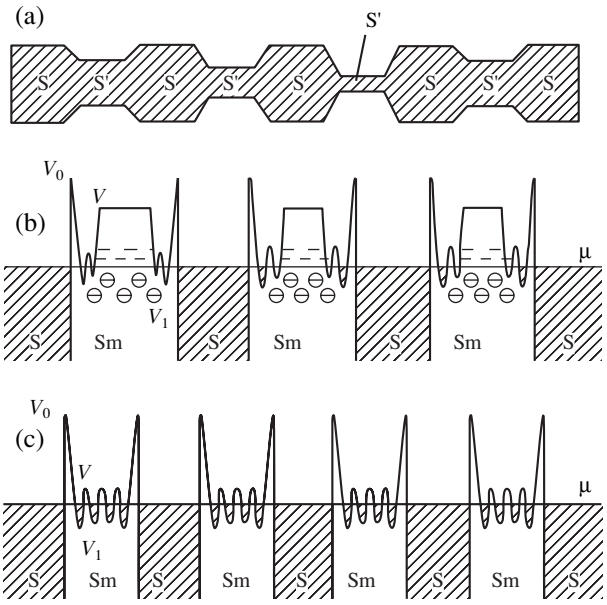



**Fig. 5.** Spatial and energy models of clusters: (a) cluster S–S'–S–S'–S… with variable-thickness bridges, (b) band structure of S–Sm–S–Sm–S… clusters with thick semiconducting bridges Sm, and (c) band structure of S–Sm–S–Sm–S… clusters with thin semiconducting bridges Sm.

The dependences of the specific resistances $\rho_\perp$ and $\rho_\parallel$ on the growth rate of eutectic compositions at various temperatures are shown in Fig. 6. As follows from this figure, $\rho_\perp$ decreases and $\rho_\parallel$ increases as the rate grows. At a $v = 200$ cm/h growth rate, the $\rho_\perp(v)$ and $\rho_\parallel(v)$ curves converge into one bundle. An interesting feature of $\rho_\perp(v)$ and $\rho_\parallel(v)$ is their approach to each other at high composition growth rates and temperatures. The $\rho_\perp(v)$ and $\rho_\parallel(v)$ curves diverge at low temperatures and composition growth rates.

As distinct from $\rho_\perp$, an increase in $\rho_\parallel$ as the rate of composition growth increases can be explained by a decrease in the diameter $d$ of metallic whiskers and the accompanying breaking of numerous bridges (N') in

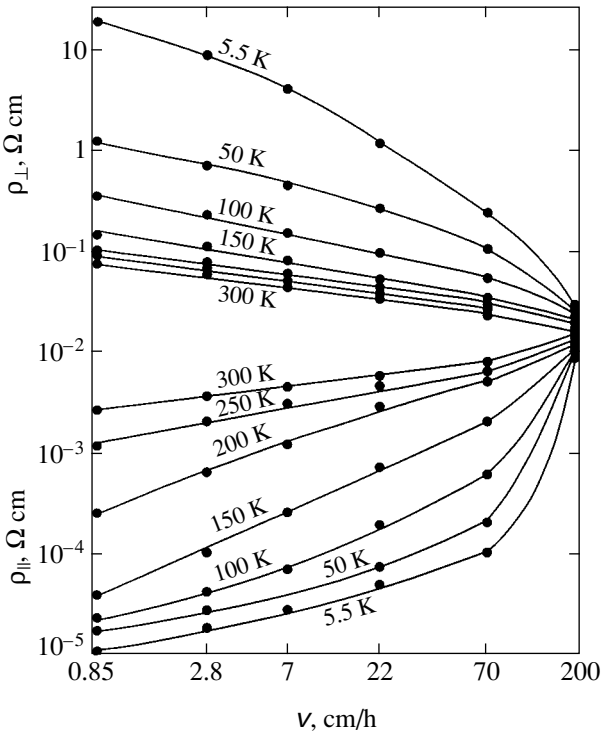



**Fig. 6.** Dependences of $\rho_\perp$ and $\rho_\parallel$ on the rate $v$ of composition growth.

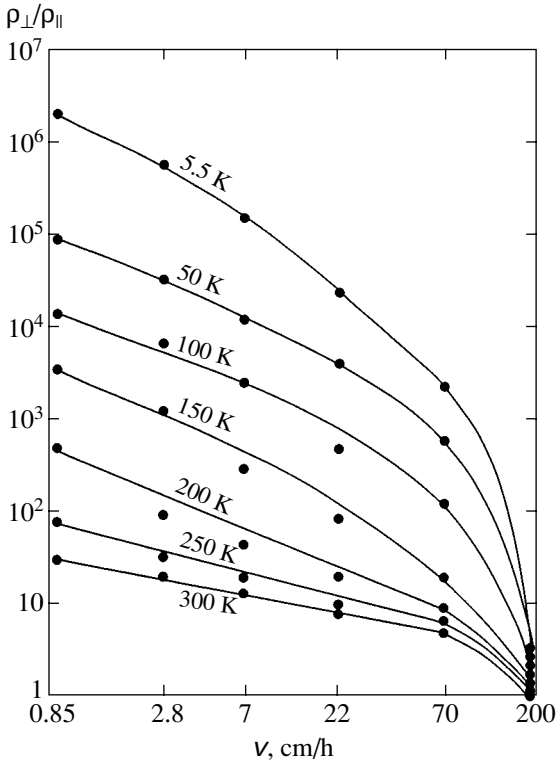



**Fig. 7.** Dependences of anisotropy coefficients $\rho_\parallel/\rho_\perp$ on the rate $v$ of composition growth.

N–N'–N–N'–N… and N–N'–N–Sm–N–N'–N… clusters. In both clusters, semiconducting interlayers appear in place of the broken bridges. Part of both clusters therefore transform into N–Sm–N–Sm–N… clusters. The disappearance of the orientation relation in composition 6 results in the gathering of the $\rho_\perp(v)$ and $\rho_\parallel(v)$ curves into one bundle. Note that, as distinct from the other compositions, orientation relations between the whiskers and matrix in composition 6 are strongly distorted. The composition consists of separate cells. Inside each cell, whiskers are parallel and arranged more closely. For this reason, the probabilities of the appearance of infinite clusters or resonance-percolation trajectories are higher in composition 6 than in the other compositions at an arbitrary percolation threshold.

These peculiarities are also observed for the dependences of the anisotropy coefficients $\rho_\perp/\rho_\parallel$ on the rate of composition growth. It is clearly seen in Fig. 7 that the anisotropy coefficient curves $\rho_\perp/\rho_\parallel$ gather into a bundle as the growth rate increases. The anisotropy coefficients $\rho_\perp/\rho_\parallel$ decrease substantially in magnitude as the temperature grows. The bundle of anisotropy coefficient curves diverges as the growth rate decreases. Maximum anisotropy coefficients are observed at low temperatures.

The temperature dependences of the anisotropy coefficients of specific resistances $\rho_\perp/\rho_\parallel$ of eutectic compositions obtained at various growth rates are shown in Fig. 8. The anisotropy coefficients $\rho_\perp/\rho_\parallel$ in samples 1–5 strongly decrease as the temperature increases. This observation can be explained by the opposite temperature dependences of $\rho_\perp$ and $\rho_\parallel$. That is, the specific resistances $\rho_\parallel$ decrease and $\rho_\perp$ increase as the temperature grows. Because of the disturbance of the orientation relation in composition 6, $\rho_\perp$ and $\rho_\parallel$ approach each other, and identical clusters participate in conductivity in both cases. We can therefore write the following inequalities:

$$v_1 < v_2 < \ldots < v_6,$$
$$\rho_{\perp 1}/\rho_{\parallel 1} > \rho_{\perp 2}/\rho_{\parallel 2} > \ldots > \rho_{\perp 6}/\rho_{\parallel 6}. \quad (5)$$

A change in the superconducting character of the $\rho_\perp(v)$ curves at $I \perp X$, a decrease in their absolute value at $T > T_c$ (the classical limit), steep drops in $\rho_\perp(v)$ resistances and their transitions to the superconducting state at $T < T_c$ (Fig. 4) below the traditional percolation threshold $\eta_c \approx 0.16$, and arbitrary percolation threshold values ($0 < \eta_p < \eta_c$) prove the inexhaustible possibilities of interphase boundaries for the example of semiconductor–superconductor eutectic compositions.

Note one more interesting feature of classical and quantum percolation conductivity. The normal $I_N$ and superconducting $I_S$ electric currents in the eutectic

compositions under consideration in both classical ($T > T_c$) and quantum ($T < T_c$) limits are the sums

$$I_N = I_{N\text{per}} + I_{N\text{Sm}}, \quad (6)$$

$$I_S = I_{S\text{per}} + I_{N\text{Sm}}, \quad (7)$$

where $I_{N\text{per}}$ is the normal percolation current component, $I_{S\text{per}}$ is the superconducting percolation current component, and $I_{N\text{Sm}}$ is the normal nonpercolation current component.

For $I \parallel X$, we have $I_{N\text{per}} \gg I_{N\text{Sm}}$ and $I_{S\text{per}} \gg I_{N\text{Sm}}$. The normal and superconducting currents are then largely determined by the percolation components; that is, $I_N \approx I_{N\text{per}}$ and $I_S \approx I_{S\text{per}}$.

For $I \perp X$ and low growth rates, $I_{N\text{Sm}} \gg I_{N\text{per}}$ and $I_{N\text{Sm}} \gg I_{S\text{per}}$, and normal $I_{N\text{per}}$ and superconducting $I_{S\text{per}}$ percolation currents only exist in discrete N–Sm–N–Sm–N… and S–Sm–S–Sm–S… cascades. The contribution of $I_{N\text{per}}$ and $I_{S\text{per}}$ begins to grow as the rate of composition growth increases. The normal $I_{N\text{per}}$ and superconducting $I_{S\text{per}}$ percolation currents exist in both finite and infinite composition clusters and are then determined by (6) and (7).

Note that, at low growth rates, the number of parallel infinite clusters per unit area of the cross section of a composition is approximately equal to the number of infinite whiskers. The supercurrent through an area of 1 mm$^2$ is then

$$I = i_1 + i_2 + \ldots + i_n. \quad (8)$$

If the statistical mean diameters of whiskers are assumed to be equal at equal growth rates, we have

$$i_1 = i_2 = \ldots = i_n, \quad (9)$$

$$I = ni. \quad (10)$$

We used (10) to approximately estimate the critical current density per whisker at 2 K. It follows from the table that a change in the critical current density by three orders of magnitude corresponds to a 100-fold change in growth rate $v$ as it decreases from 70 to 0.85 cm/h. Such a change in the critical current density cannot be explained by a tenfold increase in the diameter of whiskers. An increase in the diameter of whiskers by one order of magnitude in infinite crystals comprising S–S'–S–S'–S… chains is accompanied by an increase in the thickness of many Josephson microbridges. Many of these microbridges cease to exist. The other microbridges transform into bridges corresponding to stronger bonds. For this reason, the critical current density increases as the growth rate of compositions decreases.

Note that whiskers are arranged more regularly in the eutectic compositions grown at rates of 0.85–22 cm/h. At a 70 cm/h growth rate, strong irregularities are observed. At 200 cm/h, periodicity completely disappears. For this reason, the $\rho$ value near the critical temperature is two orders of magnitude larger at $v$ = 200 cm/h than at $v$ = 70 cm/h.

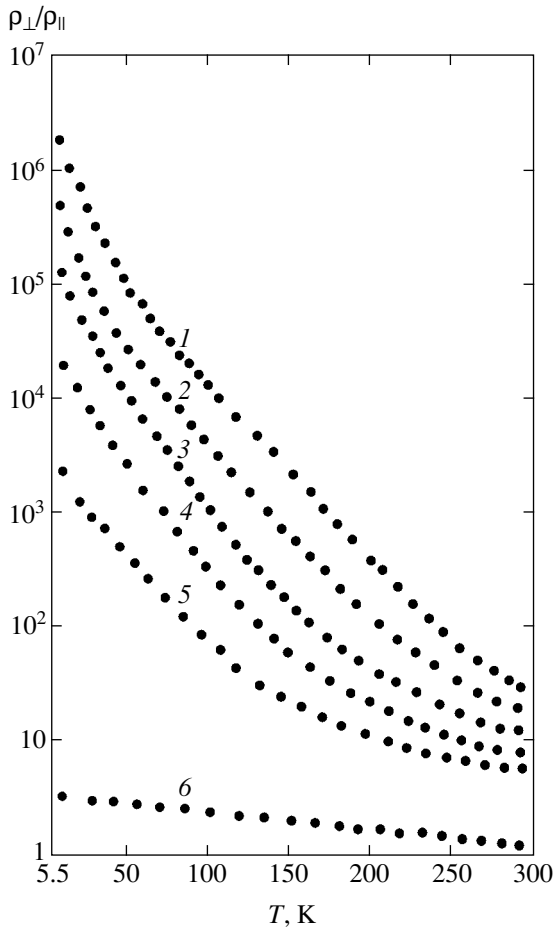



**Fig. 8.** Temperature dependences of anisotropy coefficients $\rho_\parallel/\rho_\perp$ of eutectic compositions grown at different rates $v$.

## 4. CONCLUSIONS

(1) We found that, depending on the rate of growing semiconductor (GaSb)–superconductor ($V_2Ga_5$) eutectic compositions ($v$ = 0.85, 2.8, 7, 22, 70, and 200 cm/h), the superconducting $V_2Ga_5$ phase is formed in the semiconducting GaSb matrix in the form of oriented whiskers at all growth rates except $v$ = 200 cm/h. The density of whiskers, the distances between them, and their lengths and diameters can be controlled by varying the rate of growing compositions.

(2) At $T < T_c$, $I \parallel X$, and low growth rates, supercurrent is largely carried by infinite clusters comprising S–S'–S–S'–S… Josephson chains. Part of these clusters are replaced by S–S'–S–Sm–S–Sm–S–S'–S… infinite clusters in compositions grown at higher rates. At $T < T_c$, $I \perp X$, and low growth rates, changes in the shape of $\rho_\perp$ curves are caused by the appearance of discrete cascades of S–Sm–S–Sm–S… Josephson junctions. These

cascades transform into finite clusters as $v$ increases. At large $v$, the superconducting transition is caused by the joining up of finite clusters into infinite S–Sm–S–Sm–S… clusters.

(3) At $T > T_c$, $I \parallel X$, and low $v$, normal electric current is largely carried by infinite N–N'–N–N'–N… clusters. Part of these clusters are replaced by infinite N–N'–N–Sm–N–N'–N… clusters as $v$ increases. At $T > T_c$ and $I \perp X$, normal electric current is largely carried over all infinite N–Sm–N–Sm–N… clusters that fill the whole composition.

(4) The critical supercurrent density per infinite cluster can be controlled by varying the rate of growing the composition. The diameter $d$ of superconducting whiskers decreases as the growth rate increases; this breaks numerous Josephson microbridges (S') in S–S'–S–S'–S… and S–S'–S–Sm–S–Sm–S–S'–S… clusters. In both clusters, broken microbridges are replaced by Sm semiconducting interlayers corresponding to weaker bonds than those formed by S' microbridges.

(5) We considered possible mechanisms of two- and one-particle currents over S–Sm–S–Sm–S… and N–Sm–N–Sm–N… clusters. The passage of supercurrent and normal current in semiconductor–superconductor compositions was shown to mainly occur as resonance tunneling of Cooper pairs and single electrons, respectively, along resonance-percolation trajectories at arbitrary percolation threshold values ($0 < \eta_p \le \eta_c$).

(6) The results of this work can be used in modeling and developing a theory of classical and quantum phenomena in similar low-temperature and high-temperature superconductors and Josephson structures on their base.

## ACKNOWLEDGMENTS

The author thanks F.M. Gashimzade for valuable advice and discussions and M.A. Tikhonovskiĭ for the electron microscopic images.

## REFERENCES

1. M. I. Aliev, G. I. Isakov, É. A. Isaeva, *et al.*, Fiz. Tekh. Poluprovodn. (St. Petersburg) **30**, 1871 (1996) [Semiconductors **30**, 978 (1996)].
2. G. I. Isakov, Pis'ma Zh. Tekh. Fiz. **22** (12), 70 (1996) [Tech. Phys. Lett. **22**, 1032 (1996)].
3. G. I. Isakov, Izv. Ross. Akad. Nauk, Neorg. Mater. **39**, 677 (2003).
4. M. I. Aliev and G. I. Isakov, Izv. Akad. Nauk SSSR, Neorg. Mater. **16**, 782 (1980).
5. M. I. Aliev, G. I. Isakov, and A. T. Éminzade, Izv. Akad. Nauk SSSR, Neorg. Mater. **21**, 1890 (1985).
6. G. I. Isakov, M. I. Aliev, and A. T. Éminzade, Izv. Akad. Nauk SSSR, Neorg. Mater. **24**, 682 (1988).
7. G. I. Isakov, M. I. Aliev, and A. T. Éminzade, Izv. Akad. Nauk SSSR, Neorg. Mater. **24**, 404 (1988).
8. G. I. Isakov, M. I. Aliev, and A. T. Éminzade, Izv. Akad. Nauk SSSR, Neorg. Mater. **25**, 926 (1989).
9. M. I. Aliyev, A. A. Khalilova, D. H. Arasly, *et al.*, J. Phys. D: Appl. Phys. **36**, 2626 (2003).
10. Yu. V. Kislinskiĭ, E. A. Stepantsov, Z. G. Ivanov, *et al.*, Fiz. Tverd. Tela (St. Petersburg) **43**, 581 (2001) [Phys. Solid State **43**, 602 (2001)].
11. J. A. Alarco and E. Olsson, Phys. Rev. B **52**, 13625 (1995).
12. D. Winkler, V. M. Zhang, P. A. Nilsson, *et al.*, Phys. Rev. Lett. **72**, 1260 (1994).
13. L. G. Aslamazov and M. V. Fustul', Zh. Éksp. Teor. Fiz. **81**, 382 (1981) [Sov. Phys. JETP **54**, 206 (1981)].
14. L. G. Aslamazov and M. V. Fustul', Zh. Éksp. Teor. Fiz. **83**, 1170 (1982) [Sov. Phys. JETP **56**, 666 (1982)].
15. L. G. Aslamazov and M. V. Fustul', Zh. Éksp. Teor. Fiz. **86**, 1516 (1984) [Sov. Phys. JETP **59**, 887 (1984)].
16. Th. Schapers, R. P. Muller, A. Kaluza, *et al.*, Appl. Phys. Lett. **75**, 391 (1999).
17. M. I. Petrov, D. A. Balaev, K. A. Shaĭkhutdinov, *et al.*, Fiz. Tverd. Tela (St. Petersburg) **39**, 829 (1997) [Phys. Solid State **39**, 735 (1997)].
18. M. I. Petrov, D. A. Balaev, K. A. Shaihutdinov, and K. S. Aleksandrov, Supercond. Sci. Technol. **14**, 798 (2001).
19. M. I. Aliev, G. I. Isakov, F. Yu. Aliev, *et al.*, Dokl. Akad. Nauk SSSR **306**, 583 (1989) [Sov. Phys. Dokl. **34**, 454 (1989)].
20. G. I. Isakov, Pis'ma Zh. Tekh. Fiz. **29** (10), 40 (2003) [Tech. Phys. Lett. **29**, 810 (2003)].
21. A. I. Somov and M. A. Tikhonovskiĭ, *Eutectic Compounds* (Metallurgiya, Moscow, 1975), p. 304 [in Russian].
22. R. Roberge and J. L. Fihey, J. Appl. Phys. **48**, 1327 (1977).
23. B. I. Shklovskiĭ and A. L. Éfros, Usp. Fiz. Nauk **117**, 401 (1975) [Sov. Phys. Usp. **18**, 845 (1975)].
24. I. M. Lifshits and V. Ya. Kirpichenkov, Zh. Éksp. Teor. Fiz. **77**, 989 (1979) [Sov. Phys. JETP **50**, 499 (1979)].

*Translated by V. Sipachev*

*Journal of Experimental and Theoretical Physics, Vol. 99, No. 4, 2004, pp. 801–810.*
*Translated from Zhurnal Éksperimental'noĭ i Teoreticheskoĭ Fiziki, Vol. 126, No. 4, 2004, pp. 926–935.*


**SOLIDS**
**Electronic Properties**

# Partial Moments of Electric Field Strength in the Problem of Conductivity of Binary Composites: Numerical Experiment on a Plane Disordered Lattice

**B. Ya. Balagurov and V. A. Kashin**

*Emanuel Institute of Biochemical Physics, Russian Academy of Sciences, Moscow, 119991 Russia*
*e-mail: balagurov@deom.chph.ras.ru*

Received February 5, 2004

**Abstract**—Partial (calculated by the volume of separate components) moments of electric field strength of different orders are calculated within a linear problem on the electric conductivity of binary composites. A numerical experiment is carried out on a square lattice with a size of $401 \times 401$ sites with randomly distributed bonds. The moments (of orders $n = 2, 3, \ldots, 6$) are calculated as a function of concentration for a number of fixed values of the parameter $h$—the ratio of the conductivities of the components. The behavior of partial moments near the metal–dielectric phase-transition point is studied, and the corresponding critical indices are determined. © *2004 MAIK "Nauka/Interperiodica".*

## 1. INTRODUCTION

In a number of studies [1–5], it has been pointed out that various effective characteristics of two-component media (in particular, binary composites) can be expressed in terms of partial (calculated by the volumes of separate components) moments of electric field strength of different orders. For example, the effective conductivity $\sigma_e$ can be expressed in terms of moments of both the first and second orders. The knowledge of the second-order moments allows one to determine the derivatives of $\sigma_e$ with respect to its arguments—the conductivities of the components [2]. The simultaneous determination of the effective conductivity and its derivatives allows one to analyze in sufficient detail the critical behavior of $\sigma_e$ in the neighborhood of a metal–dielectric phase-transition point (see [6, 7]). The derivatives of $\sigma_e$ also enter expressions for the low-frequency permittivity, the magnetoresistance in a weak magnetic field [2], and, for a certain relation between parameters, the formula for the thermal emf. Structural fluctuations of the field and current [1, 2], as well as the Joule heat produced in each component, are expressed in terms of second-order moments.

When studying nonlinear phenomena, one faces the problem of determining higher order moments. According to [3, 4], the calculation of the first nonlinear correction to the effective conductivity requires the knowledge of the partial moments of the fourth order. Further approximations in nonlinearity involve the moments of the sixth, eighth, etc., orders (see [5]). The fourth-order moments arise in the problem on the spectrum of low-frequency noises [3, 4]. Finally, note that, sometimes, one has to investigate partial moments of odd orders.

Thus, the partial moments of electric field strength play an important role in the theory of transport phenomena in binary composites. These quantities, determined in a linear problem of electric conductivity, represent functions of two arguments, the concentration $p$ and the ratio of the conductivities of the components, $h = \sigma_2/\sigma_1$. Thus, the initially multiparameter effective characteristics of a binary medium, being expressed in terms of appropriate moments, reduce to two-parameter functions. For systems with a metal–dielectric phase transition, the critical behavior of the two-parameter functions can be described within the standard similarity hypothesis, like the electric conductivity [8] and some other parameters [2]. For partial moments, this procedure is described in [5], where appropriate critical indices are introduced and a relation between them is determined. It is shown in [5] that each partial moment of order greater than two is characterized by a new (in addition to the indices of effective conductivity) critical index. According to [5], the number of critical indices in the two-dimensional case is half of that in the three-dimensional case.

In the present paper, we investigate a number of partial moments of electric field strength for a two-dimensional two-component medium within a linear problem of conductivity. We carry out a numerical experiment on a square lattice with a size of $401 \times 401$ sites with randomly distributed bonds over the whole range of concentrations $p$ for six values of the parameter $h$: $h = 10^{-m}$, where $m = 1, \ldots, 6$. For a specific "realization" (i.e., for a given concentration of bonds), we apply numerical methods to solve an appropriate system of Kirchhoff equations and determine potentials $V_{\mathbf{r}}$ at all sites of the lattice. Simultaneously, we solve a similar

problem on determining the potentials $\tilde{V}_{\mathbf{r}}$ at the sites of the dual lattice.

Using $V_{\mathbf{r}}$ and $\tilde{V}_{\mathbf{r}}$, we calculate lattice analogs of the partial moments of electric field strength. These quantities are determined as averages over six realizations. In the critical domain, which is the most difficult case for the analysis and where the results significantly fluctuate when passing from one realization to another, the number of realizations amounts to ten. For each moment, we investigate its behavior in the neighborhood of a metal–dielectric phase-transition point and estimate both the critical indices and the coefficients of appropriate expansions.

## 2. EFFECTIVE CHARACTERISTICS OF A MEDIUM

The problem on the conductivity of an inhomogeneous medium is posed as follows. Consider the following system of equations for a direct current:

$$\mathrm{curl}\mathbf{E} = 0, \quad \mathrm{div}\mathbf{j} = 0, \tag{1}$$

where $\mathbf{E} = \mathbf{E}(\mathbf{r})$ is the electric field strength and $\mathbf{j} = \mathbf{j}(\mathbf{r})$ is the current density. In the linear (with respect to the field) statement of the problem, $\mathbf{j}$ and $\mathbf{E}$ are related by Ohm's law (isotropic medium):

$$\mathbf{j} = \sigma(\mathbf{r})\mathbf{E}, \tag{2}$$

where $\sigma = \sigma(\mathbf{r})$ is a local conductivity of a sample that depends on coordinates. The system of equations (1), (2) is solved under the condition that there is a homogeneous field of strength $\langle\mathbf{E}\rangle$ in the medium, where $\langle\ldots\rangle$ stands for averaging over the sample volume $V$.

The effective conductivity $\sigma_e$ of the medium is defined as a coefficient of proportionality between the average current density $\langle\mathbf{j}\rangle$ and the field $\langle\mathbf{E}\rangle$:

$$\langle\mathbf{j}\rangle = \sigma_e\langle\mathbf{E}\rangle \tag{3}$$

under the condition that $V \longrightarrow \infty$. In the case of two-component systems, it is convenient to introduce a dimensionless effective conductivity $f$:

$$\sigma_e = \sigma_e(p; \sigma_1, \sigma_2) \equiv \sigma_1 f(p, h), \quad h = \sigma_2/\sigma_1, \tag{4}$$

where $p$ is the concentration (volume fraction) of the first component.

Let us define partial moments of the electric field strength of order $2n$ as follows:

$$\psi_i^{(2n)} = \langle\mathbf{e}^{2n}\rangle^{(i)}, \quad n = 1, 2, \ldots, \tag{5}$$

where

$$\mathbf{e}(\mathbf{r}) = \frac{\mathbf{E}(\mathbf{r})}{|\langle\mathbf{E}\rangle|}, \tag{6}$$

$$\langle(\ldots)\rangle^{(i)} = \frac{1}{V}\int_{V_i}(\ldots)dV. \tag{7}$$

The integration in (7) is performed over the volume $V_i$ of the $i$th component. The functions $\psi_i^{(2n)}$, just as the conductivity $\sigma_e$ in (3), are effective (self-averaging as $V \longrightarrow \infty$) characteristics of the medium and depend neither on the magnitude nor (in the isotropic case) on the direction of the applied field $\langle\mathbf{E}\rangle$. The quantities $\psi_i^{(2n)}$ are determined solely by the properties of the medium and, for two-component systems, depend on the same arguments as the function $f$ from (4):

$$\psi_i^{(2n)} = \psi_i^{(2n)}(p, h). \tag{8}$$

According to [1], the effective conductivity $\sigma_e$ can also be expressed in terms of the quadratic characteristics of the field:

$$\sigma_e = \langle\sigma\mathbf{e}^2\rangle \tag{9}$$

with $\mathbf{e} = \mathbf{e}(\mathbf{r})$ from (6). For binary media, taking into account (4), we obtain the following relation from (9):

$$f = \psi_1^{(2)} + h\psi_2^{(2)}, \tag{10}$$

where $\psi_i^{(2)}$ are second-order partial moments. Quadratic structural fluctuations of the field and current are also expressed in terms of the moments $\psi_i^{(2)}$ (see [5]).

As is shown in [3–5], when studying the nonlinear properties of inhomogeneous media, one faces the problem of calculating higher order moments of the electric field strength. For a weakly nonlinear isotropic medium, we have

$$\mathbf{j} = \{\sigma(\mathbf{r}) + \chi^{(3)}(\mathbf{r})\mathbf{E}^2 + \chi^{(5)}(\mathbf{r})\mathbf{E}^4 + \ldots\}\mathbf{E} \tag{11}$$

instead of (2). The mean values $\langle\mathbf{j}\rangle$ and $\langle\mathbf{E}\rangle$ are related by a similar formula:

$$\langle\mathbf{j}\rangle = \{\sigma_e + \chi_e^{(3)}(\langle\mathbf{E}\rangle)^2 + \chi_e^5(\langle\mathbf{E}\rangle)^4 + \ldots\}\langle\mathbf{E}\rangle, \tag{12}$$

where $\chi_e^{(3)}$, $\chi_e^{(5)}$, … are effective nonlinearity factors. According to [3, 4], the factor $\chi_e^{(3)}$ can be expressed in

terms of the electric field strength from the linear problem:

$$\chi_e^{(3)} = \langle \chi^{(3)} \mathbf{e}^4 \rangle, \quad (13)$$

where $\mathbf{e} = \mathbf{e}(\mathbf{r})$ is the same as in (6). For a two-component medium, from (13) we obtain

$$\chi_e^{(3)} = \chi_1^{(3)} \psi_1^{(4)} + \chi_2^{(3)} \psi_2^{(4)}. \quad (14)$$

Here, $\chi_i^{(3)}$ is the nonlinearity factor in the $i$th component and $\psi_i^{(4)}$ are the fourth-order partial moments. According to [5], expressions for $\chi_e^{(5)}$, $\chi_e^{(7)}$, … contain moments of the sixth, eighth, … orders, respectively. Note that the quantities $\psi_i^{(4)}$ also arise in the problem on the spectrum of low-frequency noises [3, 4].

The partial moments $\psi_i^{(2n+1)}$ of odd orders are determined by the formula

$$\langle \mathbf{e}^{2n+1} \rangle^{(i)} = \psi_i^{(2n+1)} \langle \mathbf{e} \rangle. \quad (15)$$

This formula takes into account that, for an isotropic medium, the vector quantity $\langle \mathbf{e}^{2n} \mathbf{e} \rangle^{(i)}$ can only be directed along the unit vector $\langle \mathbf{e} \rangle = \langle \mathbf{E} \rangle / |\langle \mathbf{E} \rangle|$. It follows from (15) that

$$\psi_i^{(2n+1)} = \langle \mathbf{e}^{2n} e_\parallel \rangle^{(i)}, \quad (16)$$

where $e_\parallel(\mathbf{r})$ is the component of $\mathbf{e}(\mathbf{r})$ that is parallel to $\langle \mathbf{E} \rangle$. The odd moments also depend on $p$ and $h$. Note that, according to (3),

$$f = \psi_1^{(1)} + h \psi_2^{(1)}. \quad (17)$$

For a binary randomly inhomogeneous system, simultaneous substitutions $\sigma_1 \rightleftharpoons \sigma_2$ and $p \longrightarrow 1 - p$ do not change the macroscopic properties of the medium (see [1]), so that [5]

$$f(p, h) = h f\left(1 - p, \frac{1}{h}\right), \quad (18)$$

$$\psi_1^{(n)}(p, h) = \psi_2^{(n)}\left(1 - p, \frac{1}{h}\right), \quad (19)$$

$$\psi_2^{(n)}(p, h) = \psi_1^{(n)}\left(1 - p, \frac{1}{h}\right). \quad (20)$$

Equalities (18)–(20) allow one to determine the functions $f(p, h)$ and $\psi_i^{(n)}(p, h)$ for $h > 1$, provided that they are known for $h < 1$ in the whole range of concentrations $p$.

In the two-dimensional case, the functions $f$ and $\psi_i^{(n)}$ satisfy the so-called reciprocity relations [1, 5]:

$$f(p, h) f\left(p, \frac{1}{h}\right) = 1, \quad (21)$$

$$\psi_1^{(n)}\left(p, \frac{1}{h}\right) = \frac{1}{[f(p, h)]^n} \psi_1^{(n)}(p, h), \quad (22)$$

$$\psi_2^{(n)}\left(p, \frac{1}{h}\right) = \left[\frac{h}{f(p, h)}\right]^n \psi_2^{(n)}(p, h). \quad (23)$$

Formulas (21)–(23) are valid for isotropic two-dimensional two-component systems of arbitrary structure, both for periodic and disordered ones.

For a randomly inhomogeneous medium, taking into account (18)–(20), we can rewrite equalities (21)–(23) as

$$f(p, h) f(1 - p, h) = h, \quad (24)$$

$$\psi_1^{(n)}(1 - p, h) = \left[\frac{h}{f(p, h)}\right]^n \psi_2^{(n)}(p, h), \quad (25)$$

$$\psi_2^{(n)}(1 - p, h) = \frac{1}{[f(p, h)]^n} \psi_1^{(n)}(p, h). \quad (26)$$

For $p = 1/2$, formula (24) implies the well-known result $f(1/2, h) = \sqrt{h}$ [1]. Formulas (25) and (26) allow one to relate the function $\psi_1^{(n)}$ for $p > p_c$ to $\psi_2^{(n)}$ for $p < p_c$ (and vice versa), where $p_c = 1/2$ is a critical concentration.

For a binary isotropic system, the moments of the first and second order can be expressed in terms of the function $f$ and its derivative [2]:

$$\psi_1^{(1)} = \frac{f - h}{1 - h}, \quad \psi_2^{(1)} = \frac{1 - f}{1 - h}, \quad (27)$$

$$\psi_1^{(2)} = f - h f', \quad \psi_2^{(2)} = f', \quad f' \equiv \frac{\partial f(p, h)}{\partial h}. \quad (28)$$

One can easily verify that the functions $\psi_i^{(1)}$ and $\psi_i^{(2)}$ identically satisfy all the relations given above.

## 3. ORGANIZATION OF A NUMERICAL EXPERIMENT

The partial moments of the electric field strength for a two-dimensional two-component system were calculated within a standard linear problem on the conductivity of a plane disordered lattice [9] (see also [6]). In the present paper, we carry out a numerical experiment on a square lattice with a size of $N \times N = 401 \times 401$ sites over the whole range of concentrations for six values of the parameter $h$: $h = 10^{-m}$, $m = 1, \ldots, 6$. Each lattice site $\mathbf{r} = (k, j)$, where $k$ is the number of a row and $j$ is the number of a column ($k = 1, \ldots, N$; $j = 1, \ldots, N$), is assigned a potential $V_\mathbf{r} = V_{k,j}$. For all lattice sites,

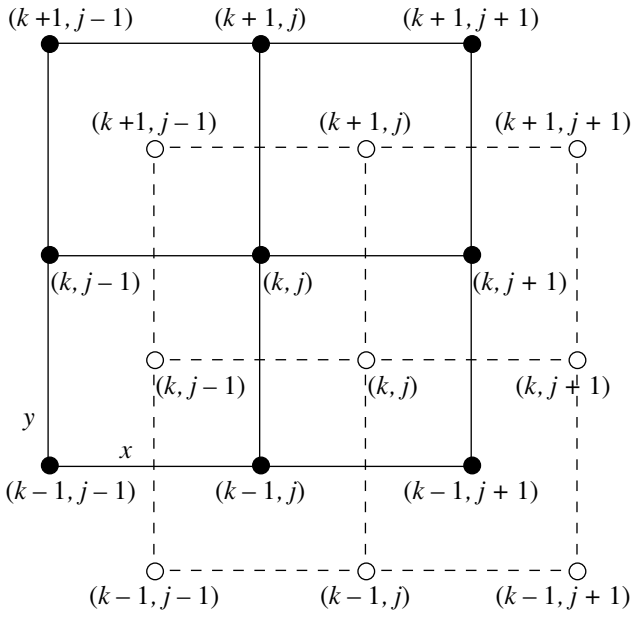



**Fig. 1.** (Solid lines) basic and (dashed lines) dual lattices.

except for certain boundary sites, these potentials satisfy the following system of Kirchhoff equations:

$$\sum_{\Delta} \sigma_{\mathbf{r}, \mathbf{r}+\Delta}(V_{\mathbf{r}} - V_{\mathbf{r}+\Delta}) = 0, \tag{29}$$

where the summation is over the four vectors $\Delta = \pm\Delta_x = \pm(1, 0)$ and $\Delta = \pm\Delta_y = \pm(0, 1)$. In the problem of bonds considered, $\sigma_{\mathbf{r}, \mathbf{r}+\Delta}$ in formula (29) is the conductivity of the bond between the sites $\mathbf{r}$ and $\mathbf{r} + \Delta$, which takes the value $\sigma_1 = 1$ with probability $p$ ("pure" bonds) and $\sigma_2 = h$ with probability $1 - p$ ("defective" bonds). On two opposite sides of the lattice, that are perpendicular to the $x$ axis, the potentials should satisfy the boundary values equal to 0 and 1, respectively. Along the $y$ axis, the potentials should satisfy periodic boundary conditions, so that the sites with $k = 1$ and $k = N$ are identified.

The numerical experiment is carried out as follows. For a certain fixed concentration $p$, a realization is chosen such that the $(1 - p)$th part of bonds in the initially pure lattice are randomly replaced by defective bonds. Then, under the boundary conditions formulated above, the system of equations (29) is solved with prescribed values of $\sigma_{\mathbf{r}, \mathbf{r}+\Delta}$, and the potentials $V_{\mathbf{r}}$ are determined at all lattice sites. Simultaneously, a similar problem is solved to determine the potentials $\tilde{V}_{\mathbf{r}}$ at the sites of the so-called dual lattice (see Fig. 1) under the same boundary conditions.

It is usually assumed that, in the problem of bonds on a square lattice with randomly distributed defects, the macroscopic properties (conductivity and other effective characteristics) of the basic and dual lattices coincide. The calculation performed in the present paper confirms this assumption as regards the effective conductivity to within the computation accuracy. However, the primary goal of considering the problem of the conductivity of the dual lattice is the application of the potentials $\tilde{V}_{\mathbf{r}}$ (together with $V_{\mathbf{r}}$) to calculating the discrete analogs of partial moments of electric field strength.

In the lattice problem, one has to take the difference of potentials at adjacent (along the $x$ axis) sites as $E_x(\mathbf{r})$:

$$E_x(\mathbf{r}) \longrightarrow V_{k,j} - V_{k,j+1}. \tag{30}$$

For the $y$ component of the electric field strength, we take $\tilde{V}_{k,j} - \tilde{V}_{k+1,j}$, so that

$$E^2(r) \longrightarrow (V_{k,j} - V_{k,j+1})^2 + (\tilde{V}_{k,j} - \tilde{V}_{k+1,j})^2. \tag{31}$$

Formula (31) applies to a horizontal bond $(k, j; k, j + 1)$ of the basic lattice. For a vertical bond $(k, j; k + 1, j)$ of this lattice, we have

$$E^2(r) \longrightarrow (\tilde{V}_{k+1,j-1} - \tilde{V}_{k+1,j})^2 + (\tilde{V}_{k,j} - \tilde{V}_{k+1,j})^2. \tag{32}$$

Thus, we have the following expression for the function $\psi_i^{(2n)}$ in the discrete problem:

$$\psi_i^{(2n)} = \frac{1}{2}L^{2n-2} \times \sum_{\{i\}} \{[(V_{k,j} - V_{k,j+1})^2 + (\tilde{V}_{k,j} - \tilde{V}_{k+1,j})^2]^n + [(\tilde{V}_{k+1,j-1} - \tilde{V}_{k+1,j})^2 + (V_{k,j} - V_{k+1,j})^2]^n\}, \tag{33}$$

where the summation is over all bonds with conductivities $\sigma_i$ and $L = N - 1$. Formula (33) takes into account that $\langle E_x \rangle = 1/L$ for a given difference of potentials $U = 1$. The summation in (33) is performed over the doubled number of bonds (on the basic and dual lattices), which is compensated for by the factor 1/2. For an odd-order partial moment $\psi_i^{(2n+1)}$, we obtain

$$\psi_i^{(2n+1)} = \frac{1}{2}L^{2n-1} \times \sum_{\{i\}} \{[(V_{k,j} - V_{k,j+1})^2 + (\tilde{V}_{k,j} - \tilde{V}_{k+1,j})^2]^n \times (V_{k,j} - V_{k,j+1}) + [(\tilde{V}_{k+1,j-1} - \tilde{V}_{k+1,j})^2 + (V_{k,j} - V_{k+1,j})^2]^n \times (\tilde{V}_{k+1,j-1} - \tilde{V}_{k+1,j})\}. \tag{34}$$

In the present paper, we apply formulas (33) and (34) to calculate the partial moments $\psi_1^{(n)}$ and $\psi_2^{(n)}$ for $n = 2$, 3, …, 6 (see Tables 1 and 2). The results, averaged over six realizations, are presented in Figs. 2–9. Figure 10

shows the logarithm of the dimensionless effective conductivity as a function of concentration.

## 4. CRITICAL BEHAVIOR OF FUNCTIONS $\psi_i^{(n)}(p, h)$

The critical behavior of the effective conductivity of a randomly inhomogeneous two-component medium in the neighborhood of a metal–dielectric phase transition point is described within the similarity hypothesis [8]. According to [8] (see also, for example, [2]), in the critical domain $h \ll 1$, $|\tau| \ll 1$, where $\tau = (p - p_c)/p_c$, $p_c$ is a critical concentration, the function $f$ behaves as follows:

$$f = \tau^t \left\{ A_0 + A_1 \frac{h}{\tau^{t/s}} + \ldots \right\} \tag{35}$$

for $\tau > 0$, $\Delta_0 \ll \tau \ll 1$,

$$f = h^s \left\{ a_0 + a_1 \frac{\tau}{h^{s/t}} + \ldots \right\} \tag{36}$$

for $|\tau| \ll \Delta_0$, and

$$f = \frac{h}{(-\tau)^q} \left\{ B_1 + B_2 \frac{h}{(-\tau)^{t/s}} + \ldots \right\} \tag{37}$$

for $\tau < 0$, $\Delta_0 \ll |\tau| \ll 1$. Here,

$$\Delta_0 = h^{s/t} \tag{38}$$

is the size of the smearing region [8]. The critical indices $t$, $s$, and $q$ are related by the formula [8]

$$q = \frac{t}{s} - t. \tag{39}$$

In the two-dimensional case considered here, we

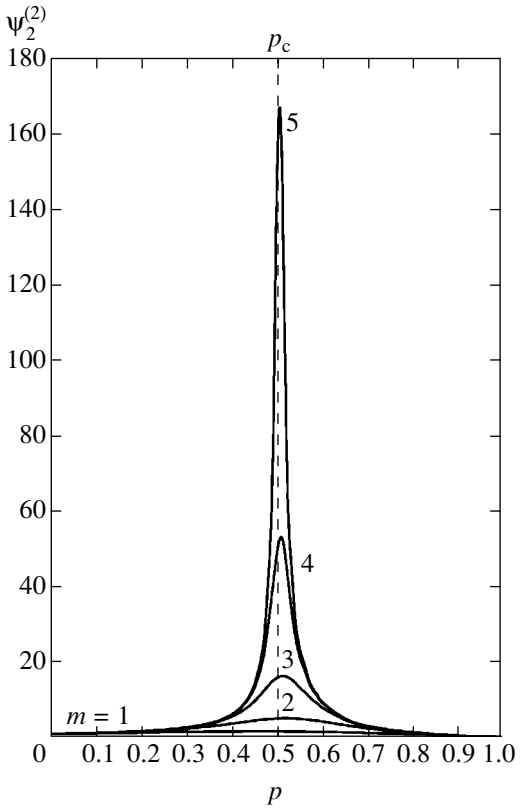



**Fig. 2.** The function $\psi_2^{(2)}$ versus $p$ for $h = 10^{-m}$, $m = 1, 2, \ldots, 5$.

**Table 1.** Critical indices of the partial moments $\psi_1^{(n)}$ and $\psi_2^{(n)}$

| $n$ | $t_n$ | $s_n$ | $q_n$ | $\mu_n$ | $\lambda_n$ |
|---|---|---|---|---|---|
| 2 | 1.30 ± 0.01 | 0.500 ± 0.004 | 3.90 ± 0.08 | 1.30 ± 0.05 | 0.500 ± 0.001 |
| 3 | 1.67 ± 0.03 | 0.666 ± 0.005 | 6.2 ± 0.2 | 2.13 ± 0.03 | 0.823 ± 0.002 |
| 4 | 1.42 ± 0.06 | 0.641 ± 0.006 | 8.8 ± 0.2 | 3.56 ± 0.06 | 1.349 ± 0.004 |
| 5 | 1.64 ± 0.05 | 0.734 ± 0.008 | 11.1 ± 0.3 | 4.55 ± 0.08 | 1.749 ± 0.006 |
| 6 | 1.5 ± 0.2 | 0.68 ± 0.01 | 13.6 ± 0.4 | 6.0 ± 0.1 | 2.307 ± 0.009 |

**Table 2.** Numerical coefficients for $\psi_1^{(n)}$ and $\psi_2^{(n)}$

| $n$ | $A_1^{(n)}$ | $a_1^{(n)}$ | $B_1^{(n)}$ | $A_2^{(n)}$ | $a_2^{(n)}$ | $B_2^{(n)}$ |
|---|---|---|---|---|---|---|
| 2 | 1.4 ± 0.6 | 0.500 ± 0.005 | 0.6 ± 0.1 | 1.1 ± 0.1 | 0.500 ± 0.001 | 0.7 ± 0.1 |
| 3 | 1.4 ± 0.1 | 0.450 ± 0.006 | 1.0 ± 0.3 | 2.0 ± 0.3 | 0.460 ± 0.003 | 0.8 ± 0.1 |
| 4 | 1.6 ± 0.3 | 0.340 ± 0.005 | 3.4 ± 1.6 | 4.8 ± 1.1 | 0.340 ± 0.003 | 0.7 ± 0.1 |
| 5 | 2.2 ± 0.2 | 0.300 ± 0.007 | 11.4 ± 7.5 | 13.3 ± 4.5 | 0.310 ± 0.005 | 0.9 ± 0.1 |
| 6 | 3.7 ± 0.8 | 0.230 ± 0.006 | $(0.8 \pm 0.5) \times 10^2$ | $(0.5 \pm 0.2) \times 10^2$ | 0.240 ± 0.005 | 1.1 ± 0.2 |

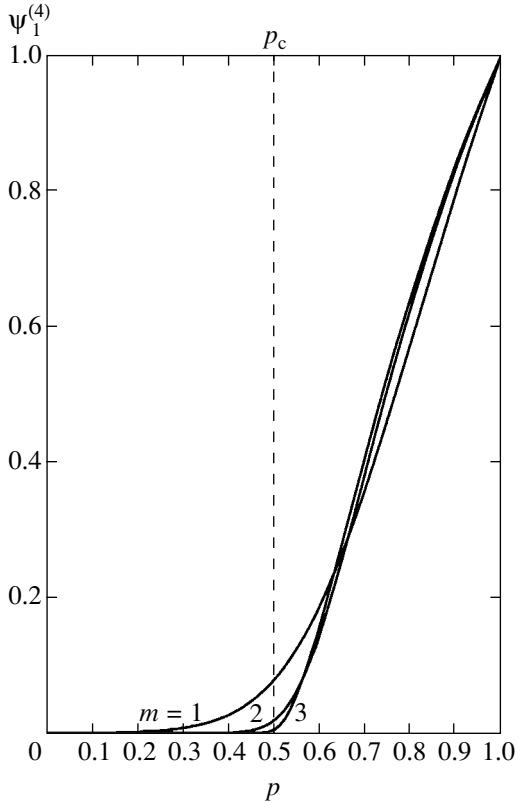



**Fig. 3.** The function $\psi_1^{(4)}$ versus $p$ for $h = 10^{-m}$, $m = 1, 2, 3$.

have [1, 8]

$$s = 1/2, \quad q = t. \tag{40}$$

It is natural to expect that the representations of the similarity hypothesis, including formulas (35)–(40), are also applicable to finely dispersed binary composites.

The critical behavior of the functions $\psi_i^{(2)}$ is determined by substituting formulas (35)–(37) into relations (28). In this case, all critical indices of the functions $\psi_i^{(2)}$ are expressed in terms of the conductivity indices (see (53)). At the same time, we do not know any relations of the type (28) for the functions $\psi_i^{(n)}$ with $n > 2$. Therefore, the behavior of the functions $\psi_i^{(n)}(p, h)$ with $n > 2$ in the neighborhood of a metal–insulator phase-transition point should be determined from their general properties. A similar procedure was performed, for example, in [2, 6, 7] for two-parameter functions that arise in the problem on the galvanomagnetic properties of binary systems in a weak magnetic field.

Restricting ourselves to the leading terms of appropriate expansions, we obtain the following expressions

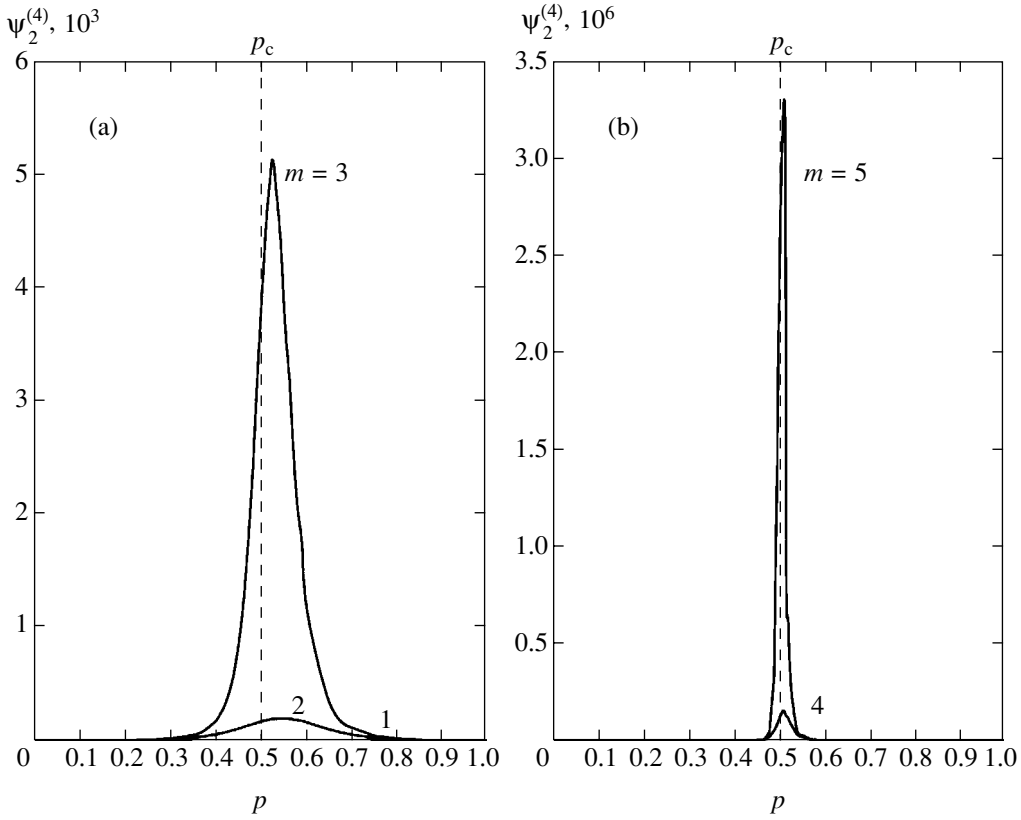



**Fig. 4.** The function $\psi_2^{(4)}$ versus $p$ for $h = 10^{-m}$, (a) $m = 1, 2, 3$ and (b) 4, 5.

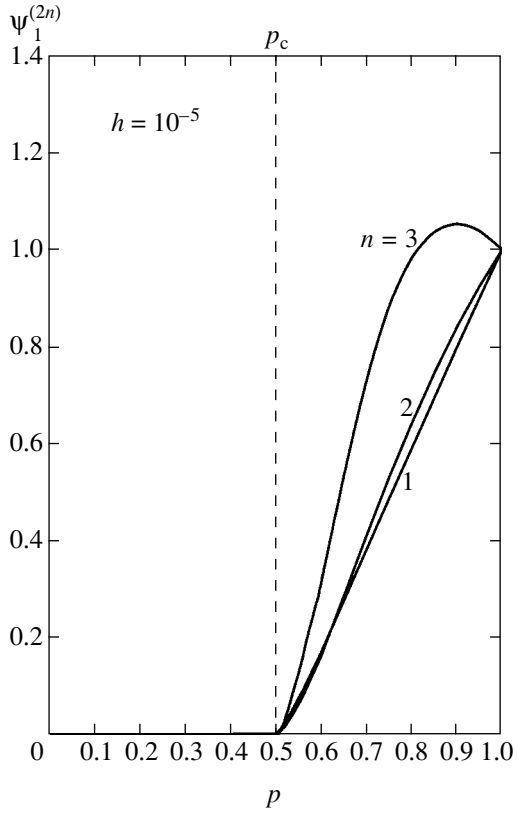



**Fig. 5.** Even partial moments $\psi_1^{(2n)}$, $n = 1, 2, 3$, as a function of $p$ for $h = 10^{-5}$.

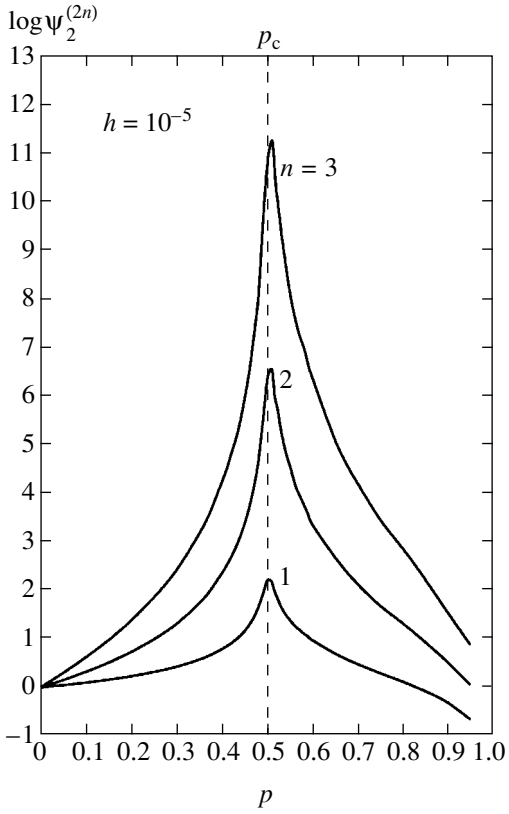



**Fig. 6.** Common logarithm of the partial moment $\psi_2^{(2n)}$, $n = 1, 2, 3$, as a function of $p$ for $h = 10^{-5}$.

for the functions $\psi_i^{(n)}(p, h)$ with $n \geq 2$ in the critical domain [5]:

$$\psi_1^{(n)} \approx A_1^{(n)}\tau^{t_n} \text{ and } \psi_2^{(n)} \approx A_2^{(n)}\tau^{-\mu_n} \tag{41}$$

for $\tau > 0$, $\Delta_0 \ll \tau \ll 1$,

$$\psi_1^{(n)} \approx a_1^{(n)}h^{s_n} \text{ and } \psi_2^{(n)} \approx a_2^{(n)}h^{-\lambda_n} \tag{42}$$

for $|\tau| \ll \Delta_0$, and

$$\psi_1^{(n)} \approx B_1^{(n)}\frac{h^n}{(-\tau)^{q_n}} \text{ and } \psi_2^{(n)} \approx B_2^{(n)}\frac{h^n}{(-\tau)^{\mu_n}} \tag{43}$$

for $\tau < 0$, $\Delta_0 \ll |\tau| \ll 1$. Here,

$$q_n = n\frac{t}{s} - t_n, \quad s_n = \frac{s}{t}t_n, \quad \lambda_n = \frac{s}{t}\mu_n. \tag{44}$$

In the two-dimensional case considered here, taking into account (40), we obtain

$$q_n = 2nt - t_n, \quad s_n = \frac{1}{2t}t_n, \quad \lambda_n = \frac{1}{2t}\mu_n. \tag{45}$$

The arguments that allow one to derive formulas (41)–(45) are presented in [5].

The numerical experiment meets certain difficulties in the neighborhood of a metal–insulator phase-transition points: at $h = 10^{-5}$, and especially at $h = 10^{-6}$, the results start to strongly fluctuate from one realization to another. Therefore, the number of realizations in the corresponding range of concentrations has increased up to ten. The results of the numerical experiment for the functions $\psi_1^{(n)}$ and $\psi_2^{(n)}$ in the critical domain, processed by formulas (41)–(43), are presented in Tables 1 and 2.

For the two-dimensional randomly inhomogeneous systems considered here, the reciprocity relation (25) (or (26)) for $p = 1/2$ with regard to $f(1/2, h) = \sqrt{h}$ is rewritten as

$$\psi_1^{(n)}\left(\frac{1}{2}, h\right) = h^{n/2}\psi_2^{(n)}\left(\frac{1}{2}, h\right). \tag{46}$$

The substitution of expressions (42) into (46) yields

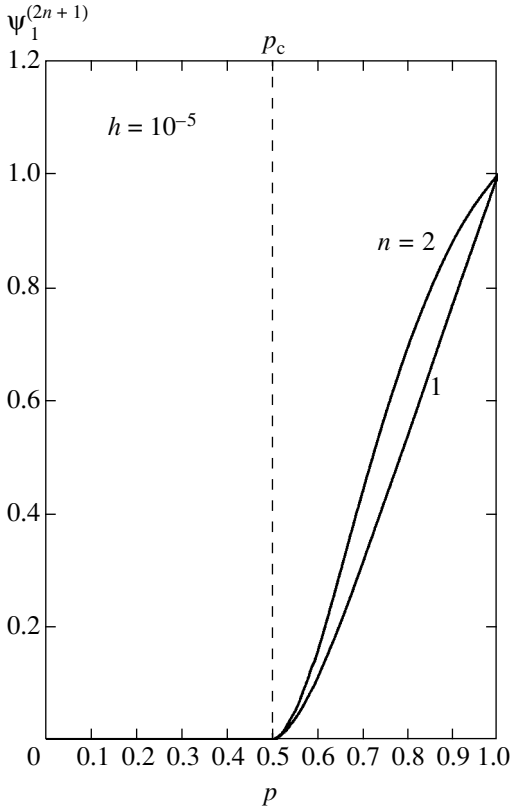



**Fig. 7.** Odd partial moments $\psi_1^{(2n+1)}$, $n = 1, 2$, as a function of $p$ for $h = 10^{-5}$.

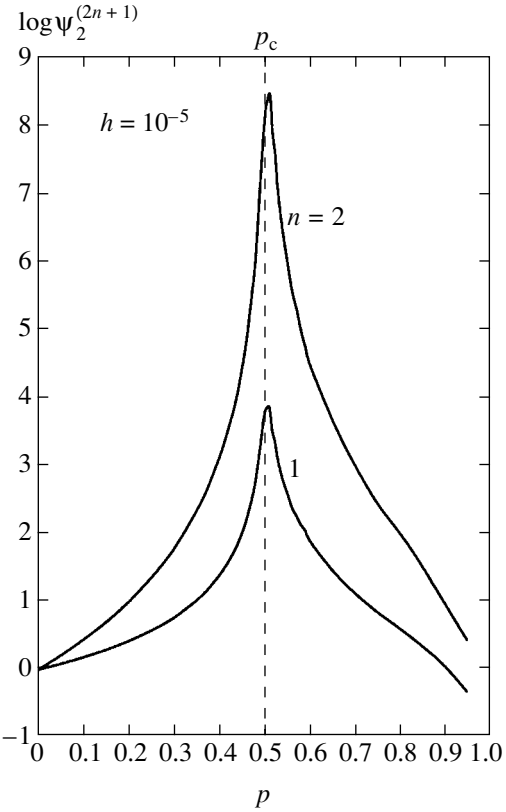



**Fig. 8.** Common logarithm of the partial moment $\psi_2^{(2n+1)}$, $n = 1, 2$, as a function of $p$ for $h = 10^{-5}$.

one more relation between critical indices:

$$s_n = \frac{n}{2} - \lambda_n, \tag{47}$$

where

$$a_1^{(n)} = a_2^{(n)}. \tag{48}$$

In view of relations (45) and (47), in the two-dimensional case, only one of five new critical indices (for fixed $n$) for the functions $\psi_1^{(n)}$ and $\psi_2^{(n)}$, for example, $t_n$, are independent; other critical indices are expressed in terms of $t_n$ as follows:

$$\begin{aligned} q_n &= 2nt - t_n, \quad s_n = \frac{1}{2t}t_n, \\ \lambda_n &= \frac{n}{2} - \frac{1}{2t}t_n, \quad \mu_n = nt - t_n. \end{aligned} \tag{49}$$

The reciprocity relations (25), (26) do not give any new relations between indices beyond the smearing region. However, the following relations between coefficients arise:

$$A_1^{(n)} = (A_0)^n B_2^{(n)}, \quad A_2^{(n)} = (A_0)^n B_1^{(n)}, \tag{50}$$

$$A_0 B_1 = 1, \quad A_0 B_2 + A_1 B_1 = 0. \tag{51}$$

The latter formula represents relations between $A_0$, $A_1$, $B_1$, and $B_2$ that follow from (24). According to (50),

$$\frac{A_1^{(n)}}{B_2^{(n)}} = \frac{A_2^{(n)}}{B_1^{(n)}} = (A_0)^n. \tag{52}$$

These equalities can be used both for verifying the correctness of computations and for the independent determination of the coefficient $A_0$.

For the second-order moments $\psi_i^{(2)}$, the substitution of (35)–(37) into (28) yields

$$\begin{aligned} t_2 &= t, \quad \mu_2 = q, \quad s_2 = s, \\ \lambda_2 &= 1 - s, \quad q_2 = \frac{t}{s} + q, \end{aligned} \tag{53}$$

$$\begin{aligned} A_1^{(2)} &= A_0, \quad A_2^{(2)} = A_1, \quad a_1^{(2)} = (1-s)a_0, \\ a_2^{(2)} &= sa_0, \quad B_1^{(2)} = -B_2, \quad B_2^{(2)} = B_1. \end{aligned} \tag{54}$$

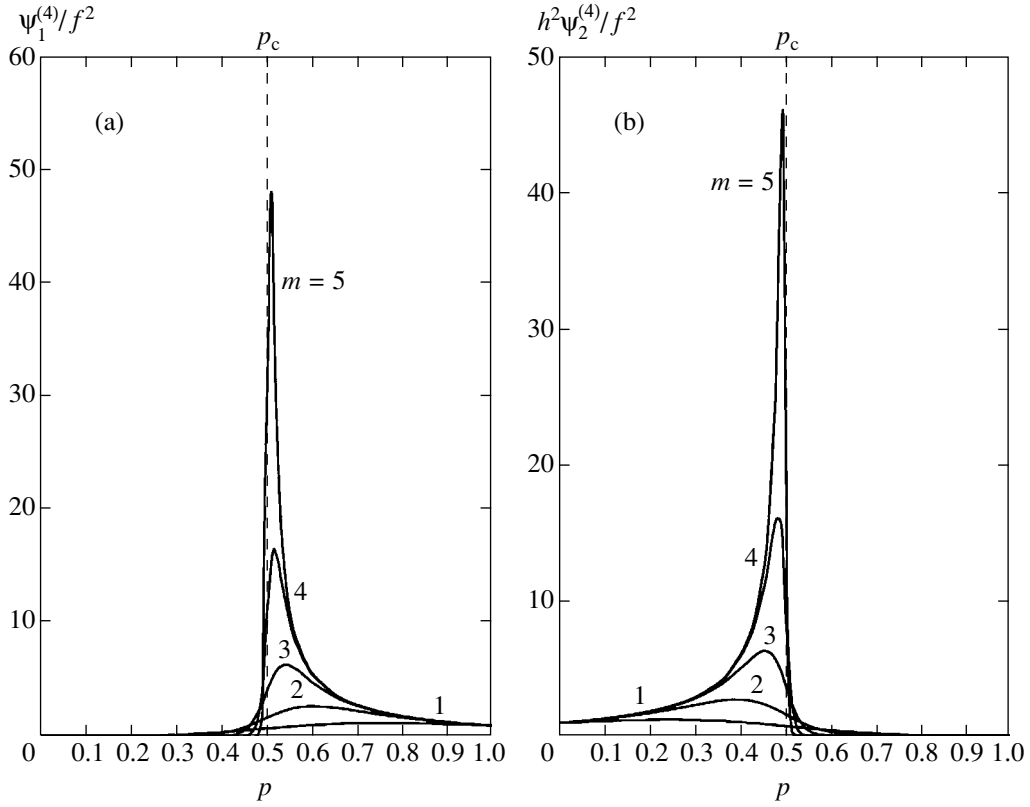



**Fig. 9.** The ratios (a) $\psi_1^{(4)}/f^2$ and (b) $h^2\psi_2^{(4)}/f^2$ as a function of $p$ for $h = 10^{-m}$, $m = 1, 2, \ldots, 5$.

The index $\lambda_2$ is positive in view of the inequality $s < 1$, and $B_1^{(2)} > 0$ since $B_2 < 0$ [2]. For a two-dimensional system ($s = 1/2$), we have

$$t_2 = \mu_2 = t, \quad s_2 = \lambda_2 = 1/2, \quad q_2 = 3t. \qquad (55)$$

Processing the results of the numerical experiment for the effective conductivity, we obtain

$$t = 1.30 \pm 0.06, \qquad (56)$$

$$A_0 = 1.4 \pm 0.6, \quad A_1 = 1.1 \pm 0.1,$$
$$B_1 = 0.7 \pm 0.1, \quad B_2 = -0.6 \pm 0.1. \qquad (57)$$

A comparison with the data of Tables 1 and 2 shows that formulas (55) and (54) (for $s = 1/2$) are valid to within the computation accuracy. Note that the value of $t$ given by (56) is in agreement with the results obtained in [3, 6].

In the case of the fourth-order moments, the critical indices are usually introduced for the function $\chi_e^{(3)}/\sigma_e^2$ (see [3, 4]). In the notation of the present paper, the following functions correspond to the relevant functions of [3, 4]:

$$\frac{\psi_1^{(4)}}{f^2} \approx \frac{1}{\tau^k} \qquad (58)$$

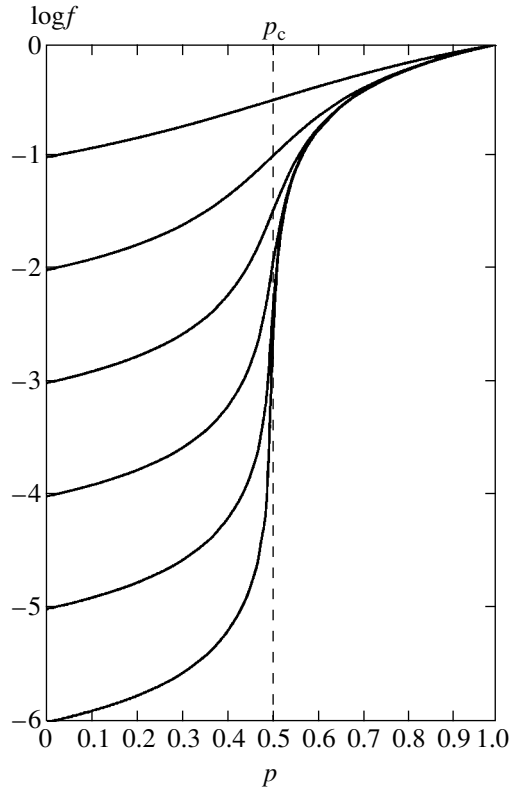



**Fig. 10.** Common logarithm of the dimensionless effective conductivity $f$ as a function of concentration $p$ for $h = 10^{-m}$, $m = 1, 2, \ldots, 6$.

for $\tau > 0$, $\Delta_0 \ll \tau \ll 1$ and

$$\frac{h^2\psi_2^{(4)}}{f^2} \approx \frac{1}{(-\tau)^{k'}} \quad (59)$$

for $\tau < 0$, $\Delta_0 \ll |\tau| \ll 1$. The substitution of expressions from (35), (41) and (37), (43) into (58) and (59) yields

$$k = 2t - t_4, \quad k' = \mu_4 - 2q. \quad (60)$$

For $n = 4$, the last relation in (49) can be rewritten as $2t - t_4 = \mu_4 - 2t$. In the two-dimensional case ($q = t$), comparing this equality with (60), we obtain

$$k = k'. \quad (61)$$

The calculation of $k$ and $k'$ by formulas (58) and (59) using the results of the numerical experiment yields

$$k = 1.02 \pm 0.02, \quad k' = 1.01 \pm 0.03, \quad (62)$$

so that equality (61) holds to within the computation accuracy.

One can easily verify that formulas (49) for the critical indices and formulas (48), (50), and (52) for the coefficients hold to the same accuracy for $n = 2, 3, \ldots$, 6. Note also that the value of the index $\lambda_4$ from Table 1 actually coincides with the value $\lambda_4 = 1.33 \pm 0.05$ obtained in [10].

## REFERENCES

1. A. M. Dykhne, Zh. Éksp. Teor. Fiz. **59**, 110 (1970) [Sov. Phys. JETP **32**, 63 (1970)].
2. B. Ya. Balagurov, Zh. Éksp. Teor. Fiz. **93**, 1888 (1987) [Sov. Phys. JETP **66**, 1079 (1987)].
3. D. Stroud and P. M. Hui, Phys. Rev. B **37**, 8719 (1988).
4. D. J. Bergman, Phys. Rev. B **39**, 4598 (1989).
5. B. Ya. Balagurov, Zh. Éksp. Teor. Fiz. **120**, 945 (2001) [JETP **93**, 824 (2001)].
6. B. Ya. Balagurov and V. A. Kashin, Zh. Éksp. Teor. Fiz. **106**, 811 (1994) [JETP **79**, 445 (1994)].
7. B. Ya. Balagurov and V. A. Kashin, Zh. Éksp. Teor. Fiz. **110**, 1001 (1996) [JETP **83**, 553 (1996)].
8. A. L. Efros and B. I. Shklovskii, Phys. Status Solidi B **76**, 475 (1976).
9. S. Kirkpatrick, Rev. Mod. Phys. **45**, 574 (1973).
10. A. M. Satanin, S. V. Khor'kov, and A. Yu. Ugol'nikov, Pis'ma Zh. Éksp. Teor. Fiz. **62**, 301 (1995) [JETP Lett. **62**, 322 (1995)].

*Translated by I. Nikitin*

*Journal of Experimental and Theoretical Physics, Vol. 99, No. 4, 2004, pp. 811–814.*
*Translated from Zhurnal Éksperimental'noĭ i Teoreticheskoĭ Fiziki, Vol. 126, No. 4, 2004, pp. 936–939.*


**SOLIDS**
**Electronic Properties**

# Light-Induced Elastic Waves: A Mechanism of the Optical Magnetic Transition in Manganese Arsenide

**G. A. Govor[a] and I. V. Semchenko[b]**

[a]*Institute of Solid State and Semiconductor Physics, National Academy of Sciences of Belarus, Minsk, 220726 Belarus*
*e-mail: govor@ifttp.bas-net.by*
[b]*Gomel State University, Gomel, 246699 Belarus*
Received September 26, 2003

**Abstract**—Dynamical characteristics of the optical magnetic transition in a manganese arsenide film have been studied. The results are considered with allowance for the possible generation of light-induced longitudinal elastic waves at the surface, which produce compression of the manganese arsenide film and cause the transition from the ferromagnetic to a disordered magnetic state. © *2004 MAIK "Nauka/Interperiodica".*

It was recently established [1, 2] that a magnetic transition in thin-film structures comprising a ferromagnetic layer of manganese arsenide and semiconductor layers of indium arsenide and gallium arsenide can be initiated by changing the density of carriers in the magnetic semiconductor layer. In connection with this, it was of interest to return to the previously discovered phenomenon called optical magnetic transition, whereby the magnetic state of manganese arsenide exhibited a change under the action of light pulses [3–5]. Several models were proposed for explaining the optical magnetic transition in manganese arsenide. In particular, Tralle [6] considered the possibility that the magnetic moment decreases due to a diamagnetic contribution to the magnetization.

This paper reports on the results of investigation of the dynamical characteristics of the optical magnetic transition in a manganese arsenide film, which passes from the ordered ferromagnetic to a disordered magnetic state under the action of light pulses. We have analyzed the results with allowance for the possible generation of longitudinal elastic waves at the film surface as a result of light absorption.

## EXPERIMENTAL

The films of manganese arsenide with equiatomic composition and a thickness from 0.5 to 2 μm were prepared by chemical vapor deposition onto an amorphous glass substrate. We obtained both textured films, having the predominant growth in the [110] direction and the magnetization vector oriented perpendicularly to the film plane, and isotropic films with predominant magnetization in the film plane.

The dynamical characteristics of the optical magnetic transition in a manganese arsenide film was studied using the experimental setup described elsewhere [3]. Light pulses were generated by a pulsed xenon lamp. In order to eliminate various interferences, the sample film was irradiated from the side of glass substrate by light pulses transmitted via an optical fiber.

The magnetic properties of a manganese arsenide film were measured using a high-frequency magnetometer. The measuring circuit was based on a thin-film permalloy core, with a sample film placed in the gap. The high-inductance magnetizing circuit and the low-inductance measuring circuit (detecting a response signal caused by a change in the magnetic state of the sample film) were spatially separated and oriented at an angle relative to each other. The shapes of the light pulse and the signal related to a change in the magnetic state of the sample film were measured by a high-frequency double-beam oscillograph.

## RESULTS AND DISCUSSION

Figure 1 shows the light pulse (curve *1*) and the emf pulse detected in the measuring coil in the presence of a constant magnetizing field (curves *2* and *3*). The magnetizing field strength was $H = 4$ kA/m. As can be seen, a transition from the ordered ferromagnetic to a disordered magnetic state (curve *2*) is related to the leading front of the light pulse, that is, to the increasing light flux. The relaxation process, whereby the film returns to the initial ferromagnetic state (curve *3*), occurs at a maximum of the light pulse, which is evidence of the nonthermal character of the optical magnetic transition in manganese arsenide.

Figure 2 presents the temperature dependence of the magnetic induction $B$ of a 1-μm-thick film of manganese arsenide and the change $\Delta B$ of this induction under the action of light pulses with a power density of 4 W/cm$^2$. As can be seen from these data, manganese

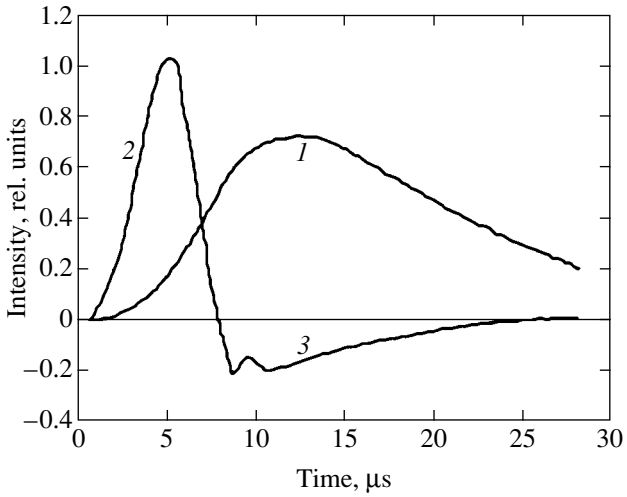



**Fig. 1.** Shapes of (*1*) the light pulse, (*2*) the emf pulse in the measuring coil of the magnetometer (corresponding to a transition of the manganese arsenide film from the ferromagnetic to a disordered magnetic state), and (*3*) the signal due to a relaxation transition to the ferromagnetic state in a constant magnetizing field of $H = 4$ kA/m.

arsenide in this film completely passes from the ferromagnetic to a disordered magnetic state at temperatures above 290 K.

The optical magnetic transition in a manganese arsenide film is accompanied by a change in the sample volume, $\Delta V/V_0 = -0.018$. As a result, there appear acoustic oscillations propagating from the light-activated surface inward the film. Using an additional layer of barium titanate deposited in vacuum onto the surface of the manganese arsenide film, it is possible to measure the piezoelectric effect and detect these acoustic oscillations in the sample film (Fig. 3, curve *2*). The initial region of this response displays a negative pulse reflecting compression of the film related to its transition from the ferromagnetic to a disordered magnetic state. Subsequently, the relaxation gives rise to the intrinsic acoustic oscillations at the fundamental frequency.

In addition to the piezoelectric effect, the barium titanate layer exhibits a pyroelectric effect related to heating of the film under the action of the light pulse. As can be seen in Fig. 3 (curve *2*), the maximum heating of the film does not exceed $\Delta T = 0.02$ K.

We have studied the influence of the rate of the light flux buildup on the optical magnetic transition in manganese arsenide. The rate of increase in the incident light energy was controlled by changing the duration of the leading pulse front from 10 to 100 μs at a constant light energy density in the pulse.

Figure 4 shows a plot of the relative volume of a disordered magnetic phase formed under the action of light pulses versus the rate of the light flux buildup. As can be seen, when this rate is below $\Delta E/\Delta t = 2 \times 10^5$ W/(cm$^2$ s), the optical magnetic transition in manganese arsenide virtually does not take place. Therefore, there exists a certain threshold value of the rate of the light energy buildup below which the effect under consideration is not observed. When the light energy buildup rate exceeds $\Delta E/\Delta t = 5 \times 10^5$ W/(cm$^2$ s) and the light power density in the pulse is on the order of 4 W/cm$^2$ (see Fig. 2), the film with a thickness within 0.5–2 μm completely passes from the ferromagnetic to a disordered magnetic state.

The nature and mechanism of structural-magnetic phase transitions in single crystal manganese arsenide were studied [6–11] taking into account changes in the lattice statics and dynamics, the magnetic and galvanomagnetic properties of crystals, and the effects of vari-

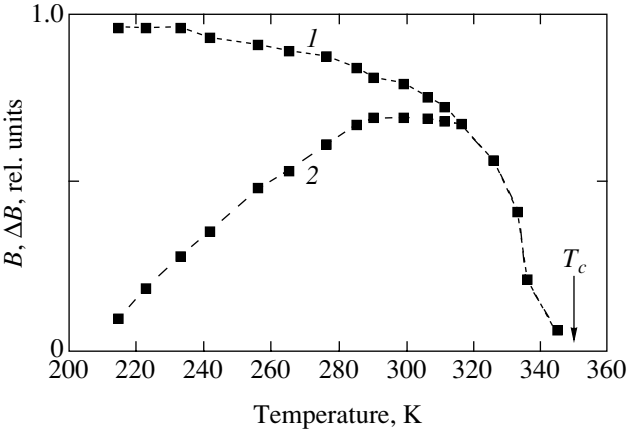



**Fig. 2.** Temperature dependence of (*1*) the magnetic induction $B$ of a 1-μm-thick film of manganese arsenide and (*2*) the change $\Delta B$ of this induction under the action of light pulses with a power density of 4 W/cm$^2$.

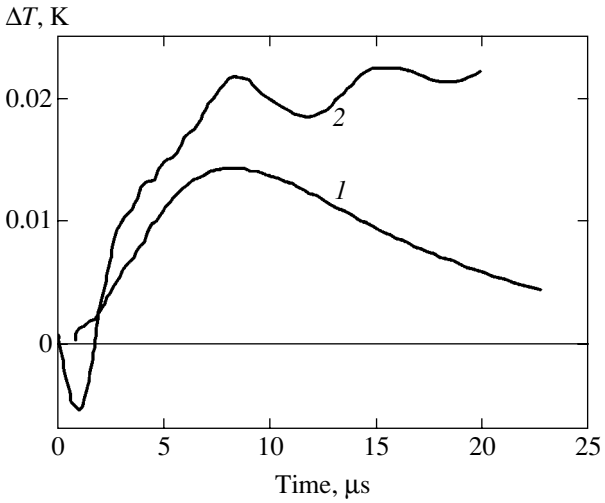



**Fig. 3.** Shapes of (*1*) the light pulse and (*2*) the emf pulse measured in a piezoelectric barium titanate layer deposited onto the manganese arsenide film.

ous external factors. It was established that, at the magnetic phase transition temperature $T_c$ = 318 K, manganese arsenide crystals exhibit a structural-magnetic phase transition from the ferromagnetic to a disordered magnetic state. In the latter state, the crystal is divided into ferromagnetically ordered microscopic blocks with the easy axis orientations in the neighboring blocks differing by 120°. Application of the magnetic field with a strength exceeding the crystallographic magnetic anisotropy leads to the appearance of a single-phase ferromagnetic state. This state is broken upon a transition to the paramagnetic state at $T_f$ = 400 K, which corresponds to a structural-magnetic phase transition of the second order. The ferromagnetic state in a manganese arsenide film breaks with the formation of a disordered magnetic phase at a temperature of $T_c$ = 350 K. The character of this transition (Fig. 2) is somewhat different from that observed in crystals.

The process of the optical magnetic transition in a manganese arsenide film passing from the ferromagnetic to a disordered magnetic state can be divided into two stages. In the first stage, optical activation of the phase transition takes place in a thin (50–100 Å thick) layer [1, 2]. In the second stage, a disordered magnetic state spreads over the entire film depth.

In order to explain the observed phenomenon, whereby a disordered magnetic state spreads in depth of the film, let us consider the possibility that longitudinal elastic oscillations are generated in the medium as a result of light absorption. The excitation of such oscillations leads to compression of the manganese arsenide film, thus decreasing the temperature of the transition to a disordered magnetic state.

The longitudinal elastic oscillations of the film propagate predominantly along the normal to the surface of the manganese arsenide film ($z$ axis) because this direction is parallel to the vector of group velocity of the light waves incident onto the sample structure. In this case, the strain tensor of the manganese arsenide film has the only nonzero component,

$$\gamma_{33} = -k_3 U_{03} \sin[k_3 z - \omega(t - t_1)], \qquad (1)$$

where $U_{03}$ is the amplitude of the longitudinal elastic oscillations, $k_3$ is the wavenumber, $\omega$ is the cyclic frequency of ultrasonic oscillations, and $t_1$ is the moment of acoustic wave excitation. Thus, we assume that relation (1) is valid for the time $t > t_1$.

For manganese arsenide with the crystallographic symmetry class *mmm*, the volume density of the elastic energy is

$$w = \frac{1}{2} C_{33} k_3^2 U_{03}^2 \sin^2[k_3 z - \omega(t - t_1)], \qquad (2)$$

where $C_{33}$ is the component of the tensor of elastic moduli (we use conventional notation for the pairs of symmetric indices [12]).

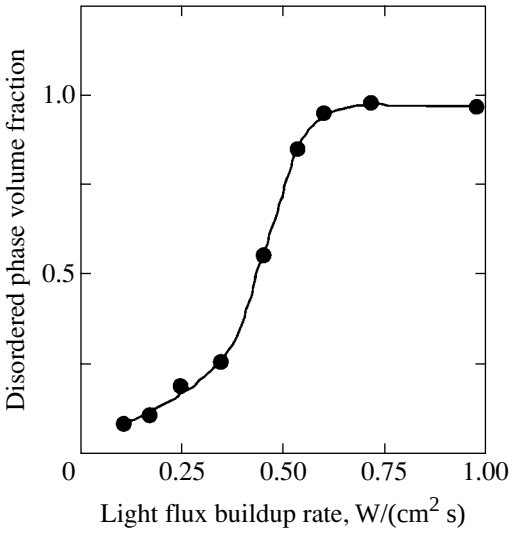



**Fig. 4.** Plot of the relative volume of a disordered magnetic phase formed due to the light-induced optical magnetic transition versus the rate of the energy flux buildup in the light pulse.

The acoustic pulse appears at the initial moment of the light flux buildup: $t_1 = 10^{-6}$ s (see Fig. 1). The energy balance equation for the system under consideration is

$$\beta E = \frac{1}{2} C_{33} k_3^2 U_{03}^2 \langle \sin^2[k_3 z - \omega(t - t_1)] \rangle_t v, \qquad (3)$$

where $v$ is the group velocity of elastic waves, $\beta$ is the coefficient of energy transfer from the light flux to the film, and angle brackets denote averaging with respect to time (which accounts for the factor 1/2). For the Young modulus $C = 2 \times 10^{11}$ N/m$^2$ and the intrinsic cyclic resonance frequency $\omega = 6.28 \times 10^5$ s$^{-1}$, we have $\beta = 0.46$.

At the initial time moment $t_1$, when the process of propagation of the longitudinal elastic oscillations has not been established, elastic displacement takes place predominantly on the front (exposed) surface of the manganese arsenide film. The resulting pressure exceeds 30 kbar, which is sufficient for activating a phase transition in manganese arsenide from the ferromagnetic to a disordered magnetic state [13, 14].

## REFERENCES

1. D. D. Awschalom and R. K. Kawakami, Nature **408**, 923 (2000).
2. D. K. Young, J. A. Gupta, E. Johnston-Galperin, *et al.*, Semicond. Sci. Technol. **17**, 275 (2002).
3. G. A. Govor and N. N. Sirota, Pis'ma Zh. Éksp. Teor. Fiz. **16**, 137 (1972) [JETP Lett. **16**, 95 (1972)].

4. G. A. Govor, Dokl. Akad. Nauk BSSR **23**, 515 (1979).
5. G. A. Govor, Dokl. Akad. Nauk BSSR **29**, 1098 (1985).
6. I. Tralle, Eur. Phys. J. B **22**, 3 (2001).
7. G. A. Govor, Fiz. Tverd. Tela (Leningrad) **23**, 1444 (1981) [Sov. Phys. Solid State **23**, 841 (1981)].
8. G. A. Govor, Fiz. Tverd. Tela (Leningrad) **28**, 38 (1986) [Sov. Phys. Solid State **28**, 18 (1986)].
9. G. A. Govor, K. Barner, and J.-W. Schunemann, Phys. Status Solidi A **113**, 403 (1989).
10. G. A. Govor, Fiz. Tverd. Tela (Leningrad) **33**, 3510 (1991) [Sov. Phys. Solid State **33**, 1971 (1991)].
11. J.-W. Shunemann, A. Lange, G. A. Govor, *et al.*, J. Alloys Compd. **178**, 237 (1992).
12. Yu. I. Sirotin and M. P. Shaskol'skaya, *Fundamentals of Crystal Physics* (Nauka, Moscow, 1979) [in Russian].
13. K. Barner, R. Steans, and R. Braunstein, Phys. Status Solidi B **95**, 483 (1979).
14. C. P. Bean and D. S. Rodbell, Phys. Rev. **126**, 104 (1962).

*Translated by P. Pozdeev*

*Journal of Experimental and Theoretical Physics, Vol. 99, No. 4, 2004, pp. 815–819.*
*Translated from Zhurnal Éksperimental'noĭ i Teoreticheskoĭ Fiziki, Vol. 126, No. 4, 2004, pp. 940–945.*


**SOLIDS**
**Electronic Properties**

# Nonlinear Resonances of Contact Photoionization in Heterogeneous Nanostructures

**B. P. Kashnikov, V. V. Makarov, E. V. Makarov, and G. I. Smirnov**

*Institute of Semiconductor Physics, Siberian Division, Russian Academy of Sciences, Novosibirsk, 630090 Russia*

*e-mail: smirnov.g.i.@iae.nsk.su*

Received March 2, 2004

**Abstract**—A theory of nonlinear interference effects is constructed for a heterogeneous charge transfer between atoms in polycrystalline films or heterogeneous nanostructures of the semiconductor-insulator type, which interact with resonance radiation, and a metallic contact surface. The probability of resonance contact photoionization in heterogeneous solid nanostructures is determined, which makes it possible to use this process in nanotechnologies and nonlinear information systems. Nonlinear resonances of contact photoionization are asymmetric due to interference of a radiative transition to an excited state and a transition to the continuum induced by the metal surface. The probability of resonance contact photoionization abruptly decreases with increasing distance between an atom in the semiconductor and the metal. © *2004 MAIK "Nauka/Interperiodica".*

## 1. INTRODUCTION

The interest in processes of heterogeneous charge transfer in the field of resonance radiation is stimulated by a variety of their possible applications in nanotechnologies, nonlinear information systems, charge coupling technologies, collective laser accelerators of ions, laser-plasma technologies, and optoelectronics [1–4]. Surface effects in photoabsorption and photodesorption spectra for particles interacting with the surface have been studied quite actively (see, for example, [5–7]). An interesting example of such effects is the emergence of asymmetry in the vibrational spectrum due to energy transfer between a molecule and a surface in accordance with the electron-hole pair mechanism [5]. The observation of resonance surface photoionization effects was reported for the first time in [8, 9]. Vast information on the electron–atom interaction problem, which is important in solid-state physics, optoelectronics, and a number of branches of atomic physics can be obtained from analysis of contact photoionization of atoms in polycrystalline films or solid heterostructures. In this study, we construct a theory of charge transfer between atoms in semiconducting crystallites or heterogeneous nanostructures of the semiconductor-insulator type, which resonantly interact with radiation, and the metallic surface of the contact. The probability of contact laser-induced photoionization of atoms localized in polycrystalline films or semiconductor-insulator nanostructures in the vicinity of a metal surface is determined.

## 2. FORMULATION OF THE PROBLEM

We assume that crystallites forming a photosensitive film possess the *n*-type conductivity and are surrounded by tunnel-transparent insulating oxide interlayers. If the excited state of an atom in an *n*-type semiconductor corresponds to the region of the allowed band of a metal above the Fermi level, the activationless charge exchange between this atom and the metal is the main photoionization mechanism. In [10], calculation of the probability of activationless charge exchange between an atom at the surface and the metal is reduced to the problem of ionization of the atom by an external electric field [11]. In the framework of the model of contact photoionization that will considered here, it will be shown that the state of an atom excited by a photon with an energy approximately equal to half the forbidden gap in a semiconductor located in the immediate vicinity of the metal surface is identified with an autoionization level against the background of the continuum formed by the quasi-continuous electron spectrum of normal metals [12]. It should be noted that resonance contact photoionization in heterogeneous nanostructures of the semiconductor-insulator type, which are in contact with the metal surface, also occurs in a similar way.

The resonance photoabsorption and surface photoionization spectra in this situation are determined by nonlinear interference effects of interaction of the given continuum and the energy states of crystallites, which correspond to the top of the valence band and the bottom of the conduction band. For photoabsorption in doped semiconductors, the role of the ground state or excited states can be played by the levels of donor impurity centers in the forbidden gap.

We will analyze the dependence of nonlinear interference effects of resonance surface photoionization on the distance between an atom in the semiconductor and

the metal. Analogous interference effects appear in the case of resonance scattering of laser radiation [13].

For definiteness, we assume that the temperature of the semiconductor is low enough for the energy states at the bottom of the conduction band and the top of the valence band to be approximately regarded as discrete. The quantum-mechanical description of contact photoionization process accompanied by absorption of radiation during the transition between the ground state $n$ of the atom and the excited state $m$ (which, in contrast to the ground state, lies above the Fermi level of the metal) is based on the total Hamiltonian

$$H = H_a + H_M + \hbar(U + V), \quad (1)$$

where the sum of three terms $H_a$, $H_M$, and $\hbar U$ is represented by the model Anderson Hamiltonian [14, 15], which defines the electron states of a system consisting of a metal and a two-level subsystem interacting with it. The Hamiltonian operators of the two-level subsystem,

$$H_a = \sum_{j = m, n} E_j, \quad (2)$$

and of the unperturbed metal,

$$H_M = \sum_p E_p c_p^\dagger c_p, \quad (3)$$

can be expressed in terms of the fermion operators of creation, $c_j^\dagger$ and $c_p^\dagger$, and annihilation, $c_j$ and $c_p$, of electrons in the atomic and metallic states; and $E_j$ and $E_p$ are the energies of the electron states in an atom and in the semi-infinite metal, respectively. Operator

$$U = \sum_{p, j} U_{pj} c_p^\dagger c_j + \text{h.c.} \quad (4)$$

describes the tunnel interaction between the metal state $p$ and the states $j = m, n$ of a crystallite. It is analogous to the Fano configurational interaction between the continuum and the autoionization state [16, 17]. It is assumed that electromagnetic radiation interacts only with atomic states; this interaction is taken into account by operator $V$.

## 3. QUANTUM KINETIC EQUATION

The general solution to the Schrödinger equation for the given problem,

$$i\hbar\partial_t \Psi = H\Psi, \quad (5)$$

can be represented by a linear combination of wavefunctions $\Psi_j$ and $\Psi_p$ of stationary states of an unperturbed system with Hamiltonian $H_0 = H_a + H_M$; taking into account the quasi-continuous nature of the electron spectrum of the metal, the summation over states differing in energy $E_p$ can be approximately replaced by integration:

$$\Psi = \sum_j a_j \Psi_j + \int d\omega_p a_p \Psi_p, \quad \omega_p = E_p/\hbar. \quad (6)$$

However, in calculating the spectral characteristics of photoionization in the field of a running wave, it is convenient to use the density matrix apparatus instead of the system of equations for the probability amplitudes of the states $a_{j, p}$ of the discrete and continuous spectra. Specific features of the procedure of transition from the given system of equations to the kinetic equation for the density matrix, which are determined by the existence of the continuum of electronic states, are described in [16]. The quantum kinetic equation suitable for any interaction of a particle with an external field was derived earlier in [18].

The nondiagonal density matrix element $\rho_{mn}$ corresponding to the Bohr frequency $\omega_{mn}$ of a radiative transition has the oscillating form in the resonance approximation,

$$\rho_{mn} = r_{mn}\exp[-i(\Omega t - \mathbf{k}\cdot\mathbf{r})], \quad \Omega = \omega - \omega_{mn}, \quad (7)$$

where $\omega = kc$ is the frequency of the light field. We assume that the interaction of an atom in the ground state lying below the Fermi level with the metal is noticeably weaker than that for an excited atom. The corresponding matrix elements of hybridization satisfy the inequality $|U_{pm}| > |U_{pn}|$.

Disregarding the effect of radiation on the population $\rho_{nn}$ of the lower atomic level, we can confine our analysis to the system of equations for $r_{mn}$ and the population $\rho_{mm}$ of the upper level:

$$(\partial_t + v_z\partial_z + \Gamma - i\Omega')r_{mn} = -i(G_{mn} + \delta_{mn} + i\gamma_{mn})\rho_{nn}, \quad \Omega' = \Omega - \delta_{mm} - \mathbf{k}\cdot\mathbf{v}, \quad (8)$$

$$(\partial_t + v_z\partial_z + \Gamma)\rho_{mm} = \text{Im}[(G_{mn} + \delta_{mn} + i\gamma_{mn})r_{nm}], \quad \Gamma = \Gamma_{mn} + \gamma_{mm}, \quad (9)$$

$$\rho_{nn} = N, \quad (10)$$

where $N$ is the total number of atoms per unit volume.

We denote by $\Gamma_{mn}$ the spontaneous decay constant for the excited state of the atom; $G_{mn} = E_0 d_{mn}/\hbar$, $E_0$ being the amplitude of the running wave and $d_{mn}$ the matrix element of the electric dipole transition moment. Parameters

$$\gamma_{mj} = \pi U_{mp} U_{pj}\big|_{\omega_p = \omega}, \quad (11)$$

$$\delta_{mj} = \sum_l \frac{2G_{ml}G_{lj}\omega_l}{\omega^2 - \omega_l^2} + \frac{1}{\pi}\int \frac{\gamma_{mj}(\omega_p)d\omega_p}{\omega - \omega_p} \quad (12)$$

take into account relaxation processes and energy shifts associated with the interference effect of tunnel interac-

tion of states $j = m, n$ with the metal surface. In expression (12), summation is extended to all nonresonance states and quantities $\gamma_{mj}$ and $\delta_{mj}$ depend on the distance $z$ between the atom and the contact surface.

## 4. NONLINEAR INTERFERENCE EFFECTS IN CONTACT IONIZATION

The linear susceptibility of the medium,

$$\chi = 2\hbar|E_0|^{-2}G_{nm}\langle r_{mn}\rangle \quad (13)$$

and the resonance light absorption coefficient

$$\alpha = 4\pi k\mathrm{Im}\chi \quad (14)$$

can be expressed in terms of the element $r_{mn}$ of the density matrix averaged over the ensemble of particles. To calculate the surface photoionization cross section

$$\sigma = 16\pi\hbar r|E_0|^{-2}N^{-1}\langle\gamma_{mm}\rho_{mm} + 2\mathrm{Re}(\gamma_{mn}r_{nm})\rangle \quad (15)$$

it is necessary to determine, along with $r_{mn}$, the population $\rho_{mm}$ of the excited state. The solutions to Eqs. (8) and (9) have the form

$$r_{mn} = i\rho_{nn}(G_{mn} + \delta_{mn} + i\gamma_{mn})[\Gamma - i(\Omega - \delta_{mm})]^{-1}, \quad (16)$$

$$\rho_{mm} = \rho_{nn}\gamma_{mn}^2(1 + q_{mn}^2)[\Gamma^2 + (\Omega - \delta_{mm})^2]^{-1},$$
$$q_{mn} = \frac{G_{mn} + \delta_{mn}}{\gamma_{mn}}. \quad (17)$$

Parameter $q_{mn}$ reflects the presence of two interfering radiation absorption channels such as the transition to excited state $m$ and the transition to a continuous-spectrum band of width $\gamma_{mm}$; the quantity $q_{mn}^2$ is proportional to the ratio of the probabilities of these two processes. In the limit $z \longrightarrow \infty$, when $\gamma_{mj}, \delta_{mj} = 0$, relations (16) and (17) can be reduced to the known results for particles that do not interact with the surface:

$$r_{mn} = i\rho_{nn}G_{mn}(\Gamma_{mn} - i\Omega)^{-1}, \quad (18)$$

$$\rho_{mm} = \rho_{nn}G_{mn}^2(\Gamma_{mn}^2 + \Omega^2)^{-1}. \quad (19)$$

Let us consider the resonant interaction of radiation with atoms locates at the same distance $z$ = const from the metal surface. This version is also realized when absorbing atoms are deposited over a thin insulating film coating the metal. In this case, in accordance with relations (14) and (16), we obtain the following relations for the contour of the frequency dependence of the resonance absorption coefficient:

$$\alpha = 8\pi\hbar kN\frac{G_{mn}\gamma_{mn}(q_{mn} - x)}{|E_0|^2\Gamma(1 + x^2)},$$
$$x = \frac{\Omega - \delta_{mm}}{\Gamma}. \quad (20)$$

The asymmetry of contour $\alpha(\Omega)$ is due to interference of processes of electron transition to discrete state $m$ and to the continuum formed by the metal surface.

In accordance with relations (15)–(17), the resonance surface photoionization spectrum is defined by the relation

$$\sigma = 16\pi\hbar|E_0|^{-2}\gamma_{mn}^2\Gamma^{-1}\frac{\theta(1 + q_{mn}^2) + 2 - 2q_{mn}x}{1 + x^2},$$
$$\theta = \gamma_{mn}\Gamma^{-1}. \quad (21)$$

The shape of surface photoionization resonance $\sigma(x)$ strongly depends on parameters $q_{mn}$ and $\theta$. In the whole, the frequency contour of the surface photoionization cross section $\sigma(x)$ resembles the Fano autoionization resonance profile [19, 20]. Depending on the distance between absorbing atoms and the metal surface, the value of parameter $\theta$ can change from 0 to 1 (in particular, we have $\theta \approx 1$ for $\Gamma_{mn} \ll \gamma_{mm}$).

As a rule, a decrease in the values of relaxation constants $\gamma_{mj}$ and energy shifts $\delta_{mj}$ with increasing distance between the atom and the metal surface is approximated by the exponential model [14, 20]

$$\gamma_{mj} = \tilde{\gamma}_{mj}e^{-az}, \quad \delta_{mj} = \tilde{\delta}_{mj}e^{-az}. \quad (22)$$

In this situation, the asymptotic behavior of the decrease in the surface photoionization cross section for $z \longrightarrow \infty$ is also represented by the exponential,

$$\sigma(z) \propto e^{-az}. \quad (23)$$

Using approximation (22), we can define the electron flux,

$$N_e = N\int_0^\infty W(z)dz = N\langle W\rangle, \quad (24)$$

induced by radiation in the vicinity of the contact surface. It is proportional to the averaged value of the probability of charge exchange between the metal and resonantly excited atoms in the semiconductor,

$$W(z) = \frac{|E_0|^2\sigma(z)}{8\pi\hbar k}, \quad (25)$$

where cross section $\sigma(z)$ is defined by formula (21). Taking into account the inequality

$$\gamma_{mj}, \delta_{mj} \gg \Gamma_{mn}, G_{mn}, \quad (26)$$

which normally holds in a surface photoionization layer of width $a^{-1}$, the asymmetric spectral distribution

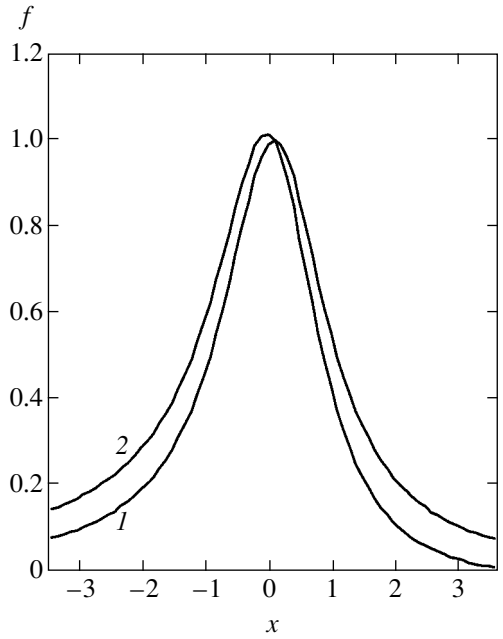



**Fig. 1.** Resonance photoabsorption spectrum: $q_{mn}^{-1} = 0$ (*1*) and $q_{mn} = 4$ (*2*).

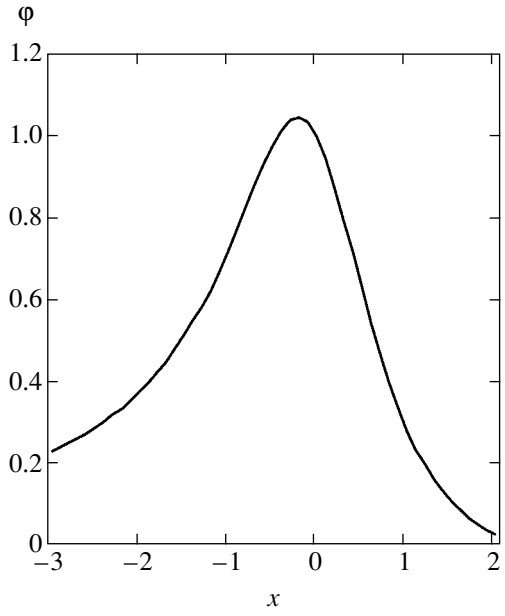



**Fig. 2.** Contour of the resonance contact photoionization cross section; $q_{mn} = 4$, $\theta = 0.8$.

of the mean probability $\langle W \rangle$ can be written in the form

$$\langle W \rangle = 2[(\tilde{\delta}_{mn}^2 - \tilde{\gamma}_{mn}^2) \times (\cos\beta - \tilde{x}\arctan\tilde{\varphi}\cos 2\beta + \tilde{x}\sin 2\beta \ln r) - \tilde{\gamma}_{mn}\tilde{\delta}_{mn}(\sin\beta - \tilde{x}\arctan\tilde{\varphi}\sin 2\beta - \tilde{x}\ln r\cos 2\beta)]/\gamma a, \quad (27)$$

$$\gamma = (\tilde{\gamma}_{mm}^2 + \tilde{\delta}_{mm}^2)^{1/2}, \quad \tilde{x} = \frac{\Omega}{\gamma},$$

$$\tilde{\varphi} = \tilde{\gamma}_{mm}(\Omega - \tilde{\delta}_{mm})^{-1},$$

$$r = [\tilde{\gamma}_{mm}^2 + (\Omega - \tilde{\delta}_{mm})^2]^{1/2}|\Omega|^{-1}, \quad (28)$$

$$\beta = \arctan\frac{\tilde{\delta}_{mm}}{\tilde{\gamma}_{mm}}.$$

The proposed model of resonance contact photoionization is directly realized, for example, when silicon is used as the semiconducting sample and the metallic contact is made of sodium. For the parameters of photoionization induced by high-power radiation emitted by semiconductor atoms at a distance $z < a^{-1} \sim 1$ nm from the metal surface, we have the relation $\tilde{\gamma}_{mn} \sim \tilde{\delta}_{mn} \lesssim \tilde{\gamma}_{mm}$, $\tilde{\delta}_{mm} \sim 10^{13}$ s$^{-1}$ [15, 21], and the resonance photoabsorption contour is represented in Fig. 1 by curve *2* for function

$$f(x) = \frac{1 - x/q_{mn}}{1 + x^2} \propto \alpha(x) \quad (29)$$

at $q_{mn} = 4$. Symmetric curve *1* in this case corresponds to the limiting case $q_{mn}^{-1} \longrightarrow 0$, when the width of the insulator layer considerably exceeds the characteristic width $a^{-1}$ of the photoionization layer. Figure 2 shows the plot of the function

$$\varphi(x) = \frac{1 - 2q_{mn}x[\theta(1 + q_{mn}^2) + 2]^{-1}}{(1 + x^2)} \propto \sigma(x), \quad (30)$$

which, accordingly, describes the profile of the resonance contact photoionization cross section for parameters $q_{mn} = 4$ and $\theta = 0.8$.

On the frequency scale, the photoionization resonance has a width on the order of $\tilde{\gamma}_{mm}$ and the degree of its asymmetry is determined by the ratio $\tilde{\delta}_{mn}/\tilde{\gamma}_{mn}$. Relations (25) and (27) lead to the following simple expression for electron flux $N_e$ corresponding to the center of the absorption line ($|\Omega| \ll \gamma$):

$$N_e = N[\tilde{\gamma}_{mm}(\tilde{\delta}_{mn}^2 - \tilde{\gamma}_{mn}^2) - 2\tilde{\gamma}_{mn}\tilde{\delta}_{mn}\tilde{\delta}_{mm}]/\gamma^2 a. \quad (31)$$

In particular, such resonance tunneling of photoelectrons from the contact layer in short-period superlattices might lead to the emergence of photocurrent resonances [22].

## 5. CONCLUSIONS

Thus, the main result of this study is the calculation of the probability of charge exchange between atoms in semiconducting crystallites or heterogeneous nanostructures of the semiconductor-insulator type, which resonantly interact with radiation, and the metal surface with allowance made for nonlinear interference effects associated with transitions to the excited state of the semiconductor and to the band in the continuous spectrum formed by the metal.

The spectral contours of photoabsorption and photoionization in semiconducting crystallites located in the vicinity of the metal surface are asymmetric like the Fano resonance profile [19, 20]. The degree of asymmetry of photoabsorption and photoionization resonances is determined by the ratio of the probability of a radiative transition to the excited state to the probability of transition to the continuum induced by the metal surface. Quite naturally, for atoms at large distances from the metal surface, the photoabsorption resonance becomes symmetric and the photoionization probability sharply decreases.

The properties of resonance contact photoionization considered here can be used for developing new technologies of data transformation and transmission of the type of photon–charge coupling.

## ACKNOWLEDGMENTS

We are grateful to E.A. Kuznetsov and G.G. Telegin for fruitful discussions.

This study was partly supported by the Program of Fundamental Studies (Presidium of the Russian Academy of Sciences) “Mathematical Methods in Nonlinear Dynamics” and the State Program Supporting Junior Russian Scientists and Leading Scientific Schools of the Russian Federation (grant no. NSh-1716.2003.1).

## REFERENCES

1. V. V. Antsiferov, G. I. Smirnov, and G. G. Telegin, Pis’ma Zh. Tekh. Fiz. **20** (8), 74 (1994) [Tech. Phys. Lett. **20**, 338 (1994)].
2. V. V. Antsiferov, G. I. Smirnov, and G. G. Telegin, Pis’ma Zh. Tekh. Fiz. **21** (14), 43 (1995) [Tech. Phys. Lett. **21**, 556 (1995)].
3. V. V. Antsiferov and G. I. Smirnov, *Coherent Radiation Processes in Plasma* (CISP, Cambridge, 1998), p. 236.
4. V. V. Antsiferov, B. P. Kashnikov, and G. I. Smirnov, Pis’ma Zh. Tekh. Fiz. **25** (3), 14 (1999) [Tech. Phys. Lett. **25**, 89 (1999)].
5. D. S. Langreth, Phys. Rev. Lett. **54**, 126 (1985).
6. D. Agassi, Phys. Rev. B **33**, 3873 (1986).
7. P. T. Leung and T. F. George, Chem. Phys. Lett. **134**, 375 (1987).
8. B. Auschwitz and K. Lacmann, Chem. Phys. Lett. **113**, 230 (1985).
9. G. F. Malyshev and G. G. Telegin, Zh. Tekh. Fiz. **56**, 1195 (1986) [Sov. Phys. Tech. Phys. **31**, 700 (1986)].
10. A. V. Chaplik, Zh. Éksp. Teor. Fiz. **54**, 332 (1968) [Sov. Phys. JETP **27**, 178 (1968)].
11. L. D. Landau and E. M. Lifshitz, *Course of Theoretical Physics*, Vol. 3: *Quantum Mechanics: Non-Relativistic Theory*, 3rd ed. (Nauka, Moscow, 1974; Pergamon, New York, 1977).
12. E. M. Lifshitz and L. P. Pitaevskiĭ, *Course of Theoretical Physics*, Vol. 5: *Statistical Physics* (Nauka, Moscow, 1978; Pergamon, New York, 1980).
13. S. G. Rautian, G. I. Smirnov, and A. M. Shalagin, *Nonlinear Resonances in Atomic and Molecular Spectra* (Nauka, Novosibirsk, 1979) [in Russian].
14. P. W. Anderson, Phys. Rev. **124**, 41 (1961).
15. L. A. Bol’shov, A. P. Napartovich, A. G. Naumovets, and A. G. Fedorus, Usp. Fiz. Nauk **122**, 125 (1977) [Sov. Phys. Usp. **20**, 432 (1977)].
16. Yu. I. Geller and A. K. Popov, *Laser Induction of Resonances in Continuum Spectra* (Nauka, Novosibirsk, 1981), p. 160 [in Russian].
17. V. V. Antsiferov, A. S. Vartazaryan, and G. I. Smirnov, Zh. Éksp. Teor. Fiz. **93**, 762 (1987) [Sov. Phys. JETP **66**, 430 (1987)].
18. S. G. Rautian and G. I. Smirnov, Zh. Éksp. Teor. Fiz. **74**, 1295 (1978) [Sov. Phys. JETP **47**, 678 (1978)].
19. V. Fano, Phys. Rev. **124**, 1866 (1961).
20. S. Ravi and G. S. Agarval, Phys. Rev. A **35**, 3354 (1987).
21. M. L. Yu and N. D. Lang, Phys. Rev. Lett. **50**, 127 (1983).
22. L. A. Al’perovich, N. T. Moshegov, A. S. Terekhov, *et al.*, Fiz. Tverd. Tela (St. Petersburg) **41**, 159 (1999) [Phys. Solid State **41**, 143 (1999)].

*Translated by N. Wadhwa*

*Journal of Experimental and Theoretical Physics, Vol. 99, No. 4, 2004, pp. 820–826.*
*Translated from Zhurnal Éksperimental'noĭ i Teoreticheskoĭ Fiziki, Vol. 126, No. 4, 2004, pp. 946–953.*


**SOLIDS**
**Electronic Properties**

# The Giant Volume Magnetostriction and Colossal Magnetoresistance in $Eu_{0.55}Sr_{0.45}MnO_3$ Manganite

**A. I. Abramovich, O. Yu. Gorbenko, A. R. Kaul', L. I. Koroleva*, and A. V. Michurin**

*Moscow State University, Moscow, 119992 Russia*
**e-mail: koroleva@ofef343.phys.msu.ru*

Received March 11, 2004

**Abstract**—A doped manganite with the composition $Eu_{0.55}Sr_{0.45}MnO_3$ exhibits giant negative magnetostriction and colossal negative magnetoresistance at temperatures in the vicinity of the magnetic phase transformation ($T \sim 41$ K). In the temperature interval 4.2 K $\le T \le$ 40 K, the isotherms of magnetization, volume magnetostriction, and resistivity exhibit jumps at the critical field strength $H_{c1}$, which decreases with increasing temperature. At 70 K $\le T \le$ 120 K, the jumps on the isotherms are retained, but the shapes of these curves change and the $H_{c1}$ value increases with the temperature. At $H < H_{c1}$, the magnetoresistance is positive and exhibits a maximum at 41 K; at $H > H_{c1}$, the magnetoresistance becomes negative, passes through a minimum near 41 K and then reaches a colossal value. The observed behavior is explained by the existence of three phases in $Eu_{0.55}Sr_{0.45}MnO_3$, including a ferromagnetic (in which the charge carriers concentrate due to a gain in the *s–d* exchange energy) and two antiferromagnetic phases (of the *A* and *CE* types). The volumes of these phases at low temperatures are evaluated. It is shown that the colossal magnetoresistance and the giant volume magnetostriction are related to the ferromagnetic phase formed as a result of the magnetic-field-induced transition of the *CE*-type antiferromagnetic phase to the ferromagnetic state. © *2004 MAIK "Nauka/Interperiodica".*

## 1. INTRODUCTION

Until recently, the interest in manganites was related primarily to the phenomenon of colossal magnetoresistance observed in some compounds at room temperature. However, our previous investigation [1–4], showed that the colossal magnetoresistance in $La_{1-x}Sr_xMnO_3$ ($0.1 \le x \le 0.3$), $Nd_{1-x}Sr_xMnO_3$ ($x = 0.33$ and 0.45), and $Sm_{1-x}Sr_xMnO_3$ ($x = 0.33$, 0.4, and 0.45) compounds is accompanied by a large negative volume magnetostriction ω. Moreover, the temperature and field dependences of the magnetostriction, ω(*T*) and ω(*H*), have proved to behave very much like the analogous curves of the magnetoresistance Δρ/ρ versus *T* and *H*. This behavior was explained by the coexistence of ferromagnetic (FM) and antiferromagnetic (AFM) phases in the above crystalline compounds, which was related to a strong *s–d* exchange interaction and the concentration of charge carriers (in this case, holes) in the FM part of the two-phase crystal [5]. It should be noted that the Curie temperature of a sample occurring in such a state is rather conditional and refers to the Curie temperature of the FM part of the system.

As is known, a compound of the $Sm_{1-x}Sr_xMnO_3$ system with $x = 0.5$ exhibits a charge–orbit ordering leading to the so-called *CE*-type AFM order. The results of the neutron diffraction and resistivity measurements showed that the compound $^{154}Sm_{0.6}Sr_{0.4}MnO_3$ contains the AFM clusters of *A* and *CE* types. The latter clusters, characterized by the charge–orbit ordering, occur in the conducting FM matrix [6]. In the composition with $x = 0.25$, the clusters with charge–orbit ordering are absent [6]. In the system with $x = 0.4$, an increase in the temperature leads to breakage, first, of the FM order at $T = T_C$, then of the *A*-rype AFM order at $T = T_N \ge T_C$, and eventually of the *CE*-type AFM order at $T = T_{co}$ ($T_N$ is the Néel temperature and $T_{co}$ is the temperature of breakage of the charge–orbit ordering). This behavior is consistent with the theoretical predictions of Dagotto *et al.* [7] based on the results of numerical modeling, according to which the concentration transition from the FM to charge–orbit ordering observed in manganites at $x = 0.5$ is a first-order phase transition and the systems with compositions close to $x = 0.5$ contain the magnetic clusters of three types: FM, *A*-type AFM, and charge–orbit ordered *CE*-type AFM.

Previously [8, 9], we have studied the magnetic, electrical, and galvanomagnetic properties of the compound $Eu_{0.7}Sr_{0.3}MnO_3$, which exhibits conductivity of the semiconductor type and is characterized by giant maxima in the temperature dependences of both the resistivity ρ and the magnetoresistance Δρ/ρ. The existence of the insulating two-phase magnetic state in this compound was confirmed by the following experimental results. The isotherms of magnetization σ(*H*) in the region of low temperatures represented the sum of a small spontaneous magnetization and a component linearly increasing with the field strength (which is characteristic of the AFM state). There was a difference in magnetizations of a sample cooled with and without

applied magnetic field, which was observed up to the maximum field strength studied (45 kOe), and the magnetic hysteresis loops of a sample cooled in the field were shifted along the $H$ axis. The contribution due to the FM clusters significantly increased the paramagnetic Curie temperature $\theta$ (from 100 K for the undoped manganite $EuMnO_3$ to 175 K for $Eu_{0.7}Sr_{0.3}MnO_3$). It was practically impossible to determine the Curie temperature of the FM part of this compound. Indeed, the $T_C$ value determined in a field of 45 kOe (by extrapolating the steepest part of the $\sigma(T)$ curve to intersection with the $T$ axis) amounted to 90 K, which was three times the value ($T_C$ = 30 K) obtained for $H$ = 0.5 kOe.

This work was aimed at a complex study of the magnetic, electrical, galvanomagnetic, and magnetoelastic properties of the compound $Eu_{0.55}Sr_{0.45}MnO_3$, which, according to the predictions of Dagotto *et al.* [7], must contain the magnetic clusters of three types: FM, *A*-type AFM, and *CE*-type AFM. It was of interest to reveal differences between the properties of this compound and those of $Sm_{1-x}Sr_xMnO_3$ (with $x$ = 0.33, 0.4, and 0.45) and $Eu_{0.7}Sr_{0.3}MnO_3$. The magnetic and galvanomagnetic properties of the latter compound [8, 9] are indicative of the absence of charge–orbit ordered *CE*-type AFM clusters. In this context, knowledge of the properties of $Eu_{0.55}Sr_{0.45}MnO_3$ can provide new information on the influence of the charge–orbit ordered phase clusters on these properties and on the nature of observed peculiarities.

## 2. EXPERIMENTAL

The sample of $Eu_{0.55}Sr_{0.45}MnO_3$ was synthesized according to the standard ceramic technology. The phase composition and crystal lattice parameters were monitored by X-ray diffraction measured on a Siemens D5000 diffractometer. According to these data, the sample represented a single-phase perovskite with an orthorhombic structure (*Pnma* space group). The orthorhombicity parameter of 0.2% (calculated using the lattice parameters) indicates that the structure is close to cubic. The sample had a tolerance factor of $\tau$ = 0.924, while the degree of disorder was $d^2$ = 0.00893. The single-phase state of the synthesized samples was confirmed by the Raman spectroscopy measurements performed on a Jobin-Yvon T64000 spectrometer with a triple monochromator. The Raman spectra showed the presence of only the phonon modes characteristic of the orthorhombic manganites with *Pnma* symmetry.

The magnetization was studied using a vibrating-sample magnetometer in a range of temperatures from 1.5 to 150 K and magnetic field strength up to 130 kOe (the measurements were performed in the Laboratory of Strong Magnetic Fields, Wroclaw, Poland). The initial magnetic susceptibility in an alternating magnetic field with an amplitude of 1 Oe and the frequency varied from 0.8 to 8 kHz was measured on an F-5063 fer-

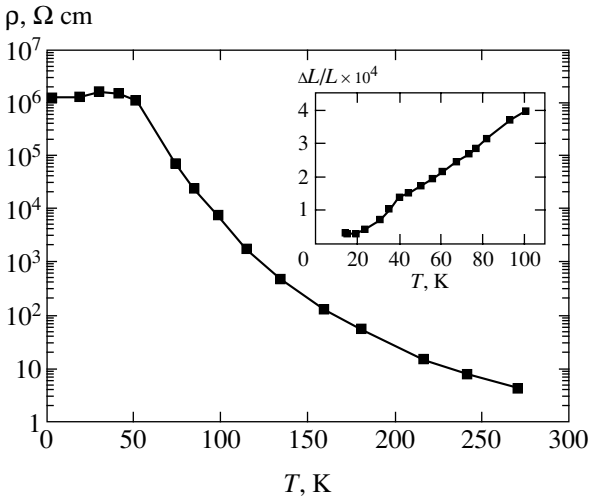



**Fig. 1.** Temperature dependences of the resistivity $\rho$ of $Eu_{0.55}Sr_{0.45}MnO_3$. The inset shows the temperature dependence of the linear thermal expansion.

rometer, while the paramagnetic susceptibility was studied by a weighing technique with electromagnetic compensation. The electric resistance was determined by the four-point-probe technique. The magnetostriction and thermal expansion in the temperature range from 4.2 to 150 K were measured using strain gauges with a resistance of 92.30 ± 0.01 Ω and a tension sensitivity coefficient of 2.26. One transducer was glued to a sample and another, to a quartz crystal, so as to have the same orientation relative to the applied magnetic field. The magnetostriction components parallel ($\lambda_\parallel$) and perpendicular ($\lambda_\perp$) to the magnetic field were measured, after which the volume ($\omega = \lambda_\parallel + 2\lambda_\perp$) and anisotropic ($\lambda_t = \lambda_\parallel - \lambda_\perp$) magnetostrictions were calculated.

## 3. RESULTS AND DISCUSSION

The temperature dependence of the initial (zero-field) magnetic susceptibility $\chi(T)$ of the sample measured in an alternating magnetic field with an amplitude of 1 Oe and a frequency of 8 kHz exhibits a maximum at $T_N$ = 41 K, which is close to the Néel temperature of undoped $EuMnO_3$. Near this temperature, we have also observed a bending point in the curve of $\rho(T)$ (Fig. 1) and a jump in the temperature dependence of the linear thermal expansion $\Delta L/L$ (see the inset to Fig. 1). As can be seen from Fig. 1, the sample in the absence of applied magnetic field behaves as an insulator when the temperature is decreased down to the minimum value used in this study ($\rho = 10^6$ Ω cm at 4.2 K).

The behavior of the isotherms of magnetization $\sigma$, magnetostriction $\omega$, and resistivity $\rho$ as functions of the magnetic field is different in the temperature intervals 4.2–40 and 70–120 K, while having much in common in each of these intervals. Figure 2 shows the $\sigma(H)$, $\omega(H)$, and $\rho(H)$ curves measured at 20 K (in the first

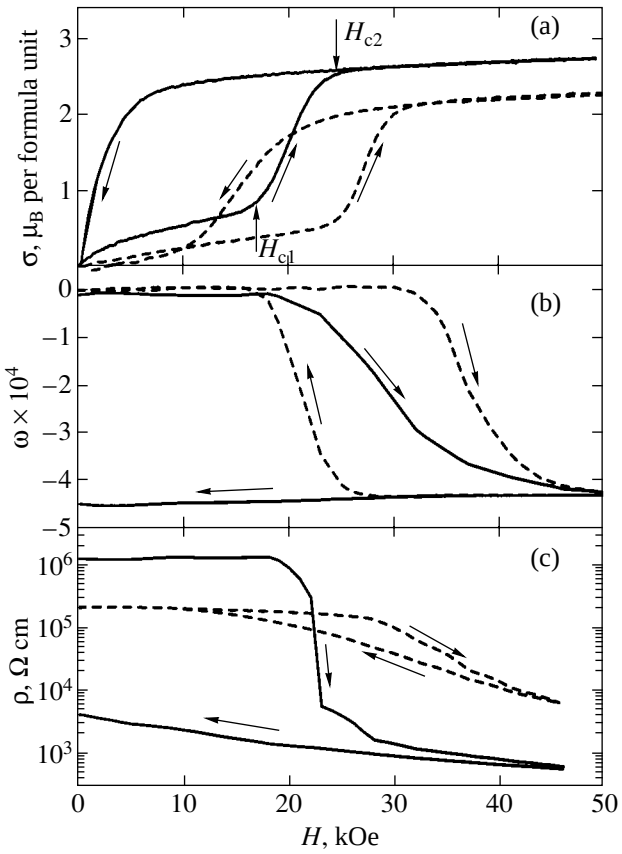



**Fig. 2.** The isotherms of (a) magnetization σ, (b) volume magnetostriction ω, and (c) resistivity ρ of $Eu_{0.55}Sr_{0.45}MnO_3$ measured at $T = 20$ K (solid curves) and 70 K (dashed curves) with increasing and decreasing applied field strength as indicated by arrows.

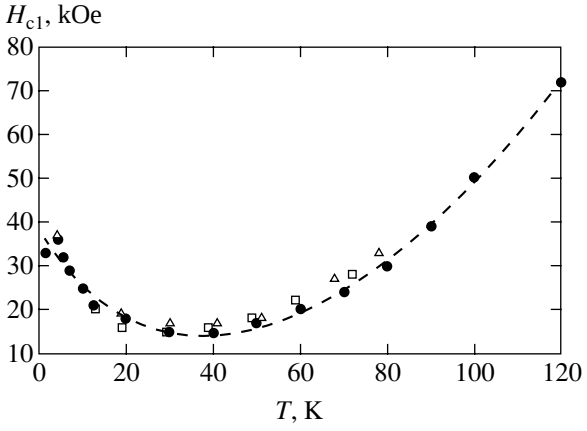



**Fig. 3.** Plot of the critical field strengths $H_{c1}$ versus temperature for $Eu_{0.55}Sr_{0.45}MnO_3$. Symbols shows the experimental data obtained from the measurements of (●) magnetization σ, (□) volume magnetostriction ω, and (△) resistivity ρ.

interval) and 70 K (in the second interval). In the first temperature interval, we observed a jumplike increase in the magnetization and in the absolute value of the magnetostriction and a sharp drop in the resistivity in a certain interval of magnetic fields, $H_{c1} < H < H_{c2}$. Judging from the value of magnetization at $H > H_{c2}$, there is a transition to the FM state in a rather large part of the sample. In the region of $H > H_{c2}$, all three parameters exhibit linear growth with the field strength *H*, showing no saturation even at the maximum magnetic fields used in this study (80 kOe for the magnetization and 50 kOe for the magnetostriction and magnetoresistance). The magnitude of magnetization at 4.2 K in a field of 80 kOe was $\sigma = 2.50\mu_B$ per formula unit, which is significantly lower (~70%) than the value ($3.55\mu_B$ per formula unit) corresponding to the FM ordering of magnetic moments of the $Mn^{3+}$ and $Mn^{4+}$ ions. In the region of jumplike variation, the curves of σ(*H*), ω(*H*), and ρ(*H*) exhibit a hysteresis whose width decreases with increasing temperature. If the magnetic field is switched off at a low temperature, the volume magnetostriction and resistivity do not return to the initial values (at least during the observation time, which was about 180 s). These initial values could be restored only by heating the sample up to about 100 K followed by cooling.

In the temperature interval 70–120 K, the jumps in the σ(*H*), ω(*H*), and ρ(*H*) curves are also observed, but the shapes of these curves change and the sample returns to the initial state after switching off the magnetic field (Fig. 2). It should be noted that behavior of the curves in this temperature interval is analogous to that observed previously in ceramics of the $Sm_{1-x}Sr_xMnO_3$ system with $x = 0.45$ and 0.40 [1, 10]. At 40 K < *T* < 70 K, the σ(*H*), ω(*H*), and ρ(*H*)) curves exhibit peculiarities typical of both temperature intervals (the sample does not return to the initial state after switching off the magnetic field; the critical field $H_{c1}$ increases with temperature).

Figure 3 shows the temperature dependence of the first critical field, $H_{c1}$, determined from the measurements of magnetization σ, magnetostriction ω, and resistivity ρ. As can be seen, the $H_{c1}$ values coincide at each particular temperature, decrease with increasing temperature in the first temperature interval, and increase in the second temperature interval, thus exhibiting a broad minimum with the lowest point at about 41 K.

Figure 4 shows the temperature dependences of the magnetization σ, magnetostriction ω, and magnetoresistance Δρ/ρ of the sample measured in various magnetic fields. As can be seen, the σ(*H*) curves exhibit a maximum at a temperature near 41 K, while the temperature dependences of ω and Δρ/ρ at $H > H_{c1}$ exhibit minima in the same region. The magnetic transition is strongly smeared under the action of applied magnetic field. In a magnetic field of 45 kOe, the volume magnetostriction reaches a giant value of $4.5 \times 10^{-4}$ in the temperature interval 10–75 K. The magnetoresistance of the sample also behaves rather unusually. At $H < H_{c1}$, the Δρ/ρ value is positive and exhibits a maximum near

41 K, where it reaches a level of 6% (Fig. 5). At $H > H_{c1}$, the magnetoresistance becomes negative, passes through a minimum approximately at 41 K, and then reaches a colossal value of $3 \times 10^5$% in a magnetic field of 50 kOe (Fig. 4).

As was noted above, the behavior of magnetization, magnetostriction, and magnetoresistance at $T \geq 70$ K in the $Eu_{0.55}Sr_{0.45}MnO_3$ sample under consideration is analogous to that observed previously at $T > T_C$ in ceramics of the $Sm_{1-x}Sr_xMnO_3$ system with $x = 0.45$ and 0.40 [1, 10]. Our investigation of a single crystal with the composition $Sm_{0.55}Sr_{0.45}MnO_3$ showed that its properties are identical to those of ceramics with the same composition. As was indicated in the introduction section, the sample of $Sm_{0.6}Sr_{0.4}MnO_3$ at 1.5 K $\leq T \leq T_C$ comprises a simply connected FM matrix containing AFM clusters of *A* and *CE* (charge–orbit ordered) types. Note that $T_C < T_N < T_{co}$ [6], where $T_{co}$ (defined as the temperature of thermal breakage of the charge–orbit ordering) coincides with the temperature of breakage of the AFM order in clusters with charge–orbit ordering. The behavior of magnetization, magnetoresistance, and volume magnetostriction at $T > T_N$ was explained [1, 10] in terms of the magnetic-field-induced transition of the *CE*-type AFM clusters to the FM state. In a *CE*-type AFM state, the magnetic moments of most nearest neighbors of manganese ions exhibit an AFM ordering and, hence, it is necessary to apply a certain threshold field for the transition from AFM to FM state. The *CE*-type AFM clusters are fully converted to the FM state in the field $H_{c2}$.

Since the isotherms of σ, ω, and Δρ/ρ observed in this study for $Eu_{0.55}Sr_{0.45}MnO_3$ in the temperature interval 1.5 K $\leq T \leq$ 120 K are very much like those observed previously at $T > T_C$ in the samples of $Sm_{0.55}Sr_{0.45}MnO_3$ and $Sm_{0.6}Sr_{0.4}MnO_3$ [1], we may suggest that our sample of $Eu_{0.55}Sr_{0.45}MnO_3$ ceramics also comprises a mixture of FM, *A*-type AFM, and *CE*-type AFM phases. However, the relative volumes of these phases in the $Eu_{0.55}Sr_{0.45}MnO_3$ sample studied is different from that in $Sm_{0.55}Sr_{0.45}MnO_3$ and $Sm_{0.6}Sr_{0.4}MnO_3$ ceramics. As can be seen from Fig. 1, the temperature dependence of the resistivity of $Eu_{0.55}Sr_{0.45}MnO_3$ exhibits a semiconductor character with a rather large value of ρ at 4.2 K (on the order of $10^6$ Ω cm), in contrast to the behavior observed for $Sm_{0.55}Sr_{0.45}MnO_3$ characterized by conductivity of the metallic type. Therefore, we suggest that the FM phase in the $Eu_{0.55}Sr_{0.45}MnO_3$ sample studied is multiply connected (in contrast to the simply connected FM phase in $Sm_{0.55}Sr_{0.45}MnO_3$). This difference is related to the fact that the tolerance factor τ = 0.924 in our $Eu_{0.55}Sr_{0.45}MnO_3$ ceramics is lower than that in $Sm_{0.55}Sr_{0.45}MnO_3$ (where τ = 0.927), while the degree of disorder $d^2 = 0.00893$ in $Eu_{0.55}Sr_{0.45}MnO_3$ is greater than that in $Sm_{0.55}Sr_{0.45}MnO_3$ (where $d^2 = 0.00784$),

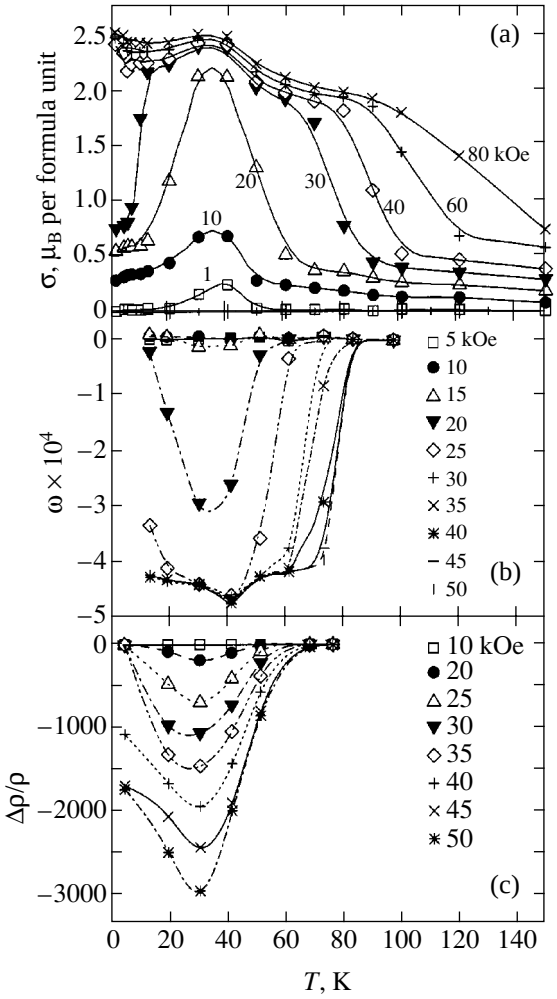



**Fig. 4.** Temperature dependences of the (a) magnetization σ, (b) volume magnetostriction ω, and (c) magnetoresistance Δρ/ρ of $Eu_{0.55}Sr_{0.45}MnO_3$ measured in various magnetic fields.

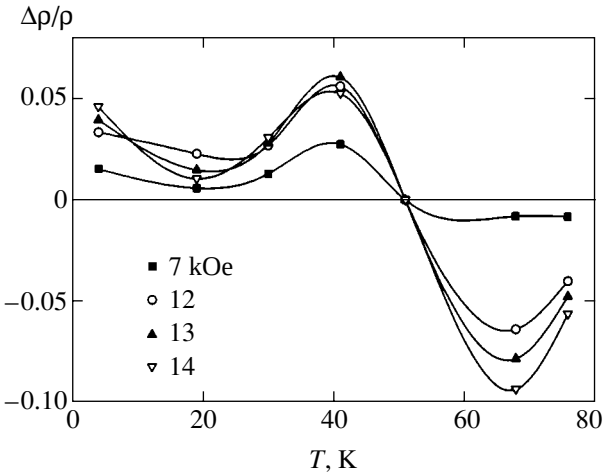



**Fig. 5.** Temperature dependences of the magnetoresistance Δρ/ρ of $Eu_{0.55}Sr_{0.45}MnO_3$ measured in various fields $H < H_{c1}$.

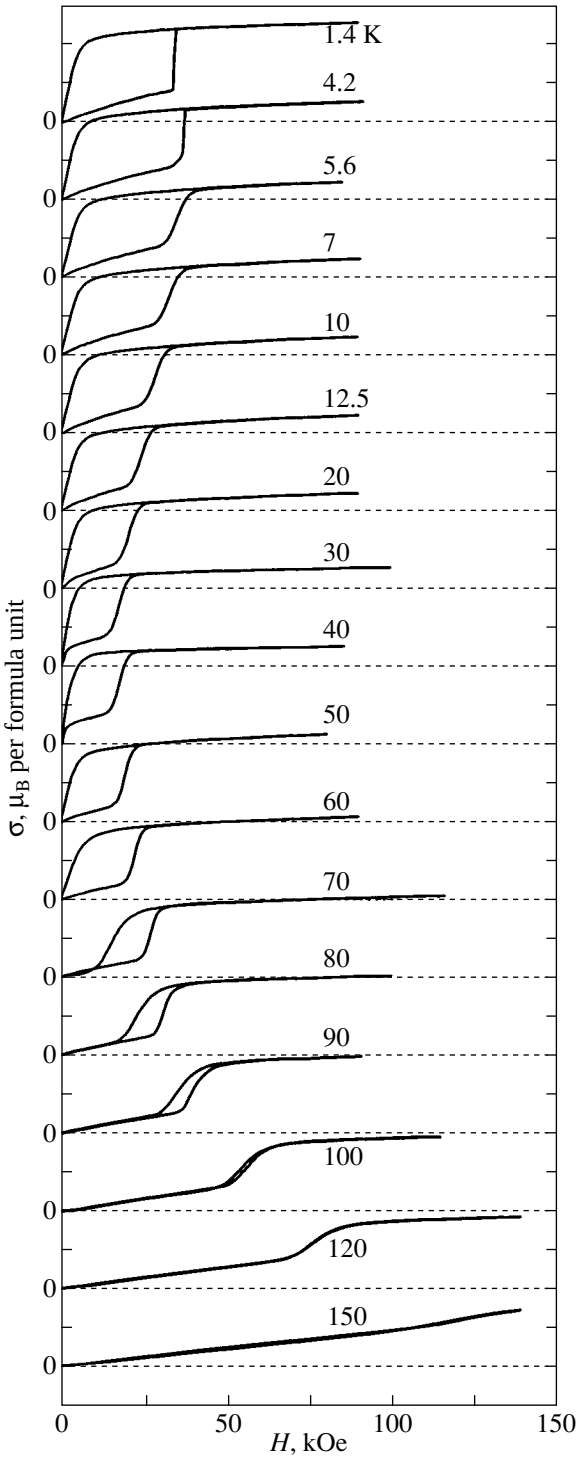



**Fig. 6.** The isotherms of magnetization σ of $Eu_{0.55}Sr_{0.45}MnO_3$ measured at various temperatures (one division on the ordinate axis corresponds to 1 $\mu_B$ per formula unit). Lower and upper curves were obtained in the field increase and decrease mode, respectively.

Nagaev [11] showed that an AFM semiconductor may occur, depending on the charge carrier density, in either insulating or conducting two-phase (FM–AFM) magnetic state. In this state, the carriers concentrate in the FM part of the crystal due to a gain in the *s–d* exchange energy, while the AFM part is depleted of carriers. It was also shown [12] that the volume fraction of the FM phase in a sample occurring in the insulating two-phase magnetic state is much smaller than that in the conducting state. For example, in doped EuSe, the FM phase volume in the conducting two-phase magnetic state is about ten times that in the insulating state [13]. It can be suggested that $Eu_{0.55}Sr_{0.45}MnO_3$ features an insulating two-phase magnetic state, whereas $Sm_{0.55}Sr_{0.45}MnO_3$ occurs in the conducting two-phase magnetic state. According to this, the magnetization in $Sm_{0.55}Sr_{0.45}MnO_3$ at $T < T_C$ proceeds via increase in the volume of the FM phase occupying most of the sample and growing by the FM planes of the *A*-type AFM phase. On this background, a jumplike increase in the magnetization related to the *CE*-type AFM clusters is not manifested because these clusters have a significantly smaller volume than the FM phase. The jumps on the $\sigma(H)$ curves are manifested only at high temperatures, where the FM and *A*-type AFM phases are thermally broken and only the *CE*-type AFM clusters are retained.

For $Eu_{0.55}Sr_{0.45}MnO_3$, the jumps in the isotherms of magnetization take place both below and above the Néel temperature ($T_N$ = 41 K) determined at the maximum of the initial susceptibility measured in alternating magnetic field. This temperature corresponds to breakage of the *A*-type AFM magnetic phase, since it coincides with $T_N$ of the undoped compound $EuMnO_3$ featuring the *A*-type AFM order. As can be seen from the data presented in Fig. 6, the jumps on the magnetization isotherms are still observed at $T$ = 120 K and are missing at $T$ = 150 K, so that the Néel temperature of the *CE*-type AFM phase falls between 120 and 150 K. It is also seen that, in the temperature interval 1.4 K ≤ $T$ ≤ 40 K (Fig. 6) and the magnetic fields $0 \le H < H_{c1}$, the $\sigma(H)$ curves can be represented as the sum of a small spontaneous magnetization ($\sigma \sim 0.1\mu_B$) and a component linearly increasing with the field H (which is characteristic of the AFM state). Thus, the volume fraction of the FM phase in this case is small, reaching only about 3% of the sample volume (as can be judged from the ratio of the magnetization of the FM phase and the saturation magnetization $\sigma_s = 3.55\mu_B$ corresponding to complete FM ordering of the sample).

As can be seen in Fig. 6, the spontaneous magnetization component disappears in the interval 40 K < $T$ < 50 K. This implies that the Curie temperature $T_C$ falls in this interval and is close to the Néel temperature $T_N$ for the composition studied. As was noted above, the magnetization at 4.2 K after a jump observed on the magnetization isotherms amounts to about 70% of the saturation magnetization, while the jump is approximately equal to $1.45\mu_B$ and changes only slightly with temperature. Assuming that the jump is due to the *CE*-type AFM phase, we infer that the volume of this phase is about 41% of the sample volume and the remaining

56% are occupied by the *A*-type AFM phase. After the jump, the magnetization exhibits a slow increase with field *H*, which is related to the magnetic-field-induced growth of the FM phase by the FM layers in the *A*-type AFM phase.

As is known, the *CE*-type AFM phase with a charge-orbit ordering is most stable at low temperatures, while at elevated temperatures it is least stable than the FM and *A*-type AFM phases [7]. For this reason, the critical field $H_{c1}$ decreases with increasing temperature in the interval below 40 K (Fig. 3), where the *CE*-type AFM phase coexists with the FM and *A*-type AFDM phases. Beginning at $T = 50$ K and up to 120 K, the $H_{c1}$ value exhibits a growth. This behavior is probably related to the fact that, at $T > T_C$, maintaining the FM order (formed as a result of the magnetic-field-induced transition of the *CE*-type AFM phase to the FM state) requires higher magnetic fields as compared to those sufficient at $T < T_C$. As was indicated above, the *CE*-type AFM phase is characterized by the orbital ordering and the related charge ordering. If the magnetic-field-induced transition from this state to the FM state were related only to breakage of the charge ordering, we might expect the appearance of a paramagnetic state at $T > T_C$. However, the results of our experiments show that the *CE*-type AFM phase converts into the FM phase. This is related to the fact that the charge carriers, being disordered under the action of the magnetic field $H > H_{c1}$, retain the FM order due to a gain in the *s–d* exchange energy.

Figures 2 and 6 show that the $\sigma(H)$, $\rho(H)$, and $\omega(H)$ curves measured with increasing and decreasing the magnetic field differ, and the isotherms measured in the field decrease mode exhibit no jumps. This means that the FM state, formed in a magnetic field above $H_{c1}$ as a result of transformation of the *CE*-type AFM phase, is retained when the field decreases down to values negligibly small as compared with $H_{c1}$. The giant magnitude of $\omega$ and the sharp drop in $\rho$ as a result of this transformation are retained upon switching off the magnetic field. The initial state can be restored only by heating the sample to $T > 100$ K followed by cooling to a required temperature. This fact indicates that the free energies of the FM and *CE*-type AFM phases are close and the magnetic-field-induced transition of the *CE*-type AFM phase to the FM state is a first-order phase transition. Beginning at $T = 60$ K or higher (i.e., above $T_C$), switching off the magnetic field reduces the magnetostriction to $\omega = 0$ and restores the resistivity $\rho$ on the initial level (before field application), although there is still a difference between the $\rho(H)$ and $\omega(H)$ curves measured with increasing and decreasing magnetic field (Fig. 2).

The negative sign of the colossal magnetoresistance observed in the $Eu_{0.55}Sr_{0.45}MnO_3$ sample studied is mostly related to the FM phase formed as a result of the magnetic-field-induced transition of the *CE*-type AFM phase to the FM state. Since a semiconductor type of the conductivity after this transformation is retained, we may conclude that the FM phase comprises an insulating *A*-type AFM matrix containing FM clusters in which the charge carriers (holes) are concentrated due to a gain in the *s–d* exchange energy. The phenomenon of colossal magnetoresistance in the compound studied is explained as follows. The external magnetic field increases the radii of the FM clusters, thus facilitating the tunneling of charge carriers between these clusters. In addition, ordering of the magnetic moments of the FM clusters in the applied field also favors the tunneling of charge carriers. Finally, the external field tends to break the FM clusters by increasing the energy of holes concentrated in these clusters and favoring their transition into a delocalized state [5, 12]. The large positive magnetoresistance observed for $H < H_{c1}$ and $T < T_C$ is probably related to the FM phase and can be explained in the same way as in nondegenerate FM semiconductors [14].

The giant volume magnetostriction is also related to the FM phase formed as a result of the magnetic-field-induced transition of the *CE*-type AFM phase to the FM state. Yanase and Kasuya [15] showed that the crystal lattice parameters inside the FM clusters are decreased as a result of the charge redistribution leading to a decrease in the energy at the expense of an increase in the overlap between the clouds of charge of the central impurity ion and the nearest-neighbor magnetic ions. Apparently, thermal breakage of the two-phase magnetic state must be accompanied by an excessive (as compared to the linear in *T*) thermal expansion of the sample, which was actually observed in our experiments (see the inset to Fig. 1). As is known, the temperature dependence of the thermal expansion in dia- and paramagnets is almost linear. The excess thermal expansion can be suppressed by applying an external magnetic field restoring the two-phase magnetic state. This must give rise to a negative volume magnetostriction, also in agreement with what was observed in our experiments (Figs. 2 and 4).

## 4. CONCLUSIONS

We have found, in addition to the previously studied $Sm_{1-x}Sr_xMnO_3$ compounds with $0.4 \le x \le 0.45$, another compound, $Eu_{0.55}Sr_{0.45}MnO_3$, in which the colossal magnetoresistance and the giant volume magnetostriction appear as a result of the magnetic-field-induced transition of the charge–orbit ordered *CE*-type AFM phase to the FM state. These systems clearly demonstrate that the phenomena of colossal magnetoresistance and giant volume magnetostriction are related to modification of the FM phase, which appears at $T > T_C$ and exhibits a sharp growth at $T < T_C$ as a result of the aforementioned magnetic-field-induced transition. Although both manganite systems comprise a mixture of three phases—FM, *A*-type AFM, and *CE*-type AFM

with charge–orbit ordering—the ratios of these phases in the two systems are different. Apparently, $Sm_{1-x}Sr_xMnO_3$ samples with $0.4 \le x \le 0.45$ predominantly contain at $T < T_C$ the FM phase, while $Eu_{0.55}Sr_{0.45}MnO_3$ under these conditions consists for the most part of the *CE*- and *A*-type AFM phases. This assumption is confirmed by the fact that the first system exhibits a metallic conductivity (the FM phase, in which the charge carriers concentrate due to a gain in the *s–d* exchange energy, is simply connected), while the second system has a conductivity of the semiconductor type (the FM phase, comprising separate clusters in the AFM matrix, is multiply connected). For this reason, the jumps in the isotherms of magnetization, magnetoresistance, and volume magnetostriction (indicative of the magnetic-field-induced transition of the charge–orbit ordered *CE*-type AFM phase to the FM state) are observed in $Eu_{0.55}Sr_{0.45}MnO_3$ beginning at the minimum temperatures studied (1.4 K), whereas in $Sm_{1-x}Sr_xMnO_3$ ($0.4 \le x \le 0.45$) these peculiarities are manifested only after breakage of the FM phase. Apparently, it is a small volume fraction of the charge–orbit ordered *CE*-type AFM phase (compared to the FM phase volume in $Sm_{1-x}Sr_xMnO_3$ with $0.4 \le x \le 0.45$) that explains the absence of such jumps at $T < T_C$. In $Eu_{0.55}Sr_{0.45}MnO_3$, the values of the magnetization, volume magnetostriction, and resistivity after the aforementioned magnetic-field-induced transition at $T < T_C$ are retained upon switching off the applied magnetic field. This fact indicates that the free energies of the FM and *CE*-type AFM phases in this compound are close.

## ACKNOWLEDGMENTS

This study was supported by the Russian Foundation for Basic Research, project no. 03-02-16100.

## REFERENCES

1. A. I. Abramovich, L. I. Koroleva, and A. V. Michurin, Zh. Éksp. Teor. Fiz. **122**, 1063 (2002) [JETP **95**, 917 (2002)].
2. R. V. Demin, L. I. Koroleva, R. Szymszak, *et al.*, Pis'ma Zh. Éksp. Teor. Fiz. **75**, 402 (2002) [JETP Lett. **75**, 331 (2002)].
3. L. I. Koroleva, A. I. Abramovich, R. V. Demin, *et al.*, Fiz. Nizk. Temp. **27**, 379 (2001) [Low Temp. Phys. **27**, 293 (2001)].
4. A. I. Abramovich, A. V. Michurin, O. Yu. Gorbenko, *et al.*, J. Phys.: Condens. Matter **12**, L627 (2000).
5. É. L. Nagaev, Usp. Fiz. Nauk **166**, 833 (1996) [Phys. Usp. **39**, 781 (1996)]; Phys. Rep. **346**, 387 (2001).
6. V. V. Runov, D. Yu. Chernyshov, A. I. Kurbakov, *et al.*, Zh. Éksp. Teor. Fiz. **118**, 1174 (2000) [JETP **91**, 1017 (2000)].
7. E. Dagotto, T. Hotta, and A. Moreo, Phys. Rep. **344**, 1 (2001).
8. A. I. Abramovich, R. V. Demin, L. I. Koroleva, *et al.*, Pis'ma Zh. Éksp. Teor. Fiz. **69**, 375 (1999) [JETP Lett. **69**, 404 (1999)].
9. A. Abramovich, R. Demin, L. Koroleva, *et al.*, Phys. Lett. A **259**, 57 (1999).
10. A. I. Abramovich, L. I. Koroleva, A. V. Michurin, *et al.*, Fiz. Tverd. Tela (St. Petersburg) **44**, 888 (2002) [Phys. Solid State **44**, 927 (2002)].
11. É. A. Nagaev, *Physics of Magnetic Semiconductors* (Nauka, Moscow, 1979) [in Russian].
12. E. L. Nagaev, *Colossal Magnetoresistance and Phase Separation in Magnetic Semiconductors* (Imperial College Press, London, 2002).
13. Y. Shapira, S. Foner, N. F. Oliveira, Jr., and T. B. Reed, Phys. Rev. B **10**, 4765 (1974).
14. L. I. Koroleva, *Magnetic Semiconductors* (Mosk. Gos. Univ., Moscow, 2003) [in Russian].
15. A. Yanase and T. Kasuya, J. Phys. Soc. Jpn. **25**, 1025 (1968).

*Translated by P. Pozdeev*

*Journal of Experimental and Theoretical Physics, Vol. 99, No. 4, 2004, pp. 827–833.*
*Translated from Zhurnal Éksperimental'noĭ i Teoreticheskoĭ Fiziki, Vol. 126, No. 4, 2004, pp. 954–961.*


**SOLIDS**
**Electronic Properties**

# Activation Conduction in Metallic Island Films

**A. P. Boltaev, N. A. Penin, A. O. Pogosov, and F. A. Pudonin**
*Lebedev Physical Institute, Russian Academy of Sciences, Moscow, 119991 Russia*
*e-mail: boltaev@sci.lebedev.ru; pudonin@sci.lebedev.ru*
Received March 23, 2004

**Abstract**—The differential conductivity of metallic island films of Ti, Co, W, and FeNi is investigated in the vicinity of liquid nitrogen temperatures. It is found that the temperature dependence of the conductivity of metallic island films in the insulator phase varies in accordance with the activation law $\sigma \propto T^n \exp(-E/kT)$. It is shown that the power of temperature in the preexponential factor varies from $n = 2$ to 1 upon an increase in the film thickness. In thicker films, in which a transition from the insulator to the metal conductivity phase takes place, the temperature dependence of the conductivity increases in proportion to temperature. The mechanism of conduction in metallic island films is discussed. © *2004 MAIK "Nauka/Interperiodica"*.

## 1. INTRODUCTION

The study of the structure and physical properties of ultrathin metallic films with a thickness smaller than 2–3 nm began at the end of the 1950s, when the first metal films were obtained for the first time [1]. The new stage in the study of these structures was stimulated by the interest in systems with a low dimension (quantum wells, dots, and wires) and by advances in the development of technologies for preparing various nanostructures and the methods for their investigation. It should be noted that metallic island films emit radiation in an external electric field [2] and exhibit photoconductivity in the visible and infrared spectral regions [3]. Considerable interest in these objects is due to synthesis and investigation of a new class of materials, viz., nanocomposites [4]. Nanocomposites are aggregates of small metallic grains (with a diameter from 1 to 100 nm) in an insulator matrix; thin metallic island films are special case of such structures.

For a moderate density of metallic islands in such structures, the thermally activated conductivity mode can be realized. It was found experimentally that the conductivity of metallic island structures and nanocomposites normally changes with temperature in accordance with the law $\sigma \propto \exp(-T_0/T)^{0.5}$. Thus, most experimental results can be described by the "1/2 law" [5–8]. Various theoretical models (e.g., [9, 10]) were employed for explaining the 1/2 law. These models were basically modifications of the hopping conduction theory for semiconductors, in which the 1/2 law is interpreted as the emergence of a Coulomb gap in the electron density of states in the vicinity of the Fermi level. The important role of the Coulomb interaction between charged grains and the formation of the Coulomb gap in nanocomposites was indicated in [9, 10]. Other models explaining the 1/2 law are also available; for example, the conduction in grained structures is attributed in [11] to the spread in the size of metallic granules.

It was shown in a number of publications that the exponent $x$ in the expression describing the measured value of conductivity in nanocomposites is not always equal to 1/2. Zvyagin and Keiper [12] proved that the standard theory of hopping conduction with variable range is inapplicable when the hopping range is smaller than the island size and is comparable with the thickness of interlayers between the islands. It was found that $x = 0.75$ in [13], $x = 0.72$ in [14], and $x = 1$ in [15]. The reasons for such deviations in the dependence with $x = 1/2$ remain unclear. It has been generally accepted, however, that the conduction mechanism in gained systems (in particular, in metallic island films) is associated with tunnel transitions of charge carriers between grains and is close in this respect to the hopping conduction mechanism in doped semiconductors [5].

It should be noted that all authors emphasized that the temperature dependences of the conductivity of structures were measured under voltages corresponding to the linear segment of the current-voltage characteristics. On the other hand, it was shown in some publications (see, for example, [16, 17]) that tunnel systems consisting of two metals separated by an insulating barrier do not exhibit an ohmic segment on the current–voltage characteristic if the conductivity is measured in extremely weak electric fields. This phenomenon was referred to in [16] as anomalous conduction at zero bias voltage. The effect of anomalous conduction was studied in [17] for systems in which 4–12-nm metal (Sn) particles were contained in an insulating layer; it was shown that the conductivity of these structures in weak electric fields may change by a factor of several units. With increasing temperature, the effect of anomalous conduction becomes weaker.

Thus, a unified concept of the mechanism of conduction in nanocomposites and thin metallic island

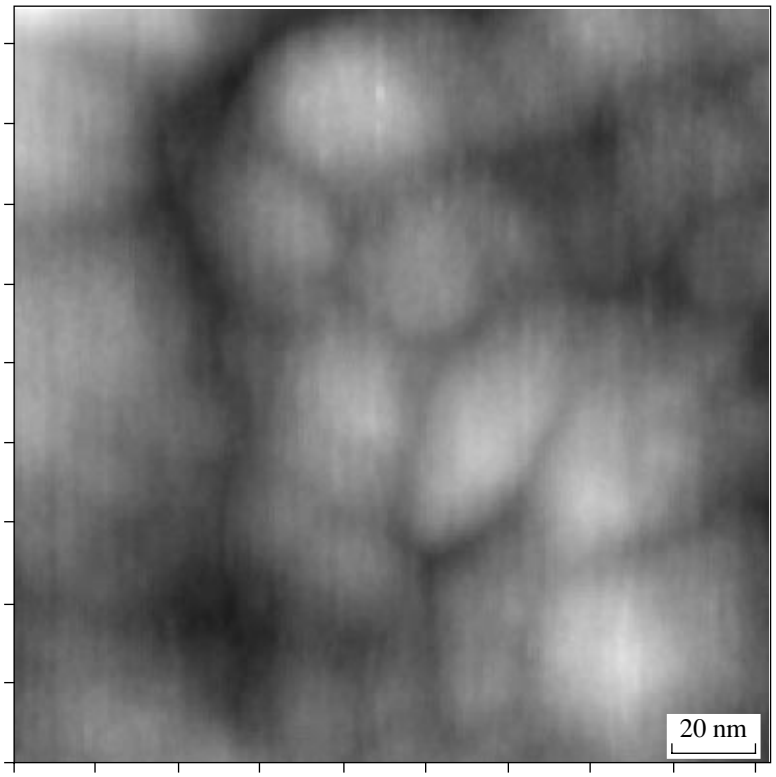



**Fig. 1.** Image of the surface of a 6.3-Å-thick tungsten film deposited on an insulating substrate.

films has not been worked out. At the same time, experimental studies of transport properties in these structures should be carried out with certain care in view of the anomalous conduction effect.

In this study, we report on an investigation of the differential conductivity of ultrathin metallic films of Ti, Co, W, and FeNi in the vicinity of liquid nitrogen temperatures, when the anomalous conduction effect does not introduce substantial errors in measured values of the conductivity of structures.

## 2. TECHNOLOGY AND RESULTS OF MEASUREMENTS

Thin films of Ti, W, FeNi, etc., were grown by microwave sputtering in argon. The details of the method for obtaining such films are given in [3]. After sputtering, each metallic structure was coated by a thin layer of an insulator ($Al_2O_3$, $d$ = 8 Å) or a semiconductor (ZnTe, SiC, $d$ = 8 Å). The thickness of the metallic films varied from 5 to 100 Å.

To measure the differential conductivity of such structures, mesostructures 0.5–1.5 mm in width and 2–3.5 mm in length were prepared. Ohmic contacts were prepared by sputtering of indium on the surface of a metallic film. The differential conductivity of the films was measured in the temperature range from 77 to 300 K. It was found that the frequency dependence of the conductivity remained unchanged for all samples up to $f$ = 1000 Hz. In the course of measurements, we monitored the variation of differential conductivity under the action of a weak electric field. For this purpose, in addition to a varying component, a constant voltage $U_0$ was supplied to the sample. The error in the conductivity measurements for island films was approximately 5%.

Analysis of anomalous conduction in metallic island films at the liquid nitrogen temperature and at a constant electric field $E_0 \sim 40$ V/cm in the samples proved that the differential conductivity of some samples changes by 30–40%. With increasing temperature, the effect of anomalous conduction was suppressed. In view of the arguments formulated above, to avoid errors associated with the anomalous conduction effect in the study of the temperature dependence of the differential conductivity of the samples, measurements were made at a frequency $f = \omega/2\pi \sim 135$ Hz and the amplitude $U_1$ of variable voltage did not exceed 10 mV ($E_1 \sim 50$ mV/cm; $U_0 = 0$). Under such conditions, the error of measurement of the differential conductivity of metallic island films did not exceed 0.1% in the temperature range under investigation.

Apart from the investigation of the conductivity of metallic films, we studied the topography of the obtained structures using atomic-force microscopy. Figure 1 shows a typical image of the surface of a 6.3-Å-thick tungsten film deposited on the insulating substrate. It can be seen that the film is of the island

type with island sizes of ~150–300 Å; the islands are separated by distances of about 15–100 Å. Experiments with film samples of the same thickness, but made of a different metal, give similar images differing only in the island size. As the film thickness increases, the islands gradually merge together and the film becomes nearly continuous when the thickness of the metallic layers is on the order of 20 Å.

Figure 2 shows the temperature dependence of the specific surface differential conductivity of a tungsten film on an insulating substrate for eight structures with various thicknesses of metallic films ($d$ = 6.3 (*1*), 7 (*2*), 7.9 (*3*), 8.5 (*4*), 9.48 (*5*), 11 (*6*), 19 (*7*), and 100 Å (*8*)). The temperature dependences for the conductivity of Ti- and FeNi-based structures differ insignificantly from the temperature dependence of the tungsten conductivity shown in Fig. 2. It can be seen from the figure that the conductivity of the tungsten film on an insulating substrate varies by seven orders of magnitude when the film thickness changes approximately by a factor of three (from 6 to 19 Å). With increasing temperature, the conductivity of the films increases. For metal film thicknesses of 19 and 100 Å, the conductivity of the films is virtually independent of temperature in the temperature range under investigation.

To study the effect of anomalous conductivity on the processes of charge carrier transfer in metallic island films, the response time for the current through the sample corresponding to a constant voltage step $U_0$ was investigated. Measurements were made at $T$ = 77 K. A voltage $U = U_0 + U_1\cos(\omega t)$ was applied to the sample and $U_0$ was changed jumpwise from zero to $U_0$ = 4 V ($E_0 \sim 20$ V/cm). The results of measurements of the current response through the sample for a tungsten film of thickness $d$ = 7.9 Å are shown in Fig. 3. The same figure presents the time dependences of alternating current through the sample in relative units ($I_\sim$) as well as variation of dc conductivity ($\sigma$) and the variation of the differential conductivity measured at frequency $f$ = 135 Hz in relative units (in the present case, these quantities coincide and are represented by the same curve in the figure). It can be seen from the figure that the time dependence of variation of the current through the sample is complex by nature. This dependence can be approximated by two exponentials with characteristic times $\tau \sim 10^{-5}$ s and $\tau \sim 10$ s. The variation of the alternating current and conductivity as a result of application of a step voltage to the sample is characterized by an exponential dependence with a characteristic time $\tau \sim 10$ s.

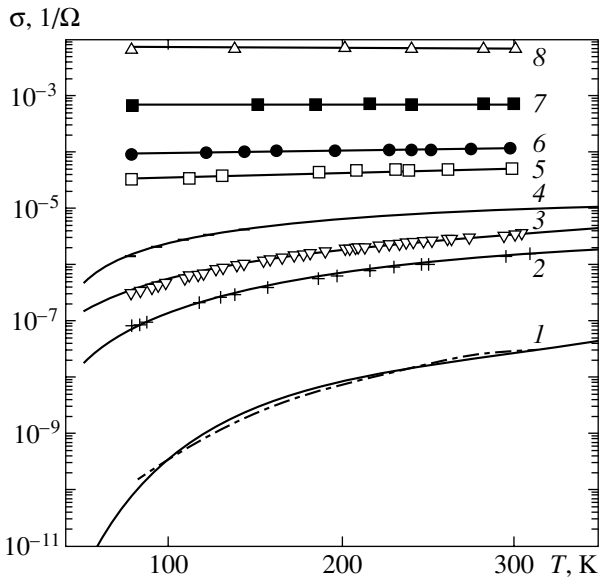



**Fig. 2.** Temperature dependence of the surface conductivity of a tungsten film for eight structures with various thicknesses of metal films: 6.3 (*1*), 7 (*2*), 7.9 (*3*), 8.5 (*4*), 9.8 (*5*), 11 (*6*), 19 (*7*), and 100 Å (*8*).

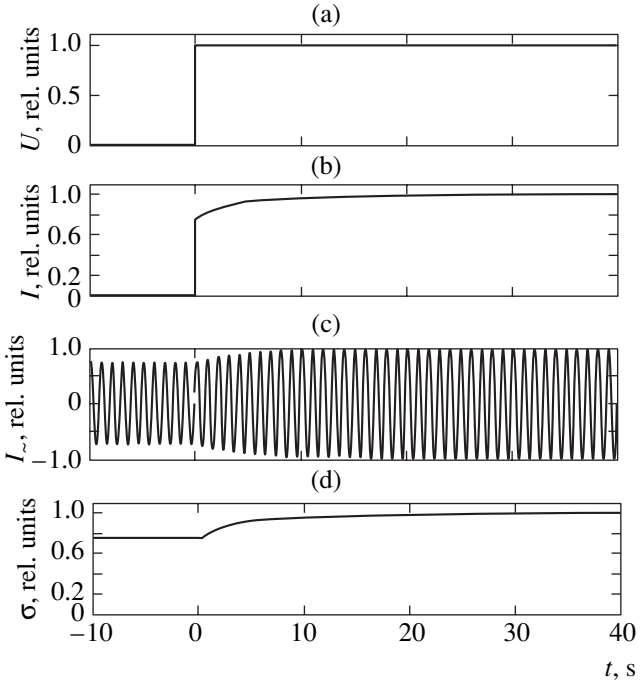



**Fig. 3.** Time dependences of variation of current through the structure and its conductivity upon a change in voltage from 0 to 4 V.

## 3. DISCUSSION OF RESULTS

The sample with a tungsten film of thickness $d$ = 6.3 Å was chosen on purpose to demonstrate the structure of the surface of an island film (Fig. 1). A tungsten film deposited of an insulating substrate becomes a conductor precisely at this value of thickness (see Fig. 2). For a smaller film thickness, no electric current through the film is observed under the present conditions.

The measured temperature dependences of the surface conductivity of metallic films shown in Fig. 2 are of the dielectric type. The sharp increase in the conduc-

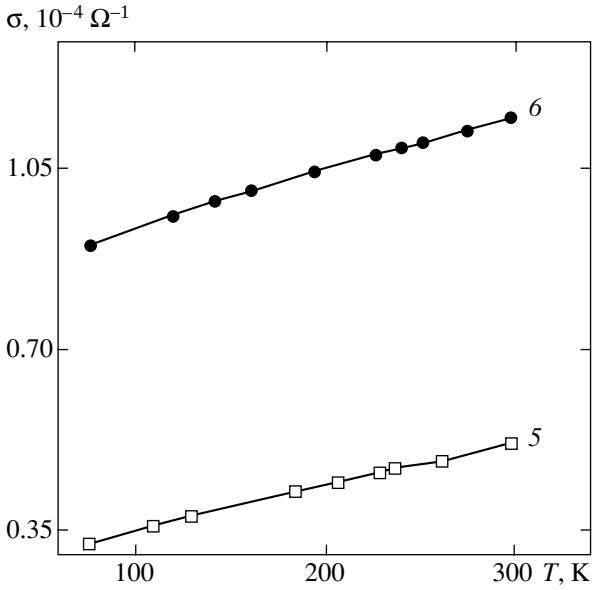



**Fig. 4.** Temperature dependence of the conductivity of a tungsten film of thickness 9.8 (*5*) and 11 Å (*6*).

tivity with increasing temperature of the samples indicates the activation nature of the conduction. It was noted above that the temperature dependence of the activation conductivity σ of island and granular films is explained most often in terms of the theory of hopping conduction with a variable hopping length (the 1/2 law). The value of this conductivity is proportional to the product of the probability of charge carrier tunneling between islands and the activation probability [5, 6, 11]:

$$\sigma \propto \sigma_0 \exp(-L/\lambda - E/kT) \propto \sigma_0 \exp(-(T_0/T)^{0.5}), \quad (1)$$

where $L$ is the hopping length, $\lambda = \hbar/(mW)^{0.5}$ is the decay length for the wavefunction of an electron in the insulator separating metallic islands, $m$ is the electron mass, $W$ is the tunnel barrier height (which nearly coincides with the half-width of the forbidden gap of the insulator), $E$ is the activation energy, $k$ is the Boltzmann constant, and $T_0$ is the parameter determining the temperature dependence of conductivity.

The measured temperature dependences of conductivity for samples *1–4* shown in Fig. 2 are correctly approximated by the 1/2 law. The values of parameter $T_0$ for curves *1–4* are given in the table. The conductivities of samples *5* and *6* vary in proportion to temperature (Fig. 4). It is difficult to explain such a temperature dependence of the conductivity in the model of hopping conduction with a variable hopping length. The question arises concerning the applicability of the 1/2 law for explaining the temperature behavior of the conductivity for samples *1–4*. To answer this question, we estimate the hopping length $L$, which is a parameter in Eq. (1), and compare the obtained value of $L$ with actual longitudinal sizes $R$ of the islands and the spacing $s$ between them, which can be estimated from Fig. 1. In accordance with Eq. (1), the hopping length $L$ is given by

$$L = [(T_0/T)^{0.5} - E/kT]\lambda. \quad (2)$$

Hopping length $L$ was estimated for sample no. 1 at temperatures $T$ = 100 and 300 K; $T_0$ = 9400 K (see table); $\lambda = \hbar/(mW)^{0.5} \approx 2$ Å; and $W \approx 1.5$ eV. Obviously, the maximal value of the hopping length ($L_{max}$) in the case when the 1/2 law is valid is observed for $E = 0$. Substituting the values of $T_0$, $E$, and λ into Eq. (2), we determine the hopping length ($L_{max}$) for temperatures $T$ = 100 and 300 K. The value of $L_{max}$ = 19.4 Å at $T$ = 100 K and $L_{max}$ = 11.2 Å at $T$ = 300 K.

It can be seen from Fig. 1 that the actual distance $s$ between islands is 15–100 Å, which is comparable to the hopping length $L_{max}$ calculated from Eq. (2). If, however, we take into account the fact that the size of the system formed by two neighboring islands exceeds 400 Å, we can assume that tunneling is possible only between adjacent islands and the standard theory of hopping conduction with a variable hopping length is inapplicable in this case. Consequently, the conduction in island films is associated with electron tunneling between neighboring islands and the hopping length is

**Table**

| | | | | | |
|---|---|---|---|---|---|
| 1 | $d$ = 6.3 Å | $\sigma \propto \exp(-T_0/T)^{1/2}$ | $T_0$ = 9400 K | | |
| | | $\sigma \propto T^n \exp(-E/kT)$ | | $E = 2.92 \times 10^{-2}$ eV | $n$ = 2 |
| 2 | $d$ = 7 Å | $\sigma \propto \exp(-T_0/T)^{1/2}$ | $T_0$ = 3400 K | | |
| | | $\sigma \propto T^n \exp(-E/kT)$ | | $E = 8.8 \times 10^{-3}$ eV | $n$ = 1.5 |
| 3 | $d$ = 7.9 Å | $\sigma \propto \exp(-T_0/T)^{1/2}$ | $T_0$ = 2100 K | | |
| | | $\sigma \propto T^n \exp(-E/kT)$ | | $E = 7.3 \times 10^{-3}$ eV | $n$ = 1.2 |
| 4 | $d$ = 8.5 Å | $\sigma \propto \exp(-T_0/T)^{1/2}$ | $T_0$ = 1000 K | | |
| | | $\sigma \propto T^n \exp(-E/kT)$ | | $E = 6 \times 10^{-3}$ eV | $n$ = 1 |
| 5, 6 | $d$ = 9.48; 11.06 Å | $\sigma \propto T$ | | | |
| 7, 8 | $d$ = 19; 100 Å | $\sigma \approx$ const | | | |

determined by the thickness of the insulator between metallic islands and is independent of temperature. With increasing film thickness, the sizes of islands increase, while hopping length $L_{max}$ decreases since parameter $T_0$ becomes smaller (see table). Obviously, electron tunneling in thicker films is also possible only between neighboring islands and the hopping length is independent of temperature. In this connection, the temperature dependence of the conductivity of metallic island films cannot be interpreted in terms of the model of hopping conductivity with a variable hopping length for all samples investigated here.

The temperature dependence of the conductivity of island films can be approximated by the empirical equation

$$\sigma = \sigma_0(T)\exp(-L/\lambda)\exp(-E/kT) \propto T^n\exp(-E/kT). \quad (3)$$

In this expression, only the factors that may be functions of temperature are taken into account. The important role of the temperature dependence of the preexponential factor was noted in a number of publications [5, 7].

We compared the measured temperature dependences of conductivities for samples *1–4* with those calculated from Eq. (3). As fitting parameters, we used the power $n$ of the temperature in the preexponential factor and the activation energy $E$. The values obtained for $n$ and $E$ are given in the table, while the calculated dependences of the conductivity for samples *1–4* are shown by solid curves in Fig. 2. It can be seen that the experimental and calculated temperature dependences of conductivity are in good agreement. Table shows that the activation energy $E$ and power $n$ of temperature in the preexponential factor decrease with increasing film thickness.

Thus, the mechanism of conduction in metallic island films is associated, as in [6, 11, 14], with tunneling of charge carriers between neighboring metallic islands. At a nonzero temperature, thermodynamic equilibrium sets in due to tunnel transition of electrons between metallic islands; in this case, a fraction of islands acquire a positive or negative charge (as a result of transition of an electron from one neutral island to another). The energy of the system in this case changes by [11]

$$E = e^2/2C(R_1, R_2, s), \quad (4)$$

where $e$ is the electron charge. The capacitance $C(R_1, R_2, s)$ of a capacitor formed by two islands depends on their longitudinal sizes ($R_1$, $R_2$) and spacing $s$. If $R_1 = R_2$ and the electrostatic induction associated with the charges of the islands can be disregarded, we have $E \approx e^2/C_R$, where $C_R$ is the capacitance of an island determined by its longitudinal size. With the formation of charged islands, the conditions are created for conduction due to tunneling of charge carriers between charged and neutral islands without a substantial change in the system energy, the conductivity being proportional to the number of charged islands [11, 14].

The validity of this mechanism of charge carrier transport in island films is confirmed by analysis of the time of current response to a step of constant voltage applied to the sample (see Fig. 3). It is noteworthy that the variations of the dc conductivity and of the differential conductivity measured at frequency $f$ = 135 Hz coincide and are described by an exponential dependence with time $\tau \sim 10$ s. Obviously, the electric field applied to the sample changes the activation energy $E$. The variation of the activation energy leads to a change in the number of charged islands and to variation of the conductivity in the films, the time of this variation being $\tau \sim 10$ s. On the other hand, the slow and identical variations of the dc conductivity and the ac differential conductivity are indications of the fact that current in island films is mainly associated with tunneling of charge carriers from charged to neutral islands. This process is faster than the tunneling of electrons between neutral islands since the former process occurs without a substantial change in the system energy with characteristic time $\tau \sim 10^{-5}$ s. If two independent (fast and slow) processes of electron transfer between islands existed simultaneously, the variation of conductivity at a higher frequency would differ from the variation of the dc conductivity. Thus, the mechanism of conduction in an metallic island film is associated with two successive processes.

(i) Electron tunneling from one neutral island to another with the formation of positively and negatively charged islands. Such a transition involves the change in the system energy by $E \approx e^2/C_R$.

(ii) Tunneling of charge carriers from a charged to a neutral island. In this case, the energy of the system remains unchanged if the islands have the same size or changes by $\Delta E \approx e^2(1/C_{R1} - 1/C_{R2})$ if the sizes of the islands are different. In this case, $E > \Delta E$. The conductivity of the film is proportional to the product of the probabilities of these two processes [14].

It can be seen from the table that the value of activation energy $E$ decreases with increasing film thickness, which can be explained by an increase in the capacitance of islands upon an increase in their size (see Eq. (4)). To explain the change in the power $n$ of temperature in the preexponential factor upon a change in the film thickness, we consider in greater detail the mechanism of charge carrier transport between islands.

The first process determines the concentration of charged islands ($N_{+-}$). We assume that the number of charged islands is determined by electron tunneling between two neutral islands without the participation of phonons and is proportional to the number of electrons in an island, which have an energy $\varepsilon$ higher than activation energy $E$ and which can tunnel to a neighboring neutral island. We also assume that the excess energy of electrons in islands relaxes during the interaction of charge carriers with phonons. Hence, the concentration

of charged islands $N_{+-} \propto n(\varepsilon)$, where $n(\varepsilon)$ is the concentration of electrons with energy $\varepsilon > E$ in a metallic island. If we assume that the energy of the conduction band of the metal is proportional to $\sqrt{\varepsilon}$, the number of electrons in the island, which have an energy higher than activation energy $E$, is proportional to

$$n(\varepsilon) \propto \int_{E+F}^{\infty} \frac{\sqrt{\varepsilon}d\varepsilon}{1 + \exp((\varepsilon - F)/kT)} \propto \begin{cases} T\exp(-E/kT) \text{ for } E \gg kT \\ T \text{ for } E \ll kT, \end{cases} \quad (5)$$

where $F$ is the Fermi energy measured from the bottom of the conduction band.

In this case, the concentration of charged islands in these metallic films (samples *1–4*) is

$$N_{+-} \propto n(\varepsilon) \propto T\exp(-E/kT) \text{ for } E \gg kT;$$

with increasing film thickness, $N_{+-} \propto T$ for $E \ll kT$ (samples *5* and *6*).

The rate of the second process is proportional to the probability $W$ of charge carrier tunneling from a charged island to a neutral one. This process also involves a change in the system energy by $\Delta E \approx e^2(1/C_{R1} - 1/C_{R2})$. As in the first process, the probability $W$ of electron tunneling between charged and neutral islands is approximately proportional to the number $n(\varepsilon)$ of electrons in a charged island, which possess energy $\varepsilon > \Delta E$ and which are capable of tunneling to a neighboring neutral island. The temperature dependence of the number of electrons $n(\varepsilon)$ in a charged island can be derived from expression (5) by substituting $\Delta E$ for $E$ in it. For $\Delta E \gg kT$, we have $W \propto T \times \exp(-\Delta E/kT)$, while for $\Delta E \ll kT$, $W \propto T$.

If electrons tunnel between islands without a change in the system energy (i.e., $\Delta E = 0$ and $E = 0$), the current in tunnel systems consisting of two metals separated by an insulating barrier is virtually independent of temperature [18].

It was noted above and proved in [14] that the conductivity of a metallic island film is proportional to the product of concentration $N_{+-}$ of charged islands and probability $W$ of electron tunneling between charged and neutral islands. It was also noted [11] that the total resistance of a sample of granular metals is proportional to the resistance between a pair of individual grains and inversely proportional to the concentration of charged grains. Consequently, we can assume that the conductivity of samples in metallic island films is

$$\sigma \propto N_{+-}W \propto T^2\exp(-(E + \Delta E)/kT)$$

$$\text{for } E \text{ and } \Delta E \gg kT,$$

$$\sigma \propto T^2\exp(-E/kT) \text{ for } E \gg kT \text{ and } \Delta E \ll kT.$$

This is in good agreement with the measured temperature dependence of conductivity for sample *1*.

As the film thickness and longitudinal sizes of islands increase, the value of $\Delta E$ tends to zero and the power of the exponential factor in the expression for probability $W$ of electron tunneling between charged and neutral islands changes from unity to zero. In this case, the temperature dependence of conductivity changes from $\sigma \propto T^2\exp(-E/kT)$ and for $\Delta E \gg kT$ to $\sigma \propto T\exp(-E/kT)$ for $E \gg kT$ and for $\Delta E = 0$. This type of the temperature dependence of conductivity is observed in samples *2–4*.

A further increase in the film thickness and in the island size is determined by the condition $E \ll kT$ and $\Delta E = 0$; in this case, $\sigma \sim T$, which is typical of samples *5* and *6*. For $E = 0$ and $\Delta E = 0$, we have $\sigma \approx$ const (samples *7* and *8*).

The temperature dependence of conductivity in samples *7* and *8* can apparently be explained not only by electron tunneling between islands. Indeed, with increasing film thickness, the sizes of islands increase and their spacing decreases, which results, on the one hand, in a sharp increase in the conductivity associated with tunneling. On the other hand, an increase in the island size and the emergence of bridges between islands in some cases create conditions under which the conductivity in the structure is limited to a considerable extent due to scattering of the momentum of charge carriers in metallic islands from defects and inhomogeneities in the film.

## 4. CONCLUSIONS

The differential conductivity of ultrathin metallic films of Ti, Co, W, and FeNi has been studied in the range of liquid helium temperatures, for which the effect of anomalous conductivity does not introduce substantial errors in the measured values of the conductivity of structures.

It has been established that the temperature dependence of the conductivity of metallic island films in the insulator phase varies in accordance with the activation law $\sigma \propto T^n\exp(-E/kT)$.

It is shown that, with increasing film thickness, the power of temperature in the preexponential factor varies from $n = 2$ to 1.

In thicker films, for which a transition from the insulator to the metal phase of conductivity is observed, the temperature dependence of conductivity increases in proportion to temperature.

The activation nature of the conductivity in thin films, the change in the power of temperature in the preexponential factor with increasing film thickness, and the increase in the conductivity proportional to temperature in thicker films can be explained in terms of the model of electron tunneling between adjacent islands taking into account the Coulomb interaction.

In films for which a transition from the insulator to the metal phase of conduction is observed, the limitation imposed on the conductivity is not only due to tunneling between islands, but also due to scattering of the momentum of charge carriers in islands from inhomogeneities.

## ACKNOWLEDGMENTS

We are grateful to N.N. Sibel'din and V.A. Chuenkov for numerous consultations and helpful advice.

This study was financed by the Russian Foundation for Basic Research, the Federal Target Program "Integration of Science and Higher Education in Russia," the Federal Target Science and Engineering Program of the Russian Federation "Complex Research Problems of the Physics of Solid Nanostructures," and the Program in Support Leading Scientific Schools of the Russian Federation (project no. NSh-1923.2003.2).

## REFERENCES

1. L. I. Maissel and R. Glang, *Handbook of Thin Film Technology* (McCraw-Hill, New York, 1970), p. 38.
2. R. D. Fedorovich, A. G. Naumovets, and P. M. Tomchuk, J. Phys.: Condens. Matter **11**, 9955 (1999).
3. A. P. Boltaev, N. A. Penin, A. O. Pogosov, and F. A. Pudonin, Zh. Éksp. Teor. Fiz. **123**, 1067 (2003) [JETP **96**, 940 (2003)].
4. V. M. Shalaev, Phys. Rep. **272**, 61 (1996).
5. B. I. Shklovskiĭ and A. L. Éfros, *Electronic Properties of Doped Semiconductors* (Nauka, Moscow, 1979; Springer, New York, 1984).
6. B. Abeles, Ping Sheng, M. D. Coutts, and Y. Arie, Adv. Phys. **24**, 407 (1975).
7. B. A. Aranzon, A. E. Varfolomeev, D. Yu. Kovalev, *et al.*, Fiz. Tverd. Tela (St. Petersburg) **41**, 944 (1999) [Phys. Solid State **41**, 857 (1999)].
8. D. A. Zakgeĭm, I. V. Rozhanskiĭ, I. P. Smirnova, and S. A. Gurevich, Zh. Éksp. Teor. Fiz. **118**, 637 (2000) [JETP **91**, 553 (2000)].
9. J. Klafter and Ping Sheng, J. Phys. C **17**, L93 (1984).
10. S. T. Chui, Phys. Rev. B **43**, 14 274 (1991).
11. E. Z. Meĭlikhov, Zh. Éksp. Teor. Fiz. **115**, 1484 (1999) [JETP **88**, 819 (1999)].
12. I. P. Zvaygin and R. Keiper, Phys. Status Solidi B **230**, 151 (2002).
13. N. Markovic, C. Christiansen, D. E. Grupp, *et al.*, Phys. Rev. B **62**, 2195 (2000).
14. C. A. Neugebauer and M. B. Webb, J. Appl. Phys. **33**, 74 (1962).
15. C. J. Adkins and E. G. Astrakharchik, J. Phys.: Condens. Matter **10**, 6651 (1998).
16. J. M. Rowell, in *Tunneling Phenomena in Solids*, Ed. by E. Burstein and S. Lundqvist (Plenum, New York, 1969; Mir, Moscow, 1973).
17. H. R. Zeller and I. Giaever, Phys. Rev. **181**, 789 (1969).
18. C. B. Duke, in *Tunneling Phenomena in Solids*, Ed. by E. Burstein and S. Lundqvist (Plenum, New York, 1969; Mir, Moscow, 1973).

*Translated by N. Wadhwa*

*Journal of Experimental and Theoretical Physics, Vol. 99, No. 4, 2004, pp. 834–849.*
*Translated from Zhurnal Éksperimental'noĭ i Teoreticheskoĭ Fiziki, Vol. 126, No. 4, 2004, pp. 962–978.*


**SOLIDS**
**Electronic Properties**

# Mesoscopic Mixing of Spin Orientation Phases in Particles with Ising Ions: Holmium:Yttrium Iron Garnets in Strong Magnetic Fields

**A. I. Popov, A. F. Popkov, and N. E. Kulagin**

*Moscow State Institute of Electronic Engineering (Technical University), Moscow, 124498 Russia*
*e-mail: popkov@nonlin.msk.ru*

Received April 6, 2004

**Abstract**—The paper presents a theoretical analysis of macroscopic quantum tunneling phenomena in small particles of a cubic ferromagnet of the $Ho_xY_{3-x}Fe_5O_{12}$ type with strongly anisotropic (Ising) impurity ions present in a low concentration $x \ll 1$ in the region of strong magnetic fields, at which many orientation phase transitions related to the competition of external and exchange field actions on the spin subsystem are observed. The theory of path integrals for the magnetic subsystem was used to calculate the instanton contributions to interphase tunneling amplitudes in the vicinity of first-order transitions for three principal orientations of an external magnetic field in a cubic crystal. It was shown that low-energy barriers separating angular phases could result in anomalously large mesoscopic volumes at which macroscopic spin tunneling effects could appear in the energy spectra of particles. The special features of spectral splitting caused by the mixing of azimuthally degenerate angular phases and phases with different polar angles of magnetization orientation were revealed.


## 1. INTRODUCTION

Macroscopic quantum coherence and macroscopic quantum tunneling of spins in magnetic mesoscopic systems predicted theoretically [1] and described in much detail in [2] were observed in experiments performed to study the frequency dependence of magnetic susceptibility and noise spectral response in small antiferromagnetic particles [3], the temperature dependence of magnetic relaxation frequencies in disperse and disordered magnetic media [4], and the relaxation of the transmission spectra of submillimeter waves in macromolecular spin clusters [5]. Macroscopic quantum effects are manifestations of the mesoscopic properties of magnetic systems and appear when there are several equilibrium positions, stable or metastable, separated by energy barriers $\Delta E$ whose height exceeds the characteristic energy $\hbar\omega$ of the quantum of magnetization oscillations near the equilibrium position by no more than one order of magnitude (usually, $\Delta E \sim 30\hbar\omega$ for ferromagnetic materials). Low-height energy barriers appear for spin modes of remagnetization of small-volume ferro- or antiferromagnetic particles [6], spins of natural magnetic defects such as magnetic vortices and domain boundaries in thin films and nanopartitions [7], spin clusters in molecular magnets [8], and in other small-volume magnetic formations. The essential difference between magnetic mesoscopic systems and nonmagnetic systems in energetically degenerate states is the dependence of the spin tunneling amplitude on the parity of their number, which is related to the Kramers theorem according to which degeneracy cannot be completely removed if the spin is half-integer [9]. This restriction is, however, is lifted in the presence of a dissipative environment, which creates dephasing of interfering tunneling amplitudes along topologically different instanton trajectories. Generally, spin-dependent macroscopic quantum coherence is easily destroyed in the presence of random spins in the system [10], for instance, in isotopes with magnetic nuclei [11], where spin transfer creates very strong fluctuations of the phase of tunneling instantons. In the absence of random spins in a magnetic system, the decay of the phase coherence of tunneling amplitudes can be caused by magnetoelastic coupling of spins with lattice vibrations. The decay of macroscopic quantum coherence may then be incomplete, the effect of the destructive interference of instantons that forbids spin tunneling is in part removed, and the effect of unquenching of macroscopic quantum coherence appears [12]. The tunnel coupling of states close in energy results in the pushing apart of levels and the splitting of degenerate states. The matrix element of energy splitting is then proportional to the transition probability amplitude. The influence of macroscopic quantum spin tunneling on the energy spectrum of a system can effectively be studied by the magnetoresonance and resonance spectroscopy methods. For instance, microwave spectroscopy was used in [5, 13] to investigate tunnel transitions and spin relaxation in the $Mn_{12}$ac molecular crystal.

1063-7761/04/9904-0834$26.00 © 2004 MAIK "Nauka/Interperiodica"

Complex-substituted rare-earth metal ferrite-garnets containing strongly anisotropic magnetic ions offer much promise as objects for studying mesoscopic manifestations of macroscopic spin tunneling in small particles. An important role in the physical properties of such materials is played by microscopic magnetic defects formed close to impurity rare-earth ion centers. These defects often play a determining role in static magnetization and magnetodynamic properties of bulk magnets [14]. An illustrative example of such materials is $Ho_xY_{3-x}Fe_5O_{12}$. Even at very low impurity Ising-type ion concentrations ($x \sim 0.001$), it exhibits magnetic phase transitions in magnetic fields comparable to the field $H \sim H_E$ of intersublattice exchange R–Fe and complex magnetic resonance behavior at low temperatures [15, 16]. At cryogenic temperatures determined by the ferromagnetic resonance frequency in an exchange field, that is, at

$$T < T^* = \hbar\gamma H_E/k_B,$$

small-volume particles with rare-earth ion impurities can exhibit cooperative quantum properties [17], which can in turn influence both magnetic saturation and phase formation and the special features of resonance properties in the region of intersection of the energy levels of impurity ions in a magnetic field. The exchange field that acts between the rare-earth metal sublattice and the iron sublattice is high ($H_E \sim 10^5$ Oe), and the expected crossover temperature in such materials can therefore be no lower than in antiferromagnets, of $T^* \sim 10$ K. A detailed analysis of macroscopic quantum phenomena during phase transitions in strong magnetic fields has not been performed thus far for materials of the specified type with rare-earth metal impurities, although the experimental data on magnetoresonance phenomena accompanying phase transitions in such materials have been reported [15].

In this work, we study macroscopic quantum tunneling and macroscopic quantum coherence effects in small particles of $Ho_xY_{3-x}Fe_5O_{12}$ cubic crystals with Ising impurity centers in a strong magnetic field oriented along the crystallographic directions $\mathbf{H} \parallel [100]$, [110], and [111]. An analysis will be performed for low holmium ion concentrations, $x \ll 1$, which are of special interest because they ensure the existence of low energy barriers between equilibrium positions. We calculated energy spectra for each field orientation taking into account the corresponding splitting close to first-order phase transition lines; these transitions can be caused by macroscopic quantum coherence effects. A detailed analysis can be very useful for performing experiments with recording the magnetic resonance spectra of such materials and in studies of spin-orientation phase transitions in strong magnetic fields.

## 2. LAGRANGIAN, PHASE TRANSITIONS, AND SPIN TUNNELING AMPLITUDES FOR HOMOGENEOUSLY MAGNETIZED $Ho_xY_{3-x}Fe_5O_{12}$ PARTICLES IN A STRONG FIELD

Yttrium:holmium iron garnet $Ho_xY_{3-x}Fe_5O_{12}$, whose unit cell contains eight formula units, has cubic symmetry; its space group is $O_h^{10}$, point group $m3m$, and cell size $a = 1.24$ nm. The rare-earth cations and yttrium ions occupy $c$ sites characterized by a dodecahedral anionic (oxygen) environment. The cell has six nonequivalent sites that can be occupied by impurity ions. These sites differ in the orientation of local symmetry axes of the environment and, as a consequence, the direction of local magnetization of strongly anisotropic holmium ions, which have the properties of pseudo-Ising ions. The rare-earth ions are in a strong exchange field generated by the ferromagnetic matrix of iron ions, which form two sublattices with octahedral ($a$ sublattice) and tetrahedral ($d$ sublattice) environments. Because of the strong intra- and intersublattice exchange interaction, two iron sublattices can be treated as a unique ferromagnetic sublattice with magnetization $M_{\mathrm{Fe}}$ up to fields that far exceed the antiferromagnetic exchange field

$$H_E = \lambda M_{\mathrm{Fe}},$$

which acts between the rare-earth metal ion spins and the iron ions. On these assumptions and with the inclusion of only the two lowest rare-earth ion levels, which is warranted at fairly low temperatures, the thermodynamic potential of the system under consideration can be written as [16]

$$E = A\sum_{k=x,y,z}\left(\frac{\nabla M_k}{M_{\mathrm{Fe}}}\right)^2 - \mathbf{M}_{\mathrm{Fe}}\cdot\mathbf{H} - \frac{x_c k_B T}{3} \times \sum_{i=1}^{3} \ln 2\cosh\frac{\mu|\mathbf{H}-\lambda\mathbf{M}_{\mathrm{Fe}}|_{z_i}}{k_B T}. \quad (1)$$

Here, index $i$ denotes actual nonequivalent positions of Ising ions, $z_i$ is the projection of the effective magnetic field onto the local magnetization axis $\mathbf{z}_i$ (in holmium:yttrium iron garnets, the local magnetization axes of holmium ions $\mathbf{z}_i$ ($i = 1, 2, 3$) coincide with the crystallographic axes [001], [010], and [100]), $\mu$ is the magnetic moment of the Ising ion, $x_c = c_0 x$ is the volume concentration of the impurity, $c_0 = 8/a^3$ is the number of moles in unit volume, and $A$ is the inhomogeneous exchange constant. In the polar coordinates, the Lagrangian of the system can be written as

$$L = \int_V \left[\frac{M}{\gamma}(1-\cos\theta)\frac{\partial\varphi}{\partial t} - E(\theta,\varphi)\right]d\mathbf{r}, \quad (2)$$

where $\gamma$ is the magnetomechanical ratio for the ferromagnetic sublattice and $\theta$ and $\varphi$ are the polar and azimuthal angles that determine the orientation of $\mathbf{M}_{\mathrm{Fe}}$. Accordingly, the action over the time interval $t_S$ is given by the equation

$$S_E = \int_0^{t_S} L\,dt = \int_0^{t_S} dt \int_V \left[\frac{M}{\gamma}(1-\cos\theta)\frac{\partial\varphi}{\partial t} - E(\theta,\varphi)\right] d\mathbf{r}. \quad (3)$$

The first term in Lagrangian (2) corresponds to the topological contribution to action (3) in the phase space of the dynamic system (the Berry phase) [9] responsible for the appearance of a phase shift in the probability amplitudes calculated below. The magnetization of Ising ions, which follows the projection of the effective magnetic field created by the iron sublattice and an external magnetic field onto the Ising axis, makes no topological contribution [18].

The tunneling probability amplitude can be found using the approach based on the theory of path integrals in the phase space of the dynamic system in imaginary time [1, 19]. We limit our consideration to instanton trajectories $\mathbf{\Omega} = (p, q)$ that make a determining contribution to the tunneling probability amplitude,

$$K_{a\to b} = \int_a^b d[\mathbf{\Omega}] \exp\left(\frac{iS_E(\mathbf{\Omega})}{\hbar}\right). \quad (4)$$

The tunneling probability amplitude is then given by the equation

$$\begin{aligned} K_{t\to t+T} &= A_K \exp(iS_E(\mathbf{\Omega}_{cl})/\hbar) \\ &= A_K \exp(-B_1 + iB_2), \end{aligned} \quad (5)$$

where

$$B_1 = \mathrm{Im} S_E(\mathbf{\Omega}_{cl})$$

determines the exponential factor of the tunneling amplitude (the Gamow factor). This factor is proportional to the imaginary action part along the instanton trajectory $\mathbf{\Omega}_{cl}(\tau)$. The action has an extremum on this trajectory, which reduces its first variation to zero; that is,

$$\delta S_E(\mathbf{\Omega}) = 0.$$

The probability amplitude phase is proportional to the real part of the action, that is,

$$B_2 = \mathrm{Re} S_E(\mathbf{\Omega}_{cl}).$$

The preexponential factor $A_K$ in the equation for the tunneling amplitude is determined by the spectrum of small fluctuation excitations in the vicinity of the instanton trajectory. We omit its calculations and only note that, in the order of magnitude, it is determined by the equation [20]

$$A_K \approx \omega_0 \sqrt{\frac{B_1}{\hbar}}, \quad (6)$$

where $\omega_0$ is the resonance oscillation frequency close to the initial equilibrium position.

## 3. MAGNETIC PHASE TRANSITIONS IN HOLMIUM:YTTRIUM IRON GARNETS

Let us briefly summarize information about magnetic phase transitions in holmium:yttrium iron garnets and the corresponding phase states of the spin subsystem necessary for the further analysis. Of greatest interest to us are low temperatures. Ignoring inhomogeneous exchange [the first term in (1)], the energy of the ground state of the system at $T = 0$ K [potential (1)] can be written in the normalized variables

$$\varepsilon = \frac{E}{\lambda M_{\mathrm{Fe}}^2}, \quad \mathbf{h} = \frac{\mathbf{H}}{\lambda M_{\mathrm{Fe}}}, \quad l = \frac{\mathbf{M}_{\mathrm{Fe}}}{M_{\mathrm{Fe}}}$$

as

$$\varepsilon = -l(\mathbf{h} - \mathbf{m}) - \mathbf{h}\cdot\mathbf{m}, \quad (7)$$

where the vector $\mathbf{m}$ is the total moment of the $Ho^{3+}$ ions in arbitrary units normalized with respect to $\mathbf{M}_{\mathrm{Fe}}$; this vector can only be directed along axes of the type [111]. The modulus of this vector

$$m = x_c \mu/\sqrt{3} M_{\mathrm{Fe}}$$

can be treated as the relative concentration of $Ho^{3+}$ ions. Magnetic phase transitions can only occur in the system as discrete reorientations of the vector $\mathbf{m}$ from one direction of the type [111] to another. The minimization of energy (7) determines the equilibrium phase states of the system. A detailed description of the main characteristics of the phase diagrams for the cases under consideration can be found, e.g., in [16].

The classification and denotations of the phases considered below for the field orientations $\mathbf{H} \parallel [001]$, $\mathbf{H} \parallel [110]$, and $\mathbf{H} \parallel [111]$ and the equations for the fields of orientation phase transitions at $m \ll 1$ are as follows.

$\mathbf{H} \parallel [001]$. Here, fourfold degenerate phases that differ in the projection of $\mathbf{m}$ onto $\mathbf{H}$ can exist:

$$\begin{aligned} &\text{phase I—}\mathbf{m} \parallel [\bar{1}\bar{1}\bar{1}],\quad [1\bar{1}\bar{1}],\quad [\bar{1}1\bar{1}],\quad [11\bar{1}], \\ &\text{phase II—}\mathbf{m} \parallel [\bar{1}11],\quad [\bar{1}11],\quad [1\bar{1}1],\quad [\bar{1}\bar{1}1]. \end{aligned} \quad (8)$$

The phase transition line separating these phases is

given by the equation

$$h_{\mathrm{I,\,II}} = 1 - \frac{m^2}{3}. \tag{9}$$

**H** ∥ [110]. In this case, three phases exist,

$$\begin{gathered}\text{phase I—}\mathbf{m} \parallel [\bar{1}\bar{1}\bar{1}],\quad [\bar{1}\bar{1}1],\\ \text{phase II—}\mathbf{m} \parallel [\bar{1}11],\quad [1\bar{1}1],\quad [\bar{1}1\bar{1}],\quad [1\bar{1}\bar{1}],\\ \text{phase III—}\mathbf{m} \parallel [111],\quad [11\bar{1}].\end{gathered} \tag{10}$$

The first-order phase transition lines for the corresponding phases have the form

$$\begin{aligned}h_{\mathrm{I,\,II}} &= 1 - \frac{m}{\sqrt{6}},\\ h_{\mathrm{II,\,III}} &= 1 + \frac{m}{\sqrt{6}}.\end{aligned} \tag{11}$$

**H** ∥ [111]. There exist four phases,

$$\begin{gathered}\text{phase I—}\mathbf{m} \parallel [\bar{1}\bar{1}\bar{1}],\\ \text{phase II—}\mathbf{m} \parallel [\bar{1}\bar{1}1],\quad [\bar{1}1\bar{1}],\quad [1\bar{1}\bar{1}],\\ \text{phase III—}\mathbf{m} \parallel [\bar{1}11],\quad [1\bar{1}1],\quad [11\bar{1}],\\ \text{phase IV—}\mathbf{m} \parallel [111].\end{gathered} \tag{12}$$

The corresponding first-order phase transition lines are

$$\begin{aligned}h_{\mathrm{I,\,II}} &= 1 - \frac{2}{3}m,\\ h_{\mathrm{II,\,III}} &= 1,\\ h_{\mathrm{III,\,IV}} &= 1 + \frac{2}{3}m.\end{aligned} \tag{13}$$

## 4. MESOSCOPIC SPIN OSCILLATIONS IN SMALL $Ho_xY_{3-x}Fe_5O_{12}$ PARTICLES AT **H** ∥ [100]

Let us consider small particles of volume

$$v_0 < (A/\lambda M_{\mathrm{Fe}}^2)^{3/2},$$

within which magnetization can be considered homogeneous in the external field **H** ∥ [100] and the first term in the equation for the energy [Eq. (1)] can be ignored. The polar axis is aligned with the axis [100], and the azimuthal angle is counted from the crystal direction [010]. The Lagrangian of unit volume then takes the form

$$\begin{aligned}L &= \frac{M_{\mathrm{Fe}}}{\gamma}(1-\cos\theta)\dot{\phi} + \lambda M_{\mathrm{Fe}}^2 h\cos\theta\\ &+ \tau\tilde{x}\ln\left[\cosh\left(\frac{\sin\theta\cos\varphi}{\tau}\right)\cosh\left(\frac{\sin\theta\sin\varphi}{\tau}\right)\right.\\ &\left.\times\cosh\left(\frac{h-\cos\theta}{\tau}\right)\right],\end{aligned} \tag{14}$$

where

$$\tilde{x} = \frac{1}{3}\lambda\mu M_{\mathrm{Fe}}x_c,\quad \tau = \frac{k_BT}{\lambda\mu M_{\mathrm{Fe}}},\quad h = \frac{H}{\lambda M_{\mathrm{Fe}}},$$

and the action is

$$S_E = \int L dt dv = v_0\int L dt. \tag{15}$$

At low temperatures $T \longrightarrow 0$ and low concentrations $x \ll 1$ (when

$$m = \frac{x_c\mu}{\sqrt{3}M_{\mathrm{Fe}}} \ll 1$$

is a small parameter), the first-order phase transition line

$$h \approx 1 - m^2/3,\quad 0 < \tau < 2m^3/3\sqrt{3}$$

appears on the $h\tau$ plane in high fields $H \sim \lambda M_{\mathrm{Fe}}$ [see (9)]. The energies of two phases fourfold degenerate with respect to the azimuthal angle

$$\varphi = \frac{\pi}{4}(2k+1),\quad k = 0, 1, 2, 3,$$

become equal on this line.

At small polar angles of deviations of $\mathbf{M}_{\mathrm{Fe}}$ from the direction of field **H**, which are always small if fields are high and concentrations $x$ low, we can, after the introduction of the new variables

$$(p, q) = (\sqrt{3}/m)(\theta\cos\varphi, \theta\sin\varphi),\quad \tilde{t} = \omega_0 t,$$

where

$$\omega_0 = \gamma\lambda M_{\mathrm{Fe}},\quad \tilde{\tau} = (\sqrt{3}/m)\tau,$$

rewrite action (15) in the normalized form

$$S_E = \frac{\mu^2 x_c^2 v_0}{9\gamma M_{\mathrm{Fe}}} f(\tilde{\tau}).$$

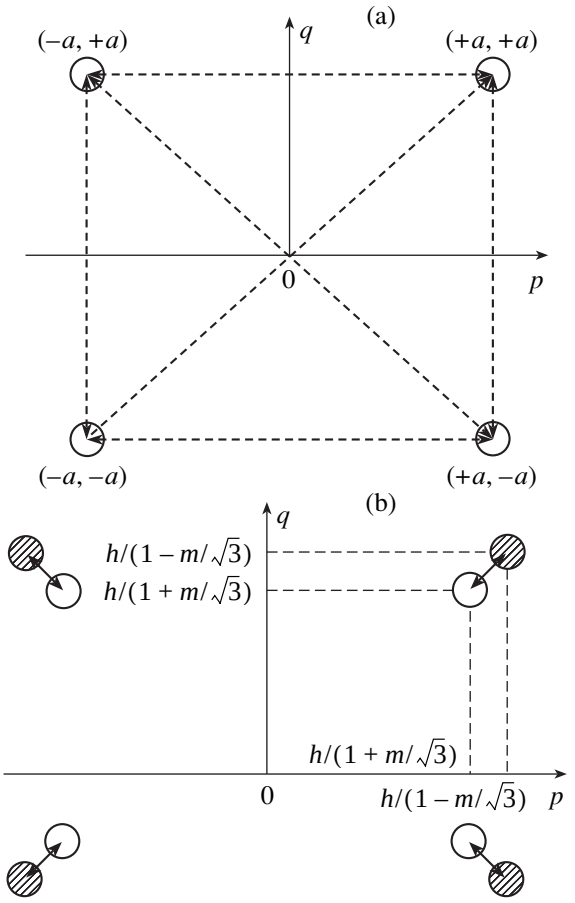



**Fig. 1.** Equilibrium points in the *pq* plane at **H** ∥ [100]: (a) instanton tunneling trajectories in phase I and (b) instanton tunneling trajectories of the I ⟶ II phase transition; open circles are phase I and hatched circles are phase II.

Here,

$$f(\tilde{\tau}) = \int d\tilde{t}\left\{\frac{1}{2}(\dot{q}p - \dot{p}q) - \tilde{H}(q, p)\right\} \tag{16}$$

is the normalized action, which contains the effective Hamiltonian

$$\tilde{H}(q, p) = \frac{h}{2}(p^2 + q^2) - \tilde{\tau}\ln\left(\cosh\frac{p}{\tilde{\tau}}\cosh\frac{q}{\tilde{\tau}} \times \cosh\left(\frac{\sqrt{3}(1-h)}{m\tilde{\tau}} - \frac{m(p^2+q^2)}{2\sqrt{3}\tilde{\tau}}\right)\right) - H(p_0, q_0), \tag{17}$$

where the constant $H(p_0, q_0)$ is the reference level for counting energy from the equilibrium state $(p_0, q_0)$. Considering that $m \ll 1$, we can ignore small corrections related to this parameter at high fields close to the phase transition line $h \sim 1$ in Hamiltonian (17). This allows us to use a simplified Hamiltonian that describes the effect of fourfold degeneracy in phases (8) on the assumption that the energy and polar angle directions in these phases are approximately equal; that is,

$$\tilde{H}(q, p) \approx \frac{1}{2}(p^2 + q^2) - \tilde{\tau}\ln\left(\cosh\frac{p}{\tilde{\tau}}\cosh\frac{q}{\tilde{\tau}}\right) - H(p_0, q_0). \tag{18}$$

We use this Hamiltonian to illustrate the method for finding spin tunneling amplitudes between equilibrium degenerate states in the limiting case of zero temperature $\tilde{\tau} \longrightarrow 0$.

The Hamilton equations

$$\dot{q} = \frac{\partial\tilde{H}(p, q)}{\partial p}, \quad \dot{p} = -\frac{\partial\tilde{H}(p, q)}{\partial q} \tag{19}$$

can be written in the form

$$\dot{q} = p - \tanh\frac{p}{\tilde{\tau}}, \quad \dot{p} = -q + \tanh\frac{p}{\tilde{\tau}}. \tag{20}$$

There are four energetically equivalent equilibrium points,

$$(q_0, p_0) = (a, \pm a); \quad (-a, \pm a),$$

where $a$ is the positive root of the transcendental equation

$$\tanh\frac{a}{\tilde{\tau}} = a.$$

These points are shown in Fig. 1 in the *pq* plane. Spontaneous transitions related to macroscopic spin tunneling (macroscopic quantum tunneling) can occur between these equilibrium points.

The tunneling probability amplitude between the equilibrium states of the system will be found by means of the passage to imaginary time $\tau = it$ and the complex variables

$$p = p_1 + ip_2, \quad q = q_1 + iq_2.$$

The initial system of equations written in these variables takes the form

$$\dot{q}_1 = -p_2 + \frac{\sin\frac{p_2}{\tilde{\tau}}\cos\frac{p_2}{\tilde{\tau}}}{\cosh^2\frac{p_1}{\tilde{\tau}} - \sin^2\frac{p_2}{\tilde{\tau}}},$$

$$\dot{q}_2 = p_1 - \frac{\sinh\frac{p_1}{\tilde{\tau}}\cosh\frac{p_1}{\tilde{\tau}}}{\cosh^2\frac{p_1}{\tilde{\tau}} - \sin^2\frac{p_2}{\tilde{\tau}}}, \tag{21}$$

$$\dot{p}_1 = q_2 - \frac{\sin\frac{q_2}{\tilde{\tau}}\cos\frac{q_2}{\tilde{\tau}}}{\cosh^2\frac{q_1}{\tilde{\tau}} - \sin^2\frac{q_2}{\tilde{\tau}}},$$

$$\dot{p}_2 = -q_1 + \frac{\sinh\frac{q_1}{\tilde{\tau}}\cosh\frac{q_1}{\tilde{\tau}}}{\cosh^2\frac{q_1}{\tilde{\tau}} - \sin^2\frac{q_2}{\tilde{\tau}}}.$$

In the complex variables, Hamiltonian (18) becomes

$$\begin{aligned} H_1 &= \mathrm{Re}\tilde{H}(q,p) = \frac{1}{2}(p_1^2 - p_2^2 + q_1^2 - q_2^2) \\ &- \frac{\tilde{\tau}}{2}\ln\left(\cosh^2\frac{p_1}{\tilde{\tau}}\cos^2\frac{p_2}{\tilde{\tau}} + \sinh^2\frac{p_1}{\tilde{\tau}}\sin^2\frac{p_2}{\tilde{\tau}}\right) \\ &- \frac{\tilde{\tau}}{2}\ln\left(\cosh^2\frac{q_1}{\tilde{\tau}}\cos^2\frac{q_2}{\tilde{\tau}} + \sinh^2\frac{q_1}{\tilde{\tau}}\sin^2\frac{q_2}{\tilde{\tau}}\right) \\ &- a^2 + 2\tilde{\tau}\ln\left(\cosh\frac{a}{\tilde{\tau}}\right), \end{aligned} \tag{22}$$

$$\begin{aligned} H_2 &= \mathrm{Im}\tilde{H}(q,p) \\ &= p_1p_2 + q_1q_2 + \arctan\left(\tanh\frac{p_1}{\tilde{\tau}}\tan\frac{p_2}{\tilde{\tau}}\right) \\ &+ \arctan\left(\tanh\frac{q_1}{\tilde{\tau}}\tan\frac{q_2}{\tilde{\tau}}\right). \end{aligned} \tag{23}$$

We therefore obtain the complex Hamiltonian

$$H(p,q) = H_1(p_1, p_2, q_1, q_2) + iH_2(p_1, p_2, q_1, q_2), \tag{24}$$

analytic with respect to each of its complex variables $p_k$, $q_k$, $k = 1, 2$. Here, all the functions ($p_k$, $q_k$, and $H_k$, $k = 1, 2$) are real functions of their real arguments ("time" $\tau$ is also real).

Partitioning the real and imaginary parts in the Hamilton equations yields the real dynamic system

$$\begin{aligned} \frac{dp_1}{d\tau} &= -\frac{\partial H_1}{\partial q_1}, \quad \frac{dp_2}{d\tau} = -\frac{\partial H_2}{\partial q_1} = \frac{\partial H_1}{\partial q_2}, \\ \frac{dq_1}{d\tau} &= \frac{\partial H_1}{\partial p_1}, \quad \frac{dq_2}{d\tau} = \frac{\partial H_2}{\partial p_1} = -\frac{\partial H_1}{\partial p_2}, \end{aligned} \tag{25}$$

which is Hamiltonian in the variables

$$\mathbf{p} = (p_1, q_2), \quad \mathbf{q} = (q_1, p_2)$$

with Hamiltonian $H_1$ and additional integral $H_2$. The system of equations obtained this way is integrable.

Instanton trajectories are found by numerically integrating the equations by the Runge–Kutta method using the asymptotic behavior of the solutions close to the equilibrium positions,

$$\tilde{q} - q_0 = \delta q_1 + i\delta q_2 = \delta\exp(i\varphi + \lambda t),$$

$$\tilde{p} - p_0 = \delta p_1 + i\delta p_2 = \delta\exp(i(\varphi - \pi/2) + \lambda t),$$

where $\delta$ and $\varphi$ are the arbitrary amplitude and phase of the asymptotic linear solution and

$$\lambda = \pm\left(1 - \frac{1}{\tilde{\tau}\cosh^2\frac{a}{\tilde{\tau}}}\right)$$

are the roots of the characteristic equation of linearized system (21).

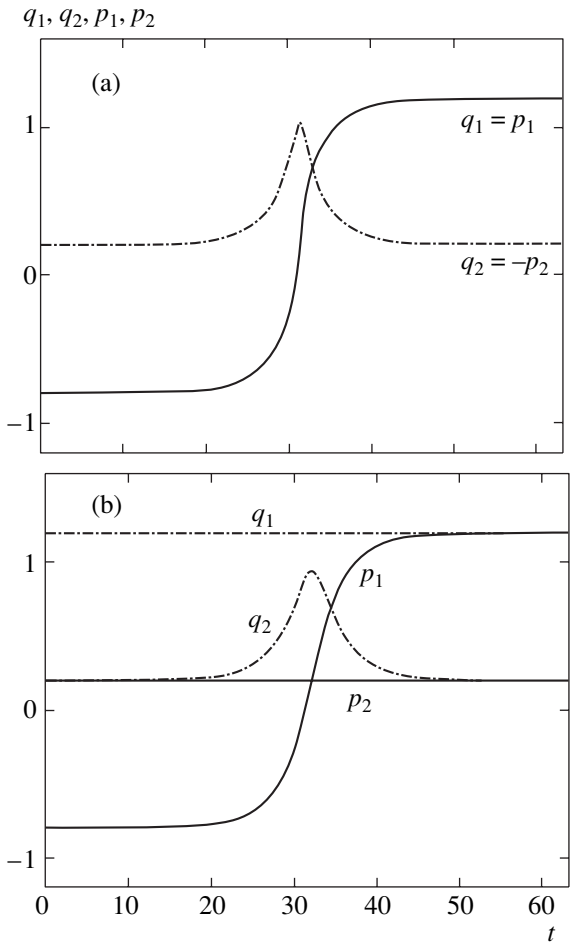



**Fig. 2.** Examples of instanton solutions at $\tilde{\tau} = \tanh 3/3$, $a = \tanh 3$ between the points (a) $(q, p) = (+a, -a)$ and $(q, p) = (-a, +a)$ and (b) $(q, p) = (+a, -a)$ and $(q, p) = (+a, +a)$. Phase variables $q_1$, $q_2$, $p_1$, and $p_2$ and time $t$ are normalized with respect to $\tilde{\tau}$.

The system under consideration has two groups of instanton solutions that connect equilibrium points situated on the diagonals and sides of a quadrangle. The obtained numerical instanton solution for the transition between the points $(q_-, p_-) = (-a, -a)$ and $(q_+, p_+) = (+a, +a)$ is shown in Fig. 2a. The instanton solution for the transition between the points $(q_-, p_-) = (+a, -a)$ and $(q_+, p_+) = (+a, +a)$ (Fig. 2b) is found in a similar way.

The behavior of the instanton trajectories when temperature varies is shown in Fig. 3. To find the trajectory

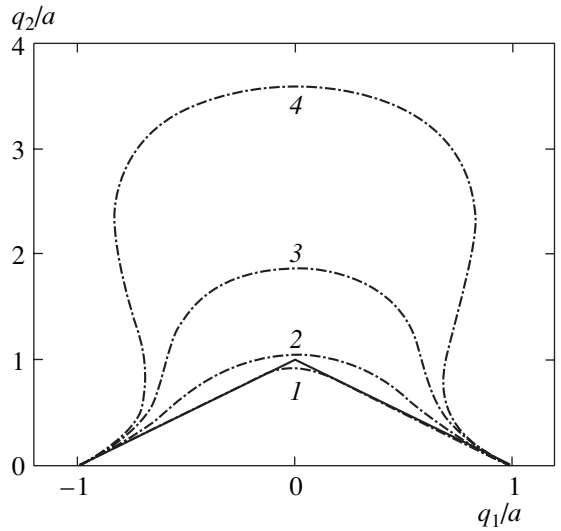



**Fig. 3.** Projection of normalized instanton trajectories (dash-and-dot lines) between the points $(q, p) = \pm(a, a)$ onto the phase plane $q_1q_2$ at various temperature parameter values $\tilde{\tau}$ = (*1*) 0.167, (*2*) 0.482, (*3*) 0.763, and (*4*) 0.924. The solid line is the projection of the limiting trajectory $q_2 = 1 - |q_1|$ that describes the instanton in the limit $\tilde{\tau} = 0$.

in the zero temperature limit, when $\tilde{\tau} \longrightarrow 0$ and $a \longrightarrow 1$, we must again consider the tunnel transitions between the points $(q_-, p_-) = (-a, -a)$ and $(q_+, p_+) = (a, a)$. In conformity with the numerical solutions already found, we will seek a trajectory with the symmetry property $(q_1, q_2) = (p_1, -p_2)$. On this trajectory, the second integral $H_2(q_1, q_2, p_1, p_2)$ is identically equal to zero. The vanishing of the first integral $H_1(q_1, q_2, p_1, p_2) = 0$ implicitly describes the trajectory of motion. Taking into account the selected symmetry, we obtain

$$H_1(q_1, q_2, p_1, p_2) = (1 - |q_1|)^2 - p_2^2 = 0$$

in the limit $\tilde{\tau} \longrightarrow 0$, $a \longrightarrow 1$.

It follows that

$$p_1 = q_1, \quad p_2 = -q_2 = \pm(1 - |q_1|). \tag{26}$$

The above numerical solution of the corresponding type for $\tilde{\tau} \neq 0$ transforms into asymptotic solution (26) in the phase space of the system in the limit $\tilde{\tau} \longrightarrow 0$, $a \longrightarrow 1$. This tendency is shown in Fig. 3, which contains instanton trajectories on the $q_2q_1$ plane for various $\tilde{\tau}$ parameter values. The trajectories are normalized with respect to the equilibrium value

$$|p_0| = |q_0| = a(\tilde{\tau}).$$

The tunneling probability amplitude is determined by (5), where the exponential factor is found as

$$-B_1 + iB_2 = \frac{iS_{\text{inst}}}{\hbar} = B_0 \tilde{f}_{\text{inst}}(\tilde{\tau}). \tag{27}$$

Here,

$$B_0 = \frac{\mu^2 x_c^2 \nu_0}{9\hbar\gamma M}$$

is the normalizing factor. The normalized exponential factor is determined at zero temperature as the limit

$$\alpha_{\text{inst}}(0) = \lim_{\tau \to 0} \tilde{f}_{\text{inst}}(\tilde{\tau}).$$

As the instanton trajectory corresponds to the $H(q, p) = 0$ zero level of the Hamiltonian, the normalized action along the instanton trajectory can be written as

$$\tilde{f}_{\text{inst}}(\tilde{\tau}) = \frac{i}{2}\int_{-\infty}^{\infty} (\dot{\tilde{q}}\tilde{p} - \dot{\tilde{p}}\tilde{q}) d\tilde{t}$$

$$= \frac{1}{2}\int_{-\infty}^{\infty} \{(-\dot{q}_1 p_2 + q_1\dot{p}_2 - \dot{q}_2 p_1 + q_2\dot{p}_1)$$

$$+ i(-q_1\dot{p}_1 + \dot{q}_1 p_1 - \dot{q}_2 p_2 + q_2\dot{p}_2)\} d\tilde{t} \tag{28}$$

$$= \frac{1}{2}\int_{q_a}^{q_b} \left\{\left(-p_2 + q_1\frac{\partial p_2}{\partial q_1} - p_1\frac{\partial q_2}{\partial q_1} + q_2\frac{\partial p_1}{\partial q_1}\right)\right.$$

$$\left. + i\left(-q_1\frac{\partial p_1}{\partial q_1} + p_1 - \frac{\partial q_2}{\partial q_1}p_2 + q_2\frac{\partial p_2}{\partial q_1}\right)\right\} dq_1.$$

After the integration of (28), we obtain

$$\alpha_{\text{inst}} = \lim_{\tilde{\tau} \to 0} \tilde{f}(\tilde{\tau}) = \alpha_1 + i\alpha_2 = -2 \tag{29}$$

for the case under consideration.

Next, let us consider the transition between the points $(q_-, p_-) = (a, -a)$ and $(q_+, p_+) = a, a$. For this transition, the second integral $H_2(q_1, q_2, p_1, p_2)$ vanishes in the limit $\tilde{\tau} \longrightarrow 0$, $a \longrightarrow 1$ at $q_1 = a$, $p_2 = 0$, whereas the first integral gives $q_2 = \pm(1 - |p_1|)$. The integration of (28) therefore yields

$$\alpha_{\text{inst}} = \lim_{\tilde{\tau} \to 0} \tilde{f}(\tilde{\tau}) = \alpha_1 + i\alpha_2 = -1 + i \tag{30}$$

for the limiting trajectory under consideration. A comparison of (29) and (30) shows that the tunneling frequency between neighboring energetically degenerate equilibrium positions is much higher than for the instanton transition between opposite states. A similar procedure can be used to find the other possible instan-

ton trajectories and their contributions to the tunneling probability at zero temperature.

We will now consider the case of a finite value of the small parameter $m$, when there are two fourfold degenerate phases (8) described by Hamiltonian (17) with different polar orientations of the magnetization vector of the iron sublattice, in the limit of low temperatures $\tilde{\tau} \longrightarrow 0$. These phases are defined as follows.

Phase I: $|p_0| = |q_0| = 1/(h+r)$, where $r = m/\sqrt{3}$. This phase exists in the fields

$$h \le 1 - \frac{r^2}{(1+r)^2}.$$

The dynamics of this phase is described by the Hamiltonian

$$\tilde{H}(q, p) = \frac{h+r}{2}(p^2+q^2) - |p| - |q| - \frac{1-h}{r} - H(p_0, q_0). \quad (31)$$

Phase II: $|p_0| = |q_0| = 1/(h-r)$. This phase exists in the fields

$$h \ge 1 - \frac{r^2}{(1-r^2)}.$$

The corresponding Hamiltonian is

$$\tilde{H}(q, p) = \frac{h-r}{2}(p^2+q^2) - |p| - |q| + \frac{1-h}{r} - H(p_0, q_0). \quad (32)$$

The energies of the two phases are compared at the first-order phase transition point [see (9)], approximately in the field

$$h = h^* \approx 1 - r^2.$$

We will consider the transition probability between the nearest equilibrium points of these two phases

$$p_0^{\mathrm{I}} = q_0^{\mathrm{I}} = \frac{1}{h+r}, \quad p_0^{\mathrm{II}} = q_0^{\mathrm{II}} = \frac{1}{h-r}$$

(see Fig. 1b) at the field value at which they become energetically degenerate, that is, at $h = h^*$. The instanton solution can be found from the condition that Hamiltonians (31) and (32) should remain at the zeroth level taking into account that

$$H(p_0, q_0) = -\frac{1}{h^*-r} + \frac{1-h^*}{r} = -\frac{1}{h^*+r} - \frac{1-h^*}{r}.$$

The particular constant instanton solution

$$p = p_1 + ip_2, \quad q = p_1 - ip_2,$$

will be sought using Hamiltonian (17) at

$$\frac{1}{h+r} < p_1 < p^*$$

for phase I and Hamiltonian (17) at

$$p^* < p_1 < \frac{1}{h-r}$$

for phase II. The $p^*$ parameter is determined by the condition of continuity of the variable $p_2(p^*)$ for the angular phases under consideration in the complex phase space of dynamic system (21).

The imaginary part of the complex Hamiltonian identically equals zero in the limit $\tilde{\tau} \longrightarrow 0$ for both angular phases I and II; that is,

$$H_2(p_1, p_2) = 0.$$

The zeroth level of the real part of the Hamiltonian for phase I is

$$H_1(p_1, p_2) = (h+r)\left(\left(p_1 - \frac{1}{h+r}\right)^2 - p_2^2\right) = 0, \quad (33)$$

which yields

$$|p_2| = p_1 - \frac{1}{h+r} \text{ for } \frac{1}{h+r} < p_1 < p^*.$$

For phase II, we find in a similar way that

$$H_1(p_1, p_2) = (h-r)\left(\left(p_1 - \frac{1}{h-r}\right)^2 - p_2^2\right) = 0, \quad (34)$$

which yields

$$|p_2| = \frac{1}{h-r} - p_1 \text{ for } p^* < p_1 < \frac{1}{h-r}.$$

The condition of sewing together these two solutions at the point $p_1 = p^*$ yields

$$p^* = \frac{1}{2}\left(\frac{1}{h-r} + \frac{1}{h+r}\right) = \frac{h}{h^2-r^2}, \quad p_2^* = \frac{1}{2}\left(\frac{1}{h-r} - \frac{1}{h+r}\right) = \frac{r}{h^2-r^2}. \quad (35)$$

Using this instanton solution for the zeroth Hamiltonian level and (28), we find

$$\alpha_{\mathrm{inst}} = \lim_{\tilde{\tau}\to 0} \tilde{f}(\tilde{\tau}) = -\frac{r^2}{2(h^{*2}-r^2)^2} \approx -\frac{r^2}{2} \quad (36)$$

for the case under consideration. The normalized action between two angular phases I and II [see (8)] that we found is much smaller than the tunneling factors obtained above for the orientation macroscopic quantum coherence transitions within each phase. This is

explained by a much lower energy barrier between phases I and II compared with the barrier for an azimuthal change in the orientation of the magnetization of the iron sublattice within one phase. As a consequence, intraphase tunneling effects become strongly suppressed as the mesoscopic volume increases. Low energy barriers substantially increase the mesoscopic volume in which the quantum properties of magnetization oscillations close to the first-order phase transition line can manifest themselves. On the other hand, transitions between azimuthal equilibrium directions of the magnetization of the iron sublattice can become essential to macroscopic quantum coherence as the volume of particles decreases, the separating barrier for the interphase transition I–II becomes smaller than the oscillation quantum $\Delta E \ll \hbar\omega_0$, and the iron sublattice ceases to distinguish between phases I and II because of strong quantum fluctuations. The barrier to the transition I–II is suppressed much more strongly as the temperature increases than intraphase barriers that determine the stability of these phases with respect to the other possible transitions.

## 5. ENERGY SPLITTING OF THE GROUND STATE AT **H** ∥ [100] TAKING INTO ACCOUNT MACROSCOPIC QUANTUM COHERENCE SPIN OSCILLATIONS

The special features of the splitting of the ground state of a system in the presence of energy degeneracy are determined by solving the problem of Hamiltonian diagonalization taking into account the matrix elements of tunneling. If fourfold azimuthal degeneracy is unimportant for macroscopic quantum coherence when states I and II are mixed, the splitting of the ground state occurs trivially, into two levels with the tunnel transition width determined by Gamow factor (36).

However if azimuthal macroscopic quantum coherence transitions acquire importance, the ground state of the mesoscopic system under consideration with four degenerate equilibrium positions is found by solving the eigenvalue problem for the mixed wavefunction

$$\Psi = c_1\phi_1 + c_2\phi_2 + c_3\phi_3 + c_4\phi_4.$$

The matrix elements of the Hamiltonian to be diagonalized are determined by the tunneling amplitudes

$$\Delta_m \approx A_m \exp(-B_{m1} + iB_{m2}), \quad m = 1, 2,$$

between the corresponding equilibrium positions. We can therefore write

$$\begin{pmatrix} -E & \Delta_1 & \Delta_2 & \Delta_1^* \\ \Delta_1^* & -E & \Delta_1 & \Delta_2 \\ \Delta_2 & \Delta_1^* & -E & \Delta_1 \\ \Delta_1 & \Delta_2 & \Delta_1^* & -E \end{pmatrix} \begin{pmatrix} c_1 \\ c_2 \\ c_3 \\ c_4 \end{pmatrix} = 0. \tag{37}$$

The energy eigenvalues are found by solving the equation

$$\begin{aligned} E^4 - 2E^2(\Delta_2^2 + 2|\Delta_1|^2) - 4E\Delta_2(\Delta_1^2 + \Delta_1^{*2}) \\ - 4\Delta_2^2|\Delta_1|^2 - (\Delta_1^2 - \Delta_1^{*2})^2 = 0. \end{aligned} \tag{38}$$

We see that, in the presence of the destructive interference of the tunneling amplitudes, that is, if

$$\Delta_1^2 + \Delta_1^{*2} = 0,$$

pairwise degeneracy of energy levels appears,

$$E_\pm = \pm\sqrt{\Delta_2^2 + 2|\Delta_1|^2}.$$

Generally, this equation has four different roots.

## 6. PHASE TRANSITIONS AND SPIN TUNNELING AT **H** ∥ [110]

Next, let us consider a small $Ho_xY_{3-x}Fe_5O_{12}$ particle of volume $v_0$ in field **H** ∥ [110]. The Lagrangian of unit volume in the coordinate system $[\bar{1}10]$, [001], [110] then has the form [16]

$$\begin{aligned} L = \frac{M_{\mathrm{Fe}}}{\gamma}(1 - \cos\theta)\dot{\phi} + M_{\mathrm{Fe}}H\cos\theta \\ + \tau\tilde{x}\ln\left\{\cosh\left(\frac{\sin\theta\sin\varphi}{\tau}\right)\right\} \\ + \tau\tilde{x}\ln\left[\cosh\left(\frac{\sqrt{2}(h - \cos\theta)}{\tau}\right) + \cosh\left(\frac{\sqrt{2}\sin\theta\cos\varphi}{\tau}\right)\right]. \end{aligned} \tag{39}$$

Here, the notation is the same as in (14). After the introduction of new variables by analogy with the case considered above, namely,

$$(q, p) = \sqrt{3}m^{-1}(\theta\cos\varphi, \theta\sin\varphi), \quad \tilde{\tau} = \omega_0 t,$$

$$\omega_0 = \gamma\lambda M_{\mathrm{Fe}}, \quad \tilde{\tau} = \sqrt{3}m^{-1}\tau,$$

we obtain the action

$$S_E = B_0 f(\tilde{\tau}) = \frac{\mu^2 x^2 v_0}{9\gamma M_{\mathrm{Fe}}} f(\tilde{\tau})$$

in fields $h \sim 1$. Here,

$$f(\tilde{\tau}) = \int d\tilde{t}\left\{\frac{1}{2}(\dot{p}q - \dot{q}p) - \tilde{H}(q, p)\right\}$$

contains the Hamiltonian

$$\tilde{H}(q, p) = \frac{1}{2}(p^2 + q^2) - \tilde{\tau}\ln\left(\cosh\frac{p}{\tilde{\tau}}\right)$$

$$-\tilde{\tau}\ln\left(\cosh\frac{\sqrt{2}\tilde{h}}{\tilde{\tau}} + \cosh\frac{\sqrt{2}q}{\tilde{\tau}}\right) - \frac{1}{2}(p_0^2 + q_0^2) \qquad (40)$$

$$+\tilde{\tau}\ln\left(\cosh\frac{p_0}{\tilde{\tau}}\right) + \tilde{\tau}\ln\left(\cosh\frac{\sqrt{2}\tilde{h}}{\tilde{\tau}} + \cosh\frac{\sqrt{2}q_0}{\tilde{\tau}}\right),$$

where

$$\tilde{h} = \frac{\sqrt{3}(1-h)}{m}$$

and the points $(q_0, p_0)$ are the equilibrium points of the system. These equilibrium points determine phase types [see (10)] in the limit $\tilde{\tau} \longrightarrow 0$, namely, phases I and III,

$$p_0 = \pm 1, \quad q_0 = 0,$$

and phase II,

$$p_0 = \pm 1, \quad q_0 = \pm\sqrt{2}.$$

The energies of the phases are compared at the first-order phase transition point, $\tilde{h} = \pm 1/\sqrt{2}$. The equilibrium points and possible types of instanton trajectories in phases I (III) and at the I–II phase transition point are shown in Fig. 4.

First, consider the tunnel intraphase macroscopic quantum coherence transition $\Delta_{I\to I}$ between the points $(q_0, p_0) = (0, \pm 1)$ in the limit of zero temperature $\tilde{\tau} \longrightarrow 0$. For this purpose, we select the instanton trajectory in the imaginary phase space

$$p = p_1, \quad q = iq_2.$$

The Hamiltonian takes the form

$$\hat{H} = \frac{1}{2}(1 - |p_1|)^2 - \frac{1}{2}q_2^2$$

in the new variables in the $\tilde{\tau} \longrightarrow 0$ limit. The instanton trajectory in the phase space is found as

$$q_2 = 1 - |p_1|$$

from the condition of the conservation of energy (the zeroth level of the first integral $H = 0$). The integration along this trajectory yields

$$\alpha[(0, -1) \longrightarrow (0, +1)] = \lim_{\tilde{\tau}\to 0}\tilde{f}(\tilde{\tau})$$

$$= i\int\frac{1}{2}\left(iq_2 - i\frac{\partial q_2}{\partial p_1}p_1\right)dp_1 = -\int_{-1}^{1} q_2 dp_1 = -1. \qquad (41)$$

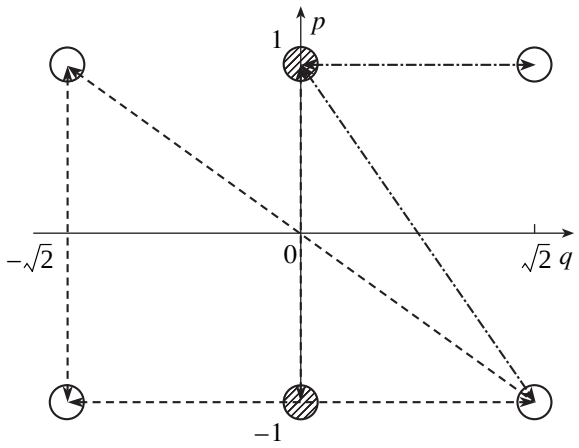



**Fig. 4. H ||** [110]. Equilibrium points and possible types of tunnel transitions in the *pq* plane for phases I (hatched circles) and II (open circles) and for the I ⟶ II phase transition.

Equation (41) also determines the instanton contribution to the probability amplitude of the tunnel intraphase transition III–III. This tunnel interaction results in the splitting of the twofold degenerate energy levels in phases I and III into two levels. The splitting value is

$$\Delta E = 2A_K\exp(-B_0).$$

Next, consider the instanton transitions $\Delta_{I\to II}$ in field $h_1 = 1/\sqrt{2}$ for the transition between the points

$$(q_0, p_0) = (0, +1) \longrightarrow (+\sqrt{2}, +1).$$

Here, we put

$$p = 1 + ip_2, \quad q = q_1.$$

The Hamiltonian then takes the form

$$\hat{H} = \frac{1}{2}q_1^2 - \frac{1}{2}p_2^2$$

at $|q_1| < 1/\sqrt{2}$ in the limit $\tilde{\tau} \longrightarrow 0$. Accordingly, $p_2 = -|q_1|$ in this region. At $|q_1| > 1/\sqrt{2}$ in the limit $\tilde{\tau} \longrightarrow 0$, we have

$$\hat{H} = \frac{1}{2}(\sqrt{2} - |q_1|)^2 - \frac{1}{2}p_2^2.$$

For this reason, the instanton is described by the equation

$$p_2 = -\sqrt{2} + |q_1|$$

in the region $1/\sqrt{2} < q_1 < \sqrt{2}$. Integration along this

instanton trajectory yields

$$\alpha[(0,1) \longrightarrow (\sqrt{2},1)] = \lim_{\tilde{\tau}\to 0}\tilde{f}(\tilde{\tau})$$

$$= i\int d\tilde{t}\left\{\frac{1}{2}(\dot{p}q - \dot{q}p) - \tilde{H}(q,p)\right\} = \frac{i}{2}q_1(1+ip_2)\Big|_0^{\sqrt{2}} \quad (42)$$

$$-i\int_0^{\sqrt{2}}(1+ip_2)dq_1 = -\frac{1}{2} - \frac{1}{\sqrt{2}}i.$$

The instanton contribution $\Delta_{\mathrm{II}\to\mathrm{II}}$ of the transition between the points $(q_0, p_0) = (\sqrt{2}, \pm 1)$ is calculated in a similar way. For this trajectory, we must use the substitution

$$p = p_1, \quad q = \sqrt{2} + iq_2.$$

The instanton contribution then equals

$$\alpha[(\sqrt{2},-1) \longrightarrow (\sqrt{2},+1)] = \lim_{\tilde{\tau}\to 0}\tilde{f}(\tilde{\tau})$$

$$= i\int d\tilde{t}\left\{\frac{1}{2}(\dot{p}q - \dot{q}p) - \hat{H}(q,p)\right\} = -\frac{i}{2}(\sqrt{2}+iq_2)p_1\Big|_{-1}^{1} \quad (43)$$

$$+i\int_{-1}^{1}(\sqrt{2}+iq_2)dp_1 = -1 + \frac{1}{\sqrt{2}}i.$$

For the transition between the points

$$(q,p) = (-\sqrt{2},-1) \longrightarrow (0,1),$$

the trajectory is obtained using the relations

$$q_1 = (p_1 - 1)/\sqrt{2}, \quad q_2 = -p_2/\sqrt{2},$$

which allow the instanton trajectory and its contribution to the exponent of tunneling to be found in the limit $\tilde{\tau} \longrightarrow 0$:

$$\alpha[(\sqrt{2},-1) \longrightarrow (0,1)] = -\sqrt{2} + i. \quad (44)$$

For the transition

$$(q,p) = (-\sqrt{2},-1) \longrightarrow (\sqrt{2},+1)$$

we perform the substitution

$$q_1 = \sqrt{2}p_1, \quad q_2 = -\sqrt{2}p_2.$$

The Hamiltonian then reduces to

$$\hat{H} = \frac{3}{2}(1-|p_1|)^2 - \frac{3}{2}p_2^2$$

in the limit $\tilde{\tau} \longrightarrow 0$. This yields $p_2 = 1 - |p_1|$ and the instanton contribution

$$\alpha[(-\sqrt{2},-1) \longrightarrow (+\sqrt{2},-1)] = -2\sqrt{2}. \quad (45)$$

The transition of the type

$$(q,p) = (0,1) \longrightarrow (\sqrt{2},1)$$

has the smallest Gamow exponential factor and is therefore of the greatest importance for macroscopic quantum coherence. The presence of this tunnel interaction results in partial removal of sixfold degeneracy of levels at the first-order phase transition point under consideration, $h = 1 - m/\sqrt{6}$ (or $\tilde{h} = 1/\sqrt{2}$). In this approximation, the sixfold degenerate level splits into three equidistant twofold degenerate levels separated by the intervals

$$\Delta E = \sqrt{2}A_K \exp(-B_0/2).$$

Similar reasoning is also valid for the second first-order phase transition point at $h = 1 + m/\sqrt{6}$ (or $\tilde{h} = -1/\sqrt{2}$).

A similar analysis for **H** ∥ [110] in the intermediate field region at $h = 1$ ($\tilde{h} = 0$) reveals the presence of three types of instanton transitions, for which

$$\alpha[(-\sqrt{2},-1) \longrightarrow (\sqrt{2},1)] = -2\sqrt{2}, \quad (46)$$

$$\alpha[(-\sqrt{2},1) \longrightarrow (\sqrt{2},1)] = -2 + i\sqrt{2}, \quad (47)$$

$$\alpha[(\sqrt{2},-1) \longrightarrow (\sqrt{2},1)] = -1 + i\sqrt{2}. \quad (48)$$

Among these transitions, of greatest importance are the instanton transitions between two pairs of the nearest neighbors [see (48)], which result in pairwise removal of fourfold degeneracy of levels at this field value. The interval between the doublets is then

$$\Delta E = \sqrt{2}A_K \exp(-B_0),$$

as with the tunnel transitions in phases I and III.

## 7. MACROSCOPIC QUANTUM COHERENCE IN THE REGION OF PHASE TRANSITIONS AT **H** ∥ [111]

We will now consider the case of **H** ∥ [111]. The initial Lagrangian of unit volume of homogeneously magnetized $Ho_xY_{3-x}Fe_5O_{12}$ ferrimagnet written in the coordinates $[11\bar{2}]$, $[\bar{1}10]$, [111] has the form [16]

$$L = \frac{M_{\mathrm{Fe}}}{\gamma}(1-\cos\theta)\dot{\varphi} + M_{\mathrm{Fe}}H\cos\theta + \tau\tilde{x}\sum_{n=0,\pm1}\ln\left[\cosh\left(\frac{1-h-\sqrt{2}\sin\theta\cos(\varphi+2\pi n/3)}{\tau}\right)\right], \quad (49)$$

where the notation corresponds to (14). The calculations will be performed using the normalized time

$$\tilde{t} = \omega_0 t, \quad \omega_0 = \gamma\lambda M_{\mathrm{Fe}}.$$

Consider the region of high magnetic fields $h \sim 1$. Let us introduce the substitutions

$$(p, q) = \frac{3}{2\sqrt{2}m}(\theta\cos\varphi, \sin\varphi), \quad \tilde{\tau} = \frac{3\tau}{2m},$$

$$h_2 = \frac{3(1-h)}{2m}.$$

Since the region of importance of normalized field $|h-1| \sim m$ changes is of the same smallness as the nonzero equilibrium deviation angle $\theta \sim m/h$ values, the Lagrangian of the problem can be simplified and the action can be written as

$$S_E = B_0 f(\tilde{\tau}). \quad (50)$$

Here, as previously,

$$B_0 = \frac{1}{9}\frac{\mu^2 x_c^2 \nu_0}{\gamma M_{\mathrm{Fe}}},$$

and

$$f(\tilde{\tau}, a) = \frac{8}{3}\int d\tilde{t}\left\{\frac{1}{2}(\dot{q}p - \dot{p}q) - \tilde{H}(q, p)\right\}$$

is the effective normalized action with the effective Hamiltonian

$$\tilde{H}(q, p) = \frac{1}{2}(p^2+q^2) - \frac{1}{4}\tilde{\tau}\ln\left[\cosh\left(\frac{h_2-2p}{\tilde{\tau}}\right)\right.$$
$$\left.\times\cosh\left(\frac{h_2+p-\sqrt{3}q}{\tilde{\tau}}\right)\cosh\left(\frac{h_2+p+\sqrt{3}q}{\tilde{\tau}}\right)\right] \quad (51)$$
$$= \frac{1}{2}(p^2+q^2) - \frac{1}{4}\tilde{\tau}\left[\ln\left(\cosh\frac{h_2-2p}{\tilde{\tau}}\right)\right.$$
$$\left.+\ln\left(\cosh\frac{2(h_2+p)}{\tilde{\tau}}\right) + \cosh\left(\frac{2\sqrt{3}q}{\tilde{\tau}}\right)\right],$$

where the $h$ parameter in the first term is set at zero, because the transitions under consideration lie in a narrow neighborhood of fields $|h-1| \sim m \ll 1$.

In the limit $\tilde{\tau} \longrightarrow 0$, the Hamiltonian (system energy) transforms into

$$\tilde{H}(q, p) = \frac{1}{2}(p^2+q^2)$$
$$-\frac{1}{4}(|h_2-2p| + |h_2+p-\sqrt{3}q| + |h_2+p+\sqrt{3}q|). \quad (52)$$

Its minimization gives the equilibrium points shown in Fig. 5.

Phases I and IV [see (12)] are in the center of the unit circle; that is, $p = 0$ and $q = 0$ for them. They correspond to the magnetization of the iron sublattice parallel to the external magnetic field applied along the [111] axis.

Phase II contains three points that lie opposite each other on the same circle,

$$p = 1, \quad q = 0; \quad p = -\frac{1}{2}, \quad q = \pm\frac{\sqrt{3}}{2}.$$

Phase III [see (12)] contains three equivalent equi-

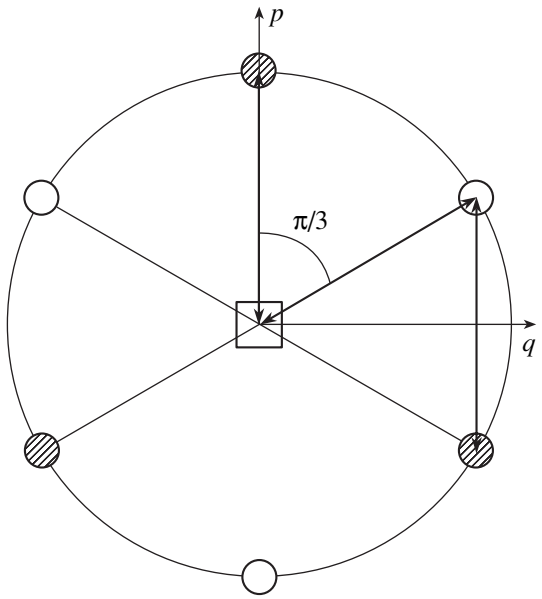



**Fig. 5.** Phases of the orientation states of $\mathbf{M}_{\mathrm{Fe}}$ at $\mathbf{H} \parallel [111]$ in the $pq$ plane and the types of tunnel transitions between them: square, phases I and IV; hatched circles, phase II; and open circles, phase III.

librium points on the circle of unit radius $\sqrt{p^2 + q^2} = 1$,

$$p = -1, \quad q = 0; \quad p = \frac{1}{2}, \quad q = \pm\frac{\sqrt{3}}{2}.$$

At the phase transition points and the $h$ parameter values $h = \pm 1, 0$, the level with the lowest energy includes the largest number of equilibrium points. Let us consider the instanton trajectories at the phase transition points of most interest to us. We introduce the complex variables

$$p = p_1 + ip_2, \quad q = q_1 + iq_2.$$

The Hamiltonian of the problem has imaginary and real components that take the form

$$\begin{aligned}\mathrm{Re}\tilde{H}(q, p) &= \frac{1}{2}(p_1^2 + q_1^2 - p_2^2 - q_2^2) \\ -\frac{1}{4}(|h_2 - 2p_1| &+ |h_2 + p_1 - \sqrt{3}q_1| + |h_2 + p_1 + \sqrt{3}q_1|),\end{aligned} \tag{53}$$

$$\begin{aligned}\mathrm{Im}\tilde{H}(q, p) &= p_1p_2 + q_1q_2 - \frac{1}{4}(2p_2\mathrm{sgn}(2p_1 - h_2) \\ &+ (p_2 - \sqrt{3}q_2)\mathrm{sgn}(h_2 + p_1 - \sqrt{3}q_1) \\ &+ (p_2 + \sqrt{3}q_2)\mathrm{sgn}(h_2 + p_1 + \sqrt{3}q_1))\end{aligned} \tag{54}$$

in the zero-temperature limit $\tilde{\tau} \longrightarrow 0$.

Using an approach similar to that described in the preceding sections, we can find the instanton contributions to the exponential factor of tunneling for all types of interphase transitions. For instance, let us find the radial instanton at $h = 1 - 2m/3$ ($h_2 = 1$) between the points $p = 0$, $q = 0$ and $p = 1$, $q = 0$, which are the equilibrium points of phases I and II. For this purpose, set $p = p_1$ and $q = iq_2$. The imaginary part of the complex Hamiltonian then identically equals zero along the trajectory under consideration; that is,

$$\mathrm{Im}\tilde{H}(p, q) = 0.$$

The remaining real Hamiltonian part is

$$\begin{aligned}&\mathrm{Re}\tilde{H}(q, p) \\ &= \begin{cases}\frac{1}{2}(p_1^2 - q_2^2) - \frac{3}{4}, & 0 < p_1 < \frac{1}{2} \\ \frac{1}{2}((p_1 - 1)^2 - q_2^2) - \frac{3}{4}, & \frac{1}{2} < p_1 < 1.\end{cases}\end{aligned} \tag{55}$$

At the zero Hamiltonian level that passes through the equilibrium points under consideration, the real part of the action identically equals zero, which gives the sought instanton solution in the phase space,

$$q_2 = \begin{cases}p_1, & 0 < p_1 < \frac{1}{2} \\ 1 - p_1, & \frac{1}{2} < p_1 < 1,\end{cases} \tag{57}$$

and allows the contribution to the action to be calculated in the form

$$\begin{aligned}\tilde{f}_{\mathrm{I}\to\mathrm{II}}(\tilde{\tau} = 0, h_2 = 1) &= \frac{4}{3}\int(\dot{q}_2p_1 - \dot{p}_1q_2)dt \\ &= -\frac{8}{3}q_2dp_1 = -\frac{2}{3}.\end{aligned} \tag{58}$$

Similar instanton calculations between the points $p = -1/2$, $q = \sqrt{3}/2$ (phase II) and $p = 1/2$, $q = \sqrt{3}/2$ (phase III) at the first-order phase transition point and $h = 1$ ($h_2 = 0$) give the instanton contribution to the normalized Gamow factor,

$$\begin{aligned}\tilde{f}_{\mathrm{II}\to\mathrm{III}}(\tilde{\tau} = 0) &= \frac{4i}{3}\int\left[p_1\left(\frac{\sqrt{3}}{2} + iq_2\right) - q_2p_1\right]dt \\ &= i\frac{4}{\sqrt{3}} - \frac{8}{3}\int_{-1/2}^{1/2} q_2dp_1 = -\frac{2}{3} + \frac{4}{\sqrt{3}}i.\end{aligned} \tag{59}$$

We see that, as distinct from the previous case, the instanton transition introduces the phase shift into the tunneling amplitude, which is determined by the imaginary term. The transition in the opposite direction changes the phase shift sign. Similar phase shifts appear in the tunneling amplitudes under the instanton transition between the equilibrium points of phases II and III. Let us analyze the special features of the tunnel splitting of the ground state introduced by macroscopic quantum coherence transitions in the vicinity of the phase transitions considered above.

## 8. THE DIAGONALIZATION OF THE INSTANTON INTERACTION HAMILTONIAN CLOSE TO PHASE TRANSITION POINTS AT **H** ∥ [111]

Let us consider phase transitions close to $h = h^* = 1 \pm m$. The wavefunction for the four energetically equivalent equilibrium points of phases I and II (III and IV) can be written as

$$\begin{aligned}\Psi(p, q) &= C_1\Psi_{\mathrm{I}}(0, 0) + C_2\Psi_{\mathrm{II}}(1, 0) \\ &+ C_3\Psi_{\mathrm{II}}\left(\frac{1}{2}, \frac{\sqrt{3}}{2}\right) + C_4\Psi_{\mathrm{II}}\left(\frac{1}{2}, -\frac{\sqrt{3}}{2}\right),\end{aligned}$$

where the coefficients $C_i$ are determined from the equation

$$\begin{pmatrix} E_{\rm I}(h)-E & \Delta_{\rm II\to I} & \Delta_{\rm II\to I} & \Delta_{\rm II\to I} \\ \Delta_{\rm II\to I} & E_{\rm II}(h)-E & 0 & 0 \\ \Delta_{\rm II\to I} & 0 & E_{\rm II}(h)-E & 0 \\ \Delta_{\rm II\to I} & 0 & 0 & E_{\rm II}(h)-E \end{pmatrix}\begin{pmatrix} C_0 \\ C_1 \\ C_2 \\ C_3 \end{pmatrix} = 0. \quad (60)$$

Here,

$$\Delta_{\rm II\to I} = A_K \exp\left(-\frac{2}{3}B_0\right).$$

At the phase transition point, we have $E_{\rm I}(h^*) = E_{\rm II}(h^*)$. Near the transition ($h - h^* \ll 1$), the change in the off-diagonal element can be considered small (if $\Delta_{\rm II\to I} \ll \hbar\omega_0$). The solution to the characteristic equation for system (60) yields the following result: in the vicinity of the I ⟶ II first-order phase transition, two split singlet levels with a degenerate doublet between them are formed close to the ground state. At the phase transition point, the splitting of the levels has the form

$$\Delta E_{1,2} = \pm\sqrt{3}\Delta_{\rm II\to I}, \quad \Delta E_{3,4} = 0, \quad (61)$$

where

$$\Delta E_i = E_i - E_{\rm I}.$$

Next, consider the III ⟶ II first-order phase transition, which occurs at $h = h^* = 1$. At the transition point, six states energetically degenerate with respect to the azimuthal angle are formed, three states in each phase. We will write the wavefunction of the system in the form

$$\Psi(p,q) = C_1\Psi_{\rm III}(-1,0) + C_2\Psi_{\rm III}\left(\frac{1}{2},\frac{\sqrt{3}}{2}\right) + C_3\Psi_{\rm II}\left(\frac{1}{2},-\frac{\sqrt{3}}{2}\right) + C_4\Psi_{\rm III}\left(-\frac{1}{2},-\frac{\sqrt{3}}{2}\right) + C_5\Psi_{\rm II}\left(-\frac{1}{2},-\frac{\sqrt{3}}{2}\right) + C_6\Psi_{\rm III}(1,0).$$

The solution to the eigenvalue problem for the tunneling interaction Hamiltonian is obtained as previously. It follows from the corresponding calculations that, at the phase transition point

$$E_{\rm III}(h^*) = E_{\rm II}(h^*) = E(1) = 0,$$

the solution to the characteristic equation describes three pairwise degenerate doublets,

$$\Delta E_{1,2} = \pm|\Delta_{\rm III\to II}|, \quad \Delta E_{3,4} = |\Delta_{\rm III\to II}|\left(\frac{5}{2}-\frac{\sqrt{13}}{2}\right)^{1/2}, \quad \Delta E_{5,6} = |\Delta_{\rm III\to II}|\left(\frac{5}{2}+\frac{\sqrt{13}}{2}\right)^{1/2}, \quad (62)$$

where $|\Delta_{\rm III\to II}| = A_K\exp(-2B_0/3)$.

**Table**

| Field direction | Orientation of $\mathbf{M}_{\rm Fe}(\theta,\varphi)$ in phases | Phase transition fields | Normalized action $\tilde{f}(0)$ for the most important transitions | Level splitting |
|---|---|---|---|---|
| **H** ∥ [100] | Phases I, II<br>$\theta_{\rm I,\,II} = \sqrt{\frac{2}{3}}\frac{m}{h}\left(1 \pm \frac{m}{\sqrt{3}h}\right)$,<br>$\varphi_{\rm I} = \varphi_{\rm II} = \pm\pi/4; \pm 3\pi/4$ | $h_{\rm I\to II} = 1 - \frac{m^2}{3}$ | $\tilde{f}_{\rm I\to II} = -\frac{m^2}{6}$ | $\{4\} \longrightarrow 2+2$ |
| **H** ∥ [110] | Phases I, II, III<br>$\theta_{\rm I,\,III} = \frac{m}{\sqrt{3}h}$, $\varphi_{\rm I,\,III} = \pm\pi/2$,<br>$\theta_{\rm II} = m$, $\tanh\varphi_{\rm II} = \pm 1/\sqrt{2}$ | $\tilde{h}_{\rm I\to II} = 1 - \frac{m}{\sqrt{6}}$<br>$\tilde{h}_{\rm II\to III} = 1 + \frac{m}{\sqrt{6}}$ | $\tilde{f}_{\rm I\to II} = \tilde{f}_{\rm I\to III} = -\frac{1}{2} \mp \frac{i}{\sqrt{2}}$ | $\{2+4\} \longrightarrow 2+2+2$ |
| **H** ∥ [111] | Phases I, II, III, IV<br>$\theta_{\rm I,\,IV} = 0$, $\theta_{\rm II,\,III} = \frac{2\sqrt{2}}{3}m$,<br>$\varphi_{\rm II} = 0, \frac{2\pi}{3}, \frac{4\pi}{3}$,<br>$\varphi_{\rm III} = \varphi_{\rm II} + \pi$ | $\tilde{h}_{\rm I\to II} = 1 - \frac{2}{3}m$<br>$\tilde{h}_{\rm II\to III} = 1$<br>$\tilde{h}_{\rm III\to IV} = 1 + \frac{2}{3}m$ | $\tilde{f}_{\rm I\to II} = -\frac{2}{3}$<br>$\tilde{f}_{\rm II\to III} = -\frac{2}{3} \pm \frac{4}{\sqrt{3}}i$<br>$\tilde{f}_{\rm III\to IV} = -\frac{2}{3}$ | $\{1+3\} \longrightarrow 1+2+1$<br>$\{3+3\} \longrightarrow 2+2+2$<br>$\{3+1\} \longrightarrow 1+2+1$ |

## 9. CONCLUSIONS

The results of the preceding analysis are summarized in the table. The table contains the normalized critical field values for various field orientations in the crystal at which the first-order orientation phase transitions that we are considering occur. For each transition, angular variables for angular phase equilibrium points and the normalized action for the corresponding instanton transitions are given. The ground state of a small particle (the lowest energy level among the possible states) is split near the phase transition point because of the mesoscopic mixing of phases caused by instanton interactions. The scheme of ground state splitting is given in the last column of the table. The number of degenerate energy states in the phase separated in energy from the angular phases between which the first-order transition occurs is given in braces. The orientation phase transition at **H** ∥ [100] attracts attention. The energy barrier between two orientation phase states with different polar angles is then substantially different from that for intraphase energetically degenerate states with different azimuthal directions of the magnetization of the iron sublattice. The instanton contributions to the exponential tunneling factor for these two cases differ by a factor of $2/r^2$, where

$$r = x_c\mu/3M_{\mathrm{Fe}} = x\mu/15\mu_B$$

is a small value for dilute compounds with $x \ll 1$. This means that the mesoscopic volume at which the influence of the mesoscopic mixing of phases I and II on the splitting of degenerate ground state levels for this field orientation can be noticeable is substantially larger than for the other orientations.

Let us estimate the mesoscopic volume for the tunnel transition between phases I and II at **H** ∥ [100], when the macroscopic quantum coherence splitting amounts to $\Delta\omega = 1000$ MHz. This can be done using the equation

$$\Delta\omega = A_\omega \exp(B_0\alpha_{\mathrm{inst}}),$$

where

$$A_\omega = \omega_0\sqrt{B_0\alpha_{\mathrm{inst}}}, \quad \omega_0 = \gamma\lambda M, \quad \alpha_{\mathrm{inst}} = \mathrm{Re}\,\tilde{f}(0).$$

It follows from this equation that

$$v_0 \approx \frac{4}{c_0\alpha_{\mathrm{inst}}}\left(\frac{15\mu_B}{x\mu}\right)^2 \ln\frac{\omega_0}{\Delta\omega} = \frac{a^3}{4}\left(\frac{15\mu_B}{x\mu}\right)\ln\frac{\omega_0}{\Delta\omega}.$$

With $x = 0.1$, $\mu = 7.5\mu_B$, $\omega_0 = 1000$ GHz, and $a = 1.24$ nm, we obtain $v_0 \approx 10^5 a^3 \approx 2 \times 10^5$ nm$^3$. The frequency splitting exponentially increases up to the main resonance frequency as the volume decreases. An increase in splitting is also favored by a decrease in the concentration of impurity ions. A similar frequency splitting related to other instanton transitions and occurring at other field orientations, when $\alpha_{\mathrm{inst}} = -1/2$, can be observed for the same volume but at substantially lower impurity concentrations, namely, at $x = 0.01$.

Decreasing the concentration to $x < 0.001$ causes the formation of magnetically inhomogeneous states in the vicinity of impurity centers [14]. The mesoscopic volume of spin clusters in the vicinity of Ising ions is then $v_{cl} \sim a^2\sqrt{A/\lambda M^2}$ and is virtually independent of the concentration of impurities at such dilution values. Only the number of spin clusters in the given crystal volume then changes. Mesoscopic magnetization oscillations caused by the tunneling of the magnetization of the iron sublattice between orientation phases in a cluster are described like macroscopic quantum coherence oscillations in small particles of volume $v_{cl}$ with the effective concentration $x_c = 1/v_{cl}$ in the equation for action (18), (44). The normalizing exponential factor is then very small,

$$B_0 \sim \frac{a}{48\sqrt{A/\lambda M^2}}\left(\frac{\mu}{15\mu_B}\right)^2 \ll 1.$$

As a result, phase states with different orientations of equilibrium positions become physically indistinguishable because of quantum fluctuations, when the distance between mesoscopic spin inhomogeneities increases until they cease to interact. In the transition region of concentrations, we can observe macroscopic quantum coherence phenomena in mesoscopic spin inhomogeneities formed by clusters of Ising impurity ions [17].

## ACKNOWLEDGMENTS

This work was financially supported by the Russian Foundation for Basic Research (project nos. 02-02-16704, 03-02-17166), the Ministry of Education of Russia (project no. E02-3.4-116), and INTAS (project no. 99-01-839).

## REFERENCES

1. M. Enz and R. Shilling, J. Phys. C **19**, 1765 (1986); J. L. van Hemmen and A. Suto, Phys. Rev. B **141**, 37 (1986); E. M. Chudnovsky and L. Gunter, Phys. Rev. Lett. **60**, 661 (1988).
2. E. M. Chudnovsky and J. Tejada, *Macroscopic Quantum Tunneling of the Magnetic Moment* (Cambridge Univ. Press, Cambridge, 1998).
3. D. D. Awshalom, M. A. McCord, and G. Grinstein, Phys. Rev. Lett. **65**, 783 (1990); D. D. Awshalom, J. F. Smith, G. Grinstein, *et al.*, Phys. Rev. Lett. **68**, 3092 (1992).
4. M. Uehara and B. Barbara, J. Phys. (Paris) **47**, 235 (1986); C. Paulsen, L. C. Sampaio, B. Barbara, *et al.*, Phys. Lett. A **161**, 319 (1991).
5. M. Dressel, B. Gorshunov, K. Rajagopal, *et al.*, Phys. Rev. B **67**, 060405(R) (2003).

6. A. Garg, Europhys. Lett. **22**, 205 (1993); V. Yu. Golyshev and A. F. Popkov, Europhys. Lett. **29**, 327 (1995).
7. T. Egami, Phys. Status Solidi A **20**, 157 (1973); E. M. Chudnovsky, O. Iglesias, and P. C. E. Stamp, Phys. Rev. B **46**, 5392 (1992); V. V. Dobrovitskii and A. K. Zvezdin, J. Magn. Magn. Mater. **156**, 205 (1996); A. F. Popkov, Fiz. Tverd. Tela (St. Petersburg) **44**, 135 (2002) [Phys. Solid State **44**, 140 (2002)].
8. D. Gatteschi, A. Caneschi, L. Paedi, and R. Sessoli, Science **265**, 1054 (1994); F. Lionti *et al.*, J. Appl. Phys. **81**, 4608 (1997); I. Tupitsin and B. Barbara, in *Magnetism: Molecules to Materials*, Ed. by A. Muller and M. Drillon (Wiley–VCH, Weinheim, 2000); P. Politi, A. Rettori, F. Hartmann-Boutron, and J. Villian, Phys. Rev. Lett. **75**, 537 (1995).
9. D. Loss, D. P. DiVincenzo, and G. Grinstein, Phys. Rev. Lett. **69**, 3232 (1992); J. Von Delft and C. L. Henley, Phys. Rev. Lett. **69**, 3236 (1992).
10. N. V. Prokof'ev and P. C. E. Stamp, J. Phys.: Condens. Matter **5**, L667 (1994).
11. P. C. E. Stamp, Physica B (Amsterdam) **197**, 133 (1994).
12. V. Yu. Golyshev and A. F. Popkov, Phys. Rev. B **56**, 2712 (1997).
13. A. A. Mukhin, A. S. Prokhorov, B. P. Gorshunov, *et al.*, Usp. Fiz. Nauk **172**, 1306 (2002) [Phys. Usp. **45**, 1186 (2002)].
14. A. K. Zvezdin, A. A. Mukhin, and A. I. Popov, Zh. Éksp. Teor. Fiz. **72**, 1097 (1977) [Sov. Phys. JETP **45**, 573 (1977)].
15. R. Z. Levitin, A. I. Popov, and V. V. Snegirev, Fiz. Tverd. Tela (Leningrad) **24**, 3138 (1982) [Sov. Phys. Solid State **24**, 1777 (1982)].
16. K. P. Belov, A. K. Zvezdin, A. M. Kadomtseva, and R. Z. Levitin, *Reorientational Transitions in Rare-Earth Magnets* (Nauka, Moscow, 1980), p. 271 [in Russian]; A. K. Zvezdin, V. M. Matveev, A. A. Mukhin, and A. I. Popov, *Rare-Earth Ions in Magnetic-Ordered Crystals* (Nauka, Moscow, 1985), p. 142 [in Russian].
17. A. F. Popkov and A. I. Popov, Pis'ma Zh. Éksp. Teor. Fiz. **65**, 445 (1997) [JETP Lett. **65**, 470 (1997)].
18. A. K. Zvezdin and A. F. Popkov, Pis'ma Zh. Éksp. Teor. Fiz. **57**, 548 (1993) [JETP Lett. **57**, 562 (1993)].
19. R. P. Feynman and A. R. Hibbs, *Quantum Mechanics and Path Integrals* (McGraw-Hill, New York, 1965; Mir, Moscow, 1968); S. Coleman, *Aspects of Symmetry* (Cambridge Univ. Press, Cambridge, 1985); A. J. Legett, Rev. Mod. Phys. **59**, 1 (1987).
20. A. Garg and G.-H. Kim, Phys. Rev. B **45**, 12 921 (1992).

*Translated by V. Sipachev*

*Journal of Experimental and Theoretical Physics, Vol. 99, No. 4, 2004, pp. 850–855.*
*From Zhurnal Éksperimental'noĭ i Teoreticheskoĭ Fiziki, Vol. 126, No. 4, 2004, pp. 979–985.*


**SOLIDS**
**Electronic Properties**

# Collective Excitations in Exciton Crystal¶

**D. V. Kulakovskii[a], Yu. E. Lozovik[b], and A. V. Chaplik[c]**

[a]*Institute of Solid State Physics, Russian Academy of Sciences, Chernogolovka, Moscow oblast, 142432 Russia*
*e-mail: kulakovd@issp.ac.ru*

[b]*Institute of Spectroscopy, Russian Academy of Sciences, Troitsk, Moscow oblast, 142092 Russia*
*e-mail: lozovik@isan.troitsk.ru*

[c]*Institute of Semiconductor Physics, Siberian Division, Russian Academy of Sciences, Novosibirsk, 630090 Russia*

Received May 19, 2004

**Abstract**—Crystal phase of indirect excitons formed by spatially separated electrons and holes in coupled quantum wells is analyzed. The collective mode spectrum of the exciton crystal at zero and nonzero magnetic fields is found. The spectrum consists of two optical and two acoustical modes (transverse and longitudinal in each case). We also study changes of the dipole crystal collective excitations at the transition exciton crystal–electron–hole plasma. *© 2004 MAIK "Nauka/Interperiodica".*

In a bilayer system with spatially separated electrons and holes, the overlapping of wavefunctions and, consequently, the recombination rate is suppressed and therefore the lifetime of laser-pumped electrons and holes can be sufficient for the appearance of different interesting quasi-equilibrium electron–hole phases [1, 2]. In particular, a superfluid phase with anomalous transport and optical properties [3], quasi-Josephson phenomena [4], and nonordinary behavior in strong magnetic fields [5, 6] can be observed (we note that there is strong analogy and some mapping between properties of bilayer electron–hole and electron–electron systems). The essential feature of the system of interwell excitons is the existence of parallel electric dipole moments of the excitons. Dipole–dipole repelling of the excitons suppresses the exchange interaction of the excitons and stabilizes exciton phases [2] in comparison with the three-dimensional case [7]. Modern manufacturing techniques of layer structures make it possible to obtain high-quality nanostructures and easily vary their parameters. In recent years, very interesting experimental results have been obtained in such bilayer electron–electron and electron–hole [8–13] systems. Presently, is possible to obtain a system with spatially separated electrons and holes of such low densities that the distance between charge carriers in a layer is much greater than the interlayer separation. In this case, formation of indirect excitons (dipoles) [14–20] and the appearance of new exciton phase, the exciton crystal, becomes possible at low temperature in some intermediate region of exciton density [2, 21, 22]. This causes the great interest in the theoretical research of the dipole crystal collective excitations at zero and nonzero normal magnetic fields, which can be used for experimental verification of the exciton crystal existence by optical methods. It is also interesting to study the behavior of the collective excitation spectrum of a dipole crystal at the transition into the phase of the electron–hole plasma.

We consider a double-layer two-dimensional system. The first layer contains an electron channel with the density of carriers $N_e$, and the second layer contains holes with the density $N_h$. The distance between the layers is $D$. We are interested the case where $N_e = N_h$. At low temperatures and small concentrations of electrons and holes ($N_{e(h)}a_B^2 \ll 1$, where $a_B$ is the effective radius of indirect excitons along layers), the system is a weakly nonideal excitonic gas of the areal density $N_{\text{exc}} = N_e = N_h$ with dipole moments $d$ perpendicular to the layers; in the ground state, $d = eD$ and it increases with the distance between layers. In a spatially separated electron–hole system, contrary to ordinary (single-layer) electron–hole systems, direct dipole–dipole repulsion makes the main contribution to the total energy (in contrast to the ordinary single-layer electron–hole system). At the same time, van der Waals attraction between excitons and the exchange interaction are insignificant (less than 1% compared with the contribution of the direct dipole–dipole interaction). At low temperatures, in a certain region of low concentrations $N_{\text{exc}}$ and interlayer separations $D$, dipole–dipole repulsion leads to the formation of an indirect exciton crystal. The small contribution of the exchange interaction in a spatially separated electron–hole system is connected with the exponentially small probability of tunnelling through the barrier of the dipole–dipole interaction.

¶ This article was submitted by authors in English.

We start with the exciton crystal phase. The kinetic energy of the system is

$$T_{\rm kin} = \sum_{n,m=-\infty}^{\infty}\left(\frac{m_e\mathbf{r}_{enm}^2}{2}+\frac{m_h\mathbf{r}_{hnm}^2}{2}\right) = \sum_{n=-\infty}^{\infty}\left(\frac{M}{2}\dot{\mathbf{R}}_{nm}^2+\frac{\mu}{2}\dot{\rho}_{nm}^2\right), \quad (1)$$

where $M = m_e + m_h$ is the exciton mass, $\mu = (m_e m_h)/(m_e + m_h)$ is the reduced mass, $\mathbf{R}_{nm} = (m_e\mathbf{r}_{enm} + m_h\mathbf{r}_{hnm})M^{-1}$ is the coordinate of the center of the exciton mass, and $\boldsymbol{\rho}_{nm} = \mathbf{r}_{enm} - \mathbf{r}_{hnm}$ represents the coordinates of relative motion.

The potential energy of the system consists of two terms. The first term is the Coulomb interaction of a spatially separated electron and hole in one exciton, which, in the case where $\rho \ll D$, can be transformed as

$$U_{\rm exc} = -\frac{e^2}{\epsilon\sqrt{\rho^2+D^2}} \approx -\frac{e^2}{\epsilon D}+\frac{1}{2}\frac{e^2}{\epsilon D^3}\rho^2. \quad (2)$$

Hence, for small oscillation amplitudes, this term can be approximated by the parabolic potential,

$$U_1 = \sum_{n,m}\frac{\mu\omega_0^2\boldsymbol{\rho}_{nm}^2}{2}, \quad (3)$$

$$\omega_0^2 = \frac{e^2}{\epsilon\mu D^3}.$$

For a real system of finite thickness, the quantity $\omega_0$ must be found using $z$-dependent wavefunctions in each layer.

The second term in the potential energy is the inter-exciton interaction along layers. In the system under consideration, the interaction of indirect excitons is the dipole–dipole repulsion,

$$U_2 = \sum_{n,n',m,m'}\frac{e^2\boldsymbol{\rho}_{nm}\boldsymbol{\rho}_{n'm'}\mathbf{R}^2-3e^2(\boldsymbol{\rho}_{nm}\mathbf{R})(\boldsymbol{\rho}_{n'm'}\mathbf{R})}{\epsilon R^5}. \quad (4)$$

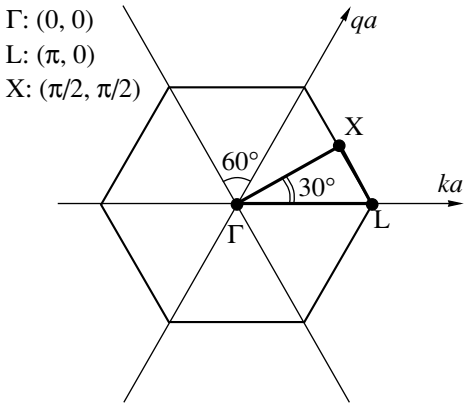



**Fig. 1.** The hexagonal reciprocal lattice. Shown are the Γ–L, Γ–X, and L–X directions. Presented coordinates of symmetry points are coefficients attached to basis vectors of the oblique-angle coordinate system.

Here, $\mathbf{R}$ is the vector between dipoles at different cells ($a$ is the lattice period)

$$\mathbf{R}^2 = [a(n-n')+(x_n-x_{n'})]^2 + [a(m-m')+(y_m-y_{m'})]^2 + 2[a(n-n')+(x_n-x_{n'})]\times[a(m-m')+(y_m-y_{m'})]\cos\alpha, \quad (5)$$

$x_n(y_m)$ are the displacements to the $x$- ($y$-) direction of the dipole ($n$, $m$) from the equilibrium, and $\alpha$ is the angle between basis vectors of an arbitrary lattice. We consider a triangular lattice of the dipole (exciton) crystal. It corresponds to a minimum of the potential energy, and the spectrum of its frequencies is stable in the two-dimensional case, in contrast to other configurations (see, e.g., [21, 23]). Hence, $\alpha = \pi/3$. The corresponding reciprocal hexagonal lattice is presented in Fig. 1.

In the harmonic approximation, we expand the potential energy $U_2$ to the second order in displacements. As a result, it is easy to obtain the equations of motion for the center-of-mass coordinates $x_n$ and $y_m$, as well as for the coordinates of relative motion $u_n$, $v_m$:

$$\ddot{u}_n = -\omega_0^3 u_n + \Omega^2\sum_{n',m'=-\infty}^{\infty}\frac{[(n-n')^2-2(m-m')^2]u_{n'}-3[(n-n')(m-m')]v_{m'}}{[(n-n')^2+(m-m')^2+(n-n')(m-m')]^{5/2}}, \quad (6)$$

$$\ddot{v}_m = -\omega_0^2 v_m + \Omega^2\sum_{n',m'=-\infty}^{\infty}\frac{[(m-m')^2-2(n-n')^2]v_{m'}-3[(n-n')(m-m')]u_{n'}}{[(n-n')^2+(m-m')^2+(n-n')(m-m')]^{5/2}}, \quad (7)$$

$$\ddot{x}_n = -\Omega_1^2 \sum_{n', m' = -\infty}^{\infty} \frac{\left[-3(n-n')^2 + \frac{45}{4}(m-m')^2\right](x_n - x_{n'})}{[(n-n')^2 + (m-m')^2 + (n-n')(m-m')]^{7/2}}$$
$$- \Omega_1^2 \sum_{n', m' = -\infty}^{\infty} \frac{\left[15\frac{\sqrt{3}}{2}(m-m')\left[(n-n') + \frac{1}{2}(m-m')\right]\right](y_m - y_{m'})}{[(n-n')^2 + (m-m')^2 + (n-n')(m-m')]^{7/2}}, \quad (8)$$

$$\ddot{y}_m = -\Omega_1^2 \sum_{n', m' = -\infty}^{\infty} \frac{\left[-3(m-m')^2 + 15\left[(n-n') + \frac{1}{2}(m-m')\right]^2\right](y_m - y_{m'})}{[(n-n')^2 + (m-m')^2 + (n-n')(m-m')]^{7/2}}$$
$$- \Omega_1^2 \sum_{n', m' = -\infty}^{\infty} \frac{\left[15\frac{\sqrt{3}}{2}(m-m')\left[(n-n') + \frac{1}{2}(m-m')\right]\right](x_n - x_{n'})}{[(n-n')^2 + (m-m')^2 + (n-n')(m-m')]^{7/2}}. \quad (9)$$

At zero magnetic field, the variables $u_n$, $v_m$, describing optical (out-of-phase) vibrations, and $x_n$, $y_m$, describing acoustic (in-phase) vibrations, are separated. From Eqs. (6)–(9), we therefore obtain two independent determinants of the second rank, which give the spectrum of collective excitations of the two-dimensional dipole crystal.

We search for the normal mode in the form $u_n = u\exp[-i\omega t + ikan]$, $v_m = v\exp[-i\omega t + iqam]$. The final equations for optical modes are

$$\begin{bmatrix} \omega^2 - \omega_0^2 + \Omega^2 f_{11}(k, q) & -3\Omega^2 f_{12}(k, q) \\ -3\Omega^2 f_{21}(k, q) & \omega^2 - \omega_0^2 + \Omega^2 f_{22}(k, q) \end{bmatrix} \times \begin{bmatrix} u \\ v \end{bmatrix} = 0, \quad (10)$$

where $\Omega^2 = 2e^2/\mu\epsilon a^3$, $(k, q)$ are the $x$, $y$ components of wave vector $\mathbf{k}$,

$$f_{11}(f, q) = \sum_{N, J = 1}^{\infty} \frac{N^2 - 2J^2}{(N^2 + J^2 + NJ)^{5/2}} \cos(kaN)\cos(qaJ),$$

$$f_{22}(k, q) = \sum_{N, J = 1}^{\infty} \frac{J^2 - 2N^2}{(N^2 + J^2 + NJ)^{5/2}} \cos(kaN)\cos(qaJ),$$

$$f_{12}(k, q) = f_{21}(k, q) = \sum_{N, J = 1}^{\infty} \frac{NJ}{(N^2 + J^2 + NJ)^{5/2}} \sin(kaN)\sin(qaJ),$$

$N = n - n'$, $J = m - m'$.

Correspondingly, for acoustic modes, we obtain

$$\begin{bmatrix} \omega^2 - \Omega_1^2 \tilde{f}_{11}(k, q) & \Omega_1^2 \tilde{f}_{12}(k, q) \\ \Omega_1^2 \tilde{f}_{21}(k, q) & \omega^2 - \Omega_1^2 \tilde{f}_{22}(k, q) \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = 0, \quad (11)$$

where $\Omega_1^2 = 8e^2D^2/M\epsilon a^5$,

$$\tilde{f}_{11}(k, q) = \sum_{N, J = 1}^{\infty} \frac{-3N^2 + 15\frac{3}{4}J^2}{(N^2 + J^2 + NJ)^{7/2}}[1 - \cos(kaN)\cos(qaJ)],$$

$$\tilde{f}_{22}(k, q) = \sum_{N, J = 1}^{\infty} \frac{-3J^2 + 15\left(N + \frac{1}{2}J\right)^2}{(N^2 + J^2 + NJ)^{7/2}}[1 - \cos(kaN)\cos(qaJ)],$$

$$\tilde{f}_{12}(k, q) = \tilde{f}_{21}(k, q) = \sum_{N, J = 1}^{\infty} \frac{-15\frac{\sqrt{3}}{2}J\left(N + \frac{1}{2}J\right)}{(N^2 + J^2 + NJ)^{7/2}} \sin(kaN)\sin(qaJ).$$

The dispersion laws $\omega(k, q)$ given by (10) and (11) are presented in Fig. 2 (optical) and Fig. 3 (acoustical). These dispersion laws are obtained, respectively, for the following directions: (a) Γ–X, (b) Γ–L, and (c) L–X, where Γ, L, and X are symmetry points in the first Brillouin zone. The hexagonal reciprocal lattice that corresponds to the triangular lattice in coordinate space and the symmetry directions are shown in Fig. 1. In our calculations, the dimensionless energy units $\omega = \omega/\omega_0$, $\Omega = \Omega/\omega_0 = 0.2$, $\Omega_1 = \Omega_1/\omega_0 = 0.01$ and the dimensionless wave vector units $k = ka$, $q = qa$ are used. Two modes are found in each of the spectra. One of the modes corresponds to lon-

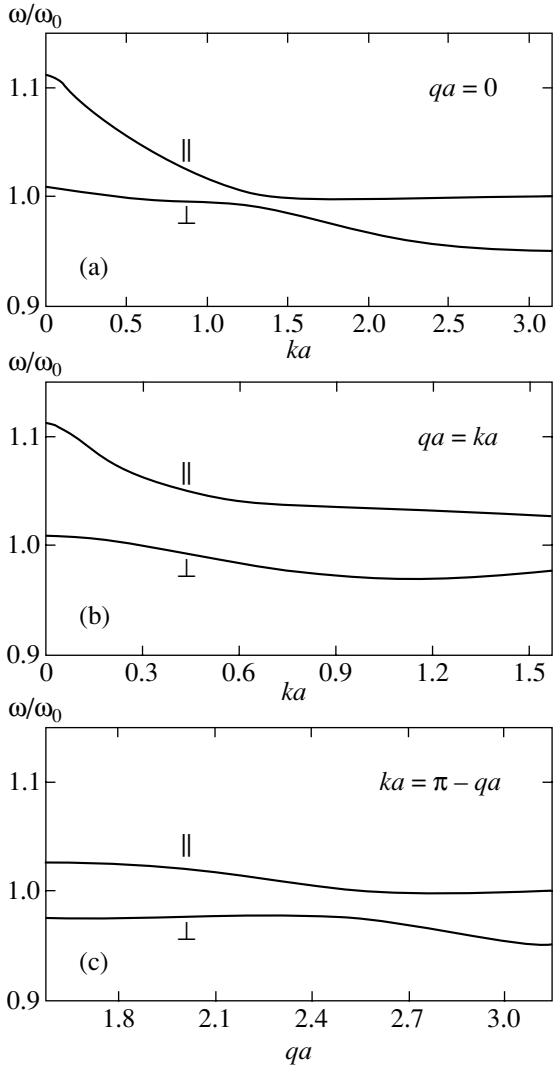



**Fig. 2.** Dispersion of the optical longitudinal and transverse modes: LXLX. (a) Γ–L direction, (b) Γ–X, and (c) L–X.

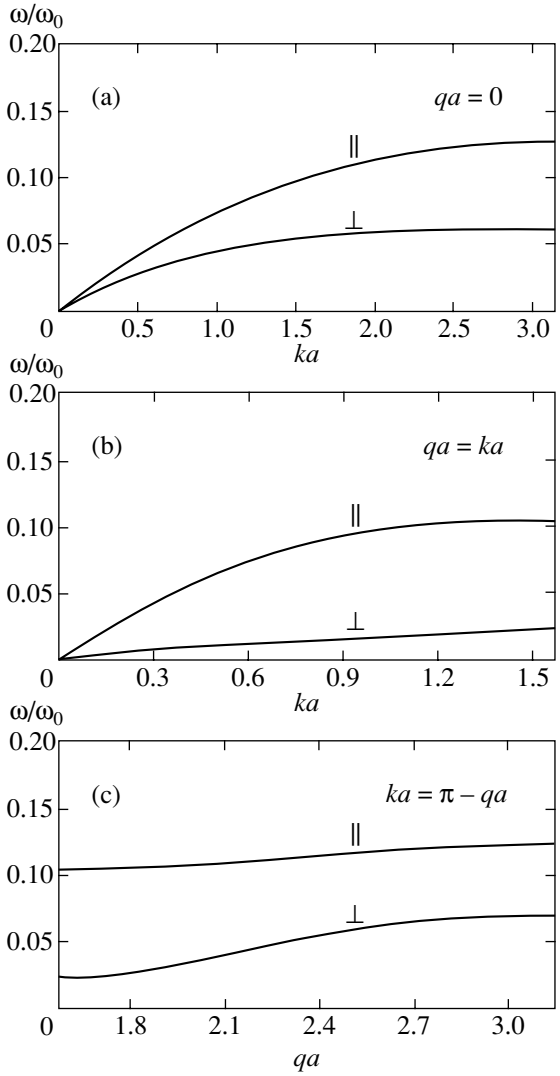



**Fig. 3.** Dispersion of the acoustical longitudinal and transverse modes: LXLX. (a) Γ–L direction, (b) Γ–X, and (c) L–X.

gitudinal (in-phase or out-of-phase) vibrations, and the other, to transverse vibrations. The transverse modes must undergo strong changes during crystal melting.

In the presence of the perpendicular magnetic field, we add the Lorentz components of acceleration and obtain the system of four equations

$$\begin{bmatrix} \omega^2 - \omega_{11}^2 & \omega_{12}^2 + i\omega(\omega_e - \omega_h) & 0 & i\omega\dfrac{eB}{\mu c} \\ \omega_{12}^2 - i\omega(\omega_e - \omega_h) & \omega^2 - \omega_{22}^2 & -i\omega\dfrac{eB}{\mu c} & 0 \\ 0 & -i\omega\dfrac{eB}{Mc} & \omega^2 - \omega_{33}^2 & \omega_{34}^2 \\ i\omega\dfrac{eB}{Mc} & 0 & \omega_{34}^2 & \omega^2 - \omega_{44}^2 \end{bmatrix} \begin{bmatrix} u \\ v \\ x \\ y \end{bmatrix} = 0, \tag{12}$$

where $B$ is the magnetic field and $\omega_e = eB/m_e c$ and $\omega_h = eB/m_h c$ are the respective cyclotron energies of an electron and a hole. Dispersion laws for the exciton crystal in the magnetic field are presented in Fig. 4.

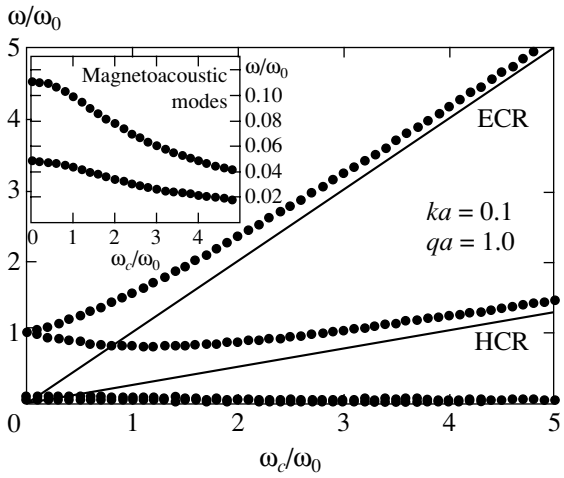



**Fig. 4.** The dipole crystal collective excitations as a function of magnetic field. ECR and HCR are the electron and hole cyclotron energies in the GaAs/AlGaAs structure. The inset shows the magnetoacoustic modes in enlarged scale.

In strong magnetic fields, optical modes approach the electron and hole cyclotron energies. Such behavior occurs because the cyclotron energy of electrons (holes) in this field is much greater than the characteristic Coulomb interactions, and therefore the spectrum of our system is defined by the magnetic field, while the *e–h* interaction is a small perturbation. This result agrees with [24], where a system of spatially separated electrons and holes under strong magnetic fields was considered. In this case, the frequencies of the acoustic modes decrease (see the inset to Fig. 4) because of magnetic localization electrons and holes in the layers. This behavior is similar to one of the acoustic modes of a Wigner crystal (see, e.g., [25]), with the dispersion law in the region of strong magnetic fields of the form

$$\omega_{-} \approx \frac{\alpha^{1/2} s k^{3/2}}{\omega_e}, \tag{13}$$

where $\alpha \equiv 2\pi N_e e^2/\epsilon m_e$ and $s^2 = 4e^2/m_e \epsilon a$.

We now discuss the phase transition between the dipole crystal and the electron–hole liquid at which the long-range order in the electron–hole system disappears. More precisely, we are interested in the change of the mode properties under the phase transition. The acoustic transverse mode, therefore, sharply disappears when the long-range order in the electron–hole system vanishes, and a damping transverse mode appears at wavevectors $k \sim 1/L$, where $L$ is the radius of the short-range crystal order (see, e.g., [26]). The optical transverse mode becomes nondispersive at small wavevectors $k$. Two other modes (longitudinal optical and acoustic ones) have the following dispersion laws at zero magnetic field:

$$\omega_{op}^2 = \omega_0^2 + \frac{2\pi e^2 N_{\text{exc}}}{\epsilon\mu} k, \tag{14}$$

$$\omega_{ac}^2 = \frac{4\pi e^2 N_{\text{exc}}^{3/2}}{\epsilon M} (kD)^2. \tag{15}$$

This result was obtained for a continuous medium (that is, in the long-wavelength limit $k[N_{\text{exc}}]^{-1/2} \ll 1$) and for a small distance between the layers $D \ll [N_{\text{exc}}]^{-1/2}$. The same result can be obtained by expanding in the small parameter $k$ ($ka \ll 1$) and omitting the summation over the shift coordinates in expressions (10) and (11), which formally means the absence of shear modulus. As follows from formulas (14), (15), and (3), two optical branches go up as the distance between layers decreases, but the acoustic mode goes down.

Transition from the interacting indirect excitons to the electron–hole plasma occurs upon further increase in the concentration of spatially separated electrons and holes [27, 28]. Characteristic values of the concentration at which such a phase transition occurs are on the order of $10^9$ cm$^{-2}$. Changes in the collective excitation spectra take on the character of a principle at this transition. Systems of the electron–hole plasma have two collective modes (we neglect the tunneling between layers). One of them is optical, with the square-root dispersion law ($\omega_{op}^{pl} \propto \sqrt{k}$), and the other is acoustic, with the linear law ($\omega_{ac}^{pl} \propto k$) [29, 30]. This means that the frequency of out-of-phase oscillations at $k = 0$ is equal to zero, in contrast to the exciton phase of the system, where out-of-phase branches start from the dipole transition energy $\omega_0$. In a perpendicular magnetic field, the hybrid magnetoplasma excitations are described as

$$\omega_{\text{mp}}^{op} = \sqrt{\omega_{op}^{pl^2} + \omega_e^2}, \tag{16}$$

$$\omega_{\text{mp}}^{ac} = \sqrt{\omega_{ac}^{pl^2} + \omega_h^2}. \tag{17}$$

We can see that such field dependences differ from those of the dipole crystal (see Fig. 4).

In summary, in this work, the properties of a dipole crystal, a new exciton phase in a two-layer electron–hole system with low densities of spatially separated electrons and holes, are investigated. Collective modes of the two-dimensional exciton triangular lattice at zero magnetic field are found. The spectrum consists of four branches: two optical (longitudinal and transverse) and two acoustic (longitudinal and transverse). The frequencies of optical modes at zero wave vector are nonzero (in contrast to a Wigner crystal), and the frequencies of the optical modes for $D \ll [N_{\text{exc}}]^{-1/2}$ purely in the two-dimensional case are $\omega_0 = \sqrt{e^2/\epsilon\mu D^3}$.

The spectra of in-phase and out-of-phase vibrations in the normal magnetic field are also found. In strong

magnetic fields, the optical modes approach the electron and hole cyclotron modes, and the acoustic ones decrease as the field increases.

We have discussed changes in the collective excitations of a dipole crystal at the transition to the electron–hole plasma phase. Plasma and magnetoplasma vibrations of that phase are considered in the system of two parallel infinite planes.

## ACKNOWLEDGMENTS

The authors are grateful to I.V. Kukushkin for useful discussion of the results. This work was supported by grants from the Russian Foundation for Basic Research and INTAS, as well as by the Presidential Grant for Scientific Schools and by the Program of the Russian Ministry of Science and Technology.

## REFERENCES

1. Yu. E. Lozovik and V. I. Yudson, JETP Lett. **22** (11), 274 (1975); Sov. Phys. JETP **44**, 389 (1976).
2. Yu. E. Lozovik and O. L. Berman, JETP **84**, 1027 (1997); Phys. Solid State **40**, 1228 (1998); J. Phys. C **14**, 12457 (2002).
3. Yu. E. Lozovik and I. V. Ovchinnikov, JETP Lett. **79** (2), 76 (2004).
4. A. V. Klyuchnik and Yu. E. Lozovik, Zh. Éksp. Teor. Fiz. **76**, 670 (1979) [Sov. Phys. JETP **49**, 335 (1979)]; Yu. E. Lozovik and V. I. Yudson, JETP Lett. **25**, 14 (1977); Yu. E. Lozovik and M. Willander, Appl. Phys. A **71**, 379 (2000).
5. Yu. E. Lozovik and A. M. Ruvinskiĭ, Zh. Éksp. Teor. Fiz. **112**, 1791 (1997) [JETP **85**, 979 (1997)].
6. Yu. E. Lozovik, O. L. Berman, and V. G. Tsvetus, Phys. Rev. B **59**, 5627 (1999).
7. L. V. Keldysh and Yu. V. Kopaev, Sov. Phys. Solid State **6**, 2791 (1964); L. V. Keldysh and A. N. Kozlov, Zh. Éksp. Teor. Fiz. **54**, 978 (1968) [Sov. Phys. JETP **27**, 521 (1968)]; B. I. Halperin and T. M. Rice, Solid State Phys. **21**, 115 (1968).
8. S. V. Tovstonog, L. V. Kulik, I. V. Kukushkin, *et al.*, Phys. Rev. B **66**, 241308 (2002).
9. M. T. Bootsmann, C. M. Hu, Ch. Heyn, *et al.*, Phys. Rev. B **67**, 121309(R) (2003).
10. A. V. Larionov, V. B. Timofeev, I. Hvam, and K. Soerensen, JETP Lett. **75**, 200 (2002).
11. L. V. Butov, L. S. Levitov, A. V. Mintsev, *et al.*, Phys. Rev. Lett. **92**, 117404 (2004); L. V. Butov, C. W. Lai, D. S. Chemla, *et al.*, Phys. Rev. Lett. **87**, 216804 (2001).
12. D. W. Snoke, S. Denev, Y. Liu, *et al.*, Nature **418**, 754 (2002).
13. Y. B. Vasilyev, V. A. Solov'ev, B. Y. Mel'tser, *et al.*, Solid State Commun. **124** (9), 323 (2002).
14. S. Charbonneau, M. L. W. Thewalt, E. S. Koteles, and B. Elman, Phys. Rev. B **38**, 6287 (1988).
15. J. E. Golub, P. F. Liao, D. J. Eilenberger, *et al.*, Solid State Commun. **72**, 372 (1989).
16. J. E. Golub, K. Kash, J. P. Harbison, and L. T. Florez, Phys. Rev. B **41**, 8564 (1990).
17. R. P. Leavitt and J. W. Little, Phys. Rev. B **42**, 11784 (1990).
18. M. M. Dignam and J. E. Sipe, Phys. Rev. B **43**, 4084 (1991).
19. F. M. Peeters and J. E. Golub, Phys. Rev. B **43**, 5159 (1991).
20. G. W. Bryant, Phys. Rev. B **46**, 1893 (1992).
21. Yu. E. Lozovik and E. A. Rakoch, Phys. Lett. A **235**, 55 (1997).
22. Z. Donkó, G. J. Kalman, P. Hartmann, *et al.*, Phys. Rev. Lett. **90**, 226 804 (2003).
23. V. I. Falko, Phys. Rev. B **49**, 7774 (1994).
24. Yu. E. Lozovik and A. M. Ruvinskiĭ, Zh. Éksp. Teor. Fiz. **112**, 1791 (1997) [JETP **85**, 979 (1997)].
25. A. V. Chaplik, Zh. Éksp. Teor. Fiz. **62**, 746 (1972) [Sov. Phys. JETP **35**, 395 (1972)]; Yu. E. Lozovik and V. I. Yudson, JETP Lett. **22**, 11 (1975); Yu. E. Lozovik, D. R. Musin, and V. I. Yudson, Solid State Phys. **21**, 1132 (1979); Yu. E. Lozovik, V. M. Farztdinov, and B. Abdullaev, J. Phys. C **18**, L807 (1985).
26. D. R. Nelson and J. M. Kosterlitz, Phys. Rev. Lett. **39**, 1201 (1977).
27. D. V. Kulakovskiĭ and Yu. E. Lozovik, JETP **98**, 1205 (2004); D. V. Kulakovskiĭ, S. I. Gubarev, and Yu. E. Lozovik, Pis'ma Zh. Éksp. Teor. Fiz. **74**, 123 (2001) [JETP Lett. **74**, 118 (2001)]; Zh. Éksp. Teor. Fiz. **121**, 915 (2002) [JETP **94**, 785 (2002)].
28. S. I. Gubarev, I. V. Kukushkin, S. V. Tovstonog, *et al.*, JETP Lett. **72**, 460 (2000).
29. R. Z. Vitlina and A. V. Chaplik, Sov. Phys. JETP **54**, 536 (1981).
30. S. Das Sarma and A. Madhukar, Phys. Rev. B **23**, 805 (1981); S. Das Sarma, Phys. Rev. B **28**, 2240 (1983).

*Journal of Experimental and Theoretical Physics, Vol. 99, No. 4, 2004, pp. 856–874.*
*Translated from Zhurnal Éksperimental'noĭ i Teoreticheskoĭ Fiziki, Vol. 126, No. 4, 2004, pp. 986–1005.*


**MISCELLANEOUS**

# Quantization of the Potential Amplitude in the One-Dimensional Schrödinger Equation

**B. Ya. Balagurov**

*Emanuel Institute of Biochemical Physics, Russian Academy of Sciences, Moscow, 119991 Russia*
*e-mail: balagurov@deom.chph.ras.ru*

Received March 1, 2004

**Abstract**—A consistent scheme is proposed for quantizing the potential amplitude in the one-dimensional Schrödinger equation in the case of negative energies (lying in the discrete-spectrum domain). The properties of the eigenfunctions $\varphi_n(x)$ and eigenvalues $\alpha_n$ corresponding to zero, small, and large absolute values of energy $E < 0$ are analyzed. Expansion in the set $\{\varphi_n(x)\}$ is used to develop a regular perturbation theory (for $E < 0$), and a general expression is found for the Green function associated with the time-independent Schrödinger equation. A similar method is used to solve several physical problems: the polarizability of a weakly bound quantum-mechanical system, the two-center problem, and the tunneling of slow particles through a potential barrier (or over a potential well). In particular, it is shown that the transmission coefficient for slow particles is anomalously large (on the order of unity) in the case of an attractive potential is characterized by certain critical values of well depth. The proposed approach is advantageous in that it does not require the use of continuum states. © *2004 MAIK "Nauka/Interperiodica".*

## 1. INTRODUCTION

One of the basic problems in nonrelativistic quantum mechanics is the determination of the spectrum of energy states for a particle in a prescribed potential (see [1, 2]). Various characteristics of this quantum-mechanical system can be analyzed by solving the problem, including its response to external perturbations (e.g. polarizability). Accordingly, a considerable part of the formal apparatus of quantum mechanics relies on perturbation theory (see [1, 2]), i.e., on expansion of the desired wavefunction in a complete set of "zeroth-order" (unperturbed) wavefunctions. Perturbation methods offer regular schemes for solving various problems. These schemes are particularly simple and instrumental when the prescribed potential function is such that the problem has only bound-state solutions (as in the harmonic oscillator problem).

However, a totally different situation arises in the case of a realistic potential vanishing at infinity (see Fig. 1), when the problem involves a continuous spectrum of energy $E$: the bound-state wavefunctions (corresponding to $E < 0$) must be combined with the continuum wavefunctions (corresponding to $E > 0$) to obtain a complete basis (see [1, 2]). Therefore, states of both types must be taken into account in perturbation theory, which substantially complicates the problem. It would obviously be advantageous to find an approach that can be used to solve problems concerning negative-energy states without invoking continuum states.

In the approach of this kind developed this paper, a feasible quantization scheme is applied to the amplitude of the potential (i.e., the well depth). The idea behind the scheme can be explained as follows. The bound state corresponding to a certain energy $E < 0$ of a particle in a potential $U(x) = U^0 v(x)$ of prescribed form can be associated with an infinite set of amplitudes $U_n^0$ ($n = 0, 1, 2, \ldots$). The values of $U_n^0 = U_n^0(E)$ and the corresponding wavefunctions $\varphi_n(x)$ are treated as eigenvalues and eigenfunctions, respectively. Accordingly, an arbitrary function $\psi(x)$ corresponding to a state with $E < 0$ can be represented as an expansion in the set $\{\varphi_n(x)\}$, which is expected to be complete.

In this study, the basic properties of the eigenfunctions $\varphi_n(x)$ and the corresponding eigenvalues $\alpha_n$ are

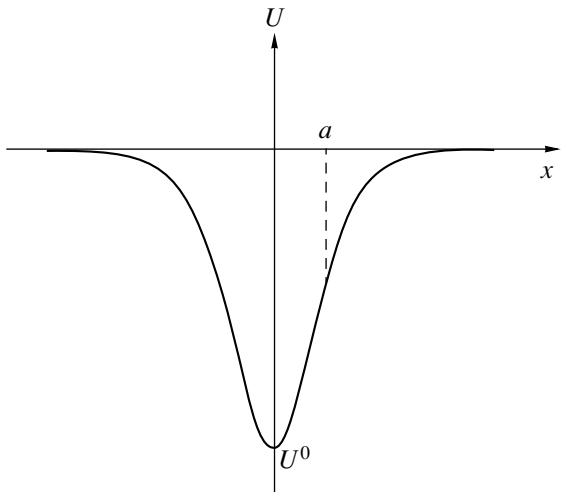



**Fig. 1.**

1063-7761/04/9904-0856$26.00 © 2004 MAIK "Nauka/Interperiodica"

analyzed for the one-dimensional Schrödinger equation. The orthonormality and completeness of $\{\varphi_n(x)\}$ are established. The asymptotic behavior of $\varphi_n(x)$ (as $x \longrightarrow \pm\infty$) is determined, including the case of $E = 0$. The properties of $\alpha_n$ and $\varphi_n(x)$ are analyzed in detail for zero, small, and large absolute values of $E$ ($E < 0$) and in the semiclassical approximation for $n \gg 1$. Expansion in the eigenfunction basis $\{\varphi_n(x)\}$ is used to develop a regular perturbation theory (for $E < 0$) without invoking continuum states. A general expression is found for the Green function associated with the time-independent Schrödinger equation (for $E < 0$).

To elucidate the scope of the proposed method, several physical problems are analyzed: the polarizability of a weakly bound quantum-mechanical system, the two-center problem, and the tunneling of slow particles through a potential barrier (or over a potential well). In each case, a consistent scheme is developed for calculating the desired physical quantities by means of expansion in the eigenfunction basis $\{\varphi_n(x)\}$. In particular, it is shown that the transmission coefficient for a particle having a small energy ($E \longrightarrow 0$) is anomalously large (on the order of unity) for certain critical values of well depth. To illustrate the key points of the proposed method, Section 12 presents an exactly solvable example in which both eigenfunctions and eigenvalues can be expressed in relatively simple analytical form. In this case, the completeness of $\{\varphi_n(x)\}$ is proved directly.

Note that quantization has been applied previously to potentials in some particular cases. For example, the Sturm functions corresponding to the three-dimensional Coulomb potential were analyzed and used to solve some quantum-mechanical problems in [3–5]. However, the analysis presented in [3–5] does not offer a general approach to the problem in the case of an arbitrary potential function. Note also that a somewhat analogous scheme for "quantizing" the permittivities of macroscopic bodies relies on the eigenfunction method proposed in [6, 7] for various problems in electrostatics. Some techniques used in [6, 7] may prove helpful in quantizing the amplitudes of finite-range potentials.

## 2. EIGENFUNCTION BASIS

Consider a particle in a one-dimensional potential of the kind illustrated by Fig. 1. If the potential function $U(x)$ is represented as $U(x) = U^0 v(x)$, where $U^0$ is the potential amplitude ($U^0 < 0$) and $v(x)$ ($0 \le v(x) \le 1$) defines its shape (see Fig. 2), then the time-independent Schrödinger equation can be written as

$$\psi''(x) + \varepsilon\psi(x) = \alpha v(x)\psi(x),$$
$$\varepsilon = \frac{2mE}{\hbar^2}, \quad \alpha = \frac{2mU^0}{\hbar^2}. \tag{1}$$

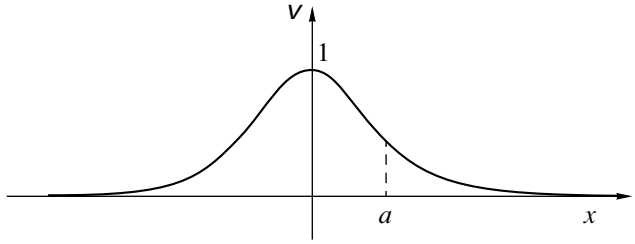



**Fig. 2.**

Suppose that $E$ is negative, i.e., $\varepsilon < 0$. Let $v(x)$ be rapidly (say, exponentially) decreasing as $x \longrightarrow \pm\infty$. The wavefunction $\psi(x)$ and its derivative are subject to standard conditions, i.e., bounded and continuous at any $x$ and vanish as $x \longrightarrow \pm\infty$. In the conventional quantization scheme [1], these conditions are used to determine the bound-state energy levels $E_n$ and wavefunctions $\psi_n(x)$.

However, an alternative approach can be developed in which the quantized parameter in Eq. (1) is $\alpha$. Accordingly, the eigenvalues are discrete values $\alpha_n$ ($n = 0, 1, 2, \ldots$) of $\alpha$ (i.e., the potential amplitude). The corresponding eigenfunctions $\varphi_n(x)$ satisfy the equation

$$\varphi_n''(x) + \varepsilon\varphi_n(x) = \alpha_n v(x)\varphi_n(x). \tag{2}$$

In this approach, the energy $\varepsilon$ is a parameter: $\alpha_n = \alpha_n(\varepsilon)$ and $\varphi_n = \varphi_n(\varepsilon; x)$. As noted in the Introduction, the set $\{\varphi_n(x)\}$ can be used to solve various quantum-mechanical problems (for $E < 0$). Therefore, the basic properties of $\varphi_n(x)$ and $\alpha_n$ should be examined.

First of all, it follows from (2) that $\varphi_n(x) \propto \exp(-\kappa|x|)$ as $x \longrightarrow \pm\infty$ ($\kappa = \sqrt{-\varepsilon}$). Furthermore, multiplying (2) by the complex conjugate $\varphi_n^*(x)$ and integrating the result with respect to $x$ from $-\infty$ to $+\infty$, we obtain

$$\alpha_n = -\left\{\int_{-\infty}^{+\infty} |\varphi_n'(x)|^2 dx - \varepsilon\int_{-\infty}^{+\infty} |\varphi_n(x)|^2 dx\right\} \times \left[\int_{-\infty}^{+\infty} |\varphi_n(x)|^2 v(x) dx\right]^{-1}. \tag{3}$$

It follows from (3) that $\alpha_n$ are real numbers that are negative if $\varepsilon < 0$ and $v(x) \ge 0$:

$$\alpha_n < 0.$$

Since $\alpha_n$ are real, it can be assumed from here on that $\varphi_n(x)$ are real.

After the changes $\varepsilon \longrightarrow \tilde{\varepsilon}$ and $n \longrightarrow m$ in (2), the equation for the corresponding $\tilde{\varphi}_m(x)$ has the form

$$\tilde{\varphi}_m''(x) + \tilde{\varepsilon}\tilde{\varphi}_m(x) = \tilde{\alpha}_m v(x)\tilde{\varphi}_m(x), \quad \tilde{\alpha}_m = \alpha_m(\tilde{\varepsilon}).$$

Multiplying this equation by $\varphi_n(x)$ and Eq. (2) by $\tilde{\varphi}_m(x)$, subtracting one from the other, and integrating the result over the $x$ axis, we obtain

$$\begin{aligned}(\alpha_n - \tilde{\alpha}_m)&\int_{-\infty}^{+\infty} \varphi_n(x)\tilde{\varphi}_m(x)v(x)dx \\ &= (\varepsilon - \tilde{\varepsilon})\int_{-\infty}^{+\infty} \varphi_n(x)\tilde{\varphi}_m(x)dx.\end{aligned} \tag{4}$$

Setting $\tilde{\varepsilon} = \varepsilon$ in (4), we find that the functions $\varphi_n(x)$ and $\varphi_m(x)$ are orthogonal with the weight function $v(x)$ if $\alpha_n \neq \alpha_m$. Accordingly, the orthonormality condition for $\{\varphi_n(x)\}$ is

$$\int_{-\infty}^{+\infty} \varphi_n(x)\varphi_m(x)v(x)dx = \delta_{nm}, \tag{5}$$

where $\varphi_n(x)$ and $\varphi_m(x)$ correspond to the same energy.

It can be argued that the eigenfunction set $\{\varphi_n(x)\}$ is complete for a broad class of potential functions (see below). Then, an arbitrary function $f(x)$ can be expanded in the basis $\{\varphi_n(x)\}$, and the coefficients $f_n$ in the expansion can be determined by using orthonormality condition (5):

$$\begin{aligned}f(x) &= \sum_n f_n\varphi_n(x), \\ f_n &= \int_{-\infty}^{+\infty} f(x)\varphi_n(x)v(x)dx.\end{aligned} \tag{6}$$

This expansion converges to $f(x)$ only if the completeness condition

$$v(x)\sum_n \varphi_n(x)\varphi_n(x') = \delta(x - x') \tag{7}$$

holds for $\{\varphi_n(x)\}$, where $\varphi_n(x)$ and $\varphi_n(x')$ correspond to the same energy. Note that (7) is valid for finite-range potentials only for $x$ and $x'$ such that $v(x) \neq 0$ and $v(x') \neq 0$.

Setting $m = n$ in (4), taking the limit as $\tilde{\varepsilon} \longrightarrow \varepsilon$, and using (5), we obtain

$$\int_{-\infty}^{+\infty} [\varphi_n(x)]^2 dx = \frac{d\alpha_n(\varepsilon)}{d\varepsilon}. \tag{8}$$

Hence,

$$\frac{d\alpha_n}{d\varepsilon} > 0. \tag{9}$$

By virtue of (5) and (8), it follows from (3) that

$$\int_{-\infty}^{+\infty} [\varphi_n'(x)]^2 dx = \varepsilon\frac{d\alpha_n(\varepsilon)}{d\varepsilon} - \alpha_n(\varepsilon),$$

which entails the inequality

$$\varepsilon\frac{d\alpha_n}{d\varepsilon} > \alpha_n. \tag{10}$$

Let us introduce the zeroth-order Green function $G_0(x - x')$:

$$\begin{aligned}\frac{d^2}{dx^2}G_0(x - x') + \varepsilon G_0(x - x') &= \delta(x - x'), \\ G_0(x - x') = -\frac{1}{2\kappa}\exp(-\kappa|x - x'|), &\quad \kappa = \sqrt{-\varepsilon}.\end{aligned} \tag{11}$$

With the use of (11), Eq. (2) can be rewritten in integral form as

$$\varphi_n(x) = -\frac{\alpha_n}{2\kappa}\int_{-\infty}^{+\infty} \exp(-\kappa|x - x'|)\varphi_n(x')v(x')dx', \tag{12}$$

which yields an asymptotic expression for $\varphi_n(x)$:

$$\begin{aligned}x \longrightarrow \pm\infty: \quad \varphi_n(x) &\approx -\frac{\alpha_n u_n(\pm)}{2\kappa}\exp(-\kappa|x|), \\ u_n(\pm) &= \int_{-\infty}^{+\infty} e^{\pm\kappa x}\varphi_n(x)v(x)dx.\end{aligned} \tag{13}$$

Note that (12) is a homogeneous Fredholm integral equation of the second kind with a polar kernel. It can be transformed into an equation with a symmetric kernel (for $v(x) > 0$) by changing to the functions

$$\Omega_n(x) = \sqrt{v(x)}\varphi_n(x). \tag{14}$$

As a result, (12) becomes

$$\Omega_n(x) = \mu_n\int_{-\infty}^{+\infty} K(x, x')\Omega_n(x')dx'; \quad \mu_n = -\alpha_n,$$

where

$$K(x, x') = \frac{1}{2\kappa}\sqrt{v(x)}\exp(-\kappa|x - x'|)\sqrt{v(x')} \tag{15}$$

is a symmetric and positive definite kernel. Under certain conditions, the eigenvalues of the corresponding integral equation are positive (e.g., see [8]): $\mu_n > 0$ (in

which case $\alpha_n < 0$). Moreover, the corresponding eigenfunctions constitute an orthonormal and complete set $\{\Omega_n(x)\}$:

$$\int_{-\infty}^{+\infty} \Omega_n(x)\Omega_m(x)dx = \delta_{nm},$$
$$\sum_n \Omega_n(x)\Omega_n(x') = \delta(x-x'). \tag{16}$$

It follows from (16) and (14) that the set $\{\varphi_n(x)\}$ satisfies the orthonormality and completeness conditions given by (5) and (7), respectively.

Combining (13) with (6), we find that $u_n(\pm)$ are the coefficients in the expansion of $f(x) = \exp(\pm\kappa x)$ in the basis $\{\varphi_n(x)\}$:

$$e^{\pm\kappa x} = \sum_n u_n(\pm)\varphi_n(x),$$

where both $u_n(\pm)$ and $\varphi_n(x)$ correspond to $\varepsilon = -\kappa^2$. As $x \longrightarrow -\infty$, the expansion of $\exp(\kappa x)$ combined with (13) yields

$$-\frac{1}{2\kappa}\sum_n \alpha_n u_n(+)u_n(-) = 1.$$

Replacing $x'$ with $x''$ in Eq. (12), multiplying the equation by $(\alpha_n)^{-1}\varphi_n(x')$, summing the results over $n$, and using (7), we obtain

$$G_0(x-x') = -\frac{1}{2\kappa}\exp(-\kappa|x-x'|) = \sum_n \frac{\varphi_n(x)\varphi_n(x')}{\alpha_n}. \tag{17}$$

Expansion (17) of $f(x) = G_0(x-x')$ in the basis $\{\varphi_n(x)\}$ can also be obtained by using (6). The corresponding coefficients $f_n$ are calculated by using (2) and (11).

If $x' = 0$, then expansion (17) reduces to

$$\exp(-\kappa|x|) = -2\kappa\sum_n \frac{\varphi_n(0)}{\alpha_n}\varphi_n(x).$$

As $x \longrightarrow \pm\infty$, we can use (13) to obtain

$$\sum_n \varphi_n(0)u_n(\pm) = 1.$$

Setting $x = x'$ in (17), we have the relation

$$\sum_n \frac{1}{\alpha_n}[\varphi_n(x)]^2 = -\frac{1}{2\kappa},$$

which is valid at any $x$. Multiplying it by $v(x)dx$, integrating the result over the $x$ axis, and using orthonormality condition (5), we obtain the "sum rule"

$$\sum_n \frac{1}{\alpha_n} = -\frac{a}{\kappa}, \quad a = \frac{1}{2}\int_{-\infty}^{+\infty} v(x)dx. \tag{18}$$

## 3. BOUND STATES

Let us apply the proposed approach to find the discrete energy spectrum for a particle in a potential $\alpha v(x)$, where $\alpha$ is a prescribed amplitude ($\alpha < 0$). Expand the bound-state wavefunction $\psi(x)$ in the eigenfunction basis $\{\varphi_m(x)\}$ corresponding to this potential function:

$$\psi(x) = \sum_m C_m\varphi_m(x).$$

Substitute this expression into (1) and use (2) to obtain

$$\sum_m C_m(\alpha_m - \alpha)v(x)\varphi_m(x) = 0.$$

Multiplied by $\varphi_n(x)$ and integrated over the $x$ axis, this relation reduces to $C_n(\alpha_n - \alpha) = 0$ Therefore, $C_n \neq 0$ if

$$\alpha_n(\varepsilon) = \alpha \tag{19}$$

and the remaining $C_m$ (with $m \neq n$) vanish. Relations (19) (with $n = 0, 1, 2, \ldots$) provide equations for the bound-state energies $\varepsilon_n = \varepsilon_n(\alpha)$.

First of all, note that the inequality $\alpha_n < 0$ implies that Eqs. (19) have solutions only if $\alpha < 0$ (i.e., if the potential is attractive), which is obvious from physical considerations. Furthermore, since (9) implies that $|\alpha_n|$ is a monotone increasing function of $|\varepsilon|$, Eq. (19) has a unique solution $\varepsilon_n = \varepsilon_n(\alpha)$ for each particular $n$. Therefore, the wavefunction $\psi_n(x)$ of a nondegenerate bound-state with energy $\varepsilon_n$ normalized to unity is

$$\psi_n(x) = C_n\varphi_n(\varepsilon_n; x),$$
$$C_n = \left(\left[\frac{d\alpha_n}{d\varepsilon}\right]_{\varepsilon=\varepsilon_n}\right)^{-1/2} = \left[\frac{d\varepsilon_n(\alpha)}{d\alpha}\right]^{1/2}. \tag{20}$$

Expression (20) is derived by using (8) and the relation

$$\left[\frac{d\alpha_n}{d\varepsilon}\right]_{\varepsilon=\varepsilon_n}\left[\frac{d\varepsilon_n}{d\alpha}\right] = 1,$$

which follows from (19). As $|x| \longrightarrow \pm\infty$, we can combine (13) with (20) to find an asymptotic expres-

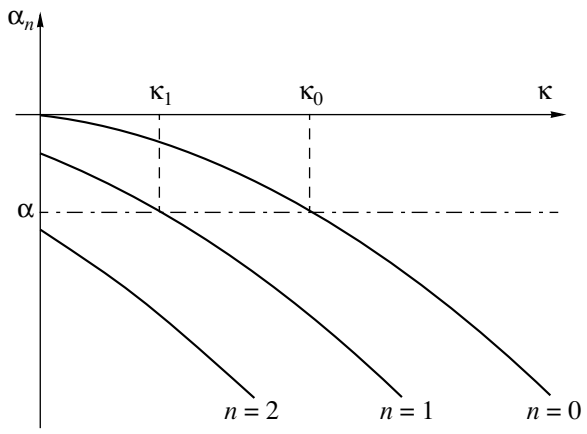



**Fig. 3.**

sion for $\psi_n(x)$:

$$\psi_n(x) \approx A_n(\pm)e^{-\kappa_n|x|},$$
$$A_n(\pm) = -\left[C_n\frac{\alpha_n u_n(\pm)}{2\kappa}\right]_{\varepsilon=\varepsilon_n}, \quad (21)$$

where $\kappa_n = \sqrt{-\varepsilon_n}$.

Figure 3 is a schematic illustration of the first three eigenvalues $\alpha_n$ ($n = 0, 1, 2$) as functions of $\kappa = \sqrt{-\varepsilon}$. (The fact that $\alpha_0(\varepsilon)$ vanishes as $\varepsilon \longrightarrow 0$ is specific to the one-dimensional case; see Sections 4 and 5.) In the case illustrated by Fig. 3, Eqs. (19) have two solutions (corresponding to $n = 0$ and $n = 1$): $\varepsilon_0 = -\kappa_0^2$ and $\varepsilon_1 = -\kappa_1^2$. The state corresponding to $n = 2$ is a virtual level that becomes a real one as $|\alpha|$ is slightly increased.

If there exist several bound states in a potential well and $\varepsilon_n \neq \varepsilon_m$ (while $\alpha_n(\varepsilon_n) = \alpha_m(\varepsilon_m) = \alpha$), then it follows from (4) that the corresponding wavefunctions given by (20) are mutually orthogonal. Therefore, the set $\{\psi_n(x)\}$ satisfies the standard orthonormality condition

$$\int_{-\infty}^{+\infty}\psi_n(x)\psi_m(x)dx = \delta_{nm}.$$

However, the wavefunctions $\psi_n(x)$ do not constitute a complete set, in contrast to $\{\varphi_n(x)\}$.

## 4. STATES WITH ZERO ENERGY

Before proceeding to the important case of small energies, the basic properties of states with $\varepsilon = 0$ should be established.

**4.1.** The eigenfunctions $\zeta_n(x) = \varphi_n(0; x)$ and eigenvalues $\lambda_n = \alpha_n(0)$ corresponding to zero-energy states satisfy the equation obtained from (2) as $\varepsilon \longrightarrow 0$:

$$\zeta_n''(x) = \lambda_n v(x)\zeta_n(x). \quad (22)$$

Assume that $\zeta_n'(x) \longrightarrow 0$ as $|x| \longrightarrow \infty$. Under this assumption, a standard analysis can be applied to (22) to show that $\zeta_n(x)$ and $\zeta_m(x)$ are orthogonal with the weight function $v(x)$ if $\lambda_n \neq \lambda_m$. Therefore, an orthonormality condition analogous to (5) is valid:

$$\int_{-\infty}^{+\infty}\zeta_n(x)\zeta_m(x)v(x)dx = \delta_{nm}. \quad (23)$$

Moreover, this assumption implies that $\zeta_n(x)$ does not vanish at infinity: $\zeta_n(x) \longrightarrow \zeta_n(\pm\infty) \neq 0$ as $x \longrightarrow \pm\infty$. It is also natural to assume that completeness condition (7) valid for $\varepsilon \neq 0$ holds as $\varepsilon \longrightarrow 0$:

$$v(x)\sum_n \zeta_n(x)\zeta_n(x') = \delta(x - x'). \quad (24)$$

Note that the potential amplitudes $U_n^0$ corresponding to $\lambda_n$ are the critical values of well depth at which zero-energy states appear (or disappear).

Define the zeroth-order Green function for $\varepsilon = 0$ as $g_0(x - x') = |x - x'|/2$ (satisfying Eq. (11) with $\varepsilon = 0$). It can be used to write Eq. (22) in integral form:

$$\zeta_n(x) = \lambda_n\int_{-\infty}^{+\infty}\frac{|x - x'|}{2}\zeta_n(x')v(x')dx' + \frac{1}{2}[\zeta_n(+\infty) + \zeta_n(-\infty)]. \quad (25)$$

Note that two possibilities arise as $\kappa \longrightarrow 0$ in Eq. (12). In particular, the eigenvalue associated with a certain state (interpreted as a ground one and assigned $n = 0$) may scale linearly with $\kappa$ as $\kappa \longrightarrow 0$ ($\lambda_0 = 0$). This agrees with the well-known fact that there exists a bound state of a one-dimensional system even if the attractive potential is infinitesimally weak (see [1]). By virtue of (12), (22), and (25), the ground-state wavefunction is a constant (independent of $x$):

$$\zeta_0(x) = \zeta_0 = \frac{1}{\sqrt{2a}}, \quad \lambda_0 = 0. \quad (26)$$

The function (26) is normalized by (23), and $a$ is given by (18). In the linear approximation in $\kappa$, the bound-state eigenvalue can be determined from (12): $\alpha_0^{(1)} = -\kappa/a$ (see Section 5 for details).

The other possibility arises for $n \neq 0$, when $\lambda_n = \alpha_n(0) \neq 0$. In this case, as $\kappa \longrightarrow 0$, the limit of (12) is well defined only if

$$\int_{-\infty}^{+\infty} \zeta_n(x) v(x) dx = 0, \quad n \neq 0.$$

This relation is satisfied automatically, because it means that $\zeta_n(x)$ (with $n \neq 0$) is orthogonal to $\zeta_0 = \text{const}$, in accordance with (23). It can be shown that the correct limit of (12) as $\kappa \longrightarrow 0$ is Eq. (25).

Setting $x = 0$ in (25) and using the resulting equation to eliminate the term $[\zeta_n(+\infty) + \zeta_n(-\infty)]/2$, we rewrite (25) as

$$\zeta_n(x) - \zeta_n(0) = \lambda_n \int_{-\infty}^{+\infty} \frac{1}{2}[|x - x'| - |x| - |x'|]\zeta_n(x') v(x') dx',$$

where $|x|/2$ is added to symmetrize the kernel without changing the integral.

**4.2.** By virtue of (23) and (24), any function $f(x)$ can be represented as (6) with $\varphi_n(x)$ replaced by $\zeta_n(x)$. In particular, $f(x) = x$ and $f(x) = |x|$ are expanded as follows:

$$x = \frac{1}{2a}\int_{-\infty}^{+\infty} t v(t) dt + \sum_{n \neq 0} \frac{1}{\lambda_n}[\zeta_n(-\infty) - \zeta_n(+\infty)]\zeta_n(x), \tag{27}$$

$$|x| = \frac{1}{2a}\int_{-\infty}^{+\infty} |t| v(t) dt - \sum_{n \neq 0} \frac{1}{\lambda_n}[\zeta_n(-\infty) + \zeta_n(+\infty) - 2\zeta_n(0)]\zeta_n(x). \tag{28}$$

Here, the integral terms correspond to $n = 0$, and the expansion coefficients corresponding to $n \neq 0$ are calculated by multiplying Eq. (22) by $x$ or $|x|$ and integrating the result over the $x$ axis.

Combining (27) with (28), we have

$$\frac{x - |x|}{2} = \frac{1}{4a}\int_{-\infty}^{+\infty} (t - |t|) v(t) dt + \sum_{n \neq 0} \frac{1}{\lambda_n}[\zeta_n(-\infty) - \zeta_n(0)]\zeta_n(x), \tag{29}$$

where the left-hand side vanishes if $x > 0$. In particular, as $x \longrightarrow +\infty$, (29) yields

$$\frac{1}{4a}\int_{-\infty}^{+\infty} (t - |t|) v(t) dt + \sum_{n \neq 0} \frac{1}{\lambda_n}[\zeta_n(-\infty) - \zeta_n(0)]\zeta_n(+\infty) = 0. \tag{30}$$

Setting $x = 0$ in (29) and subtracting the resulting equation from (30), we obtain

$$\sum_{n \neq 0} \frac{1}{\lambda_n}[\zeta_n(-\infty) + \zeta_n(+\infty) - \zeta_n(0)]\zeta_n(0) = \sum_{n \neq 0} \frac{1}{\lambda_n}\zeta_n(-\infty)\zeta_n(+\infty). \tag{31}$$

In a similar manner, we can expand the function

$$f(x) \equiv \int_{-\infty}^{+\infty} |x - t| v(t) dt = \frac{1}{2a} I - 2a \sum_{n \neq 0} \frac{1}{\lambda_n}[\zeta_n(-\infty) + \zeta_n(+\infty)]\zeta_n(x), \tag{32}$$

$$I = \int_{-\infty}^{+\infty}\int_{-\infty} |t - t'| v(t) v(t') dt dt'.$$

Setting $x = 0$ in (32) and using (28) with $x = 0$ to eliminate the integral on the left-hand side of the resulting equation, we obtain

$$\sum_{n \neq 0} \frac{1}{\lambda_n}\zeta_n(-\infty)\zeta_n(+\infty) = \frac{1}{8a^2} I, \tag{33}$$

where $I$ is defined in (32).

The expansion of $f(x) = |x - x'|$ in the basis $\{\zeta_n(x)\}$ is

$$|x - x'| = \frac{1}{2a}\int_{-\infty}^{+\infty} |t - x'| v(t) dt - \sum_{n \neq 0} \frac{1}{\lambda_n}[\zeta_n(-\infty) + \zeta_n(+\infty) - 2\zeta_n(x')]\zeta_n(x).$$

When $x = x'$, it yields

$$\frac{1}{2a}\int_{-\infty}^{+\infty} |t - x| v(t) dt = \sum_{n \neq 0} \frac{1}{\lambda_n}[\zeta_n(-\infty) + \zeta_n(+\infty) - 2\zeta_n(x)]\zeta_n(x).$$

Multiplying this equation by $v(x)dx$ and integrating the

result over the $x$ axis, we find another useful relation,

$$\sum_{n\neq 0}\frac{1}{\lambda_n} = -\frac{1}{4a}I, \tag{34}$$

where $I$ is the same integral defined in (32).

**4.3.** For $n \gg 1$, a semiclassical approximation analogous to the standard one [1] can be used to calculate both $\lambda_n$ and $\zeta_n(x)$. However, the case of $\varepsilon = 0$ is special in that the turning points are located at $x = \pm\infty$; i.e., there is no "outer region." Note also that the semiclassical approximation becomes invalid as $|x| \longrightarrow \infty$, and the exact solution to Eq. (22) with $V(x)$ given by an asymptotic expression.

Let $V(x)$ be an even function characterized by the following asymptotic behavior:

$$|x| \longrightarrow \infty : V(x) \approx V_0 e^{-2\gamma|x|}, \tag{35}$$

where $V_0 \sim 1$ and $\gamma \sim 1/a$. The semiclassical expression for $\zeta_n(x)$ valid for $\exp(-\gamma|x|) \gg 1/n$ has the form

$$\zeta_n(x) = \sqrt{\frac{2}{b}}\frac{(-1)^n}{\sqrt[4]{V(x)}}\cos\left\{\sqrt{|\lambda_n|}\int_{-\infty}^{x}\sqrt{V(t)}dt + \theta\right\}. \tag{36}$$

The exact solution to Eq. (22) with potential (35) is

$$\begin{aligned}\zeta_n(x) &= \zeta_n(\pm\infty)J_0(\mu_n e^{-\gamma|x|}),\\ \mu_n &= \frac{1}{\gamma}\sqrt{|\lambda_n|V_0},\end{aligned} \tag{37}$$

where $J_0(z)$ is a Bessel function, and the plus and minus signs in $\zeta_n(\pm\infty)$ correspond to $x > 0$ and $x < 0$, respectively. The constant $b$ in (36) is defined in (38) below.

Expressions (36) and (37) must be matched at $x > 0$ and $x < 0$ such that $1/n \ll \exp(-\gamma|x|) \ll 1$. By matching them at a negative $x$, the constant phase is found: $\theta = -\pi/4$. The matching at a positive $x$ leads to the quantization rule

$$\int_{-\infty}^{+\infty}\sqrt{|\lambda_n|V(x)}dx = \left(n+\frac{1}{2}\right)\pi,$$

which yields

$$\sqrt{|\lambda_n|} = \left(n+\frac{1}{2}\right)\frac{\pi}{b},$$

and

$$\lambda_n \approx -n(n+1)\left(\frac{\pi}{b}\right)^2, \quad b = \int_{-\infty}^{+\infty}\sqrt{V(x)}dx, \tag{38}$$

where $b \sim 1/\gamma \sim a$. In (38), a term $\sim n^0$ is neglected since its exact calculation requires that a term $\sim 1/n$ be retained in the quantization rule.

The matching of (36) and (37) also leads to an expression for $\zeta_n(\pm\infty)$:

$$\zeta_n(+\infty) = \frac{\pi}{b}\sqrt{\frac{1}{\gamma}\left(n+\frac{1}{2}\right)},$$

$$\zeta_n(-\infty) = (-1)^n\zeta_n(+\infty).$$

Finally, since $V(0) = 1$, we have

$$\zeta_n(0) = \sqrt{\frac{2}{b}}\cos\frac{n\pi}{2}.$$

Note that the normalization integral is calculated by using expression (36), and the squared fast-oscillating cosine is replaced with 1/2 (see [1]).

When the potential function asymptotically behaves as a power,

$$|x| \longrightarrow \infty : V(x) \approx \frac{g}{|x|^\nu}$$

(where $g \sim a^\nu$ with $\nu > 2$), semiclassical approximation (36) is valid if $(a/|x|)^{(\nu-2)/2} \gg 1/n$, and the exact solution to (22) at $|x| \gg a$ is

$$\zeta_n(x) = \zeta_n(\pm\infty)\Gamma\left(\frac{\nu-1}{\nu-2}\right)\left(\frac{2}{\mu_n}\right)^{1/(\nu-2)}\sqrt{|x|}\,J_\tau\left(\frac{\mu_n}{|x|^{(\nu-2)/2}}\right),$$

$$\tau = \frac{1}{\nu-2}, \quad \mu_n = \frac{2\sqrt{|\lambda_n|g}}{\nu-2},$$

where $J_\tau(z)$ is a Bessel function. Matching these solutions under the condition that $1/n \ll (a/|x|)^{(\nu-2)/2} \ll 1$, we obtain

$$\sqrt{|\lambda_n|} = \left[n+\frac{\nu}{2(\nu-2)}\right]\frac{\pi}{b}, \quad \theta = -\frac{\nu\pi}{4(\nu-2)},$$

$$\zeta_n(+\infty) = \frac{1}{\Gamma\left(\frac{\nu-1}{\nu-2}\right)}\left[\frac{2\pi}{b\sqrt{g}}\right]^{1/2}\left[\frac{\sqrt{|\lambda_n|g}}{\nu-2}\right]^\rho,$$

$$\rho = \frac{\nu}{2(\nu-2)},$$

with $b$ given by (38).

## 5. STATES WITH SMALL ENERGIES

Now, let us examine the properties of $\alpha_n$ and $\varphi_n(x)$ corresponding to small energies ($\kappa a \ll 1$). Equation (2) with $\varepsilon = -\kappa^2$ and $\varphi_n(x)$ expressed as

$$\varphi_n(x) = B_n(x)e^{-\kappa|x|}, \tag{39}$$

becomes an equation for $B_n(x)$:

$$\begin{aligned}B_n''(x) - 2\kappa\,\mathrm{sgn}\,x\,B_n'(x) - \alpha_n V(x)B_n(x)\\ = 2\kappa B_n(0)\delta(x),\end{aligned} \tag{40}$$

where $\operatorname{sgn} x = +1$ if $x > 0$ and $\operatorname{sgn} x = -1$ if $x < 0$. The solution to Eq. (40) is sought as an expansion in the zero-energy eigenfunction basis:

$$B_n(x) = \sum_m b_{nm}\zeta_m(x). \tag{41}$$

Substituting (41) into Eq. (40), multiplying the equation by $\zeta_k(x)$, integrating the result from $-\infty$ to $+\infty$, and making use of (22) and (23), we obtain

$$(\alpha_n - \lambda_k)b_{nk} + 2\kappa\sum_m b_{nm}M_{mk} = 0, \tag{42}$$

where

$$\begin{aligned} M_{mn} &= \zeta_m(0)\zeta_n(0) + \int_{-\infty}^{+\infty}\zeta_m'(x)\zeta_n(x)\operatorname{sgn} x\,dx, \\ M_{nn} &= \frac{1}{2}\{[\zeta_n(-\infty)]^2 + [\zeta_n(+\infty)]^2\}, \\ M_{n0} &= \zeta_0[\zeta_n(-\infty) + \zeta_n(+\infty) - \zeta_n(0)], \quad M_{0n} = \zeta_0\zeta_n(0). \end{aligned} \tag{43}$$

When $\kappa$ is small, system (42) can be solved by a perturbation method using expansions in powers of $\kappa$:

$$\begin{aligned} b_{nm} &= \delta_{nm} + b_{nm}^{(1)} + b_{nm}^{(2)} + \ldots, \\ \alpha_n &= \lambda_n + \alpha_n^{(1)} + \alpha_n^{(2)} + \ldots. \end{aligned}$$

In the linear approximation in $\kappa$, (42) yields

$$\alpha_n^{(1)} = -2\kappa M_{nn} = -\kappa\{[\zeta_n(-\infty)]^2 + [\zeta_n(+\infty)]^2\}, \tag{44}$$

$$n \neq m\colon\ b_{nm}^{(1)} = \frac{2\kappa M_{nm}}{\lambda_m - \lambda_n} \tag{45}$$

with $M_{nm}$ given by (43). The approximation is valid if $|\alpha_n^{(1)}| \sim \kappa/a$ is much smaller than $|\lambda_n| \sim 1/a^2$, i.e., if $\kappa a \ll 1$. It follows from the results obtained in Section 4 that the corresponding condition is $\kappa a \ll n + 1/2$ for large $n$.

In the same approximation, orthonormality condition (5) yields

$$b_{nm}^{(1)} + b_{mn}^{(1)} = 2\kappa\int_{-\infty}^{+\infty}|x|\zeta_n(x)\zeta_m(x)v(x)dx, \tag{46}$$

and hence

$$b_{nn}^{(1)} = \kappa\int_{-\infty}^{+\infty}|x|[\zeta_n(x)]^2 v(x)dx.$$

The expression for $M_{nm}$ in (43) can be rewritten as

$$\begin{aligned} M_{nm} &= \frac{1}{2}[\zeta_n(-\infty)\zeta_m(-\infty) + \zeta_n(+\infty)\zeta_m(+\infty)] \\ &\quad -\frac{1}{2}(\lambda_n - \lambda_m)\int_{-\infty}^{+\infty}|x|\zeta_n(x)\zeta_m(x)v(x)dx. \end{aligned} \tag{47}$$

It can readily be shown that relation (46) holds for $b_{nm}^{(1)}$ given by (45) with $M_{nm}$ defined in (47).

In the quadratic approximation with respect to $\kappa$, we have

$$\alpha_n^{(2)} = 4\kappa^2\sum_{m\neq n}\frac{M_{nm}M_{mn}}{\lambda_n - \lambda_m}, \tag{48}$$

where $M_{nm}$ is defined in (43). Both coefficients $b_{nm}^{(2)}$ and approximations of higher orders in $\kappa$ can be calculated by a similar method.

For the ground state ($n = 0$), expression (44) combined with (26) yields

$$\alpha_0^{(1)} = -\frac{\kappa}{a}, \tag{49}$$

where $a$ is defined in (18). Solving Eq. (19) with $\alpha_0^{(1)}$ given by (49), we obtain $\varepsilon_0 = -(\alpha a)^2$. In standard units, this result is equivalent to the well-known expression

$$E_0 = -\frac{m}{2\hbar^2}\left[\int_{-\infty}^{+\infty}U(x)dx\right]^2$$

for the energy of a weakly bound state [1]. When $n = 0$, (48) yields an expression for $\alpha_0^{(2)}$ that can be rewritten, by using (31) and (33), as

$$\alpha_0^{(2)} = -\frac{2\kappa^2}{a}\sum_{n\neq 0}\frac{1}{\lambda_n}\zeta_n(-\infty)\zeta_n(+\infty) = -\frac{\kappa^2}{4a^3}I, \tag{50}$$

where $I$ is the integral defined in (32).

Note that the lowest order terms in the expansion of $\alpha_0(\kappa)$ in powers of $\kappa$ can be determined from (18). In the linear approximation in $\kappa$, (49) is obtained. In the quadratic approximation, we have

$$\alpha_0^{(2)} = \frac{\kappa^2}{a^2}\sum_{n\neq 0}\frac{1}{\lambda_n},$$

which is equivalent to (50) by virtue of (34). The third-order correction is

$$\alpha_0^{(3)} = \frac{[\alpha_0^{(2)}]^2}{\alpha_0^{(1)}} - [\alpha_0^{(1)}]^2\sum_{n\neq 0}\frac{\alpha_n^{(1)}}{\lambda_n^2},$$

where $\alpha_n^{(1)}$ is given by (44).

The quantities $u_n(\pm)$ defined in (13) are calculated, by using (39) and (41), as

$$u_n(\pm) = -\frac{2\kappa}{\alpha_n}\sum_m b_{nm}\zeta_m(\pm\infty).$$

Using results obtained above, we obtain

$$u_0^{(0)}(\pm) = \sqrt{2a},$$

$$u_0^{(1)}(+) = -u_0^{(1)}(-)$$

$$= \kappa\sqrt{2a}\sum_{n\neq 0}\frac{1}{\lambda_n}[\zeta_n(+\infty) - \zeta_n(-\infty)]\zeta_n(0),$$

$$n \neq 0\colon u_n^{(0)}(\pm) = 0, \quad u_n^{(1)}(\pm) = -\frac{2\kappa}{\lambda_n}\zeta_n(\pm\infty).$$

The expressions for $u_0^{(1)}(\pm)$ are derived by eliminating the integral contained in $b_{00}^{(1)}$ with the use of relation (28) for $x = 0$ and by taking into account (31) and (33). By virtue of relation (27) for $x = 0$, the expressions for $u_0^{(1)}(\pm)$ can be rewritten as

$$u_0^{(1)}(+) = -u_0^{(1)}(-) = \frac{\kappa}{\sqrt{2a}}\int_{-\infty}^{+\infty} x v(x)dx.$$

For an even potential function, it holds that $u_0^{(1)}(+) = u_0^{(1)}(-) = 0$.

## 6. STATES WITH LARGE $|\varepsilon|$

As $|\varepsilon| \longrightarrow \infty$, the values of $|\alpha_n|$ increase indefinitely. According to (10), the growth of $|\alpha_n|$ cannot be steeper than linear in $|\varepsilon|$. However, if $v(x)$ is finite at any $x$, then $|\alpha_n|$ cannot grow slower than $|\varepsilon|$, since a state with sufficiently large $|\varepsilon|$ cannot otherwise "fit" in the potential well. Thus, $|\alpha_n|$ scales linearly with $|\varepsilon|$ as $|\varepsilon| \longrightarrow \infty$. If the potential function has a minimum at $x = 0$ (see Fig. 1) and $v(0) = 1$, then the corresponding scaling factor is unity, and we have (for $\varepsilon = -\kappa^2$)

$$|\varepsilon| \longrightarrow \infty\colon \alpha_n = -\kappa^2(1 + \Delta_n + \ldots), \qquad (51)$$

where $|\Delta_n| \ll 1$. To simplify calculations, the analysis that follows is developed for an even potential function.

To calculate $\Delta_n$, we note that the value of $\alpha_n$ (for moderate $n$) is determined by the behavior of $v(x)$ in the neighborhood of $x = 0$ as $|\varepsilon| \longrightarrow \infty$. Suppose that, as $x \longrightarrow 0$,

$$v(x) \approx 1 - x^2\delta^2 \quad (|x| \ll 1/\delta),$$

where $\delta \sim 1/a$. Thus, the eigenvalue problem is reduced to the harmonic oscillator problem [1]. In this approximation, we have

$$\Delta_n = (2n + 1)\frac{\delta}{\kappa},$$

$$|x| \ll \frac{1}{\delta}\colon \varphi_n(x) = \left(\frac{\kappa\delta}{\pi}\right)^{1/4}\frac{1}{\sqrt{2^n n!}} \qquad (52)$$

$$\times \exp(-(\kappa\delta/2)x^2)H_n(\sqrt{\kappa\delta}x),$$

where $H_n(z)$ is a Hermitian polynomial. The approximation is valid if the correction $\Delta_n$ is much smaller than unity, i.e., if $\kappa \gg (2n + 1)\delta$ or, since $\delta \sim 1/a$: $\kappa a \gg n + 1/2$ ($n = 0, 1, 2, \ldots$).

Oscillator approximation (52) for $\varphi_n(x)$ is valid if $|x| \ll 1/\delta$. To find $\varphi_n(x)$ at $|x| > 1/\delta$ a semiclassical approximation is used ($x \gg x_0 > 0$):

$$\varphi_n(x) = \frac{D_n}{\sqrt{k(x)}}\exp\left\{-\int_{x_0}^{x} k(y)dy\right\},$$

$$k(x) = \kappa[1 - (1 + \Delta_n)v(x)]^{1/2}, \quad x_0 = \frac{\sqrt{\Delta_n}}{\delta}, \qquad (53)$$

$$D_n = \frac{1}{(2\pi)^{1/4}}\sqrt{\frac{\kappa\delta}{n!}}\left(\frac{2n+1}{2e}\right)^{(2n+1)/4},$$

where $x_0 \ll 1/\delta$, $e = 2.718\ldots$ is the base of natural logarithms, and $D_n$ is calculated by matching expressions for $\psi_n(x)$ in (52) and (53) under the condition that $x_0 \ll x \ll 1/\delta$. It should be emphasized that semiclassical approximation (53) is valid for any $n$ such that $n + 1/2 \ll \kappa/\delta$, including $n = 0$.

As $x \longrightarrow +\infty$, the expression for $\psi_n(x)$ in (53) yields (13) with

$$u_n(+) = \frac{2}{\kappa}\frac{1}{(2\pi)^{1/4}}\sqrt{\frac{\delta}{n!}}\left(\frac{2\kappa}{\delta}\right)^{(2n+1)/4}\exp\left\{\frac{\kappa}{\delta}\left(I_1 + \frac{\Delta_n}{2}I_2\right)\right\},$$

$$\frac{I_1}{\delta} = \int_0^\infty [1 - \sqrt{1 - v(x)}]dx, \qquad (54)$$

$$\frac{I_2}{\delta} = \lim_{s\to 0}\left[\frac{1}{\delta}\ln(s\delta) + \int_s^\infty \frac{v(x)dx}{\sqrt{1 - v(x)}}\right],$$

and $\Delta_n$ given by (52). For the even potential function, it holds that $u_n(-) = (-1)^n u_n(+)$.

It follows from the results obtained above that the eigenvalues $\alpha_n$ with $n \gg 1$ that correspond to the potential functions considered here should be expanded in powers of $\kappa a/n$. In particular, "small" energy means

that $\kappa a \ll n$ (including the case when $1 \ll \kappa a \ll n$). The case of an "intermediate" energy, when $\kappa a \sim n \gg 1$, can be analyzed in a semiclassical approximation analogous to the standard one [1]. The corresponding quantization rule is

$$\int_0^{x_1} \sqrt{\varepsilon - \alpha_n v(x)}\,dx = \left(n + \frac{1}{2}\right)\frac{\pi}{2},$$

where $x_1$ is the zero of the radicand.

## 7. PERTURBATION THEORY

Consider the bound state characterized by wavefunction $\psi_n^{(0)}(x)$ and energy $\varepsilon_n^{(0)}$ in a potential well with $U(x) = U^0 v(x)$. We seek $\psi_n(x)$ and $\varepsilon_n$ in the case when the Hamiltonian contains a small perturbation $V'(x)$. The equation for $\psi_n(x)$ is written as

$$\begin{aligned} \psi_n''(x) + \varepsilon_n \psi_n(x) - \alpha v(x)\psi_n(x) &= V(x)\psi_n(x), \\ V(x) &= \frac{2m}{\hbar^2}V'(x). \end{aligned} \tag{55}$$

To find a solution to Eq. (55), each $\psi_n(x)$ is expanded in the complete eigenfunction set $\{\varphi_m(x)\}$ defined in Section 2:

$$\psi_n(x) = \sum_m A_{nm}\varphi_m(x). \tag{56}$$

Each $\varphi_m(x)$ satisfies Eq. (2) with $n$ replaced with $m$ and $\varepsilon = \varepsilon_n^{(0)}$, where the energy $\varepsilon_n^{(0)}$ of the bound state is determined by solving Eq. (19). Accordingly, the expression for the bound-state wavefunction $\psi_n^{(0)}(x)$ in terms of $\varphi_n(x)$ (corresponding to $\varepsilon = \varepsilon_n^{(0)}$) is given by (20).

Substituting (56) into Eq. (55), multiplying the equation by $\varphi_k(x)$ (also corresponding to $\varepsilon = \varepsilon_n^{(0)}$), integrating the result over the $x$ axis, and using Eqs. (2) and (5), we obtain

$$\begin{aligned} (\varepsilon_n - \varepsilon_n^{(0)})\sum_m \langle k|m\rangle A_{nm} + (\alpha_k - \alpha_n)A_{nk} &= \sum_m \langle k|V|m\rangle A_{nm}, \\ \langle k|m\rangle &= \int_{-\infty}^{+\infty} \varphi_k(x)\varphi_m(x)\,dx, \\ \langle k|V|m\rangle &= \int_{-\infty}^{+\infty} \varphi_k(x)V(x)\varphi_m(x)\,dx, \end{aligned} \tag{57}$$

where $\langle n|n\rangle = (C_n)^{-2}$ by virtue of (8) and (20), and $\alpha$ is replaced with $\alpha_n = \alpha_n(\varepsilon_n^{(0)})$ by virtue of (19).

System (57) is solved by using expansions in powers of $V(x)$:

$$\begin{aligned} A_{nm} &= A_{nm}^{(0)} + A_{nm}^{(1)} + A_{nm}^{(2)} + \ldots, \\ \varepsilon_n &= \varepsilon_n^{(0)} + \varepsilon_n^{(1)} + \varepsilon_n^{(2)} + \ldots, \end{aligned}$$

where $A_{nm}^{(0)} = C_n\delta_{nm}$ with $C_n$ defined in (20) for $\varepsilon = \varepsilon_n^{(0)}$. In the linear approximation in $V(x)$, we obtain

$$\varepsilon_n^{(1)} = \frac{\langle n|V|n\rangle}{\langle n|n\rangle} = \int_{-\infty}^{+\infty} \psi_n^{(0)}(x)V(x)\psi_n^{(0)}(x)\,dx, \tag{58}$$

which is, expectedly, identical to the well-known result (see [1]). In the same approximation, (57) yields

$$n \neq m\colon A_{nm}^{(1)} = \frac{C_n}{\alpha_m - \alpha_n}[\langle m|V|n\rangle - \varepsilon_n^{(1)}\langle m|n\rangle], \tag{59}$$

where $\varepsilon_n^{(1)}$ is given by (58). In the linear approximation in $V(x)$, the normalization condition for $\psi_n(x)$ entails

$$\sum_m \langle n|m\rangle A_{nm}^{(1)} = 0,$$

which can be combined with (59) to find $A_{nn}^{(1)}$.

In the quadratic approximation in $V(x)$, (57) yields

$$\begin{aligned} \varepsilon_n^{(2)} = (C_n)^2 {\sum_m}' \frac{1}{\alpha_m - \alpha_n}[\langle n|V|m\rangle - \varepsilon_n^{(1)}\langle n|m\rangle] \\ \times [\langle m|V|n\rangle - \varepsilon_n^{(1)}\langle m|n|\rangle], \end{aligned} \tag{60}$$

where the primed sum skips the term with $m = n$. Both coefficients $A_{nm}^{(2)}$ and higher order approximations of $\varepsilon_n$ and $A_{nm}$ can be calculated by a similar method. Note that both functions $\varphi_n(x)$ and $\varphi_m(x)$ and values of $\alpha_n$ and $\alpha_m$ contained in (56)–(60) correspond to the energy $\varepsilon = \varepsilon_n^{(0)}$.

A perturbation theory can also be developed for the eigenvalues $\alpha_n$ and eigenfunctions $\varphi_n(x)$. The corresponding equation for $\varphi_n(x)$ is

$$\varphi_n''(x) + \varepsilon\varphi_n(x) - \alpha_n v(x)\varphi_n(x) = V(x)\varphi_n(x).$$

The linear and quadratic in $V(x)$ corrections to an eigenvalue are

$$\alpha_n^{(1)} = -\langle n|V|n\rangle, \quad \alpha_n^{(2)} = -{\sum_m}' \frac{\langle n|V|m\rangle\langle m|V|n\rangle}{\alpha_m^{(0)} - \alpha_n^{(0)}},$$

where $\langle n|V|m\rangle$ is defined in (57), and both the eigenvalues $\alpha_n^{(0)}$ and $\alpha_m^{(0)}$ and the eigenfunctions $\varphi_n(x)$ and $\varphi_m(x)$ are calculated for $V(x) = 0$ and an arbitrary $\varepsilon$. In

some cases, it may be easier to calculate corrections to bound-state energies by solving Eq. (19) for $\alpha_n = \alpha_n^{(0)} + \alpha_n^{(1)} + \alpha_n^{(2)} + \ldots$.

## 8. POLARIZABILITY

Consider the ground bound state of a point charge $e$ in the uniform electric field of strength $\mathscr{E}$. For this system, $V'(x) = -ex\mathscr{E}$ and $V(x) = -(2me\mathscr{E}/\hbar^2)x$, and expression (60) (with $\langle 0|x|0\rangle = 0$) can be used to find the polarizability of the system:

$$\Lambda = \frac{4me^2}{\hbar^2}(C_0)^2\sum_{n=1}^{\infty}\frac{(\langle 0|x|n\rangle)^2}{\alpha_0 - \alpha_n} \tag{61}$$

with $(C_0)^2 = d\varepsilon_0/d\alpha$ and both $\alpha_n$ and $\varphi_n(x)$ corresponding to $\varepsilon = \varepsilon_0$, where $\varepsilon_0$ is determined by solving the equation $\alpha_0(\varepsilon) = \alpha$.

Suppose that there exists only a weakly bound state of a one-dimensional quantum-mechanical system, $\varepsilon_0 = -\kappa_0^2$, where $\kappa_0 = |\alpha|a \ll 1/a$, and the corresponding wavefunction is $\varphi_0(x) \approx \zeta_0\exp(-\kappa_0|x|)$ with $\zeta_0$ given by (26). Replacing $\alpha_0$ with $\alpha$, rewrite (61) as

$$\Lambda = -\frac{4me^2}{\hbar}\frac{d\varepsilon_0(\alpha)}{d\alpha}\int_{-\infty}^{+\infty}xdx\varphi_0(x)$$
$$\times\int_{-\infty}^{+\infty}x'dx'\varphi_0(x')\sum_{n=1}^{\infty}\frac{\varphi_n(x)\varphi_n(x')}{\alpha_n - \alpha}.$$

Using representation (17) of $G_0(x - x')$, rearrange the sum contained in this expression as follows:

$$\sum_{n=1}^{\infty}\frac{\varphi_n(x)\varphi_n(x')}{\alpha_n - \alpha} = G_0(x - x')$$
$$-\frac{\varphi_0(x)\varphi_0(x')}{\alpha} + \alpha\sum_{n=1}^{\infty}\frac{\varphi_n(x)\varphi_n(x')}{\alpha_n(\alpha_n - \alpha)}.$$

The second term on the right-hand side does not contribute to $\Lambda$ (since $\langle 0|x|0\rangle = 0$), and the third term is smaller than the first one by the factor $(\kappa_0 a)^2 \ll 1$. Thus,

$$\Lambda \approx -\frac{4me^2}{\hbar^2}\frac{d\varepsilon_0(\alpha)}{d\alpha}\int_{-\infty}^{+\infty}xdx\varphi_0(x)$$
$$\times\int_{-\infty}^{+\infty}x'dx'\varphi_0(x')G_0(x - x').$$

Substituting $\varphi_0(x) = (2a)^{-1/2}\exp(-\kappa_0|x|)$, $\varepsilon_0 = -\kappa_0^2 = -(\alpha a)^2$, and $G_0(x - x') = -(2\kappa_0)^{-1}\exp(-\kappa_0|x - x'|)$ into this expression, we obtain

$$\Lambda \approx \frac{5}{4}\frac{me^2}{\hbar^2\kappa_0^4}, \tag{62}$$

which is consistent with the corresponding result obtained in the three-dimensional case (see [1, Section 76, Problem 5]).

Alternatively, expression (62) can be derived by a method similar to that applied in solving Problems 4 and 5 in Section 76 of [1]. In the linear approximation in $V(x)$, the correction $\psi^{(1)}(x)$ to the unperturbed wavefunction $\psi_0(x)$ satisfies the equation

$$\psi^{(1)''}(x) + \varepsilon_0\psi^{(1)}(x) - \alpha v(x)\psi^{(1)}(x) = V(x)\psi_0(x),$$
$$V(x) = -Bx, \quad B = \frac{2me\mathscr{E}}{\hbar^2}.$$

Here, use is made of the fact that $\varepsilon^{(1)} = 0$ if $\langle 0|x|0\rangle = 0$. Representing $\psi^{(1)}(x) = f(x)\psi_0(x)$, we obtain the following equation for $f(x)$:

$$f''(x) + 2\frac{\psi_0'(x)}{\psi_0(x)}f'(x) = -Bx.$$

For a weakly bound state, we have $\psi_0(x) \approx (\kappa_0)^{1/2} \times \exp(-\kappa_0|x|)$ and, therefore,

$$f''(x) - 2\kappa_0 f'(x)\operatorname{sgn}x = -Bx.$$

The solution to this equation of interest for the present analysis is

$$f(x) = \frac{B}{4\kappa_0}\left(x|x| + \frac{x}{\kappa_0}\right).$$

Calculating the average dipole moment $d = ex$,

$$\bar{d} \approx 2e\int_{-\infty}^{+\infty}x\psi_0(x)\psi^{(1)}(x)dx = 2e\int_{-\infty}^{+\infty}xf(x)[\psi_0(x)]^2dx,$$

we obtain (62).

## 9. TWO-CENTER PROBLEM

A special perturbation theory can be developed to find the bound states of a particle interacting with the field generated by two identical potential wells separated by a distance $l \gg a$. Denote their centers by $x_1$ and $x_2$ ($l = x_1 - x_2 > 0$). The eigenfunctions $\varphi_\nu(x)$ in the two-center problem satisfy the equation

$$\varphi_\nu''(x) + \varepsilon\varphi_\nu(x) = \alpha_\nu[v(x - x_1) + v(x - x_2)]\varphi_\nu(x). \tag{63}$$

The eigenfunctions $\varphi_n^{(i)}(x) \equiv \varphi_n(x - x_i)$ corresponding to the $i$th isolated well obey Eq. (2) with $x \longrightarrow x - x_i$.

Since both sets $\{\varphi_n^{(i)}(x)\}$ ($i = 1, 2$) are complete at any $x$ (if $v(x) \neq 0$ at any $x$), we can expand $\varphi_\nu(x)$ either in $\{\varphi_n^{(1)}(x)\}$ or in $\{\varphi_n^{(2)}(x)\}$.

However, it is reasonable to seek $\varphi_\nu(x)$ represented as

$$\varphi_\nu(x) = \sum_m \{A_{\nu m}\varphi_m^{(1)}(x) + B_{\nu m}\varphi_m^{(2)}(x)\}. \quad (64)$$

Substituting (64) into Eq. (63), multiplying the equation by $\varphi_n^{(1)}(x) = \varphi_n(x - x_1)$, integrating the result over the $x$ axis, and performing similar operations on $\varphi_n^{(2)}(x) = \varphi_n(x - x_2)$, we obtain a set of equations for the coefficients $A_{\nu n}$ and $B_{\nu n}$:

$$\begin{aligned}(\alpha_n - \alpha_\nu)A_{\nu n} + \sum_m (\alpha_m - \alpha_\nu)B_{\nu m}J_{mn}^{(1)}(-l)\\ = \alpha_\nu \sum_m \{A_{\nu m}J_{nm}^{(2)}(-l) + B_{\nu m}J_{nm}^{(1)}(l)\},\\ \sum_m (\alpha_m - \alpha_\nu)A_{\nu m}J_{mn}^{(1)}(l) + (\alpha_n - \alpha_\nu)B_{\nu n}\\ = \alpha_\nu \sum_m \{A_{\nu m}J_{nm}^{(1)}(-l) + B_{\nu m}J_{nm}^{(2)}(l)\}.\end{aligned} \quad (65)$$

where

$$\begin{aligned}J_{nm}^{(1)}(\pm l) &= \int_{-\infty}^{+\infty} \varphi_n(x)\varphi_m(x \pm l)v(x)dx,\\ J_{nm}^{(2)}(\pm l) &= \int_{-\infty}^{+\infty} \varphi_n(x \pm l)\varphi_m(x \pm l)v(x)dx.\end{aligned} \quad (66)$$

Since $l \gg a$, asymptotic expression (13) for $\varphi_n(x)$ can be used to write

$$\begin{aligned}J_{nm}^{(1)}(\pm l) &\approx -\frac{\alpha_m u_m(\pm)u_n(\mp)}{2\kappa}e^{-\kappa l},\\ J_{nm}^{(2)}(\pm l) &\approx \frac{\alpha_n u_n(\pm)}{2\kappa}\frac{\alpha_m u_m(\pm)}{2\kappa}e^{-2\kappa l}\int_{-\infty}^{+\infty} e^{\mp 2\kappa x}v(x)dx.\end{aligned} \quad (67)$$

In the case of short-range potentials $v(x)$ whose asymptotic behavior is described by (35), they can be treated as nonoverlapping if $\kappa < \gamma \sim 1/a$. Note also that this conditions ensures that the integral in (67) is convergent.

Consider a state $\nu$ "close" to the unperturbed state $n$. Assuming that $\kappa l \gg 1$, we can neglect the terms containing $(\alpha_n - \alpha_\nu)J_{nm}^{(1)}$ and $J_{nm}^{(2)}$ in (65), as well as those with $m \neq n$, since they are proportional to $\exp(-2\kappa l)$. As a result, (65) yields two solutions $\alpha_{n\sigma}$ ($\sigma = 1, 2$) for $\alpha_\nu$, which correspond to $B_{\nu n}/A_{\nu n} = -(-1)^\sigma$:

$$\begin{aligned}\alpha_{n\sigma} &= \alpha_n\{1 + (-1)^\sigma J_n\} \quad (\sigma = 1, 2),\\ J_n &= -\frac{1}{2\kappa}\alpha_n u_n(+)u_n(-)e^{-\kappa l}.\end{aligned} \quad (68)$$

Higher order corrections to (68) can be obtained in a consistent manner by taking into account the neglected terms in (65).

Substituting (68) into Eq. (19), we find the corresponding energy levels:

$$\varepsilon_{n\sigma}(l) = \varepsilon_n + (-1)^\sigma 2\kappa_n A_n(+)A_n(-)e^{-\kappa_n l} \quad (\sigma = 1, 2). \quad (69)$$

Here, $A_n(\pm)$ are the coefficients in the asymptotic expressions for the bound-state wavefunction $\psi_n(x)$ given by (21) in the case of an isolated potential well with $\varepsilon_n = -\kappa_n^2$ determined by solving the equation $\alpha_n(\varepsilon) = \alpha$. Energy levels (69) are associated with the wavefunctions

$$\psi_{n\sigma}(x) = \frac{1}{\sqrt{2}}\{\psi_n(x - x_1) - (-1)^\sigma \psi_n(x - x_2)\},$$

which are identical to those used in the standard analysis of the two-center problem.

Consider the splitting of the level with $n = 0$ when $l \gg a$ and $\kappa a \ll 1$, while $\kappa l$ is arbitrary. In this case, the results presented in Section 5 can be used to reduce (67) to

$$J_{00}^{(1)}(\pm l) \approx e^{-\kappa l}, \quad J_{00}^{(2)}(\pm l) \approx e^{-2\kappa l}.$$

Then, we drop the terms with $m \neq 0$ in (65), since they scale with the parameter $\kappa a \ll 1$, to obtain

$$\alpha_{0\sigma} \approx \frac{\alpha_0}{1 - (-1)^\sigma \exp(-\kappa l)}, \quad (70)$$

where $\alpha_0 = -\kappa/a$. It follows from (70) that the system has a single bound state if $|\alpha|a^2 < a/l \ll 1$ and two bound states if $a/l < |\alpha|a^2 \ll 1$. If $\kappa l \gg 1$, then (70) is equivalent to (68) for $n = 0$ and $\kappa a \ll 1$. Expression (70) is exact in the case of two singular potentials $v(x) = 2a\delta(x)$ separated by a distance $l$.

## 10. GREEN'S FUNCTION

The Green function $G(x, x')$ satisfies the equation

$$\left\{\frac{\partial^2}{\partial x^2} + \varepsilon - a v(x)\right\}G(x, x') = \delta(x - x'). \quad (71)$$

Representing the solution to Eq. (71) as an expansion in

the basis $\{\varphi_n(x)\}$, we obtain

$$G(x, x') = \sum_n \frac{\varphi_n(x)\varphi_n(x')}{\alpha_n - \alpha}. \quad (72)$$

If $\alpha = 0$, then (72) reduces to expansion (17) of the zeroth-order Green function $G_0(x - x')$. Using (17), we can rewrite expression (72) for $G(x, x')$ as

$$G(x, x') = G_0(x - x') + \alpha\sum_n \frac{\varphi_n(x)\varphi_n(x')}{\alpha_n(\alpha_n - \alpha)}. \quad (73)$$

In the case of a finite-range potential, i.e., when $V(x) = 0$ outside the interval $(x_1, x_2)$, expansions (72) and (17) are valid if at least one of the coordinates $x$ and $x'$ belongs to $(x_1, x_2)$, while the other coordinate may be arbitrary. If both coordinates lie outside $(x_1, x_2)$, then expansions (72) and (17) do not exist and expression (73) for $G(x, x')$ with $G_0(x - x')$ given by (11) is valid (cf. [7]).

Rewrite Eq. (71) in the integral form

$$G(x, x') = G_0(x - x') + \alpha\int_{-\infty}^{+\infty} G_0(x - x_1)V(x_1)G(x_1, x')dx_1.$$

When $\alpha$ is sufficiently small, this equation can be solved by iteration:

$$G(x, x') = G_0(x - x') + G^{(1)}(x, x') + G^{(2)}(x, x') + \ldots, \quad (74)$$

$$G^{(k)}(x, x') = \alpha\int_{-\infty}^{+\infty} G_0(x - x_1)V(x_1)G^{(k-1)}(x_1, x')dx_1 \quad (75)$$

$$(k \geq 1),$$

where $G^{(k)}(x, x')$ scales with the $k$th power of $\alpha$. The expansion defined by (74) and (75) corresponds to a standard perturbation that can be applied in the one-dimensional case if $|\alpha|a^2 \ll \kappa a$ or, in standard units, if $|U| \ll (\hbar^2/ma^2)\kappa a$ (see [1, Section 45]).

As $\alpha \longrightarrow 0$, the general expression for the Green function given by (72) or (73) yields (74) with

$$G^{(k)}(x, x') = (\alpha)^k\sum_n \frac{\varphi_n(x)\varphi_n(x')}{(\alpha_n)^{k+1}} \quad (k \geq 1). \quad (76)$$

To elucidate the correspondence between (75) and (76), replace $x'$ with $x_1$ in Eq. (12), multiply the equation by $\alpha^k\varphi_n(x')(\alpha_n)^{-(k+1)}$, and perform the summation over $n$. Since the result is recursive relation (75), the expansions defined by (74) and (75) and by (74) and (76) are equivalent. The expansion defined by (74) and (76) is valid only if $|\alpha/\alpha_n| \ll 1$ for all $n$, including $n = 0$. According to the results presented in Section 5, if $\kappa a \ll 1$, then $|\alpha_0| \approx \kappa/a \ll |\alpha_n| \sim 1/a^2$. Therefore, the expansion defined by (74) and (76) is valid if $|\alpha| \ll |\alpha_0|$ or $|\alpha|a^2 \ll \kappa a$, which is equivalent to the criterion formulated above.

Thus, potential energy can be treated as a perturbation (for $\kappa a \ll 1$) in the one-dimensional case only under the very restrictive condition that $|\alpha|a^2 \ll \kappa a \ll 1$. However, it is clear that the expansion in which the state with $n = 0$ is singled out has a wider scope corresponding to the three-dimensional case: $|\alpha|a^2 \ll 1$ (for $\kappa a \ll 1$).

Let single out the contribution of the state with $n = 0$ in (73):

$$G(x, x') = G_0(x - x') + \frac{\alpha}{\alpha_0(\alpha_0 - \alpha)}\varphi_0(x)\varphi_0(x') + \tilde{G}(x, x'), \quad (77)$$

$$\tilde{G}(x, x') = \alpha\sum_{n\neq 0} \frac{\varphi_n(x)\varphi_n(x')}{\alpha_n(\alpha_n - \alpha)}.$$

For $n \neq 0$ and $\kappa a \ll 1$, it holds that $|\alpha_n| \sim 1/a^2$, and $\tilde{G}(x, x')$ can be represented as a series in powers of $\alpha$ if $|\alpha|a^2 \ll 1$:

$$\tilde{G}(x, x') = \alpha\sum_{n\neq 0} \frac{\varphi_n(x)\varphi_n(x')}{(\alpha_n)^2} + \alpha^2\sum_{n\neq 0} \frac{\varphi_n(x)\varphi_n(x')}{(\alpha_n)^3} + \ldots.$$

By adding and subtracting the contribution of the state with $n = 0$, the sums contained in this expansion can be expressed in terms of $G^{(k)}(x, x')$ defined by (75). The resulting perturbative series for the Green function is

$$G(x, x') = G_0(x - x') + \frac{\alpha}{\alpha_0(\alpha_0 - \alpha)}\varphi_0(x)\varphi_0(x') + \alpha \times\left\{\int_{-\infty}^{+\infty} G_0(x - x_1)V(x_1)G_0(x_1 - x')dx_1 - \frac{\varphi_0(x)\varphi_0(x')}{(\alpha_0)^2}\right\} + \ldots \quad (78)$$

with higher order terms in $\alpha$ "renormalized" in a similar manner. Expansion (78) is valid if $|\alpha|a^2 \ll 1$ for $\kappa a \ll 1$.

When $V(x) = 2a\delta(x)$, there is a single state, for which $\alpha_0 = -\kappa/a$ and $\varphi_0(x) = (2a)^{-1/2}\exp(-\kappa|x|)$. In this case, it can readily be shown that an exact expression for the Green function is given by (77) with $\tilde{G}(x, x') = 0$ (cf. Problems 2.45 and 2.46 in [9]).

## 11. TRANSMISSION COEFFICIENT FOR SLOW PARTICLES

Consider a plane wave passing through a potential barrier (or over a potential well). Represent the wavefunction as

$$\psi(x) = e^{ikx} + F(x)e^{ik|x|}. \quad (79)$$

The amplitudes of the transmitted and reflected waves, $d$ and $r$, are expressed in terms of $F(x)$ as follows:

$$d = 1 + F(+\infty), \quad r = F(-\infty).$$

Substituting (79) into (1) (with $\varepsilon = k^2 > 0$), we have

$$\begin{aligned} &F''(x) + 2ikF'(x)\operatorname{sgn}x - \alpha v(x)F(x) \\ &+ 2ikF(0)\delta(x) = \alpha v(x)\exp(ik(x-|x|)). \end{aligned} \quad (80)$$

The solution to Eq. (80) is sought in the form of an expansion in the eigenfunction set $\{\zeta_n(x)\}$ corresponding to zero-energy states:

$$F(x) = \sum_m F_m \zeta_m(x).$$

Using a standard procedure, we obtain a set of equations for the coefficients $F_n$:

$$\begin{aligned} &(\lambda_n - \alpha)F_n + 2ik\sum_m M_{mn}F_m \\ &= \alpha\int_{-\infty}^{+\infty} e^{ik(x-|x|)}\zeta_n(x)v(x)dx, \end{aligned} \quad (81)$$

where $M_{mn}$ is defined in (43).

For $k \longrightarrow 0$ ($ka \ll 1$), we represent $F_n$ as expansions in powers of $k$: $F_n = F_n^{(0)} + F_n^{(1)} + F_n^{(2)} + \ldots$. In the zeroth approximation, we have $F_n^{(0)} = -\sqrt{2a}\,\delta_{n0}$ and $F^{(0)}(x) = -1$; i.e., the transmitted wave obviously vanishes. In the linear approximation in $k$,

$$d^{(1)} = -\frac{ik}{\alpha a}\left\{1 + \frac{\alpha}{2}\int_{-\infty}^{+\infty}(x-|x|)v(x)dx\right.$$

$$\left. - 2\alpha a\sum_{n\neq 0}\frac{1}{\lambda_n}\left[\zeta_n(0) + \frac{\alpha}{\lambda_n - \alpha}\zeta_n(-\infty)\right]\zeta_n(+\infty)\right\}.$$

Using (30) to eliminate the integral, we finally obtain

$$d^{(1)} = 2ik\sum_{n=0}^{\infty}\frac{\zeta_n(-\infty)\zeta_n(+\infty)}{\lambda_n - \alpha}. \quad (82)$$

According to (82), the transmission coefficient is $D = |d|^2 \propto k^2$ as $k \longrightarrow 0$, in agreement with [1, Section 25, Problem 5].

When $|\alpha|a^2 \ll 1$, expression (82) reduces to $d \approx -ik/(\alpha a)$, which is valid if $ka \ll |\alpha|a^2$. If $ka \sim |\alpha|a^2 \ll 1$, then $F_0$ should be calculated by solving Eq. (81) with a term $\sim k/a$ added to $\alpha$ on the left-hand side. The result

$$d \approx \frac{ik}{ik - \alpha a}$$

obtained in this case is exact when $v(x) = 2a\delta(x)$.

Expression (82) for $d$ corresponding to an attractive potential ($\alpha < 0$) exhibits singular behavior as $\alpha \longrightarrow \lambda_n$ (with $n \neq 0$). In this case, $F_n$ (with $n \neq 0$) should be calculated by solving Eq. (81) with a term $\sim k/a$ retained in addition to $(\lambda_n - \alpha)$ on the left-hand side. For $|\alpha - \lambda_n|a^2 \ll 1$ and $ka \ll 1$, the result is

$$d \approx 2ik\frac{\zeta_n(-\infty)\zeta_n(+\infty)}{\lambda_n - \alpha + 2ikM_{nn}}, \quad (83)$$

where $M_{nn}$ is defined in (43). If $ka \ll |\alpha - \lambda_n|a^2 \ll 1$, then $|d| \ll 1$. If $|\alpha - \lambda_n|a^2 \ll ka \ll 1$, then (83) implies that $|d| \sim 1$; for an even potential function, $|d| \approx 1$. Thus, an anomalously large transmission coefficient is obtained, even though the potential amplitude is not small: $|\alpha|a^2 \approx |\lambda_n|a^2 > 1$.

In the same approximation ($|\alpha - \lambda_n|a^2 \ll ka \ll 1$), we have (for an even potential function)

$$x < 0\colon \psi(x) \approx e^{ikx} - \left[1 - \frac{\zeta_n(x)}{\zeta_n(-\infty)}\right]e^{-ikx}. \quad (84)$$

Thus, the reflected wave vanishes in the asymptotic expression for $\psi(x)$ ($x \longrightarrow -\infty$), whereas the wavefunction corresponding to $|x| < a$ may substantially differ from the incident plane wave. Note also that expression (83) (more precisely, its analytic continuation to the complex $\varepsilon$ plane) has a pole at $\varepsilon = \varepsilon_n$, where the bound-state energy $\varepsilon_n$ is determined by solving Eq. (19) with $\alpha_n(\varepsilon) = \lambda_n + \alpha_n^{(1)}(\varepsilon)$, where $\alpha_n^{(1)}(\varepsilon)$ is defined by (44).

In the case of a potential barrier ($\alpha > 0$), the transmission coefficient decreases with increasing $\alpha$. When $\alpha a^2 \gg 1$, the sum in (82) is dominated by terms with $n \gg 1$. The semiclassical expressions for $\lambda_n$ and $\zeta_n(\pm\infty)$ corresponding to potentials with exponential asymptotics described by (35) are presented in Section 4. Substituting them into (82) and using the well-known expansion [10]

$$\frac{1}{\cosh\dfrac{\pi x}{2}} = \frac{4}{\pi}\sum_{n=0}^{\infty}(-1)^n\frac{2n+1}{(2n+1)^2 + x^2},$$

we obtain

$$d^{(1)} \approx -ik\frac{\pi}{\gamma}\frac{1}{\cosh(b\sqrt{\alpha})}$$

with $b$ defined in (38). When $b \sim a$, it holds that $b\sqrt{\alpha} \gg 1$, and the final expression for $d^{(1)}$ is

$$d^{(1)} \approx -ik\frac{2\pi}{\gamma}\exp\left\{-\sqrt{\alpha}\int_{-\infty}^{+\infty}\sqrt{v(x)}dx\right\} \quad (\sqrt{\alpha}a \gg 1, ka \ll 1). \tag{85}$$

## 12. EXACTLY SOLVABLE EXAMPLE

As an illustration of the approach proposed here, consider the potential function

$$v(x) = \frac{1}{\cosh^2\gamma x}, \tag{86}$$

for which both eigenfunctions and eigenvalues can be found in a relatively simple form. Equation (2) with $v(x)$ defined by (86) is solved by analogy with Problem 5 in Section 23 of [1]. The eigenvalues are

$$\alpha_n = -\gamma^2\left(n + \frac{\kappa}{\gamma}\right)\left(n + 1 + \frac{\kappa}{\gamma}\right) \quad (n = 0, 1, 2, \ldots). \tag{87}$$

The corresponding eigenfunctions normalized according to (5) are

$$\varphi_n(x) = B_n(1-\xi^2)^{(2\rho-1)/4}C_n^\rho(\xi), \quad \xi = \tanh\gamma x, \quad B_n = 2^\rho\Gamma(\rho)\left[\gamma\frac{(n+\rho)n!}{2\pi\Gamma(n+2\rho)}\right]^{1/2}, \quad \rho = \frac{1}{2} + \frac{\kappa}{\gamma}, \tag{88}$$

where $C_n^\rho(\xi)$ are Gegenbauer polynomials [10–12].

According to [10–12], the Gegenbauer polynomials $C_n^\rho(\xi)$ are orthogonal with the weight function $w(\xi) = (1-\xi^2)^{(2\rho-1)/2}$ and the functions

$$\Phi_n^\rho(\xi) = 2^\rho\Gamma(\rho)\left[\frac{(n+\rho)n!}{2\pi\Gamma(n+2\rho)}\right]^{1/2}(1-\xi^2)^{(2\rho-1)/4}C_n^\rho(\xi)$$

constitute a complete orthonormal (closed) set on the interval $(-1 < \xi < +1)$:

$$\sum_{n=0}^{\infty}\Phi_n^\rho(\xi)\Phi_n^\rho(\xi') = \delta(\xi - \xi').$$

Therefore, the set $\{\varphi_n(x)\}$ of the functions defined in (88) satisfies the condition

$$\sum_{n=0}^{\infty}\varphi_n(x)\varphi_n(x') = \gamma\delta(\xi - \xi') = \frac{\delta(x-x')}{v(x)}, \tag{89}$$

which is equivalent to completeness condition (7).

Condition (89) can be derived as follows. First, the sum on the left-hand side of (89) is calculated from $n = 0$ to $n = N$ by invoking the Christoffel–Darboux formula (see [10–12]). Then, an asymptotic expression for $C_n^\rho(\xi)$ valid at $N \gg 1$ (see [11]) is used to find the limit as $N \longrightarrow \infty$. As a result, we obtain (89).

As $x \longrightarrow +\infty$, we have $\xi \approx 1 - 2\exp(-2\gamma x)$, and (88) yields

$$x \longrightarrow +\infty: \quad \varphi_n(x) \approx B_nC_n^\rho(1)2^{(2\rho-1)/2}\exp(-\kappa x).$$

Combined with (13) and an expression for $C_n^\rho(1)$ borrowed from [10, 11], we obtain

$$u_n(+) = -\frac{\kappa}{\alpha_n}2^{(4\rho+1)/2} \times \frac{\Gamma(\rho)}{\Gamma(2\rho)}\left[\gamma\frac{(n+\rho)\Gamma(n+2\rho)}{2\pi n!}\right]^{1/2} \tag{90}$$

with $\rho$ defined in (88). Accordingly, we have $u_n(-) = (-1)^n u_n(+)$. As $\kappa \longrightarrow 0$, expression (90) reduces to

$$u_0^{(0)}(\pm) = \sqrt{\frac{2}{\gamma}}, \quad u_0^{(1)}(\pm) = 0, \quad n \neq 0: \ u_n^{(0)}(\pm) = 0, \quad u_n^{(1)}(+) = -\frac{2\kappa}{\lambda_n}\sqrt{\gamma\left(n + \frac{1}{2}\right)}. \tag{91}$$

As $\kappa \longrightarrow \infty$, (90) yields

$$\kappa \longrightarrow \infty: \quad u_n(+) \approx \frac{2}{\kappa}\frac{1}{(2\pi)^{1/4}}\sqrt{\frac{\gamma}{n!}}\left(\frac{2\kappa}{\gamma}\right)^{(2n+1)/4}\exp\left(\frac{\kappa}{\gamma}\ln 2\right). \tag{92}$$

Since

$$C_n^\rho(\xi) \approx \frac{\rho^{n/2}}{n!}H_n(\xi\sqrt{\rho}),$$

as $\rho \longrightarrow \infty$ and $\xi \longrightarrow 0$ (see [10, 11]), where $H_n(z)$ denotes Hermitian polynomials, (88) becomes

$$\varphi_n(x) \approx \left(\frac{\kappa\gamma}{\pi}\right)^{1/4}\frac{1}{\sqrt{2^n n!}} \times \exp(-(\kappa\gamma/2)x^2)H_n(x\sqrt{\kappa\gamma}). \tag{93}$$

Expression (93) is valid if $\kappa/\gamma \gg 1$ for $n = 0$ and if $\kappa/\gamma \gg n$ for $n > 0$, provided that $\gamma|x| \ll 1$.

As $\kappa \longrightarrow 0$, we use the fact that $C_n^{1/2}(\xi) = P_n(\xi)$ (see [10, 11]), where $P_n(\xi)$ denotes Legendre polyno-

mials, to derive an expression for $\zeta_n(x) = \varphi_n(0; x)$ from (88),

$$\zeta_n(x) = \sqrt{\gamma\left(n+\frac{1}{2}\right)}P_n(\xi), \quad \xi = \tanh\gamma x \tag{94}$$

and find $\lambda_n = -\gamma^2 n(n+1)$. The functions in the set defined by (94) are orthonormal with weight function (86) and complete (closed). Since $P_n(1) = 1$ and $P_n(-1) = (-1)^n$ [10, 11], it follows from (94) that

$$\zeta_n(+\infty) = \sqrt{\gamma\left(n+\frac{1}{2}\right)}, \qquad \zeta_n(-\infty) = (-1)^n\zeta_n(+\infty). \tag{95}$$

By invoking an expression for $P_n(0)$ (see [10, 11]), we use (94) to obtain

$$\zeta_n(0) = \sqrt{\frac{\gamma}{\pi}\left(n+\frac{1}{2}\right)}\frac{\Gamma\left(\frac{n}{2}+\frac{1}{2}\right)}{\Gamma\left(\frac{n}{2}+1\right)}\cos\frac{n\pi}{2}.$$

Hence,

$$n \gg 1\colon \zeta_n(0) \approx \sqrt{\frac{2\gamma}{\pi}}\cos\frac{n\pi}{2}. \tag{96}$$

For potential function (86), the constants defined in (18), (32), (38), (52), and (54) are

$$a = \frac{1}{\gamma}, \quad b = \frac{\pi}{\gamma}, \quad I = \frac{4}{\gamma^3}, \qquad \delta = \gamma, \quad I_1 = \ln 2, \quad I_2 = 0. \tag{97}$$

When combined with (97), expressions (87), (91)–(93), (95), and (96) are equivalent to the corresponding results presented in Sections 4–6.

The tunneling of a particle through the potential barrier defined by (86) was considered in [1, Section 25, Problem 4]. The solution obtained can be used to obtain the transmitted-wave amplitude $d$:

$$d = \frac{\Gamma(s+1-i(k/\gamma))\Gamma(-s-i(k/\gamma))}{\Gamma(1-i(k/\gamma))\Gamma(-i(k/\gamma))}, \qquad s = \frac{1}{2}\left(-1+\sqrt{1-\frac{4\alpha}{\gamma^2}}\right). \tag{98}$$

In the linear approximation in $k$, (98) yields

$$d^{(1)} = -ik\frac{\pi}{\gamma}\frac{1}{\cos\left[\frac{\pi}{2}\sqrt{1-\frac{4\alpha}{\gamma^2}}\right]}. \tag{99}$$

This result can also be obtained by substituting $\lambda_n = -\gamma^2 n(n+1)$ and expression (95) for $\zeta_n(\pm\infty)$ into (82). If $\alpha > 0$ and $\alpha/\gamma^2 \gg 1$, then (99) is equivalent to the expression preceding (85), with $b = \pi/\gamma$.

ACKNOWLEDGMENTS

I thank V.Ya. Krivnov for discussion of this work.

## *APPENDIX A*

The eigenvalue problem can be solved for a number of specific potential functions. In addition to the example discussed in Section 12, the results obtained for three exactly examples are presented here.

**1.** For the potential function

$$v(x) = e^{-2\gamma|x|}, \tag{A.1}$$

the subsystem of symmetric eigenfunctions (representing even states) is

$$\varphi_{sn}(x) = A_{sn}J_\nu(\mu_{sn}e^{-\gamma|x|}), \qquad \nu = \frac{\kappa}{\gamma}, \quad \mu_{sn} = \frac{1}{\gamma}\sqrt{-\alpha_{sn}}, \tag{A.2}$$

where $J_\nu(z)$ is the Bessel function. The corresponding eigenvalues $\alpha_{sn}$ are found by solving the equation

$$J'_\nu(\mu_{sn}) = 0 \tag{A.3}$$

with $\mu_{sn}$ defined in (A.2). The normalization factor is determined by condition (5):

$$A_{sn} = \frac{\mu_{sn}}{J_\nu(\mu_{sn})}\sqrt{\frac{\gamma}{\mu_{sn}^2-\nu^2}}. \tag{A.4}$$

The antisymmetric eigenfunctions (representing odd states) are

$$\varphi_{an} = A_{an}J_\nu(\mu_{an}e^{-\gamma|x|})\operatorname{sgn}x, \qquad \nu = \frac{\kappa}{\gamma}, \quad \mu_{an} = \frac{1}{\gamma}\sqrt{-\alpha_{an}}. \tag{A.5}$$

The corresponding eigenvalues $\alpha_{an}$ are found by solving the equation

$$J_\nu(\mu_{an}) = 0 \tag{A.6}$$

with $\mu_{an}$ defined in (A.5). The normalization factor is

$$A_{an} = -\frac{\sqrt{\gamma}}{J'_\nu(\mu_{an})} = \frac{\sqrt{\gamma}}{J_{\nu+1}(\mu_{an})}. \tag{A.7}$$

In this case,

$$a = \frac{1}{2\gamma}, \quad b = \frac{2}{\gamma}, \quad I = \frac{3}{4\gamma^3}.$$

The orthonormality of the system defined by (A.2)–(A.7) can be verified directly. Its completeness follows

from the convergence of the Fourier–Bessel and Dini expansions [11].

**2.** As an example of finite-range potential, consider the potential well

$$v(x) = \theta(a - |x|), \tag{A.8}$$

where $\theta(x) = 1$ at $x > 0$ and $\theta(x) = 0$ at $x < 0$.

The symmetric eigenfunctions are

$$\begin{gathered}|x| < a\colon\ \varphi_{sn}^{(i)}(x) = A_{sn}\cos(\mu_{sn}x),\\ \mu_{sn} = \sqrt{-\alpha_{sn} - \kappa^2},\\ |x| > a\colon \varphi_{sn}^{(e)}(x) = B_{sn}\exp(-\kappa(|x| - a)),\\ B_{sn} = A_{sn}\cos(\mu_{sn}a),\\ A_{sn} = \left[a - \frac{\kappa}{\alpha_{sn}}\right]^{-1/2}.\end{gathered} \tag{A.9}$$

The corresponding eigenvalues $\alpha_{sn}$ are found by solving the equation

$$\mu_{sn}\tan(\mu_{sn}a) = \kappa \tag{A.10}$$

with $\mu_{sn}$ defined in (A.9).

The antisymmetric eigenfunctions are

$$\begin{gathered}|x| < a\colon\ \varphi_{an}^{(i)}(x) = A_{an}\sin(\mu_{an}x),\\ \mu_{an} = \sqrt{-\alpha_{an} - \kappa^2},\\ |x| > a\colon\ \varphi_{an}^{(e)}(x) = B_{an}\exp(-\kappa(|x| - a))\operatorname{sgn}x,\\ B_{an} = A_{an}\sin(\mu_{an}a),\\ A_{an} = \left[a - \frac{\kappa}{\alpha_{an}}\right]^{-1/2}.\end{gathered} \tag{A.11}$$

The corresponding eigenvalues $\alpha_{an}$ are found by solving the equation

$$\mu_{an}\cot(\mu_{an}a) = -\kappa \tag{A.12}$$

with $\mu_{an}$ defined in (A.11).

The system defined by (A.9)–(A.12) is orthonormal and complete on the interval $(-a < x < a)$.

**3.** For the potential function

$$v(x) = \frac{a_0^4}{(x^2 + a_0^2)^2} \tag{A.13}$$

the eigenfunctions and eigenvalues corresponding to zero-energy states can be found.

The even states with $\varepsilon = 0$ are associated with the symmetric eigenfunctions

$$\zeta_{sn}(x) = \sqrt{\frac{2}{\pi a_0^3}}\sqrt{x^2 + a_0^2}\cos\left[(2n+1)\arctan\frac{x}{a_0}\right] \tag{A.14}$$

and eigenvalues

$$\begin{gathered}\lambda_{sn} = -\frac{1}{a_0^2}[(2n+1)^2 - 1] = -\frac{4}{a_0^2}n(n+1)\\ (n = 0, 1, 2, \ldots).\end{gathered} \tag{A.15}$$

It follows from (A.14) that

$$\begin{gathered}\zeta_{sn}(+\infty) = \zeta_{sn}(-\infty) = (-1)^n(2n+1)\sqrt{\frac{2}{\pi a_0}},\\ \zeta_{sn}(0) = \sqrt{\frac{2}{\pi a_0}}.\end{gathered} \tag{A.16}$$

The odd states with $\varepsilon = 0$ are associated with the antisymmetric eigenfunctions

$$\zeta_{an}(x) = \sqrt{\frac{2}{\pi a_0^3}}\sqrt{x^2 + a_0^2}\sin\left[2n\arctan\frac{x}{a_0}\right] \tag{A.17}$$

and eigenvalues

$$\lambda_{an} = -\frac{1}{a_0^2}[(2n)^2 - 1] \quad (n = 1, 2, \ldots). \tag{A.18}$$

It follows from (A.17) that

$$\begin{gathered}\zeta_{an}(+\infty) = -\zeta_{an}(-\infty) = (-1)^{n+1} \times 2n\sqrt{\frac{2}{\pi a_0}},\\ \zeta_{an}(0) = 0.\end{gathered} \tag{A.19}$$

The eigenfunction set $\{\zeta_{sn}(x), \zeta_{an}(x)\}$ is orthonormal (by (23)) and complete.

The results presented in Sections 5 and 6 can be used to generalize the analysis of the case with $\varepsilon = 0$. In the present example, we have

$$a = \frac{\pi a_0}{4}, \quad b = \pi a_0, \quad I = \frac{3\pi}{4}a_0^3,$$

$$I_1 = 2 - \sqrt{2}\ln(1 + \sqrt{2}) \approx 0.75,$$

$$I_2 = 2\ln 2 - \sqrt{2}\ln(1 + \sqrt{2}) \approx 0.14.$$

## *APPENDIX B*

The states with large $|\varepsilon|$ for potential functions with cusps at their minimum points call for a separate analysis. Let $v(x)$ tend to a linear function of $x$ as $x \longrightarrow 0$:

$$v(x) = 1 - 2\gamma|x| + \ldots \quad \left(|x| \ll \frac{1}{\gamma}\right), \tag{B.1}$$

where $\gamma \sim 1/a$. For $|x| \ll a$, Eq. (2) combined with (51) and (B.1) becomes

$$\varphi_n''(x) - 2\gamma\kappa^2\left(|x| - \frac{\Delta_n}{2\gamma}\right)\varphi_n(x) = 0. \quad \text{(B.2)}$$

The eigenfunctions $\varphi_{sn}(x)$ associated with even states are symmetric in $x$:

$$\begin{aligned}\varphi_{sn}(x) &= B_{sn}\Phi(\xi_s(x)),\\ \xi_s(x) &= (2\gamma\kappa^2)^{1/3}\left(|x| - \frac{\Delta_{sn}}{2\gamma}\right),\\ B_{sn} &= \frac{(2\gamma\kappa^2)^{1/6}}{\sqrt{2z_n'}\Phi(-z_n')},\end{aligned} \quad \text{(B.3)}$$

where $\Phi(z)$ denotes the Airy function [1]:

$$\Phi(z) = \sqrt{\pi}\mathrm{Ai}(z) = \frac{1}{\sqrt{\pi}}\int_0^\infty \cos\left(zt + \frac{1}{3}t^3\right)dt. \quad \text{(B.4)}$$

The correction $\Delta_{sn}$ (see (51)) is determined by the condition $\Phi'(\xi_s(0)) = 0$:

$$\Delta_{sn} = \left(\frac{2\gamma}{\kappa}\right)^{2/3} z_n'. \quad \text{(B.5)}$$

The positive numbers $z_n'$ in (B.3) and (B.5) are the zeros of the derivative $\Phi'(-z)$.

The eigenfunctions $\varphi_{an}(x)$ associated with odd states are antisymmetric:

$$\begin{aligned}\varphi_{an}(x) &= B_{an}\Phi(\xi_a(x))\,\mathrm{sgn}\,x,\\ \xi_a(x) &= (2\gamma\kappa^2)^{1/3}\left(|x| - \frac{\Delta_{an}}{2\gamma}\right),\\ B_{an} &= \frac{(2\gamma\kappa^2)^{1/6}}{\sqrt{2}\Phi'(-z_n)}.\end{aligned} \quad \text{(B.6)}$$

The correction $\Delta_{an}$ (see (51)) is determined by the condition $\Phi(\xi_a(0)) = 0$:

$$\Delta_{an} = \left(\frac{2\gamma}{\kappa}\right)^{2/3} z_n. \quad \text{(B.7)}$$

The positive numbers $z_n$ in (B.6) and (B.7) are the zeros of $\Phi(-z)$. The normalization factors in (B.3) and (B.6) are calculated by using the equation

$$\int [\Phi(z)]^2 dz = z[\Phi(z)]^2 - [\Phi'(z)]^2, \quad \text{(B.8)}$$

which is easily verified by means of the equation $\Phi''(z) - z\Phi(z) = 0$ for $\Phi(z)$.

The function $\varphi_n(x)$ can be calculated in the semiclassical approximation at $x \gg x_0 > 0$:

$$\begin{aligned}\varphi_n(x) &= \frac{D_n}{\sqrt[4]{\alpha_n v(x) - \varepsilon}}\\ &\times \exp\left\{-\int_{x_0}^x \sqrt{\alpha_n v(t) - \varepsilon}\,dt\right\},\end{aligned} \quad \text{(B.9)}$$

where $x_0$ is the zero of the radicand: $x_0 \approx \Delta_n/(2\gamma) \ll a$. The factor $D_n$ is determined by matching (B.9) with (B.3) and (B.6) in the interval $x_0 \ll x \ll a$:

$$D_n = \frac{1}{2}(2\gamma\kappa^2)^{1/6} B_n. \quad \text{(B.10)}$$

The calculation of (B.10) relies on the asymptotic formula for $\Phi(z)$ [1]:

$$z \longrightarrow +\infty: \quad \Phi(z) \approx \frac{1}{2z^{1/4}}\exp\left(-\frac{2}{3}z^{3/2}\right). \quad \text{(B.11)}$$

As $x \longrightarrow +\infty$, function (B.9) reduces to expression (13) with $\alpha_n \approx -\kappa^2$ and

$$u_n(+) \approx \frac{1}{\kappa}\left(\frac{2\gamma}{\kappa}\right)^{1/6} B_n \exp\left\{\kappa a\left(I_3 + \frac{\Delta_n}{2}I_4\right)\right\}, \quad \text{(B.12)}$$

where

$$\begin{aligned}I_3 &= \frac{1}{a}\int_0^\infty [1 - \sqrt{1 - v(x)}]\,dx,\\ I_4 &= \frac{1}{a}\int_0^\infty \frac{v(x)dx}{\sqrt{1 - v(x)}}.\end{aligned} \quad \text{(B.13)}$$

Here, the value of $a$ can be set equal to that defined in (18). For the potential function $v(x) = \exp(-2\gamma|x|)$, we use $a = 1/(2\gamma)$ to find

$$I_3 = 2(1 - \ln 2), \quad I_4 = 2.$$

Formulas (B.9)–(B.13) apply to both even and odd states.

## REFERENCES

1. L. D. Landau and E. M. Lifshitz, *Course of Theoretical Physics*, Vol. 3: *Quantum Mechanics: Non-Relativistic Theory*, 4th ed. (Oxford Univ. Press, Oxford, 1980; Nauka, Moscow, 1989).

2. A. I. Baz', Ya. B. Zel'dovich, and A. M. Perelomov, *Scattering, Reactions, and Decays in Nonrelativistic Quantum Mechanics*, 2nd ed. (Israel Program for Scientific Translations, Jerusalem, 1966; Nauka, Moscow, 1971).
3. M. Rotenberg, Ann. Phys. (N.Y.) **19**, 262 (1962).
4. S. V. Khristenko and S. I. Vetchinkin, Opt. Spektrosk. **37**, 1017 (1974).
5. S. M. Blinder, Phys. Rev. A **29**, 1674 (1984).
6. B. Ya. Balagurov, Zh. Éksp. Teor. Fiz. **93**, 316 (1987) [Sov. Phys. JETP **66**, 182 (1987)].
7. B. Ya. Balagurov, Zh. Éksp. Teor. Fiz. **94** (7), 95 (1988) [Sov. Phys. JETP **67**, 1351 (1988)].
8. I. I. Privalov, *Integral Equations* (ONTI, Moscow, 1935) [in Russian].
9. V. M. Galitskiĭ, B. M. Karnakov, and V. I. Kogan, *Problems on Quantum Mechanics* (Nauka, Moscow, 1992) [in Russian].
10. I. S. Gradshteĭn and I. M. Ryzhik, *Table of Integrals, Series, and Products*, 4th ed. (Fizmatlit, Moscow, 1962; Academic, New York, 1980).
11. *Higher Transcendental Functions (Bateman Manuscript Project)*, Ed. by A. Erdelyi (McGraw-Hill, New York, 1953; Nauka, Moscow, 1974), Vol. 2.
12. G. Szego, *Orthogonal Polynomials* (Am. Math. Soc., New York, 1959; Fizmatgiz, Moscow, 1962).

*Translated by A. Betev*

*Journal of Experimental and Theoretical Physics, Vol. 99, No. 4, 2004, pp. 875–888.*
*From Zhurnal Éksperimental'noĭ i Teoreticheskoĭ Fiziki, Vol. 126, No. 4, 2004, pp. 1006–1020.*


**MISCELLANEOUS**

# Bilayer Membrane in Confined Geometry: Interlayer Slide and Entropic Repulsion¶

**S. V. Baoukina* and S. I. Mukhin****

*Moscow State Institute for Steel and Alloys (Technological University), Moscow, 119049 Russia*
**e-mail: svt_lana19@yahoo.com*
***e-mail: sergeimoscow@online.ru*

Received April 28, 2004

**Abstract**—We derive the free energy functional of a bilayer lipid membrane from the first principles of elasticity theory. The model explicitly includes position-dependent mutual slide of monolayers and bending deformation. Our free energy functional of a liquid-crystal membrane allows for incompressibility of the membrane and vanishing of the in-plane shear modulus and obeys reflectional and rotational symmetries of the flat bilayer. Interlayer slide at the midplane of the membrane results in local difference of surface densities of the monolayers. The slide amplitude directly enters the free energy via the strain tensor. For small bending deformations, the ratio between the bending modulus and the area compression coefficient, $K_b/K_A$, is proportional to the square of monolayer thickness $h$. Using the functional, we perform self-consistent calculation of the entropic potential acting on a bilayer between parallel confining walls separated by distance $2d$. We find that at the minimum of the confining potential, the temperature-dependent curvature $\alpha \propto T^2/K_b d^4$ is enhanced four times for a bilayer with slide as compared with a unit bilayer. We also calculate viscous modes of a bilayer membrane between confining walls. Pure bending of the membrane is investigated, which is decoupled from area dilation at small amplitudes. Three sources of viscous dissipation are considered: water and membrane viscosities and interlayer drag. The dispersion relation gives two branches $\omega_{1,2}(q)$. © *2004 MAIK "Nauka/Interperiodica".*

## 1. INTRODUCTION

A cell membrane is characterized by complex structural and dynamical properties [1–4]. Theoretical modeling and description of lipid membranes is of great fundamental and practical interest and has quite a long history. The phenomenological model introduced in [5] treated a lipid membrane as a single sheet with bending rigidity and spontaneous curvature. This model was later used for calculation of the frequency spectrum of the membrane in a water solution [6] and for investigation of entropic interactions of membranes in multilayer systems [7]. The bilayer structure of a lipid membrane was analyzed in [3, 8], where the dynamic coupling between the monolayers and the interlayer slide was considered. The frequency spectrum of a membrane in the bulk water was recalculated in [9] with the coupling between local curvature and local densities of lipids within the monolayers taken into account. Afterwards, viscous modes of a bilayer adhering to a substrate were found in [10] using the density-difference model [9], supplemented with a binding potential [11].

In this paper, we derive a new free energy functional of a bilayer membrane with interlayer slide. The interlayer slide function, membrane stretching, and bending amplitude directly enter the strain tensor of the membrane. Our functional is a generalization of the density-difference model used in [9, 10]. In our model, two lateral deformation fields (interlayer slide and stretching) generate the change in the local density and density difference of the monolayers, used in free energy functional in [9, 10]. However, unlike in [9, 10], we do not require the presence of neutral surfaces in each monolayer in the general case. As a step towards understanding intermembrane interactions, we study dynamics of a bilayer membrane in a water solution confined between parallel walls. The effect of confinement is modeled by the entropic potential [12].

This paper is organized as follows. In Section 2, we introduce an anisotropic elastic moduli tensor, initially containing 21 independent components. The reflection and rotation symmetries of the flat bilayer reduce the number of components to 5. Next, we impose the zero shear stress modulus and incompressibility constraint. We restrict ourselves to the case of small bending deformations and exclude the corresponding strain and elastic tensor components. Thus, the number of independent components of the elastic tensor in the free energy functional is reduced to two. The derived free energy functional of a bilayer membrane contains three fields describing area dilation and bending deformation coupled to interlayer slide.

In Section 3, a parabolic entropic potential acting on the membrane between the confining walls is introduced. We self-consistently calculate the curvature of

¶ This article was submitted by authors in English.

the confining potential at its minimum for a bilayer with a slide and for a unit bilayer. Using our model, we analytically derive the four-time decrease in bending rigidity due to interlayer slide.

In Section 4, we use the derived functional to study dynamic properties and dissipative mechanisms of the bilayer membrane in water solution confined between parallel walls. We investigate only pure bending deformations of the membrane (zero total lateral stretching), which decouple from area dilation. The velocity field in the surrounding water is found by solving the Stokes equations for an incompressible fluid. Fluid velocity vanishes at the walls. Equations of motion are determined as force balance conditions on the membrane surfaces with inertial effects neglected. Three sources of dissipation are included in the dynamic equations: water and membrane viscosities and interlayer drag.

In Section 5, we discuss limitations and possible improvements of our model and correspondence with earlier results [10]. In Appendix A, static behavior of a membrane in the axially symmetric case is studied. Analytic solutions are obtained for a circular membrane bent by external pressure. Membrane bending, interlayer slide, and lateral stress distribution are found as functions of pressure across the membrane. In Appendix B, we rederive the dispersion relation [9] for a membrane in the bulk water solution using our free energy functional.

## 2. FREE ENERGY FUNCTIONAL

The free energy density of an anisotropic medium can be written to the lowest order in the elastic strain tensor as [13, 14]

$$F = \frac{1}{2}\lambda_{iklm}u_{ik}u_{lm}, \tag{1}$$

where summation over the repeated indices $i, k, l, m$ is performed. The indices $i, k, l, m$ take values 1, 2, 3, which denote the respective space axes $x, y, z$; $u_{ik}$ is the strain tensor; and $\lambda_{iklm}$ is the elastic (modulus) tensor. By definition, the elastic tensor is symmetric under the transpositions $i \rightleftharpoons k$, $l \rightleftharpoons m$, and $i, k \rightleftharpoons l, m$,

$$\lambda_{iklm} = \lambda_{kilm} = \lambda_{ikml} = \lambda_{lmik},$$

and has 21 independent coefficients.

With (1), the (symmetric) stress tensor $\sigma_{ik}$ is defined as

$$\sigma_{ik} = \frac{\partial F}{\partial u_{ik}} = \lambda_{iklm}u_{lm}. \tag{2}$$

In a symmetric medium, there is a correlation between different components $\lambda_{iklm}$ and the number of independent elements of the tensor of elastic modulus is reduced.

We introduce a Cartesian coordinate system with the $z$ axis perpendicular to the unperturbed (flat) membrane plane and with the monolayer interface (i.e., the bilayer midplane) positioned in the $xy$ plane (at $z = 0$). The membrane thickness is equal to $2h$, and the flat membrane is modeled as a thin bilayer plate bounded by the $z = -h$ and $z = h$ planes with in-plane linear dimension $R \gg 2h$. The $xy$ plane is a plane of reflection symmetry. This implies that the free energy must be invariant under the transformation $x \longrightarrow x$, $y \longrightarrow y$, $z \longrightarrow -z$. Therefore, all the components $\lambda_{iklm}$ with an odd number of $z$ indices are equal to zero [10]. The membrane can be considered laterally isotropic. Then the $z$ axis is an axis of rotational symmetry. The expression for the elastic energy density $F$ then becomes [11]

$$\begin{aligned} F &= \frac{1}{2}\lambda_{xxxx}(u_{xx}^2 + u_{yy}^2) + \frac{1}{2}\lambda_{zzzz}u_{zz}^2 \\ &+ \lambda_{xxyy}u_{xx}u_{yy} + (\lambda_{xxxx} - \lambda_{xxyy})u_{xy}^2 \\ &+ \lambda_{xxzz}(u_{xx}u_{zz} + u_{yy}u_{zz}) + 2\lambda_{xzxz}(u_{xz}^2 + u_{yz}^2). \end{aligned} \tag{3}$$

Assuming that the membrane is in a liquid state, we require that the in-plane shear modulus (the coefficient in front of $u_{xy}^2$) vanish, and thus obtain

$$\lambda_{xxxx} = \lambda_{xxyy}.$$

Hence, expression (3) further simplifies and acquires the form

$$\begin{aligned} F &= \frac{1}{2}\lambda_{xxxx}(u_{xx} + u_{yy})^2 + \frac{1}{2}\lambda_{zzzz}u_{zz}^2 \\ &+ \lambda_{xxzz}(u_{xx}u_{zz} + u_{yy}u_{zz}) + 2\lambda_{xzxz}(u_{xz}^2 + u_{yz}^2). \end{aligned} \tag{4}$$

Let an external force applied perpendicular to the membrane plane induce a small bending deformation along the $z$ axis. Allowing for a typical experimental situation, we consider a thin membrane with the ratio of its thickness $2h$ to the lateral linear dimension (effective radius) $R$ of the order $10^{-3}$. Hence, we neglect the applied external stresses on the top and bottom membrane surfaces compared to the internal lateral stresses in it. Due to the smallness of the membrane thickness, zero stresses on the surface also vanish in the bulk of the membrane. We therefore impose the condition usually implied for the thin plates [13],

$$\sigma_{xz}(\mathbf{r}) = \sigma_{yz}(\mathbf{r}) = \sigma_{zz}(\mathbf{r}) \equiv 0, \tag{5}$$

where $r$ spans the membrane bulk. This condition is justified by the fact that small external pressure normal to a thin membrane induces relatively high lateral stresses in it [13]. Indeed, we show in Appendix A that the ratio of the normal stress to the lateral stress is on the order

of $(h/R)^2$. In accordance with (2) and (4), the $z$ components (5) of the stress tensor are related to the strain tensor components as

$$\sigma_{xz} = 4\lambda_{xzxz}u_{xz}, \quad \sigma_{yz} = 4\lambda_{yzyz}u_{yz}, \tag{6}$$

$$\sigma_{zz} = \lambda_{xxzz}(u_{xx} + u_{yy}) + \lambda_{zzzz}u_{zz}. \tag{7}$$

Combining relations (5) and (7), we find

$$u_{zz} = -\frac{\lambda_{xxzz}}{\lambda_{zzzz}}(u_{xx} + u_{yy}). \tag{8}$$

It is interesting to mention that, as follows from (6), the first two conditions in (5) require vanishing of the strain tensor components $u_{xz}$ and $u_{yz}$.

Condition (5) allows omission of the terms containing $u_{xz}$ and $u_{yz}$ in (4). Also using (8) and expressing $u_{zz}$ via $u_{xx} + u_{yy}$ in (4), we find the expression for the free energy density:

$$F = \frac{1}{2}\left(\lambda_{xxxx} - \frac{\lambda_{xxzz}^2}{\lambda_{zzzz}}\right)(u_{xx} + u_{yy})^2. \tag{9}$$

In addition, we impose the "incompressibility" condition, i.e., the constancy of the bulk density of the membrane:

$$u_{xx} + u_{yy} + u_{zz} = 0. \tag{10}$$

Condition (10) is satisfied simultaneously with (8) if $\lambda_{zzzz} = \lambda_{xxzz}$.

Finally, the free energy density is written as

$$F = \frac{1}{2}K_1(u_{xx} + u_{yy})^2, \tag{11}$$

where $K_1$ denotes a superposition of anisotropic elastic moduli:

$$K_1 = (\lambda_{xxxx} - \lambda_{zzzz}).$$

In the linear approximation for the strain tensor, we have

$$u_{ik} = \frac{1}{2}\left(\frac{\partial u_i}{\partial x_k} + \frac{\partial u_k}{\partial x_i}\right), \tag{12}$$

where $u_i$ is the $i$th component of the distortion field.

To introduce the essentials of our model in a simple way, we limit the following discussion to the case of a small bending amplitude, i.e., we impose condition

$$|u_z| \ll h,$$

where $u_z(\mathbf{r})$ is the $z$ component of displacement describing the deformed membrane. Also, we neglect the $z$ dependence of the component $u_z(\mathbf{r})$ in the thin plate approximation [13], thus defining the "shape" function $\xi(x, y) \approx u_z(\mathbf{r})$independent of the depth $z$. Substituting $\xi(x, y)$ in definition (12) and then in relations (6) and conditions (5), we obtain the partial differential equations

$$\frac{\partial u_x}{\partial z} = -\frac{\partial \xi}{\partial x}, \quad \frac{\partial u_y}{\partial z} = -\frac{\partial \xi}{\partial y}. \tag{13}$$

In integrating Eqs. (13), we introduce two functions: the (inhomogeneous) lateral stretching of the membrane $\mathbf{a}(x, y)$ and the in-plane slide $\pm\mathbf{f}(x, y)$ of the lower ($z < 0$) and upper ($z > 0$) monolayers at the midplane $z = 0$ of the membrane. Thus, the in-plane distortions $u_x$ and $u_y$ of each monolayer have the form

$$\begin{aligned} u_x &= -z\frac{\partial \xi(x, y)}{\partial x} \\ &+ (\Theta(z) - \Theta(-z))\, f_x(x, y) + a_x(x, y), \\ u_y &= -z\frac{\partial \xi(x, y)}{\partial y} \\ &+ (\Theta(z) - \Theta(-z))\, f_y(x, y) + a_y(x, y), \end{aligned} \tag{14}$$

where the step function is defined as

$$\Theta(z > 0) \equiv 1, \quad \Theta(z < 0) \equiv 0,$$

and the choice of the sign of $\Theta$ and of its argument is made for later convenience. The step functions in (14) model splitting of the membrane into two separate monolayers and describe a discontinuity of in-plane distortions across the interface between the monolayers.

Here it is worth emphasizing the limitations of the validity of relations (14). Expressions (14) are clearly distinct from the usual expressions for thin plates [13]. In the latter case, the displacements $u_x$ and $u_y$ are set to zero at $z = 0$, implying the presence of a neutral (not stretched) surface at the midplane of the plate in the small bending approximation $\xi \ll h$ [13]. It is shown in Appendix A (see Eq. (A.11)) that the second term in (14) is of the same order as the first one,

$$f_{x,y} \sim h\xi/R,$$

where $R$ is the effective radius of the membrane. The small bending approximation is justified when the term quadratic in $\xi$ is negligibly small compared to linear terms in the expressions for in-plane distortions $u_x$ and $u_y$,

$$O(\xi^2/R) \ll hx/R.$$

This condition is fulfilled as long as $\xi \ll h$. On the other hand, for a strongly bent thin plate, the $\xi^2$-term domi-

nates over the $\xi$-term and therefore higher-order terms should be added on the right-hand side of Eqs. (14).

We now discuss the physical meaning of expressions (14). The membrane stretching $\mathbf{a}(x, y)$ defines the position-dependent shift of the neutral surface (along the $z$ coordinate), while the slide function $\mathbf{f}(x, y)$ multiplied by step functions leads to the splitting of this neutral surface into two surfaces belonging to the upper and lower monolayers. These surfaces are determined from the conditions

$$u_x(x, y, z) \equiv 0, \quad u_y(x, y, z) \equiv 0.$$

The function $\mathbf{f}(x, y)$ provides an additional degree of freedom in comparison with a bilayer without slide (or a single monolayer). Under the condition of zero total lateral stretching (i.e., pure bending deformation, $\mathbf{a} \equiv 0$), the presence of the function $\mathbf{f}$ means that the neutral surface splits into two such surfaces located in each monolayer symmetrically with respect to the midplane $z = 0$. The total amplitude of the common interlayer slide at each point $(x, y)$ of the midplane is then given by $2\mathbf{f}(x, y)$, which signifies discontinuity of in-plane distortions $u_x$ and $u_y$ across the midplane $z = 0$. In the opposite case where $\mathbf{f} \equiv 0$, the monolayers are coupled together (no interlayer slide) and the distortion field is the sum of bending and stretching (for small deformations), the latter being continuous across the midplane $z = 0$. In general, distortion field (14) includes bending, stretching, and mutual interlayer slide.

Substituting (14) in (12), we proceed to determine the strain tensor components for each monolayer ($z > 0$, $z < 0$):

$$\begin{aligned} u_{xx} &= -z\frac{\partial^2\xi(x, y)}{\partial x^2} \\ &+ (\Theta(z) - \Theta(-z))\frac{\partial f_x(x, y)}{\partial x} + \frac{\partial a_x(x, y)}{\partial x}, \\ u_{yy} &= -z\frac{\partial^2\xi(x, y)}{\partial y^2} \\ &+ (\Theta(z) - \Theta(-z))\frac{\partial f_y(x, y)}{\partial y} + \frac{\partial a_y(x, y)}{\partial y}, \end{aligned} \tag{15}$$

and $u_{zz}$ can be expressed via $u_{xx}$ and $u_{yy}$ using (8).

The above expressions allow for a free (static) mutual slide of the monolayers. The jump of the lateral strain across the interface between the monolayers does not cost elastic energy. Hence, this jump does not introduce any additional spatial scale smaller than $h$ into the problem.

The free energy functional of the whole membrane $F_v$ is obtained by an integration of the free energy density $F$ over the membrane volume stepwise: first over the thickness coordinate ($-h < z < h$), and then over the membrane plane $\{x, y\}$. Using expressions (11) and (15), we finally find

$$\begin{aligned} F_v &= \frac{K_1}{2}\int(u_{xx} + u_{yy})^2 dV = \frac{K_1}{2} \\ &\times\left\{\frac{2h^3}{3}\iint(\tilde\nabla^2\xi)^2 dxdy - 2h^2\iint(\tilde\nabla^2\xi)(\tilde\nabla\cdot\mathbf{f})dxdy\right. \\ &\left. + 2h\iint((\tilde\nabla\cdot\mathbf{a})^2 + (\tilde\nabla\cdot\mathbf{f})^2)dxdy\right\}, \end{aligned} \tag{16}$$

where the tilde refers to two-dimensional differentiation:

$$\tilde\nabla = \left\{\frac{\partial}{\partial x}, \frac{\partial}{\partial y}\right\}.$$

Equation (16) is actually quite remarkable. The mean curvature of the interlayer surface $H$ is expressed as

$$\tilde\nabla^2\xi = \frac{\partial^2\xi}{\partial x^2} + \frac{\partial^2\xi}{\partial y^2} \approx 2H. \tag{17}$$

Therefore, the first term on the right-hand side gives the effective bending energy, i.e., the extrinsic curvature-bending energy functional $F_c$ of the "standard" form [5, 13]:

$$F_c = \frac{K_b}{2}\int(2H - c_0)^2 dS, \tag{18}$$

with zero spontaneous curvature $c_0$. Here, $K_b$ is the bending rigidity (modulus). Comparing (16) and (18), we find

$$2h^3K_1/3 = K_b.$$

The last term in (16) accounts for the elastic energy of area dilation with the area compression coefficient defined as

$$K_a = 2hK_1.$$

In general, the local relative area dilation $\Delta S/S$ equals $u_{xx} + u_{yy}$ [13]. According to Eq. (15), the relative area dilation is given by $\tilde\nabla\cdot\mathbf{a}$, while the difference of relative area dilations between the monolayers is given by $2(\tilde\nabla\cdot\mathbf{f})$. Hence, the $(\tilde\nabla\cdot\mathbf{a})^2$ term in (16) arises due to continuous (across the monolayers interface $z = 0$) lateral stretching of the membrane, which leads to the change in the average lipid density. The $(\tilde\nabla\cdot\mathbf{f})^2$ term represents the energy of local area difference of the monolayers (area-difference elasticity [2]), which is equivalent to the difference of lipid densities in monolayers (density-difference model [9]). In principle, this energy is not related to the presence of neutral surfaces within the monolayers (at large membrane stretch-

ing/compression, there are no neutral surfaces that would obey $u_x \equiv u_y \equiv 0$, see expression (14)). As is apparent from Eq. (16), the relation between bending and area compression coefficients (see [2])

$$K_b/K_A \sim h^2$$

occurs naturally in our derivation.

Next, the second term on the right-hand side of Eq. (16) expresses coupling between bending deformation and interlayer slide producing a local area dilation difference between monolayers. We note that in the lowest-order approximation, bending is decoupled from (continuous) area dilation caused by lateral stretching. Due to the hydrophobic effect, the monolayers, while sliding, are forced to stick together and to follow the same shape defined by $\xi(x, y)$ on the monolayer interface. Mutual interlayer slide along the interface leads to relaxation of stretching/compression of the monolayers caused by bending deformation and thus permits the free energy decrease.

Finally, our free energy functional is invariant with respect to transversal slide of monolayers such that $\mathrm{div}\mathbf{f} = 0$. Hence, the energy does not change under a mutual rotation of the monolayers (as a whole) or a position-independent shift of one of the monolayers with respect to the other.

We consider pure bending deformations of the membrane with no overall stretching. Therefore, we require the lateral strain integrated over the thickness to be zero at each point of the membrane. This imposes a restriction on the form of $u_x$ and $u_y$: the function $\mathbf{a}(x, y)$ must be equal to zero at every point of the bilayer. Hence, this function is omitted everywhere below. Then the strain tensor components can be written as

$$\begin{aligned} u_{xx} &= -z\frac{\partial^2\xi}{\partial x^2} + (\Theta(z) - \Theta(-z))\frac{\partial f_x}{\partial x}, \\ u_{yy} &= -z\frac{\partial^2\xi}{\partial y^2} + (\Theta(z) - \Theta(-z))\frac{\partial f_y}{\partial y}, \end{aligned} \tag{19}$$

and $u_{zz}$ can again be expressed via $u_{xx}$ and $u_{yy}$ using (8).

The free energy functional of the membrane acquires the form

$$F_v = \frac{K_1}{2}\int(u_{xx} + u_{yy})^2 dV = \frac{K_1}{2}\left\{\frac{2h^3}{3}\iint(\tilde\nabla^2\xi)^2 dxdy\right.$$
$$\left. - 2h^2\iint(\tilde\nabla^2\xi)(\tilde\nabla\cdot\mathbf{f})dxdy + 2h\iint(\tilde\nabla\cdot\mathbf{f})^2 dxdy\right\}. \tag{20}$$

To study the properties of functional (20) in detail, a simple problem with cylindrically symmetric deformation is discussed in Appendix A. The equilibrium of the membrane is defined by the Euler–Lagrange equations, which are obtained by equating to zero the first variational derivatives of the elastic energy functional $F(\xi, f)$ with respect to the functions $\xi(r)$ and $f(r)$.

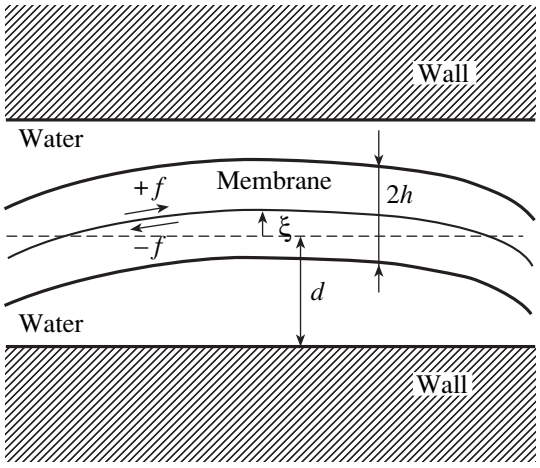



**Fig. 1.** Membrane in the confined geometry. A bilayer membrane (each monolayer of thickness $h$) is placed in a water solution between parallel walls separated by distance $2d$. The bending amplitude $\xi = u_z$ is defined at the midplane and is independent of the depth in the membrane. The interlayer slide function $f$ parameterizes position-dependent mutual slide of the monolayers at their interface.

## 3. CONFINING POTENTIAL FOR A BILAYER WITH SLIDE

Direct influence of confined geometry on the membrane behavior manifests itself in the reduction of the manifold of accessible membrane conformations. Entropic interactions of the membrane with confining walls (see Fig. 1) can be modeled [12] by introduction of an extra potential energy $W$ dependent on the bending amplitude,

$$W = \frac{\alpha}{2}\xi^2.$$

The free energy functional (20) appended with a confining potential $W$ acquires the form

$$F_v = \frac{K_1}{2}\left\{\frac{2h^3}{3}\iint(\tilde\nabla^2\xi)^2 dxdy\right.$$
$$- 2h^2\iint(\tilde\nabla^2\xi)(\tilde\nabla\cdot\mathbf{f})dxdy \tag{21}$$
$$\left. + 2h\iint(\tilde\nabla\cdot\mathbf{f})^2 dxdy\right\} + \frac{\alpha}{2}\iint\xi^2 dxdy.$$

The curvature of the confining potential at its minimum,

$$\alpha = \left.\frac{d^2W}{d\xi^2}\right|_{\xi=0},$$

is calculated below using a self-consistent procedure.

In the Fourier space $\mathbf{q} = \{q_x, q_y\}$, free energy functional (21) is written as

$$F_v = \int_0^\infty\int_0^\infty \left(K_1\frac{2h^3}{3}q^4 + \alpha\right)|\xi_q|^2\frac{dq_x dq_y}{(2\pi)^2}$$
$$-\int_0^\infty\int_0^\infty K_1 h^2 i q^2(\xi_q \mathbf{q}\cdot\mathbf{f}_q^* - \xi_q^*\mathbf{q}\cdot\mathbf{f}_q)\frac{dq_x dq_y}{(2\pi)^2} \qquad (22)$$
$$+\int_0^\infty\int_0^\infty 2K_1 h|\mathbf{q}\cdot\mathbf{f}_q|^2\frac{dq_x dq_y}{(2\pi)^2},$$

where

$$q^2 = q_x^2 + q_y^2.$$

We diagonalize the quadratic form in (22) with respect to $\xi_q$ and $\mathbf{q}\cdot\mathbf{f}_q$ by the linear transformation

$$\mathrm{Re}\tilde{\xi}_q = \mathrm{Re}\xi_q - \frac{3}{2h}\frac{q^2}{\tilde{q}^4}\mathrm{Im}(\mathbf{q}\cdot\mathbf{f}_q),$$
$$\mathrm{Re}\tilde{\xi}_q = \mathrm{Im}\xi_q + \frac{3}{2h}\frac{q^2}{\tilde{q}^4}\mathrm{Re}(\mathbf{q}\cdot\mathbf{f}_q), \qquad (23)$$

where

$$\tilde{q}^4 = q^4 + \frac{\alpha}{K_b}, \quad K_b = \frac{2h^3}{3}K_1.$$

In terms of the variables $\tilde{\xi}_q$ and $\mathbf{f}_q$, energy functional (22) becomes

$$F_v = \int_0^\infty\int_0^\infty \frac{K_1}{2}\left\{\frac{4h^3}{3}\tilde{q}^4|\tilde{\xi}_q|^2\right.$$
$$\left. + h\left(4 - 3\left(\frac{q}{\tilde{q}}\right)^4\right)|\mathbf{q}\cdot\mathbf{f}_q|^2\right\}\frac{dq_x dq_y}{(2\pi)^2}. \qquad (24)$$

Using relations (23) and functional (24), we calculate the thermodynamic average

$$\langle|\xi_q|^2\rangle = \frac{k_B T}{\frac{2h^3}{3}K_1 q^4 + \alpha}$$
$$+ \frac{3q^4 k_B T}{\left(\frac{2h^3}{3}K_1 q^4 + \alpha\right)\left(q^4 + \frac{6\alpha}{K_1 h^3}\right)}, \qquad (25)$$

where $k_B$ is the Boltzmann constant and $T$ is the temperature.

In the absence of interlayer slide, only the first term remains in Eq. (25), as obtained in [12, 15]. The second term in (25) signifies enhancement of the bending fluctuations caused by interlayer slide. The latter leads to relaxation of the lateral stresses (see Appendix A and Fig. 3 below) and thus to a decrease of the free energy of the bent membrane.

The mean-square fluctuations of the bending amplitude are found as

$$\langle\xi^2(\bar{r})\rangle = \int_0^\infty \langle|\xi_q|^2\rangle q\frac{dq}{2\pi} = \sqrt{\frac{3}{32}}\frac{k_B T}{\sqrt{\alpha K_1 h^3}}. \qquad (26)$$

In the confined geometry, the average bending amplitude is restricted to finite two-dimensional space between the walls (neglecting the volume occupied by the membrane itself, i.e., $h \ll d$), thus providing the self-consistency condition for determination of the effective rigidity $\alpha$,

$$\langle\xi^2\rangle = \mu d^2, \qquad (27)$$

where $\mu \le 1$.

Substituting (26) in (27), we obtain a self-consistent solution for $\alpha$:

$$\alpha = \frac{(k_B T)^2}{16\mu^2 d^4 K_b}. \qquad (28)$$

We here also evaluate the curvature $\alpha_0$ of the confining potential for a unit bilayer (without interlayer slide). In this case, the second term on the right-hand side of (25) is zero, and hence

$$\alpha_0 = \frac{(k_B T)^2}{64\mu^2 d^4 K_b}. \qquad (29)$$

Thus, interlayer slide results in considerable enhancement ($\alpha/\alpha_0 = 4$) of the curvature of the confining potential.

## 4. BILAYER DYNAMICS: VISCOUS MODES

To study the dynamical properties of the introduced model of bilayer membrane with interlayer slide, we here determine the equations of motion and find the eigenmodes of the membrane surrounded by water solution. We are interested in the behavior of the membrane confined between parallel walls (see Fig. 1).

Let a flat membrane lie in the *xy* plane with the normal pointed along the *z* axis. We treat each monolayer constituting the membrane as a (unit) two-dimensional condensed structure. We require the equilibrium between viscous stresses exerted on the membrane surface by water solution and the membrane restoring force. We neglect inertial effects and introduce three

sources of viscous dissipation: water and membrane viscosities and interlayer drag. The force balance equations are expressed as

$$-\frac{\delta F_s}{\delta \xi} + \Pi_{zz}(z = +0) - \Pi_{zz}(z = -0) = 0, \tag{30}$$

$$\frac{\delta F_s}{\delta f_x} - 2\eta_m h \frac{\partial}{\partial t}(\tilde{\nabla}^2 f_x) + 2b_s \frac{\partial f_x}{\partial t} - \Pi_{xz}(z = +0) - \Pi_{xz}(z = -0) = 0, \tag{31}$$

$$\Pi_{xz}(z = +0) - \Pi_{xz}(z = -0) = 0. \tag{32}$$

Here, the fluid stress tensor is defined as

$$\Pi_{ik} = -p\delta_{ik} + \eta_w\left(\frac{\partial v_i}{\partial x_k} + \frac{\partial v_k}{\partial x_i}\right),$$

where $p$ denotes hydrostatic pressure, $\mathbf{v}$ is the velocity, and $\eta_w$ is the viscosity of water solution. The fluid stress tensor is evaluated at the upper ($z = +0$) and lower ($z = -0$) membrane surfaces and carries the sign of the normal. The first term on the left-hand side of Eq. (30) is the elastic restoring membrane force, which is balanced by viscous stress of the fluid normal to the membrane surface. Equation (31) represents force balance in the lateral direction and contains the following contributions [3, 9]: (a) tangential traction on the interlayer surface due to differential flow of monolayers; (b) coherent surface flow of the monolayers as unit surfaces (with the dynamic viscosity $\eta_m$); (c) viscous drag between monolayers (characterized by the coefficient $b_s$) that arises at finite velocity of their mutual slide; (d) traction of the surrounding fluid. Equation (32) accounts for the absence of total stretching forces exerted by water on the membrane because we here discuss only pure bending deformations of the membrane, i.e., when the total area dilation is zero.

Navier–Stokes equations for water solutions surrounding the membrane should be added to balance Eqs. (30)–(32). In the small-velocity limit, treating fluid as incompressible and neglecting inertia, we write the "creeping flow" equations as

$$\nabla p = \eta_w \Delta v, \quad \nabla \cdot \mathbf{v} = 0. \tag{33}$$

The nonslip boundary conditions at the membrane–water interface provide the continuity of normal and tangential velocities of the fluid and the membrane:

$$\frac{\partial \xi}{\partial t} = v_z(\pm 0), \tag{34}$$

$$\frac{\partial f_i}{\partial t} = v_i(z = +0), \quad -\frac{\partial f_i}{\partial t} = v_i(z = -0), \quad i = x, y. \tag{35}$$

Confinement between parallel walls at the distance $2d$ implies vanishing of water velocity (normal and tangential components) at the walls surfaces:

$$v_j(z = \pm d) = 0, \quad j = x, y, z. \tag{36}$$

To find the dispersion relation, we make the Fourier transform of free energy functional (21) and of the force balance and creeping flow equations. For this, we expand the vibration in plane waves propagating along the $x$ direction. The free energy density $F_s(q, \omega)$ then takes the form

$$F_s(q, \omega) = \frac{K_1}{2}\left\{\frac{4h^3}{3}q^4|\xi_q|^2 - 2h^2(\xi_q f_q^* - \xi_q^* f_q)q^3 i + 4h|f_q|^2 q^2\right\} + \alpha|\xi_q|^2 \tag{37}$$

and

$$F_v = \frac{L_y}{(2\pi)^2}\int_{-\infty}^{+\infty}\int_{0}^{+\infty} F_s(q, \omega)dqd\omega, \tag{38}$$

where $L_y$ is the system dimension along the $y$ axis, and we have omitted the index $\omega$ in the subscripts of Fourier components.

Restoring membrane forces are given by functional derivatives of the free energy,

$$\frac{\delta F_s}{\delta \xi_q^*} = \left(K_1\frac{2h^3}{3}q^4 + \alpha\right)\xi_q + K_1 h^2 f_q q^3 i, \tag{39}$$

$$\frac{\delta F_s}{\delta f_q^*} = -K_1 h^2 \xi_q q^3 i + 2K_1 h q^2 f_q. \tag{40}$$

Fourier transforms of creeping flow Eqs. (33) for the components of water velocity and pressure

$$v_x = w_x(z)\exp(iqx - i\omega t), \quad v_z = w_z(z)\exp(iqx - i\omega t), \quad p = p_q(z)\exp(iqx - i\omega t)$$

are written as

$$iqw_x + \frac{\partial w_z}{\partial z} = 0,$$

$$iqp_q = \eta_w\left(-q^2 w_x + \frac{\partial^2 w_x}{\partial z^2}\right),$$

$$\frac{\partial p_q}{\partial z} = \eta_w\left(-q^2 w_z + \frac{\partial^2 w_z}{\partial z^2}\right).$$

We find the following solutions of differential equations (41) with normal velocity continuous at $z = 0$, also

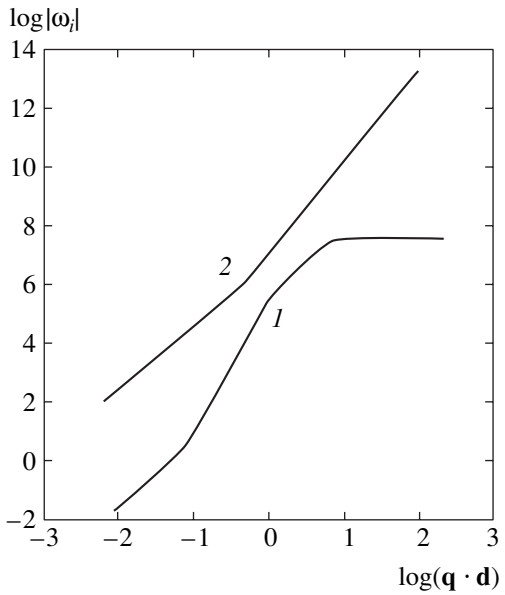



**Fig. 2.** Viscous modes of a bilayer membrane in a water solution confined between parallel walls in the case of pure bending deformations. Damping rates $|\omega_i|$ (1/s) are plotted as functions of the dimensionless parameter ($\mathbf{q} \cdot \mathbf{d}$), where $\mathbf{q}$ is the wave vector and $2d$ is the distance between the walls. Two branches *1* and *2* originate from bending and interlayer slide. The following values of parameters are used: $d = 10^{-6}$ cm, $h = 2 \times 10^{-7}$ cm, $\eta_w = 10^{-2}$ dyn s/cm$^2$, $\eta_m = 1$ dyn s/cm$^2$, $b_s = 10^7$ dyn s/cm$^3$, $K_1 = 2 \times 10^8$ erg/cm$^3$.

obeying the condition of zero lateral stretching force acting on the membrane (Eq. (32)) and the condition

$$v_x(z = +0) = -v_x(z = -0) \tag{41}$$

resulting from Eq. (35):

$$p_q = \begin{cases} 2\eta_w(C_1 e^{qz} + C_2 e^{-qz}), & z > 0 \\ 2\eta_w(-C_1 e^{-qz} - C_2 e^{qz}), & z < 0, \end{cases} \tag{42}$$

$$w_z = \begin{cases} [C_1 z + C_3]e^{qz} + [C_2 z + C_4]e^{-qz}, & z > 0 \\ [-C_1 z + C_3]e^{-qz} + [-C_2 z + C_4]e^{qz}, & z > 0, \end{cases} \tag{43}$$

$$w_x = \begin{cases} \left[C_1 z + C_3 + \dfrac{C_1}{q}\right] i e^{qz} \\ + \left[-C_2 z - C_4 + \dfrac{C_2}{q}\right] i e^{-qz}, \quad z > 0, \\ \left[-C_2 z + C_4 - \dfrac{C_2}{q}\right] i e^{qz} \\ + \left[C_1 z - C_3 - \dfrac{C_1}{q}\right] i e^{-qz}, \quad z < 0. \end{cases} \tag{44}$$

This solution maintains the symmetry relations compatible with the confined geometry:

$$\begin{aligned} w_x\left(z, x + \frac{\pi}{q}\right) &= w_x(-z, x), \\ w_z\left(z, x + \frac{\pi}{q}\right) &= -w_z(-z, x). \end{aligned} \tag{45}$$

The physical meaning of Eq. (45), according to the definitions given before (41), is that the $x/z$ component of water velocity around a vibrating membrane behaves symmetrically/antisymmetrically under simultaneous translation by the half-period ($x \longrightarrow x + \pi/q$) along the wave propagation direction $x$ and mirror reflection in the midplane between the confining walls ($z \longrightarrow -z$).

We next eliminate the unknown coefficients $C_2$ and $C_4$ using stick boundary conditions at the walls, Eq. (36). We then substitute the solutions in form (42)–(44) into Fourier-transformed force balance equations (30) and (31) (exploiting (39) and (40)) and into nonslip conditions (34) and (35) at the water–membrane interface. Thus, we finally obtain the algebraic system of four linear homogeneous equations for the unknowns $C_1$, $C_3$, $\xi_q$, and $f_q$:

$$\begin{aligned} &\xi_q\left[-\frac{2h^3}{3}K_1 q^4 - \alpha\right] + f_q[-K_1 h^2 i q^3] \\ &+ C_1[-4\eta_w 2q^2 d^2 e^{2qd}] \\ &+ C_3[4\eta_w q(1 + e^{2qd} - 2qde^{2qd})] = 0, \end{aligned} \tag{46}$$

$$\begin{aligned} &\xi_q(-h^2 K_1 q^3 i - 4\eta_w q\omega) \\ &+ f_q(2K_1 hq^2 - 2\eta_m hq^2 i\omega - 2b_s i\omega) \\ &+ C_1[-4\eta_w i(1 + e^{2qd} + 2qde^{2qd})] \\ &+ C_3[-4\eta_w i 2qe^{2qd}] = 0, \end{aligned} \tag{47}$$

$$i\omega\xi_q + 2qd^2 e^{2qd} C_1 + C_3(1 + 2qde^{2qd} - e^{2qd}) = 0, \tag{48}$$

$$\begin{aligned} &\omega f_q + \frac{C_1}{q}(1 - e^{2qd} - 2qde^{2qd} - 2q^2 d^2 e^{2qd}) \\ &+ C_3(1 - 2qde^{2qd} - e^{2qd}) = 0. \end{aligned} \tag{49}$$

The dispersion relation $\omega(q)$ is found by equating the determinant of system (46)–(49) to zero. The latter gives a quadratic equation for $\omega(q)$ that results in two branches $\omega_1(q)$ and $\omega_2(q)$ (see Fig. 2). Two viscous modes—the hydrodynamically damped bending mode and the intermonolayer slipping mode—mix, and the power law $\omega(q)$ changes with the wavelength of fluctuations. For pure bending deformation of the membrane, there exist up to four hydrodynamic regimes (depending on the parameters of the system), separated by three crossover wave vectors.

We use the result in (28) to estimate the upper limit $q_0$ of the smallest $q$ interval where the eigenmodes are modified by the confining potential, i.e., where the induced rigidity term ($\sim\alpha$) dominates over the bending term ($\sim K_b q^4$) in (22) (and in the first bracket in Eq. (46)):

$$q \ll q_0 \equiv \left(\frac{\alpha}{K_b}\right)^{1/4} = \left(\frac{k_B T}{K_b}\right)^{1/2} \frac{1}{2d}\left(\frac{1}{\mu}\right)^{1/2}. \quad (50)$$

For a typical value of bending rigidity at room temperature [2]

$$K_b \approx 25 k_B T,$$

we obtain

$$q_0 \sim \frac{0.1}{d}.$$

The second crossover wave vector $1/d$ bounds the long-wavelength regime where confinement of the surrounding water between the walls affects membrane dynamics. For $q \gg 1/d$, the membrane behaves as in the bulk water solution. We assume that the distance between confining walls is much greater than the monolayer thickness ($2d/h \sim 10$). The crossover wave vector for the bulk fluid $q_1$ (see Appendix B) for the chosen parameters $h = 2 \times 10^{-7}$ cm, $\eta_w = 10^{-2}$ dyn s/cm$^2$, and $b_s = 10^7$ dyn s/cm$^3$ acquires the value

$$q_1 = \frac{\eta_w}{b_s h^2} \sim 10^5 \text{ cm}^{-1}$$

and therefore obeys the condition $q_1 \ll 1/d$. Therefore, it does not influence dynamic behavior of the membrane in the confined geometry. In the interval of even shorter wavelengths, there is one more crossover wave vector

$$q_2 = \sqrt{\frac{b_s}{\eta_m h}} \approx \frac{10^7}{\sqrt{2}} \text{ cm}^{-1}$$

($\eta_m = 1$ dyn s/cm$^2$), which obeys the condition $1/d \ll q_2$. Hence, we investigate four intervals of wave vector values:

$$q \ll q_0,$$

$$q_0 \ll q \ll 1/d,$$

$$1/d \ll q \ll q_2,$$

$$q_2 \ll q.$$

For long wavelengths, $q \ll 1/d$, confinement between the walls modifies the bending mode with respect to the membrane in the bulk solution (see Appendix B),

$$\omega_1^B = -iq^3 \frac{K_1 h^3}{24\eta_w} \approx -iq^3 \frac{K_b}{\eta_w}, \quad (51)$$

and results in either $q^2$- or $q^6$-dependences of $\omega_1$ instead of the $q^3$-dependence of the "bulk" mode. For $q \ll q_0$, the bending mode becomes (to be compared with [16])

$$\omega_1 = -iq^2 \frac{\alpha d^3}{24\eta_w}. \quad (52)$$

The mode $\omega_1(q)$ is driven by the entropic potential, characterized by curvature $\alpha$, and is damped by viscous losses in the surrounding fluid. It is interesting to mention that, for a bilayer with interlayer slide, $\alpha$ is four times greater than for a unit bilayer (see Eqs. (28) and (29)). Thus, interlayer slide leads to faster dynamics of the membrane.

For $q_0 \ll q \ll 1/d$, the hydrodynamically damped bending mode is given by

$$\omega_1 = -iq^6 \frac{K_1 h^3 d^3}{144\eta_w} \approx -iq^6 \frac{K_b d^3}{\eta_w}. \quad (53)$$

In this wave vector interval, the finite thickness $d$ of water layers effectively enhances water viscosity from $\eta_w$ to $\eta_w/(qd)^3 \gg \eta_w$. Result (47) coincides (modulo a numeric coefficient) with the damped vibration mode of erythrocyte walls consisting of two membranes with liquid between them [6].

On the other hand, intermonolayer slipping mode $\omega_2(q)$, damped by viscous drag at the monolayer mutual interface, remains unchanged by confinement (see Appendix B):

$$\omega_2 = -iq^2 \frac{K_1 h}{b_s} \approx -iq^2 \frac{K_A}{b_s}. \quad (54)$$

For a membrane in the bulk solution, mixing of the bending and slipping modes occurs at $q \approx q_1$ (see Appendix B). The relative order of the parameters $q_1$, $1/d$, and $q_2$ by increasing value depends on the choice of characteristic parameters of the system. With our choice, $q_1 \ll 1/d$ and the mixing of the modes is delayed up to $q \approx 1/d$ (see Fig. 2). We speculate that this happens because confinement hinders bending fluctuations and therefore the bending mode remains slower than the slipping mode up to $q \approx 1/d$.

In the short-wavelength limit $q \gg 1/d$ we recover, as expected, the result for a membrane in the bulk water. Confinement is not revealed in this case because membrane-induced vibrations of water decay exponentially before reaching the walls. Namely, for $q \gg 1/d$, the branch $\omega_2(q)$ now corresponds to the bending mode damped by viscous losses in the surrounding fluid:

$$\omega_2^B = -iq^3 \frac{K_1 h^3}{6\eta_w} \approx -iq^3 \frac{K_b}{\eta_w}. \quad (55)$$

The renormalized bending rigidity ($\sim K_1 h^3$) arises for high-frequency fluctuations (to be compared with the numeric coefficients in (51) and (55)) because the bending mode is faster than the interlayer slipping

mode [9, 10]; interlayer slide leading to relaxation of lateral stresses in monolayers is retarded. In the interval $1/d \ll q \ll q_2$, the branch $\omega_1(q)$ becomes the interlayer slipping mode with a renormalized area compression modulus (superscript $B$ below indicates that the solution coincides with the bulk water case):

$$\omega_1^B = -iq^2\frac{K_1 h}{4b_s} \approx -iq^2\frac{K_A}{b_s}. \quad (56)$$

Finally, for $q \gg q_2$, the $\omega_1(q)$ mode is driven by the (high-frequency) effective rigidity $K_1$ and is damped by the monolayer surface viscosity $\eta_m$, which dominates over interlayer drag as the monolayers are dynamically coupled:

$$\omega_1^B = -i\frac{K_1}{4\eta_m}. \quad (57)$$

Viscous modes for a membrane in confined geometry obtained in this paper qualitatively agree with the results for a membrane bound to substrate [10]. We have not included the membrane tension into our free energy functional, because in the considered limit of small bending deformations of the bilayer, the term proportional to gradient of the bending amplitude is negligible [15].

The dispersion relation for a bilayer membrane in the bulk water based on our free energy functional (20) is derived in Appendix B and also agrees with earlier results obtained using the density-difference model [9].

## 5. CONCLUSIONS

A novel free energy functional of a bilayer fluid membrane derived in this paper reflects important physical properties of the membrane defining its dynamic behavior. The functional allows for a two-dimensional liquid-crystalline structure of the membrane and weak adherence between the monolayers constituting it, leading to their mutual slide under (bending) deformations. Our free energy functional contains three coupled fields parameterizing degrees of freedom related to bending of the membrane, interlayer mutual slide, and area dilation.

Using this functional, we have self-consistently calculated the curvature of the effective entropic potential acting on the membrane between two parallel confining walls. We found that the curvature at the potential minimum (located at the middle between the walls) is enhanced four times for a bilayer with interlayer slide in comparison with a unit membrane (with forbidden slide) of the same thickness. This leads to faster dynamics. This increase can be ascribed to a (partial) decrease of the lateral stress in the bent membrane via interlayer slide (static softening of the membrane). The relaxation of stresses effectively lowers the energetic "cost" of membrane bending and increases the thermodynamic probability for conformations with greater bending amplitudes. This in turn amplifies entropic repulsion.

We have also calculated the dispersion relations for a membrane confined between parallel walls. Our results are in qualitative agreement with those for a membrane bound to a substrate [10]. Confinement modifies the viscous modes $\omega(q)$ at long wavelengths compared to the bulk water case. We have found four wave-vector intervals separated by three characteristic wave-vector values, $q_0 \ll 1/d \ll q_2$, defined in Section 4. The inverse of the half-distance $d$ between the confining walls divides the $q$ axis into two intervals with confined ($q \ll 1/d$) and bulk ($q \gg 1/d$) behavior, respectively. The wave vector $q_0$ delimits the interval of $q$ values in which the entropic potential modifies the spectrum of bending modes (see also [16]). In the interval $q_0 \ll q \ll 1/d$, we found the dependence of bending mode

$$\omega(q) \propto q^6,$$

similar to peristaltic modes of a soap film [6]. Unlike in [10], we do not obtain the dependence

$$\omega(q) \propto q^4,$$

because the overall membrane tension is not included in our free energy functional. Since we consider the limit of small bending deformations of a flat bilayer, the term proportional to the gradient of the bending amplitude vanishes [15]. In the interval $q \gg 1/d$, confinement is not important because membrane-induced vibrations of water decay exponentially before reaching the walls. At $q > q_2$, as in the bulk case, the monolayer surface viscosity $\eta_m$ dominates over interlayer drag and the monolayers become dynamically strongly coupled.

Finally, we mention some limitations and possible improvements of our approach. Our functional respects reflectional symmetry of a flat bilayer and therefore implies that spontaneous curvature is zero. We assumed a thin-plate approximation for each monolayer with constant elastic moduli. In other words, we developed a phenomenological effective medium model. Hence, only fluctuations with wavelengths larger than the intermolecular distance in a lipid monolayer are considered. We have exploited smallness of the bending-to-thickness ratio using linear approximation for the stress tensor. In the small-bending approximation, area dilation is decoupled from bending. In this paper, we discussed only pure bending deformations, but the area dilation dynamics can also be studied using our functional. We found only damped eigenmodes of the membrane in a confined water solution. The propagating modes will be considered elsewhere.

## ACKNOWLEDGMENTS

The authors are grateful to Prof. R. Bruinsma for the formulation of the problem and to Prof. Yu.A. Chizmadzhev and his coworkers for useful comments. The

work of S.B. is supported by the Nonprofit Foundation "Dynasty" and a PhD student grant (A03-2.9-283) from the Russian Ministry of Education.

## APPENDIX A

### *Analytic Solutions in the Axially Symmetric Case*

We can obtain analytic results describing the equilibrium shape of and mutual monolayer slide in the bilayer lipid membrane under constant external pressure in the cylindrically symmetric case. We consider a flat (unperturbed) circular membrane in the plane *xy* of the radius *R*. We search for an equilibrium solution independent of the polar angle $\phi$,

$$\xi = \xi(r), \tag{A.1}$$

where *r* is the radial coordinate in the reference system with the origin at the center of the unperturbed membrane midplane and with the *z* axis directed along the membrane normal. Hence, the slide functions take the form

$$f_x(x, y) = f(r)\cos\phi, \quad f_y(x, y) = f(r)\sin\phi, \tag{A.2}$$

which then leads to the following expression for the radial component of the distortion field:

$$u_r(r, z) = -z\frac{\partial\xi(r)}{\partial r} + (\Theta(z) - \Theta(-z))\, f(r). \tag{A.3}$$

Because the deformation is purely radial, the angular component of the distortion is zero:

$$u_\phi = 0.$$

The symmetry of distortion fields (A.1) and (A.2) allows expressing the free energy density (11) in the cylindrical coordinates as follows:

$$F = \frac{K_1}{2}(u_{rr} + u_{\phi\phi})^2, \tag{A.4}$$

where

$$u_{rr} = \frac{\partial u_r}{\partial r}, \quad u_{\phi\phi} = \frac{u_r}{r} + \frac{1}{r}\frac{\partial u_\phi}{\partial\phi} = \frac{u_r}{r}. \tag{A.5}$$

The equilibrium of the membrane under pressure is defined by the Euler–Lagrange equations, which are obtained by equating to zero the first variational derivatives of the elastic energy functional $F(\xi, f)$ with respect to the functions $\xi(r)$ and $f(r)$ entering $u_{rr}$ and $u_{\phi\phi}$ in accordance with (A.5) and (A.3),

$$\begin{aligned}\frac{\partial F_{sr}}{\partial\xi(r)} - 2\pi rP &= 0,\\ \frac{\delta F_{sr}}{\delta f(r)} &= 0,\end{aligned} \tag{A.6}$$

where

$$F_v = \int_0^R F_{sr}\, dr,$$

*P* is the *z* component of the external pressure difference applied to the opposite sides of the membrane.

Equations (A.6) can be decoupled by the introduction of the new unknown functions $p(r)$ and $g(r)$ instead of $\xi$ and $f$:

$$p = \frac{4h}{3}\frac{\partial\xi}{\partial r} - 2f, \quad g = h\frac{\partial\xi}{\partial r} - 2f. \tag{A.7}$$

In the new basic set of functions $\{p, q\}$, Eqs. (A.6) become form

$$\begin{aligned}r^3p''' + 2r^2p'' - rp' + p &= P_1r^3,\\ r^2g'' + rg' - g &= 0,\end{aligned} \tag{A.8}$$

where

$$P_1 = \frac{2P}{K_1h^2}.$$

Both equations in (A.8) belong to the Euler class of equations and can be solved analytically using the transformation of the variable,

$$r = e^x,$$

where $-\infty < x < \infty$ is the new variable.

The following boundary conditions are imposed.

(1) $(p''(r)r + p'(r) - p(r)/r)|_{r=0} = 0$, the bending amplitude $\xi(r)$ is arbitrary at $r = 0$;

(2) $\xi(R) = 0$, membrane is fixed at the edge (no vertical displacement);

(3) $(p'(r)r + p(r))|_{r=R} = 0$, zero torque at the membrane edge;

(4) $\partial\xi/\partial r|_{r=0} = 0$, the slope at the center is zero;

(5) $f(0) = 0$, no intermonolayer slide at the center (axial symmetry);

(6) $g'(r)r + g(r)|_{r=R} = 0$, the intermonolayer slide at the edge is arbitrary.

These conditions have transparent physical meaning. Conditions (1) and (3) originate from the expression for the variational derivative $\partial F_{sr}/\partial\xi$, and condition (6) arises in the variational derivative $\delta F_{sr}/\delta f$; both derivatives include integration by parts in the segment $[0 \le r \le R]$. In particular, condition (1) is obtained by equating the prefactor in front of $\delta\xi(r = 0)$ to zero. Condition (3) is derived by equating the prefactor in front of $\partial\xi/\partial r|_{r=R}$ to zero, which in turn corresponds to zero

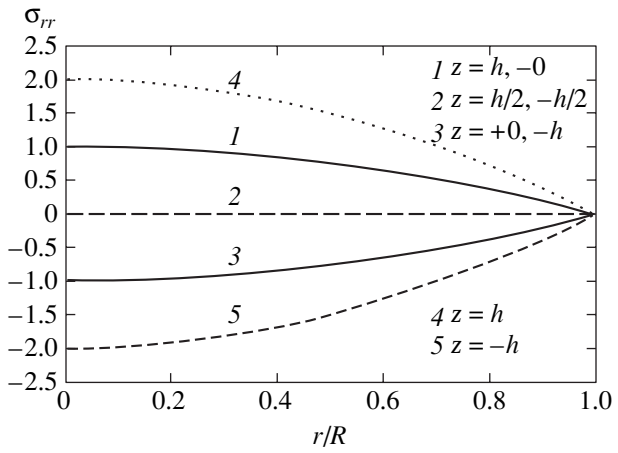



**Fig. 3.** Lateral stress $\sigma_{rr}$ normalized by $3PR^2/4h^2$ for various $z$-positions inside the membrane ($z > 0$ for the upper monolayer and $z < 0$ for the lower) is plotted as a function of the radial coordinate $r$ (in dimensionless units). The solid lines show stresses in the upper and lower monolayers (the stress profiles along the $z$ axis in both monolayers coincide due to interlayer slide, see text). The dashed line represents two neutral surfaces (at $z = \pm h/2$) in the lower and upper monolayers. The dotted lines ($z = h$, $z = -h$) characterize the stresses in the membrane of the same thickness $2h$ with forbidden slide.

torque $M$ at the membrane edge (hence, the membrane slope at the edge is arbitrary):

$$M = K_1\pi h^2 \times\left(\frac{4}{3}h\left(\frac{\partial^2\xi}{\partial r^2}r + \frac{\partial\xi}{\partial r}\right) - 2\left(\frac{\partial f}{\partial r}r + f\right)\right). \quad \text{(A.9)}$$

Condition (2) models the fixation of the membrane at the periphery. Condition (4) implies a smooth shape at the center of the curved membrane. The resulting solutions are

$$\xi(r) = \frac{3P}{32K_1h^3}(r^4 - 4R^2r^2 + 3R^4),$$
$$f(r) = \frac{3P}{16K_1h^2}(r^3 - 2R^2r). \quad \text{(A.10)}$$

The bending amplitude $\xi(r) = u_z(r)$ is defined at the interface (midplane) of the membrane and is $z$-independent (for small bending of the membrane considered here). The function $f(r)$ characterizes the amplitude of mutual slide of the monolayers at the interface of the membrane ($z = 0$) (the total amplitude is given by $2f$). As a result of this slide, the bottom surface of the upper monolayer is compressed, and the top surface of the lower monolayer is expanded. In the present approximate approach, $f$ is constant along the thickness (along the $z$ axis) of the monolayers and depends on the position in the plane of the membrane. It is apparent from (A.10) that

$$f \sim h\xi/R.$$

Substituting (A.10) in expression (A.3) for the radial distortion $u_r$, we find

$$u_r = \frac{3P}{16K_1h^2}\left(-z\frac{2}{h} - \Theta(-z) + \Theta(z)\right)(r^3 - 2R^2r). \quad \text{(A.11)}$$

The radial stress component corresponding to the distortion given by (A.11) is readily found as

$$\sigma_{rr}(r, z) = \frac{3P}{4h^2}\left(-z\frac{2}{h} - \Theta(-z) + \Theta(z)\right)(r^2 - R^2). \quad \text{(A.12)}$$

It is important to mention here that the lateral stress component $\sigma_{rr}$ in (A.12) proves to be independent of the elastic modulus ($K_1$) in our weak bending approximation. On the other hand, the distortion and slide fields and the strain tensor components depend on the elastic modulus.

In the considered case of a small bending amplitude, there is no overall stretch of the deformed membrane (i.e., pure bending occurs) and thus at any $r$,

$$\int_{-h}^{h}\sigma_{rr}(r)dz \approx \int_{-h}^{h}[u_{rr}(r) + u_{\phi\phi}(r)]dz = 0. \quad \text{(A.13)}$$

The $z$-dependent factor in Eq. (A.12) guarantees that Eq. (A.13) is satisfied. Condition (A.13) is kept by the equality of the factors in front of $\Theta(+z)$ and $\Theta(-z)$ (i.e., the function $f$ is taken to be the same in both monolayers). Simultaneously, stretching deformation of the monolayers is equal to zero, $\mathbf{a} \equiv 0$, in the definitions of the distortion field components (see expression (14) in Section 3). In general, if the problem is not restricted to the weak bending deformation and/or if there are additional forces acting in the lateral direction (stretching the membrane), we may introduce $\mathbf{a}(\mathbf{r}) \neq 0$ or use two functions $f_1 \neq f_2$ in front of $\Theta(+z)$ and $\Theta(-z)$, respectively.

Results of analytic solution of the static equations in the cylindrically symmetric case are presented in Fig. 3. The lateral stress $\sigma_{rr}(r, z)$ is shown for several values of the $z$ coordinate for a bilayer with mutual interlayer slide (solid lines and dashed lines) and for a unit bilayer with forbidden slide but of the same thickness $2h$ (lines *4*, *5*). Relaxation of lateral stresses in both monolayers is induced by mutual interlayer slide. The neutral (not stretched) surface at the interface of the membrane splits into two. Consequently, a neutral surface (with vanishing lateral stress) appears in the middle of each monolayer: at $z = +h/2$ (the upper monolayer) and at $z = -h/2$ (the lower monolayer) (line *2*). The monolayers are deformed as if they were disconnected, independent layers, but still adjusted to the same shape defined at their mutual interface inside the membrane. Therefore,

the stress profiles along the $z$ axis coincide with each other in both monolayers. As a result, the stresses at the top and bottom external surfaces of the membrane ($z = \pm h$, lines *1*, *3*) decrease by half compared to the case without slide ($z = \pm h$, lines *4*, *5*). Simultaneously, as follows from (A.12), the lateral stresses at the boundary $r = R$ turn to zero through the whole depth at the membrane, $\sigma_{rr}(R, z) = 0$, corresponding to the absence of the applied external stretching forces.

## *APPENDIX B*

### *Bilayer Modes in the Bulk Water*

To test the relevance of our approach to describe the dynamical properties of a bilayer, we here rederive the dispersion relation for a membrane in the bulk water solution using our free energy functional (20) introduced in Section 2. Our results agree with the previous ones obtained for a membrane in the bulk fluid using the curvature elastic model [6] and the density-difference model [9].

For the surrounding bulk fluid, we search for the solution of creeping flow equations (33) (Section 4) satisfying the nonslip conditions at the membrane–water interface, Eqs. (34) and (35). In addition, we impose the boundary conditions for fluid velocity components $v_i$,

$$v_j(z = \pm\infty) = 0, \quad j = x, y, z, \tag{B.1}$$

which require the fluid velocity field to vanish at large distances from the membrane.

As in Section 4, we expand vibrations in plane waves propagating along the $x$ axis. We make a Fourier transform of free energy functional (20). The free energy density $F_s(q, \omega)$ is written as

$$F_s(q, \omega) = \frac{K_1}{2}\left\{\frac{4h^3}{3}q^4|\xi_q|^2 - 2h^2(\xi_q f_q^* - \xi_q^* f_q)q^3 i + 4h|f_q|^2 q^2\right\}. \tag{B.2}$$

The components of water velocity and pressure in the form

$$v_x = w_x(z)\exp(iqx - i\omega t),$$
$$v_z = w_z(z)\exp(iqx - i\omega t),$$
$$p = p_q(z)\exp(iqx - i\omega t)$$

are substituted in Fourier-transformed creeping flow equations (41) (see Section 4). The solutions of differential Eqs. (41) satisfying boundary conditions (B.1), with the normal velocity continuous at $z = 0$ and also obeying the condition of zero lateral stretching force acting on the membrane (Eq. (32) in Section 4), are given by

$$p_q = \begin{cases} -2\eta_w C_1 e^{-qz}, & z > 0 \\ 2\eta_w C_1 e^{qz}, & z < 0, \end{cases} \tag{B.3}$$

$$w_z = \begin{cases} [-C_1 z + C_3]e^{-qz}, & z > 0 \\ [C_1 z + C_3]e^{qz}, & z < 0, \end{cases} \tag{B.4}$$

$$w_x = \begin{cases} \left[C_1 z - C_3 - \dfrac{C_1}{q}\right] i e^{-qz}, & z > 0 \\ \left[C_1 z + C_3 + \dfrac{C_1}{q}\right] i e^{qz}, & z < 0, \end{cases} \tag{B.5}$$

where the constants $C_2$ and $C_4$, which are present in Eqs. (42)–(44), turn to zero due to boundary conditions (B.1).

The unknown coefficients $C_1$ and $C_3$ are determined from nonslip conditions (34) and (35). Then, we substitute solutions (B.3)–(B.5) in Fourier transforms of force balance Eqs. (30) and (31) and obtain an algebraic system of two linear homogeneous equations for the components $\xi_q$ and $f_q$,

$$\xi_q\left[-\frac{2h^3}{3}K_1 q^4 + 4\eta_w i q w\right] + f_q[-K_1 h^2 i q^3] = 0, \tag{B.6}$$

$$\xi_q(-h^2 K_1 q^3 i) + f_q(2K_1 h q^2 - 2i\omega(\eta_m h q^2 + 2\eta_w q + b_s)) = 0. \tag{B.7}$$

Equating the determinant of this system to zero, we obtain a quadratic equation for $\omega(q)$, which results in two branches $\omega_1(q)$ and $\omega_2(q)$. There are three hydrodynamic regimes,

$$q \ll q_1, \quad q_1 \ll q \ll q_2, \quad q_2 \ll q,$$

separated by crossover wave vectors $q_1$ and $q_2$ [9]:

$$q_1 = \frac{\eta_w}{b_s h^2}, \quad q_2 = \sqrt{\frac{b_s}{\eta_m h}}. \tag{B.8}$$

For long wavelengths, $q \ll q_1$, the dispersion relations are given by

$$\omega_1^B = -iq^3\frac{K_1 h^3}{24\eta_w} \approx -iq^3\frac{K_b}{\eta_w}, \tag{B.9}$$

$$\omega_2^B = -iq^2\frac{K_1 h}{b_s} \approx -iq^2\frac{K_A}{b_s}, \tag{B.10}$$

which describe respectively the hydrodynamically damped bending mode $\omega_1^B(q)$ and the intermonolayer slipping mode $\omega_2^B(q)$ damped by viscous drag at the

membrane midplane. Here, the superscript "*B*" is introduced to label membrane modes in the bulk fluid.

For wave vectors in the interval $q_1 \ll q \ll q_2$, the bending and slipping modes mix [9],

$$\omega_1^B = -iq^2\frac{K_1 h}{4b_s} \approx -iq^2\frac{K_A}{b_s}, \quad \text{(B.11)}$$

$$\omega_2^B = -iq^3\frac{K_1 h^3}{6\eta_w} \approx -iq^3\frac{K_b}{\eta_w}. \quad \text{(B.12)}$$

The branch $\omega_2^B(q)$ now corresponds to the bending mode damped by viscous losses in the surrounding fluid, and the branch $\omega_1^B(q)$ describes the damping of the slipping mode. The elastic moduli in (B.11) and (B.12) differ in general from those in (B.9) and (B.10), because high-frequency (bending) fluctuations occur at nonrelaxed monolayer surface densities [9].

In the short-wavelength limit, $q \gg q_2$, we obtain

$$\omega_1^B = -i\frac{K_1}{4\eta_m}, \quad \text{(B.13)}$$

$$\omega_2^B = -iq^3\frac{K_1 h^3}{6\eta_w} \approx -iq^3\frac{K_b}{\eta_w}. \quad \text{(B.14)}$$

The $\omega_1^B(q)$ mode is driven by the (high-frequency) effective rigidity $K_1$ and is damped by the monolayer surface viscosity $\eta_m$. Effective rigidity is induced by dynamic coupling of monolayers [3]. Monolayer surface viscosity overwhelms interlayer drag and becomes the main source of dissipation.

## REFERENCES


1. A. Ben-Shaul, in *Structure and Dynamics of Membranes*, Ed. by R. Lipowsky and E. Sackmann (Elsevier, Amsterdam, 1995), Chap. 7.
2. U. Seifert and R. Lipowsky, in *Structure and Dynamics of Membranes*, Ed. by R. Lipowsky and E. Sackmann (Elsevier, Amsterdam, 1995), Chap. 8.
3. A. Yeung and E. Evans, J. Phys. II **5**, 1501 (1995).
4. E. I. Kats and V. V. Lebedev, *Fluctuational Effects in the Dynamics of Liquid Crystals* (Springer, New York, 1993).
5. W. Helfrich, Z. Naturforsch. **28**, 693 (1973).
6. F. Brochard and J. F. Lennon, J. Phys. (Paris) **36**, 1035 (1975).
7. W. Helfrich, Z. Naturforsch. **33**, 305 (1977).
8. E. Evans and A. Yeung, Lipids Chem. Phys. **73**, 39 (1994).
9. U. Seifert and S. A. Langer, Europhys. Lett. **23**, 71 (1993).
10. M. Kraus and U. Seifert, J. Phys. II **4**, 1117 (1994).
11. U. Seifert, Phys. Rev. E **49**, 3124 (1994).
12. F. C. MacKintosh, Phys. Rev. E **50**, 2891 (1994).
13. L. D. Landau and E. M. Lifshitz, *Course of Theoretical Physics*, Vol. 7: *Theory of Elasticity*, 4th ed. (Nauka, Moscow, 1987; Pergamon, New York, 1980).
14. H. Kleinert, *Gauge Fields in Condensed Matter*, Vol. 2: *Stresses and Defects* (World Sci., Singapore, 1989).
15. S. A. Safran, *Statistical Thermodynamics of Surfaces, Interfaces and Membranes* (Perseus, New York, 1994).
16. E. I. Kats, V. V. Lebedev, and S. V. Malinin, Zh. Éksp. Teor. Fiz. **113**, 2096 (1998) [JETP **86**, 1149 (1998)].