Journal of Experimental and Theoretical Physics, Vol. 91, No. 3, 2000, pp. 433–448.
From Zhurnal Éksperimental'noĭ i Teoreticheskoĭ Fiziki, Vol. 118, No. 3, 2000, pp. 499–515.

NUCLEI, PARTICLES, AND THEIR INTERACTION

Testing Chiral Models of Vector Mesons in Multipion Decays of ρ(770) and ω(782)¶

N. N. Achasov* and A. A. Kozhevnikov**

Sobolev Institute for Mathematics, Siberian Division, Russian Academy of Sciences, Novosibirsk, 630090 Russia
**e-mail: achasov@math.nsc.ru*
***e-mail: kozhev@math.nsc.ru*

Received April 11, 2000

Abstract—Predictions of the models based on effective chiral Lagrangians of vector mesons ρ(770) and ω(782) are considered for such measurable quantities as the partial widths of various $\rho \longrightarrow 4\pi$ and $\omega \longrightarrow 5\pi$ decays and their corresponding excitation curves. Analogous curves for the channel $\rho^- \longrightarrow (4\pi)^-$ are obtained for τ lepton decays. The angular distributions for various combinations of the final pions from the $\rho \longrightarrow 4\pi$ and $\omega \longrightarrow 5\pi$ decays in the e^+e^--annihilation and photoproduction are also found. A simple method for the order-of-magnitude estimation of the partial widths based essentially on the threshold character of the above multipion decays is presented. © *2000 MAIK "Nauka/Interperiodica".*

1. INTRODUCTION

In early times, the chiral pion dynamics was tested in the low-energy ππ scattering, with the scattering data extracted from the studies of the one pion exchange–dominated reaction $\pi N \longrightarrow \pi\pi N$. The main difficulty in interpreting the results was due to the necessity of extrapolating to the pion pole. The sources of soft pions that are feasible at present are unconventional from the point of view of the chiral pion dynamics and do not have the drawbacks of the classical ππ scattering. Indeed, progress in increasing the intensity of low-energy e^+e^- colliders (φ factories) and photon beams, as well as the huge number of the specific hadronic decays of τ leptons, offers naturally controlled sources of soft pions, provided the sufficiently low invariant mass regions of hadronic systems are isolated. Since the yield of pions is considerably larger when they are produced through the proper vector-resonance states, the theoretical study of the multipion decay channels of the low-lying vector mesons ρ(770) and ω(782) becomes important [1].

In this paper, based on the Weinberg Lagrangian [2] obtained under the nonlinear realization of the chiral symmetry—or in modern terms, the Lagrangian of hidden local symmetry (HLS) [3]—we systematically consider the predictions resulting from the approach of [2, 3]. Specifically, the partial widths and resonance excitation curves are calculated for the reactions

$$e^+e^- \longrightarrow \rho^0 \longrightarrow 2\pi^+2\pi^- \text{ and } e^+e^- \longrightarrow \rho^0 \longrightarrow \pi^+\pi^-2\pi^0.$$

It is shown that the intensities of the above decays change twice as fast as the phase space variation when the energy varies inside the ρ width. This means that the e^+e^- annihilation offers an ideal tool for the study of such effects. The decay widths of the charged ρ meson

$$\rho^\pm \longrightarrow \pi^\pm 3\pi^0 \text{ and } \rho^\pm \longrightarrow 2\pi^\pm\pi^\mp\pi^0,$$

and of the ω meson

$$\omega \longrightarrow 2\pi^+2\pi^-\pi^0 \text{ and } \omega \longrightarrow \pi^+\pi^-3\pi^0$$

are also evaluated.

The paper is organized as follows. Section 2 contains the expressions for the $\rho \longrightarrow 4\pi$ amplitudes and the order-of-magnitude estimates of the $\rho \longrightarrow 4\pi$ partial widths. The results of calculation of the excitation curves and partial widths for different isotopic states of four pions are presented in Section 3. This is done for the e^+e^- annihilation, τ decays, and photoproduction. In addition, the angular distributions of the emitted pions are obtained for the e^+e^- annihilation and photoproduction. The partial widths of the decays $\omega \longrightarrow 5\pi$ are discussed in Section 4. The angular distributions of various combinations, chosen among five emitted pions, are obtained for the e^+e^- annihilation and photoproduction. Section 5 contains concluding remarks.

2. THE $\rho \longrightarrow 4\pi$ DECAY AMPLITUDES

2.1. Historical Background and General Remarks

The decay $\rho \longrightarrow 4\pi$ is a unique source of soft ($|\mathbf{p}| \sim m_\pi$) pions. It attracts much attention [4–7] because it can be used to study the chiral dynamics of the ρ meson interaction with multipion systems. As was found in [4, 5, 6], this decay must be rather strong, $B(\rho \longrightarrow 4\pi) \sim 10^{-4}$. A detailed analysis given in [7] revealed a number of shortcomings of the calculations in [5, 6] related with the actual violation of chiral invariance, in particular, the Adler condition [8] for soft

¶This article was submitted by the authors in English.

pions. The correct results based on the amplitudes obeying the Adler condition and obtained in [1, 7] correspond to $B(\rho \longrightarrow 4\pi) \approx 10^{-5}$. The large magnitude of the branching ratio $B(\rho \longrightarrow 4\pi) \sim 10^{-4}$ obtained in [4] is related, by all appearance, with a very rough method of calculation. A common drawback of papers [4–7] is that their authors evaluate the partial width at the energy level equal to the mass of the ρ meson, as if this were a genuine narrow resonance. However, because the width of the ρ resonance is rather large and $\Gamma_{\rho \to 4\pi}(E)$ rises rapidly with the energy increase even at energies inside the ρ peak, one tends to think that the magnitude of the 4π partial width at the ρ mass cannot adequately characterize the dynamics of the process. In this respect, the resonance excitation curve in the channel $e^+e^- \longrightarrow \rho^0 \longrightarrow 4\pi$ is of much interest, being a testing ground of various chiral models of the decay under consideration.

The HLS approach [3] permits one to include the axial mesons as well.[1] In the ideal treatment, under the assumption that

$$m_\rho \sim E \ll m_{a_1},$$

the difference between the models with and without the a_1 meson would amount to taking the higher derivatives into account, and the latter were small.[2] In real life, one has

$$m_{a_1}^2 - m_\rho^2 \sim m_\rho^2,$$

and the correction may appear to be appreciable even at the ρ mass. In fact, the calculation of [7] shows that the corrections amount to ~20–30% in the width. This implies, in particular, that the left shoulder of the ρ peak, where the contributions of higher derivatives vanish rapidly, is the best place to work. In this work, we do not take the a_1 meson into account.

2.2. The General Expression for the $\rho \longrightarrow 4\pi$ *Decay Amplitude*

The $\rho \longrightarrow 4\pi$ decay amplitudes are obtained from the Weinberg Lagrangian [2]

$$\begin{aligned}\mathcal{L} &= -\frac{1}{4}(\partial_\mu \boldsymbol{\rho}_\nu - \partial_\nu \boldsymbol{\rho}_\mu + g[\boldsymbol{\rho}_\mu \times \boldsymbol{\rho}_\nu])^2 \\ &+ \frac{m_\rho^2}{2}\left[\boldsymbol{\rho}_\mu + \frac{\boldsymbol{\pi} \times \partial_\mu \boldsymbol{\pi}}{2g f_\pi^2(1 + \boldsymbol{\pi}^2/4f_\pi^2)}\right]^2 \\ &+ \frac{(\partial_\mu \boldsymbol{\pi})^2}{2(1 + \boldsymbol{\pi}^2/4f_\pi^2)^2} - \frac{m_\pi^2 \boldsymbol{\pi}^2}{2(1 + \boldsymbol{\pi}^2/4f_\pi^2)},\end{aligned} \tag{2.1}$$

where $\boldsymbol{\pi}$, m_π and $\boldsymbol{\rho}$, m_ρ are, respectively, the isovector field and mass of the pion and the ρ meson and f_π = 92.4 MeV is the pion decay constant. The cross denotes the vector product in the isotopic space. The ρρρ coupling constant g and the ρππ coupling constant $g_{\rho\pi\pi}$ are related to the ρ mass and the pion decay constant f_π via the hidden local symmetry parameter a as [3]

$$\begin{aligned} g &= m_\rho / f_\pi \sqrt{a}, \\ g_{\rho\pi\pi} &= \sqrt{a} m_\rho / 2 f_\pi. \end{aligned} \tag{2.2}$$

We note that imposing the universality condition $g = g_{\rho\pi\pi}$ gives $a = 2$. This then leads to the so-called KSRF relation [10]

$$2 g_{\rho\pi\pi}^2 f_\pi^2 / m_\rho^2 = 1, \tag{2.3}$$

which beautifully agrees with the experiment. The ρππ coupling constant resulting from this relation is $g = g_{\rho\pi\pi} = 5.89$.

We first obtain the $\pi \longrightarrow 3\pi$ transition amplitudes that are necessary for the calculation of the multipion decays of vector mesons. They are found from the diagrams in Fig. 1a and are given by

$$\begin{aligned} M(\pi^+ \longrightarrow \pi^+_{q_1}\pi^+_{q_2}\pi^-_{q_3}) &= (1 + P_{12})\frac{1}{2f_\pi^2} \\ &\times \left\{-2(q_1, q_2) + a(q_1, q_2 - q_3)\left[1 - \frac{m_\rho^2}{D_\rho(q_2 + q_3)}\right]\right\}, \\ M(\pi^+ \longrightarrow \pi^+_{q_1}\pi^0_{q_2}\pi^0_{q_3}) &= (1 + P_{23})\frac{1}{2f_\pi^2} \\ &\times \left\{(q_3, q_1 - 2q_2) + a(q_3, q_1 - q_2)\left[1 - \frac{m_\rho^2}{D_\rho(q_1 + q_2)}\right]\right\}, \\ M(\pi^0 \longrightarrow \pi^+_{q_1}\pi^-_{q_2}\pi^0_{q_3}) &= (1 + P_{12})\frac{1}{2f_\pi^2} \\ &\times \left\{(q_1, q_2 - 2q_3) - a(q_1, q_2 - q_3)\left[1 - \frac{m_\rho^2}{D_\rho(q_2 + q_3)}\right]\right\}, \\ M(\pi^0 \longrightarrow \pi^0_{q_1}\pi^0_{q_2}\pi^0_{q_3}) &= -\frac{1}{f_\pi^2}[(q_1, q_2) + (q_1, q_3) + (q_2, q_3)], \end{aligned} \tag{2.4}$$

where P_{ij} is the interchange operator of the pion momenta q_i and q_j and

$$D_\rho(q) = m_\rho^2 - q^2 - i m_\rho^2 \left(\frac{q^2 - 4m_\pi^2}{m_\rho^2 - 4m_\pi^2}\right)^{3/2} \frac{\Gamma_{\rho\pi\pi}(m_\rho^2)}{\sqrt{q^2}} \tag{2.5}$$

[1] The problem of the inclusion of vector and axial mesons and photons in the framework of chiral theories has required considerable efforts; for example, [9]. It is solved in an elegant way in the approach based on hidden local symmetry [3].

[2] Taking the higher derivatives into account also requires accounting for the chiral loops, the task which is not yet fulfilled for vector mesons.

is the inverse ρ meson propagator.[3]

The diagrammatic representations of the $\rho \longrightarrow 4\pi$ decay amplitudes for different combinations of charges of the final pions are shown in Figs. 1b and 1c. By introducing the 4-vector ε_μ of the decaying ρ meson polarization, we can write the general expression for the amplitude as

$$M = \frac{g_{\rho\pi\pi}}{f_\pi^2}\varepsilon_\mu J_\mu,$$

where

$$\frac{g_{\rho\pi\pi}}{f_\pi^2} = \frac{\sqrt{a}m_\rho}{2f_\pi^3} \tag{2.6}$$

is obtained from Eq. (2.2). We now give the expressions for the current J_μ for all the decay modes considered here.

(1) The decay $\rho^0(q) \longrightarrow \pi^+(q_1)\pi^+(q_2)\pi^-(q_3)\pi^-(q_4)$. One has

$$\begin{aligned} J_\mu &= (1+P_{12})(1+P_{34}) \\ &\times\left\{-q_{1\mu}\left[\frac{1}{2}+\frac{a(q_2,q_3)-(a-2)(q_3,q_4)}{D_\pi(q-q_1)}\right]\right. \\ &+ q_{3\mu}\left[\frac{1}{2}+\frac{a(q_1,q_4)-(a-2)(q_1,q_2)}{D_\pi(q-q_3)}\right] \\ &\left.+ am_\rho^2(1+P_{13})\frac{q_{1\mu}(q_3,q_2-q_4)}{D_\pi(q-q_1)D_\rho(q_2+q_4)}\right\}, \end{aligned} \tag{2.7}$$

where $D_\pi(q) = m_\pi^2 - q^2$ is the inverse pion propagator.

(2) The decay $\rho^0(q) \longrightarrow \pi^+(q_1)\pi^-(q_2)\pi^0(q_3)\pi^0(q_4)$. In this case, one has $J_\mu = J_\mu^{nan} + J_\mu^{an}$, where

$$\begin{aligned} J_\mu^{nan} &= -(1-P_{12})(1+P_{34})q_{1\mu}\left\{\frac{1}{4}+\frac{1}{D_\pi(q-q_1)}\right. \\ &\times\left.\left[(a-1)(q_3,q_4)-(a-2)(q_2,q_3)+am_\rho^2\frac{(q_3,q_2-q_4)}{D_\rho(q_2+q_4)}\right]\right\} \\ &+(1+P_{34})\frac{m_\rho^2}{2D_\rho(q_1+q_3)D_\rho(q_2+q_4)} \\ &\times[(q_1+q_3-q_2-q_4)_\mu(q_1-q_3,q_2-q_4) \\ &-2(q_1-q_3)_\mu(q_1+q_3,q_2-q_4) \\ &+2(q_2-q_4)_\mu(q_2+q_4,q_1-q_3)] \end{aligned} \tag{2.8}$$

[3] Our notation for the Lorentz-invariant scalar product of two 4-vectors a and b is $(a, b) = a_0b_0 - \mathbf{a}\cdot\mathbf{b}$.

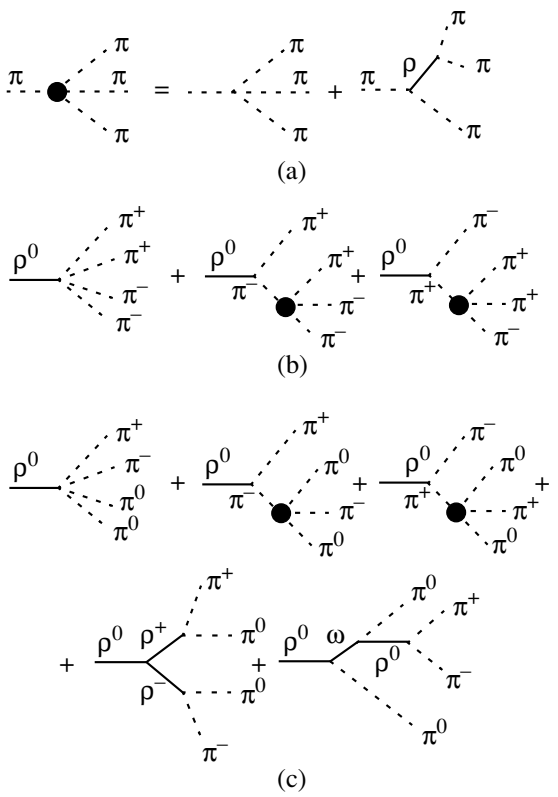

Fig. 1. (a) The diagrams describing the $\pi \longrightarrow 3\pi$ transition amplitude. The symmetrization over the momenta of identical pions is understood whenever necessary. (b), (c) the diagrams describe the amplitudes of the decays $\rho^0 \longrightarrow \pi^+\pi^-\pi^+\pi^-$ and $\rho^0 \longrightarrow \pi^+\pi^-\pi^0\pi^0$, respectively. The shaded circles in the $\pi \longrightarrow 3\pi$ vertices in diagrams (b) and (c) refer to the sum of the diagrams shown in (a). The symmetrization over the momenta of identical pions emitted from different vertices is implied. The diagrams for the decays $\rho^+ \longrightarrow \pi^+\pi^0\pi^0\pi^0$ and $\rho^+ \longrightarrow \pi^+\pi^+\pi^-\pi^0$ are similar to those of (b) and (c), respectively.

is obtained from Eq. (2.1) and

$$\begin{aligned} J_\mu^{an} &= 2\left(\frac{N_c g^2}{8\pi^2}\right)^2(1+P_{34})[q_{1\mu}(1-P_{23})(q,q_2)(q_3,q_4) \\ &+q_{2\mu}(1-P_{13})(q,q_3)(q_1,q_4) \\ &+q_{3\mu}(1-P_{12})(q,q_1)(q_2,q_4)] \\ &\times\left[\frac{1}{D_\rho(q_1+q_2)}+\frac{1}{D_\rho(q_1+q_3)}+\frac{1}{D_\rho(q_2+q_3)}\right]\frac{1}{D_\omega(q-q_4)}, \end{aligned} \tag{2.9}$$

with $D_\omega(q) = m_\omega^2 - q^2$ being the inverse ω meson propagator and $N_c = 3$ (the number of colors) is the contribution of the term induced by the anomalous Wess–Zumino Lagrangian [3, 11]

$$\mathcal{L}_{\omega\rho\pi} = \frac{N_c g^2}{8\pi^2 f_\pi}\varepsilon_{\mu\nu\lambda\sigma}\partial_\mu\omega_\nu(\boldsymbol{\pi}\cdot\partial_\lambda\boldsymbol{\rho}_\sigma), \tag{2.10}$$

manifesting in the process $\rho^0 \longrightarrow \omega\pi^0 \longrightarrow \pi^+\pi^-\pi^0\pi^0$. In general, J_μ^{an} is attributed to the contribution of higher derivatives. We nevertheless take it into account to show the effect of higher derivatives and the dynamical effect of the opening of the $\rho \longrightarrow \omega\pi \longrightarrow 4\pi$ channel. In agreement with [3], the contribution of the point vertex $\omega \longrightarrow 3\pi$ is omitted. The following amplitudes of the charged ρ decay are necessary for obtaining the $\omega \longrightarrow 5\pi$ decay amplitude and are interesting on their own for studying the reactions of peripheral ρ meson production and τ decays.

(3) The decay $\rho^+(q) \longrightarrow \pi^+(q_1)\pi^0(q_2)\pi^0(q_3)\pi^0(q_4)$. One has

$$J_\mu = (1 + P_{24} + P_{34})\left\{2q_{1\mu}\left[\frac{1}{3} + \frac{(q_2, q_3)}{D_\pi(q - q_1)}\right]\right.$$

$$- \frac{q_{4\mu}}{D_\pi(q - q_4)}\left[2(a-1)(q_2, q_3)\right. \qquad (2.11)$$

$$\left.\left. - (a-2)(q_1, q_2 + q_3) + am_\rho^2(1 + P_{23})\frac{(q_2, q_1 - q_3)}{D_\rho(q_1 + q_3)}\right]\right\}.$$

(4) The decay $\rho^+(q) \longrightarrow \pi^+(q_1)\pi^+(q_2)\pi^-(q_3)\pi^0(q_4)$. Here, the contribution induced by the anomalous Wess–Zumino Lagrangian is also possible; hence, $J_\mu = J_\mu^{nan} + J_\mu^{an}$, where

$$J_\mu^{nan} = (1 + P_{12})\left\{\frac{1}{2}(q_1 - q_4)_\mu\right.$$

$$+ \frac{q_{1\mu}(1 + P_{23})}{D_\pi(q - q_1)}[(a-1)(q_2, q_3) - (a-2)(q_2, q_4)]$$

$$- \frac{q_{4\mu}}{D_\pi(q - q_4)}[a(q_1, q_3) - (a-2)(q_1, q_2)]$$

$$- am_\rho^2\left[\frac{q_{1\mu}}{D_\pi(q - q_1)}(1 + P_{23})\frac{(q_2, q_3 - q_4)}{D_\rho(q_3 + q_4)}\right.$$

$$\left. + \frac{q_{4\mu}(q_1, q_2 - q_3)}{D_\pi(q - q_4)D_\rho(q_2 + q_3)}\right] + \frac{m_\rho^2}{2D_\rho(q_1 + q_3)D_\rho(q_2 + q_4)} \qquad (2.12)$$

$$\times[(q_1 + q_3 - q_2 - q_4)_\mu(q_1 - q_3, q_2 - q_4)$$

$$- 2(q_1 - q_3)_\mu(q_1 + q_3, q_2 - q_4)$$

$$\left. + 2(q_2 - q_4)_\mu(q_1 - q_3, q_2 + q_4)]\right\}$$

is obtained from Eq. (2.1), while the anomaly-induced term is

$$J_\mu^{an} = 2\left(\frac{N_c g^2}{8\pi^2}\right)^2(1 + P_{23})[q_{1\mu}(1 - P_{24})(q, q_4)(q_2, q_4)$$

$$+ q_{2\mu}(1 - P_{14})(q, q_1)(q_3, q_4) \qquad (2.13)$$

$$+ q_{4\mu}(1 - P_{12})(q, q_2)(q_1, q_3)]$$

$$\times\left[\frac{1}{D_\rho(q_1 + q_2)} + \frac{1}{D_\rho(q_1 + q_4)} + \frac{1}{D_\rho(q_2 + q_4)}\right]\frac{1}{D_\omega(q - q_3)}.$$

One can verify that up to corrections of the order m_π^2/m_ρ^2, the above amplitudes vanish in the limit of the vanishing 4-momentum for each final pion. In different words, they obey the Adler condition.

The amplitudes for the $\rho^- \longrightarrow (4\pi)^-$ decays are obtained from the above expressions for the $\rho^+ \longrightarrow (4\pi)^+$ decay amplitudes by inverting the overall sign.

2.3. Nonrelativistic Limit and an Order-of-Magnitude Estimate of the Partial Widths

As will be clear in what follows, the nonrelativistic expressions for the above amplitudes are needed. They are obtained by neglecting the space components of the pion 4-momenta. For the $\pi \longrightarrow 3\pi$ transition, they are given by

$$\begin{aligned}
M(\pi^+ &\longrightarrow \pi^+_{q_1}\pi^+_{q_2}\pi^-_{q_3}) = -2m_\pi^2/f_\pi^2,\\
M(\pi^+ &\longrightarrow \pi^+_{q_1}\pi^0_{q_2}\pi^0_{q_3}) = -m_\pi^2/f_\pi^2,\\
M(\pi^0 &\longrightarrow \pi^+_{q_1}\pi^-_{q_2}\pi^0_{q_3}) = -m_\pi^2/f_\pi^2,\\
M(\pi^0 &\longrightarrow \pi^0_{q_1}\pi^0_{q_2}\pi^0_{q_3}) = -3m_\pi^2/f_\pi^2.
\end{aligned} \qquad (2.14)$$

We note that the HLS parameter a is dropped from the expressions in the nonrelativistic limit. The nonrelativistic expressions for the $\rho \longrightarrow 4\pi$ decay amplitudes are relevant for the four-pion invariant mass below 700 MeV. They are obtained by neglecting the space components of the pion momenta in the Lorentz invariant scalar products (q_i, q_j). One can convince oneself that in this limit, a enters the expressions for the amplitudes as an overall factor, Eq. (2.6), and the amplitudes therefore become

$$\begin{aligned}
M(\rho^0 &\longrightarrow \pi^+_{q_1}\pi^+_{q_2}\pi^-_{q_3}\pi^-_{q_4}) \approx -\frac{g_{\rho\pi\pi}}{2f_\pi^2}(\varepsilon, q_1 + q_2 - q_3 - q_4),\\
M(\rho^0 &\longrightarrow \pi^+_{q_1}\pi^-_{q_2}\pi^0_{q_3}\pi^0_{q_4}) \approx -\frac{g_{\rho\pi\pi}}{4f_\pi^2}(\varepsilon, q_1 - q_2),\\
M(\rho^+ &\longrightarrow \pi^+_{q_1}\pi^+_{q_2}\pi^-_{q_3}\pi^0_{q_4}) \approx \frac{g_{\rho\pi\pi}}{4f_\pi^2}(\varepsilon, q_1 + q_2 - 2q_4),
\end{aligned} \qquad (2.15)$$

$$M(\rho^+ \longrightarrow \pi^+_{q_1}\pi^0_{q_2}\pi^0_{q_3}\pi^0_{q_4}) \approx \frac{g_{\rho\pi\pi}}{f_\pi^2}(\varepsilon, q_1).$$

These simplified expressions[4] are convenient, especially, in the calculation of the $\omega \longrightarrow 5\pi$ decay amplitude, because the typical invariant masses of the four-pion system in the above decay are in the vicinity of 620 MeV (see Section 4 for more detail). Moreover, as the direct numerical calculations of the subsequent Section 3.1 show, the nonrelativistic expressions (2.15) are valid with an accuracy up to 15% even at the invariant mass near m_ρ, provided the terms induced by the anomalous Wess–Zumino Lagrangian are neglected[5] in the $\rho^0 \longrightarrow \pi^+\pi^-2\pi^0$ and $\rho^\pm \longrightarrow 2\pi^\pm\pi^\mp\pi^0$ decay amplitudes.

We now give very simple estimates of the partial $\rho \longrightarrow 4\pi$ widths that are valid up to the factor of three. They are based on approximating the true branching ratio by the product of the squared modulus of the matrix element averaged over the polarizations of the decaying ρ and by the nonrelativistic expression for the Lorentz-invariant phase space volume of four pions. The latter is obtained by setting $n = 4$ and $M = m_\rho$ in the general expression

$$\Phi_n(M) = \frac{R_n(M)}{2M(2\pi)^{3n-4}N_s}, \tag{2.16}$$

where M is the decaying particle mass; n is the number of final particles with the masses m_i ($i = 1, \ldots, n$); N_s is the factor accounting for indistinguishable particles in the final state; and [12]

$$R_n(M) = \frac{(2\pi^3)^{(n-1)/2}}{2\Gamma\left[\frac{3}{2}(n-1)\right]}\frac{\left(\prod_{i=1}^n m_i\right)^{1/2}}{\left(\sum_{i=1}^n m_i\right)^{3/2}}\left(M - \sum_{i=1}^n m_i\right)^{(3n-5)/2}, \tag{2.17}$$

with Γ being the gamma function. We note that at $m_i = m_\pi = 139.6$ MeV ($i = 1, \ldots, 4$), the numerical value of $\Phi_4(m_\rho)$ agrees within 7% with the exact value obtained via numerical integration of Eq. (3.3) in the next section.

The results of the estimation are as follows. The answer depends on which combination of the pion momenta is retained in the expression for the amplitude. The relation

$$(\varepsilon, q_1 + q_2 + q_3 + q_4) = 0 \tag{2.18}$$

expressing the transverse character of the ρ polarization four-vector ε permits one to eliminate one combination of the pion momenta in favor of another one. This results in the above-mentioned uncertain factor up to 3, which illustrates the role of angular correlations among the momenta of the final pions.

(1) The $\rho^0 \longrightarrow 2\pi^+2\pi^-$ decay. Here, using the first line of Eq. (2.15), we find

$$\frac{1}{3}|M|^2 = \frac{1}{12}\left(\frac{g_{\rho\pi\pi}}{f_\pi^2}\right)^2(\mathbf{q}_1 + \mathbf{q}_2 - \mathbf{q}_3 - \mathbf{q}_4)^2 \approx \frac{1}{3}\left(\frac{g_{\rho\pi\pi}}{f_\pi^2}\right)^2\mathbf{q}_\pi^2, \tag{2.19}$$

where the approximate equality follows by setting all the pion three-momenta to the equilibrium one ($|\mathbf{q}_\pi| = 133$ MeV), obtained by taking the energy of each pion equal to $m_\rho/4$ and neglecting the angular correlations among different momenta. Another choice is to eliminate, for example, $\mathbf{q}_3$ and $\mathbf{q}_4$, which results in multiplying the right-hand side of Eq. (2.19) by the factor 2. Taking $N_s = 4$ and $M = m_\rho$, we then find

$$\Gamma_{\rho^0 \to 2\pi^+2\pi^-} \approx 0.8\text{–}1.6 \text{ keV}.$$

(2) The $\rho^0 \longrightarrow \pi^+\pi^-2\pi^0$ decay. Equation (2.18) allows us to eliminate q_2 in the second equation in (2.15) and replace $q_1 - q_2$ with $2q_1 + q_3 + q_4$. This replacement results in the factor of three multiplying the approximate expression for $|M|^2$ obtained from the second line of Eq. (2.15). With $N_s = 2$, the estimate is then

$$\Gamma_{\rho^0 \to \pi^+\pi^-2\pi^0} \approx 0.2\text{–}0.6 \text{ keV}.$$

(3) The $\rho^\pm \longrightarrow 2\pi^\pm\pi^\mp\pi^0$ decay. Here, $N_s = 2$, and the estimate is

$$\Gamma_{\rho^\pm \to 2\pi^\pm\pi^\mp\pi^0} \approx 0.6\text{–}1.0 \text{ keV}.$$

The uncertainty within the factor 5/3 comes from the above arbitrariness of the choice among the final pion momenta.

(4) The $\rho^\pm \longrightarrow \pi^\pm3\pi^0$ decay. Here, $N_s = 6$, and the estimate is

$$\Gamma_{\rho^\pm \to \pi^\pm3\pi^0} \approx 0.5\text{–}1.5 \text{ keV},$$

[4] Amplitudes (2.15) are obtained from the expressions given in Section 2.2 obeying the Adler condition. However, they do not contain the factor m_π^2/f_π^2 typical for the $\pi\pi$ scattering amplitudes that also obey this condition, see Eq. (2.14), except the natural kinematical smallness $\propto \varepsilon\sum_i c_i\mathbf{q}_i$, where c_i are numerical coefficients from Eq. (2.15). This is because the pointlike $\rho \longrightarrow 4\pi$ vertex in the exact $\rho \longrightarrow 4\pi$ amplitude Fig. 1b up to the terms $O(m_\pi^2/m_\rho^2)$ is compensated in the Adler limit of vanishing of one of the pion momenta by the π pole diagrams, while in the nonrelativistic limit the latter compensate only one-half of the former.

[5] The terms induced by the anomalous Wess–Zumino Lagrangian involved in the $\rho^0 \longrightarrow \pi^+\pi^-2\pi^0$ and $\rho^\pm \longrightarrow 2\pi^\pm\pi^\mp\pi^0$ decay amplitudes are appreciable at the ρ mass, see Section 3.1. However, they vanish in the nonrelativistic limit.

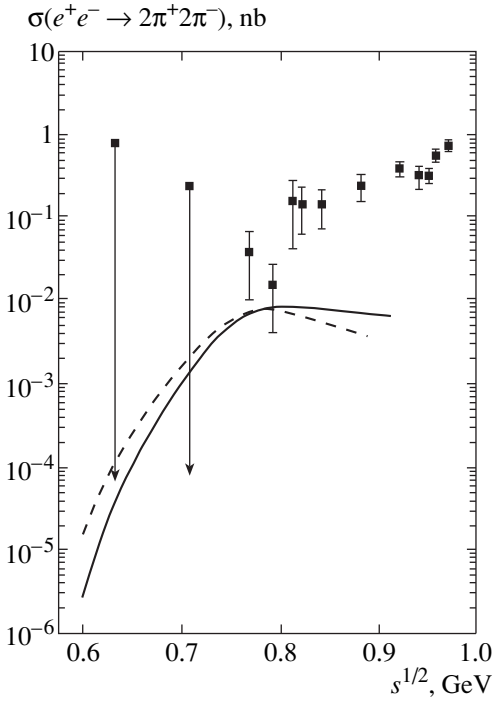

Fig. 2. The energy dependence of the $e^+e^- \longrightarrow \rho^0 \longrightarrow \pi^+\pi^-\pi^+\pi^-$ reaction cross section in the model based on the chiral Lagrangian due to Weinberg. Experimental points are from [16]. The solid (dashed) line refers to the dynamic (phase space) model of the decay.

with the above arbitrariness of the choice among the pion momenta resulting in the uncertain factor 3. We note that the order-of-magnitude estimates in cases 2 and 3 should be compared to the exact calculations in the models where the terms induced by the anomalous Lagrangian are neglected (see footnote 5). We now proceed to the exact numerical evaluation of the above branching ratios and show that they agree with the above very simple estimates sufficiently well.

3. RESULTS FOR VARIOUS $\rho \longrightarrow 4\pi$ DECAYS

In evaluating the partial widths of the 4π decays of ρ meson, the modulus squared of the matrix element is expressed via the Kumar variables [13]. The idea of speeding up the numerical integration suggested in [14] is realized in the numerical algorithm. The results of evaluating the partial widths at $\sqrt{s} = m_\rho = 770$ MeV are as follows:

$$\Gamma_{\rho^0 \to 2\pi^+2\pi^-}(m_\rho) = 0.89 \text{ keV},$$

$$\Gamma_{\rho^0 \to \pi^+\pi^-2\pi^0}(m_\rho) = 0.24 \text{ keV, and } 0.44 \text{ keV},$$

respectively, without and with the induced anomalous term taken into account. This coincides with the results obtained in [7]. For the charged ρ meson decays, the estimates are new:

$$\Gamma_{\rho^\pm \to \pi^\pm 3\pi^0}(m_\rho) = 0.41 \text{ keV},$$

$$\Gamma_{\rho^\pm \to 2\pi^\pm \pi^m p \pi^0}(m_\rho) = 0.71 \text{ keV, and } 0.90 \text{ keV},$$

respectively, without and with the anomaly-induced term taken into account. A comparison with the model estimates in Section 2.3 shows that they are rather reliable despite their extreme simplicity. In obtaining the exact figures, the narrow ρ width approximation is used. This is equivalent to using $\Gamma_{\rho\pi\pi} \longrightarrow 0$ in Eq. (2.5). Keeping the physical value of the ρ width gives results that deviate from those obtained in the narrow-width approximation by a quantity that does not exceed a few percent of the values obtained in that approximation. This is true for the invariant mass of the four-pion state lying below the $\rho\pi$ threshold energy, $m_{4\pi} < 910$ MeV. We recall that allowing for the finite width effects is in fact equivalent to taking the loop correction into account.

The above results are obtained for a = 2. The variation of a within 20% around this value leads to the variation of the branching ratios within 20% around the above values. This fact can easily be traced in the nonrelativistic limit, where the parameter a enters the expressions for the amplitudes as an overall factor $\sqrt{a}$, see Eqs. (2.2) and (2.15).

3.1. The $\rho^0 \longrightarrow 4\pi$ Decay Manifested in e^+e^- Annihilation

The results of the 4π state production cross section in the reaction $e^+e^- \longrightarrow \rho^0 \longrightarrow 4\pi$,

$$\sigma_{e^+e^- \to \rho \to 4\pi}(s) = \frac{12\pi m_\rho^3 \Gamma_{\rho e^+e^-}(m_\rho)\Gamma_{\rho \to 4\pi}(E)}{E^3 |D_\rho(s)|^2}, \quad (3.1)$$

where $s = E^2$ is the square of the total center-of-mass energy and $D_\rho(s)$ is obtained from Eq. (2.5) by substituting $q^2 \longrightarrow s$, are shown in Figs. 2 and 3. The values of the vector meson parameters taken from [15] are used here and in what follows. To demonstrate the effects of chiral dynamics, the energy dependence of the cross section evaluated in the pure phase space model for the four-pion decay is also shown. In this model, the 4π partial width normalized to the width at the ρ mass calculated in the dynamical model is given by

$$\Gamma^{LIPS}_{\rho \to 4\pi}(\sqrt{s}) = \Gamma_{\rho \to 4\pi}(m_\rho)\frac{W_{4\pi}(s)}{W_{4\pi}(m_\rho^2)}, \quad (3.2)$$

where the four-pion phase-space volume is [12, 13]

$$W_{4\pi}(s) = \frac{\pi^3}{16(2\pi)^8 s^{3/2} N_s} \int\limits_{(3m_\pi)^2}^{(\sqrt{s}-m_\pi)^2} \frac{ds_1}{s_1}\lambda^{1/2}(s, s_1, m_\pi^2) \times \int\limits_{(2m_\pi)^2}^{(\sqrt{s_1}-m_\pi)^2} \frac{ds_2}{s_2}\lambda^{1/2}(s_1, s_2, m_\pi^2)\lambda^{1/2}(s_2, m_\pi^2, m_\pi^2). \quad (3.3)$$

In this formula, $N_s = 4$ (2) is the factor that accounts for the identity of final pions in the final state $2\pi^+2\pi^-$ (respectively, $\pi^+\pi^-2\pi^0$), and

$$\lambda(x, y, z) = x^2 + y^2 + x^2 - 2(xy + xz + yz). \quad (3.4)$$

As the evaluation shows, the ratio

$$r(s) = \frac{\Gamma_{\rho \to 2\pi^+2\pi^-}(\sqrt{s})}{\Gamma^{LIPS}_{\rho \to 2\pi^+2\pi^-}(\sqrt{s})}$$

changes from 0.4 at $\sqrt{s} = 650$ MeV to 1 at $\sqrt{s} = m_\rho$. As can be seen from Figs. 2 and 3, the rise of the $\rho \longrightarrow 4\pi$ partial width with the energy increase is so fast that it completely compensates the falling of the ρ meson propagator and the electron width. Also noticeable is the dynamical effect in the decay $\rho^0 \longrightarrow \pi^+\pi^-2\pi^0$ at $\sqrt{s} > 850$ MeV resulting from the anomaly-induced Lagrangian ωπ threshold (see Fig. 3). To quantify the above-mentioned effect of the vanishing higher derivative contribution at the left shoulder of the ρ resonance, we note that the difference between the magnitudes of $\Gamma_{\rho \to \pi^+\pi^-2\pi^0}(\sqrt{s})$ with and without the term originating from the anomaly-induced Lagrangian, which is equal to 100% at $\sqrt{s} = m_\rho$, decreases rapidly with an energy decrease, amounting to 8% at $\sqrt{s} = 700$ MeV and 0.3% at $\sqrt{s} = 650$ MeV.

It should be pointed out that the evaluation of the partial widths with the nonrelativistic expressions for the $\rho \longrightarrow 4\pi$ amplitudes, Eq. (2.15), gives values that deviate from those obtained with the exact expressions within 7–15%, depending on the energy in the interval from 610 to 770 MeV. This corresponds to the case where the terms induced by the anomalous Lagrangian are neglected (see footnote 5).

As seen from Fig. 2, the predictions of chiral symmetry for the $e^+e^- \longrightarrow 2\pi^+2\pi^-$ reaction cross section do not contradict the experimental points of the CMD-2 detector [16] attributed to the energies $\sqrt{s} < 800$ MeV. At $\sqrt{s} > 800$ MeV, however, one can observe a substantial deviation of the predictions of Lagrangian (2.1) from the CMD-2 data. By all appearances, this is the result of neglecting the contribution of higher derivatives and chiral loops. It is expected that the left shoulder of the ρ peak is practically free of such contributions and is therefore preferable for studying the dynamical effects of chiral symmetry. Even at $\sqrt{s} = 650$ MeV, where the contribution of higher derivatives is negligible, one can hope to gather one event of the reaction $e^+e^- \longrightarrow 2\pi^+2\pi^-$ per day and up to 10 events of this reaction per day at $\sqrt{s} = 700$ MeV, provided the luminosity $L = 10^{32}$ cm^{-2} s^{-1} is achieved (i.e., to have a factory for a comprehensive study of the chiral dynamics of multipion systems).

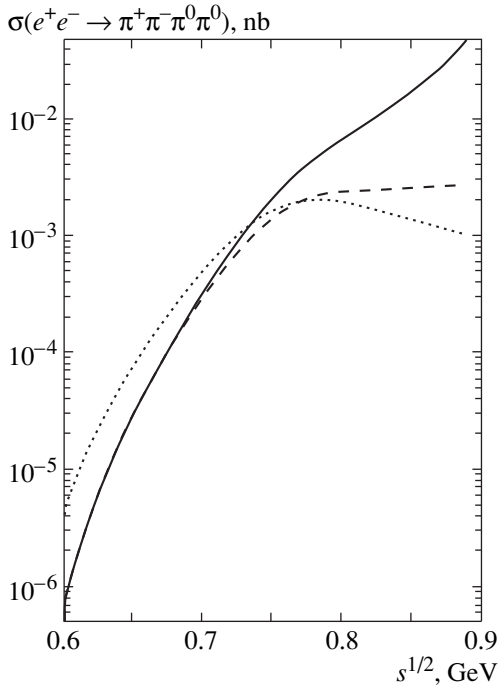

Fig. 3. The energy dependence of the $e^+e^- \longrightarrow \rho^0 \longrightarrow \pi^+\pi^-\pi^0\pi^0$ reaction cross-section in the model based on the chiral Lagrangian due to Weinberg, added with the terms induced by the anomalous Wess–Zumino Lagrangian. Solid, dashed, and dotted lines refer to the respective models with the anomaly induced term, without this term, and phase space.

Because of the helicity conservation, the ρ meson is produced in the states with the spin projections $\lambda = \pm 1$ on the e^+e^- beam axis characterized by a unit vector $\mathbf{n}_0$ that is assumed to be directed along the z-axis. Using the expressions for the total $\rho \longrightarrow 4\pi$ amplitudes, one can then obtain the angular distributions for the final pions. They are expected to be cumbersome. However, a good approximation for these distributions can be obtained from the approximate nonrelativistic expression in Eq. (2.15), at least at energies less than 700 MeV. Taking θ_i and ϕ_i to be the polar and azimuthal angles of the pion three-momentum $\mathbf{q}_i$, where the momentum assignment corresponds to Eq. (2.15); assuming the ρ meson to be at rest, one arrives at the following results.

(1) The $\rho^0 \longrightarrow 2\pi^+2\pi^-$ decay. The probability density of the emission of four charged pions can be found directly from the first line in Eq. (2.15):

$$\begin{aligned} w &\propto (\mathbf{q}_1+\mathbf{q}_2-\mathbf{q}_3-\mathbf{q}_4)^2-(\mathbf{n}_0\cdot(\mathbf{q}_1+\mathbf{q}_2-\mathbf{q}_3-\mathbf{q}_4))^2 \\ &= \sum_{i=1}^{4}\mathbf{q}_i^2\sin^2\theta_i + 2|\mathbf{q}_1|(1-P_{23}-P_{24}) \\ &\times|\mathbf{q}_2|\sin\theta_1\sin\theta_2\cos(\phi_1-\phi_2) \\ &-2|\mathbf{q}_2|(1+P_{34})|\mathbf{q}_3|\sin\theta_2\sin\theta_3\cos(\phi_2-\phi_3) \\ &+2|\mathbf{q}_3||\mathbf{q}_4|\sin\theta_3\sin\theta_4\cos(\phi_3-\phi_4). \end{aligned} \tag{3.5}$$

As before, P_{ij} interchanges the momenta q_i and q_j. One can use Eq. (2.18) to eliminate the momenta of negatively charged pions q_3 and q_4. Then, the probability density of the emission of two π^+'s found from the first line in Eq. (2.15) is

$$\begin{aligned} w &\propto (\mathbf{q}_1+\mathbf{q}_2)^2-(\mathbf{n}_0\cdot(\mathbf{q}_1+\mathbf{q}_2))^2 = \mathbf{q}_1^2\sin^2\theta_1 \\ &+\mathbf{q}_2^2\sin^2\theta_2+2|\mathbf{q}_1||\mathbf{q}_2|\sin\theta_1\sin\theta_2\cos(\phi_1-\phi_2). \end{aligned} \tag{3.6}$$

With Eq. (2.18) taken into account, the angular distribution for the emission of two π^-'s is obtained from Eq. (3.6) by the respective replacements $\mathbf{q}_{1,2} \longrightarrow \mathbf{q}_{3,4}$.

(2) The $\rho^0 \longrightarrow \pi^+\pi^-2\pi^0$ decay. The probability density of the emission of $\pi^+\pi^-$ pair is found from the second line in Eq. (2.15) as

$$\begin{aligned} w &\propto (\mathbf{q}_1-\mathbf{q}_2)^2-(\mathbf{n}_0\cdot(\mathbf{q}_1-\mathbf{q}_2))^2 = \mathbf{q}_1^2\sin^2\theta_1 \\ &+\mathbf{q}_2^2\sin^2\theta_2-2|\mathbf{q}_1||\mathbf{q}_2|\sin\theta_1\sin\theta_2\cos(\phi_1-\phi_2). \end{aligned} \tag{3.7}$$

Eliminating the momentum q_2, one finds the corresponding expression for the final state combination $\pi^+2\pi^0$:

$$\begin{aligned} w &\propto (2\mathbf{q}_1-\mathbf{q}_3-\mathbf{q}_4)^2-(\mathbf{n}_0\cdot(2\mathbf{q}_1-\mathbf{q}_3-\mathbf{q}_4))^2 \\ &= 4\mathbf{q}_1^2\sin^2\theta_1+\mathbf{q}_3^2\sin^2\theta_3+\mathbf{q}_4^2\sin^2\theta_4 \\ &-(1+P_{34})4|\mathbf{q}_1||\mathbf{q}_3|\sin\theta_1\sin\theta_3\cos(\phi_1-\phi_3) \\ &+2|\mathbf{q}_3||\mathbf{q}_4|\sin\theta_3\sin\theta_4\cos(\phi_3-\phi_4). \end{aligned} \tag{3.8}$$

In view of Eq. (2.18), the angular distribution for the state $\pi^-2\pi^0$ is obtained from the above by replacing $\mathbf{q}_1 \longrightarrow \mathbf{q}_2$ and changing the signs in front of the terms containing $(1+P_{34})$.

3.2. The $\rho \longrightarrow 4\pi$ Decay in τ Decays

Based on the vector current conservation, the partial width of the decay $\tau^- \longrightarrow \nu_\tau(4\pi)^-$ [17, 18] can be written as the integral over the invariant mass of the four-pion state m extended up to some mass m_0, whose maximal value is $m_{0\max} = m_\tau$,

$$\begin{aligned} &B_{\tau^-\to\nu_\tau(4\pi)^-}(m_0) \\ &= T_\tau\int_{4m_\pi}^{m_0} dm\frac{2m^2\Gamma_{\tau^-\to\nu_\tau\rho^-}(m)}{\pi|D_\rho(m^2)|^2}\Gamma_{\rho^-\to(4\pi)^-}(m), \end{aligned} \tag{3.9}$$

where T_τ and

$$\begin{aligned} \Gamma_{\tau^-\to\nu_\tau\rho^-}(m) &= \frac{G_F^2\cos\theta_C}{8\pi f_\rho^2} \\ &\times m_\tau^3 m_\rho^2\left(1-\frac{m^2}{m_\tau^2}\right)^2\left(1+2\frac{m^2}{m_\tau^2}\right) \end{aligned} \tag{3.10}$$

are, respectively, the lifetime of the τ lepton and the partial width of its decay $\tau^- \longrightarrow \nu_\tau\rho^-$ [17], with m being the invariant mass of the four-pion state. Using the numerical values of the $\rho \longrightarrow 4\pi$ decay widths, one can evaluate the branching ratios of the four-pion τ decays for various values of the upper invariant mass m_0. The results of the evaluation of the branching ratios of the decays $\tau^- \longrightarrow \nu_\tau 2\pi^-\pi^+\pi^0$ and $\tau^- \longrightarrow \nu_\tau\pi^-3\pi^0$ for the invariant mass values of the four-pion system from 600 to 850 MeV are plotted in Fig. 4. In particular, taking $m_0 = 740$ MeV, one obtains

$$B_{\tau^-\to\nu_\tau 2\pi^-\pi^+\pi^0}(m_0) = 7.6\times10^{-8}(8.4\times10^{-8})$$

without (with) the anomaly induced term, respectively, and

$$B_{\tau^-\to\nu_\tau 3\pi^0\pi^-}(m_0) = 4.6\times10^{-8}.$$

The corresponding values for the upper integration mass $m_0 = 640$ MeV are

$$B_{\tau^-\to\nu_\tau 2\pi^-\pi^+\pi^0}(m_0) = 2.895\times10^{-10}(2.900\times10^{-10})$$

without (with) the anomaly induced term, respectively, and

$$B_{\tau^-\to\nu_\tau 3\pi^0\pi^-}(m_0) = 1.8\times10^{-10}.$$

The comparison of both curves in Fig. 4a again demonstrates that the contributions of higher derivatives represented by the terms induced by the anomalous Wess–Zumino Lagrangian vanish rapidly as the mass decreases. Unfortunately, the domains in the low four-pion invariant mass, where the chiral dynamics effects are clean, are hardly accessible with τ factories.

Indeed, guided by the expression for the cross section of the τ lepton pair production in the e^+e^- annihilation,

$$\sigma_{\tau^+\tau^-}(s) = \frac{4\pi\alpha^2}{3s}\sqrt{1-\frac{4m_\tau^2}{s}}\left(1+2\frac{m_\tau^2}{s}\right), \tag{3.11}$$

one can find that up to $N = 25 \times 10^7$ τ lepton pairs with the total energy $\sqrt{s} = m_{\psi(2S)}$ can be produced per season at a τ-charm factory with the expected luminosity $L = 10^{34}$ cm^{-2} s^{-1} [15]. This implies that one can detect only 2 to 4 events per season in the four-pion mass range below 700 MeV. However, the event counting rate rises rapidly with the increase of the upper integration mass m_0 in Eq. (3.9), reaching the figure about 60 to 120 events per season at $m_0 = m_\rho$, depending on the charge combination of the final pions.

3.3. The $\rho \longrightarrow 4\pi$ Decay in Photoproduction, $\pi N \longrightarrow \rho\pi N$, and So On

To conclude on the possibility of studying the $\rho \longrightarrow 4\pi$ decays in photoproduction, we calculate the quantity

$$B^{\text{aver}}_{\rho \to 4\pi}(m_0) = \frac{2}{\pi}\int\limits_{4m_\pi}^{m_0} dm \frac{m^2 \Gamma_{\rho \to 4\pi}(m)}{\left|D_\rho(m^2)\right|^2}, \qquad (3.12)$$

which is the average of the branching ratio against the invariant mass of the four-pion state. In the limit $m_0 \longrightarrow \infty$, Eq. (3.12) serves as the definition of the branching ratio for a wide resonance. Equation (3.12) should be confronted with the familiar definition of the branching ratio at the ρ mass,

$$B_{\rho \to 4\pi}(m_\rho) = \Gamma_{\rho \to 4\pi}(m_\rho)/\Gamma_\rho, \qquad (3.13)$$

which follows from Eq. (3.12) upon the replacement

$$m\Gamma_\rho/\pi\left|D_\rho(m^2)\right|^2 \longrightarrow \delta(m^2 - m_\rho^2),$$

which is valid in the narrow-width limit. With the partial widths evaluated here, one finds

$$B_{\rho^0 \to 2\pi^+ 2\pi^-}(m_\rho) = 5.9 \times 10^{-6}$$

and

$$B_{\rho^0 \to \pi^+\pi^- 2\pi^0}(m_\rho) = 1.6 \times 10^{-6}(2.9 \times 10^{-6})$$

without (with) the anomaly-induced term, respectively. The plot of $B^{\text{aver}}_{\rho \to 4\pi}(m_0)$ is shown in Fig. 5. In particular, the evaluation gives

$$B^{\text{aver}}_{\rho^0 \to 2\pi^+ 2\pi^-}(m_0) = 4.4 \times 10^{-6}, 6.1 \times 10^{-8},$$

$$\text{and } 1.4 \times 10^{-9}$$

at the respective values $m_0 = 850$, 700, and 640 MeV. For other four-pion decay modes of ρ^0, the results are as follows. In the model with the vanishing term induced by the anomalous Wess–Zumino Lagrangian, one obtains

$$B^{\text{aver}}_{\rho^0 \to \pi^+\pi^- 2\pi^0}(m_0) = 1.3 \times 10^{-6}, 1.58 \times 10^{-8},$$

$$\text{and } 3.66 \times 10^{-10}$$

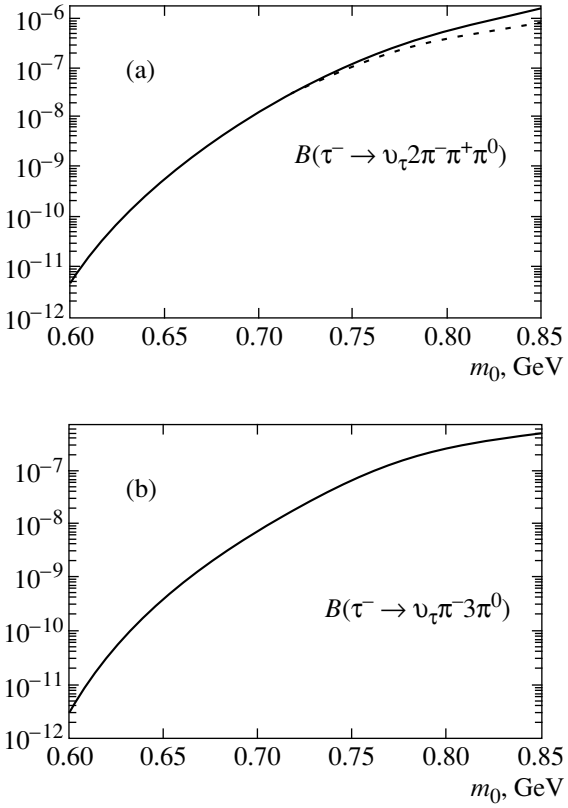

Fig. 4. The branching ratio of the decays (a) $\tau^- \longrightarrow \nu_\tau 2\pi^-\pi^+\pi^0$ and (b) $\tau^- \longrightarrow \nu_\tau \pi^- 3\pi^0$ vs. the invariant mass of the four-pion system, see Eq. (3.9). The solid (dotted) line in (a) refers to the model with (without) the term induced by the anomalous Lagrangian.

at $m_0 = 850$, 700, and 640 MeV. In the model that includes the above term, one obtains

$$B^{\text{aver}}_{\rho^0 \to \pi^+\pi^- 2\pi^0}(m_0) = 4.9 \times 10^{-6}, 1.65 \times 10^{-8},$$

$$\text{and } 3.63 \times 10^{-10}$$

at the same respective values of m_0. As is expected, the branching ratios in the two models mentioned above converge to because of the rapid vanishing of the contributions of higher derivative terms. The difference between the two definitions of the branching ratio is seen by comparing $B^{\text{aver}}_{\rho \to 4\pi}$ ($m_0 = 850$ MeV) evaluated for various charge combinations of the final pions with $B_{\rho \to 4\pi}(m_\rho)$ evaluated above.

With the total number of ρ mesons $N \approx 6 \times 10^9$ expected to be produced on a nucleon at the Jefferson Laboratory [19], one may hope to observe about 100 and 360 events of the ρ decays into the respective states $\pi^+\pi^-2\pi^0$ and $2\pi^+2\pi^-$ in the mass range $m_0 < 700$ MeV, where the chiral dynamics effects are cleanest. The photoproduction on heavy nuclei results in increasing the number of the produced ρ mesons faster than $A^{2/3}$, where A is the mass number. The generally adopted

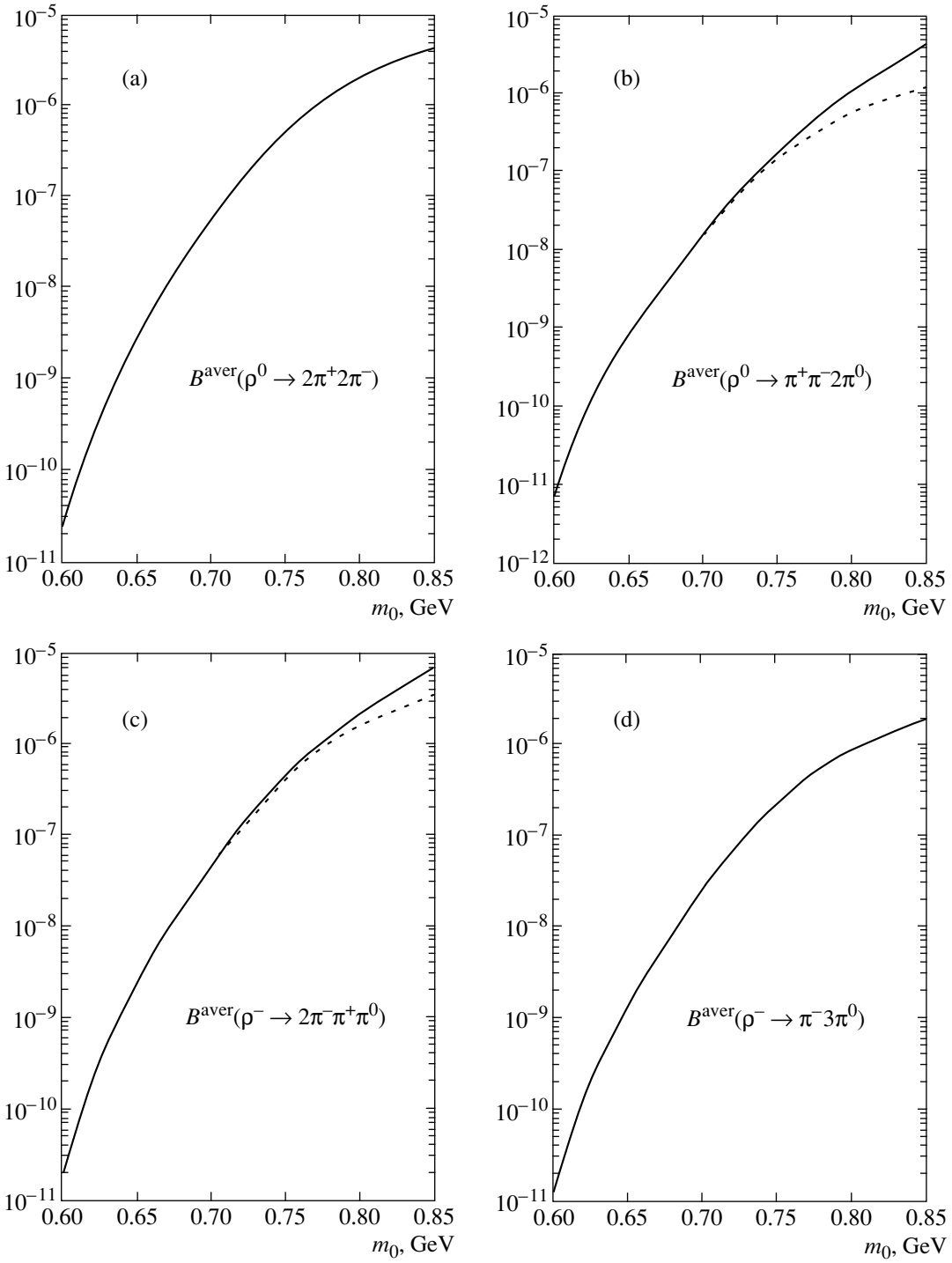

Fig. 5. The branching ratio of the decays (a) $\rho^0 \longrightarrow 2\pi^+2\pi^-$, (b) $\rho^0 \longrightarrow \pi^+\pi^-2\pi^0$, (c) $\rho^- \longrightarrow 2\pi^-\pi^+\pi^0$, (d) $\rho^0 \longrightarrow \pi^+\pi^-2\pi^0$ vs. the invariant mass of the four-pion system, see Eq. (3.12). The solid (dotted) line in (b) and (c) refers to the model with (without) the term induced by the anomalous Lagrangian.

behavior is in accordance with the function $A^{0.8–0.95}$ [20]. Thus, the photoproduction of four-pion states on heavy nuclei would give the possibility for the high statistics study of chiral dynamics effects in the four-pion decays of $\rho(770)$. It should be recalled once more that the counting rate rises rapidly with the increase of m_0.

The conclusions about the angular distributions of the final pions with zero net charge in photoproduction

are as follows. The general expression should certainly be deduced from the full decay amplitudes given in Section 2 together with the detailed form of the photoproduction mechanism. The qualitative picture, however, can be obtained by noting that the *s*-channel helicity conservation is a good selection rule for photoproduction reactions. In the helicity reference frame (which is the frame where ρ is at rest and its spin quantization axis is directed along the ρ momentum in the center-of-mass system), the expressions for the angular distributions coincide with the corresponding expressions for the production of these states in e^+e^- annihilation. The latter can be found in Section 3.1. For the photoproduction on heavy nuclei at high energies, the direction of the final ρ momentum lies at the scattering angle less than 0.5° and the vector $\mathbf{n}_0$ can therefore be treated as pointed along the photon-beam direction.

We note that other peripheral reactions can provide sufficiently intense sources of ρ mesons. For example, the diffractive production of the ρπ state in π*N* collisions are currently under study with the VES detector in Protvino. The regions of the four-pion invariant mass spectrum larger than m_ρ, namely, $m_0 \approx 850$ MeV with $B^{\text{aver}}_{\rho\to 4\pi}(m_0) \sim 10^{-5}$, should be included to reliably measure the $\rho \longrightarrow 4\pi$ branching ratio. As explained in the Introduction, this would require including the contributions of the a_1 meson and higher derivatives to the total amplitude. Nevertheless, the results of the present paper shown in Figs. 5c and 5d, obtained while neglecting these contributions, can be regarded as a guess in the experimental work in this direction.

4. THE $\omega \longrightarrow 5\pi$ DECAY

4.1. The $\omega \longrightarrow 5\pi$ *Decay Amplitudes*

One can be convinced that the $\omega \longrightarrow \rho\pi \longrightarrow 5\pi$ decay amplitude unambiguously follows from the anomaly-induced Lagrangian (2.10). This amplitude is represented by the diagrams shown in Fig. 6. As could be expected, its general expression is cumbersome. However, it can be considerably simplified by noting that although $|\mathbf{q}_\pi|/m_\pi \simeq 0.5$, the nonrelativistic expressions in Eq. (2.15) for the $\rho \longrightarrow 4\pi$ decay amplitudes in the diagrams in Fig. 6a are valid within 5% in the relevant 4π mass range [1]. Similarly, the expression for the combination $D_\pi^{-1} M(\pi \longrightarrow 3\pi)$ involved in the expression for the diagrams in Fig. 6b can be replaced, with the same accuracy, by $(8m_\pi^2)^{-1}$ times the nonrelativistic $\pi \longrightarrow 3\pi$ amplitudes in Eq. (2.14). Thus, using Eq. (2.15) to obtain the expression for the sum of the diagrams shown in Fig. 6a and Eq. (2.14) to obtain the expression for the sum of the diagrams shown in Fig. 6b,[6] one obtains, upon neglecting the corrections

[6] The contribution of the diagrams Fig. 6b was neglected in [1].

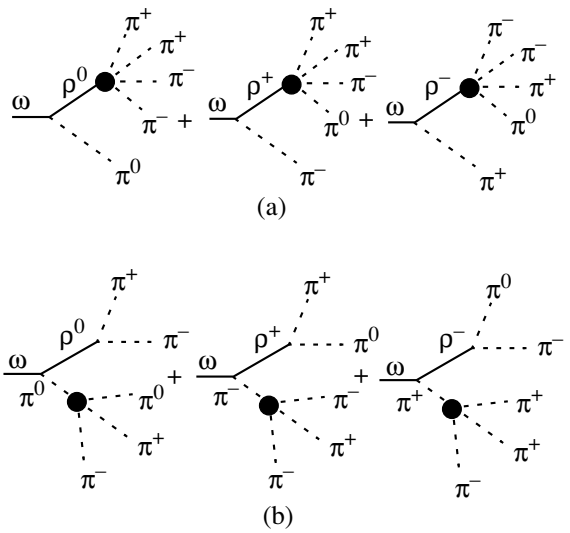

Fig. 6. The diagrams describing the amplitudes of the decays $\omega \longrightarrow \pi^+\pi^-\pi^+\pi^-\pi^0$. The shaded circles in set (a) denote the entire set of the $\rho \longrightarrow 4\pi$ diagrams shown in Fig. 1b, c. The shaded circles in the $\pi \longrightarrow 3\pi$ vertices in set (b) refer to the sum of diagrams shown in Fig. 1a. The symmetrization over the momenta of identical pions emitted from different vertices is understood. The diagrams for the decay $\omega \longrightarrow \pi^+\pi^-\pi^0\pi^0\pi^0$ are obtained from those shown here via evident replacements.

of the order $O(|\mathbf{q}_\pi|^4/m_\pi^4)$ or higher, the decay amplitudes

$$M(\omega \longrightarrow 2\pi^+2\pi^-\pi^0) = \frac{N_c g_{\rho\pi\pi} g^2}{8(2\pi)^2 f_\pi^3}\varepsilon_{\mu\nu\lambda\sigma} q_\mu \epsilon_\nu$$

$$\times\left\{(1+P_{12})q_{1\lambda}\left[\frac{(q_2+3q_4)_\sigma}{D_\rho(q-q_1)} - \frac{2q_{4\sigma}}{D_\rho(q_1+q_4)}\right]\right.$$

$$-(1+P_{35})q_{3\lambda}\left[\frac{(q_5+3q_4)_\sigma}{D_\rho(q-q_3)} - \frac{2q_{4\sigma}}{D_\rho(q_3+q_4)}\right] \tag{4.1}$$

$$\left.-(1+P_{12})(1+P_{35})q_{3\lambda}\left[\frac{2q_{4\sigma}}{D_\rho(q-q_4)} + \frac{q_{1\sigma}}{D_\rho(q_1+q_3)}\right]\right\},$$

with the final momentum assignment according to $\pi^+(q_1)\pi^+(q_2)\pi^-(q_3)\pi^-(q_5)\pi^0(q_4)$ and

$$M(\omega \longrightarrow \pi^+\pi^-3\pi^0) = \frac{N_c g_{\rho\pi\pi} g^2}{8(2\pi)^2 f_\pi^3}(1-P_{12})$$

$$\times(1+P_{34}+P_{35})\varepsilon_{\mu\nu\lambda\sigma} q_\mu \epsilon_\nu q_{1\lambda}$$

$$\times\left\{q_{3\sigma}\left[\frac{1}{D_\rho(q-q_3)} - \frac{1}{D_\rho(q_1+q_3)}\right]\right. \tag{4.2}$$

$$-q_{2\sigma}\left[\frac{4}{3D_\rho(q-q_1)}-\frac{1}{2D_\rho(q_1+q_2)}\right]\Bigg\},$$

with the final momentum assignment according to $\pi^+(q_1)\pi^-(q_2)\pi^0(q_3)\pi^0(q_4)\pi^0(q_5)$. In both these formulas, ϵ_ν and q_μ stand for the ω meson polarization and momentum 4-vectors. The first term in each square bracket refers to the specific diagram shown in Fig. 6a; the second term, to the diagram in Fig. 6b.

Even in this simplified form, however, the expressions for the $\omega \longrightarrow 5\pi$ amplitudes are not easy to use for the evaluation of branching ratios. To go further, one should note the following. The contribution of the diagrams shown in Fig. 6a depends on the invariant mass of the 4π system, which changes in a very narrow range (558 MeV $< m_{4\pi} <$ 642 MeV). In all the ρ propagators involved in the first terms in every set of square brackets in Eqs. (4.1) and (4.2), this invariant mass can be set to the equilibrium value $\overline{m_{4\pi}^2}^{1/2}$ = 620 MeV within the width accuracy of 20%. This equilibrium value is evaluated for the pion energy $E_\pi = m_\omega/5$ that gives the dominant contribution. The ρ propagators involved in the last terms in the square brackets of the above expressions depend on the invariant mass of the pion pairs, for which a similar statement is true. This invariant mass varies in a narrow range (280 MeV $< m_{2\pi} <$ 360 MeV). With the same accuracy, one can set it to $\overline{m_{2\pi}^2}^{1/2}$ = 295 MeV in all relevant propagators. On the other hand, the amplitude of the $\omega \longrightarrow \rho^0\pi^0 \longrightarrow (2\pi^+2\pi^-)\pi^0$ is

$$M[\omega \longrightarrow \rho^0\pi^0 \longrightarrow (2\pi^+2\pi^-)\pi^0]$$
$$= 4\frac{N_c g_{\rho\pi\pi} g^2}{8(2\pi)^2 f_\pi^3}\varepsilon_{\mu\nu\lambda\sigma}q_\mu\epsilon_\nu(q_1+q_2)_\lambda\frac{q_{4\sigma}}{D_\rho(q-q_4)}, \quad (4.3)$$

where the momentum assignment is the same as in Eq. (4.1). The other relevant amplitude corresponding to the first diagram in Fig. 6b is

$$M[\omega \longrightarrow \rho^0\pi^0 \longrightarrow (\pi^+\pi^-)(\pi^+\pi^-\pi^0)]$$
$$= \frac{N_c g_{\rho\pi\pi} g^2}{8(2\pi)^2 f_\pi^3}\varepsilon_{\mu\nu\lambda\sigma}q_\mu\epsilon_\nu(1+P_{12})(1+P_{35})\frac{q_{1\lambda}q_{3\sigma}}{D_\rho(q_1+q_3)}. \quad (4.4)$$

Then, taking the above consideration concerning the invariant masses into account and comparing Eqs. (4.1), (4.3), and (4.4), one can see that

$$M(\omega \longrightarrow 2\pi^+2\pi^-\pi^0) \quad (4.5)$$
$$\approx \frac{5}{2}M[\omega \longrightarrow \rho^0\pi^0 \longrightarrow (2\pi^+2\pi^-)\pi^0]\left[1-\frac{D_\rho(\overline{m_{4\pi}^2})}{2D_\rho(\overline{m_{2\pi}^2})}\right].$$

A similar analysis shows that

$$M(\omega \longrightarrow \pi^+\pi^-3\pi^0) \quad (4.6)$$
$$\approx \frac{5}{2}M[\omega \longrightarrow \rho^+\pi^- \longrightarrow (\pi^+3\pi^0)\pi^-]\left[1-\frac{D_\rho(\overline{m_{4\pi}^2})}{2D_\rho(\overline{m_{2\pi}^2})}\right],$$

where

$$M[\omega \longrightarrow \rho^+\pi^- \longrightarrow (\pi^+3\pi^0)\pi^-]$$
$$= -4\frac{N_c g_{\rho\pi\pi} g^2}{8(2\pi)^2 f_\pi^3}\frac{\varepsilon_{\mu\nu\lambda\sigma}q_\mu\epsilon_\nu q_{1\lambda}q_{2\sigma}}{D_\rho(q-q_2)}, \quad (4.7)$$

and the final momentum assignment is the same as in Eq. (4.2). The numerical values of $\overline{m_{4\pi}^2}^{1/2}$ and $\overline{m_{2\pi}^2}^{1/2}$ found above are such that the correction factor in parentheses of Eqs. (4.5) and (4.6) amounts to 20% in magnitude. In what follows, the above correction is taken into account as the overall factor 0.64 in front of the branching ratios of the decays $\omega \longrightarrow 5\pi$. In making this estimate, the imaginary part of the ρ propagators in the square brackets of Eq. (4.5) and (4.6) is neglected. This assumption is valid with an accuracy better than 1% in width.

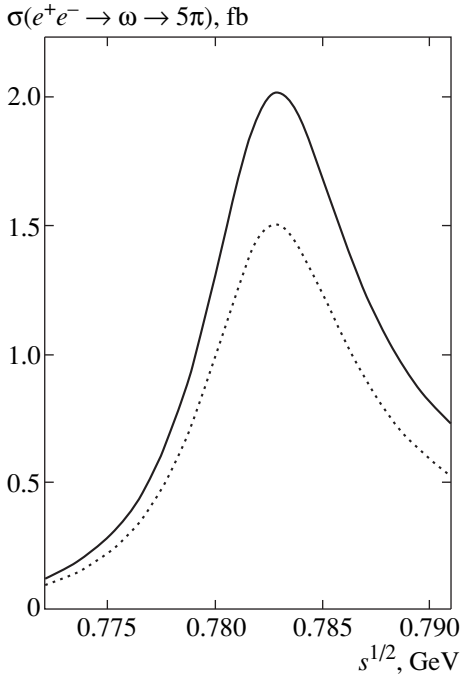

Fig. 7. The $\omega \longrightarrow 5\pi$ excitation curves in e^+e^- annihilation in the vicinity of the ω resonance. The solid (dotted) line refers to the $2\pi^+2\pi^-\pi^0$ ($\pi^+\pi^-3\pi^0$) final state.

4.2. An Approximate Model Estimate of the $\omega \longrightarrow 5\pi$ Branching Ratio

It is useful to derive the model estimate of the $\omega \longrightarrow 5\pi$ partial widths as follows. The corresponding equilibrium pion momenta are $|\mathbf{q}_{\pi^+}|$ = 70 MeV and $|\mathbf{q}_{\pi^0}|$ =

79 MeV The integrations over the angles of final pions can be evaluated under the assumption that they are independent. Using the nonrelativistic expression for the phase space volume of five pions obtained from Eq. (2.17) with $n = 5$ and $M = m_\omega$, one finds

$$R_5 = \frac{\pi^6\left(\prod_{i=1}^{5} m_{\pi i}\right)^{1/2}}{60\left(\sum_{i=1}^{5} m_{\pi i}\right)^{3/2}}\left(m_\omega - \sum_{i=1}^{5} m_{\pi i}\right)^5. \qquad (4.8)$$

Its numerical value coincides with the numerically evaluated exact expression to within 1%. Introducing the branching ratio at the ω mass as

$$B_{\omega\to 5\pi} = \Gamma_{\omega\to 5\pi}(m_\omega)/\Gamma_\omega, \qquad (4.9)$$

one finds

$$B_{\omega\to 5\pi} \approx \left[\frac{5N_c}{2\pi^2}\left(\frac{g_{\rho\pi\pi}g^2}{8f_\pi^3}\right)\frac{m_\omega|\mathbf{q}_{\pi^+}|}{\left|D_\rho(\overline{m^2_{4\pi}})\right|}\right]^2$$

$$\times \frac{R_5}{18(2\pi)^{11}m_\omega\Gamma_\omega}\left|1-\frac{D_\rho(\overline{m^2_{4\pi}})}{2D_\rho(\overline{m^2_{2\pi}})}\right|^2 \qquad (4.10)$$

$$\times\begin{cases}|\mathbf{q}_{\pi^0}|^2, & \text{for } 2\pi^+2\pi^-\pi^0,\\ \dfrac{|\mathbf{q}_{\pi^+}|^2}{3}, & \text{for } \pi^+\pi^-3\pi^0.\end{cases}$$

The above formula is obtained with the Lorentz structures chosen as in Eqs. (4.3) and (4.7). Using the antisymmetry property of $\varepsilon_{\mu\nu\lambda\sigma}$ and the 4-momentum conservation relation

$$q = \sum_{i=1}^{5} q_i,$$

one can eliminate one combination of the pion momenta in favor of another (see Section 4.4 for more detail). This arbitrariness results in an uncertain factor ranging from 2 for the $\omega\to 2\pi^+\pi^-\pi^0$ decay to 3 for the $\omega\to\pi^+\pi^-3\pi^0$ decay. As in the case of the $\rho\to 4\pi$ decay, this factor illustrates the role of the angular correlations among the final pions. The calculation gives

$$B(\omega\to 2\pi^+2\pi^-\pi^0) = (2.5\text{–}3.9)\times 10^{-9},$$

$$B(\omega\to\pi^+\pi^-3\pi^0) = (1.0\text{–}3.8)\times 10^{-9},$$

where, in addition to the above uncertain factor due to the choice of momenta, the correction for the mass difference of $\pi^\pm$ and π^0 is taken into account.

4.3. A More Careful Evaluation of the Branching Ratios and Excitation Curves

The partial widths can be evaluated with an accuracy of a 20% by taking the amplitude of each relevant decay to be 5/2 times the ρπ state production amplitude with the subsequent decay $\rho\to 4\pi$ and using the calculated widths of the latter:

$$B_{\omega\to 2\pi^+2\pi^-\pi^0} = \left|1-\frac{D_\rho(\overline{m^2_{4\pi}})}{2D_\rho(\overline{m^2_{2\pi}})}\right|^2\left(\frac{5}{2}\right)^2\frac{2}{\pi\Gamma_\omega} \qquad (4.11)$$

$$\times\int_{4m_{\pi^+}}^{m_\omega-m_{\pi^0}} dm\frac{m^2\Gamma_{\omega\to\rho^0\pi^0}(m)\Gamma_{\rho\to 2\pi^+2\pi^-}(m)}{\left|D_\rho(m^2)\right|^2} = 1.1\times 10^{-9},$$

where

$$\Gamma_{\omega\to\rho^0\pi^0}(m) = g^2_{\omega\rho\pi}q^3(m_\omega, m, m_{\pi^0})/12\pi,$$

$$g_{\omega\rho\pi} = \frac{N_c g^2}{8\pi^2 f_\pi} = 14.3\ \text{GeV}^{-1}.$$

The notation is such that

$$q(m_i, m_j, m_k) = \lambda^{1/2}(m_i^2, m_j^2, m_k^2)/2m_i, \qquad (4.12)$$

with the function λ given by Eq. (3.4) standing for the momentum of final particle j (or k) in the rest frame system of decaying particle i. We note the a^{-1} dependence of the $\omega\to 5\pi$ width on the HLS parameter a through Eq. (2.2). The branching ratio $B_{\omega\to\pi^+\pi^-3\pi^0}$ is obtained from Eq. (4.11) by inserting the lower integration limit $m_{\pi^+} + 3m_{\pi^0}$, replacing $m_{\pi^0}\to m_{\pi^+}$ in the expression for the momentum q, and then inserting the $\rho^+\to\pi^+3\pi^0$ decay width corrected for the mass difference of charged and neutral pions. Obviously, the main correction of this type comes from the phase-space volume of the final 4π state. One obtains

$$B_{\omega\to\pi^+\pi^-3\pi^0} = \left|1-\frac{D_\rho(\overline{m^2_{4\pi}})}{2D_\rho(\overline{m^2_{2\pi}})}\right|^2\left(\frac{5}{2}\right)^2\frac{2}{\pi\Gamma_\omega} \qquad (4.13)$$

$$\times\int_{m_{\pi^+}+3m_{\pi^0}}^{m_\omega-m_{\pi^+}} dm\frac{m^2\Gamma_{\omega\to\rho^+\pi^-}(m)\Gamma_{\rho^+\to\pi^+3\pi^0}(m)}{\left|D_\rho(m^2)\right|^2} = 8.5\times 10^{-10},$$

where

$$\Gamma_{\omega\to\rho^+\pi^-}(m) = g^2_{\omega\rho\pi}q^3(m_\omega, m, m_{\pi^+})/12\pi.$$

As pointed out in [3], the inclusion of the direct $\omega\to\pi^+\pi^-\pi^0$ vertex reduces the 3π decay width of ω by 33%. This implies that the effect of the pointlike $\omega\to 3\pi\to 2\pi(3\pi)$ diagrams shown in Fig. 6b must be

taken into account by the following replacement in the expression for the suppression factor:

$$\left|1-\frac{D_\rho(\overline{m^2_{4\pi}})}{2D_\rho(\overline{m^2_{2\pi}})}\right|^2 \longrightarrow \left|1-\frac{D_\rho(\overline{m^2_{4\pi}})}{2}\left[\frac{1}{D_\rho(\overline{m^2_{2\pi}})}-\frac{1}{3m^2_\rho}\right]\right|^2 \approx \left|1-\frac{D_\rho(\overline{m^2_{4\pi}})}{3D_\rho(\overline{m^2_{2\pi}})}\right|^2 \approx 0.75. \quad (4.14)$$

Using this value instead of 0.64 results in the increase of the above branching ratios by a factor of 1.17. One can see that the results of the order-of-magnitude estimates given in Section 4.2 agree with the careful calculations by a factor lower than 4.5.

The numerical value of the $\omega \longrightarrow 5\pi$ decay width changes by a factor of 2 when the energy varies within $\pm\Gamma_\omega/2$ around the ω mass. In other words, the dependence of this partial width on the energy is very strong. This is illustrated in Fig. 7 with the $\omega \longrightarrow 5\pi$ excitation curves in e^+e^- annihilation,

$$\sigma_{e^+e^-\to\omega\to 5\pi}(s) = 12\pi\left(\frac{m_\omega}{E}\right)^3\Gamma_{\omega e^+e^-}(m_\omega)\times\frac{\Gamma_\omega B_{\omega\to 5\pi}(E)}{[(s-m^2_\omega)^2+(m_\omega\Gamma_\omega)^2]}, \quad (4.15)$$

where $B_{\omega\to 2\pi^+2\pi^-\pi^0}(E)$ ($B_{\omega\to\pi^+\pi^-3\pi^0}(E)$) is given by Eq. (4.11) (respectively, (4.13)) with the substitution $m_\omega \longrightarrow E$. The strong energy dependence of the partial width results in the asymmetric shape of the ω resonance and the shift of its peak position by +0.7 MeV. As can be seen from Fig. 7, the peak value of the 5π state production cross section is about 1.5–2.0 fb. The decays $\omega \longrightarrow 5\pi$ can nevertheless be observable on e^+e^- colliders. Indeed, with the luminosity $L = 10^{33}$ cm^{-2} s^{-1} near the ω peak, which seems to be feasible, one may expect to detect about 2 events per week for the above decays.

4.4. Angular Distributions of the Final Pions in the $\omega \longrightarrow 5\pi$ Decay

The angular distributions of the final pions must of course be deduced from the full amplitudes in Eqs. (4.1) and (4.2). However, some qualitative conclusions about the angular distributions can be drawn from the simplified expressions in Eqs. (4.3)–(4.7). We find these distributions in the rest frame of the decaying ω meson. Since helicity is conserved, only the $\omega(782)$ states with the spin projections $\lambda = \pm 1$ on the e^+e^- beam axis given by the unit vector $\mathbf{n}_0$ are populated. The notation for the vector product of the pion 3-momenta used in what follows is

$$\mathbf{q}_i\times\mathbf{q}_j = |\mathbf{q}_i||\mathbf{q}_j| \times\sin\theta_{ij}(\sin\Theta_{ij}\cos\Phi_{ij},\ \sin\Theta_{ij}\sin\Phi_{ij},\ \cos\Theta_{ij}). \quad (4.16)$$

In other words, θ_{ij} is the angle between the pion momenta $\mathbf{q}_i$ and $\mathbf{q}_j$, with Θ_{ij} and Φ_{ij} being the polar and azimuthal angles of the normal to the plane spanned by $\mathbf{q}_i$ and $\mathbf{q}_j$. Choosing $\mathbf{n}_0$ to be the unit vector along the z axis, the probability density of the emission of two π^+'s with the momenta $\mathbf{q}_1$ and $\mathbf{q}_2$ and of π^0 with the momentum $\mathbf{q}_4$ is represented as

$$w \propto [\mathbf{q}_4\times(\mathbf{q}_1+\mathbf{q}_2)]^2-(\mathbf{n}_0\cdot[\mathbf{q}_4\times(\mathbf{q}_1+\mathbf{q}_2)])^2 = \mathbf{q}_4^2[\mathbf{q}_1^2\sin^2\theta_{41}\sin^2\Theta_{41}+\mathbf{q}_2^2\sin^2\theta_{42}\sin^2\Theta_{42} + 2|\mathbf{q}_1||\mathbf{q}_2|\sin\Theta_{41}\sin\Theta_{42}\sin\theta_{41} \times\sin\theta_{42}\cos(\Phi_{41}-\Phi_{42})] \quad (4.17)$$

in the case of the final state $2\pi^+2\pi^-\pi^0$. Here, the momentum assignment is the same as in Eq. (4.1). The angular distribution of π^0 and of two π^-'s with the momenta $\mathbf{q}_3$, $\mathbf{q}_5$ is obtained from Eq. (4.17) by the replacement $\mathbf{q}_{1,2} \longrightarrow \mathbf{q}_{3,5}$, because the identity

$$\varepsilon_{\mu\nu\lambda\sigma}q_\mu\epsilon_\nu(q_1+q_2)_\lambda q_{4\sigma} = -\varepsilon_{\mu\nu\lambda\sigma}q_\mu\epsilon_\nu(q_3+q_5)_\lambda q_{4\sigma}$$

is satisfied. Another identity

$$\varepsilon_{\mu\nu\lambda\sigma}q_\mu\epsilon_\nu(q_1+q_2)_\lambda q_{4\sigma} = -\varepsilon_{\mu\nu\lambda\sigma}q_\mu\epsilon_\nu(q_1+q_2)_\lambda(q_3+q_5)_\sigma$$

allows one to write the angular distribution that includes four charged pions as

$$\begin{aligned} w &\propto [(\mathbf{q}_1+\mathbf{q}_2)\times(\mathbf{q}_3+\mathbf{q}_5)]^2 \\ &-(\mathbf{n}_0\cdot[(\mathbf{q}_1+\mathbf{q}_2)\times(\mathbf{q}_3+\mathbf{q}_5)])^2 \\ &= (1+P_{12})(1+P_{35})\mathbf{q}_1^2\mathbf{q}_3^2\sin^2\theta_{13}\sin^2\Theta_{13} \\ &+2|\mathbf{q}_1||\mathbf{q}_2|(1+P_{35})\mathbf{q}_3^2 \\ &\times\sin\theta_{13}\sin\theta_{23}\sin\Theta_{13}\sin\Theta_{23}\cos(\Phi_{13}-\Phi_{23}) \\ &+2|\mathbf{q}_3||\mathbf{q}_5|(1+P_{12})\mathbf{q}_1^2 \\ &\times\sin\theta_{13}\sin\theta_{15}\sin\Theta_{13}\sin\Theta_{15}\cos(\Phi_{13}-\Phi_{15}) \\ &+2|\mathbf{q}_1||\mathbf{q}_2||\mathbf{q}_3||\mathbf{q}_5|(1+P_{35}) \\ &\times\sin\theta_{13}\sin\theta_{25}\sin\Theta_{13}\sin\Theta_{25}\cos(\Phi_{13}-\Phi_{25}), \end{aligned} \quad (4.18)$$

where P_{ij} interchanges the indices i and j.

For the final state $\pi^+\pi^-3\pi^0$, the corresponding probability density can be obtained from Eqs. (4.6) and (4.7) as

$$w \propto [\mathbf{q}_1 \times \mathbf{q}_2]^2 - (\mathbf{n}_0 \cdot [\mathbf{q}_1 \times \mathbf{q}_2])^2 = \mathbf{q}_1^2 \mathbf{q}_2^2 \sin^2\theta_{21} \sin^2\Theta_{21}, \tag{4.19}$$

where the momentum assignment is the same as in Eq. (4.2). The corresponding angular distribution of one charged (e.g., π^+) and three neutral pions can be obtained from Eqs. (4.6) and (1.7) with the help of the identity

$$\varepsilon_{\mu\nu\lambda\sigma} q_\mu \epsilon_\nu q_{1\lambda} q_{2\lambda} = -\varepsilon_{\mu\nu\lambda\sigma} q_\mu \epsilon_\nu q_{1\lambda} (q_3 + q_4 + q_5)_\sigma$$

as

$$\begin{aligned} w \propto & \left[\mathbf{q}_1 \times \sum_i \mathbf{q}_i\right]^2 - \left(\mathbf{n}_0 \cdot \left[\mathbf{q}_1 \times \sum_i \mathbf{q}_i\right]\right)^2 \\ = & \mathbf{q}_1^2 \Bigg[\sum_i \mathbf{q}_i^2 \sin^2\theta_{i1} \sin^2\Theta_{i1} \\ & + 2 \sum_{i \neq j} |\mathbf{q}_i||\mathbf{q}_j| \sin\theta_{i1} \sin\theta_{j1} \\ & \times \sin\Theta_{i1} \sin\Theta_{j1} \cos(\Phi_{i1} - \Phi_{j1}) \Bigg], \end{aligned} \tag{4.20}$$

where the indices i and j run over 3, 4, 5.

The strong energy dependence of the five-pion partial width of ω implies that the branching ratio at the ω mass (Eq. (4.9), evaluated above) is slightly different from that determined by the expression

$$B^{\text{aver}}_{\omega \to 5\pi}(E_1, E_2) = \frac{2}{\pi} \int_{E_1}^{E_2} dE \frac{E^2 \Gamma_\omega B_{\omega \to 5\pi}(E)}{(E^2 - m_\omega^2)^2 + (m_\omega \Gamma_\omega)^2}. \tag{4.21}$$

Taking $E_1 = 772$ MeV and $E_2 = 792$ MeV, one finds

$$B^{\text{aver}}_{\omega \to 2\pi^+ 2\pi^- \pi^0}(E_1, E_2) = 9.0 \times 10^{-10}$$

and

$$B^{\text{aver}}_{\omega \to \pi^+ \pi^- 3\pi^0}(E_1, E_2) = 6.7 \times 10^{-10},$$

to be compared with Eq. (4.11) and (4.13), respectively. In particular, the quantity $B^{\text{aver}}_{\omega \to 2\pi^+ 2\pi^- \pi^0}(E_1, E_2)$ is the relevant characteristic of this specific decay mode in photoproduction experiments. The Jefferson Lab [19] could also be suitable for detecting the five-pion decays of ω. However, in view of the suppression of the ω photoproduction cross section by a factor of 1/9 as compared with the ρ one, the total number of ω mesons will amount to 7×10^8 per nucleon. Hence, the increase of intensity of this machine by a factor of 50 is highly desirable in order to observe the $\omega \longrightarrow 5\pi$ decay and to measure its branching ratio. Evidently, the ω photoproduction on heavy nuclei is preferable because the cross section grows with the mass number A as $A^{0.8–0.95}$ [20].

The conclusions about the angular distributions in photoproduction are as follows. Of course, their general expression should be deduced from the full decay amplitudes in Eqs. (4.1) and (4.2) together with the detailed form of the photoproduction mechanism. The qualitative picture, however, can be obtained by noting that the *s*-channel helicity conservation is a good selection rule for the photoproduction reactions. Then, in the helicity reference frame (which is the frame where ω is at rest and its spin quantization axis is directed along the ω momentum in the center-of-mass system), the expressions for the angular distributions coincide with the expressions in Eqs. (4.17)–(4.20). Since the direction of the final ω momentum lies at a scattering angle that is less than 0.5° for the photoproduction on heavy nuclei at high energies, the vector $\mathbf{n}_0$ can be treated as pointing along the photon-beam direction.

5. CONCLUSION

The results presented in this paper show that the left shoulder of the ρ peak is, by all appearance, the best place to study the effects of chiral dynamics as compared to the classical low-energy ππ scattering. The e^+e^- colliders with a sufficiently large luminosity at energies below the ρ mass provide the controlled source of soft pions. The role of higher derivatives and loop corrections in the low-energy effective Lagrangian for soft pions, as well as various schemes of incorporation of vector mesons into the chiral approach, can be successfully tested with such machines. The intense beams of photons from the Jefferson Laboratory are also of great importance in achieving the theoretical goals mentioned above. The decay $\omega \longrightarrow 5\pi$ is of special interest, because its kinematics is such that the final pions are essentially nonrelativistic, so that the effects of chiral dynamics are manifested most clearly. The measurements of the branching ratio of the five-pion decays of ω would be a challenge to experimenters, because this task would help in rigorously testing chiral theories that incorporate the vector mesons.

ACKNOWLEDGMENTS

We are grateful to G.N. Shestakov and A.M. Zaĭtsev for discussion. This work is supported in part by the RFBR-INTAS, grant no. IR-97-232.

REFERENCES

1. N. N. Achasov and A. A. Kozhevnikov, Phys. Rev. D **61**, 077904 (2000).
2. S. Weinberg, Phys. Rev. **166**, 1568 (1968).

3. M. Bando, T. Kugo, S. Uehara, *et al.*, Phys. Rev. Lett. **54**, 1215 (1985); M. Bando, T. Kugo, and K. Yamawaki, Nucl. Phys. B **259**, 493 (1985); Prog. Theor. Phys. **73**, 1541 (1985); Phys. Rep. **164**, 217 (1988).
4. L. Bányai and V. Rittenberg, Phys. Rev. **184**, 1903 (1969).
5. A. Bramon, A. Grau, and G. Pancheri, Phys. Lett. B **317**, 190 (1993).
6. S. I. Eidelman, Z. K. Silagadze, and E. A. Kuraev, Phys. Lett. B **346**, 186 (1995).
7. R. S. Plant and M. C. Birse, Phys. Lett. B **365**, 292 (1996).
8. S. L. Adler, Phys. Rev. B: Condens. Matter **137**, 1022 (1965); **139**, B1638 (1965).
9. J. Schwinger, Phys. Lett. B **24**, 473 (1967); J. Wess and B. B. Zumino, Phys. Rev. **163**, 1727 (1967); S. Gasiorowicz and D. A. Geffen, Rev. Mod. Phys. **41**, 531 (1969); O. Kaymakcalan, S. Rajeev, and J. Schechter, Phys. Rev. D: Part. Fields **30**, 594 (1984); U.-G. Meissner, Phys. Rep. **161**, 213 (1988).
10. K. Kawarabayashi and M. Suzuki, Phys. Rev. Lett. **16**, 255 (1966); Riazzuddin and Fayyazuddin, Phys. Rev. **147**, 1071 (1966).
11. J. Wess and B. B. Zumino, Phys. Lett. B **37**, 95 (1971).
12. E. Byckling and K. Kajantie, *Particle Kinematics* (Wiley, London, 1973).
13. R. Kumar, Phys. Rev. **185**, 1865 (1969).
14. T. W. Sag and G. Szekeres, Math. Comput. **18**, 245 (1964).
15. C. Caso, G. Conforto, A. Gurtu, *et al.*, Eur. Phys. J. C **3**, 1 (1998).
16. R. R. Akhmetshin, F. V. Anashkin, M. Arpagaus, *et al.*, hep-ex/9912020; Phys. Lett. B **475**, 190 (2000).
17. Y. S. Tsai, Phys. Rev. D: Part. Fields **4**, 2821 (1971).
18. F. J. Gilman and D. H. Miller, Phys. Rev. D: Part. Fields **17**, 1846 (1978); F. J. Gilman and S. H. Rhye, Phys. Rev. D: Part. Fields **31**, 1066 (1985).
19. A. R. Dzierba, in *Proceedings from Jefferson Lab/NCSU. Workshop on Hybrids and Photoproduction Physics, North Carolina State University, November 13–15, 1997*, p. 661.
20. D. W. G. S. Leith, in *Hadronic Interactions of Electrons and Photons* (Academic, New York, 1971), p. 195.

Journal of Experimental and Theoretical Physics, Vol. 91, No. 3, 2000, pp. 449–457.
Translated from Zhurnal Éksperimental'noĭ i Teoreticheskoĭ Fiziki, Vol. 118, No. 3, 2000, pp. 516–526.

ATOMS, SPECTRA, RADIATION

Forms of Localization of Rydberg Wave Packets

E. A. Shapiro

Institute of General Physics, Russian Academy of Sciences, Moscow, 117942 Russia
e-mail: zheka@gon.ran.gpi.ru

Received February 29, 2000

Abstract—A graphic method for describing the localization of quantum wave packets in several dimensions is proposed. The method employs classical action–angle coordinates. Different kind of wave packets arising in the two- and three-dimensional Coulomb problems are described from a unified point of view. Specifically, a more precise interpretation of the structure of "angular" Rydberg wave packets is given. Certain additional features of the proposed approach are discussed. © *2000 MAIK "Nauka/Interperiodica".*

1. INTRODUCTION

Localized quantum states (wave packets) have been studied since quantum mechanics came into existence. Nonetheless, such states have been attracting increasing attention in the last few years. The reason is that wave packets are quantum states which are closest to classical particles. These are objects which manifest both classical and exclusively quantum features. Consequently, as the experimental and technical possibilities increase, the question of the production and control of wave packets in practice is becoming increasingly more important in many fields ranging from microelectronics to chemistry. Specifically, wave packets arising in Rydberg (highly excited) atoms have been attracting a great deal of attention in the last few years. Such packets have been investigated theoretically and experimentally [1–3].

Wave packets arise in many processes as a superposition of quantum states with different quantum numbers. Often, when a superposition of quantum states in a specific process is under study it is desirable to have a simple method for answering questions such as: Is a wave function in a system with several degrees of freedom localized? What is the wave function like? Does such localization exist for any classical model? What are the dynamics of a localized state? In the present paper a quite simple method is proposed for giving a qualitative answer to these questions. Such an answer can often make it easier to understand the physics of a problem, underscoring the relation between its classical and quantum features. Such a method has not been discussed in the literature.

In classical mechanics the behavior of integrable systems is often described using action–angle coordinates. Then, as is well known, a transformation is made into a coordinate system where the momenta are the integrals of motion and the conjugate coordinates—the angles—vary linearly in time. If the actions and their conjugate angles are known, then everything is known about the motion of a classical system. However, a strict definition of the quantum action–angle coordinates encounters mathematical difficulties. Quantum action–angle coordinates are too difficult to use, and as a result they cannot be used as a simple tool for analyzing the properties of a wave function.

It is shown in the present paper how the classical action–angle coordinates can be conveniently used to analyze the localization and dynamics of quantum states. For this, the sense in which classical action–angle coordinates can describe a quantum state is discussed in Section 2. Specifically, the wave function is represented in a form that singles out the classical action and angles and it is shown how such a description can be applied to wave packets. This approach is especially convenient for describing systems with several degrees of freedom. The proposed method is explained in greater detail in Section 3 for a two-dimensional Coulomb problem. The localization and dynamics of various Rydberg wave packets in a two-dimensional problem are analyzed in the same section. Almost all these packets have been studied previously, but a unified approach to their description has never been proposed. Specifically, a simplified understanding of the structure of certain packets sometimes led to an incorrect conception of their localization. Wave packets arising in a three-dimensional Coulomb problem with $m = 0$ are studied in Section 4. Certain additional features of the proposed approach—semisiclassical quantization, the Wigner function, and the calculation of matrix elements—are discussed in Section 5. The results obtained in this paper are briefly summarized in Section 6.

2. SEMICLASSICAL REPRESENTATION OF ANGLES

It is well known [4, 5] that a semiclassical wave function of an integrable system can be constructed as follows. Let Γ be the Lagrangian manifold of the corresponding classical system, i.e., the set of points in phase space which are reached during motion with the

same integrals of motion and with different initial conditions. For finite motion a Lagrangian manifold is equivalent to a multidimensional torus [5, 6]. For one-dimensional motion, Γ is a trajectory in phase space. For the system to be quantizable Γ must satisfy the Bohr–Sommerfeld quantization rule (see below). A semiclassical, unnormalized wave function at an arbitrary point $\mathbf{r}$ in the classically allowed region far from singularities (turning points) is determined as [5]

$$\Psi(\mathbf{r}) = \sum_{a_m} \det\left|\frac{\partial^2 S}{\partial \mathbf{r}\partial \mathbf{J}}\right|^{1/2} \times \exp\left[i\left(S_{a_m} - Et - \frac{\pi\gamma_{a_m}}{2}\right)\right] \equiv \sum_{a_m} \Psi_{a_m}(\mathbf{r}), \quad (1)$$

where S is the action, $\mathbf{J} = \{J_i\}$ is a set of integrals of motion, γ_{a_m} is an index characterizing the point a_m and depends on the properties of Γ (see [4]), and $\{a_m\}$ is a set of points on Γ that project into a fixed point $\mathbf{r}$. It is obvious that there are $m = 2^{N_{osc}}$ such points, where N_{osc} is the number of degrees of freedom along which the system executes an oscillatory. Here and below $\hbar = 1$.

Action–angle coordinates can be introduced into the system being integrated. In this case, the integrals of motion (action) $\mathbf{J}$ are the momenta and the coordinates conjugate to them—the angles—are

$$\lambda = \{\lambda_i\} = \left\{\left.\frac{\partial S}{\partial J_i}\right|_{r = \text{const}}\right\}$$

(see [5, 6]). The action and angles can be normalized in a manner so that in classical motion the λ_i vary linearly from 0 to 2π. According to the Bohr–Sommerfeld quantization rule

$$\mathbf{J} = \mathbf{n} + \gamma/4,$$

where $\boldsymbol{\gamma} = \{\gamma_i\}$ is a set of Maslov indices corresponding to different basic cycles [4].

We shall now study a superposition of states with quantum numbers $\mathbf{n}$, closely spaced around $\mathbf{n}_0$, and amplitudes $C_{\mathbf{n}}$. For every such state the action at the points of the corresponding Lagrangian manifold $a_m(\mathbf{r})$ that project into the point $\mathbf{r}$ is

$$S^{\mathbf{n}}_{a_m} \approx S^{\mathbf{n}_0}_{a_m} + (\mathbf{n} - \mathbf{n}_0)\lambda, \quad (2)$$

where $S^{\mathbf{n}}_{a_m}$ is the action at the point a_m in classical motion with $\mathbf{J} = \mathbf{n} + \gamma/4$.

The phase of classical motion along the orbit λ at a given point $\mathbf{r}$ remains essentially unchanged for small variations of $\mathbf{J}$. Consequently, it can be assumed that the wave function of a superposition of states is approximately

$$\Psi(\mathbf{r}, t) = \sum_{a_m} \left|\det\frac{\partial\lambda}{\partial\mathbf{r}}\right|^{1/2} \exp[iF_{a_m}]\Psi_\lambda. \quad (3)$$

In Eq. (3) $\lambda(\mathbf{r})$ is estimated at $\mathbf{n} = \mathbf{n}_0$, and

$$\Psi_\lambda(\lambda, t) = \sum_{\mathbf{n}} C_{\mathbf{n}} \exp(i(\mathbf{n}\lambda - E_{\mathbf{n}}t)), \quad (4)$$

where $E_{\mathbf{n}}$ is the energy for classical motion with $\mathbf{J} = \mathbf{n} + \gamma/4$. The function

$$F_{a_m} = S^{\mathbf{n}_0}_{a_m} - \mathbf{n}_0\lambda - \gamma_{a_m}\pi/2$$

incorporates the generating function of a classical canonical transformation to action–angle coordinates with $\mathbf{J} = \mathbf{n} + \gamma/4$, equal to $S^{\mathbf{n}_0}_{a_m} - \mathbf{J}\lambda$ [6], and the additional phase

$$\gamma_{a_m}\pi/2 + \gamma\lambda/4.$$

In Eq. (3) the expression for the wave function for each point a_m is divided into two parts. Ψ_λ characterizes only a quantum state given by the set of amplitudes $C_{\mathbf{n}}$. The second term, containing the preexponential factor and the exponential in Eq. (3), characterizes the r-representation in which the wave function is written. We assume that it is independent of the set of quantum numbers, since $\lambda(\mathbf{r})$ changes very little with small variations of $\mathbf{J}$.

In what follows we shall say that Ψ_λ gives a wave function in the angle λ_i representation. This terminology, which is not absolutely precise, corresponds to the general principles of the construction of action–angle quantum coordinates [7–9].

Since the function Ψ_λ is found as a Fourier transform of the amplitudes of the states $C_{\mathbf{n}}\exp(-iE_nt)$, its localization is quite easily described. If the function Ψ_λ is strongly localized near some values $\lambda(t)$, then, evidently, $\Psi(\mathbf{r})$ is localized near $\mathbf{r}(\lambda(t))$.

If the function $\Psi_\lambda(\lambda)$ is not strongly localized (for example, the localization width $\Delta\lambda > \pi$), it is more convenient to express the quantum state as

$$\Psi(\mathbf{r}, t) = \sum_{a_m} \Psi^{\mathbf{n}_0}_{a_m}(\mathbf{r}, t) f(\lambda_{a_m}, t), \quad (5)$$

where

$$f(\lambda, t) = \Psi_\lambda \exp(-i(\mathbf{n}_0\lambda - E_{n_0}t)),$$

and $\Psi^{\mathbf{n}_0}_{a_m}$ are the terms corresponding to different points a_m in Eq. (1) for states with the quantum numbers $\mathbf{n}_0$.

The formula (5) can be interpreted as follows. A semiclassical wave function can be represented as

being distributed on Γ (in the one-dimensional case—on the phase trajectory) and then projected into the r space by adding a preexponential factor. Then, oscillations of the wave function as a function of $\mathbf{r}$ arise because of the interference of terms corresponding to different sheets of Γ. This approach (Einstein–Brilloui–Keller approach [5]) is incorporated in Eq. (1), where the action can be calculated at the points of Γ that are projectable into $\mathbf{r}$. According to Eq. (5), however, the "wave packet" is the envelope moving along Γ and modulating the central state of the packet. The function $f(\lambda, t)$ determines the form and dynamics of this envelope. Specifically, if $f(\lambda)$ is localized on one sheet of the Langrangian manifold (on one branch of the phase trajectory), then there is no interference of the terms corresponding to different a_m and the wave function of the packet has no oscillations which are characteristic for the eigenstates.

Thus, the prescription for describing the spatial localization of the wave function of a superposition of quantum states is as follows. The geometric meaning of the classical angles corresponding to the quantum numbers of the problem must be determined. Next, the wave function of the superposition of states must be written in the angle representation (4), or (which is actually the same thing) the modulating function f in the Eq. (5) must be determined. The localization of Ψ_λ or f in some range of angles determines the localization of the wave function in the coordinate region corresponding to these angles.

In what follows we shall illustrate the application of this principle for various Rydberg wave packets. The method will be used to analyze the localization of a wave function in various packets. Then this analysis will be illustrated by means of numerical calculations.

3. TWO-DIMENSIONAL COULOMB PROBLEM

In this section the various Rydberg wave packets arising in a two-dimensional Coulomb problem will be considered. Such a model problem is often used to describe electronic states with large quantum numbers $m \simeq l$ in fields where the motion of the electron remains almost planar [10–14].

The motion of an electron in the classical two-dimensional Coulomb problem can be described using the following two pairs of action–angle variables [15]. The first pair consists of the angular momentum $J_m = p_\phi$ and its conjugate angle ϕ_{xy}, which is the angle between the direction to the perihelion of the electron orbit and a fixed direction in the xy plane. The second pair consists of the action

$$J_n = p_\phi + \frac{1}{2\pi}\oint p_r dr$$

and its conjugate angle θ—the average anomaly of the electron, giving its position on an ellipse. The orbit of a classical electron in a two-dimensional Coulomb problem is an ellipse with semiaxes J_n^2 and J_nJ_m and with focus located at the nucleus of the atom. The average anomaly θ is proportional to the area swept out by the radius vector of an electron in the classical motion. The value $\theta_0 = 0$ corresponds to an interior turning point and $\theta_0 = \pi$ corresponds to an exterior turning point.

In classical motion

$$H = -\frac{J_n^2}{2}, \quad \dot{\phi}_{xy} = 0, \quad \dot{\theta} = \omega_k = 1/J_n^3. \tag{6}$$

For quantization according to the Bohr–Sommerfeld rule

$$J_n = n + 1/2, \quad n = 0, 1, 2, \ldots, \tag{7}$$

$$J_m = m, \quad m = -n, \ldots, n. \tag{8}$$

In classical motion, the two-dimensional electron undergoes oscillations only along the radius, one of the two polar coordinates. The semiclassical wave function at in arbitrary point $\mathbf{r}$ of a classically allowed region is given by the sum of two terms corresponding to the motion of an electron along two ellipses passing through the point $\mathbf{r}$ (Fig. 1). In Eq. (1) the indices γ_{a_m} for these two terms differ by 1. The wave function Ψ_λ of these two terms is, correspondingly,

$$\Psi_\lambda^{n,m,1,2} = \exp[i(n\theta_{1,2} + l\phi_{xy\,1,2} - E_n t)], \tag{9}$$

where

$$E_n = -\frac{1}{2(n+1/2)^2},$$

$\phi_{xy\,1}$ and $\phi_{xy\,2}$ are the slope angles of the two ellipses, and $\theta_1 = -\theta_2$ are the average anomalies characterizing the point $\mathbf{r}$ in the motion along the first or second ellipse. The function F in Eq. (3) can be found by calculating numerically or analytically the action S and the coordinates θ and φ for each ellipse at each point.

To calculate the semiclassical wave function of the superposition of states with different values of n and m, it would be necessary to draw two ellipses with semiaxes J_n^2 and J_nJ_m through the point $\mathbf{r}$ for each pair n, m and then find on each ellipse the value of the action at the point $\mathbf{r}$, calculate the preexponential factor in Eq. (1), and sum all values of the characteristic wave functions with the required coefficients.

In the approximate expression (3) it is assumed that all ellipses for different values of n and m are the same as for the central state n_0, m_0. Then it is sufficient to draw through the point $\mathbf{r}$ only two ellipses, to calculate on each ellipse

$$\Psi_\lambda = \sum_{n,m} C_{n,m}\exp[i(n\theta + m\phi_{xy} - E_n t)],$$

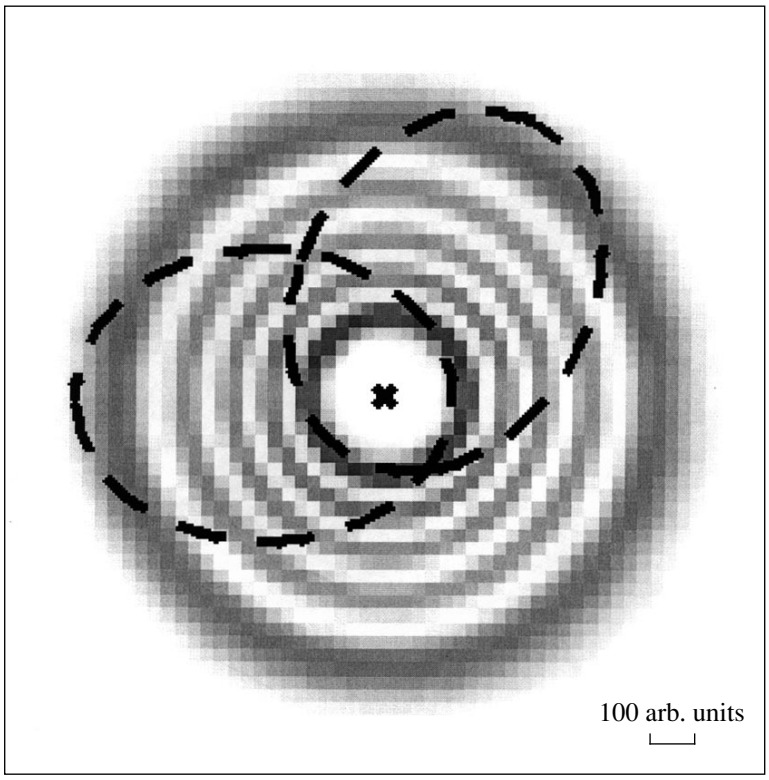

Fig. 1. Squared wave function of the state with $n = 19$, $m = 15$. The broken lines show two ellipses corresponding to classical motion with the actions $J_n = 19.5$ and $J_m = 15$.

and substitute this expression into Eq. (3), using the function F found for the central state.

However, according to the Eq. (5), there is no need to calculate the semiclassical wave function of the superposition in order to describe the wave packet. It is sufficient to know the wave function of one central state $\Psi^{n_0, m_0}(\mathbf{r})$, the function whose exact form is well known [16] and which can be easily found numerically without using the semiclassical approximation. The wave function of the packet $\Psi(\mathbf{r}, t)$ is $\Psi^{n_0, m_0}(\mathbf{r})$, normalized by the envelope

$$f(\theta, \phi_{xy}, t) = \Psi_\lambda \exp[-i(n_0\theta + m_0\phi_{xy} - E_{n_0}t)].$$

The function $\Psi(\mathbf{r}, t)$ is more localized at locations, where $f(\theta, \phi_{xy}, t)$ is localized. Given the geometric meaning of the angles θ and ϕ_{xy} it is easy to predict the region and degree of localization of the wave function. The presence or absence of an interference pattern in the region of localization depends on the relative contribution made to the wave function at a particular point by both sets of angles θ_1, $\phi_{xy\,1}$ and θ_2, $\phi_{xy\,2}$ corresponding to this point, i.e., on the degree of localization of the function f. This latter approach will be assumed below in the interpretation of the form of the localization of the wave function of superpositions of states.

In this and the next sections various wave packets in two- and three-dimensional Coulomb problems will be studied. The form of the localization of various packets will be predicted qualitatively using the analysis in the angle representation. As an illustration, we shall find numerically the wave functions of the packets using the exact (not semiclassical) expressions for the characteristic states of an atom.

As a first example of a wave packet, we shall study the superposition of quantum states with the same angular momentum m_0 and with amplitudes having a Gaussian distribution over n:

$$\Psi^{rwp} = \sum_n \exp\left[-\frac{(n-n_0)^2}{\sigma^2} - i\theta_0(n-n_0)\right]\Psi^{n, m_0}. \quad (10)$$

Since the Fourier transform of a Gaussian is a Gaussian, the function $\Psi_\lambda(\theta, \phi)$ (and hence $f(\theta, \phi)$ also) is localized as a function of θ near θ_0 and is not localized as a function of ϕ_{xy}. All possible ellipses contribute to the wave function of such a state, and on each ellipse the regions θ close to θ_0 contribute. This means that $\Psi^{rwp}(\mathbf{r})$ should be localized along the radius near $r(\theta_0)$ and should not be localized along the polar coordinate. Such states are called “radial wave packets” [1, 2, 17].

Figure 2a shows an example of the wave function of a radial wave packet. Here $n_0 = 19$, $m_0 = 15$, $\sigma = 3$, and

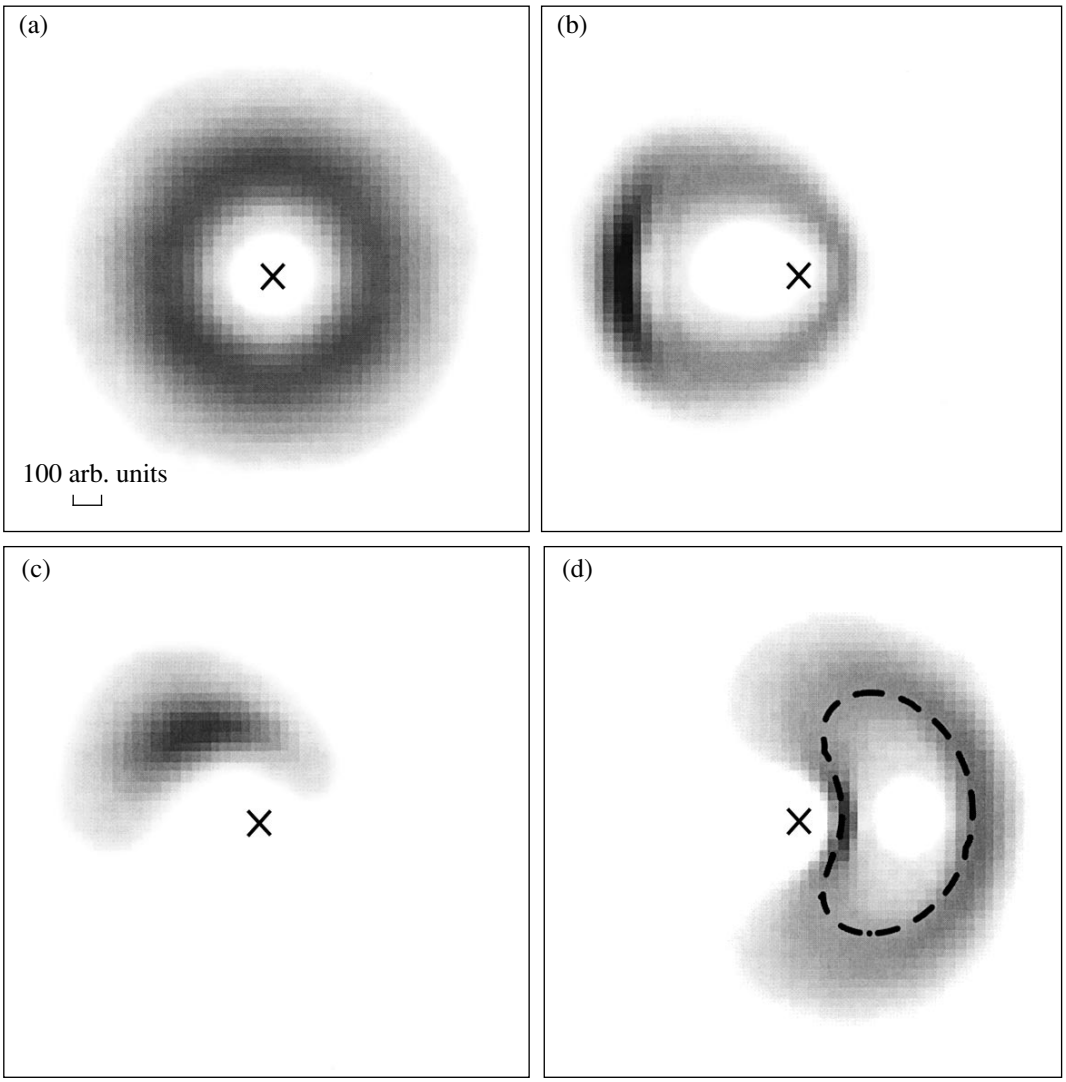

Fig. 2. Wave packets of a two-dimensional Coulomb problem: (a) radial wave packet, (b) elliptical state, (c) wave packet on an elliptical orbit, (d) the state Ψ^{CP}. The broken line corresponds to the curve $\theta + \phi_{xy} = 0$.

$\theta_0 = \pi/3$. The superposition wave function, calculated without using any approximations, in Fig. 2a is localized along the radius and is not localized along the angle. This corresponds to the prediction of our theory.

The radial wave packets were produced experimentally using short optical pulses with a wide spectrum [1, 2]. However, such states are excited in a strong field as a result of Raman bound–bound transitions via the continuum [17]. Moreover, such states can be produced in a microwave field with frequency close to ω_k [10, 18].

To describe the dynamics of a radial wave packet we shall expand, as usual, the energy E_n up to second order around E_{n_0}:

$$E_n = E_{n_0} + E'N + \frac{E''}{2}N^2,$$

where

$$N = n - n_0, \quad E' = -\omega_k = -\frac{1}{(n_0 + 1/2)^3},$$

$$E'' = \frac{3}{(n_0 + 1/2)^4}.$$

According to Eq. (4), we obtain for Ψ_λ

$$\Psi_\lambda^{rwp}(t) = \exp(i[n_0\theta + m_0\varphi - E_{n_0}t])$$
$$\times \sum_N \exp\left[-N^2\left(\frac{1}{\sigma^2} - \frac{iE''t}{2}\right) + iN(\theta - \theta_0 - \omega_k t)\right]. \quad (11)$$

Hence, it is evident that the wave function Ψ^{rwp} is an unperturbed wave function of a state with quantum

numbers n_0, m_0 that is modulated by a Gaussian wave packet moving along θ (i.e., oscillating along the radius) with frequency ω_k. This packet spreads in a characteristic time $t_{spr} = 2/\sigma^2 E''$ and is completely restored in a time $t_{rev} = \pi/E''$.

The expression (11) also describes the fractional restoration of the radial wave packet in a time $t_{rev} i/j$, when the term $E''t$ adds the corresponding phases to all components of the sum (11). Indeed, at the time $t = t_{rev} i/j$ the phase $N^2 E''t/2$ is identical for the states N, $N \pm k$, $N \pm 2k$, …, where k depends on i/j [19]. In this case the sum (11) decomposes into k subsums, each of which contains only terms with the numbers $N + s$, $N + s \pm k$, … and reproduces the initial localized distribution k times on the classical trajectory. According to the formula (11), if at $t = 0$ Ψ_λ is given by a localized 2π periodic function $G_1(\theta)$ (1), then at $t = t_{rev} i/j$

$$\Psi_\lambda^{rwp}(t_{rev} i/j) = \frac{1}{k} \times \sum_{s=0}^{k-1} \exp\left[i\left(\frac{2\pi s^2 i}{j} + s(\theta - \omega_k t) \right) \right] G_k(\theta - \omega_k t), \tag{12}$$

where G_k is a $2\pi/k$ periodic function which is identical to G_1 in the region of localization. As a result of interference of terms with different values of s, depending on i/j, the distribution (12) can represent from 0 to k packets on a classical trajectory. The expression (12) is identical to the known expression describing the partial restorations of the wave packets [19], and its derivation is clearer than the corresponding arguments in the r representation. In addition, the arguments in the angle representation (4) transfer directly to the motion of a packet along several angles simultaneously.

We shall now consider the superposition of states with the same principal quantum number n_0 and with a Gaussian distribution over m near m_0. The wave function of such a state

$$\Psi_\lambda^{e2d} = \exp(i[n_0\theta - E_{n_0} t]) \times \sum_m \exp\left[-\frac{(m - m_0)^2}{\sigma^2} + im(\phi_{xy} - \phi_{xy\,0}) \right] \tag{13}$$

is localized near the angle $\phi_{xy\,0}$. Only ellipses with slope angles close to $\phi_{xy\,0}$ contribute to the wave function of such a state. This means that $\Psi^{e2d}(\mathbf{r})$ is localized near an ellipse corresponding to classical motion with $J_n = n_0 + 1/2$, $J_m = m_0$, and $\phi_{xy} = \phi_{xy\,0}$. An example of such a state with $n_0 = 19$, $m_0 = 15$, $\sigma = 3$, and $\phi_{xy\,0} = 0$ is shown in Fig. 2b. This state can be called "elliptical."

In the literature localized superpositions with respect to the quantum number m are usually called "angular wave packets" [1, 12] in view of the somewhat simplified understanding of their angular distribution. It is obvious from the analysis performed here that the so-called "angular packets" are localized not so much along the polar angle as along the angle ϕ_{xy} near one central ellipse. When the wave function of such a state is analyzed, it is usually easy to see the elliptical type localization (see, for example, the "angular packet" in [1]).

States with the sharpest localization near a Keplerian ellipse were described as coherent quantum states corresponding to the hidden symmetry of a two-dimensional Coulomb problem [12]. They can be produced, for example, by using a combination of pulses of crossed electric or electric and magnetic fields [20]. If Ψ_λ is localized near several angles $\phi_{xy\,k}$, the wave function $\Psi(\mathbf{r})$ is localized near several Keplerian ellipses (see Fig. 1 in [10]).

If the superposition of states is localized along m and n, $\Psi(\mathbf{r})$ is one or several wave packets revolving around the nucleus along elliptical orbits. Examples of such states can be found in [1, 2, 10, 12]. Such a packet can be produced, for example, from an elliptical state using a microwave field in resonance with respect to ω_k or from a low-lying state by using a combination of a short optical pulse and a microwave or constant field. Figure 2c shows such a state with $n_0 = 19$, $m_0 = 15$, $\sigma_n = \sigma_l = 3$, $\theta_0 = \pi/3$, and $\phi_{xy\,0} = 0$. The dynamics of such a state in the absence of a field is identical to that of a radial packet. A resonance microwave field strongly suppresses the spreading along the angle θ and adds a very slow precession and spreading along the angle ϕ_{xy} [10].

Finally, we shall consider localized states of the form

$$\Psi^{CP}(\mathbf{r}) = \sum_N \exp\left[-\frac{N^2}{\sigma^2} \right] \Psi^{n_0 + N,\, m_0 + N}. \tag{14}$$

Such a state can be produced, for example, from a state with quantum numbers n_0, m_0 by a circularly polarized field with frequency close to ω_k [10, 13]. Substituting the expression (14) into Eq. (4), we find that Ψ_λ^{CP} (and hence also $\Psi^{CP}(\mathbf{r})$) possesses Gaussian localization around the curve $\theta + \phi_{xy} = 0$. This curve is a stable trajectory of a classical electron in a coordinate system rotating together with the field. Figure 2d shows the state with $n_0 = 19$, $m_0 = 15$, and $\sigma = 3$, found numerically, and the corresponding classical trajectory.

If in the expansion (14) the state (n_0, m_0) is circular, i.e., $m_0 = n_0$, $\Psi^{CP}(\mathbf{r})$ is a completely localized wave packet revolving around the nucleus along a circular trajectory. The dynamics of such wave packets has been studied many times in the literature [13, 14].

4. WAVE PACKETS IN A THREE-DIMENSIONAL COULOMB PROBLEM

We shall now consider a three-dimensional Coulomb problem. The corresponding actions are quantized according to the Bohr–Sommerfeld rule as

$$L_m = p_{\varphi_{\text{polar}}} = m, \quad m = -l, \ldots, l,$$

$$J_l = \frac{1}{2\pi}\oint p_{\psi_{\text{polar}}} d\psi_{\text{polar}} + J_m = l + \frac{1}{2}, \quad l = 0, \ldots, n, \tag{15}$$

$$J_n = \frac{1}{2\pi}\oint p_r dr + J_l = n + 1, \quad n = 0, 1, 2, \ldots$$

The angles corresponding to these actions are the average anomaly θ and two other Keplerian elements of the orbit—the argument of the pericenter and the longitude of the ascending angle [15]. The notation for the principal quantum number n, introduced here according to the Bohr–Sommerfeld quantization rule, differs by 1 from the conventional value.

The description of quantum states is especially simple when the numbers m are very large or very small. If $m \sim l \gg 1$, the electron wave function is localized near the xy plane and the problem is essentially equivalent to a two-dimensional problem.

Let us consider the opposite case, $m = 0$. In this case the wave function possesses radial symmetry relative to the z axis. In the xz plane the position of a classical electron is characterized by the average anomaly θ and the angle ϕ_z (the argument of the pericenter of the orbit) characterizing the slope of the orbit with respect to the z axis.

The classical motion includes oscillations along two polar coordinates (ψ_{polar} and r), and the wave function at the point $\mathbf{r}$ is given by a sum of four terms corresponding to motion along four ellipses and characterized by $\Psi_\lambda$of the form

$$\Psi_\lambda^{n, l, 1, 2, 3, 4} = \exp[i(n\theta_{1, 2, 3, 4} + l\phi_{z\,1, 2, 3, 4} - E_n t)], \tag{16}$$

where

$$E_n = -\frac{1}{2(n+1)^2}, \quad \theta_4 = \theta_1, \quad \theta_2 = \theta_3 = -\theta_1,$$

$$\phi_{z\,3} = \phi_{z\,1}, \quad \phi_{z\,4} = \phi_{z\,2}.$$

Compared with the two-dimensional problem, here there are two additional ellipses with the opposite sign of ϕ_z and simultaneously with the plane of classical motion rotated by π relative to the z axis. Consequently, the interference pattern in the xz plane is more complicated than in the xy plane. Moreover, the preexponential factor in Eq. (1) is larger for points $\mathbf{r}$ near the z axis, which corresponds to classical motion with $J_m = 0$. Figure 3a shows the probability distribution in the xz plane for the state $n = 19$, $l = 15$, and $m = 0$.

Just as in the two-dimensional case, the superposition of states with different n and the same value of l is a radial wave packet. The angular structure of a radial wave packet is identical to that of a state with the quantum numbers n_0, l in the corresponding formula (5). Figure 3b shows the numerically constructed radial wave packet with $n_0 = 19$, $l = 15$, $m = 0$, $\sigma = 3$, and $\theta_0 = \pi$. The dynamics of such a packet is identical to the two-dimensional case.

Let us consider the superposition of states with the same n but different l:

$$\Psi^{e3d} = \sum_l \exp\left[-\frac{(l-l_0)^2}{\sigma^2} - i\phi_{z\,0}(l-l_0)\right]\Psi^{n_0, l}. \tag{17}$$

Such a state is localized around the angle $\phi_{z\,0}$. In the xz plane this localization corresponds to a superposition of two ellipses, one of which is obtained from the other by reflection relative to the z axis. At the locations where these ellipses intersect the wave function is an interference pattern corresponding to the state n_0, l_0. Figures 3c, 3d show the state Ψ^{e3d} with $n_0 = 19$, $l_0 = 15$, $m = 0$, $\sigma = 3$, $\phi_{z\,0} = \pi$, and $\phi_{z\,0} = \pi/2$. Such states are also often called "angular wave packets," having in mind localization with respect to the polar angle. In agreement with our analytical predictions, Figs. 3c, 3d show that such an interpretation is not completely correct. Such states can be produced from a single Rydberg state by using a constant electric-field pulse [21].

Just as in the two-dimensional case, simultaneous localization along θ and ϕ_z corresponds to wave packets moving along ellipses in the xz plane. In this case the complete wave function is a wave packet which is symmetric with respect to rotations around the z axis and oscillates along the z axis. Such a state can arise, for example, when a linearly polarized microwave field acts on the state (n_0, l_0) [18].

In order for a three-dimensional quantum state to be localized near a single ellipse localization with respect to ϕ_z and ϕ_{xy} is required, i.e., a superposition with respect to both quantum numbers l and m is necessary. The generalized coherent states of the three-dimensional Coulomb problem possess this property [22]. Just as the two-dimensional elliptical states, they can be produced, making use of symmetry conservation in electric or crossed electric and magnetic fields.

5. ADDITIONAL FEATURES OF THE PROPOSED APPROACH

We have shown that the angle representation (4) makes it possible to describe in a simple manner the localization and dynamics of multidimensional semiclassical wave packets. We shall consider several important additional features of this representation.

The first important feature is the possibility of describing the system by using intuitively obvious "unified" semiclassical quantization [23–26]. Then the actions J_i are replaced by the operators $-i\partial/\partial\lambda_i + \gamma_i/4$ and periodic boundary conditions are imposed on the wave function. In this approach $\mathbf{J}$ and λ commute correctly from the standpoint of quantum mechanics, the energy and action of the system are determined to semiclassical accuracy, and Ψ_λ is the wave function of the

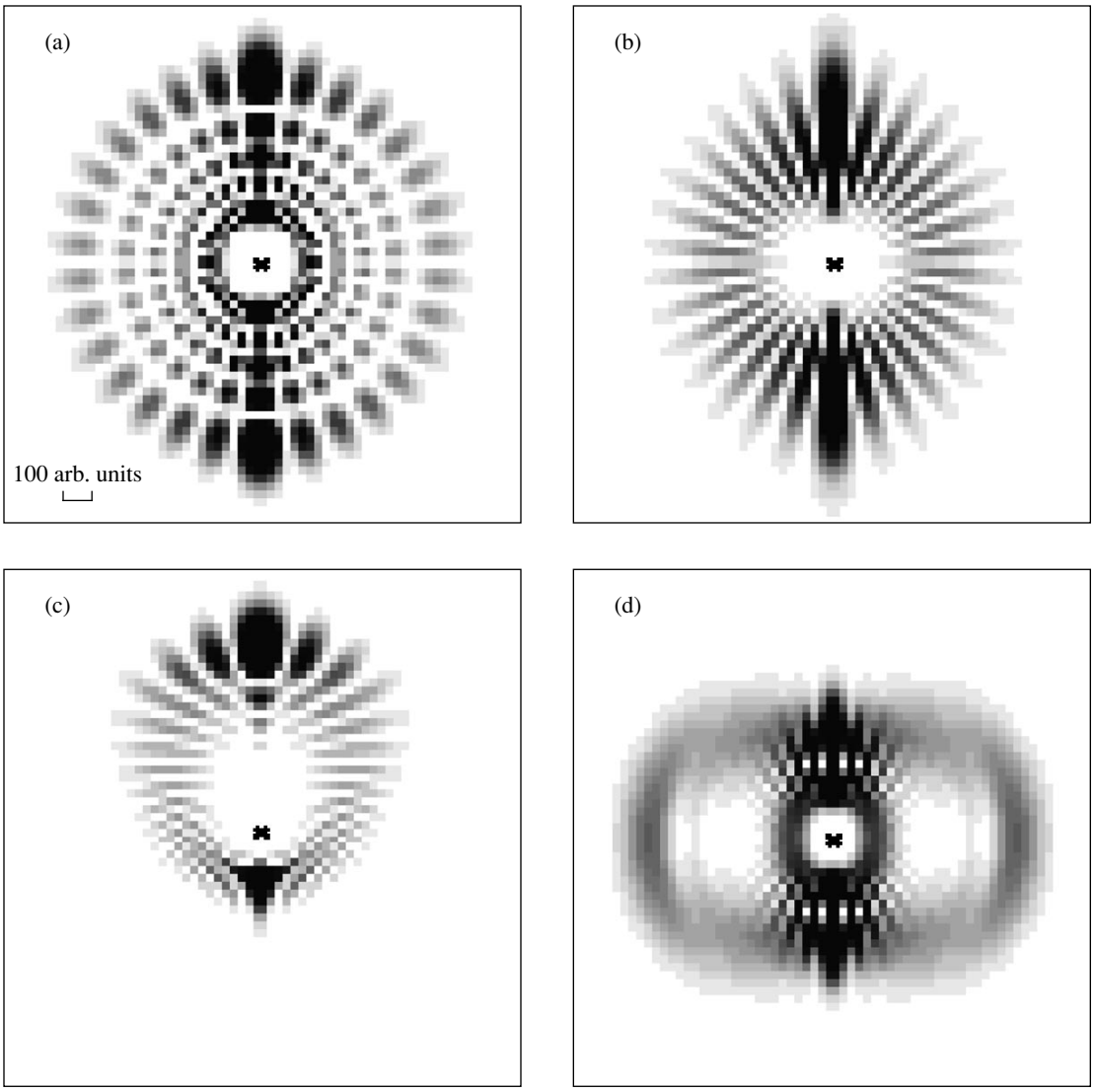

Fig. 3. (a) Distribution of the squared wave function in the *xz* plane for the state $n = 19$, $l = 15$, $m = 0$; (b) three-dimensional radial wave packet in the *xz* plane *z*; (c) and (d) packets corresponding to a superposition of states with different values of *l*.

system. This semiclassical method, which gives a unified description of all points of a classical trajectory, has been used to find the energies of semiclassical systems. According to the Eqs. (3) and (5), it can also describe the localization of a wave function.

An important feature of such an approach follows from the following fact. The quantum theory of nonlinear resonance [9] reproduces all corresponding steps of the classical perturbation theory with respect to the action [6] and actually uses Ψ_λ as a formal auxiliary tool. This tool is used to describe an isolated resonance and in regions where several resonances can interact. Consequently, "unified" semiclassical quantization is applicable not only to strictly integrable systems but also to a wide class of perturbed systems.

In such systems the form of the localization of a wave function can be most easily determined directly from the classical approximate (averaged) Hamiltonian.

A second important feature of the representation (4) is its convenience for calculating semiclassical matrix elements. Indeed, let a perturbation $V(\mathbf{r}, t)$ act on a system. When calculating the matrix element of the perturbation between states with close quantum numbers $\mathbf{n}_1$ and $\mathbf{n}_2$, the fact that the values of the angles $\lambda(\mathbf{r})$ for them are almost the same must be taken into account, just as in the description of wave packets. Consequently, only the expression

$$\Psi_\lambda^{\mathbf{n}_2 *} V \Psi_\lambda^{\mathbf{n}_1} = \exp\{i[(\mathbf{n}_1 - \mathbf{n}_2)\lambda - (E_{n_1} - E_{n2})t]\} V(\lambda(\mathbf{r}), t). \quad (18)$$

remains in the integral determining the matrix element. This approach immediately answers the question of which of the multidimensional Fourier components of the perturbation $V(\mathbf{r}, t)$ describes particular transitions in a system with several degrees of freedom.

The latter remark refers to the form of the Wigner function in the representation (4). This function, determined by the expression

$$W(\mathbf{r}, \mathbf{p}) = \int d\mathbf{Q} \int d\mathbf{P} \exp(-i(\mathbf{p} \cdot \mathbf{Q} + \mathbf{r} \cdot \mathbf{P})) \times \langle \Psi | \exp(i(\hat{\mathbf{p}} \cdot \mathbf{Q} + \hat{\mathbf{r}} \cdot \mathbf{P})) | \Psi \rangle, \quad (19)$$

is often used as a quantum-mechanical analog of a classical representation in phase space [8, 25]. For large quantum numbers it is very difficult to calculate W in the r representation. Moreover, the semiclassical Wigner function possesses singularities that depend on the quantum state and not only on the representation [25]. These facts make it difficult to use W for qualitative analysis of semiclassical quantum systems. At the same time, it is easy to calculate W in the representation (4). For a state with quantum numbers $\mathbf{n}$ it has the form of a δ function on the torus $\mathbf{J} = \mathbf{J}_\mathbf{n}$, corresponding to classical motion [25]. The representation (4) can strongly simplify the analysis of a quantum system in phase space when the structure of the system, and not the structure of the singularities, is important.

6. CONCLUSIONS

In this paper a quite simple approach was proposed for describing wave packets in multidimensional problems. In this approach attention is focused primarily on the question of the angular localization corresponding to combinations with a particular set of quantum numbers. The angular representation (3)–(5) was used to describe quantum states. It was shown that the expression for the wave function of a wave packet in this representation determines the envelope modulating the central state in the packet.

As an application, various Rydberg wave packets arising in two- and three-dimensional Coulomb problems were described in a unified manner. This analysis made it possible to elucidate the form of the localization of the so-called "angular wave packets" as well as some other localized states. Accurate (not semiclassical) numerical calculations support the qualitative predictions obtained using the proposed approach.

The representation (4) can be obtained by means of "unified" quantization of the corresponding classical Hamiltonian. The approach in some cases can make it easier to describe perturbed multidimensional systems and the dynamics of systems in phase space.

ACKNOWLEDGMENTS

I thank M.V. Fedorov for consultations and for reading this paper. This work was supported by the Russian Foundation for Basic Research (project no. 00-02-17058) and the Russian-American program CRDF.

REFERENCES

1. G. Alber and P. Zoller, Phys. Rep. **199**, 231 (1991).
2. M. Nauenberg, C. R. Stroud, Jr., and J. A. Yeazell, Sci. Am. **270**, 44 (1994).
3. T. C. Weinacht, J. Ahn, and P. H. Bucksbaum, Nature **397**, 233 (1999).
4. V. P. Maslov, *Asymptotic Methods and Perturbation Theory* (Nauka, Moscow, 1988).
5. A. M. O. de Almedia, *Hamiltonian Systems – Chaos and Quantization* (Cambridge Univ. Press, Cambridge, 1988).
6. V. I. Arnold, *Mathematical Methods of Classical Mechanics* (Nauka, Moscow, 1978; Springer-Verlag, New York, 1989).
7. S. D. Augustin and H. Rabitz, J. Chem. Phys. **71**, 4956 (1979).
8. L. Mandel and E. Wolf, *Optical Coherence and Quantum Optics* (Cambridge Univ. Press, Cambridge, 1995).
9. G. P. Berman and A. R. Kolovskiĭ, Usp. Fiz. Nauk **162**, 95 (1992) [Sov. Phys. Usp. **35**, 303 (1992)].
10. E. A. Shapiro and Paolo Bellomo, Phys. Rev. A **60**, 1403 (1999).
11. J. A. Yeazell and C. R. Stroud, Jr., Phys. Rev. A **35**, 2806 (1987).
12. M. Nauenberg, Phys. Rev. A **40**, 1133 (1989).
13. I. Bialynicki-Birula, M. Kalinski, and J. H. Eberly, Phys. Rev. Lett. **73**, 1777 (1994).
14. Z. D. Gaeta and C. R. Stroud, Jr., Phys. Rev. A **42**, 6308 (1990).
15. H. Goldstein, *Classical Mechanics* (Addison-Wesley, Reading, 2nd ed.; 1980; Nauka, Moscow, 1975).
16. X. L. Yang, S. H. Guo, and F. T. Chan, Phys. Rev. A **43**, 1186 (1991).
17. M. V. Fedorov, *Electron in Intense Light Field* (Nauka, Moscow, 1991).
18. A. Buchleitner, K. Sacha, D. Delande, and J. Zakrzewski, Eur. Phys. J. D **5**, 145 (1999).
19. I. S. Averbukh and N. F. Perelman, Phys. Lett. A **139**, 449 (1989).
20. P. Bellomo and C. R. Stroud, Jr., Phys. Rev. A **59**, 2139 (1999).
21. B. S. Mecking and P. Lambropoulos, Phys. Rev. Lett. **83**, 1743 (1999).
22. J.-C. Gay, D. Delande, and A. Bommier, Phys. Rev. A **39**, 6587 (1989).
23. P. A. M. Dirac, Proc. R. Soc. London, Ser. A **114**, 243 (1927).
24. R. A. Marcus, J. Chem. Phys. **54**, 3965 (1971).
25. M. V. Berry, Philos. Trans. R. Soc. London **287**, 237 (1977).
26. B. Mirbach and H. J. Korsch, J. Phys. A **27**, 6579 (1994).

Translation was provided by AIP

Journal of Experimental and Theoretical Physics, Vol. 91, No. 3, 2000, pp. 458–468.
Translated from Zhurnal Éksperimental'noĭ i Teoreticheskoĭ Fiziki, Vol. 118, No. 3, 2000, pp. 527–538.

ATOMS, SPECTRA, RADIATION

Effect of the Diamagnetic Interaction on the Intensity of the Radiation Lines of an Atom in a Magnetic Field

V. D. Ovsiannikov* and V. V. Chernushkin

Voronezh State University, Voronezh, 394693 Russia
**e-mail: vit@ovd.vsu.ru*

Received April 4, 2000

Abstract—Under the conditions of the total Paschen–Back effect the diamagnetic interaction determines the dependence of the intensity of the Zeeman components of atomic radiation lines on a magnetic field. The change in the matrix elements of the radiative transitions is due to the magnetically induced corrections to the wavefunctions of the initial and final states, whose contributions are of the same order of magnitude for the head lines of the optical series. For the high-frequency lines the positive corrections to the matrix element from the wavefunction of the upper level dominate. A magnetic field also induces dipole radiative transitions with selection rules for the orbital angular momentum $|\Delta l| \geq 3$. The matrix elements of such transitions increase rapidly with the energy of the upper level, making possible efficient single-photon population of the dipole-inaccessible Rydberg states in moderate magnetic fields. *© 2000 MAIK "Nauka/Interperiodica".*

1. INTRODUCTION

External electromagnetic fields alter the optical spectrum of matter. Consequently, it is important to know the properties of the atomic spectrum in a magnetic field not only for determining the basic laws of the matter–field interaction but also for detecting the field and measuring its intensity by optical methods. The Zeeman effect—the splitting of atomic levels in a magnetic field—is described in detail in the literature for the linear (in the magnetic field B) paramagnetic interaction with an atom; see, for example, [1–5]). The frequency, intensity, and polarization of individual Zeeman components of split lines have been calculated for the normal and anomalous Zeeman effect on the sublevels of the fine structure.

The effect of a diamagnetic interaction (quadratic in B) has been investigated in detail only for hydrogen atoms. Calculations of the energy have been performed using perturbation series [6—10] and numerical integration of the Schrödinger equation for an atom in strong and superstrong fields [4, 11–14]. Together with the first-order diamagnetic energy, the field-independent distribution of the oscillator strengths over the diamagnetic sublevels, including for transitions from highly excited Rydberg states, has also been calculated [6]. Higher order perturbation theory has been used [15] to calculate the field-dependent corrections to the matrix elements of dipole transitions. These show, specifically, that the transition matrix element between nondegenerate states increases and a field has a selective effect on the intensity of the diamagnetic components of transition lines from degenerate states [16].

Interest in the optical properties of multielectronic atoms in magnetic fields has increased appreciably in the last few years. The method for calculating diamagnetic corrections to the energy has been extended to arbitrary states of alkali atoms for degenerate diamagnetic sets [7, 17] and for isolated states taking account of corrections which are second-order in the diamagnetic interaction [18]. The diamagnetic first- and second-order susceptibilities have been calculated for the ground and metastable states of inert atoms [19], and asymptotic formulas have been written out and the parameters appearing in them have been calculated for determining the susceptibilities of Rydberg states of alkali atoms [15]. Hartree–Fock calculations of energy levels in a wide range of magnetic field intensities have been performed for helium [20–22], lithium [23], and beryllium [24].

Together with the frequency, an important characteristic of an atomic line is its intensity. Information about the dependence of the intensity of Zeeman lines on the magnetic field gives additional data on the Zeeman effect in atomic spectra and could be helpful in many problems of atomic physics and astrophysics.

The splitting of atomic lines into Zeeman components results in a redistribution of the line intensities, which in weak fields (anomalous Zeeman effect) are determined by intensity-independent matrix elements of dipole transitions between Zeeman sublevels of states with definite values of the total angular momentum of the atom $\mathbf{J} = \mathbf{L} + \mathbf{S}$ and its projection M, where $\mathbf{L}$ is the orbital angular momentum and $\mathbf{S}$ is the spin. In strong fields, for the total Paschen–Back effect, the states of the atom and the matrix elements of dipole transitions are characterized by definite values and corresponding selection rules for the orbital L and spin S angular momenta and their projections M_L and M_S sep-

1063-7761/00/9103-0458$20.00

arately. The difference of these states from the fine-structure states $|nJM\rangle$ results in a difference between the intensity distribution of the optical lines of an atom in the anomalous Zeeman effect and the corresponding distribution in the total Paschen–Back effect. The atomic spectra is quite easy to calculate in the limiting cases of weak and strong fields and such a calculation is presented in detail in the literature [1, 3–5]. However, the dependence of the intensity of radiative transitions on the field intensity is ordinarily ignored. For intermediate magnetic fields, qualitative information on that variation in the oscillator strengths of the Zeeman components of doublet lines which is caused by the atom–field interaction linear in B (paramagnetic interaction) is available [25].

In strong fields, when the paramagnetic interaction energy is greater than the spin–orbit interaction energy and the indicated rearrangement of the spectrum from $|nJM\rangle$ to $|nLM_LSM_S\rangle$ states is completed, the change in the matrix element of a radiative transition is determined by the diamagnetic (quadratic in B) interaction. The effect of a diamagnetic interaction under the conditions of the total Paschen–Back effect on the dependence of the intensity of radiative transitions of hydrogen atoms on the field intensity has been investigated in [16]. In the present paper this dependence is calculated for multielectronic atoms.

Analytical expressions for the diamagnetic corrections of order B^2 to the matrix elements of radiative transitions are presented in Section 2. The radial matrix elements determining these corrections are calculated in Section 3 using the wave functions and Green's functions of a model potential for the valence electron in alkali atoms and in helium [26, 27]. The numerical results and the possibility of observing experimentally the change in the intensity of atomic lines in a magnetic field are discussed in Section 4.

2. DIAMAGNETIC CORRECTIONS TO THE DIPOLE MATRIX ELEMENTS

Both the energy spectrum (eigenvalues) and wave functions (eigenvectors) of an atom change in a constant magnetic field. The change in the energies of the atomic levels is expressed in the shift and splitting of the emission and absorption lines, whereas the field dependence of the wave functions results in a change in the matrix element of the electromagnetic transition $d_{if} = \langle\psi_i|(\mathbf{e}\cdot\mathbf{d})|\psi_f\rangle$ ($\mathbf{e}$ is the polarization vector of the emitted or absorbed photon, $\mathbf{d}$ is the dipole moment operator of the atom), which together with the frequency $\omega_{if} = E_i - E_f$ determines the line intensity:[1]

$$I_{if} = \frac{4}{3}\alpha^3\omega_{if}^4|d_{if}|^2. \quad (1)$$

[1] We use atomic units.

Here $\alpha \approx 1/137$ is the fine-structure constant. In what follows, this constant is taken into account in the interaction with the magnetic field by choosing the unit of magnetic induction $B_0 = B_{at}/\alpha = 2.35 \times 10^5$ T, where $B_{at} = 1715$ T is an atomic unit.

For fields $B \ll B_0/\nu^3$ ($\nu = 1/\sqrt{-2E_{nl}}$ is the effective principal quantum number of the energy level E_{nl}) the magnetically induced corrections to the energy and the wave functions can be calculated using perturbation theory. For the central line with the magnetic field parallel to the polarization (π polarization) the frequency shift and the change in the matrix elements are quadratic in the field. For side lines with polarization perpendicular to the field $\mathbf{B}$ (σ polarization) the frequency shift contains a linear term, which is identical for all lines in a given series, $\Delta\omega_{if}^{(1)} \sim B$. The quadratic correction depends strongly on the principal quantum number of the upper level: $\Delta\omega_{if}^{(2)} \sim \nu^4 B^2$. As calculations for hydrogen atoms show [16], the analogous corrections to the matrix element can be even larger. In this connection, the effect of the diamagnetic interaction on the dipole matrix elements for optical lines of multielectronic atoms in the region of the total Paschen–Back effect is quite important. On the basis of the discussion in the Introduction it can be assumed that the diamagnetic corrections to the frequencies of atomic lines have been well studied, so that in what follows we shall confine our attention to the calculation of the corrections only to the matrix elements.

The operator describing the interaction of an atom with a constant magnetic field $\mathbf{B}$,

$$\hat{V}_B = \hat{V}_m + \hat{V}_D,$$

contains a linear paramagnetic term corresponding to the interaction of the magnetic dipole moment of an atom with the field (the z-axis is directed along the $\mathbf{B}$ vector),

$$\hat{V}_m = -(\mathbf{m}\cdot\mathbf{B}) = \frac{B}{2}(\hat{L}_z + 2\hat{S}_z), \quad (2)$$

and a quadratic diamagnetic term corresponding to the interaction with the field of the additional moment induced by the Larmor precession of the electronic shell of an atom,

$$\hat{V}_D = \frac{B^2}{12}r^2[1 - \mathbf{C}_{20}(\theta, \varphi)]. \quad (3)$$

Here r, θ, and φ are the radial and angular coordinates of the valence electron, $\mathbf{C}_{20}$ is the modified spherical function of the angular variable [28], and $L = l$ and $S = s$ are the orbital and spin angular momenta of the valence electron.

In the region of the total Paschen–Back effect the spin-orbit interaction can be neglected. The operator (2) can be included in the atomic Hamiltonian, leaving

the diamagnetic interaction (3) as a perturbation. Then the wave functions of the unperturbed basis will be the wave functions of the valence electron neglecting spin, and the energies will be split with respect to the magnetic quantum number: $E_{nlm} = E_{nl} + mB/2$. Since the levels of a valence electron with $l > 3$ are hydrogen-like, the radiative properties of the Zeeman states with $m > 3$ are completely analogous to those properties of the states of the hydrogen atom studied in [16]. Consequently, we shall confine our attention only to transitions between states with $m \le l \le 3$. Such states are nondegenerate, and perturbation theory for the nondegenerate states is sufficient to calculate the corrections to the wave functions and the matrix elements.

Let us assume that the nondegenerate initial and final states in Eq. (1) in the absence of a magnetic field are identical to the eigenstates of an atom $|1\rangle$ and $|0\rangle$ (with energies E_1 and E_0). Then the equations for the wave function and energy of an atom in a field can be written in an integral form, convenient for iteration, using the unperturbed Green's function $G_E(\mathbf{r}, \mathbf{r}')$ [29]:

$$\begin{aligned} E_{i(f)} &= E_{1(0)} + \langle 1(0)|\hat{V}_D|i(f)\rangle, \\ |i(f)\rangle &= |1(0)\rangle - G'_{E_{i(f)}}\hat{V}_D|i(f)\rangle. \end{aligned} \tag{4}$$

Here

$$G'_{E_{i(f)}}(\mathbf{r}, \mathbf{r}') = G_{E_{i(f)}}(\mathbf{r}, \mathbf{r}') - \frac{\langle \mathbf{r}|1(0)\rangle\langle 1(0)|\mathbf{r}'\rangle}{E_{1(0)} - E_{i(f)}} \tag{5}$$

is the reduced Green's function with the exact energy of an atom in a field.

Using the standard iterative procedure of perturbation theory to solve Eqs. (4) we obtain a series expansion in powers of the small parameter $B^2 \ll 1$ for the wave functions and energies, whence power series can be obtained for the transition frequency and matrix element:

$$\begin{aligned} \omega_{if}(B) &= \omega_{10}\left(1 + \sum_{s=1}^{\infty} w_{10}^{(s)} B^{2s}\right), \\ d_{if}(B) &= d_{10}\left(1 + \sum_{s=1}^{\infty} q_{10}^{(s)} B^{2s}\right). \end{aligned} \tag{6}$$

The coefficient $w_{10}^{(s)}$ in the expansion for the frequency is the ratio of the difference of the diamagnetic susceptibilities $\chi_0^{(s)} - \chi_1^{(s)}$, determining the corrections $\Delta E_{f(i)}^{(s)} = -\chi_{0(1)}^{(s)} B^{2s}/2s!$ [18] of order s to the energies of the lower and upper levels, to the unperturbed frequency ω_{10}:

$$w_{10}^{(s)} = \frac{\chi_0^{(s)} - \chi_1^{(s)}}{2s!\,\omega_{10}}. \tag{7}$$

The first-order susceptibility $\chi_{nlm}^{(1)}$ can be determined from the relation for the correction to the energy

$$\Delta E^{(1)} = \langle nlm|\hat{V}_D|nlm\rangle = -\frac{B^2}{2}\chi_{nlm}^{(1)}. \tag{8}$$

Substituting here the operator (3) and integrating with respect to the angular variables gives

$$\chi_{nlm}^{(1)} = -\frac{l(l+1) + m^2 - 1}{2(2l-1)(2l+3)}\langle nl|r^2|nl\rangle. \tag{9}$$

The mean-square radius of the electron orbit in this expression and therefore the factor $w^{(1)}$ are rapidly increasing functions of the principal quantum number of the level $|nl\rangle$.

The decomposition of the tensor $\chi^{(2)}$ into irreducible parts for nondegenerate levels of multielectronic atoms is given in [18], where the numerical values of the irreducible components of this tensor are also calculated for the s, p, and d states of alkali elements. A method for calculating $\chi^{(2)}$ in highly excited Rydberg states is proposed in [15]. Numerical calculations of the susceptibilities $\chi^{(s)}$ of very high orders (up to $s = 75$) for the hydrogen levels with principal quantum numbers $n \le 3$ were performed in [8]. Analytic expressions and the general properties of the third-order susceptibilities in arbitrary states of hydrogen were obtained in [9, 10]. Besides perturbation theory, other methods making it possible to determine the change in the radiation frequencies in fields stronger than the upper limit of applicability of perturbation theory have also been used to calculate the energies of atomic levels in a magnetic field (see, for example, [11–14, 20–24]). Thus, there is now available a variety of analytical and numerical methods for calculating the change in the frequencies of radiative transitions of atoms in a magnetic field and therefore for determining the corrections (7).

The change of the matrix element $d_{if}(B)$ in a magnetic field can be calculated on the basis of the integral equations (4). Expanding the wave functions of the initial and final states in powers of the operator (3), the coefficients $q_{10}^{(s)}$ in Eq. (6) can be represented as a ratio of the matrix element of order s to the unperturbed value. Specifically, for $s = 1$ (in what follows we confine our attention to the first-order correction $q_{10}^{(1)} \equiv q_{10}$, omitting the order index s) we have

$$q_{10} = q_{10}(1) + q_{10}(0), \tag{10}$$

where

$$q_{10}(1) = -\frac{\langle 1|\hat{V}_D G'_{E_1}\hat{d}_\mu|0\rangle}{B^2 d_{10}} \tag{11}$$

determines the contribution due to the field-induced change in the wave function of the upper level $|1\rangle$; the term

$$q_{10}(0) = -\frac{\langle 1|\hat{d}_\mu G'_{E_0}\hat{V}_D|0\rangle}{B^2 d_{10}}. \quad (12)$$

gives the contribution associated with the change in the wave function of the lower level $|0\rangle$; $\hat{d}_\mu$ is the operator of the μ projection of the electric dipole moment ($\mu = 0, \pm 1$).

After integrating with respect to the angular variables of the valence electron, the quantities (11) and (12) can be represented as a ratio of a linear combination of the second-order radial matrix elements of the electric dipole and diamagnetic interaction operators (r and r^2) to the first-order matrix element of the operator r. In the general case, we obtain for the transition $|1\rangle \equiv |nlm\rangle \longrightarrow |n'l'm'\rangle \equiv |0\rangle$ ($l' = l \pm 1$, $m' = m - \mu$)

$$q^\mu_{10}(nlm) = \sum_{p=0,\pm 2} A^\mu_p(l, m; l')a_p(nl; n'l'),$$
$$q^\mu_{10}(n'l'm') = \sum_{p=0,\pm 2} A^{-\mu}_p(l', m'; l)a_p(n'l'; nl), \quad (13)$$

where $A^\mu_p(l, m; l')$ is the ratio of the angular integrals

$$A^\mu_0(l, m; l') = -\frac{l(l+1)+m^2-1}{4(2l-1)(2l+3)},$$
$$A^\mu_{\pm 2}(l, m; l \mp 1) = 0,$$
$$A^0_2(l, m; l+1) = \frac{(l+2)^2-m^2}{8(2l+3)(2l+5)},$$
$$A^0_{-2}(l, m; l-1) = \frac{(l-1)^2-m^2}{8(2l-1)(2l-3)}, \quad (14)$$
$$A^1_2(l, m; l+1) = A^{-1}_2(l, -m; l+1)$$
$$= -\frac{(l+m+1)(l+m+2)}{8(2l+3)(2l+5)},$$
$$A^{-1}_{-2}(l, m; l-1) = A^1_{-2}(l, -m; l-1)$$
$$= -\frac{(l+m-1)(l+m)}{8(2l-1)(2l-3)};$$

$$a_p(nl; n'l') = \frac{\langle nl|r^2 g^{(nl)}_{l+p} r|n'l'\rangle}{\langle nl|r|n'l'\rangle} \quad (15)$$

is the ratio of the radial matrix elements. In these expressions $g^{(nl)}_{l_1}(r, r')$ is the reduced radial Green's function with energy E_{nl} in the subspace of the states of a valence electron with orbital angular momentum l_1 [27].

3. CALCULATION OF THE MATRIX ELEMENT BY THE MODEL POTENTIAL METHOD

For transitions between states of a valence electron all radial matrix elements in Eqs. (11)–(13) and (15) can be calculated analytically and presented as functions of the principal and orbital quantum numbers using the model-potential method [26, 27, 29]. For first-order matrix elements these expressions can be represented as a generalized hypergeometric function of two variables after calculating the radial integrals using the explicit expression for the radial wave function:

$$R_{nl}(r) = \frac{2}{\nu_{nl}^2}\sqrt{\frac{n_r! Z^3}{\Gamma(n_r + 2\lambda + 2)}} f_{n_r\lambda}\left(\frac{2Zr}{\nu_{nl}}\right). \quad (16)$$

Here Z is the charge of the core ion in whose field the valence electron moves; $\nu_{nl} = Z/\sqrt{-2E_{nl}}$ is the effective principal quantum number, $n_r = 0, 1, 2, \ldots$ is the radial quantum number, $\lambda = \nu_{nl} - n_r - 1$ is the effective orbital quantum number of the level;

$$f_{k\lambda}(x) = e^{-x/2}x^\lambda L_k^{2\lambda+1}(x) \quad (17)$$

is the Sturm function of the Schrödinger equation with the model Hamiltonian. Just like the Coulomb function, this function is expressed in terms of the Laguerre polynomials $L^\alpha_k(x)$, which are related with the confluent hypergeometric function as [30]

$$L^\alpha_k(x) = \frac{(\alpha+1)_n}{n!}{}_1F_1(-k; \alpha+1; x).$$

Here $(\alpha+1)_n = (\alpha+1)(\alpha+2)\ldots(\alpha+n)$ is the Pochhammer symbol.

These relations make it possible to express the radial matrix element in terms of the generalized hypergeometric function F_2 of two variables:

$$\langle nl|r|n'l'\rangle = \frac{\Gamma(a)}{4Z}\sqrt{n_r! n'_r! \frac{(2\lambda+2)_{n_r}(2\lambda'+2)_{n'_r}}{\Gamma(b)\Gamma(b')}} \times x^{\lambda+2}x'^{\lambda'+2}F_2(a; -n_r, -n'_r; b, b'; x, x'). \quad (18)$$

Here

$$a = \lambda + \lambda' + 4, \quad b = 2\lambda + 2, \quad b' = 2\lambda' + 2,$$
$$x = \frac{2\nu'}{\nu+\nu'}, \quad x' = \frac{2\nu}{\nu+\nu'}.$$

For negative integer values of the upper parameters, the double series representing the function F_2 reduces to a double sum—a polynomial in powers of the arguments x and x'. In practical calculations the radial quantum number of the lower level is usually small, so that it is

convenient to represent the series F_2 as a linear combination of $n_r' + 1$ hypergeometric functions:

$$F_2(a; -k, -n_r'; b, b'; x, x') = \sum_{s=0}^{n_r'} \frac{(a)_s(-n_r')_s}{s!(b')_s}(x')^s {}_2F_1(a+s, -k; b; x).$$

To calculate this sum, it is convenient to use recurrence relations between the functions ${}_2F_1$ with adjacent parameters [30]. The formula (18) can be regarded as a generalization of the Gordon formula (see, for example, [2]) for the radial matrix element of a radiative transition of an arbitrary atom, the states of whose valence electron are described by a model potential. Using the Sturm expansion of the Green's function of the model potential [27, 29],

$$g_l^{(nl)}(r, r') = \frac{4Z}{\nu_{nl}}\left\{\sum_{k\neq n_r}^{\infty}\frac{k!}{\Gamma(k+2\lambda+2)} \times \frac{f_{k\lambda}\left(\frac{2Zr}{\nu_{nl}}\right)f_{k\lambda}\left(\frac{2Zr'}{\nu_{nl}}\right)}{k+\lambda+1-\nu_{nl}} + \frac{n_r!}{\Gamma(\nu_{nl}+\lambda+1)\nu_{nl}} \times\left(\frac{5}{2}+r\frac{d}{dr}+r'\frac{d}{dr'}\right)f_{n_r\lambda}\left(\frac{2Zr}{\nu_{nl}}\right)f_{n_r\lambda}\left(\frac{2Zr'}{\nu_{nl}}\right)\right\}, \quad (19)$$

the second-order matrix element can be obtained by analogy with the radial matrix element (18) as a generalized hypergeometric series, each term of which can be expressed in terms of the function F_2. One of the negative integer parameters of this function is identical to the radial quantum number of the initial or final state, and the other is identical to the summation index of the series (19). Thus, we obtain for the ratios (15) of the second- and first-order radial matrix elements

$$a_0(nl; n'l') = \frac{\nu^4(b)_3}{8Z^4}\sum_{k=0}^{\infty}\frac{(b)_k}{k!} \times F_2(b+3; -n_r, -k; b, b; 1, 1) \times \Phi(a; k, n_r, n_r'; b, b'; x, x'), \quad (20)$$

where the cofactor Φ under the summation sign is the ratio of generalized hypergeometric functions F_2 and for $k \neq n_r$ has the form

$$\Phi(a; k, n_r, n_r'; b, b'; x, x') = \frac{F_2(a; -k, -n_r'; b, b'; x, x')}{(k-n_r)F_2(a; -n_r, -n_r'; b, b'; x, x')}. \quad (21)$$

For $k = n_r$ this cofactor is

$$\Phi(a; n_r, n_r, n_r'; b, b'; x, x') = \frac{1}{\nu}\left[\frac{(2\nu-n_r)F_2(a; -n_r-1, -n_r'; b, b'; x, x')}{2F_2(a; -n_r, -n_r'; b, b'; x, x')} - 1 - \frac{n_rF_2(a; -n_r+1, -n_r'; b, b'; x, x')}{2F_2(a; -n_r, -n_r'; b, b'; x, x')}\right]. \quad (22)$$

The generalized hypergeometric function F_2 with unit arguments in Eq. (20) is the orthogonality integral for the Laguerre polynomials, which is different from zero only for $|k - n_r| \leq 3$. Consequently, the series in k here contains no more than seven nonzero terms.

The effective orbital quantum numbers λ and λ_1 for states of the valence electron with angular momenta l and $l \pm 2$ virtually always assume values such that the difference $\nu_{nl} - \lambda_1$ is not a positive integer. In this case the Green's function in the second-order matrix element in the expression (15) for $a_{\pm 2}$ may not be reduced. The nonreduced Green's function is the first term in Eq. (19), where λ is replaced by λ_1, and the summation extends over all nonnegative integer values of k, including $k = n_r$. After integrating with respect to the radial coordinate of the valence electron Eq. (15) assumes the form

$$a_{\pm 2}(nl; n'l') = \frac{\nu^4\Gamma(\alpha)\Gamma(\alpha')}{8Z^4\Gamma(a)\Gamma(b_1)}x^{\lambda_1-\lambda} \times \sum_{k=0}^{\infty}\frac{(b_1)_k}{k!(k+\lambda_1+1-\nu)}F_2(\alpha; -n_r, -k; b, b_1; 1, 1) \times \frac{F_2(\alpha'; -k, -n_r'; b_1, b'; x, x')}{F_2(a; -n_r, -n_r'; b, b'; x, x')}, \quad (23)$$

where $\alpha = \lambda + \lambda_1 + 5$, $\alpha' = \lambda_1 + \lambda' + 4$, and $b_1 = 2\lambda_1 + 2$. The generalized hypergeometric function F_2 appearing here is not an orthogonality integral, since the difference between the upper and lower parameters is not an integer. Nonetheless, the Sturm series remains rapidly converging, and $n_r + 2$ or $n_r + 3$ terms, where n_r is the radial quantum number of the upper level, are sufficient to calculate the series with relative accuracy 10^{-4}.

4. RESULTS OF NUMERICAL CALCULATIONS AND DISCUSSION

4.1. Variation in the Matrix Elements of Dipole-Allowed Transitions

We performed calculations of the numerical values of the factor q, determining the first diamagnetic correction to the matrix elements of dipole transitions between the ground and first excited states of alkali atoms, corresponding to the most important series of lines, and the corrections to the matrix elements of singlet and triplet lines of helium.

Table 1. Diamagnetic susceptibilities q^{π} and q^{σ} for corrections to the matrix elements $d_{n_1P,\,nS}$ of radiative π and σ transitions $n_1P \longrightarrow nS$ corresponding to the Zeeman components of the lines in the principal series of alkali atoms: $(k) \equiv 10^k$

Atom	$n_r(n_1P)$	d_{10}	$q^{\pi}(1)$	$q^{\pi}(0)$	q^{π}	$q^{\sigma}(0)$	q^{σ}
Li	0	–2.40	–1.34	–1.20	–2.54	–11.2	–13.8
	1	1.29(–1)	5.56(2)	–2.02(2)	3.54(2)	–3.90(2)	7.21(2)
	2	1.14(–1)	1.16(3)	–6.94(1)	1.09(3)	–1.40(2)	2.19(3)
	3	8.51(–2)	3.87(3)	–4.62(1)	3.82(3)	–9.65(1)	7.64(3)
	4	6.57(–2)	1.13(4)	–3.72(1)	1.12(4)	–7.95(1)	2.24(4)
	5	5.23(–2)	2.83(4)	–3.27(1)	2.83(4)	–7.11(1)	5.66(4)
	6	4.32(–2)	6.31(4)	–2.96(1)	6.31(4)	–6.54(1)	1.26(5)
Na	0	–2.54	0.925	–2.31	–1.36	–1.51(1)	–1.32(1)
	1	–1.43(–1)	–4.36(2)	2.00(2)	–2.36(2)	3.70(2)	–5.02(2)
	2	–2.72(–2)	–2.24(3)	4.08(2)	–1.84(3)	7.67(2)	–3.72(3)
	3	–5.24(–3)	–1.85(4)	1.19(3)	–1.73(4)	2.27(3)	–3.47(4)
	4	8.74(–4)	1.99(5)	–4.78(3)	1.94(5)	–9.15(3)	3.88(5)
	5	2.83(–3)	1.09(5)	–1.08(3)	1.08(5)	–2.08(3)	2.17(5)
	6	3.38(–3)	1.57(5)	–7.07(2)	1.56(5)	–1.36(3)	3.12(5)
K	0	–2.95	–0.612(5)	–0.192(3)	–0.804	–2.31(1)	–2.43(1)
	1	–7.42(–2)	–1.77(3)	7.05(2)	–1.06(3)	1.38(3)	–2.16(3)
	2	1.88(–2)	8.59(3)	–9.87(2)	7.60(3)	–2.02(3)	1.52(4)
	3	2.75(–2)	1.11(4)	–3.61(2)	1.07(4)	–7.64(2)	2.14(4)
	4	2.58(–2)	2.31(4)	–2.49(2)	2.28(4)	–5.40(2)	4.56(4)
	5	2.25(–2)	4.88(4)	–2.06(2)	4.86(4)	–4.54(2)	9.72(4)
	6	1.94(–2)	9.79(4)	–1.83(2)	9.78(4)	–4.07(2)	1.96(5)
Rb	0	–3.06	–2.85(–2)	0.815	0.786	–2.68(1)	–2.69(1)
	1	–1.32(–1)	–1.09(3)	4.47(2)	–6.42(2)	8.72(2)	–1.31(3)
	2	–1.05(–2)	–1.49(4)	2.07(3)	–1.29(4)	4.23(3)	–2.56(4)
	3	8.95(–3)	2.94(4)	–1.32(3)	2.81(4)	–2.78(3)	5.61(4)
	4	1.26(–2)	3.81(4)	–6.11(3)	3.74(4)	–1.32(3)	7.48(4)
	5	1.26(–2)	6.78(4)	–4.44(2)	6.74(4)	–9.71(2)	1.35(5)
	6	1.17(–2)	1.24(5)	–3.69(2)	1.24(5)	–8.16(2)	2.48(5)
Cs	0	–3.25	–0.417	5.84	5.43	–3.44(1)	–3.52(1)
	1	–1.38(–1)	–1.35(3)	5.41(2)	–8.10(2)	1.08(3)	–1.63(3)
	2	–9.85(–3)	–2.04(4)	2.75(3)	–1.77(4)	5.81(3)	–3.50(4)
	3	1.06(–2)	3.13(4)	–1.38(3)	2.99(4)	–3.03(3)	5.95(4)
	4	1.43(–2)	4.15(4)	–6.70(2)	4.09(4)	–1.50(3)	8.16(4)
	5	1.40(–2)	7.37(4)	–4.93(2)	7.32(4)	–1.12(3)	1.46(5)
	6	1.27(–2)	1.35(5)	–4.19(2)	1.35(5)	–9.67(2)	2.69(5)

Table 1 shows the values obtained for the coefficients (10) and the individual contributions to them from field-induced changes of the wave functions of the upper level (11) and the lower level (12) for π ($n_1P0 \longrightarrow nS0$)and σ ($n_1P1 \longrightarrow nS0$) transitions in the principal series of the alkali atoms. The numerical values of the matrix element d_{10} of the dipole transition of an unperturbed atom, which in this case does not depend on the type of transition, are also presented. The absolute value of the factor for a σ transition is approximately twice the analogous value for the π transition in accordance with the relation for the corrections (11) $q^{\sigma}(1) = 2q^{\pi}(1)$, making the main contribution to Eq. (10) (the table gives only the value for $q^{\pi}(1)$). The factors q are negative for the head lines of all atoms and from one to three subsequent lines (for lithium, only for

the head line). For the high-frequency lines the factors q are positive and increase rapidly with the principal quantum number of the upper level; this is similar to the behavior of q for the lines of the Lyman and Balmer series of the hydrogen atom [16]. Thus, the intensity of the Zeeman components of the head lines of the principal series of alkali atoms decreases as the magnetic field intensity increases, while the intensity of the high-frequency lines in this series increases. For transitions from high states, the diamagnetic correction to the wave function of the upper level makes the main contribution to q—this is evident from Table 1: for transitions in the principal series from states with $n_r > 3$ the relation $|q(1)| \gg |q(0)|$ holds. The correction $q(1)$ is positive and increases rapidly with n_r, while $q(0)$ is negative and decreases smoothly in absolute magnitude, so that $q^\sigma \approx q^\sigma(1) = 2q^\pi(1) \approx 2q^\pi$, which is similar to the behavior of these quantities for hydrogen lines [16].

We note that the numerical values of the correction factors q and their rates of increase with increasing effective principal quantum number ν_n of the upper level $|nP\rangle$,

$$q = a_6\nu_n^6, \tag{24}$$

are much greater than the corresponding values for the diamagnetic susceptibilities (9):

$$\chi_n = a_4\nu_n^4.$$

This means that the change in the matrix elements makes the main contribution to the change in the intensity (1) for the high-frequency lines of a specific series of the single-electron spectrum of an atom, and the effect on the intensity is itself just as important magnetically induced optical effect as the action of the field on the frequencies of the same lines. The calculation shows that the numerical value of the coefficient a_6 for the main series is essentially identical for all alkali atoms: $a_6^\sigma = 2a_6^\pi \approx 0.53$—for potassium, rubidium, and cesium; $a_6^\sigma = 2a_6^\pi \approx 0.5$ for lithium; and, $a_6^\sigma = 2a_6^\pi \approx 0.57$ for sodium.

For the sharp series (the transitions $n_1S \longrightarrow nPm$) the values of the factors q for π and σ transitions in most cases have opposite signs because the signs of the corresponding dipole matrix elements of the unperturbed atom are opposite. The same relation between the signs of q^π and q^σ is also observed for other atomic series. The signs and absolute values of the coefficients q depend on the specific atom and on the specific transition.

As an illustration, Table 2 gives the numerical values of the susceptibilities q determining the change in the matrix elements of the diffuse series of alkali atoms, which correspond to the transitions $n_1Dm_1 \longrightarrow nPm$. The axial symmetry of the system ensures that $q(-m_1 \longrightarrow -m) = q(m_1 \longrightarrow m)$ with a simultaneous replacement of photons with right-hand circular polarization, which are emitted in transitions with $m_1 - m = \pm 1$, by photons with left-hand polarization (and vice versa). For this reason, transitions only between states with positive magnetic quantum numbers are indicated in Table 2. The ratio of the dipole matrix elements $d(m_1 \longrightarrow m)$ of an unperturbed atom, which correspond to different values of m_1 and m, is determined only by the angular factors, so that for transitions in the diffuse series $d(0 \longrightarrow 1) : d(1 \longrightarrow 0) : d(1 \longrightarrow 1) : d(0 \longrightarrow 0) : d(2 \longrightarrow 1) = 1 : \sqrt{3} : \sqrt{3} : 2 : \sqrt{6}$. The table gives data only for the smallest matrix element, $d_{n_1D,\,nP}(0 \longrightarrow 1)$. Here, just as for the principal series, the asymptotic formula (24) can be written down for each coefficient $q(m_1 \longrightarrow m)$. The ratio between the asymptotic values of the coefficients corresponding to transitions from states with $m_1 \neq 0$ is essentially identical to the ratios between the unperturbed matrix elements: $a_6(1 \longrightarrow 0) : a_6(1 \longrightarrow 1) : a_6(2 \longrightarrow 1) \approx 1 : 1 : \sqrt{2}$, where $a_6(1 \longrightarrow m) \approx 0.35$ for all alkali atoms. For transitions from states with $m_1 = 0$ the ratio of the coefficients is close in absolute value to the ratio of the unperturbed matrix elements and negative: $a_6(0 \longrightarrow 1) : a_6(0 \longrightarrow 0) \approx -2$, which is due to the large negative contribution of intermediate S states to q. This contribution is related with the strong perturbation of the D levels with zero magnetic quantum number by the close-lying S levels, which are absent in states with $m_1 \neq 0$. Consequently, the coefficients $a_6(0 \longrightarrow m)$ are much greater in absolute value than $a_6(1, 2 \longrightarrow m)$. The numerical values of these quantities are presented in Table 3.

Our computed diamagnetic corrections to the intensity of the singlet and triplet lines in helium atoms, corresponding to radiative transitions into the ground state $1s^2(^1S_0)$ and into the metastable $1s2s(^1S_0)$ and $1s2s(^3S_1)$ states, decrease the matrix element of the radiative transition for the head lines of series ending on metastable levels and increase it for all other lines of the series and the lines of the series of transitions into the ground state of the atom; this agrees qualitatively with Hartree–Fock calculations of the radiative matrix elements [20]. Just as in alkali atoms, the field-induced change of the upper level, which increases rapidly with the principal quantum number, makes the main contribution to the effect. The corrections to the matrix elements of σ transitions are approximately two times greater than the corrections to the matrix elements of π transitions in proportion to the ratio of the angular matrix elements, corresponding to $\mu = 1$ and $\mu = 0$ in Eq. (11). The asymptotic values of q are determined by the Eq. (24) with the coefficients $a_6^\sigma = 2a_6^\pi \approx 0.5$, which are equal to the analogous coefficients for the corrections to the matrix elements of the principal series of alkali atoms (see above).

The experimental observation of the dependence of atomic line intensities on the magnetic field seems to be

Table 2. Matrix element $d_{n_1D,\,nP}(0 \longrightarrow 1)$ and diamagnetic susceptibilities $q(m_1 \longrightarrow m)$ for corrections to the matrix elements of radiative transitions in the diffuse series of alkali atoms $n_1Dm_1 \longrightarrow nPm$; $(k) \equiv 10^k$

Atom	$n_r(n_1D)$	$d(0 \longrightarrow 1)$	$q(0 \longrightarrow 1)$	$q(1 \longrightarrow 0)$	$q(1 \longrightarrow 1)$	$q(0 \longrightarrow 0)$	$q(2 \longrightarrow 1)$
Li	0	1.16	5.16(2)	7.34(1)	7.48(1)	–1.65(2)	1.11(2)
	1	4.38(–1)	4.31(3)	6.63(2)	6.76(2)	–1.32(3)	1.00(3)
	2	2.54(–1)	2.25(4)	3.54(3)	3.55(3)	–6.80(3)	5.31(3)
	3	1.74(–1)	8.37(4)	1.24(4)	1.24(4)	–2.64(4)	1.86(4)
	4	1.30(–1)	2.51(5)	3.40(4)	3.40(4)	–8.28(4)	5.10(4)
	5	1.02(–1)	6.44(5)	7.98(4)	7.98(4)	–2.22(5)	1.20(5)
	6	8.31(–2)	1.47(6)	1.67(5)	1.67(5)	–5.27(5)	2.51(5)
Na	0	1.44	4.69(2)	2.81(1)	3.56(1)	–1.93(2)	4.60(1)
	1	4.24(–1)	4.27(3)	4.17(2)	4.46(2)	–1.59(3)	6.40(2)
	2	2.27(–1)	2.40(4)	2.91(3)	2.94(3)	–8.33(3)	4.37(3)
	3	1.49(–1)	9.27(4)	1.10(4)	1.11(4)	–3.26(4)	1.66(4)
	4	1.09(–1)	2.84(5)	3.15(4)	3.15(4)	–1.03(5)	4.72(4)
	5	8.45(–2)	7.37(5)	7.55(4)	7.55(4)	–2.74(5)	1.13(5)
	6	6.83(–2)	1.70(6)	1.61(5)	1.61(5)	–6.50(5)	2.41(5)
K	0	1.69	1.57(3)	–3.24(1)	–1.58(1)	–8.11(2)	–4.04(1)
	1	2.06(–1)	–4.84(5)	–2.91(2)	–2.02(2)	2.42(5)	–3.92(2)
	2	6.53(–2)	–4.94(5)	1.73(2)	2.91(2)	2.47(5)	3.19(2)
	3	2.94(–2)	–1.91(6)	3.02(3)	3.17(3)	9.56(5)	4.61(3)
	4	1.62(–2)	–6.71(6)	1.22(4)	1.24(4)	3.37(6)	1.84(4)
	5	1.01(–2)	–2.03(7)	3.49(4)	3.51(4)	1.02(7)	5.25(4)
	6	6.74(–3)	–5.38(7)	8.27(4)	8.29(4)	2.70(7)	1.24(5)
Rb	0	1.79	–3.28(3)	–5.65(1)	–3.37(1)	1.59(3)	–7.33(1)
	1	3.82(–2)	–9.34(4)	–3.25(3)	–2.75(3)	4.30(4)	–4.63(3)
	2	–2.27(–2)	3.48(5)	7.47(3)	7.23(3)	–1.65(5)	1.11(4)
	3	–2.56(–2)	7.21(5)	1.38(4)	1.38(4)	–3.43(5)	2.07(4)
	4	–2.26(–2)	1.78(6)	3.24(4)	3.23(4)	–8.51(5)	4.86(4)
	5	–1.93(–2)	4.21(6)	7.22(4)	7.21(4)	–2.01(6)	1.08(5)
	6	–1.65(–2)	9.23(6)	1.49(5)	1.49(5)	–4.43(6)	2.23(5)
Cs	0	1.85	–3.57(2)	–7.64(1)	–4.37(1)	1.07(2)	–9.82(1)
	1	–3.13(–1)	5.47(3)	4.24(2)	3.87(2)	–2.23(3)	6.17(2)
	2	–1.60(–1)	2.66(4)	1.87(3)	1.88(3)	–1.09(4)	2.81(3)
	3	–1.03(–1)	1.04(5)	7.34(3)	7.36(3)	–4.27(4)	1.10(4)
	4	–7.40(–2)	3.27(5)	2.21(4)	2.21(4)	–1.36(5)	3.31(4)
	5	–5.69(–2)	8.79(5)	5.51(4)	5.52(4)	–3.71(5)	8.27(4)
	6	–4.57(–2)	2.09(6)	1.21(5)	1.21(5)	–8.94(5)	1.82(5)

distinctly accessible for transitions between the Zeeman components of Rydberg states, similarly to the experimental study of the field dependence presented in [31] for Stark lines. The rapid increase in the corrections with increasing principal quantum number n of the upper level shows that for $n = 20$ the diamagnetic corrections to the intensity are 20–30% in the field $B = 0.5$ T.

4.2. Diamagnetically Induced Dipole Transitions

Together with the change in the matrix elements and intensities of dipole-allowed transitions, a diamagnetic interaction induces dipole-forbidden transitions. Just as the allowed transitions, these transitions occur between states with opposite parity, but in addition to the dipole selection rules the rule $\Delta l = \pm 3$ already appears in first order in the diamagnetic interaction, so that, specifically, a transition from the n_3F state into the nS state

Table 3. Numerical values of the factor $a_6(0 \longrightarrow m)$, determining the asymptotic susceptibilities (24) for corrections to the matrix elements of radiative transitions in the diffuse series of alkali atoms $n_1D0 \longrightarrow nPm$

Atom	$a_6(0 \longrightarrow 0)$	$a_6(0 \longrightarrow 1)$
Li	4.4	–1.8
Na	6.0	–2.5
K	–300	150
Rb	37	18
Cs	9.6	–4.4

Table 4. Susceptibility η for the relative intensity of magnetically induced ($n_3F \longrightarrow nS$) and allowed ($nP \longrightarrow nS$) radiative π transitions into the ground state of alkali atoms and into the ground and metastable states of the helium atom; $(k) \equiv 10^k$

Atom	$4F \longrightarrow nS$	$5F \longrightarrow nS$	$6F \longrightarrow nS$	$7F \longrightarrow nS$	$8F \longrightarrow nS$	$9F \longrightarrow nS$	a_{12}
Li 2*S*	2.30(7)	5.24(8)	5.57(9)	3.67(10)	1.80(11)	6.86(11)	27
Na 3*S*	7.40(4)	1.17(5)	7.60(5)	6.71(6)	6.82(7)	3.71(8)	3.4(–2)
K 4*S*	1.13(4)	1.13(6)	1.77(7)	1.41(8)	7.54(8)	3.17(9)	1.6(–1)
Rb 5*S*	4.00(3)	6.31(4)	2.58(6)	2.77(7)	1.73(8)	8.00(8)	5.2(–2)
Cs 6*S*	5.98(3)	5.16(4)	2.48(6)	2.72(7)	1.70(8)	7.65(8)	5.3(–2)
He 1*S*	9.64(8)	2.38(10)	3.92(11)	9.64(12)	4.48(14)	3.69(13)	3.3(2)
He 2^1S	1.27(11)	3.60(12)	6.34(13)	1.62(15)	7.69(16)	6.43(15)	6.0(4)
He 2^3S	2.12(8)	4.55(9)	4.26(10)	2.52(11)	1.09(12)	3.75(12)	1.2(2)

becomes possible. The intensity of the corresponding lines is proportional to the fourth power of the field B and is determined by the matrix element of the induced transitions, which (without the factor B^2) can be represented as

$$d^{\pi}_{n_3 n} = A_1(lm)(R_1 + R_2) \tag{25}$$

for π transitions and

$$d^{\sigma}_{n_3 n} = A_2(lm)R_1 + A_3(lm)R_2 \tag{26}$$

for σ transitions, where

$$R_1 = \left\langle n_3 l + 3 \middle| r^2 g^{(n_3)}_{l+1} r \middle| nl \right\rangle$$

is the radial matrix element of the transition induced by a magnetic field acting on the initial state $|n_3 l + 3\rangle$,

$$R_2 = \left\langle n_3 l + 3 \middle| r g^{(n)}_{l+2} r^2 \middle| nl \right\rangle$$

is the matrix element of a transition induced by the magnetic field acting on the final state. The dependence on the orbital and magnetic quantum numbers for the angular factors can be written in an explicit form similarly to the expressions (14):

$$A_1(l, m) = \frac{1}{8}\sqrt{\frac{[(l+1)^2 - m^2][(l+2)^2 - m^2][(l+3)^2 - m^2]}{(2l+1)(2l+3)^2(2l+5)^2(2l+7)}},$$
$$A_2(l, m) = A_1(l, m)\sqrt{\frac{(l+1+m)(l+2+m)}{2[(l+1)^2 - m^2]}}, \tag{27}$$
$$A_3(l, m) = A_1(l, m)\sqrt{\frac{(l+3+m)(l+4+m)}{2[(l+3)^2 - m^2]}}.$$

It is convenient to express the intensity of a magnetically induced transition in terms of the intensity of the dipole-allowed transition:

$$I_{n_3 l+3 \to nl} = \eta I_{n_1 l+1 \to nl} B^4. \tag{28}$$

The numerical parameter η is determined by the ratio of the frequencies and matrix elements of the induced and allowed transitions:

$$\eta = \left(\frac{\omega_{n_3 n}}{\omega_{n_1 n}}\right)^4 \left(\frac{d_{n_3 n}}{d_{n_1 n}}\right)^2. \tag{29}$$

For transitions into the ground state of an alkali atom, the intensity of the head line of the principal radiation series of the atom can be chosen as the normalizing intensity.

The explicit expressions for the ratios of the radial matrix elements R_1 and R_2 to $\langle n_1 l_1 | r | nl \rangle$ differ from Eq. (23) only by the parameters of the initial and intermediate states, so that the calculations of these quantities are fundamentally identical to the calculations of the corrections to the matrix elements of allowed transitions, described in Section 3.

Table 4 gives the numerical values of the parameter η for the intensity of the transition $n_3F \longrightarrow nS$, expressed in terms of the intensity of the transition $nP \longrightarrow nS$ (n is the principal quantum number of the ground state for an alkali atom, $n = 2$ for transitions into metastable states of helium). For transitions into the ground state of helium $1sn_3f(^1F_3) \longrightarrow 1s^2(^1S_0)$ the parameter η expresses the intensity in terms of $I_{2P \to 1S}$. Since the angular factors $A_1(l, m)$ and $A_2(l, m)$, determining the matrix elements of the induced π and σ transitions (24) and (25), are identical in the present case and the matrix element R_2 is small compared with R_1, the numerical values of η for π and σ transitions are essentially identical. Consequently, data only for π transitions are presented in the table.

The computational results presented in Table 4 demonstrate a rapid increase in the matrix elements of induced transitions n_3F–nS with increasing principal quantum number n_3 of the upper level. This shows, specifically, that one-photon excitation of the F states of atoms in a magnetic field by resonance radiation is possible. The probability of excitation of the Rydberg F levels with large principal quantum numbers ($n_3 > 10$) with the same intensity of the exciting light pulses is

much higher than the probability of excitation of the first F states with $n_3 = 4, 5$. In all atoms the parameter η increases monotonically with the principal quantum number of the F level, with the exception of the singlet states of helium, where this parameter decreases somewhat for transitions from $9F$ and $10F$ states because of the decrease of the third-order matrix element R_1. Subsequently (for $n_3 > 10$), the increase in η reappears.

A formula, similar to Eq. (24), for the asymptotic dependence on the principal quantum number of the upper level can be written down for the coefficients η:

$$\eta = a_{12}\nu^{12}. \quad (30)$$

The numerical values of the coefficients a_{12} are presented in the last column in Table 4. As one can see from the table, the largest values of the magnetically induced matrix element correspond to a transition into a metastable singlet state 2^1S_0.

It follows from the data obtained, specifically, that the cross section of the transition $2^1S_0 \longrightarrow 8^1F_3$ in the field $B = 10$ T is approximately one-fourth the cross section for the strongest resonance line of the transition $2^1S_0 \longrightarrow 2^1P_1$. In such a field the absolute values of the diamagnetic shift and the paramagnetic Zeeman splitting of the induced line are approximately identical and equal about 5 cm^{-1} [18, 19].

5. CONCLUSIONS

The changes due to the diamagnetic interaction and examined here in the intensity of the radiation lines of atoms give new information about the interaction of multielectronic atoms with a magnetic field under conditions of the total Paschen–Back effect, where the spin-orbit splitting is small compared with the magnetic splitting. The computed values of the diamagnetic susceptibilities, determining the corrections to the matrix elements of dipole-allowed transitions, show that the effect of a magnetic field on the wave functions of atomic levels between which a transition occurs can be more important than the change in their energies. The effect of a magnetic field on the intensity of optical lines is found to be selective, similarly to the effect on the lines in the spectrum of hydrogen-like atoms [16]. The selectivity of the diamagnetic effects is manifested not only in the difference in the absolute values and signs for the corrections to the intensity of the lines of different atoms and the lines in different series of the same atom but also in the differences in the signs of the corrections to the intensity of the head and high-frequency lines in the same series as well as to the intensity of the Zeeman π and σ components of the same line.

For practical applications the most interesting effect is the induction of transitions which are forbidden in the dipole approximation. The diamagnetic interaction of an atom with the field is completely responsible for this effect. The intensity of such transitions increases rapidly with the principal quantum number of the upper level; this shows that the transitions between Rydberg states with angular momentum $l = 3$ and the nS ground state can be observed in quite weak fields which are attainable under laboratory conditions. This effect could be helpful for single-photon pumping of F states of an atom in a magnetic field. A magnetic field can be used in the same manner for two-photon population of G states ($l = 4$), and so on. We note that magnetic-field-induction of dipole-allowed transitions in the spectrum of an atom is similar to the analogous phenomenon in an electric field, specifically, deexcitation of metastable states of an atom by an electric field [2], and can appear, for example, in the quenching of metastable $\left(np^5\ n+1s\left[\frac{3}{2}\right]_2\right)$ and $\left(np^5\ n+1s'\left[\frac{1}{2}\right]_0\right)$ states of inert atoms.

ACKNOWLEDGMENTS

This work was supported by the Ministry of Education of Russia (project no. 97-0-5.1-63) and INTAS (grant no. 97-0369).

REFERENCES

1. I. I. Sobel'man, *Introduction to the Theory of Atomic Spectra* (Fizmatgiz, Moscow, 1963).
2. H. A. Bethe and E. E. Salpeter, *Quantum Mechanics of One- and Two-Electron Atoms* (Academic, New York, 1957; Fizmatgiz, Moscow, 1960).
3. V. S. Lisitsa, Usp. Fiz. Nauk **153**, 379 (1987).
4. H. Friedrich, *Theoretical Atomic Physics* (Springer-Verlag, Berlin, 1991).
5. L. A. Bureeva and V. S. Lisitsa, *Perturbed Atom* (IzdAT, Moscow, 1997).
6. T. P. Grozdanov and H. S. Taylor, J. Phys. B **19**, 4075 (1986).
7. P. A. Braun, Rev. Mod. Phys. **65**, 115 (1993).
8. A. M. Vaĭnberg, V. A. Gani, and A. E. Kudryavtsev, Zh. Éksp. Teor. Fiz. **113**, 550 (1998) [JETP **86**, 305 (1998)].
9. V. D. Ovsiannikov, Phys. Rev. A **57**, 3719 (1998).
10. V. D. Ovsyannikov and K. V. Khalev, Zh. Éksp. Teor. Fiz. **116**, 838 (1999) [JETP **89**, 444 (1999)].
11. W. Rösner, G. Wunner, H. Herold, and H. Ruder, J. Phys. B **17**, 29 (1984).
12. M. V. Ivanov, J. Phys. B **21**, 447 (1988).
13. T.-H. Wang and C.-S. Hsue, Phys. Rev. A **52**, 4508 (1995).
14. Yu. P. Kravchenko, M. A. Liberman, and B. Johansson, Phys. Rev. A **54**, 287 (1996).
15. V. D. Ovsiannikov and S. V. Goossev, Phys. Scr. **57**, 506 (1998).
16. V. D. Ovsyannikov and V. V. Chernushkin, Zh. Éksp. Teor. Fiz. **116**, 1161 (1999) [JETP **89**, 618 (1999)].
17. P. A. Braun, Opt. Spektrosk. **69**, 1206 (1990) [Opt. Spectrosc. **69**, 713 (1990)].

18. S. V. Goossev and V. D. Ovsiannikov, J. Phys. B **28**, 5251 (1995).
19. S. V. Gusev and V. D. Ovsyannikov, Opt. Spektrosk. **83**, 893 (1997) [Opt. Spectrosc. **83**, 822 (1997)].
20. M. D. Jones, G. Ortiz, and D. M. Ceperley, Phys. Rev. A **59**, 2875 (1999).
21. M. V. Ivanov, J. Phys. B **27**, 4513 (1994).
22. X.-X. Guan and Z.-W. Wang, Phys. Lett. A **244**, 120 (1998).
23. M. V. Ivanov, Phys. Lett. A **239**, 72 (1998).
24. M. V. Ivanov and P. Schmelcher, Phys. Rev. A **57**, 3793 (1998).
25. E. U. Condon and G. H. Shortley, *The Theory of Atomic Spectra* (Macmillan, New York, 1935; Inostrannaya Literatura, Moscow, 1949).
26. G. Simons, J. Chem. Phys. **55**, 756 (1971).
27. N. L. Manakov, V. D. Ovsyannikov, and L. P. Rapoport, Opt. Spektrosk. **38**, 206 (1975) [Opt. Spectrosc. **38**, 115 (1975)].
28. D. A. Varshalovich, A. N. Moskalev, and V. K. Khersonskiĭ, *Quantum Theory of Angular Momentum* (Nauka, Leningrad, 1975; World Scientific, Singapore, 1988).
29. N. L. Manakov, V. D. Ovsiannikov, and L. P. Rapoport, Phys. Rep. **141**, 319 (1986).
30. *Heigher Transcendental Functions (Bateman Manuscript Project)*, Ed. by A. Erdelyi (McGraw-Hill, New York, 1953; Nauka, Moscow, 1965, 1966), Vol. 1, Chap. 5.7; Vol. 2, Chap. 10.2.
31. M. Bellermann, T. Bergeman, A. Haffmans, *et al.*, Phys. Rev. A **46**, 5836 (1992).

Translation was provided by AIP

Journal of Experimental and Theoretical Physics, Vol. 91, No. 3, 2000, pp. 469–478.
Translated from Zhurnal Éksperimental'noĭ i Teoreticheskoĭ Fiziki, Vol. 118, No. 3, 2000, pp. 539–549.

PLASMA, GASES

Formation, Cascade Development, and Rupture of the *X*-Pinch Neck

G. V. Ivanenkov*[, a], S. A. Pikuz[a], T. A. Shelkovenko[a], J. Greenly[b], D. B. Sinars[b], and D. A. Hammer[b]

[a]*Lebedev Physical Institute, Russian Academy of Sciences, Moscow, 117924 Russia*
[b]*Laboratory of Plasma Studies, 369 Upson Hall, Cornell University, Ithaca, New York 14853, USA*
**e-mail: ivanenkov@sci.lebedev.ru*

Received January 24, 2000

Abstract—Data from high-resolution x-ray shadow photography of an *X*-pinch in the diode of a high-power dense-plasma generator are presented. The processes leading to the formation of a minidiode, the compression of the neck arising in it, and the cutoff and subsequent emptying of the neck are studied. Cascade formation of short-lived structures, which consistently reproduce the form of the minidiode on small scales before the x-ray burst, is observed in the course of the implosion. The position of the x-ray emission points is determined.

1. INTRODUCTION

An electrical explosion of wire loads is used in modern research to obtain powerful x-ray pulses [1]. The properties of matter in such nanosecond discharges vary over a wide range from a liquid and vapor up to a strongly emitting hot and high-density plasma with multiply-charged ions, greatly complicating an already difficult diagnostics problem. In solving this problem, a method of high-resolution x-ray shadow photography has been developed over the last few years [2, 3]. This made it possible to begin the study of the previously unknown internal structure of a discharge through an exploding wire, containing together with the usually studied plasma phase of the corona, core material with heterogeneous composition (liquid–vapor) [4]. The latter could be responsible for the unusually high values of the plasma parameters (the density, temperature, and ion charge) characterizing the hot points of the necks in discharges through a wire. An *X* pinch—an important element of the new method—gives a unique experimental realization of a single neck.

The diagnostics is based on the method of shadow projection of a point source. The large, compared with laser probing, shortening of the wavelengths of the source radiation makes deep layers accessible: the limit of the achievable density, determined in the laser method by the nonuniformity of the plasma and the aperture angles of the optics employed, is now several orders of magnitude (2–3 and more) greater than for a light-range laser, and the weak sensitivity to sharp changes in density makes it possible to operate with essentially arbitrary gradients. Investigations [2, 3] have demonstrated that the bursts of the hot points in *X* pinches, placed in the diode of a high-voltage generator, are highly efficient as miniature short-pulse x-ray sources. The subject of these investigations were discharges through one or several parallel wires, while the present work is devoted to the processes occurring in an *X* pinch, which is an interesting object and an important element of the observation method.

2. EXPERIMENTAL ARRANGEMENT

In our experiments, the plasma generator consisted of an *XP* setup (Cornell University; parameters: 470 kA, 0.5 Ω, 100 ns). Pairs of crossed Mo wires with diameters 12.7, 17, 25, and 30 μm were exploded in its diode. Their radiation was detected with diamond detectors with photoconductivity, whose signals were recorded using a fast Tektronix 684B oscillograph. The time resolution of the detectors, taking account of the input cables and the oscillograph, was ≈0.5 ns, but in reality the determining resolution of the method—the duration of the burst of the hot point of a Mo *X* pinch—was even shorter. Its measurement was a difficult problem: even the latest experiments make it possible to assert only that the duration of the burst lies within the resolution of the apparatus employed—shorter than 250 ps for a high-frequency oscillograph and <100 ps for a slit-scan camera. The spatial resolution of x-ray images depended on the size of the hot spot, the distance from this spot to the object, and the hardness of the detected radiation (filters, film types and position). The geometric ratio of the distances of the object from the recording photographic film and the source (magnification) was 4–10 and gave good quality photographs: after digitizing with a Nikon LS-2000 scanner the resolution was estimated to be 1.2 μm. Detecting films with different sensitivity arranged in the form of "sandwiches" Kodak RAR 2497, DEF, Mikrat VE, and Ti filters

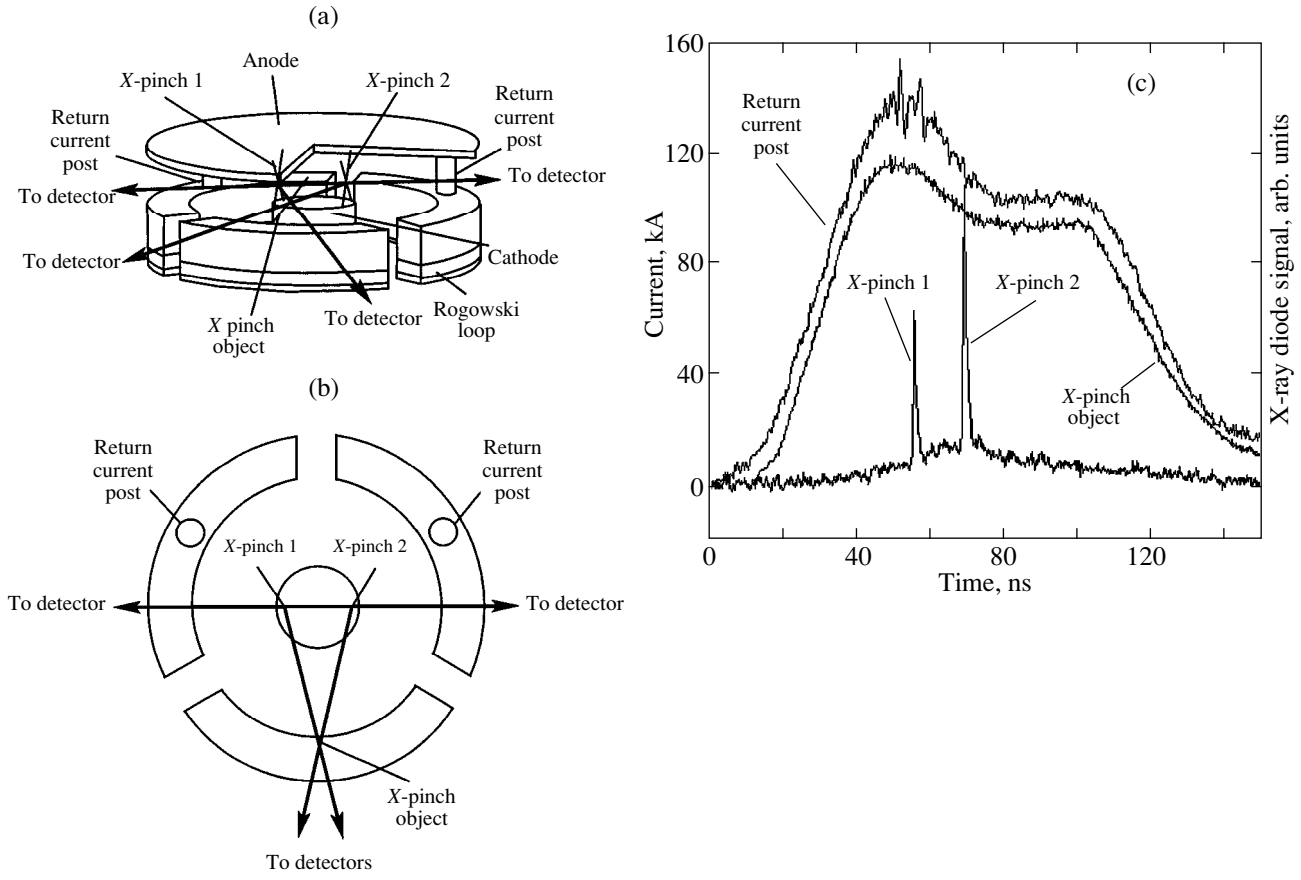

Fig. 1. (a), (b) Diagram of the experiment, (c) typical oscillograms of the current through the load and the signal from the x-ray detector.

12.5 μm thick made it possible to obtain at the same time images in radiation with different hardness.

Two *X* pinches were placed in the diode in parallel, serving as sources of x-rays for obtaining in one shot two (at different times) images of the experimental object. The time delay between the bursts of up to 30 ns was determined by the difference of the load masses, which were chosen on the basis of experience; the variance for virtually identical masses was 0.2–2 ns. Another advantage of the scheme with a pairwise arrangement of pinches in the diode was an extremely small, in contrast to a discharge through one *X* pinch, number of bursts of hot points (1–2). This effect, caused by current redistribution between the loads, improved the quality of the images obtained. Optimization of the source dimensions made it possible to operate in the spectral range 1–5 Å, including the transmission band 2.5–4 Å of the Ti filters.

Two variants of the diagnostics complex were used (Fig. 1). First, there was the standard scheme which used in our previous experiments, where the experimental load replaced outside the diode one of three return-current rods (Fig. 1a); the other two supported the anode plate. The diode current was divided into three parts corresponding to the inductances of the load and rods, the currents through which were measured with individual Rogowski loops (oscillograms in Fig. 1c). Because of the high inductance, less than one third of the total current, usually not greater than 120 kA, flowed through the load; this made it possible either to use only very thin wires or to study only the onset of an explosion. *X* pinches based on various metals, from Al to Au, with wire diameters from 7.5 to 25 μm were investigated in this scheme.

In the other scheme, which was used to obtain most of the images presented below, the *X* pinches in the diode served simultaneously as the object of investigation and the source of x-rays. Although the current in this apparatus was divided between the pinches, only the total current was measured, and the currents of the *X* pinches were assumed to be approximately the same, from 200 to 230 kA. Only *X* pinches from Mo and W were investigated in this scheme. The difficulties were exacerbated by the fact that very often the pinches still had more than one hot spot, degrading the quality of the recorded image. Typical signals from diamond photodetectors are shown in Fig. 1c.

The method of crossing the thin conductors in the diode was also important: the processes occurring in the neck depended on how the wires touched one

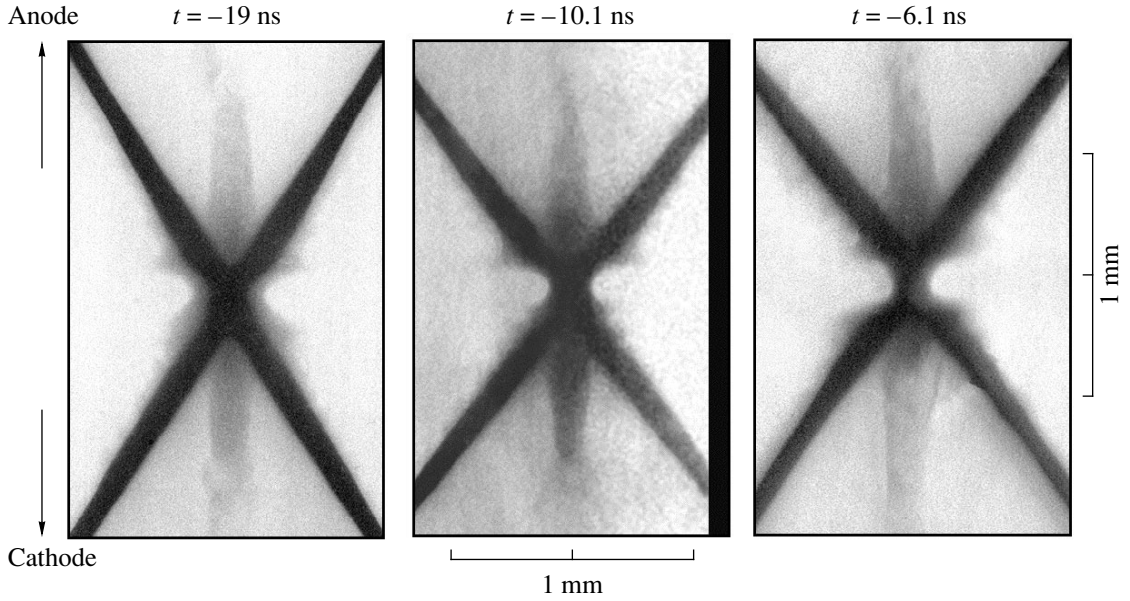

Fig. 2. Slow stage of the evolution of a *X* pinch before the x-ray burst.

another: excess twisting of the strands engendered parasitic helical disturbances, impeding good compression of the plasma. "Point" touching of the wires was best. The intensity of the x-ray burst was highest in this case. The angle 2α between the wires was also varied. The best choice was $\alpha \simeq \pi/6$.

3. FORMATION OF A MINIDIODE AND NECK IN AN *X* PINCH

The general picture of the evolution, constructed according to data from many shots, is shown in Figs. 2–4. It substantially supplements previous results of optical and x-ray spectral measurements [5]. It is evident how the slow (10 ns scale; Fig. 2) process of formation of a minidiode transforms into a rapid process, accelerating up to the moment of the burst the development of compression instabilities of the neck (Fig. 3), which is completed by a cutoff of the neck and subsequent emptying of the minidiode (Fig. 4). The time scales of these processes run successively through values from the initial values from the initial 10 ns to 1–2 ns up to a moment approximately over 5 ns up to the moment of the burst taken as 0 and then to 0.1 ns. After the burst the sequence of scales with the same values alternates in the opposite direction.

Let us examine the photographs in greater detail. The photograph in Fig. 2 shows the core metal remaining after the explosion and the corona plasma. The dense core matter behaves differently inside and outside the neck region. Inside the neck it gradually merges into a single, evidently liquid, filament, and the initial trough (frame for –19 ns) transforms into a minidiode shorted by a dense filament. No indications of boiling are noticeable here. The situation is completely different in the outer part of the cores, where the pattern is similar to parallel wires: the volume is frothy with vapor bubbles and the current can flow only along the surrounding plasma corona. It appears that the greater the distance from the wire crossing, the more intense the heating and explosion of the metal by the current are. In the corona surrounding the cores, the increase in the amplitude of the plasma perturbations as the crossing is approached draws attention. At the crossing itself, where a minidiode gradually forms, no corona is seen; but, incidentally, this does not mean that it is absent. Nonetheless, it can be concluded that in time cylindrical geometry dominates around individual wires far from the minidiode and is more noticeable inside.

The coronal plasma acquires a special structure near the crossing—these are axial jets which were observe in preceding frames. Their asymmetry is noticeable: the jet is a sharp cone in the direction of the anode and opens up toward the cathode. They can be attributed to the collision of the fluxes of matter evaporated from the surface of the metal at the start of the explosion. The inclination of the front of the vapor cloud, propagating along the normal to the surface of the wires with velocity $3c_s$ ($c_s \approx (\varepsilon_F/m_i)^{1/2} \approx (1–2) \times 10^5$ cm/s—the velocity of sound in the cold phase, ε_F is the Fermi energy of the electrons in the metal) toward the axis gives rise to cumulation, and the axial velocity of the jet is

$$3c_s/\sin\alpha \approx (0.6–1.2) \times 10^6 \text{cm/s}.$$

Therefore, in the 30–40 ns elapsed from the onset of the current the length of the jet ranges from 200 to 400 μm (in our photographs 300–400 μm). The rest of the vapor expands freely, until the braking action of the fields has an effect. The amount of matter flowing into a jet and

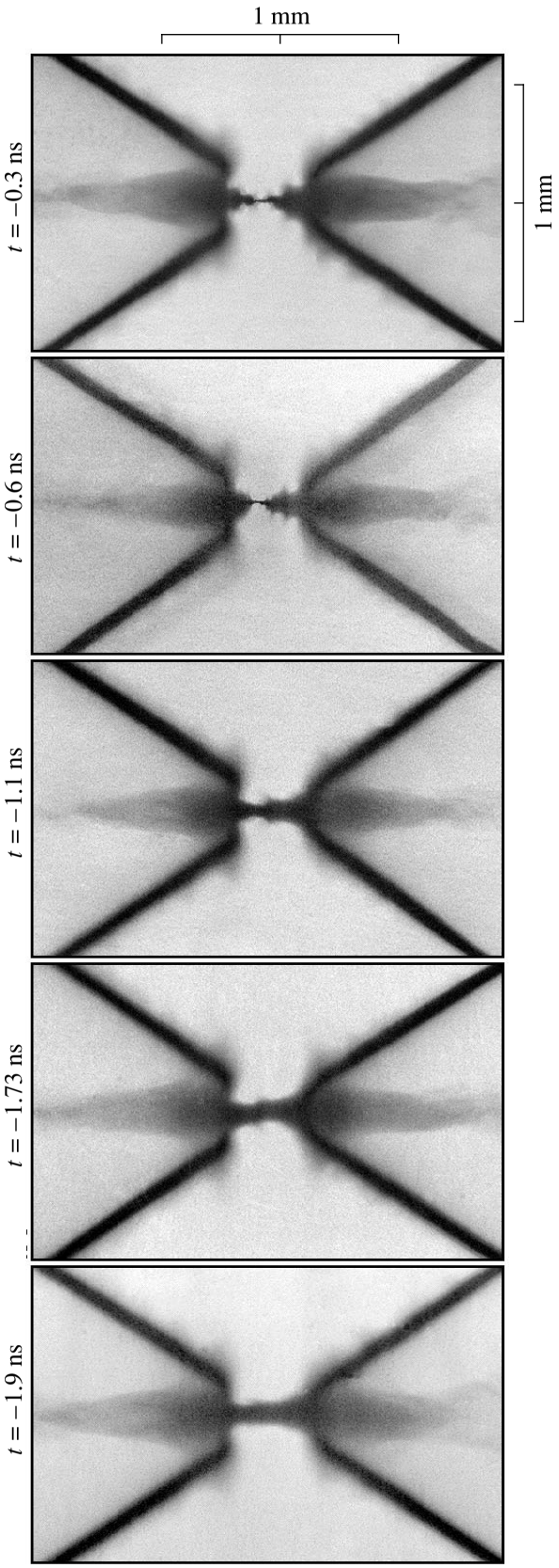

Fig. 3. Fast stage of the evolution of a *X* pinch before the x-ray burst.

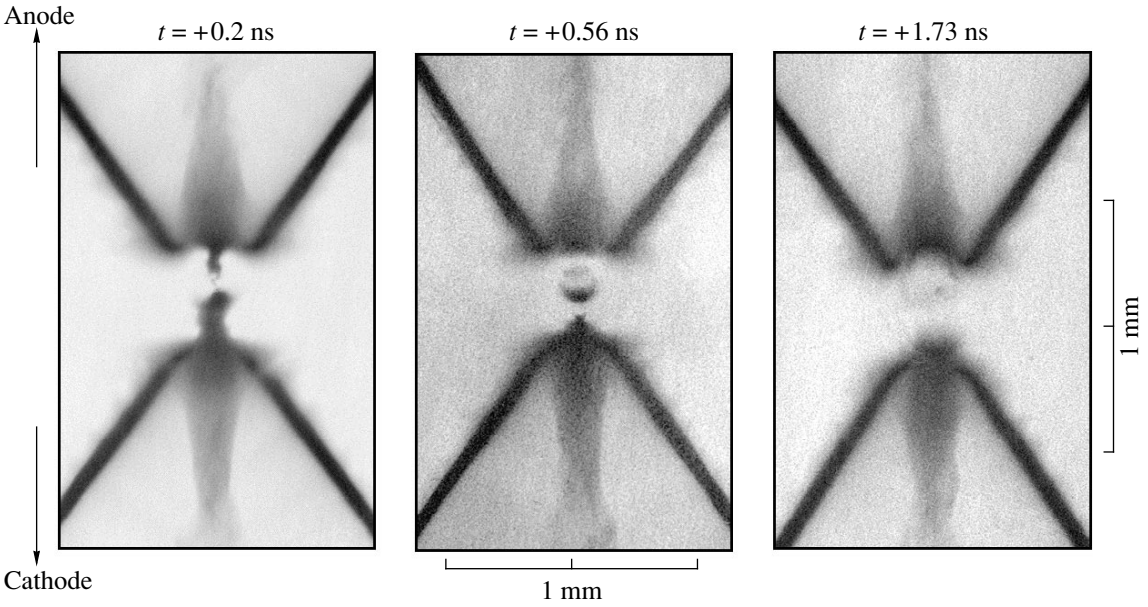

Fig. 4. Image of a *X* pinch at fast stage of emptying of the minidiode after the x-ray burst.

the role of this matter in the current distribution at the entrance and exit of the minidiode are still not known.

The origin of the longest of our time scales, 10 s, can be understood. It is the same as the heating time of the neck volume

$$\tau \approx a^2/2\pi\chi,$$

where $a \approx 100$ μm is the radius of the neck and $\chi = 2\kappa/3n_i \propto Z^{-1}T^{5/2}$ is the electronic thermal diffusivity. For this, the heating wave gradually converging into the interior must heat the plasma to $T \approx 30$ eV and $Z \approx 5$ ($Z \propto T^{1/2}$). Under such conditions the magnetic viscosity

$$\chi_m = c^2/4\pi\sigma \propto ZT^{-3/2}$$

reaches $\approx 5 \times 10^4$ cm^2/s. Since the magnetic Reynolds number

$$R_m = va/\chi_m < c_s a/\chi_m \approx 0.1$$

($c_s = (ZT/m_i)^{1/2} \approx 7 \times 10^5$ cm/s is the velocity of sound), here diffusion determines the penetration of the current and field into the interior volume of the plasma. The required time

$$\tau_m = a^2/2\pi\chi_m$$

is only ≈0.3 ns, and the power density of Joule heating with a typical current $I = 100$ kA reaches

$$j^2/\sigma \approx 2I^2/\pi^2c^2\tau_m a^2 \approx 50 \text{ TW/cm}^3.$$

Such a wave of current heating deviates increasingly from cylindrical symmetry as it penetrates.

In summary, at the crossing point the dense core matter is gradually heated and redistributed, evidently, primarily in the azimuthal direction, forming a single cylindrical neck inside the minidiode. Outside the crossing the processes resemble an explosion of a pair of parallel wires, but because of the inclination a cumulative collapse of the evaporated matter with formation of a special structure of the coronal plasma–axial jets–occurs. Their velocities are still much higher than the axial expansion rate of the dense matter in the crossing, and the latter does not yet participate in the formation of the jets. Dense matter, evidently still not compressed much, appears in the base of the jets only at the next stage; its outflow is still not substantial in the process of magnetic compression of the neck, especially since it is impeded by the dense matter forming the electrodes of the minidiode. These are the conditions at the transition to rapid compression of the neck.

4. PROCESSES NEAR AND AFTER THE X-RAY BURST

We shall now turn to the frames of the events immediately before the x-ray burst and rupture of the neck (Fig. 3; Fig. 5 shows enlarged, brighter images of the minidiode; the thinnest neck with a sharp boundary and thickness ≈6 μm was photographed at $t = -0.3$ ns almost at the moment of x-ray emission). Comparing the frames for $t = -1.9$ and -173 ns, we note how short the time is between the quite even neck and the appearance of an appreciable distortion at the center. Its smooth development between $t = -2$ and -1 ns then accelerates sharply, and in the frames for $t = -0.6$ ns and -0.3 ns narrowing is clearly evident in the anode half of the minidiode, encompassing approximately one-fourth of the total length of the filament closer to the middle. The picture is associated with the develop-

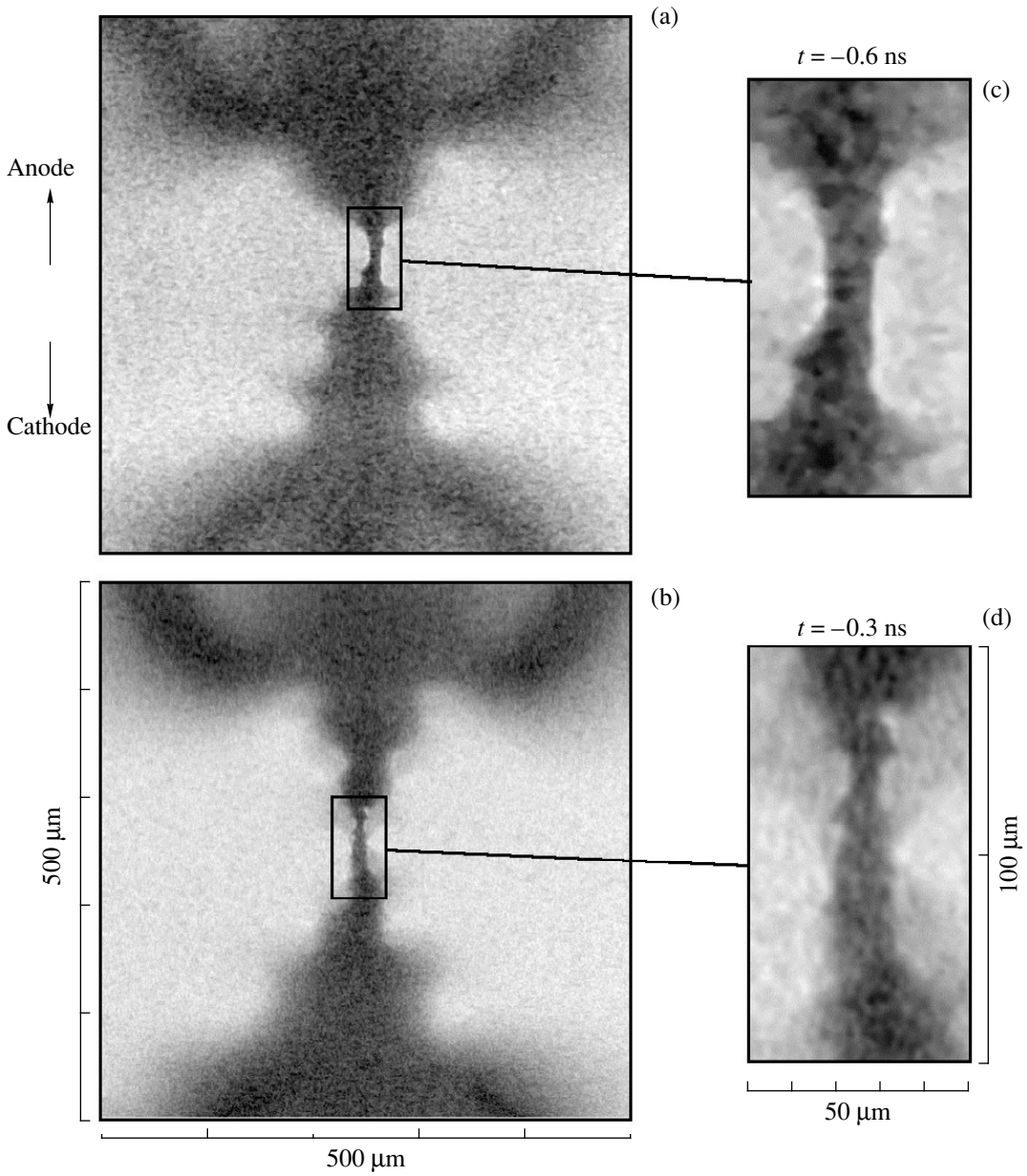

Fig. 5. Enlarged images of a minidiode in the phase of fast compression of the neck and formation of a cascade of diode structures.

ment of an instability of the neck, during which a diode structure of a smaller scale forms in the minidiode. Closer study of Fig. 5 shows the appearance of similar structures, closer to the burst (t = –0.3 ns), of a higher order inside a new neck. Cascade development of the neck, terminated by an explosion, occurs. It is visible in the frame for t = +0.2 ns (Fig. 4) and somewhat later the same thing appears on the "minicathode" side also; for $t > 0.5$ ns an almost symmetric pattern with a central plasma bunch arises.

The form of the cutoff of the filament in Fig. 4 is associated with a microexplosion of the neck. The characteristic effects are noticeable in enlarged fragments (Fig. 6) of the first three frames: shock fronts propagate away from the cutoff points in the direction of the central bunch and minielectrodes. The bunch at the center gradually dissipates in the next nanosecond, and by the end of the second nanosecond after the burst the minidiode contains almost no appreciable plasma. At the same time, indications of an increase in the mass of the dense matter in the bases of the jet cones on the minielectrodes appear. Moreover, reflection of wave fronts accompanying an interaction with stationary dense cores is observed. The shock waves penetrating into the space between the cores gradually displace the dense neck matter into the region near the symmetry axis of the X pinch.

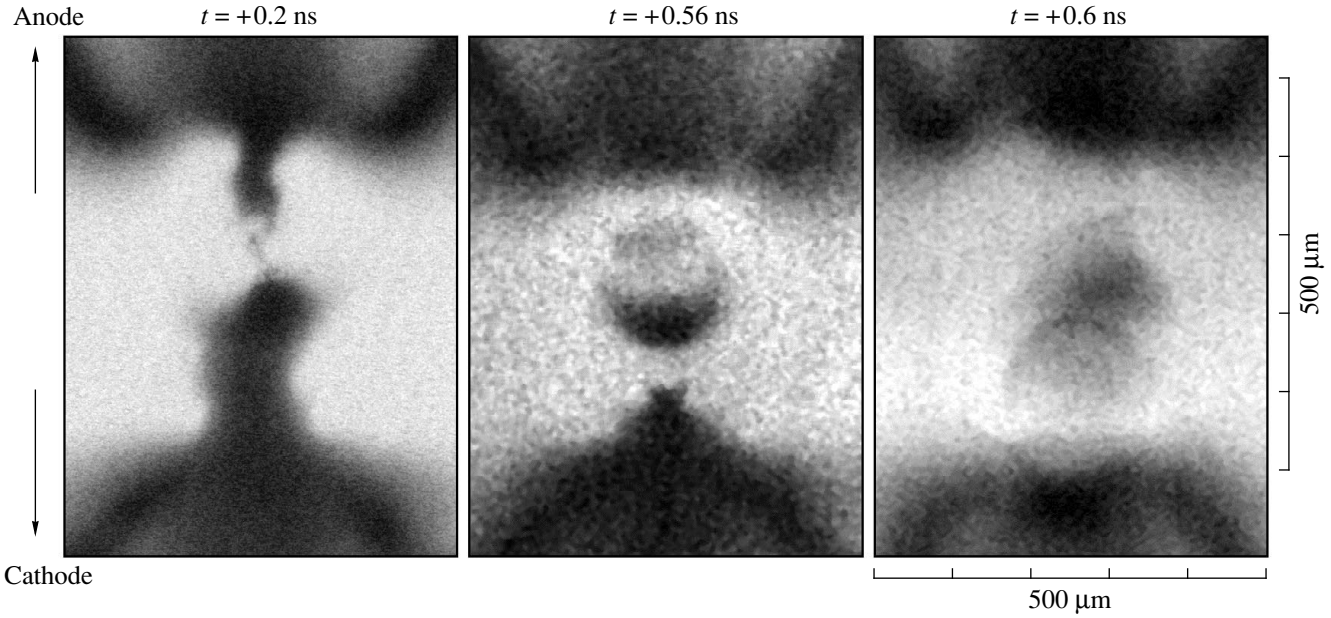

Fig. 6. Enlarged fragments of the first images in Fig. 4.

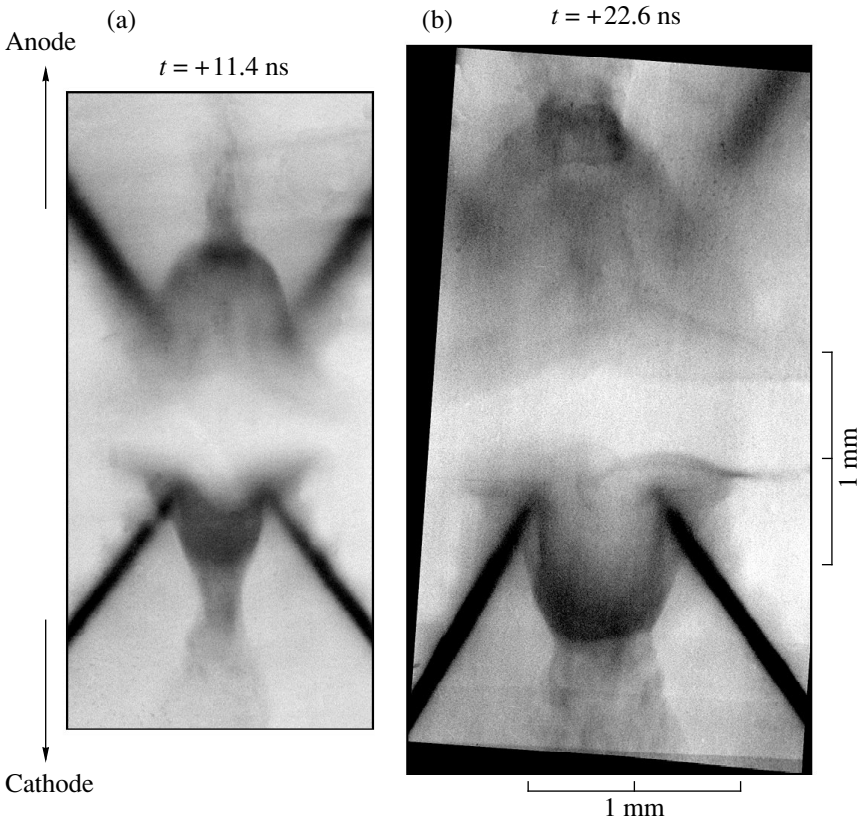

Fig. 7. The process of slow outflow of the neck matter from the minidiode.

The initial expansion rate of the plasma in the microexplosion can also be estimated. For this, we shall employ spectral data that make it possible to judge the processes occurring at the stage of the brightest emission of the plasma in the x-ray burst. The relative half-width of the lines of Ne-like Mo was

$$\Delta\lambda/\lambda = 3 \times 10^{-3}.$$

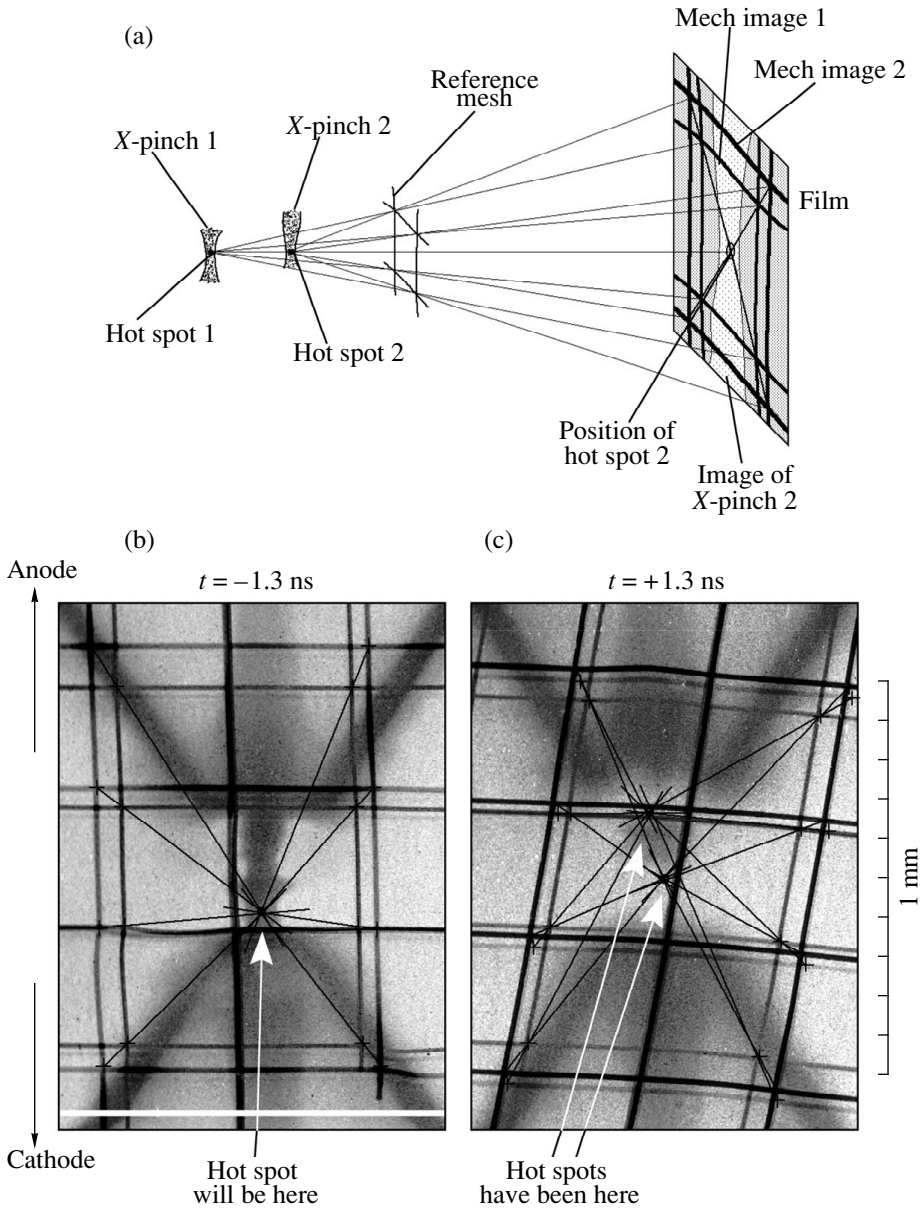

Fig. 8. (a) Scheme and (b) and (c) examples of the determination of the positions of the hot spots.

Attributing it entirely to Doppler broadening, caused by uniform spreading of the plasma in 4π directions from the point of the explosion, we estimate the velocity to be 9×10^7 cm/s. For comparison, the velocity of a fast magnetic sound

$$c_{fm} = [(ZT_e + B^2/4\pi n_i)/m_i]^{1/2}$$

under Bennet equilibrium conditions with $Z = 30$ and $T_e = 1$ kV is three times lower. In other words, the plasma of the hot points expands with this velocity seemingly into a void; i.e., the plasma does not feel the more rarefied matter in front of it. Of course, this estimate, which neglects other expansion mechanisms, could be too high. Nonetheless, it still characterizes the initial stage of the microexplosion, whose duration can be estimated by the expansion time. Even according to the maximum spatial resolution of 1 μm the latter is only 1 ps, which is very short compared to our time resolution. Consequently, most of the explosive expansion should be much slower. X-ray spectral methods were also used to measure the average (time-integrated) temperature of the hot spots $T_e = 0.8$–1.3 keV, and generation of electronic and ionic beams in the minidiode was detected.

The subsequent decomposition is even slower. This is demonstrated in Fig. 7, in the later frames of which the cores and plasma formations are appreciably larger. Here one can see how the dense flow of plasma displaced

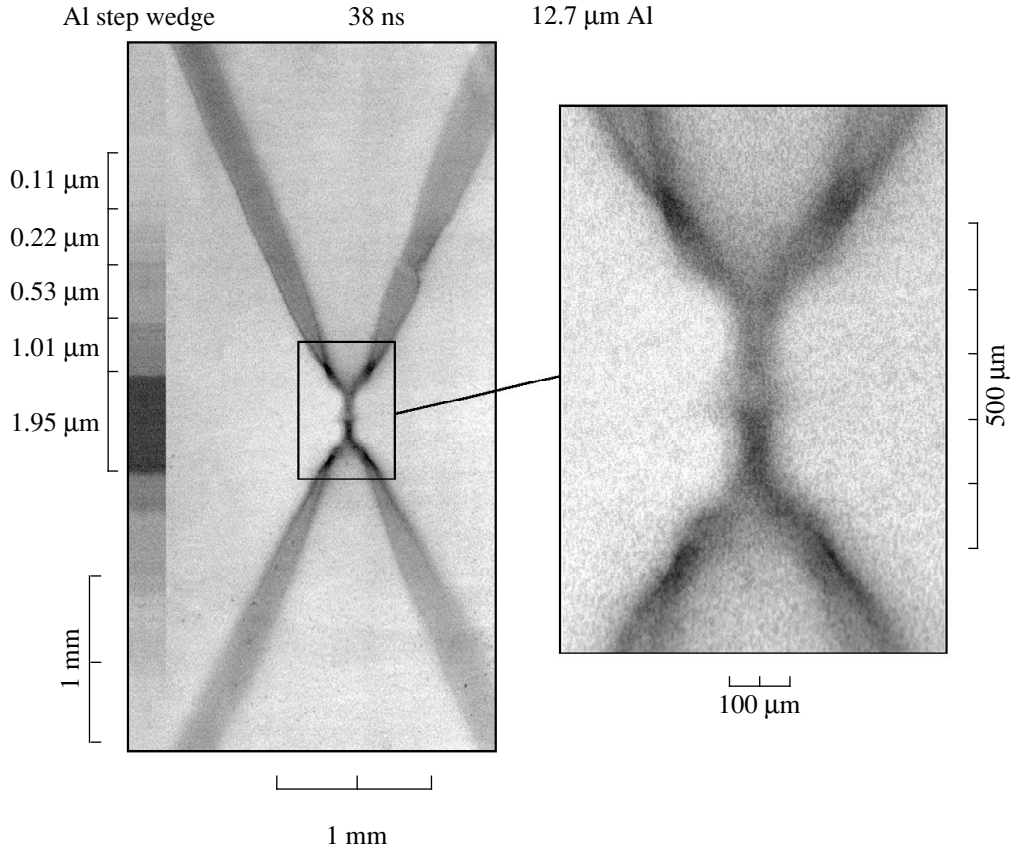

Fig. 9. Density distribution of the matter of an aluminum *X* pinch inside and near the neck.

from the minidiode squeeze through the minielectrodes. Similarly to what happens when a fast moving body strikes a barrier, splashes of unloading matter can be seen directed into the minidiode, and gradual swelling of the minielectrodes into the space between the cores is observed at the center.

The rapid x-ray burst is a cumulation of events in the evolution of an *X*-pinch type discharge. We were able to determine the positions of the burst points. The scheme shown in Fig. 8 was used for this purpose. In this scheme, an insulated grid with high transmittance was placed between the pinch object and the detecting film (the spacing of the cells is much larger than the diameter of the wires). A double image of the grid, obtained from bursts of both pinches at different times, formed on the film together with an image of the object. The position of the point of the pinch object was determined by the double mask image method according to the intersection of the lines connecting the corresponding points of two images (see scheme in Fig. 8). This made it possible to work with one- and two-point burst; the position of the hot emitting points in the neck was determined with accuracy no worse than ≈5 μm. Examples of the determination are shown in the same Fig. 8 against the background of photographs of processes occurring both before and after the burst, for one hot spot and for a pair of hot spots.

Processes occur similarly in *X* pinches from other metals, for example W, with the same arrangement as for Mo. But similar phenomena occur in the scheme with the pinches arranged in the return current lead, though one- or two-point bursts are more difficult to obtain here (such experiments are explained in greater detail in [6]). It is important that for filaments with close dimensions the length scales and the sequence of the temporal phases of the events are the same in both schemes, and as the diameter of the wires increases, the spatial and temporal scales of the phenomena also increase correspondingly, but the patterns remain similar. Axial jets of rarefied plasma can be clearly seen in tungsten, just as in molybdenum. The image of an aluminum *X* pinch at 38 ns from the onset of current flow, obtained in a scheme with a load in the return current circuit, is presented in Fig. 9. Here the current was low for the burst of the hot spot. A stepped attenuator (Al, the thicknesses are indicated to the side in the figure) made it possible to obtain a picture of the density distribution in an *X* pinch at the slow stage of formation of a minidiode. It is evident that far from the crossing the cores are in a two-phase vapor–liquid state, and the highest densities occur inside and close to the minid-

iode. No axial jets are visible here; this is typical for Al and a number of other metals, because the matter of the jets is transparent to the probe radiation from the bursts of an Mo *X* pinch.

5. CONCLUSIONS

Our observations have revealed a number of unique processes in the minidiode of a *X* pinch. It is still difficult to say how similar they are to an implosion in discharges through single wires. Probably, the azimuthal nonuniformity, caused by the presence of a distinguished plane for the arrangement of wires in the load (at least one) and the three-dimensionality of the magnetic field in the *X* pinch, is important. Even in the crossing itself the formation of a minidiode initially is different from single wires, and the more dense matter is subjected to compression in the minidiode. This could explain the high x-ray spectral parameters recorded in an *X* pinch.

The central event of the evolution of a pinch is the almost instantaneous x-ray burst. It coincides with cascade development of an increasingly more complicated minidiode structure with an increasingly thinner neck. We note that the similarity of such structures is consistently reproduced, even though the scales decrease very substantially. How strong can the compression be? According to calculations [7], the size of the neck can decrease to 0.1–0.3 μm, comparable to the range of the radiation. This makes it possible to estimate the radiation energy flux from the hot spot. According to measurements, the total yield of x-rays in the range 2.5–3.5 Å is 0.1–0.3 J, the emission time < 0.1 ns (see Section 1), and the size of the emitting plasma falls within the experimental resolution of 1 μm. Consequently, it is not an exaggeration to take ≈0.5 J for the energy emitted in the entire spectrum of the hot spot. Then the energy flux should be $>0.5/(\pi \times 10^{-8} \times 10^{-10}) \approx 10^{17}$ W/cm^2. If the maximum computed radius of the neck is taken to be 0.1 μm, then the flux is two orders of magnitude larger. In this case the radiation pressure can reach subgigabar and even gigabar values (this form of pressure was taken into account in [7]) and plays an important role in the compression dynamics. These estimates, of course, determine the upper limits, characterizing only a very small mass of the compressed matter. But they make it possible to obtain the largest values of the radiation fluxes incident on the object. Thus, for a pair of *X* pinches separated by a typical distance of 3 cm and irradiating one another, taking the energy flux emitted from the hot spot as 10^{18} W/cm^2 and assuming the emission to be isotropic, we obtain probe radiation fluxes of 10^7 W/cm^2.

Although evidence for rupture of the neck had been noted previously [8], the inadequate temporal resolution of optical interferometry still did not guarantee a final result, and the data obtained can be taken as confirmation of the effect. Especially interesting, together with the somewhat mysterious cascade development of the neck, is the microexplosion which terminates this process. We hope that subsequent investigations will shed light on the new data obtained using the technique of x-ray shadow photography of dense and hot plasma.

ACKNOWLEDGMENTS

We are grateful to S. Yu. Gus'kov and Ya. S. Dimant for a discussion of this paper. This work was partially supported by grant no. DE-FG02-98ER54496 and contract no. BD-9356 of the Sandia National Laboratory (USA).

REFERENCES

1. R. B. Spielman, C. Deeny, G. A. Chandler, *et al.*, in *Proceedings of the 4th International Conference on Dense Z-Pinches, Vancouver, 1997*, p. 101.
2. D. H. Kalantar and D. A. Hammer, Phys. Rev. Lett. **71**, 3806 (1993); T. A. Shelkovenko, S. A. Pikuz, A. R. Mingaleev, and D. A. Hammer, Bull. Am. Phys. Soc. **42**, 2051 (1997).
3. S. Yu. Gus'kov, G. V. Ivanenkov, A. R. Mingaleev, *et al.*, Pis'ma Zh. Éksp. Teor. Fiz. **67**, 531 (1998) [JETP Lett. **67**, 559 (1998)]; G. V. Ivanenkov, A. R. Mingaleev, S. A. Pikuz, *et al.*, Zh. Éksp. Teor. Fiz. **114**, 1216 (1998) [JETP **87**, 663 (1998)]; Fiz. Plazmy **25**, 851 (1999) [Plasma Phys. Rep. **25**, 783 (1999)]; T. A. Shelkovenko, S. A. Pikuz, A. R. Mingaleev, and D. A. Hammer, Rev. Sci. Instrum. **70**, 667 (1999).
4. S. A. Pikuz, G. V. Ivanenkov, T. A. Shelkovenko, and D. Hammer, Pis'ma Zh. Éksp. Teor. Fiz. **69**, 349 (1999) [JETP Lett. **69**, 377 (1999)]; S. A. Pikuz, T. A. Shelkovenko, D. B. Sinars, *et al.*, Phys. Rev. Lett. **83**, 4313 (1999).
5. G. V. Ivanenkov, A. R. Mingaleev, S. A. Pikuz, *et al.*, Fiz. Plazmy **22**, 403 (1996) [Plasma Phys. Rep. **22**, 363 (1996)].
6. T. A. Shelkovenko, S. A. Pikuz, D. A. Hammer, *et al.*, Phys. Plasmas **6**, 2840 (1999).
7. G. V. Ivanenkov and W. Stepniewski, J. Mosc. Phys. Soc. **9**, 337 (1999); G. V. Ivanenkov, S. A. Pikuz, D. B. Sinars, *et al.*, Fiz. Plazmy **26** (9), 927 (2000) [Plasma Phys. Rep. **26** (2000) (in press)].
8. S. M. Zakharov, G. V. Ivanenkov, A. A. Kolomenskiĭ, *et al.*, Pis'ma Zh. Tekh. Fiz. **8**, 1060 (1982) [Sov. Tech. Phys. Lett. **8**, 456 (1982)]; Fiz. Plazmy **13**, 206 (1987) [Sov. J. Plasma Phys. **13**, 115 (1987)].

Translation was provided by AIP

Journal of Experimental and Theoretical Physics, Vol. 91, No. 3, 2000, pp. 479–482.
Translated from Zhurnal Éksperimental'noĭ i Teoreticheskoĭ Fiziki, Vol. 118, No. 3, 2000, pp. 550–553.

PLASMA, GASES

Conservation Law for the Binormal Momentum in the Presence of Charge Drift in a Magnetic Field

B. A. Trubnikov

Russian Research Centre Kurchatov Institute, pl. Kurchatova 1, Moscow, 123182 Russia

Received March 14, 2000

Abstract—It is well-known that the motion of a charged particle in a magnetic field is described by the drift approximation, in which it is assumed that the squared velocity of the particle and the magnetic moment of the Larmor circle are conserved. It is shown that to a first approximation a third conservation law is also satisfied: the unaveraged generalized momentum in the direction of the binormal to the reference flux line around which the particle rotates is conserved. This new conservation law makes it possible to determine additional fine details of the motion, specifically, the deflection of a particle in the direction of the normal to the reference flux line, in terms of which the known drift velocity along the binormal is expressed after averaging. *© 2000 MAIK "Nauka/Interperiodica".*

1. For small Larmor radius ρ the motion of a charged particle in a magnetic field B is described well by the drift approximation, in which (1) the squared velocity v^2 of the particle is conserved, (2) the magnetic moment $\mu = mv_\perp^2/2B$ of the Larmor circle is conserved, (3) its guiding center (GC) moves along the reference flux line, and (4) the guiding center slowly shifts sideways along the binormal $\mathbf{b}$ to the reference flux line with average drift velocity

$$\mathbf{V_b} = \mathbf{b}\frac{v_\parallel^2 + (v_\perp^2/2)}{\omega R}, \tag{1}$$

where $\omega = qB/mc$ is the local Larmor frequency and R is the radius of curvature $K = 1/R$ of the reference flux line. These four features give a complete picture of the motion of a particle in the drift approximation.

2. We recall that the equations of the drift approximation which were described above are usually derived from Newton's equation (see, for example, [1])

$$\dot{\mathbf{v}} = \mathbf{v} \times \boldsymbol{\omega},$$

writing the position of the particle in the form

$$\mathbf{r} = \mathbf{R} + \boldsymbol{\rho}_\perp,$$

where $\mathbf{R}$ are the coordinates of the guiding center. Expanding the fields as

$$\mathbf{B}(\mathbf{r}) = \mathbf{B}(\mathbf{R}) + (\boldsymbol{\rho}\nabla)\mathbf{B}$$

in powers of the small radius ρ and averaging over fast Larmor oscillations gives Eq. (1). The term with $v_\parallel$ in it is due to the centrifugal force oriented along the normal $\mathbf{n}$ to the reference flux line, and the term with $v_\perp$ is due to the diamagnetic force

$$\mathbf{F}_\perp = \mu\nabla_\perp B, \tag{2}$$

where

$$\nabla_\perp B = \mathbf{n}KB.$$

However, such a "force" interpretation of Eq. (1) is not the only possible one. In the present paper it is shown that an alternative and even more detailed description of the motion can be obtained from the Lagrangian

$$L = \frac{1}{2}mv^2 + \frac{q}{c}\mathbf{v}\cdot\mathbf{A}. \tag{3}$$

3. We have discovered that for a small Larmor radius

$$\boldsymbol{\rho}_\perp = \mathbf{n}x + \mathbf{b}y$$

this Lagrangian must be regarded as being independent of the binormal component y, and consequently on the basis of the same drift approximation, together with conservation of v^2 and μ, there is a third conservation law—conservation of the generalized binormal momentum

$$P_b = \partial L/\partial\dot{y} = \text{const} = 0.$$

Subsequent averaging gives an alternative derivation of Eq. (1) with its new "nonforce" interpretation. In the process it also becomes clear that the particle motion differs from the simple picture of a drifting "circle." This difference is important for the derivation of Eq. (1).

4. As an illustration, it is helpful to examine first the cases where the exact conservation of the generalized binormal momentum is due to the symmetry of the field $\mathbf{B}$. As three "symmetric" examples, we shall choose the following cases of particle motion: (a) in the field of a wire carrying a current J_z; (b) in the field of a planar dipole $\mu_z^{(1)}$; and (c) in the field of a point dipole $\mu_z^{(2)}$.

1063-7761/00/9103-0479$20.00

Case (a). For a wire carrying a current the potential in cylindrical coordinates r, φ, z is

$$A_z = \frac{2J_z}{c}\ln\left(\frac{\text{const}}{r}\right)$$

and the Lagrangian does not depend on z, so that the generalized binormal momentum P_z is conserved. This can be rewritten as an exact expression for the velocity along the binormal:

$$v_b = \dot{z} = v_0 \ln\frac{r}{r_0} = r_0\omega_0 \ln\frac{r}{r_0}, \tag{4}$$

where

$$v_0 = \frac{2J_z q}{mc^2} = r_0\omega_0 = \text{const}, \quad \omega_0 = \frac{qB_0}{mc},$$

$$B_0 = 2J_z/cr_0.$$

We note here that the choice of a definite value of P_z means choosing a constant r_0 determining the "reference" magnetic surface near which the particle drifts.

Case (b). For a planar dipole in Cartesian coordinates x, y, z the potential is

$$A_z = \frac{\mu_z^{(1)} x}{x^2 + y^2}$$

and the Lagrangian likewise is independent of z, so that P_z = const is conserved, whence an exact expression can be obtained for the unaveraged velocity along the binormal:

$$v_b = \dot{z} = v_0\left(1 - \frac{x_0 x}{x^2 + y^2}\right), \tag{5}$$

where

$$x_0 = q\mu_z^{(1)}/cP_z, \quad v_0 = x_0\omega_0,$$

$$\omega_0 = \frac{qB_0}{mc}, \quad B_0 = \frac{\mu_z^{(1)}}{x_0^2}.$$

We note that here the choice of a definite value of P_z fixes the parameter x_0 and a reference magnetic surface, in this case with a circular section

$$[x - (x_0/2)]^2 + y^2 = (x_0/2)^2$$

with radius $x_0/2$.

Case (c). For a point dipole in cylindrical coordinates the potential

$$A_\varphi = \mu_z^{(2)} r(r^2 + z^2)^{-3/2}$$

and the Lagrangian are independent of the angle φ, so that the generalized momentum $P_\varphi = \partial L/\partial\dot{\varphi}$ is conserved. This gives an exact expression for the unaveraged angular velocity along the binormal

$$\dot{\varphi} = \omega_0\left[\frac{r_0^2}{r^2} - \frac{r_0^3}{(r^2 + z^2)^{3/2}}\right], \tag{6}$$

where

$$r_0 = \frac{q\mu^{(2)}}{cP_\varphi}, \quad \omega_0 = \frac{qB_0}{mc}, \quad B_0 = \frac{\mu_z^{(2)}}{r_0^3},$$

so that here the choice of a definite value of P_φ fixes the parameter r_0 and the reference magnetic surface, in this case with section

$$z^2 = r_0^{2/3} r^{4/3} - r^2.$$

Time-averaging of Eqs. (4)–(6) over the fast Larmor oscillations should give the average velocity (1) along the flux lines, corresponding to the binormals, of the three magnetic fields described. But, as we can see, it is not the forces (centrifugal and diamagnetic) but the particle coordinates, describing the deflection of a particle only in the direction of the corresponding normals to the reference magnetic surfaces, that appear in these formulas.

5. Since the formula (1) is applicable to any fields, an expression of the type Eqs. (4)–(6) should also hold in the general case with a small Larmor radius, which for our purposes we write in the form

$$\boldsymbol{\rho}_\perp = \mathbf{n}x + \mathbf{b}y.$$

However, we shall confine our attention only to vacuum magnetic fields, for which the vector potential and field can be represented as

$$\mathbf{A} = F\nabla G, \quad \mathbf{B} = \text{rot}\mathbf{A} = \nabla F \times \nabla G, \tag{7}$$

and therefore the flux lines of the field lie on the lines of intersection of two magnetic surfaces F = const and G = const, and the magnetic field is taken at the location of the particle.

However, we are interested only in the drift approximation, where the quantities x and y are assumed to be small, and we wish to expand the functions F and G in powers of x and y. On the reference flux line itself the field is $B_0(s)$, and its transverse gradient is

$$\nabla_\perp B = \mathbf{n}KB,$$

so that near the reference flux line the modulus of the field is approximately

$$B = B_0(1 + Kx).$$

In addition, it is easy to see that near the reference flux line the expansions of the functions F and G must have the form

$$F = x\left(1 + \frac{1}{2}Kx + \ldots\right)B_0(s), \quad G = y + \ldots, \tag{8}$$

since it is this representation that leads to the formula

$$B = B_0(1 + Kx)$$

for the field. Correspondingly, the Lagrangian in the drift approximation can be written in the simplified form

$$\mathcal{L} \approx \frac{1}{2}m[(\lambda\dot{s})^2 + \dot{x}^2 + \dot{y}^2] + \frac{qB_0}{c}x\left(1 + \frac{1}{2}Kx\right)\dot{y}, \quad (9)$$

where $\lambda = 1 - Kx$ is the longitudinal Lamé coefficient. The small terms with twisting of the reference flux line can be neglected. In addition, such a Lagrangian obviously does not depend explicitly on the coordinate y. Therefore, the unaveraged binormal generalized momentum is conserved approximately during the entire time of the motion:

$$P_b = P_y = \frac{\partial\mathcal{L}}{\partial\dot{y}} = m\dot{y} + \frac{qB_0}{c}x\left(1 + \frac{1}{2}Kx\right), \quad (10)$$

which, in addition, should be set equal to zero, which corresponds to choosing a definite reference flux line from which the coordinates x and y are measured.

This means that the binormal velocity of the particle is

$$\begin{aligned} v_b \approx \dot{y} &= -\omega x\left(1 + \frac{1}{2}Kx\right), \\ V_b = \langle\dot{y}\rangle &= -\omega\left(\langle x\rangle + \frac{1}{2}K\langle x^2\rangle\right), \end{aligned} \quad (11)$$

where $\omega_0 = qB_0/mc$. Thus the average drift velocity along the binormal **b** in the general case can be expressed in terms of the average deviations $\langle x\rangle$ and $\langle x^2\rangle$ in the direction of the normal! This is the main result of our Lagrangian description, which refines the picture of the familiar drift approximation where the quantities $\langle x\rangle \neq 0$ and $\langle x^2\rangle$ do not enter explicitly.

6. In turn, they should be determined from the equation obtained for the coordinate x from the simplified Lagrangian (9):

$$\ddot{x} = \frac{\partial\mathcal{L}}{m\partial x} = -K(1 - Kx)\dot{s}^2 + \omega(1 + Kx)\dot{y}. \quad (12)$$

Substituting here the expression (11) for $\dot{y}$, we obtain approximately the "main" equation for x:

$$\ddot{x} + \omega^2 x = \mathcal{P} = -K\left(\dot{s}^2 + \frac{3}{2}\omega^2 x^2\right), \quad (13)$$

which must be solved by successive approximations, assuming the right-hand side $\mathcal{P}$ to be small and setting $x = x_1 + x_2$. Then we have for the first term in the equation with a "semiclassical" solution

$$\begin{aligned} \ddot{x}_1 + \omega^2 x_1 &= 0, \quad x_1 = \frac{a}{\sqrt{\omega(t)}}\cos\phi, \\ \phi &= \int_0^t \omega(t')dt' + \phi_0, \end{aligned} \quad (14)$$

where a = const is the amplitude and ϕ is the angle of rotation. Substituting x_1 into $\mathcal{P}$, we write the equation for the second correction as

$$\ddot{x}_2 + \omega^2 x_2 = \mathcal{P} = C + D\cos 2\phi, \quad (15)$$

where

$$C = -K\left(\dot{s}^2 + \frac{3}{4}\omega a^2\right), \quad D = -\frac{3}{4}K\omega a^2.$$

Its approximate solution is

$$x_2 = \langle x\rangle - \frac{D}{3\omega^2}\cos 2\phi, \quad (16)$$

where

$$\langle x\rangle = \frac{C}{\omega^2} = -\frac{K}{\omega^2}\left(\dot{s}^2 + \frac{3}{4}\omega a^2\right).$$

Thus the complete expression for the deflection x of a particle from the reference flux line in the direction of the normal has the form

$$x = \frac{a}{\sqrt{\omega}}\cos\phi + \langle x\rangle + \frac{Ka^2}{4\omega}\cos 2\phi. \quad (17)$$

Here the term with $\cos 2\phi$ should not be neglected, since it is of the same order of magnitude as $\langle x\rangle$ and $\langle x^2\rangle$, without knowing which it is impossible to obtain the correct expression for the average drift velocity (1) along the binormal.

This velocity is obtained from the expression (11):

$$\begin{aligned} V_b &= -\omega\left(\langle x\rangle + \frac{1}{2}K\langle x^2\rangle\right) \\ &= \frac{K}{\omega}(\dot{s}^2 + \omega^2\langle x^2\rangle) = \frac{v_\parallel^2 + v_\perp^2/2}{\omega R} \end{aligned} \quad (18)$$

and is identical to the familiar expression (1).

7. In the opinion of the present author, the "Lagrangian" refinement, examined above, of the formulas of the standard drift approximation is not only of methodological but also of practical interest. In the first place, we have found a previously unnoticed and hence new conservation law for the generalized unaveraged binormal momentum. The reader can assess the practical usefulness of our formulas by attempting to solve any specific problem and answering the question: Where is the particle under consideration actually located?

The point is that the "stand drift approximation" gives a qualitative picture of the motion, but it does not adequately determine even the concept of a "guiding center" R of the particle under consideration. For example, in the cases (a), (b), and (c) indicated above the particle moves within the strip

$$r_{\min} < r < r_{\max},$$

and it would seem that the "guiding center" should be (mentally) placed at the center of the strip (the geometric center, $R = (r_{\min} + r_{\max})/2$). But this quantity does not appear at all in the formulas of the drift approximation, and it should be replaced by the more accurate concepts of "reference" flux line and deviation of x from it (specifically, the time average $\langle x \rangle$). Then x contains a term with $\cos 2\phi$, which takes into account the difference of the trajectory from the standard "hypocycloid picture," and even though the difference is small it must be taken into account if we wish to obtain Eq. (1) for the average velocity along the binormal. This formula, as is well known, describes the "banana" trajectories of trapped particles, which are responsible for the main energy losses in tokamaks, the most promising setups for solving the problem of controlled thermonuclear fusion.

In numerical calculations of trajectories with a "small but not too small" Larmor radius, it is helpful to compare the results with our Eqs. (11) and (17), which are more accurate than the formulas of the stand drift approximation. It should also be noted that the conservation of the invariant $\mu \sim v_\perp^2/B$ in our method is simply a consequence of the semiclassical nature of the solution for x in first order:

$$\langle x_1^2 \rangle \omega = a_0^2/2 = \text{const}.$$

ACKNOWLEDGMENTS

I am grateful to V.D. Shafranov for valuable remarks.

REFERENCES

1. B. A. Trubnikov, *Theory of Plasma* (Énergoatomizdat, Moscow, 1996).

Translation was provided by AIP

Journal of Experimental and Theoretical Physics, Vol. 91, No. 3, 2000, pp. 483–487.
Translated from Zhurnal Éksperimental'noĭ i Teoreticheskoĭ Fiziki, Vol. 118, No. 3, 2000, pp. 554–559.

PLASMA, GASES

On the Charge of Dust Particles in a Low-Pressure Gas Discharge Plasma

A. V. Zobnin, A. P. Nefedov, V. A. Sinel'shchikov*, and V. E. Fortov

Institute of Thermal Physics of Extremal States, Russian Academy of Sciences, Moscow, 127412 Russia
**e-mail: vladimir@hedric.msk.su*

Received March 29, 2000

Abstract—Self-consistent molecular-dynamics calculations of the charge of micron-size particles in a low-pressure gas-discharge plasma are performed. It is shown that charge exchange of ions on neutrals starts to affect the charge of dust particles at pressures corresponding to ion mean free paths much greater than the Debye radius. The computational results show that the potential of a particle depends nonmonotonically on the pressure and on the particle size. © *2000 MAIK "Nauka/Interperiodica".*

1. INTRODUCTION

One of the most important parameters determining the properties of a dust plasma is the charge of the dust particles. However, at present there are no direct methods for measuring the charge Q of macroparticles. In [1–3] measurements of the resonance frequency of vertical oscillations of macroparticles, levitating in the near-electrode layer of a capacitive-type radio-frequency discharge, were used to determine Q. The oscillations were excited either with low-frequency modulation of the voltage [1, 2] or with modulated laser radiation [3] illuminating one of the particles. The product Qn_i, where n_i is the ion density in the near-electrode layer, is determined from the experiment. The ion density is not measured in the experiment; it is obtained by extrapolating electron density measurements performed in the quasineutral region. In [4] the quantity Q is extracted by analyzing the trajectories of two interacting macroparticles. However, the result is sensitive to the screening length, which is determined simultaneously with Q from the same measurements. Thus, in the experiments the charge of macroparticles is determined in combination with another unknown parameter. Naturally, this affects the accuracy of the measurements of Q. Measurements performed in a stratified global discharge also suffer from similar drawbacks [5].

The charge of spherical particles with radius a is often estimated from measurements of the floating potential φ_{fl}, assuming that

$$Q = a\varphi_{fl}.$$

In so doing, the fact that the conditions under which levitation of particles is observed in a gas-discharge plasma of a particular type are such that the floating potential dependence on the probe size is neglected. Consequently, the quantity φ_{fl} determined by probe methods cannot be identified with the surface potential of a dust particle.

It should also be noted that there are no reliable published experimental data that would make it possible to judge the dependence of the charge of macroparticles on either their size or the plasma parameters.

The limited orbits model (OML) [6], which is valid in the collisionless case, is often used to determine the charge of dust particles in the theoretical description of the properties of a dust plasma and in calculations of the parameters of dust structures. However, under the conditions of most experiments on dust plasma the collisions of ions with neutrals have a large effect on the ion flux onto a dust particle and therefore on the charge of the particle also. It is shown in [7] that under the conditions of the near-electrode later of an hf discharge at the typical experimental pressures the ions which have undergone collisions with charge transfer near a dust particle make the main contribution to the ion flux onto a particle. Under conditions for which the formation of dust structures was observed in the positive column of a dc glow discharge [8] and in an hf induction-type discharge [9], the velocity of the directed motion of ions is lower than the thermal velocity and the characteristic scale of variation of the parameters of an unperturbed plasma by the particles is much larger than the Debye radius. In this connection it is of interest to investigate the question of the influence of ion–neutrals collisions on the charging of an isolated dust particle in a spatially uniform, isotropic plasma. It is not difficult to estimate the ion mean free path for which the collisionless theory no longer works. For simplicity, we shall consider a quite small dust particle with radius

$$a \ll R_D T_i/T_e,$$

where

$$R_D = \left[4\pi e^2\left(\frac{n_e}{T_e} + \frac{n_i}{T_i}\right)\right]^{-1/2} = \left[\left(\frac{1}{R_{De}}\right)^2 + \left(\frac{1}{R_{Di}}\right)^2\right]^{-1/2}$$

1063-7761/00/9103-0483$20.00 © 2000 MAIK "Nauka/Interperiodica"

is the Debye radius, T_e, n_e, T_i, and n_i are, respectively, the electron and ion temperatures and densities, and e is the electron charge. For ions surrounding a negatively charged particle with surface potential φ_s there exists a potential well, and at distances less than $\rho = ae|\varphi_s|/T_i$ the depth of this well is greater than the thermal energy of an ion. If an ion passing near a particle has undergone a collision with charge transfer inside a sphere of radius ρ, then such an ion has little chance of escaping from the potential well. Most such ions settle on the particle. To estimate the additional collisional ion flux onto a particle the ion flux through a sphere of radius ρ must be multiplied by the probability of charge exchange on the ions passing through the indicated sphere:

$$J \sim \pi\rho^2 v_i n_i \rho/\lambda_i,$$

where v_i is the average thermal velocity of the ions and λ_i is the ion mean free path.

Let us compare this flux with the ion flux onto a particle as determined according to the Langmuir formula in the OML model:

$$J_{OML} = \pi a^2 v_i n_i (1 + e|\varphi_s|/T_i).$$

Comparing shows that the ion flux due to collisions becomes of the water of the ion flux in the OML approximation when

$$\lambda_i \sim a(\varphi_s e/T_i)^2.$$

Since $e|\varphi_s| \sim T_e$, under conditions typical for a glow discharge, $T_e/T_i \sim 100$, the quantity λ_i can be much greater than not only the particle radius but also of the Debye radius ($R_D \approx R_{Di} \sim 0.08$ mm). Thus, for $a = 1$ μm collisionless have a strong effect even for $\lambda_i \sim 10$ mm, which corresponds to a pressure of the order of 1.3 Pa. Much higher pressures were used in the experiments of [8, 9]. Consequently, the calculation of the charge of dust particles under conditions intermediate between the collisionless and hydrodynamic limits is an urgent problem. Such a calculation cannot be performed analytically and requires numerical methods.

2. DESCRIPTION OF THE MATHEMATICAL MODEL

The calculation of the charge of a dust particle was performed under assumptions which are nearly always satisfied under the conditions of experiments with a gas-discharge dust plasma. Specifically, an isolated spherical particle was considered; the particle surface absorbed all ions and electrons incident on it and it did not emit charged particles; the plasma was assumed to consist of electrons and singly charged positive ions in their proper gas; the electron mean free path was much greater than the particle size, so that electron losses could be neglected and the equilibrium distribution could be used for the spatial distribution of the electrons; ion–neutrals collisions were described using the cross-section for resonance charge transfer, which plays the main role in the relaxation of the momentum of neon ions in neon gas [10]. The ion–ion and ion–electron collisions where neglected, since the degree of ionization is low. In addition, it was assumed that the unperturbed ion and electron distribution functions are Maxwellian with temperatures T_e and T_i, respectively, and the ion temperature was assumed to be the equal to the temperature of the neutral gas, whose perturbation by the ions was neglected.

The spatial distribution of the ions and the ion flux onto a particle were calculated by direct numerical simulation of the motion of each ion in the region under study in a self-consistent electric field.

The computational procedure consisted of the following. The electron density distribution $E(R)$ and the electric field $n_e(R)$ were calculated for a given initial charge of a dust particle and a given ion distribution in a sphere of radius L (which is the computational cell). For this, the following system of equations was solved numerically:

$$n_e(R) = n_{e0}\exp(e\varphi(R)/T_e), \quad (1)$$

$$E(R) = -e\left(4\pi\int_a^R n_e(r)r^2dr - N_i(R) - Q\right)/R^2, \quad (2)$$

$$\varphi(R) = \int_R^L E dr, \quad (3)$$

where $N_i(R)$ is the number of ions in a sphere with radius R and Q is the charge of a dust particle. The unperturbed ion density n_{i0} was given and the unperturbed electron density n_{e0} was chosen so as to ensure electrical neutrality of the computational cell:

$$4\pi\int_a^L n_e(r)r^2dr - N_i(L) - Q = 0. \quad (4)$$

The potential on the outer boundary of the computational cell was assumed to be zero.

The spatial distribution of the ions and the ion flux onto a particle were calculated by the molecular-dynamics method for the electric field and electron density distribution obtained in this manner. A condensing time grid was used as the particle was approached. The inflow of ions from outside the system was given by the flux

$$\frac{dN_i}{dt} = \pi L^2 n_{i0}\left(\frac{8T_i}{\pi m_i}\right)^{1/2}, \quad (5)$$

where m_i is the ion mass. The radial components of the momentum p and the absolute magnitude of the angular

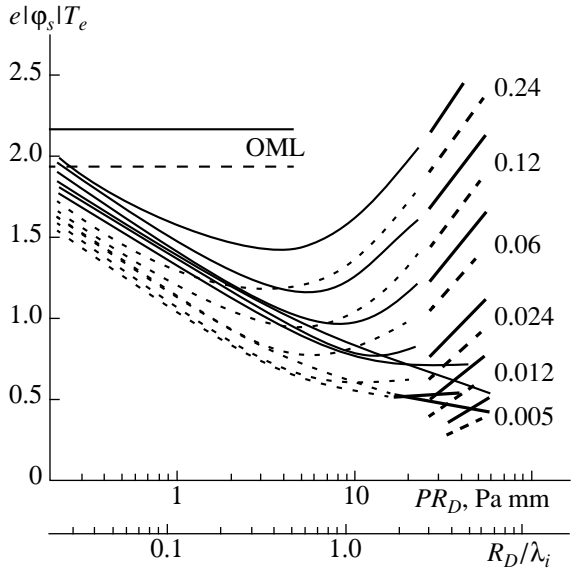

Fig. 1. The normalized surface potential of a particle versus the parameter PR_D (or R_D/λ_i), calculated for particles of various sizes (a/R_D = 0.24, 0.12, 0.06, 0.24, 0.012, and 0.005) for T_e = (solid curves) 2.5 and (dashed curves) 5 eV. The OML approximation is shown by horizontal straight lines, hydrodynamic approximation is displayed by heavy lines.

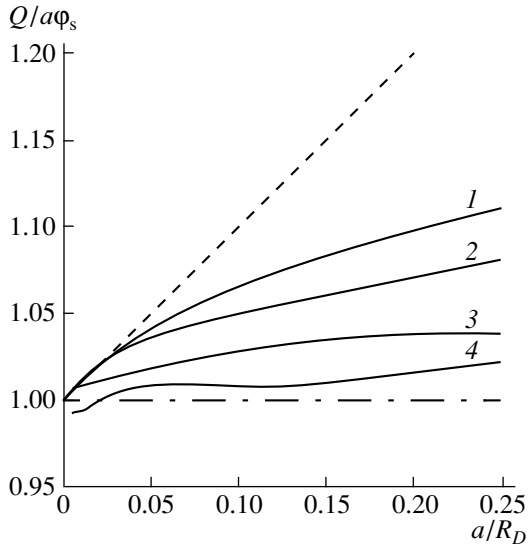

Fig. 2. Relation between the surface potential and charge of a particle for T_e = 2.5 eV: R_D/λ_i = (*1*) 0.125, (*2*) 0.6, (*3*) 1.25, and (*4*) 2.5. The dashed curve corresponds to the Debye potential and the dash–dotted curve corresponds to the Coulomb potential.

momentum M of the ions flowing in from the outside were random numbers with the distributions

$$W(p) \propto p\exp(-p^2/2m_iT_i), \tag{6}$$

$$W(M) \propto M\exp(-M^2/2L^2m_iT_i), \tag{7}$$

which corresponded to a Maxwell distribution of the ions on the outer boundary of the cell.

The algorithm for taking account of charge exchange of ions on neutrals consisted of following. For each ion the probability that the ion will collide with an atom having random values of the radial component of the momentum and angular momentum in a time equal to the integration time step Δt was determined (the velocity distribution of the atoms was assumed to be Maxwellian). At the end of each integration time step a die was rolled, and in accordance with the previously computed probability the radial component of the momentum and the angular momentum of the ion either acquired values equal to the corresponding values for the atom or remained unchanged. In the calculations, the velocity dependence of the cross-section for resonance charge transfer was used in the form [11]

$$\sigma = \sigma_0(\ln(b/v))^2,$$

where σ_0 and b are constants. Despite the logarithmic character of this dependence, it is important to it into account because the average velocity of ions close to a dust particle can be an order of magnitude higher than the thermal velocity.

The ion flux I_e^P onto a particle calculated in this manner, together with the electron flux I_e^P onto a particle calculated using the formula

$$I_e^P = \pi a^2\left(\frac{8T_e}{\pi m_e}\right)^{1/2} n_{e0}\exp\left(\frac{e\varphi_s}{T_e}\right), \tag{8}$$

where $\varphi_s = \varphi(a)$ is the surface potential of the particle, were used to correct the particle charge:

$$Q^* = Q + (I_e^P - I_i^P)\Delta t. \tag{9}$$

Then the entire procedure was repeated. The calculation was terminated when the particle charge became stationary, which is equivalent to the ion and electron fluxes onto a particle being equal to one another.

3. COMPUTATIONAL RESULTS

The calculations were performed for a neon plasma with two electron temperatures T_e = 2.5 and 5 eV. The ion temperature was assumed to be T_i = 0.025 eV, and the ion density at the outer boundary of the computational cell was $n_{i0} = 2 \times 10^8$ cm^{-3}. The radius L of the computational cell was chosen so that the ion flux onto a particle was much less than the random ion flux (5) directed into the computational cell. The radius of a dust particle was varied in the range 1–20 μm, i.e.,

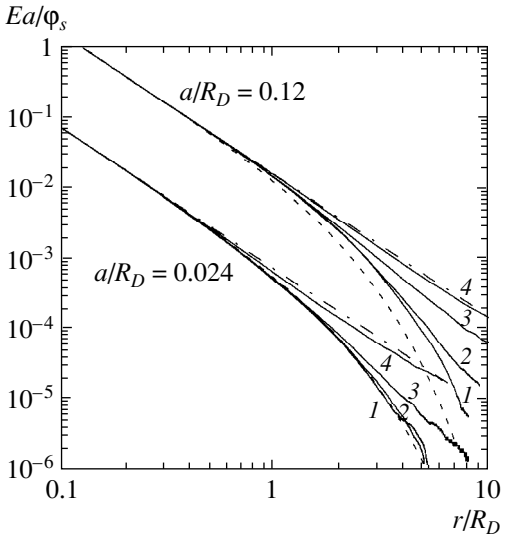

Fig. 3. Electric field distribution near a particle with T_e = 2.5 eV: R_D/λ_i = (*1*) 0.025, (*2*) 0.125, (*3*) 0.6, and (*4*) 2.5.

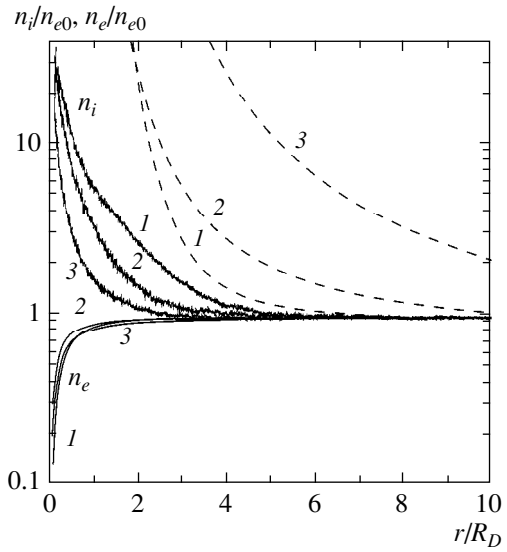

Fig. 4. Ion and electron density profiles near a particle with T_e = 2.5 eV and a/R_D = 0.12: R_D/λ_i = (*1*) 0.025, (*2*) 0.6, (*3*) 2.5. Broken curves are equilibrium ion density distribution.

approximately (0.012–0.24)R_D. The neon pressure was varied from P = 0 up to 270 Pa. The computational results are presented in Figs. 1–4. It should be noted that for an ideal plasma the surface potential of a particle depends on the density of charged particles only through the Debye radius and is determined by the ratios of the particle radius and the ion mean free path to the Debye radius. The method for representing the results was chosen on the basis of these considerations. We also note that this fact was used to test the numerical model. Specifically, a calculation was performed for $n_{i0} = 8 \times 10^8$ cm^{-3}.

Figure 1 shows, aside from the results of the numerical calculation performed using the model described above, the dependences corresponding to the OML model and the hydrodynamic approximation for ions. The hydrodynamic approximation is valid if the ion mean free path $\lambda_i \ll R_D$. In addition, if a dust particle is much smaller than the electron Debye radius (under the conditions considered $R_{De} \sim 0.8$ mm), then the ion current onto a particle is determined by the formula

$$I_i \approx 4\pi a|\varphi_s| b_i n_{i0}, \qquad (10)$$

where b_i is the ion mobility. If the pressure is still not too high and Eq. (8) is valid for the electron current, then the balance of the ion and electron fluxes onto a particle leads to the equation

$$|\varphi_s| = a\left(\frac{T_e}{2\pi b_i^2 m_e}\right)^{1/2} \exp\left(\frac{e\varphi_s}{T_e}\right). \qquad (11)$$

It is evident from Fig. 1 that for $PR_D > 30$ Pa mm, i.e., when $R_D/\lambda_i > 3$, the computational results agree well with the computational results obtained in the hydrodynamic approximation. At the same time, in the collisionless limit ($PR_D \longrightarrow 0$) the quantity $e|\varphi_s|/T_e$ tends to the value calculated in the OML model (see Fig. 1). The relation between the surface potential and charge of a particle is presented in Fig. 2. We note that for the range of values of the parameter a/R_D studied, the relation between Q and φ_s differs negligibly from the relation corresponding to a vacuum, and this difference decreases with increasing value of the parameter R_D/λ_i.

It is evident from the data presented in Fig. 1 that ion–neutrals collisions start to affect the potential of micron-size dust particles already at pressures of the order of several Pa, which corresponds to ion mean free paths of the order of several mm, i.e., much greater than the Debye radius. As long as the mean free path remains larger than the Debye radius, increasing pressure results in a higher ion current and a lower particle charge. In the hydrodynamic regime, as pressure increases, the particle charge once again starts to increase, since the ion mobility decreases. Collisions also cause the surface potential of a particle to depend weakly on the particle size. In the hydrodynamic limit the surface potential of a particle increases monotonically (in modulus) with increasing particle radius (approximately logarithmically). As follows from Fig. 1 and estimates of the ion flux which are presented in the Introduction, the opposite trend is observed at low pressures for very small particles ($a/R_D \ll T_i/T_e$), since for them the ion flux is proportional to the cubed particle radius, while the electron flux is proportional to the squared radius. For particles with $a/R_D \geq T_i/T_e$ (see Fig. 1) the absolute value of the surface potential increases with the particle size, but this dependence becomes weaker as the particle radius and pressure decrease.

It follows from the electric field distributions, presented in Fig. 3, near a particle that as the pressure or the value of the parameter R_D/λ_i increases, a transition occurs from the Debye to the Coulomb potential because the screening effect decreasing. Comparing the ion and electron density distributions, presented in Fig. 4, near a particle also attests to the fact that the screening effect decreases as the parameter R_D/λ_i increases. In turn, the perturbation of the Boltzmann distribution of the ion density (see Fig. 4) is due to be ion losses on the particle. As the particle size increases, this perturbation increases. As a result, for large particles the deviations from the Debye potential begin to appear for lower values of the parameter R_D/λ_i (see Fig. 3). The data presented in Fig. 2 also attest to this.

In summary, it was shown above that at pressures corresponding to ion mean free paths much greater than the Debye radius the limited orbits approximation for calculating the charge of dust particles levitating in a gas-discharge plasma leads to large errors. The pressure dependence of the surface potential of the particles is nonmonotonic, and for neon plasma it has a minimum at pressure ~100 Pa.

REFERENCES

1. A. Melzer, Th. Trottenberg, and A. Piel, Phys. Lett. A **191**, 301 (1994).
2. Th. Trottenberg, A. Melzer, and A. Piel, Plasma Sources Sci. Technol. **4**, 450 (1995).
3. A. Homann, A. Melzer, and A. Piel, Phys. Rev. E **59**, 3835 (1999).
4. U. Konopka, L. Ratke, and H. M. Thomas, Phys. Rev. Lett. **79**, 1269 (1997).
5. V. I. Molotkov, M. Y. Poustylnik, and V. M. Torchinsky, in *Proceedings of the International Conference on Physics of Clusters "Clusters in Plasma and Gases", Pushchino, Russia, 1999*, p. 51.
6. V. N. Tsytovich, Usp. Fiz. Nauk **167**, 57 (1997) [Phys. Usp. **40**, 53 (1997)].
7. V. A. Shveĭgert, I. V. Shveĭgert, V. M. Bedanov, *et al.*, Zh. Éksp. Teor. Fiz. **115**, 877 (1999) [JETP **88**, 482 (1999)].
8. A. M. Lipaev, V. I. Molotkov, A. P. Nefedov, *et al.*, Zh. Éksp. Teor. Fiz. **112**, 2030 (1997) [JETP **85**, 1110 (1997)].
9. Yu. V. Gerasimov, A. P. Nefedov, V. A. Sinel'shchikov, and V. E. Fortov, Pis'ma Zh. Tekh. Fiz. **24** (19), 62 (1998) [Tech. Phys. Lett. **24**, 774 (1998)].
10. Yu. P. Raĭzer, *The Physics of Gas Discharge* (Nauka, Moscow, 1987).
11. E. E. Nikitin and B. M. Smirnov, *Slow Atomic Collisions* (Énergoizdat, Moscow, 1990).

Translation was provided by AIP

Journal of Experimental and Theoretical Physics, Vol. 91, No. 3, 2000, pp. 488–496.
Translated from Zhurnal Éksperimental'noĭ i Teoreticheskoĭ Fiziki, Vol. 118, No. 3, 2000, pp. 560–569.

FLUIDS

Surface Anchoring and Temperature Variations of the Pitch in Thin Cholesteric Layers

V. A. Belyakov[a, *] and E. I. Kats[a, b]

[a] *Landau Institute for Theoretical Physics, Russian Academy of Sciences, Chernogolovka, Moscow oblast, 142432 Russia*

[b] *Laue-Langevin Institute, F-38042, Cedex 9, Grenoble, France*

**e-mail: bel@cpd.landau.ac.ru*

Received April 24, 2000

Abstract—The temperature variations of the cholesteric pitch in thin planar layers of cholesterics and their dependence on the surface anchoring force are investigated theoretically. It is shown that the temperature variations of the pitch in a layer are of a universal character. This is manifested in the fact that they depend not separately on the parameters of the sample but only on one dimensionless parameter $S_d = K_{22}/dW$, where K_{22} is the torsional modulus in the Frank elastic energy, W is the height of the surface-anchoring potential, and d is the thickness of the layer. The investigation is performed the parameter S_d in a range where the change per unit number of cholesteric half-turns within the thickness of the layer accompanying a change in the temperature is due to the slipping of the director on the surface of the layer through the potential barrier for surface anchoring. The critical values of the parameter S_d (which are most easily attained experimentally by varying the thickness of the layer), determining the region of applicability of the approach employed, are presented. The temperature variations of the free energy of the layer and the pitch of the cholesteric helix in the layer as well as the temperature hysteresis in the variations of the pitch with increasing and decreasing temperature are investigated for the corresponding values of S_d. Numerical calculations of the quantities mentioned above are performed using the Rapini anchoring potential. © *2000 MAIK "Nauka/Interperiodica"*.

1. INTRODUCTION

Interest in detailed investigations of chiral liquid crystals (LC) has increased in recent years (see, for example, [1]). This interest is partially due to the general-physical problem (not solved thus far) of choosing the chiral order parameters characterizing chiral systems (see, for example, the review in [2]). Such investigations are also of interest because the many additional advantages of chiral LC over the nematic LC, which are ordinarily used, have not been completely realized in applications. Since cells of chiral, specifically, cholesteric, LC are now widely used as sensors of various kinds as well as for information display devices controlled by optical transparencies and for many other purposes, it is extremely important to study the properties of these LC in bounded geometries. Investigations of the optical characteristics of thin layers of chiral LC yield important information (for applications and for understanding the physics of LC) about changes in the structure of a LC in thin layers and about the dynamics of these changes.

In [3, 4] the temperature behavior of the cholesteric pitch in thin planar cholesteric layers was investigated by measuring their optical transmission spectra, which on the basis of a well-developed theory of the optical properties of chiral liquid crystals [5, 6] were interpreted in terms of the temperature variations of the parameters of the cholesteric in the layer. An unusual temperature behavior of the transmission spectra of light with wavelength of the order of the pitch of the cholesteric helix and temperature hysteresis in abrupt changes of the pitch were attributed to the deviations of the director on the surface of the sample away from the direction of alignment in the potential well of the surface-anchoring forces and the abrupt transitions of the cholesteric helix in the layer between configurations differing from one another by one half-turn of the helix in the layer. Comparing the results with the theory of temperature variations of the pitch in cholesteric layers, which was developed on the basis of the continuum theory of elasticity taking account of the surface-anchoring forces, showed that the parameters of the experimental samples were such that the mechanism of the temperature jump in the pitch did not correspond to a transition between configurations of the helix with the number of half-turns differing by one by means of the director overcoming on the surface the barrier in the surface anchoring potential. Specifically, the measured angles of deflection of the director on the surface of the layer from the direction of alignment for the temperature of the jump [3, 4] were much smaller than the critical (see below) value of this angle. Consequently, the mechanism of the jump in the pitch and the reasons why a superposition of two spectra corresponding to configurations in which the number of half-turns of the helix in the layer differs by one are present in the transmission spectra remained unknown. This requires a fur-

ther study of the question, specifically, an analysis of the situations for which a reliable theoretical description is known, especially since the pitch jumps appear in various precise investigations of layers of chiral liquid crystals, for example, in the nonlinear generation of high optical harmonics [7] and investigations with a Fabry–Perot interferometer [8]. The present paper is devoted to an analysis of one such situation, specifically, the study of the temperature variations of the cholesteric pitch in thin layers. The conditions under which the simple continuum theory of elasticity, taking account of the surface-anchoring forces, is applicable are found. For these conditions a universal description of the temperature variations of the pitch in the layer is proposed, and a theory of the temperature hysteresis of jumps in the pitch is given and experimentally observable effects are found.

2. BASIC EQUATIONS

We shall consider the temperature behavior of the pitch of a helix in a thin planar cholesteric layer, assuming the surface-anchoring forces to be identical on both surfaces of the layer and assuming the alignment axis to be the same on both surfaces. We shall use for this the expression for the free energy in the form [9]

$$F(T) = 2W_s(\varphi) + \frac{K_{22}d}{2}\left(\frac{2\pi}{p_d(T)} - \frac{2\pi}{p(T)}\right)^2, \qquad (1)$$

where K_{22} is the torsional elastic modulus, $W_s(\varphi)$ is the surface anchoring potential, d is the thickness of the layer, $p(T)$ is the equilibrium value of the pitch of the cholesteric helix for temperature T in a bulk cholesteric, $p_d(T)$ is the pitch at the same temperature in a layer of thickness d, and φ is the angle of deflection of the director on the surface of the layer from the direction of alignment.

The formula (1) requires comment. The point is that the properties of a deformed cholesteric depend strongly on the ratios of the nonuniformity scale and the pitch of the helix. On scales much less than the pitch of the helix the cholesteric has the same properties as a nematic. In the opposite limit the elastic properties of the cholesteric are equivalent to those of a smectic. Consequently, the expression for the "quasi-smectic" energy (1) is valid for cells with thickness much greater than the pitch of the cholesteric helix when the surface anchoring ensures that the director is not tilted away from the normal to the cholesteric axis for the equilibrium configuration of the LC in the layer.

Since the pitch $p_d(T)$ in the layer is uniquely related with the angle φ and the equilibrium pitch $p(T)$ is related with $\varphi_0(T)$—the angle of deflection of the director on the surface of the layer away from the direction of alignment in the absence of surface anchoring, the expression for the free energy can be rewritten as a function of these angles:

$$F(T) = 2W_s(\varphi) + (2K_{22}/d)[\varphi - \varphi_0(T)]^2. \qquad (2)$$

However, φ and $\varphi_0(T)$ are interrelated and the relation between them can be found by minimizing the free energy (2). This gives an equation for the angle φ as a function of temperature:

$$\frac{\partial W_s(\varphi)}{\partial \varphi} + \frac{2K_{22}}{d}[\varphi - \varphi_0(T)] = 0. \qquad (3)$$

Using Eqs. (2) and (3), the expression for the free energy can be represented in the form

$$F(T) = 2W_s(\varphi) + \left(\frac{\partial W_s(\varphi)}{\partial \varphi}\right)^2 \frac{d}{2K_{22}}, \qquad (4)$$

where the temperature dependence of the angle φ of deflection of the director from the direction of alignment on the surface of the layer is determined by Eq. (3). In what follows, we shall assume for simplicity (and also for making estimates) that the surface anchoring potential $W_s(\varphi)$ is determined only by one characteristic energy W (depth of the potential well).

We note that dividing Eqs. (2)–(4) by the depth W of the surface potential we find that the depth of the surface anchoring potential, the elastic modulus, and the thickness of the layer enter in these equations only through the dimensionless parameter $S_d = K_{22}/dW$ (we note that the parameter S_d differs only by a factor from the parameter $S = 4\pi(d/p)S_d$ used in [3, 4]). Thus, for a fixed value of S_d the temperature dependence of the orientation of the director on the surface (the angle φ) is determined only by the form of the anchoring potential and the temperature dependence of the equilibrium pitch $p(T)$, i.e., $\varphi_0(T)$. The temperature dependence of φ and the free energy, which follow from the equations presented, for a fixed value of the parameter S_d will be universal, i.e., it will not depend on the number of half-turns of the cholesteric helix over the thickness of the layer (or, which is the same thing, it does not depend on the thickness of the layer).

The potential barrier between the configurations of the helix that differ by one half-turn also does not depend on the thickness of the layer and is determined only by the parameter S_d. The barrier height depends on the temperature (through the function $\varphi_0(T)$) and is determined by the expression

$$B(T, S_d) = F(\varphi_c, \varphi_0(T), S_d) - F(\varphi(T), \varphi_0(T), S_d), \qquad (5)$$

where φ_c is a critical angle at which an abrupt change occurs in the configuration of the cholesteric helix in the layer; the equilibrium value of the angle $\varphi(T)$ is determined by the solution of Eq. (3), and the free energy $F(\varphi, \varphi_0(T), S_d)$ is determined by Eq. (2). As should be the case, when the angle $\varphi(T)$ reaches the critical value the barrier height becomes zero and a transition between

configurations differing from one another by one half-turn of the helix occurs abruptly in the layer.

The value φ_c of the pitch in the layer immediately before the jump (of course, the angle $p_d(T_c)$ of free rotation of the director, i.e., the pitch T_c of a helix in a bulk cholesteric, can also be found from the temperature $p(T_c)$) can be expressed in terms of the critical angle $\varphi_0(T_c)$, whose value is determined by the form of the anchoring potential:

$$p_d(T_c) = \frac{2d}{N + 2\varphi_c/\pi}, \quad p(T_c) = \frac{2d}{N + 2\varphi_0(T_c)/\pi}, \quad (6)$$

where N is the number of half-turns of the helix over the thickness of the layer in the initial equilibrium configuration of the helix, T_c is the temperature of the pitch jump, and the free-rotation angle of the helix for this temperature of the jump is given by the formula

$$\varphi_0(T_c) = \varphi_c + \left(\frac{\partial W_s(\varphi)}{\partial \varphi}\right)_{\varphi = \varphi_c} \frac{1}{2WS_d}. \quad (7)$$

Thus, if the form of the surface anchoring potential and the temperature behavior of the pitch of the helix in a bulk cholesteric are known, then the relations presented above determine the temperature behavior of the pitch of the helix in a layer of finite thickness. These same formulas can be used to solve the inverse problem, specifically, to reconstruct the surface anchoring potential on the basis of the experimentally found temperature behavior of the pitch of the helix in a layer of finite thickness.

3. TEMPERATURE HYSTERESIS OF PITCH VARIATIONS IN A LAYER

To describe the temperature variations of the pitch in a layer it is necessary to know the specific form of the surface anchoring potential $W_s(\varphi)$ [10, 11]. However, the qualitative character of the temperature variations of the pitch in a cholesteric layer can be determined even without specifying the form of this potential. Since we are interested in investigating the temperature dependence of the pitch in the layer, specifically, determining the temperatures of the pitch jumps that correspond to changes in the number of half-steps of the cholesteric spiral in a layer by one, we recall the results of the corresponding analysis in [3, 4]. On the surface of the layer the director is oriented in the direction of alignment at a temperature for which an integer number of half-pitches $p(T)$ of the helix of a bulk cholesteric fit within the thickness of the layer, i.e., $\varphi_0(T) = 0$. When the temperature deviates from this value $\varphi_0(T)$ and $\varphi(T)$ become different from zero. The temperature for which $\varphi_0(T) = \pi/4$ corresponds to the situation where the free energies of both configurations of the helix which differ by one half-turn over the thickness of the layer are the same, but because of the presence of a potential barrier between these configurations, which is due to surface anchoring, the transition between them occurs not at this temperature point but with a further variation of the temperature. Thus, temperature hysteresis is observed in the transition, i.e., the transition temperatures are different for different directions of variation of the temperature. However, the temperature determined by the condition $\varphi_0(T) = \pi/4$ would correspond to the point of a pitch jump at a transition between configurations of the helix in the absence of hysteresis. In reality, however, because of hysteresis the pitch jump occurs at a temperature determined by the condition $\varphi(T) = \varphi_c$, where φ_c is the critical angle at the pitch of the helix changes abruptly.

Using the natural assumption that the potential $W_s(\varphi)$ is an even function of φ—the angle of deflection of the director from the alignment direction on the surface of the layer—it is easy to conclude that the solution of Eq. (3) for φ is an odd function of $\varphi_0(T)$ and is bounded by the values $\pm\varphi_c$. For example, when the angle φ reaches one of these limiting values φ_c the pitch of the cholesteric helix in the layer changes abruptly, and the angle φ also assumes a new value abruptly. If the jump point corresponds to the angle of free rotation of the director $\varphi_0(T_+)$, then the angle of free rotation of the director determining the value of φ after the jump is $\varphi_0(T_+) - \pi/2$ but in a new configuration of the helix differing by one half-turn over the thickness of the layer. For temperature varying in the opposite direction the pitch jump occurs when the angle φ reaches a different limiting value $-\varphi_c$, and the corresponding angle of free rotation is $\varphi_0(T_-) = -\varphi_0(T_+)$ (where T_- is the temperature corresponding to the jump for this direction of temperature variation). The latter relation results in the following relation between the values of the free pitches of the helix (temperatures) corresponding to a jump with opposite directions of temperature variation [4]:

$$\frac{1}{p(T_+)} + \frac{1}{p(T_-)} = \left(N + \frac{1}{2}\right)\frac{1}{d},$$
$$d\left(\frac{1}{p(T_+)} - \frac{1}{p(T_-)}\right) \quad (8)$$
$$= \left(\frac{\partial W_s(\varphi)}{\partial \varphi}\right)_{\varphi = \varphi_c} \frac{1}{\pi WS_d} + \frac{1}{2}\left(\frac{4\varphi_0}{\pi} - 1\right),$$

where N is the number of half-pitches in a layer for the initial configuration of the helix (before the jump at temperature T_+ and after the jump at the temperature T_-).

The first relation in Eqs. (8) is universal and at first glance in no way depends on the surface anchoring forces. In reality, the jump temperatures T_+ and T_- depend on the surface anchoring, more accurately, on the parameter S_d. The relation itself determines the relation, which does not depend on the specific form of the surface potential, between the pitch jump temperatures with opposite directions of temperature variation, i.e., temperature hysteresis in the pitch variations, and it has

predictive power, which can be checked experimentally. Specifically, if the pitch at the jump point is measured experimentally for a certain direction of temperature variation, then the pitch at the jump point with the opposite direction of temperature variation can be calculated using Eq. (8) and checked experimentally. Specifically, it follows from Eq. (8) that in the limit of vanishingly weak surface anchoring (i.e., $T_+ \longrightarrow T_-$) the pitch at the jump point is $d/(N + 1/2)$, as should be [3], and there is no hysteresis. It follows from the same formula that as the thickness of the layer increases, the hysteresis decreases (more accurately, the hysteresis decreases for fixed S_d with increasing N).

4. CRITICAL THICKNESS OF THE LAYER

We shall now discuss in somewhat greater detail the conditions of applicability of the mechanism, discussed here, for the jump in the pitch of the helix due to slipping of the director on the surface through the barrier in the surface anchoring potential. Since $\varphi_c = \pi/4$ in reasonable models of surface anchoring potentials [3, 4], we shall assume below that the jump condition is $\varphi(T) = \pi/4$, i.e., on the surface of the layer the director turns from the alignment direction by the angle $\pi/4$ when anchoring forces are present. Depending on the surface anchoring force, more precisely, the value of the parameter S_d, the free rotation angle $\varphi_0(T)$ of the director can exceed $\pi/2$ or remain less than this value. In the first case the helix is "overwound," i.e., the equilibrium configuration of the helix and the configuration under study differ by more than one half-turn. Consequently, the pitch jump accompanying a change in the number of half-turns of the helix by one is accompanied by a transition into not an equilibrium but rather a metastable configuration of the helix. In addition, depending on the anchoring force, the number of such metastable states, which differ by 1, 2, and so on half-turns from the equilibrium configuration of the helix in the layer, can be quite large. Thus, for infinitely strong surface anchoring the number of such states is infinite, and the transition into the equilibrium state by slipping of the director through the surface anchoring barrier cannot occur at all. This means that in this case the mechanism of relaxation of the configuration of a helix to its equilibrium state accompanying temperature variations is not the same as slipping of the director on the surface through the surface anchoring barrier, for example, the mechanism of fluctuation formation of a defect in the interior volume of the layer, giving rise to further relaxation of the director distribution in the layer to the equilibrium configuration.

The simplest form of such defects are the so-called χ lines [9] in cholesterics (analogous to dislocations in smectic-type layered structures). A change in the number of half-turns of the cholesteric helix and correspondingly "shedding" of excess free energy can occur when such a dislocation loop nucleates and grows (in the central part of the layer, where the elastic stresses are weakest under conditions of strong surface anchoring). The nucleation and growth of a dislocation loop are determined by the competition between two factors: loss in linear tension because of the appearance of a singular line $2\pi Rt$ (R is the radius of the dislocation loop and $t \approx K_{22}$ is the linear tension) and gain in surface energy inside the dislocation loop $\pi R^2 bB/(Ad)$ (here B is the potential barrier given by Eq. (5), b is Burgers vector (of the order of the pitch of the helix), and A is the surface area of the layer). Minimizing both of these contributions gives the activation energy of the process $E_a = \pi t^2 Ad/(bB)$, and in order for the dislocation mechanism to be efficient the value of E_a must not be too large compared with k_BT (k_B is Boltzman's constant).

On this basis it is helpful to introduce the concept of the critical thickness of a layer or the critical value of the parameter S_d, which determine the boundary value of the thickness or S_d for which the jump in the pitch of the helix in the mechanism of slipping through a barrier with temperature variations is accompanied by a transition of the helix into an equilibrium configuration. The critical values of the thickness of the layer and S_{dc} are determined by Eq. (3), if $\varphi = \varphi_c$ and $2(\varphi_c - \varphi_0(T)) = -\pi$ there, which gives

$$\left(\frac{\partial W_s(\varphi)}{\partial \varphi}\right)_{\varphi = \varphi_c} - \frac{\pi K_{22}}{d} = 0. \qquad (9)$$

As follows from Eq. (9), the critical values depend on the explicit form of the surface potential. For definiteness we shall use here and below the critical values in the Rapini potential (see, for example, [9–11]): $W_s(\varphi) = -(W/2)\cos^2\varphi$, for which $\varphi_c = \pi/4$. Then the critical thickness d_c and the parameter S_{dc} are determined by the formulas

$$d_c = \frac{2\pi K_{22}}{W}, \quad S_{dc} = \frac{K_{22}}{Wd_c} = \frac{1}{2\pi}. \qquad (10)$$

Thus, for layer thicknesses less than d_c or $S_d > S_{dc}$ metastable configurations of the helix in the layer are not drawn into the temperature jump in the pitch and the temperature variations of the pitch can be described on the basis of the mechanism of slipping of the director on the surface through the anchoring potential barrier.

5. PITCH VARIATIONS IN THE RAPINI POTENTIAL

The relations presented above can be used to reconstruct the anchoring potential from the experimental measurements of the temperature variations of the pitch in cholesteric layers. However, as demonstrated above, some regularities of the corresponding dependences are insensitive to the specific form of the surface anchoring potential, so that it is reasonable to perform theoretical calculations of these dependences on the basis of a model potential. The Rapini potential is most widely

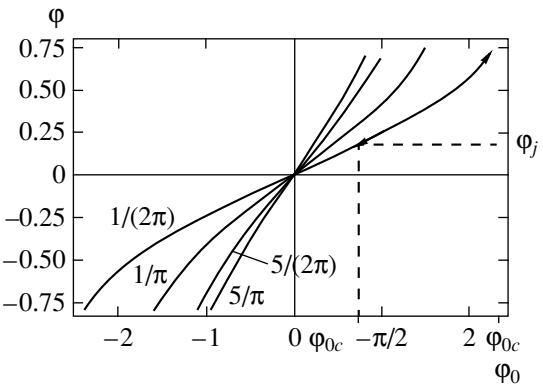

Fig. 1. Computed dependence of the deflection angle of the director at the surface of a layer for various values of the parameter S_d. The value of the angle φ_j, realized after the pitch jump, is marked on one of the curves.

used for describing the surface anchoring [9–11]. Consequently, we present below the calculations of the temperature variations of the pitch in a planar cholesteric layer for the Rapini potential.

The explicit form of the expression for the free energy is now determined by the relation

$$F(T)/W = 2S_d[\varphi(T) - \varphi_0(T)]^2 - \cos^2\varphi(T). \quad (11)$$

The relation determining the equilibrium value of the angle $\varphi(T)$ becomes

$$4S_d[\varphi(T) - \varphi_0(T)] + \sin 2\varphi(T) = 0. \quad (12)$$

The expression for the free energy in the form of a function of the deflection angle of the director at the surface from the alignment direction is

$$\frac{F(T)}{W} = \frac{\sin^2 2\varphi(T)}{8S_d} - \cos^2\varphi(T). \quad (13)$$

Finally, Eq. (5) for the barrier becomes

$$\frac{B(T)}{W} = 2S_d\left[\frac{\pi}{4}\operatorname{sgn}\varphi_0 - \varphi_0(T)\right]^2 + \cos^2\varphi(T) - \frac{\sin^2 2\varphi(T)}{8S_d} - \frac{1}{2}. \quad (14)$$

Using Eq. (12), we recast Eq. (14) into the form

$$\frac{B(T)}{W} = 2S_d\left[\frac{\pi}{4}\operatorname{sgn}\varphi_0 - \varphi(T)\right] \times \left[\frac{\pi}{4}\operatorname{sgn}\varphi_0 + \varphi(T) - 2\varphi_0(T)\right] + \cos^2\varphi(T) - \frac{1}{2}. \quad (15)$$

The temperature dependence of the pitch of the helix can be represented in a universal form. For this, only the functions $\varphi_0(T)$ (the angle of free rotation of the director on the surface of the layer) or $\varphi(T)$ (the rotation angle of the director on the surface of the layer in the presence of anchoring) should appear in Eqs. (11)–(15). As noted above, the corresponding dependences are universal, i.e., they are determined only by the dimensionless parameter S_d. The conversion from computed quantities to dependences for a real sample is made using simple transformations (see below).

For example, the free rotation angle of the director at the jump point is given by the expression

$$\varphi_0(T_c) = \varphi_c + \frac{1}{4S_d}, \quad (16)$$

and the pitch corresponding to the jump in a bulk cholesteric is given by the formula

$$p(T_c) = 2d\left(N + \frac{1}{2} + \frac{1}{2\pi S_d}\right)^{-1}. \quad (17)$$

Figure 1 shows the temperature dependence $\varphi(T)$ of the deflection angle of the director on the surface of the layer for the critical value of the parameter $S_d = S_{dc} = 1/2\pi$ and several other values of the parameter S_d for which the pitch jump cannot draw metastable states into the process leading to the change in the configuration of the helix in the layer. Possible multiple pitch jumps, alternating with smooth temperature variations of the pitch with increasing temperature, fall on the right-hand branch of the curve. As temperature increases (it is assumed that in the bulk of a cholesteric the pitch decreases with increasing temperature), the angle $\varphi(T)$ reaches the critical value $\pi/4$ (for a free rotation angle φ_{0c} that depends on S_d), after which it decreases abruptly to φ_j, corresponding to the abscissa $\varphi_{0c} - \pi/2$. As temperature increases further, $\varphi(T)$ changes from the value corresponding to the abscissa $\varphi_{0c} - \pi/2$ to $\pi/4$, where it once again changes abruptly to the same value as in the first jump. Thus, the angle changes repeatedly along the curve between the points with the abscissas $\varphi_{0c} - \pi/2$ and φ_{0c}. In the process, the number of half-turns of the helix over the thickness of the layer changes, but the range of variation of the director angles on the surface of the layer is independent of the number of half-turns of the helix. As temperature decreases, $\varphi(T)$ reaches a critical value $-\pi/4$, after which it changes abruptly to the value $-\varphi_j$ corresponding to the abscissa $-\varphi_{0c} + \pi/2$, and as temperature decreases further, it changes along the left-hand branch of the curve between the points with abscissas $-\varphi_{0c} + \pi/2$ and $-\varphi_{0c}$, just as with increasing temperature.

It follows from Eq. (16) that the case of interest for the problem at hand, where after the jump $\varphi(T) = \varphi_0(T) = 0$, i.e., a new configuration of the helix corresponds to a minimum of the free energy, corresponds to the value $S_d = 1/\pi$. In this case the angles which the director makes at the surface with the alignment direction before and after the jump differ by $\pi/4$. The maximum value of this difference is $\pi/2$ and occurs in the limit of vanishingly weak surface anchoring, i.e., in the limit of infinitely large S_d.

The values of the pitches of the helix in the layer for a specific sample immediately before and after a jump can be easily found from the computed curve. Thus, as temperature increases in a layer with N half-turns of the helix the values of the pitch before and after a jump are given by the expressions, respectively,

$$p_c = \frac{2d}{N + 1/2}, \quad p_j = \frac{2d}{N + 1/2 + 2\varphi_j/\pi}. \tag{18}$$

As temperature decreases, the corresponding values of the pitches in the layer are given by the expressions

$$p_c = \frac{2d}{N - 1/2}, \quad p_j = \frac{2d}{N - 1/2 - 2\varphi_j/\pi}. \tag{19}$$

The second relation in Eqs. (8) for the temperature hysteresis of the pitch assumes the form

$$d\left(\frac{1}{p(T_+)} - \frac{1}{p(T_-)}\right) = \frac{1}{2\pi S_d}. \tag{20}$$

Figures 2 and 3 show the computed values of the free energy of a layer for various values of the parameter S_d as a function of the free angle of rotation of the director on the surface and the angle of rotation of the director taking account of the surface anchoring, respectively. We note that for a constant direction of temperature variation the free energy of the equilibrium state of the layer varies between two values, exceeding its minimum, and the free energy of the layer can pass through the minimum value before the first pitch jump only if the direction of temperature variation changes. As could be seen above, the only exception is the case corresponding to the parameter $S_d = 1/\pi$, for which the equilibrium value of the free energy after the jump reaches a minimum value.

Figure 4 shows the computed dependence of the height of the barrier between the configurations of the helix that differ by one half-turn in the layer on the angle of rotation of the director at the surface in the presence of surface anchoring. Figure 5 shows the same dependence but as a function of the free angle of rotation of the director. It is evident that, just like the free energy, the behavior of the barrier does not depend on the number of turns of the helix within the thickness of the layer, but rather it is determined only by the value of the parameter S_d. We note here that the curves presented correspond to a real barrier only for $\varphi_0(T) > \pi/4$, while for a smaller value of this angle the energy of the initial configuration of the helix is lower than for the configuration with a different number of half-turns of the helix, and the curves presented do not represent a barrier.

Figure 6 shows the hysteresis loops for the deviation of the number $N(T)$ of half-turns of the helix in the layer versus the integer number corresponding to the orientation of the director at the surface of the layer in the direction of alignment for the case where the temperature range contains one pitch jump with temperature varying in the direct and reverse directions (a) and for the case where the temperature range contains several jumps of the pitch with temperature varying in the direct and reverse directions (b). In contrast to the preceding figures, here the reference point for the angle $\varphi_0(T)$

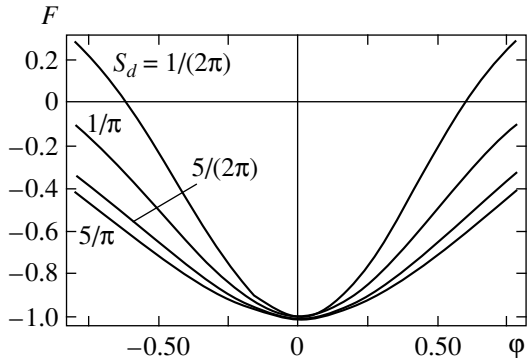

Fig. 2. Computed dependence of the free energy of the layer (arbitrary units) on the deflection angle of the director at the surface of the layer $\varphi(T)$ for various values of the parameter S_d.

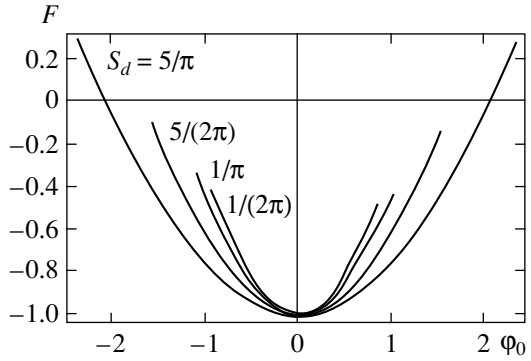

Fig. 3. Computed dependence of the free energy (arbitrary units) on $\varphi_0(T)$ for various values of the parameter S_d.

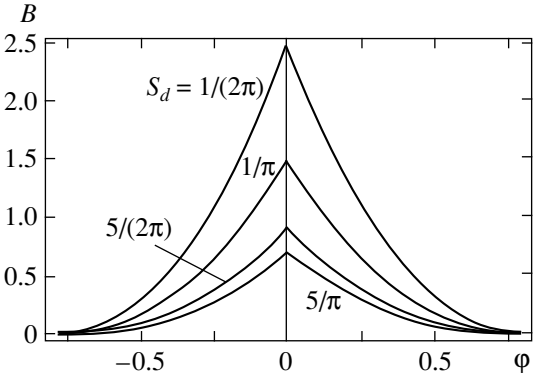

Fig. 4. Computed dependence of the barrier height (arbitrary units) between helix configurations differing by one half-turn in the layer versus the deflection angle of the director at the surface of the layer for various values of the parameter S_d.

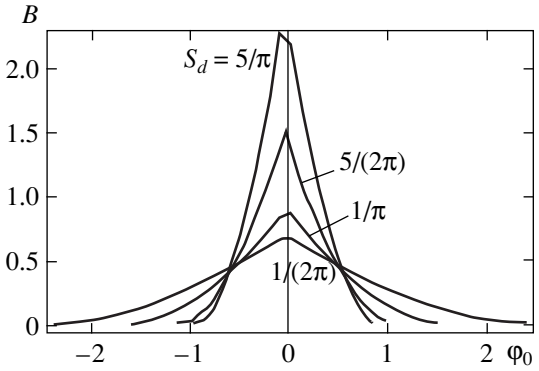

Fig. 5. Computed dependence of the barrier height (arbitrary units) between helix configurations differing by one half-turn in the layer on $\varphi_0(T)$ for various values of the parameter S_d.

does not change after a pitch jump, but rather it remains fixed in the entire temperature range. We call attention to the fact that temperature ranges where the variation of N does not depend on the direction of temperature variation occur at the limits of the ranges of variation of $\varphi_0(T)$ which were presented above.

6. TEMPERATURE FLUCTUATIONS OF THE DIRECTOR ORIENTATION

In the analysis performed above, the orientational fluctuations of the director in the layer were completely ignored, since the analysis was performed on the basis of the continuum theory of elasticity, which works with the macroscopic average characteristics of a liquid crystal. It is obvious that under definite conditions fluctuations can be very strong, and sometimes they can even determine the character of the phenomenon under study (see, for example, the investigation of orientational fluctuations in nematic layers [12, 13]). In what follows, the influence of orientational fluctuations of the director on the temperature variations of the pitch in the layer is briefly analyzed on the basis of the model free energy (1).

The quadratic correction to the free energy (with respect to small angular deflections $\Delta\varphi$ of the director away from its equilibrium orientation at the surface of the layer for the temperature under study) is determined by the expression

$$\Delta F = W(\cos 2\varphi + 2S_d)(\Delta\varphi)^2. \qquad (21)$$

Hence the mean-squared angular temperature fluctuations of the director at the surface of the layer $\langle(\Delta\varphi)^2\rangle$ can be expressed in terms of the temperature on the basis of the equipartition theorem for energy [14],

$$\langle(\Delta\varphi)^2\rangle = \frac{k_B T}{W(\cos 2\varphi + 2S_d)A}. \qquad (22)$$

The expression (22) shows that because of temperature fluctuations of the orientation angle of the director at the surface or, equivalently, the free energy of the layer the temperatures of the pitch jumps shift in the direction of decreasing hysteresis: a decrease of the jump temperature with increasing temperature and an increase of the jump temperature with decreasing temperature. If the shift of the jump points covers the temperature range of hysteresis, then hysteresis does not appear in the temperature variations of the pitch of the helix in the layer.

Specifically, a modification of Eq. (16) for the free angle of rotation of the director at the jump point of the pitch, determining the temperature of the jump, follows from Eq. (22). Now, taking account of fluctuations of the director, we obtain instead of Eq. (16)

$$\varphi_0(T_c) \approx \frac{\pi}{4} + \frac{1}{4S_d} \pm \sqrt{\frac{k_B T_c}{(\cos 2\varphi + 2S_d)WA}}, \qquad (23)$$

where $\pi/4$ is substituted for the critical angle φ_c, i.e., the value for the Rapini potential.

An obvious consequence of Eq. (23) is that hysteresis decreases with temperature and can vanish completely with a definite increase of temperature.

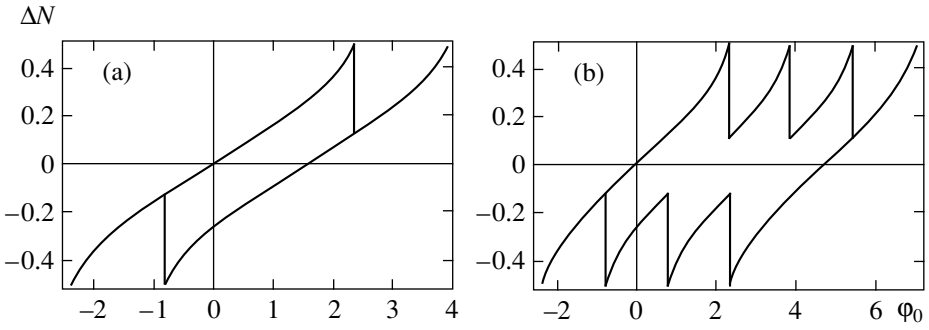

Fig. 6. Computed dependence on $\varphi_0(T)$ of the hysteresis in the deviation ΔN of the number of half-turns in a layer from an integer for the temperature range (a) with one pitch jump for both directions of variation of the temperature and (b) with three jumps for $S_d = 1/2\pi$. Here the reference point for the angle $\varphi_0(T)$ does not change after a pitch jump, but rather it remains fixed in the entire range of temperature variation.

Of course, Eqs. (21) and (22) describe only uniform fluctuations of φ on the basis of quasi-smectic (large-scale) energy (1). The general orientational fluctuation of the director can be described by two parameters that depend on three coordinates. However, the energy of such a general fluctuation cannot be found by confining attention only to the large-scale energy (1). The total Frank energy must be minimized. The solution of this problem falls outside the scope of the present work.

7. CONCLUSIONS

The analysis, given above, of the temperature variations of the pitch of a helix in a cholesteric layer shows that if the pitch jumps in the layer can be described by the mechanism of surface slipping of the director (i.e., if the parameter S_d is larger than the critical value found above), then a quite universal pattern of temperature variations of the pitch, determined by the dimensionless parameter S_d, is realized. The theory developed gives experimentally verifiable predictions, for example, the values of the pitch in the layer at the jump temperatures for temperature varying in opposite directions should be related with one another. The formulas derived make it possible to obtain from the measurements quantitative information about the surface anchoring forces (the value of the parameter S_d), if the form of the surface potential is known, for example, the Rapini potential. However, the problem of reconstructing the real potential of the surface anchoring from the measured temperature dependences of the pitch of the helix in a layer is probably no less urgent. This problem can be solved in principle using the formulas obtained in the present paper. This problem is especially important for liquid crystals bounded by perfect single-crystal substrates, for which the anchoring potential, in contrast to the Rapini potential, can be nonmonotonic with respect to angle and can even possess several local minima.

As far as the hysteresis of the pitch in cholesteric (and other chiral) liquid crystal layers is concerned, this phenomenon is promising for applications, especially since not only can the temperature hysteresis examined above appear, but hysteresis with respect to other external actions, for example, electric or magnetic fields, can also arise.

It is interesting that although formally the large-scale energy (1) used above has a quasi-smectic form, the physics of cholesteric and smectic ordering is very different. Smectic layers are associated with density modulations in the system, and consequently a change in the number of smectic layers in the film results in an abrupt change in the thickness of the film. "Cholesteric layers" are formed in a continuous distribution of the orientational order parameter (director), and consequently a change in the number of half-turns of the cholesteric spiral on the thickness of the sample is not necessarily directly related with a change in the thickness of the sample. Consequently, it can be stated that the mechanism, investigated in the present paper, for the abrupt change in the pitch of a helix in cholesteric layers is dual to the intensely studied, in recent years (see, for example, [15–17]), phase transitions of layerwise thinning of the freely suspended films of smectic liquid layers.

The results obtained in the present work can be applied directly for determining the surface anchoring force and restoring the form of the surface potential according to the angular deflections of the director.[1] The results obtained can be used to determine whether or not the mechanism of surface slipping of the director works in the case investigated. This is determined according to the difference between the angle corresponding to a pitch jump and the critical angle. If the difference exceeds the rms angle of the fluctuations, determined by Eq. (22), then the mechanism of the pitch jump is different from surface slipping.

Investigations of the dependence of the temperature hysteresis in pitch jumps on the thickness of the layer make it possible to study the effect of fluctuations of the orientation of the director on the temperature variations of the pitch. The general tendency here is as follows: the smaller the thickness of the layer, other conditions being the same, the stronger the effect of fluctuations of the director on the temperature variations of the pitch are. This is illustrated in Fig. 5, which demonstrates that for small values of the parameter S_d (i.e., large layer thicknesses) the barrier height is a very gently sloping function of the angle of free rotation of the director $\varphi_0(T)$, and consequently small fluctuation variations of the layer energy make it possible to overcome the barrier for $\varphi_0(T) < \varphi_0(T_c)$, i.e., they decrease the hysteresis. Ultimately, for sufficiently thick layers this results in no temperature hysteresis in variations of the pitch (as is observed experimentally).

On the whole, the possibility of varying the experimental parameters opens up a large scope for experimental investigation of the problem. For example, changing the temperature makes it possible to change near T_1—the transition point of the LC into an isotropic liquid—the parameter S_d, since $W \propto (T - T_1)^{1/2}$, and $K_{22} \propto (T - T_1)$, so that $S_d \propto (T - T_1)^{1/2}$ and as $T \longrightarrow T_1$ strong anchoring conditions always occur. As the surface area A of the layer decreases, fluctuations should play a larger role, since they grow as $1/A$, which decreases hysteresis and hysteresis vanishes if the surface area is sufficiently small. It could be helpful to study the phenomenon in a wedge-shaped cell.

It should also be noted that the results obtained in the present work are a necessary initial step for studying the dynamical characteristics of LC layers. For

[1] The preferred variant is the one where an absolutely rigid anchoring of the director with the surface exists on one of the surfaces of the layer, since in this case the form of the potential can be reconstructed up to the angle of deflection of the director from the alignment direction $\pi/2$ and not up to $\pi/4$, as in the case considered.

example, assuming a purely relaxation dynamics for a single variable describing the properties of the model under study, the angle φ, the temporal correlation of the phase of the light which has passed through a LC layer can be found. Comparing the correlation function, found theoretically in this manner, with the experimental data makes it possible to find an important characteristic of the LC in the layer—the relaxation time of the angular deflection of the director at the surface from the direction corresponding to a minimum of the free energy of the layer.

ACKNOWLEDGMENTS

This work was supported by INTAS (grant no. 30234). One of us (E. I. K.) is also grateful to the Russian Foundation for Basic Research (project no. 00-02-17785) for partial support.

REFERENCES

1. *Summaries of 7th International Conference on Ferroelectric Liquid Crystals, Darmstadt University of Technology, Germany, 1999.*
2. A. B. Harris, R. D. Kamien, and T. C. Lubensky, Rev. Mod. Phys. **71**, 1745 (1999).
3. H. Zink and V. A. Belyakov, Mol. Cryst. Liq. Cryst. **265**, 445 (1995); Pis'ma Zh. Éksp. Teor. Fiz. **63**, 37 (1996) [JETP Lett. **63**, 43 (1996)].
4. H. Zink and V. A. Belyakov, Zh. Éksp. Teor. Fiz. **112**, 524 (1997) [JETP **85**, 285 (1997)]; Mol. Cryst. Liq. Cryst. **329**, 457 (1999).
5. V. A. Belyakov and V. E. Dmitrienko, Sov. Sci. Rev., Sect. A **13**, 1 (1989).
6. V. A. Belyakov, *Diffraction Optics of Complex-Structured Periodic Media* (Nauka, Moscow, 1988; Springer-Verlag, New York, 1992).
7. T. Furukawa, T. Yamada, K. Ishikawa, *et al.*, Appl. Phys. B **60**, 485 (1995).
8. Z. Zhuang, Y. J. Kim, and J. S. Patel, Phys. Rev. Lett. **84**, 1168 (2000).
9. P. G. de Gennes and J. Prost, *The Physics of Liquid Crystals* (Clarendon Press, Oxford, 1993).
10. L. M. Blinov, E. I. Kats, and A. A. Sonin, Usp. Fiz. Nauk **152**, 449 (1987) [Sov. Phys. Usp. **30**, 604 (1987)].
11. L. M. Blinov and V. G. Chigrinov, *Electrooptics Effects in Liquid Crystal Materials* (Springer-Verlag, New York, 1994), Chap. 3.
12. B. Ya. Zel'dovich and N. V. Tabiryan, Zh. Éksp. Teor. Fiz. **81**, 1738 (1981) [Sov. Phys. JETP **54**, 922 (1981)].
13. T. Ya. Marusiĭ, Yu. A. Reznikov, V. Yu. Reshetnyak, *et al.*, Zh. Éksp. Teor. Fiz. **91**, 851 (1986) [Sov. Phys. JETP **64**, 502 (1986)].
14. L. D. Landau and E. M. Lifshitz, *Statistical Physics* (Nauka, Moscow, 1976; Pergamon, Oxford, 1980), Part 1.
15. P. G. de Gennes, Langmuir **6**, 1448 (1990).
16. S. Pankratz, P. M. Johnson, Y. T. Nguyen, and C. C. Huang, Phys. Rev. E **58**, R2721 (1998).
17. E. E. Gorodetskiĭ, E. S. Pikina, and V. E. Podnek, Zh. Éksp. Teor. Fiz. **115**, 61 (1999) [JETP **88**, 35 (1999)].

Translation was provided by AIP

Journal of Experimental and Theoretical Physics, Vol. 91, No. 3, 2000, pp. 497–503.
Translated from Zhurnal Éksperimental'noĭ i Teoreticheskoĭ Fiziki, Vol. 118, No. 3, 2000, pp. 570–578.

SOLIDS
Structure

Spatial Distribution of Elastic Deformations in Ge/Si Structures with Quantum Dots

A. V. Nenashev[a, *] and A. V. Dvurechenskiĭ[b]

[a]*Novosibirsk State University, Novosibirsk, 630090 Russia*
[b]*Institute of Semiconductor Physics, Siberian Division, Russian Academy of Sciences, Novosibirsk, 630090 Russia*
**e-mail: nenashev@isp.nsc.ru*

Received January 18, 2000

Abstract—A method is developed for calculating the elastic deformation in coherently strained heterostructures on the basis of the valence force field (VFF) model using the Green's function of the "atomistic" elastic problem. The spatial distribution of the elastic deformations in a Ge/Si system with pyramidal Ge quantum dots buried in a Si matrix is investigated theoretically. The deformation distribution in and around the pyramids is determined. Near quantum dots, the region near the tip of the pyramid is most strongly intensely. Inside quantum dots the region of the vertex is most relaxed, and the most strained section lies on the contour of the pyramid base. Compression occurs in the plane of the pyramid base inside quantum dots, and stretching occurs along the vertical direction. The picture is reversed near quantum dots: stretching occurs in the lateral direction and compression in the vertical direction. It is shown that the local deformations and their spatial distribution are essentially independent (to within the scaling) of the size of the quantum dots for 10–15 nm pyramid bases.

1. INTRODUCTION

The interest of researchers in semiconductor structures containing zero-dimensional objects (quantum dots) has increased substantially in the last few years. On account of their small sizes and high size and shape uniformity quantum dots formed during heteroepitaxy of elastically strained systems are most attractive for practical applications [1, 2].

The main characteristic of quantum dots, which determines their electric and optical properties, is the energy spectrum of the charge carriers. To calculate the spectrum of the states in self-forming islands it is necessary to know the elastic deformation due to the lattice mismatch between the nanocluster and substrate materials. For example, in the Ge/Si system the lattice mismatch is 4.2%. A deformation can change the energy structure by an amount of the order of 0.1 eV [3], which is comparable to the size-quantization energy of the charge carriers in quantum dots. Consequently, finding the elastic deformation fields is a necessary preliminary step for calculating the electronic structure of self-forming quantum dots.

Most theoretical and experimental works on determining the elastic deformation in self-forming quantum dots are concerned with the heterosystem InAs/GaAs [4–9]. There are two approaches to calculating the deformation in quantum dots: the continuous medium model [4, 8, 10] and the valence force field (VFF) model [5–7, 9]. The first model is macroscopic and the second is atomistic, and it is based on the potential of the elastic forces as a function of the interatomic distances and angles. Moreover, a recent work [11] employed the molecular-dynamics method to obtain the distribution of elastic stresses over the free surface in Ge/Si structures with quantum dots.

Even though there have been many experimental investigations of Ge/Si heterostructures with quantum dots (see [1] and the references cited there), no theoretical investigations of the electronic spectrum of such quantum dots have yet been performed. Our objective in the present work is to find numerically the elastic deformation fields in self-forming Ge quantum dots buried in a Si matrix; knowledge of these fields is necessary in order to analyze the electronic spectrum in such structures theoretically. We chose the atomistic approach to find the deformations. This choice is made because, in the first place, the objects under study are comparable in size with the lattice constant. In the second place, to calculate the electronic spectrum it is better to have a result in the form of the positions of the atoms and not in the form of average quantities, such as the deformation tensor. We developed a method for calculating the elastic deformation on the basis of the VFF model [12–14] with a Keating potential [13]. The Green's function, together with the atomistic approach, is used for the first time to find the deformation in a heterostructure. The main advantage of the method is that the final results are insensitive to the position of the boundary of the atomic cluster chosen.

The initial structure and the characteristic shape and size of Ge quantum dots in Si are shown in Fig. 1. The quantum dots are square pyramids with a (001) base and {105} lateral faces, and they lie on top of a thin

1063-7761/00/9103-0497$20.00

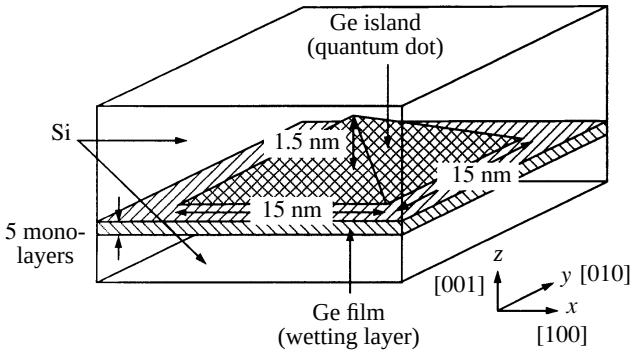

Fig. 1. Schematic diagram of a typical Ge quantum dot in Si.

continuous Ge film. The problem is to find the deformation in a system containing quantum dots and a two-dimensional (2D) layer.

2. MODEL

The VFF model is widely used to calculate the elastic deformation at the atomic level. It is postulated in this model that the elastic deformation energy of a crystal depends only on the positions of the atomic nuclei and can be expressed as a sum in which each term depends only on the position of an atom and its nearest neighbor. In the present paper we employ the Keating expression [13] for the elastic energy V extended to the case of a heterostructure (see [6]):

$$V = \frac{3}{16}\sum_i\sum_j \frac{\alpha_{ij}}{d_{0,ij}^2}\left(d_{ij}^2 - d_{0,ij}\right)^2 + \frac{3}{8}\sum_i\sum_{(j,k)} \frac{\beta_{ijk}}{d_{0,ij}d_{0,ik}}\left(\mathbf{d}_{ij}\mathbf{d}_{ik} + \frac{d_{0,ij}d_{0,ik}}{3}\right)^2, \quad (1)$$

where the indices i, j, and k enumerate the atoms, the index i runs over all atoms, j in the first sum runs over the nearest neighbors of the ith atom, and the pair of indices (j, k) in the second sum runs over all pairs of nearest neighbors of the ith atom; $\mathbf{d}_{ij}$ is a vector directed from the atom i to the atom j, and d_{ij} is the length of this vector; $d_{0,ij}$ is the length of the undeformed i–j bond; and, α_{ij} and β_{ijk} are the force constants in the Keating model. For $d_{0,ij}$, α_{ij}, and β_{ijk} we use the values

$$d_{0,ij} = \frac{d_{0,i} + d_{0,j}}{2}, \quad \alpha_{ij} = \frac{\alpha_i + \alpha_j}{2},$$

$$\beta_{ijk} = \frac{(\beta_i + \beta_j) + (\beta_i + \beta_k)}{4},$$

where $d_{0,i}$, α_i, and β_i are the corresponding values for a pure Si or Ge crystal depending on the type of ith atom. The values used in this work are presented in table.

Parameters of the VFF model for silicon and germanium [13]

	d_0, nm	α, N/m	β, N/m
Si	0.2352	48.5	13.8
Ge	0.2450	38	12

The problem is to find a set of atomic positions that minimizes V. This is equivalent to vanishing of the elastic forces acting on each atom. Ordinarily, this problem is solved by using periodic boundary conditions (see, for example, [9]). In such an approach a large number of atoms must be included in the calculations, since to avoid any influence of the finite size of the region of the crystal under study the boundaries of this region must be located at a distance which is large compared with the size of the quantum dots. This increases the required computer resources substantially.

In the method proposed for calculating the elastic deformation in a quantum dot, only atoms belonging to a quantum dot and several atomic layers around the dot are included in the analysis. For this, we shall consider the set of atomic displacements and we shall introduce the following transformation of this set:

$$u_\alpha^i \longrightarrow u_\alpha^i + \sum_j G_{\alpha\beta}^{ij} f_\beta^j[u], \quad (2)$$

where α and β are tensor indices, which run through the values x, y, z; the indices i and j run over all atoms; the brackets denote the set of displacements of the atoms; u_α^i is the αth component of the displacements of the ith atom, measured from its position in the undeformed lattice of the matrix in the absence of quantum dots (we are considering a defect-free structure); $f_\beta^j[u]$ is the βth component of the elastic force (defined as $-\partial V/\partial u_\beta^j$) acting on the jth atom with atomic displacement u; $G_{\alpha\beta}^{ij}$ is the αth component of the displacement of the ith atom under the action of a single force applied to the jth atom of the pure matrix crystal in the direction β (in the linear approximation). We shall call $G_{\alpha\beta}^{ij}$ the Green's function of the atomistic problem, by analogy to the Green's tensor, or the Green's function of the main problem of the continuum theory of elasticity [15].

We note that the transformation (2) decreases the elastic forces acting on the atoms if the force constants of the substrate and quantum-dot atoms are close. Indeed, for atomic displacements $\nu_\alpha^i = G_{\alpha\beta}^{ij}$ we have the forces $f_\alpha^i[\nu] = -\delta_{ij}\delta_{\alpha\beta}$ for a pure substrate crystal

and $f_\alpha^i[v] = f_\alpha^i[0] - \delta_{ij}\delta_{\alpha\beta} + \delta f_{\alpha\beta}^{ij}$ in the presence of quantum dots (the linear approximation is assumed). Here $f_\alpha^i[0]$ are the elastic forces for zero displacements; $\delta f_{\alpha\beta}^{ij}$ are small compared to 1. For the displacements

$$\tilde{u}_\alpha^i = u_\alpha^i + \sum_j G_{\alpha\beta}^{ij} f_\beta^j[u]$$

we obtain

$$f_\alpha^i[\tilde{u}] = f_\alpha^i[u] + \sum_j (-\delta_{ij}\delta_{\alpha\beta} + \delta_{\alpha\beta}^{ij}) f_\beta^j[u]$$

$$= \sum_j \delta f_{\alpha\beta}^{ij} f_\beta^j[u].$$

Since $\delta f_{\alpha\beta}^{ij}$ are small, the forces $f[\tilde{u}]$ will be small compared to $f[u]$. The nonlinear terms in the expression for $f[\tilde{u}]$ are dropped, since they are also small compared to $f[u]$.

Thus, by repeating the transformation (2) the elastic forces can be made as small as desired. If the initial atomic displacements are set equal to zero, then after performing the transformation (2) successively the displacements will have the form

$$u_\alpha^i = \sum_j G_{\alpha\beta}^{ij} g_\beta^j. \quad (3)$$

The quantities g_α^i in Eq. (3) vanish (in the linear approximation) for all sites i where the second coordination sphere contains only matrix atoms. This follows from the fact that the expression for the elastic force acting on such a site is identical to the analogous expression for a pure matrix crystal. Thus, according to the definition of $G_{\alpha\beta}^{ij}$ we obtain from Eq. (3)

$$f_\alpha^i[u] = -g_\alpha^i.$$

Consequently, after performing the transformation (2) the quantities g_α^i will be zero for all such sites.

Thus, the sum in Eq. (3) contains a finite number of nonzero terms, which is approximately equal to the number of atoms in a quantum dot. We now write the transformation (2) in terms of g_α^i:

$$g_\alpha^i \longrightarrow g_\alpha^i + f_\alpha^i[u]. \quad (4)$$

Since the number of nonzero coefficients g is finite, a finite number of operations must be performed in order to perform this transformation.

Thus, the proposed method consists of performing the transformation (4) successively, starting with zero atomic displacements, until the elastic forces are negligible. In this method it is sufficient to calculate only the positions of the atoms belonging to a quantum dot together with its environment within one lattice constant, because the positions of all other atoms (in the linear approximation) do not influence the nonvanishing value of g in Eq. (3).

In order to apply the above-described method to a structure which includes an infinite two-dimensional layer it is necessary to find the positions of the atoms in a system containing only a 2D layer and then to use these positions as initial values. In this case the atoms located far from the quantum dots need not be included. Strictly speaking, it is necessary to use the Green's function $\tilde{G}_{\alpha\beta}^{ij}$ for a structure containing a 2D layer together with the Green's function $G_{\alpha\beta}^{ij}$ for the pure matrix. However, for the system Ge/Si (100) the difference between $G_{\alpha\beta}^{i0}$ and $G_{\alpha x}^{i0}$ can be neglected because the thickness of the wetting layer and the difference between the force constants of Si and Ge are both small.

Finding the atomic positions in a continuous germanium film on (100) silicon reduces to finding the distances between the (100) atomic planes. The distance between two neighboring Ge monolayers can be obtained in the continuum theory of elasticity (from the deformation tensor of a two-dimensional Ge layer in a Si matrix). The distance between the Si–Si layers in the theory of elasticity remains the same as in bulk Si, and the Si–Ge layer is found as the arithmetic-mean of the distances Si–Si and Ge–Ge.

The function $G_{\alpha\beta}^{ij}$ can be calculated using the "continuum" Green's function. We note first that since all sites in a diamond-type lattice are equivalent, it is sufficient to find $G_{\alpha\beta}^{i0}$ for a site o. Next, on account of the symmetry of the three directions $|100\rangle$, it is sufficient to find only the coefficients $G_{\alpha x}^{i0}$. As a first approximation to $G_{\alpha x}^{i0}$, we shall use the Green's function for the problem of the continuum theory of elasticity, $G_{\alpha x}(\mathbf{r})$, where $\mathbf{r}$ is a vector directed from the site o to the site i. The function $G_{\alpha\beta}(\mathbf{r})$ is determined as the displacement of the elastic medium in the direction α under the action of a single force applied to the origin of coordinates in the direction β. (In what follows, we shall drop the second index x in the Green's function.) It was found in [16] for a cubic crystal. The elastic moduli of the medium, which are required in order to calculate $G_{\alpha\beta}(\mathbf{r})$, can be obtained from the force constants α and β [13].

The function $G_\alpha(\mathbf{r})$ decreases with distance as r^{-1}. To obtain the correction $G_\alpha^{(2)}(\mathbf{r})$, which decreases as r^{-2}, we shall consider the displacement of two sublattices

relative to one another in a deformation. The components *x*, *y*, and *z* of such a displacement are, respectively, $-A\xi\varepsilon_{xz}/2$, $-A\xi\varepsilon_{xz}/2$, and $-A\xi\varepsilon_{xy}/2$ [13], where A is the lattice constant, $\xi = (\alpha - \beta)/(\alpha + \beta)$ is the deformation tensor, and $\xi_{\alpha\beta}$ is the relative-displacements parameter [17]. If $G_\alpha(\mathbf{r})$ is treated as a displacement vector, then the relative displacement of the sublattices is proportional to r^{-2}—this is what gives the desired correction $G_\alpha^{(2)}(\mathbf{r})$. Assuming that the displacement is distributed uniformly over both sublattices, we obtain finally

$$G_x^{(2)}(\mathbf{r}) = \pm\frac{A\xi}{8}\left(\frac{\partial G_y(\mathbf{r})}{\partial z} + \frac{\partial G_z(\mathbf{r})}{\partial y}\right)$$

for the *x* component and similar expressions for the *y* and *z* components. Here the "+" and "–" signs are chosen depending on the sublattice: the "+" sign corresponds to the sublattice which transforms into the other sublattice as a result of a translation by the vector $\left(\frac{A}{4}, \frac{A}{4}, \frac{A}{4}\right)$.

It is easy to check numerically that for atomic displacements equal to $G_\alpha(\mathbf{r}) + G_\alpha^{(2)}(\mathbf{r})$ the elastic forces decrease with distance as r^{-4}. Thus the desired Green's function G_α^{i0} is $G_\alpha(\mathbf{r}) + G_\alpha^{(2)}(\mathbf{r})$ to within $O(r^{-4})$. Consequently, we can set $G_\alpha^{i0} = G_\alpha(\mathbf{r}) + G_\alpha^{(2)}(\mathbf{r})$ for distances $r > 10A$. For smaller distances the Green's function can be found numerically as the equilibrium displacement of atoms in a cluster surrounding the atom *o* under a single force directed toward the atom *o*, and the atoms on the boundary of the cluster must be fixed with the displacements $G_\alpha(\mathbf{r}) + G_\alpha^{(2)}(\mathbf{r})$.

The atomic displacements can be converted into local values of the deformation tensor. For this, we shall consider the deformation of a tetrahedron consisting of the nearest neighbors of a lattice site. The form of the tetrahedron is determined by six parameters, for example, the edge lengths. Thus, the deformation of such a tetrahedron uniquely determines the six components of the deformation tensor associated with a given lattice site. The elastic energy related with a definite atom can be found similarly: the expression (1) for the elastic energy is separated into terms, each of which is associated with a lattice site and consequently can be interpreted as the fraction of the energy per atom.

3. DEFORMATIONS IN A Ge/Si HETEROSYSTEM

The method described above was used to find the spatial distribution of other deformations in Ge quantum dots buried in a Si matrix (Fig. 1).

Figure 2 (isolines) shows the distribution of the elastic energy per atom in two sections for quantum dots with the base edge length equal to 28 lattice constants (about 15 nm). The thickness of the solid Ge film was taken to be 5 monolayers [18]. It is evident that the region near the tip of the pyramid is the most stressed region around quantum dots. Inside quantum dots the neighborhood of the tip is most strongly relaxed. The most strongly stressed section lies along the contour of the pyramid base. In the region of the Si–Ge interface the elastic energy per atom changes abruptly at a transition to an atom of a different kind or to an atom with a different immediate environment. Consequently, the isolines are not shown there.

For comparison, Fig. 3 shows a similar distribution of the elastic energy in a quantum dot with a ~10 nm base. It is evident that the overall picture of the spatial distribution of the elastic stresses is similar to the picture obtained for a 15 nm dot. This shows that the elastic deformation of a system with these sizes can be described by the macroscopic approximation. In the macroscopic limit structures with the same shape but different size should be deformed identically, i.e., the spatial distribution of the deformation temperature and hence the density of the elastic energy should be the same in such structures to within the scaling. This follows from dimensional considerations: the deformation tensor, being a dimensionless quantity, cannot depend on the dimensions of the structure themselves, but rather it depends only on the ratio of the dimensions, i.e., on the shape.

To analyze the quantitative dependences of the elastic stresses on the size of a nanocluster we examined the values of the elastic deformation at equivalent points of pyramids of different size. We shall say that points inside nanoclusters of different size are equivalent to one another if the nanoclusters transform into one another under a transformation that brings one cluster into coincidence with another. As the size of a nanocluster increases, the elastic deformation at equivalent points must approach a constant corresponding to the continuous medium approximation. Our results show that the elastic deformation first becomes a constant near the center of the pyramid base (curve *1* in Fig. 4). On the whole, in the central region of a Ge pyramid the deviation of the deformation tensor from the limiting macroscopic value does not exceed 10^{-3} for 10–15 mm pyramids. Thus, the values of ε_{xy} on the pyramid axis in the indicated size range of pyramids lies in the range $\pm 0.5 \times 10^{-3}$, which corresponds to a zero macroscopic value. A different behavior of the deformation as a function of the size of a nanoisland is observed near an edge of the pyramid base. As an illustration, we present the dependence of the component ε_{xx} at a point located 1.5 lattice constants from the edge center in the direction of the center of the pyramid base (curve *2* in Fig. 4). As the size of the Ge cluster increases, the elastic deformation near the edge does not approach a constant; this is because the deformation at the edge of the pyramid becomes infinite in the macroscopic limit.

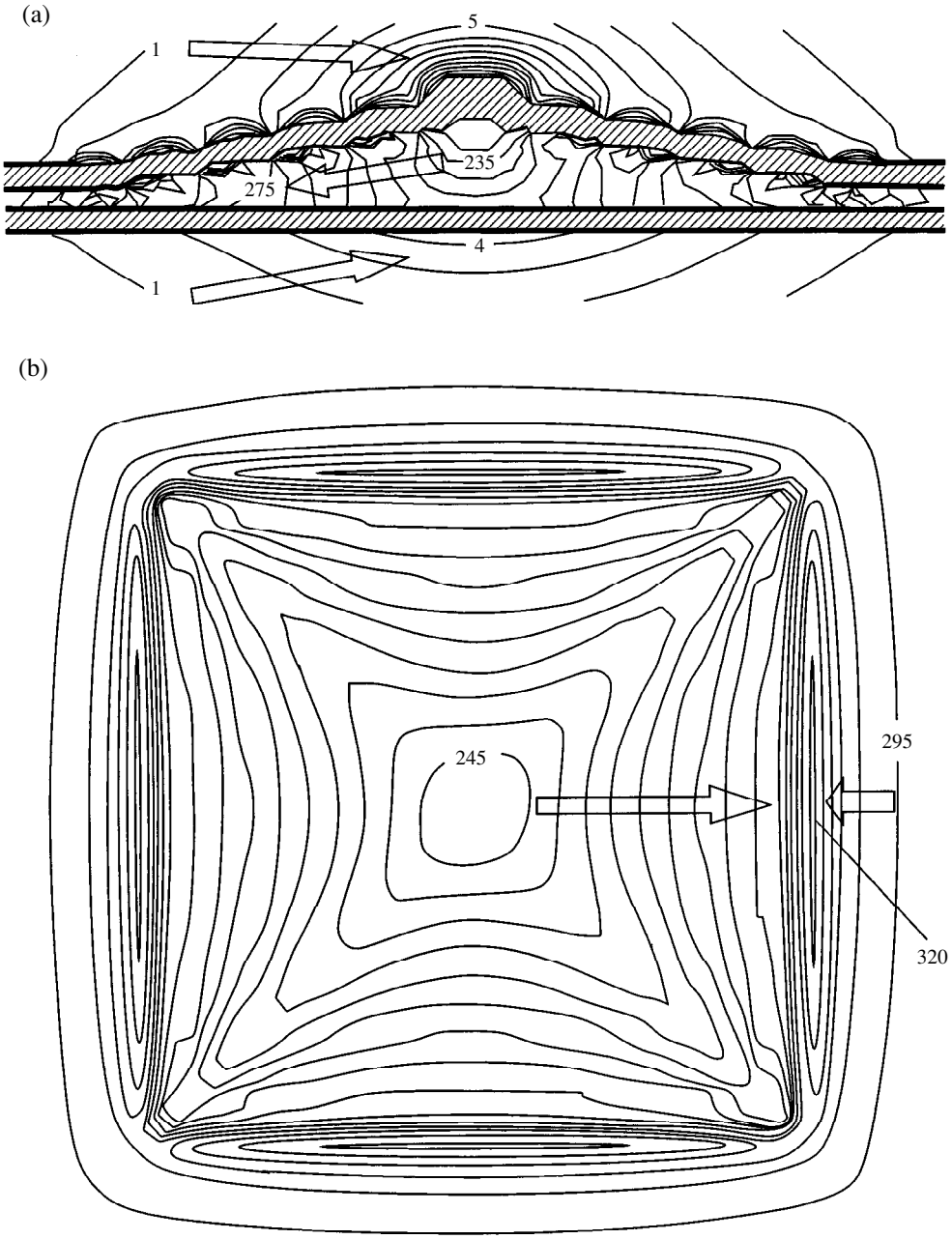

Fig. 2. Distribution of elastic energy in a Ge/Si quantum dot on the side of a 28 lattice constants base (15 nm): (a) in the (100) plane passing through the pyramid axis; (b) in the (001) plane passing through the center of a continuous Ge layer. The numbers indicate the energies in units of 10^{-4} eV per atom; the arrows show the direction of increasing energy. The spacing of the isolines is 5×10^{-4} eV inside the pyramid and 10^{-4} eV outside the pyramid. Isolines are not shown near the Si–Ge interface (≈4 monolayers).

As one can easily see (see, for example, the analytic solution for a cubic quantum dot [19]), according to the macroscopic theory of elasticity the deformation tensor at a fixed distance from the edge should grow logarithmically as a function of the cluster size; our results demonstrate this (Fig. 4).

To show the structure of the deformation field of a quantum dot and its environment, the deformation tensor profile along the pyramid axis for a 15 nm quantum dot is shown in Fig. 5. Inside the pyramid ε_{xx} and $\varepsilon_{yy} < 0$ but $\varepsilon_{zz} > 0$. This means that compression occurred in the lateral direction and stretching occurred in the vertical direction. The opposite picture is observed in the environment around a quantum dot: stretching occurs in the lateral direction and compression in vertical direction.

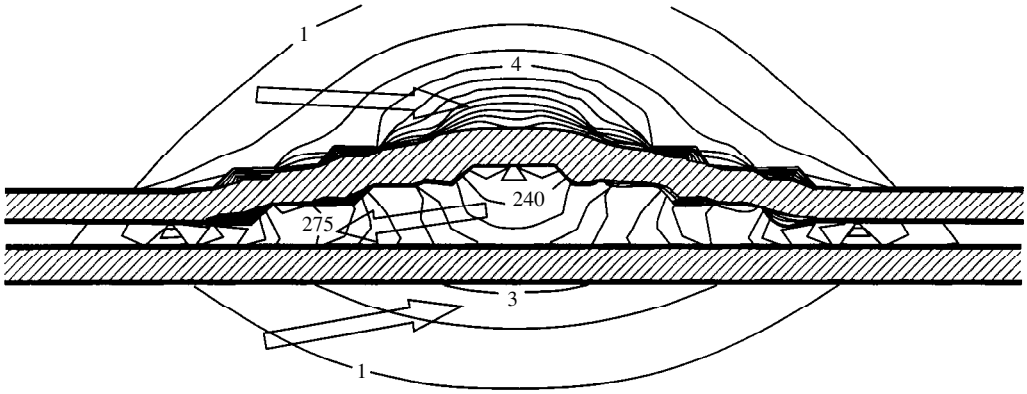

Fig. 3. Distribution of elastic energy in quantum dots on the side of an 18 lattice constants base in the (100) section. The notation is the same as in Fig. 2.

The spatial distribution obtained for the deformation can be compared with the results of similar calculations for InAs/GaAs quantum dots. It should be kept in mind that the distribution of the deformations depend strongly on the slope angle of the lateral faces [5]. The profiles of the components of the deformation tensor which are presented in Fig. 5 are similar to the profiles obtained for InAs/GaAs quantum dots with {104} lateral faces [5]. For a comparatively large slope of the faces, for example, for quantum dots bounded by {101} planes [4–6, 20], a qualitative difference from our results appears: the deformation inside quantum dots becomes strongly nonuniform and even sign-alternating. As a result, a region of purely hydrostatic compression arises approximately at the center of the quantum dots. This is important, since the band picture depends on the structure of the deformation: splitting of the light- and heavy-hole subbands does not occur under hydrostatic compression, and under anisotropic deformation, as in Fig. 5, the subbands split by ~0.15 eV (see [3]).

The computational results for the deformation in a heterostructure depend on whether the continuous medium approximation or the VFF model is used. These two models are compared in [20]. We note that near heteroboundaries the VFF model and the continuous medium approximation give substantially different results. For all other regions the difference between the two models is due primarily to the fact that the Keating model does not describe the elastic properties of the medium accurately. This discrepancy is much smaller for Ge/Si heterostructures than for InAs/GaAs structures. The errors in fitting the elastic properties using the parameters of the Keating model are 1% and 7%, respectively, for Si and Ge and 13% and 22%, respectively, for GaAs and InAs [20].

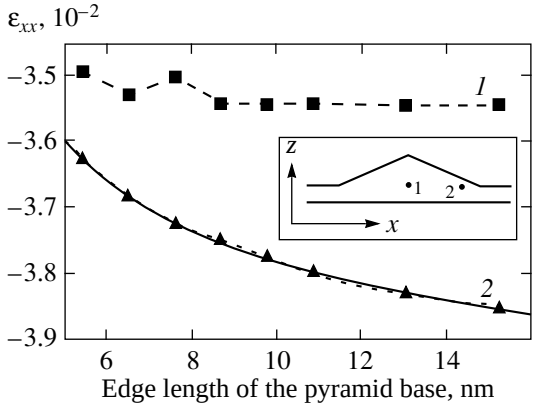

Fig. 4. The component ε_{xx} of the deformation tensor versus the size of a quantum dot: (*1*) at the center of the pyramid base; (*2*) at distance of 1.5 lattice constants from the center of the edge of the pyramid base (see inset). Solid line is result of fitting a function of the form $A + B\ln(a + C)$ to the curve *2*, where *a* is the size of the quantum dot.

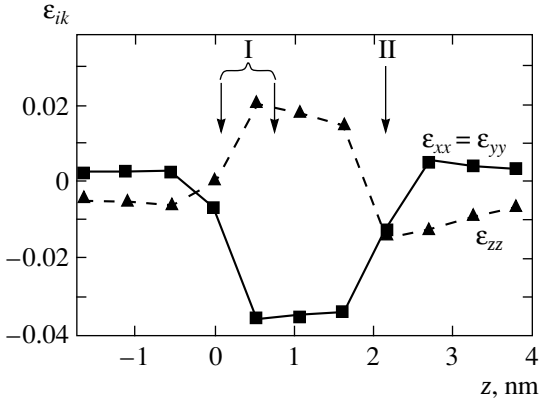

Fig. 5. Profiles of the components of the deformation tensor along the symmetry axis of a quantum dot. The step along *z* is equal to the lattice constant. The numeral I denotes the region occupied by a solid film of Ge and the numeral II denotes the tip of the pyramid.

Another difference between the two models, which was noted in [20], is that when the VFF model is used the symmetry of the deformation field decreases to C_{2v} (C_{4v} in the continuous-medium model). We add also that in the continuous-medium model the distribution of the deformation depends only on the shape of the inclusion and not its size [4]. A size dependence of the deformation appears in the VFF model.

In [21] it was concluded on the basis of an analysis of the Raman scattering spectra in a Ge/Si structure with quantum dots that directly below the tip of a quantum dot there exists a region where the deformation is much different from the deformation elsewhere in the quantum dot. The thickness of this region is 0.4 nm [21]. Our results (Fig. 5) also showed the existence of a region with a size of about 0.3 nm below the tip of the quantum dot, where the component ε_{zz} of the deformation tensor is negative, while elsewhere in the quantum dot it is positive.

4. CONCLUSIONS

In summary, in the present paper a method was developed for calculating the elastic deformation fields in nonuniform structures on the basis of an atomistic approach. The distribution of deformations in Ge/Si heterostructures with quantum dots for 6–15 nm pyramid bases was obtained. It was found that the deformations and their spatial distribution are essentially independent of the size of a quantum dot in the central region of a pyramid for quantum dots larger than 10 nm. At the same time, near the edge of a pyramid base the deformation increases logarithmically as a function of the size of the quantum dots.

We propose to use the results obtained to calculate the energy spectrum of holes in Ge/Si quantum dots. These quantum dots are so small that there is no guarantee that the effective-mass method is applicable for describing the motion of electrons in them. Consequently, following [7], the tight-binding method will be used to obtain the energy spectrum. Thus, even though the calculation of the deformation can be done in the continuous-medium approximation, the atomistic approach, which gives direct information about specific atomic bonds, is preferred. Moreover, an appreciable number of the atoms in a quantum dot lie on the Si–Ge heteroboundary, and for these atoms a description in terms of the continuous-medium approximation is unsuitable.

ACKNOWLEDGMENTS

We thank A. I. Yakimov and N. P. Stepina for a discussion of the results, valuable remarks, and unfailing interest in this work. This work was supported by the Interdepartmental Scientific Program “Universities of Russia—Basic Research” (project no. 015.01.01.34), the program “Surface atomic structures” (project no. 1.14.99), and in part by the Russian Foundation for Basic Research (project no. 00-02-17885).

REFERENCES

1. A. V. Dvurechenskiĭ and A. I. Yakimov, Izv. Vyssh. Uchebn. Zaved., Mater. Élektron. Tekh., No. 4, 4 (1999).
2. N. N. Ledentsov, V. M. Ustinov, V. A. Shchukin, *et al.*, Fiz. Tekh. Poluprovodn. (St. Petersburg) **32**, 385 (1998) [Semiconductors **32**, 343 (1998)].
3. M. M. Rieger and P. Vogl, Phys. Rev. B **48**, 14 276 (1993).
4. M. Grundmann, O. Stier, and D. Bimberg, Phys. Rev. B **52**, 11 969 (1995).
5. M. A. Cusack, P. R. Briddon, and M. Jaros, Phys. Rev. B **56**, 4047 (1997).
6. H. Jiang and J. Singh, Phys. Rev. B **56**, 4696 (1997).
7. T. Saito, J. N. Schulman, and Y. Arakawa, Phys. Rev. B **57**, 13 016 (1998).
8. S. Noda, T. Abe, and M. Tamura, Phys. Rev. B **58**, 7181 (1998).
9. L.-W. Wang, J. Kim, and A. Zunger, Phys. Rev. B **59**, 5678 (1999).
10. A. D. Andreev, J. R. Downes, D. A. Faux, and E. P. O’Reilly, J. Appl. Phys. **86**, 297 (1999).
11. I. Daruka, A.-L. Barabasi, S. J. Zhou, *et al.*, Phys. Rev. B **60**, R2150 (1999).
12. M. J. P. Musgrave and J. A. Pople, Proc. R. Soc. London, Ser. A **268**, 474 (1962).
13. P. N. Keating, Phys. Rev. **145**, 637 (1966).
14. R. M. Martin, Phys. Rev. B **1**, 4005 (1970).
15. L. D. Landau and E. M. Lifshitz, *Course of Theoretical Physics*, Vol. 7: *Theory of Elasticity* (Nauka, Moscow, 1982; Pergamon, New York, 1986).
16. I. M. Lifshits and L. N. Rozentsveĭg, Zh. Éksp. Teor. Fiz. **17**, 783 (1947).
17. W. A. Harrison, *Electronic Structure and the Properties of Solids: The Physics of the Chemical Bond* (Freeman, San Francisco, 1980; Mir, Moscow, 1983), Vol. 1.
18. A. I. Yakimov, A. V. Dvurechenskiĭ, Yu. Yu. Proskuryakov, *et al.*, Appl. Phys. Lett. **75**, 1413 (1999).
19. J. R. Downes, D. A. Faux, and E. P. O’Reilly, J. Appl. Phys. **81**, 6700 (1997).
20. O. Stier, M. Grundmann, and D. Bimberg, Phys. Rev. B **59**, 5688 (1999).
21. A. B. Talochkin, V. A. Markov, A. I. Nikiforov, and S. A. Tiĭs, Pis’ma Zh. Éksp. Teor. Fiz. **70**, 279 (1999) [JETP Lett. **70**, 288 (1999)].

Translation was provided by AIP

Journal of Experimental and Theoretical Physics, Vol. 91, No. 3, 2000, pp. 504–514.
Translated from Zhurnal Éksperimental'noĭ i Teoreticheskoĭ Fiziki, Vol. 118, No. 3, 2000, pp. 579–591.

SOLIDS
Structure

Diffraction and Channeling in Nanotubes

N. K. Zhevhago* and V. I. Glebov

Russian Science Centre Kurchatov Institute, 123182 Moscow, Russia
**e-mail: ab2211@technologist.com*

Received January 18, 2000

Abstract—A theory of the interaction of fast charged particles and gamma rays with nanotubes with different helicity is developed. Analytical expressions are obtained for the potential and the electron density of a nanotube taking account of the anisotropic thermal vibrations of the atoms. A system of equations determining the quantum states of the transverse motion of relativistic electrons, positrons, and x-ray photons in a superlattice consisting of nanotubes is formulated, and methods for solving this system are developed. Calculations of the soft x-ray Bragg reflection coefficients of a superlattice are performed in the two-wave approximation of the dynamical theory of diffraction. © *2000 MAIK "Nauka/Interperiodica".*

1. INTRODUCTION

Nanotubes can be regarded as gigantic carbon molecules (fullerenes) which are obtained by folding graphite planes into a cylinder whose diameter is measured in nanometers and whose length can reach macroscopic dimensions [1]. This linear structure of nanotubes determines their extremely high strength [2] and the strong dependence of the electrical conductivity on the diameter and helicity (the angle between the most highly packed chains of atoms and the axis of the cylinder) [3]. There exist nanotubes whose walls contain a single layer of atoms [4] and nanotubes with walls consisting of several concentric layers [5]. A remarkable feature of some single-wall nanotubes is their capability to unite, as they form from plasma, into a "rope" whose transverse cross section is a two-dimensional hexagonal superlattice consisting of nanotubes [6]. Since the discovery of nanotubes in 1991 [5], a large number of works have appeared on the problems of synthesizing nanotubes, the physical properties and possible applications in nanoelectronics [7], catalysis [8], and other fields. Besides nanotubes consisting of carbon atoms similar structures based on boron nitride also exist.

As noted in [9], channeling of fast particles in nanotubes exhibits a number of special features associated with the relatively large diameter of the channels and a weaker influence of dechanneling factors than in ordinary crystals. These features (in prospect of synthesis of quite long defect-free nanotubes) could find application in high-energy particle-beam physics. Channeling and electromagnetic radiation spectra arising during channeling of relativistic electrons and positrons with energy above 100 MeV have been analyzed previously [10] on the basis of classical mechanics by numerical simulation of particle trajectories. It was shown in [10] that channeling of neutral particles (x-ray photons and thermal neutrons) becomes possible when the wall thickness and diameter of a nanotube are sufficiently large. The present work is devoted to a further development of the theory of the interaction of fast charged and neutral particles with nanotubes. First, an analytical method for calculating the potential and electron density distribution in superlattices consisting of nanotubes will be developed, and a general approach to the problem of the quantum states of relativistic electrons (positrons) and x-ray photons, propagating at sufficiently small angles with respect to the axis of nanotubes or the planes of a superlattice, will be formulated on the basis of this method. Second, the electromagnetic radiation spectra of fast particles in nanotube superlattices will be analyzed for cases where quantum effects are substantial. Further, calculations of the Bragg reflection coefficients for soft x-rays with different wavelengths will be performed on the basis of the two-wave theory of diffraction.

2. STRUCTURE AND CONTINUOUS POTENTIAL OF NANOTUBES

As noted above, the structure of a nanotube is determined by the method of constructing the nanotube from a crystallographic plane of graphite (Fig. 1a). Let **a** and **b** denote the basis vectors of a planar lattice of graphite, whose unit cell (dashed rhombus in Fig. 1a) contains two carbon atoms with the coordinates (**a** + **b**)/3 and 2(**a** + **b**)/3, respectively. Taking into account that the angle between the vectors **a** and **b** is $\pi/3$ and the vectors have the same modulus ($a = b$) and denoting by l the length of the bond between the carbon atoms (which is usually 0.14 nm), we obtain $a = b = l\sqrt{3}$. The roll-up vector $\mathbf{r}_0$ is determined as a linear combination $\mathbf{r}_0 = n\mathbf{a} + m\mathbf{b}$ of the basis vectors, where the pair (n, m) of integers is called the indices of a nanotube. A nanotube can be constructed from a strip of width r_0, cut from a plane perpendicular to the vector $\mathbf{r}_0$ and folded into a cylinder (a row of such cylinders is shown sche-

1063-7761/00/9103-0504$20.00

matically in Fig. 1b). The angle θ between the roll-up vector $\mathbf{r}_0$ and the basis vector **a** of the graphite lattice is called the helicity of a nanotube. The helicity shows the angle under which the most closely packed chains of carbon atoms are wound on the cylindrical surface of a nanotube. It follows from simple geometric considerations that the radius R and helicity θ of a single-wall nanotube are uniquely determined by the indices of this nanotube:

$$R = \frac{l\sqrt{3}}{2\pi}\sqrt{n^2 + nm + m^2}, \qquad \theta = \arctan\left(\frac{m\sqrt{3}}{m + 2n}\right). \tag{1}$$

Since the basis vectors **a** and **b** are equivalent, it can be assumed with no loss of generality that $n \geq m$ and therefore the helicity lies in the range $0 \leq \theta \leq \pi/6$.

It is well known [11] that channeling occurs when the angles between the momenta of fast charged particles and the chains of atoms are sufficiently small. Then the potential of the chain atoms, averaged over the entire length of a chain, is the effective potential acting on a particle. The form factor $f(\mathbf{k})$ of an individual carbon atom can be represented, to a high degree of accuracy, in the form [12]

$$f(\mathbf{k}) = 4\pi Z e a_j \exp[-\mathbf{k}^2/(4b_j^2)], \tag{2}$$

where Ze is the nuclear charge, $a_j = \{3.222, 5.270, 2.012, 0.5499\}10^{-4}$ nm^2, and $b_j = \{10.330, 18.694, 37.456, 106.88\}$ nm^{-1} are dimensional parameters which are determined, according to [12], from the best fit of Eq. (2) to the most accurate calculations based on the Hartree–Fock method, and summation over repeated indices j is assumed here and below (in the present case from 1 to 4).

Small displacements of the atoms of a nanotube relative to the positions of equilibrium in an ideal lattice can be due to thermal vibrations as well as other factors (for example, structural defects or deformations of nanotubes). In general, these displacements are anisotropic and can be taken into account by introducing an additional factor of the form e^{-W} into the atomic form factor (2), where the exponent (the Debye–Waller factor) is given by

$$W = (1/2)(k_r^2 u_r^2 + k_\tau^2 u_\tau^2 + k_z^2 u_z^2), \tag{3}$$

where u_r^2, u_τ^2, and u_z^2 are the mean-square deviations of the atoms in various directions: normal to the circumference of the cylinder (y), tangential (x), and along the axis of the cylinder (z).

The average potential $V_R(\rho)$ of the atomic chain (more accurately, the potential energy of a particle with charge e), whose atoms lie on the z-axis with spacing d_R, is related with the form factor of an individual atom by the relation

$$V_R(\rho) = e(2\pi)^{-2} d_R^{-1} \int f(\mathbf{k}_\perp, 0) e^{i\mathbf{k}_\perp \boldsymbol{\rho}} d\mathbf{k}_\perp. \tag{4}$$

Here **ρ** is the radius vector in a plane orthogonal to the chain, $(\mathbf{k}_\perp, k_z)$ are the corresponding components of the vector **k**, and the integration is performed over the transverse momenta $\mathbf{k}_\perp$. Using the relations (2)–(4) we obtain the average potential of a chain of atoms

$$V_R(x, y) = \frac{4Ze^2}{d_R\sqrt{\beta_j^{(r)}\beta_j^{(\tau)}}} a_j \exp\left[-\frac{x^2}{\beta_j^{(r)}} - \frac{y^2}{\beta_j^{(\tau)}}\right], \tag{5}$$

where $\beta_j^{(r)} = b_j^{-2} + 2u_r^2$ and $\beta_j^{(\tau)} = b_j^{-2} + 2u_\tau^2$. For an ideal static structure ($u_r^2 = u_\tau^2 = 0$) Eq. (5) is identical to the corresponding result

$$V_R(\rho) = \frac{4Ze^2}{d_R} a_j b_j^2 \exp[-b_j^2 \rho^2], \tag{6}$$

obtained previously in [10]. Here $\rho = (x^2 + y^2)^{1/2}$ denotes the distance to the chain.

For subsequent calculations, it is important that a nanotube with indices (n, m) can always be represented as a collection of chains of atoms parallel to the axis of a nanotube and arranged in a definite manner along the perimeter of the cylinder. Indeed, it is easy to show that the potential of a nanotube is invariant under translations by the vector

$$\mathbf{t} = q^{-1}[(2m + n)\mathbf{a} - (2n + m)\mathbf{b}],$$

and, since $\mathbf{tr}_0 = 0$, this vector directed along the axis of the nanotube. Here q denotes the largest common denominator of $(2m + n)$ and $(2n + m)$, so that t is the smallest of all possible translation vectors of this kind. The distance d_R between neighboring atoms in each such a chain is equal to the modulus of the vector t and is given by

$$d_R = \frac{3l}{q}(n^2 + nm + m^2)^{1/2}. \tag{7}$$

We shall now take into account the fact that the unit cell contains two atoms on a crystal plane of graphite (see Fig. 1a). It can be easily calculated that the surface density σ of atoms is $\sigma = 3^{-3/2}4l^{-2}$. Thus, two sequences of atomic chains parallel to the axis of the nanotube are obtained. In each sequence the chains are equally spaced along the generatrix of the cylinder, and the distance δ between the neighboring chains in a given sequence is $\delta = 2(\sigma\delta_R)^{-1}$. Next, we choose a cylindrical coordinate system (r, φ, z) in which the radial coordinate r is measured from the axis of a nanotube, and we denote by $\varphi_k^{(\mu)}$ the azimuthal angle of the kth chain (k =

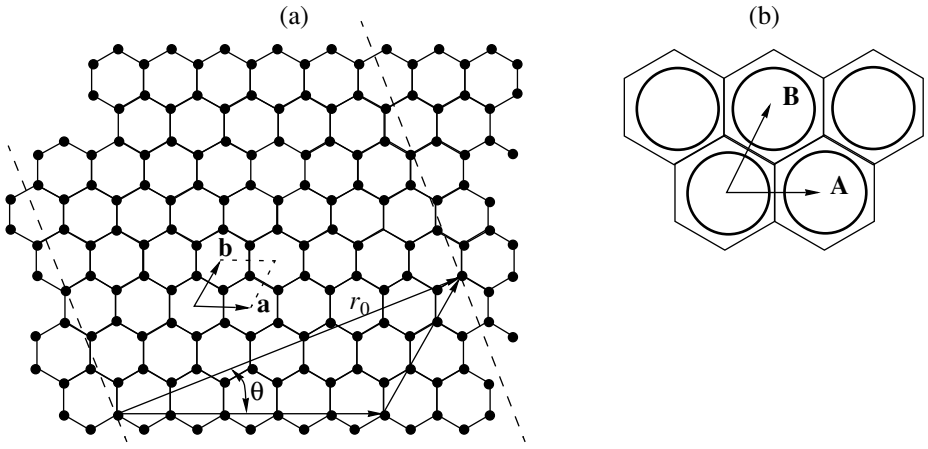

Fig. 1. (a) The atomic structure of a nanotube and (b) the structure of a nanotube superlattice.

0, 1, 2, …, $N - 1$) in the corresponding sequence of chains ($\mu = 1, 2$). Thus each sequence contains exactly

$$N = \frac{2}{q}(n^2 + nm + m^2), \tag{8}$$

atomic chains, whose azimuthal coordinates are determined by the relations

$$\varphi_k^{(1)} = k\delta\varphi, \quad \delta\varphi = \frac{\pi q}{n^2 + nm + m^2}, \tag{9}$$

$$\varphi_k^{(2)} = \varphi_k^{(1)} + \Delta\varphi, \quad \Delta\varphi = \frac{\pi(n+m)}{n^2 + nm + m^2}. \tag{10}$$

We note two limiting helicities $\theta = 0$ and $\theta = \pi/6$, where the atomic chains, parallel to the axis of a nanotube, are most densely packed and therefore the number of chains is relatively small. The first case corresponds to nanotubes (n, 0) with zero index m (which are usually called zigzag nanotubes); here, $q = n$, $N = 2n$, $d_R = 3l$, $\delta\varphi = \Delta\varphi = \pi/n$, and the two chain sequences overlap, i.e., there are actually $2n$ chains with a doubled linear atomic density $2/d_R$. The other case corresponds to nanotubes with equal indices (n, n) (which are usually called armchair nanotubes); here, $q = 3n$, $N = 2n$, $d_R = l\sqrt{3}$, $\delta\varphi = \pi/n$, and $\Delta\varphi = 2\pi/(3n)$, i.e., there are $4n$ chains arranged in pairs.

The potential of a nanotube is obtained by summing the average potentials of all chains. To perform such a summation we represent the potential of an individual chain (5) as a Fourier series

$$V_R = \sum_{\nu=-\infty}^{\infty} V_\nu(r) e^{i\nu\varphi}. \tag{11}$$

The coefficients $V_\nu(r)$ in the expansion (11) are related with the potential V_R of the chain by the relation

$$V_\nu(r) = \frac{1}{2\pi}\int_0^{2\pi} V_R e^{-i\nu\varphi} d\varphi, \tag{12}$$

where V_R is a function of the cylindrical coordinates. Specifically, when the average potential of a chain has the form (5), the coordinates x and y are related with the cylindrical coordinates r and φ by the relations $x = r\sin\varphi$ and $y = r\cos\varphi - R$; the expansion coefficients become

$$V_\nu(r) = \frac{4Ze^2}{d_R\sqrt{\beta_j^{(r)}\beta_j^{(\tau)}}} a_j \times \exp\left[-\frac{R^2}{\beta_j^{(r)}} - \frac{r^2}{2}\left(\frac{1}{\beta_j^{(r)}} + \frac{1}{\beta_j^{(\tau)}}\right)\right] \sum_{\mu=-\infty}^{\infty} I_\mu(\gamma_j) I_{\nu-2\mu}(\alpha_j). \tag{13}$$

Here $\alpha_j = 2Rr/\beta_j^{(r)}$, $\gamma_j = (1/\beta_j^{(\tau)} - 1/\beta_j^{(r)})r^2/2$, I_μ is the modified Bessel function. If it is assumed that the displacements of the atoms are distributed isotropically ($u_r^2 = u_\tau^2$), then $\beta_j^{(\tau)} = \beta_j^{(r)}$ and $\gamma_j = 0$, so that only one term with $\mu = 0$ remains in the sum (13) and a simpler expression is obtained:

$$V_\nu(r) = \frac{4Ze^2}{d_R\beta_j^{(r)}} a_j \exp\left(-\frac{R^2 + r^2}{\beta_j^{(r)}}\right) I_\nu(\alpha_j). \tag{13'}$$

In the static limit ($u_r^2 = u_\tau^2 = 0$), corresponding to the potential of the chain (6), the result acquires the simpler form

$$V_\nu(r) = \frac{4Ze^2}{d_R} a_j b_j^2 \exp[-b_j^2(r^2 + R^2)] I_\nu(2b_j^2 Rr). \tag{14}$$

Using the Fourier expansion (11), the potential produced by the first sequence of chains in the nanotube can be represented in the form

$$U^{(1)}(r, \varphi) = \sum_{\nu=-\infty}^{\infty} V_\nu(r) e^{i\nu\varphi} \sum_{k=0}^{N-1} e^{-i\nu k\delta\varphi}.$$

Taking account of Eqs. (7) and (8), the summation over the chains now reduces to calculating the sum of a geometric progression, which can be written in the form

$$\sum_{k=0}^{N-1} e^{-i\nu k\delta\varphi} = N\delta_{\nu, sN}, \quad s = 0, \pm 1, \pm 2, \ldots,$$

where $\delta_{\nu, sN}$ is the Kronecker delta. We now take account of the fact that two sequences of chains contribute to the potential $U(r, \varphi)$ of a nanotube, i.e.,

$$U(r, \varphi) = U^{(1)}(r, \varphi) + U^{(1)}(r, \varphi - \Delta\varphi).$$

Here $\Delta\varphi$ is the azimuthal shift between the sequences of chains, which was introduced in Eq. (10). As a result we arrive at the following expression for the average potential of a nanotube with arbitrary indices (n, m):

$$U(r, \varphi) = 2N\left\{V_0(r) + s\sum_{s=1}^{\infty} V_{sN}(r)\cos\left[\frac{\pi s(n+m)}{q}\right] \times \cos\left[sN\varphi - \frac{\pi s(n+m)}{q}\right]\right\}. \quad (15)$$

For the model of an atom corresponding to the form factor (2), the coefficients $V_\nu(r)$ have the form (13), (13'), or (14), and since the potential (15) is appreciably different from zero only close to ($|R - r| \ll R$) the walls of a nanotube, where the arguments of the modified Bessel function are quite large ($\alpha_j \gg 1$), the Bessel functions can be replaced by the asymptotic expressions of the form [13]

$$I_{sN}(\alpha_j) \simeq (2\pi\alpha_j)^{-1/2}(\xi^2 + 1)^{-1/4} \times \exp\{\alpha_j[\sqrt{\xi^2 + 1} - \xi\ln(\xi + \sqrt{\xi^2 + 1})]\}.$$

Here $\xi = sN/\alpha_j$ is the ratio of the index of the Bessel function to its argument.

The coefficients V can be represented in an analytical form for some other models of an atom as well. Specifically, for the Moliere model, which is commonly used for calculating the average potentials [9, 11], the expression analogous to Eq. (14) in the range $r \le R$ is

$$V_\nu(r) = \frac{2Ze^2}{d_R}\tilde{a}_j I_\nu\left(\frac{\tilde{b}_j r}{a_{TF}}\right)K_\nu\left(\frac{\tilde{b}_j R}{a_{TF}}\right). \quad (16)$$

Here summation over the repeated index j from 0 to 3 is assumed; $\{\tilde{a}_j\} = (0.35, 0.55, 0.1\}$ and $\{\tilde{b}_j\} = \{0.3, 1.2, 6.0\}$ are the constants in the Moliere model, a_{TF} is the Thomas–Fermi radius, and the potentials in the region $r \ge R$ are obtained from Eq. (16) by interchanging the arguments in the modified Bessel functions I_ν and K_ν. It should be noted that the Moliere approximation is a fit to the corresponding calculations of the atomic potential on the basis of the Thomas–Fermi model, which, as is well known, is not accurate enough for light atoms, specifically carbon atoms.

We shall now analyze the general expression (15) for the average potential of a nanotube. The zero harmonic ($s = 0$) in Eq. (15) corresponds to a potential averaged over the azimuthal angle. The remaining harmonics lead to modulation of the azimuthal distribution. As the number of chains forming the walls of a nanotube increases, the modulation frequency increases and the modulation depth decreases. The distance between the neighboring chains (in each of two sequences) is determined by the quantity $\delta = 3^{-1/2}2^{-1}ql(n^2 + nm + m^2)^{-1/2}$. If δ is much less than the range of the atomic potential, then analysis of Eq. (15) shows that the higher order harmonics ($s \ge 1$) are negligibly small. This condition holds for indices (m, n) for which the parameter q (see Eq. (7)) is much less than the number N of chains. Confining attention to the zeroth harmonic ($s = 0$) and neglecting the random deviations of atoms from the positions of equilibrium ($u_r^2 = u_\tau^2 = 0$), we arrive at the following expressions for the potential of a nanotube:

$$U(r, \varphi) = 3^{-2/3} \times 32\pi Ze^2 l^{-2} R a_j b_j^2 \times \exp[-b_j^2(r^2 + R^2)] I_0(2b_j^2 rR) \quad (17)$$

for the Doyle–Turner approximation (2) and

$$U(r, \varphi) = 3^{-3/2} \times 16\pi Ze^2 l^{-2} R\tilde{a}_j I_0\left(\frac{\tilde{b}_j r}{a_{TF}}\right)K_0\left(\frac{\tilde{b}_j R}{a_{TF}}\right) \quad (18)$$

for the Moliere approximation ($r \le R$), and close to ($|R - r| \lesssim R$) the walls of the nanotube the arguments of the cylindrical functions are quite large and the expressions (17) and (18) can be replaced by the simpler asymptotic representations:

$$U(r, \varphi) = 4\pi^{1/2}\sigma Ze^2 a_j b_j\left(\frac{R}{r}\right)^{1/2} \times \exp[-b_j^2(R - r)^2], \quad (17')$$

$$U(r, \varphi) = 2\pi Ze^2\sigma a_{TF}\frac{\tilde{a}_j}{\tilde{b}_j}\left(\frac{R}{r}\right)^{1/2}\exp\left[-\frac{\tilde{b}_j|R - r|}{a_{TF}}\right]. \quad (18')$$

According to Eqs. (17') and (18') the potential of nanotubes with intermediate helicity (when $q \ll N$) is virtually independent of the indices of the nanotube and near the top of the barrier ($r \approx R$) it is identical with the potential of a graphite plane averaged over the coordinates of the atoms on the plane (we recall that $\sigma = 3^{-3/2}4l^{-2}$ is the surface density of atoms, and $|R - r|$ must be interpreted as the distance to the plane). The depen-

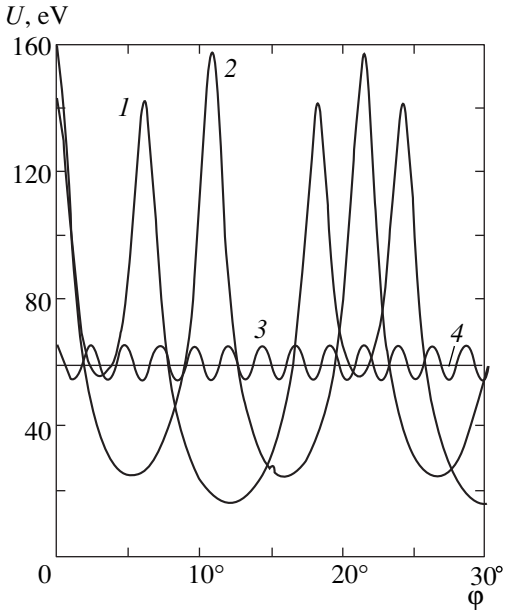

Fig. 2. The azimuthal distribution of the height of the potential barrier $U(r, \varphi)$ of nanotubes with various indices (n, m). The curves *1*, *2*, and *3* correspond to nanotube indices (10,10), (17,0), and (12,8), respectively. The horizontal straight line *4* corresponds to nanotubes with intermediate helicity, for example, (11,9).

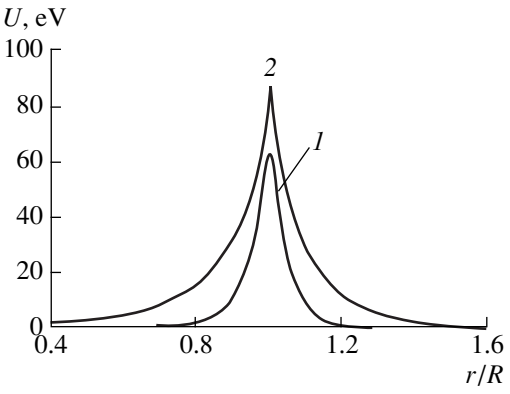

Fig. 3. The average potential of a nanotube with intermediate helicity according to (*1*) the Doyle–Turner model and (*2*) the Moliere model.

dence of the potential on the indices in this case appears only at sufficiently large distances ($|R - r| \sim R$) from the nanotube wall, where the curvature of the wall becomes substantial, but in this region the values of the potential and its gradient are relatively small.

The azimuthal distribution of the height of the potential barrier $U(R, \varphi)$ of nanotubes with different indices (n, m), based on Eqs. (14) and (15), is illustrated in Fig. 2. The curves *1*, *2*, and *3* correspond to nanotube indices (10,10), (17,0), and (12,8), respectively. The horizontal straight line *4* corresponds to nanotubes with the parameters $q \ll N$, for example (11,9).

Figure 3 shows the dependences of the potentials (17) and (18) on the distance to the axis of the nanotube (11,9). Calculations based on the more accurate Doyle–Turner model (curve *1*) give a sharper dependence of the potential on r near the wall of a nanotube and appreciably lower potential barrier than similar calculations based on the commonly used Moliere approximation (curve *2*).

To estimate the effect of thermal vibrations of the atoms it is necessary to know the temperature dependence of the mean-square deviations u_r^2 and u_τ^2. This reduces to analyzing the phonon spectra of the nanotubes. It should be noted that in contrast to graphite photons in nanotubes are localized in the transverse plane, and consequently u_r^2 and u_τ^2 can differ appreciably from the corresponding values in graphite. If the values $u_\perp = 8.5 \times 10^{-3}$ nm and $u_\parallel = 3.8 \times 10^{-3}$ nm presented in [14] for room-temperature graphite are nonetheless used to estimate u_r nm and u_τ, then thermal vibrations do not appreciably influence the average potential of a nanotube.

X-rays are scattered primarily by electrons of a substance. The Fourier component of the electron density in a carbon atom can be approximated to an adequate degree of accuracy by an expression similar to Eq. (2):

$$f^{(e)}(\mathbf{k}) = Za_j^{(e)} \exp\left[-\frac{k^2}{(2b_j^{(e)})^2}\right], \qquad (19)$$

where $a_j^{(e)} = \{0.3499, 0.3014, 0.2103, 0.0946, 0.0438\}$ and $b_j^{(e)} = \{17.300, 11.400, 75.501, 155.24, 7.596\}$ nm^{-1} are five pairs of fotting parameters, which are determined, according to [15], from the best fit of Eq. (19) to the corresponding values calculated by the Hartree–Fock method. The electron density averaged over the direction of the nanotubes or over the planes of the superlattice is obtained using Eq. (19) just as was done above for the potential.

In ordered two-dimensional structures of nanotubes (superlattices), in addition to the channeling inside individual nanotubes, a different type of channeling, which can be called planar channeling, becomes possible. The two-dimensional superlattice (see Fig. 1b) has the hexagonal symmetry characterized by a pair of basis vectors **A** and **B**, whose modulus is the period L of the superlattice. Nanotubes in a superlattice are kept in the position of equilibrium by van-der-Waals forces, and the gap between the walls of neighboring nanotubes is usually about 3.15 Å. For example, the superlattice with period $L \simeq 16.95$ Å, consisting of (10,10) nanotubes with diameter $2R \simeq 13.8$ Å, has been observed [6]. Let a charged particle move at a comparatively large angle with respect to the axis of the nanotubes (so that the channeling conditions inside the nanotubes are not satisfied), but the momentum of the par-

ticle makes a quite small angle with respect to one of the planes of the superlattice. Then, following Lindhard's arguments [11], it can be assumed that a particle is subjected to the effective potential of the nanotubes averaged over the corresponding planes of the superlattice:

$$U_p(x) = \sum_{\nu} U_1(x + \nu d_p),$$
$$U_1(x) = \frac{1}{d_n}\int_{-d_n/2}^{d_n/2} U(r, \varphi)dy. \tag{20}$$

Here $U_1(x)$ is the potential of an individual plane, x is the distance to the plane, the potential of a nanotube $U(r, \varphi)$ in the general case has the form (15), y is the distance to the nanotube, d_n is the distance between the neighboring nanotubes in a given plane, and summation extends over all planes ν which are arranged periodically (with period $d_p = 3^{3/2}L^2/2d_n$) in a direction orthogonal to the x-axis.

3. SCATTERING AND RADIATION OF ELECTRONS AND POSITRONS IN NANOTUBE SUPERLATTICES

The motion of relativistic charged particles at a small angle to the axis of the nanotubes (or to the planes of the superlattice), similarly to the case of crystals [16], can be represented as the longitudinal motion with constant momentum p_z and the transverse motion in an average potential. The wave functions of the transverse motion $\psi(\mathbf{r})$ satisfy the equation

$$[\Delta_r + p_\perp^2(\mathbf{r})]\psi(\mathbf{r}) = 0, \tag{21}$$

where Δ_r is the Laplacian in the space of transverse coordinates, $p_\perp^2(\mathbf{r}) = 2E(\varepsilon - U(r))$, E is the total particle energy, $\varepsilon = E - \sqrt{p_z^2 + 1}$ is the transverse energy, $U_s(r)$ is the potential of the nanotube superlattice, and the system of units $\hbar = m = c = 1$ is used. Electromagnetic radiation appears as a result of spontaneous radiative transitions between the states of transverse motion. According to [17], for radiative transitions between high-lying ($\varepsilon \gg U_0\max(1, EU_0)$, U_0 is the depth of the potential well) states of the continuous spectrum of transverse energies the superlattice potential $U_s(\mathbf{r})$ can be treated as a perturbation, the angles of deflection of an ultrarelativistic particle by the field are small compared with the effective angles of emission E^{-1}, and, consequently, the radiation is dipole one, the emission probability per unit path length of an electron (positron), differential with respect to the photon energy ω, can be related with the Fourier component of the potential $U_\mathbf{H}$ of a unit cell of the superlattice by the simple relation (see, for example, [18])

$$\frac{dw}{d\omega} = \frac{e^2}{E^2}\sum_{\mathbf{H}} \frac{H^2}{(\mathbf{H}\nu)^2}|U_\mathbf{H}|^2$$
$$\times\left[1 + \frac{u}{2(1+u)} - 2\Omega_\mathbf{H} + 2\Omega_\mathbf{H}^2\right]\eta(1 - \Omega_\mathbf{H}). \tag{22}$$

Here $\boldsymbol{\nu}$ is a unit vector in the direction of the initial momentum of a charged particle, $u = \omega/(E - \omega)$, $\Omega_\mathbf{H} = u/2E\mathbf{H}\boldsymbol{\nu}$, and η is the Heavside unit step function.

A unit cell can be chosen so that it contains a single nanotube (see Fig. 1b), and the reciprocal-lattice vectors $\mathbf{H} = n_1\mathbf{h}_1 + n_2\mathbf{h}_2$ are represented as a superposition of basis vectors $\mathbf{h}_1$ and $\mathbf{h}_2$ (orthogonal to the corresponding vectors of the direct superlattice) with integer coefficients n_1 and n_2. The moduli of the vectors are determined by the relations $h = 4\pi/L\sqrt{3}$ and $H = h(n_1^2 + n_2^2 + n_1n_2)^{1/2}$. As a result, the expression for $U_\mathbf{H}$ becomes

$$U_\mathbf{H} = \frac{1}{S}\int_S U(r, \varphi)e^{-i\mathbf{H}\mathbf{r}}d^2r, \tag{23}$$

where $U(r, \varphi)$ is the potential of an individual nanotube (15). The integration in Eq. (23) extends over the area of a unit cell $S = L^2\sqrt{3}/2$, but, since outside a unit cell the potential of a nanotube is negligibly small, the integration in Eq. (3) can be extended over the entire transverse plane without substantial loss of accuracy. Substitution of Eq. (15) into Eq. (20) and the subsequent integration give

$$U_\mathbf{H} = \frac{N}{S}\frac{8\pi Ze^2}{d_R}a_j\exp\left(-\frac{H^2}{(2b_j)^2}\right)$$
$$\times\left\{J_0(RH) + 2\sum_{s=1}^{\infty} J_{sN}(RH)\cos\left[\frac{\pi s(n+m)}{q}\right]\right.$$
$$\left.\times\cos\left[sN\varphi_\mathbf{H} - \frac{\pi s(n+m)}{q}\right]\right\}, \tag{24}$$

where $\varphi_\mathbf{H}$ is the azimuthal angle of the vector $\mathbf{H}$. We note that the higher order azimuthal harmonics with $s \geq 1$ can make an appreciable contribution to Eq. (24) only if the argument of the Bessel functions is comparable in magnitude to the index, i.e., when the vector $\mathbf{H}$ is a large multiple of the main reciprocal-lattice vectors ($H/h \sim N$). However, if the analysis is confined to processes where $H \ll hN$, then with adequate degree of

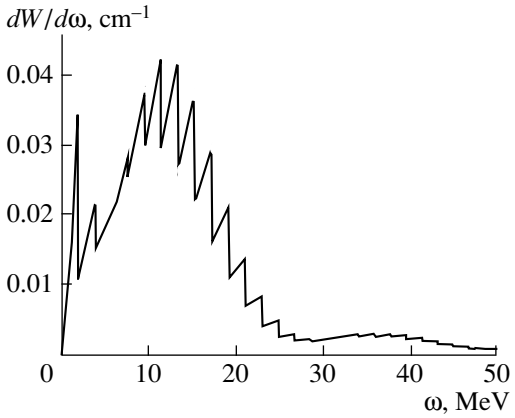

Fig. 4. Spectral distribution of the energy of coherent bremsstrahlung emitted by a 1 GeV electron (positron) entering at angle 3.04×10^{-4} with respect to the (1,0) planes of a superlattice of (10,10) nanotubes.

accuracy simpler expressions can be used instead of Eq. (24):

$$U_0 = \frac{KZe^2R}{(R+g/2)^2}, \quad U_{\mathbf{H}} = U_0\Im(\mathbf{H}),$$

$$\Sigma = \sum_j a_j \simeq 1.105 \times 10^{-3}\ \text{nm}^2, \qquad (25)$$

$$\Im(\mathbf{H}) = \Sigma^{-2} a_j \exp\left(-\frac{H^2}{(2b_j)^2}\right) J_0(RH).$$

Here we have introduced the structure factor $\Im(\mathbf{H})$ of a nanotube, normalized by the condition $\Im(0) = 1$. Ordinarily, the van-der-Waals gap $g = L - 2R$ is 3.15 Å, as noted above.

If it is assumed at the outset that there is no azimuthal correlation between different nanotubes of a superlattice, then the form factor of a nanotube (24) must be additionally averaged over $\varphi_{\mathbf{H}}$. As a result, we arrive once again at the expression (25), where, however, **H** can now be regarded as arbitrarily large compared with the modulus of the basis vector of the reciprocal lattice $h = 4\pi/L\sqrt{3}$. The nonideality of a nanotube superlattice can be taken into account by multiplying Eqs. (24) and (25) by the exponential factor $\exp(-W_s)$, where the Debye–Waller factor is $W_s = H^2 u_s^2/2$, where u_s^2 now denotes the mean-square deviation of the superlattice sites from their positions in an ideal structure. We note that in a real superlattice ($u_s^2 \neq 0$), besides coherent bremsstrahlung, incoherent background radiation corresponding to scattering of particles by individual nanotubes should also arise. The spectral distribution of the coherent bremsstrahlung energy $dW/d\omega = \omega du/d\omega$, corresponding to a 1 GeV electron (positron) and entrance angle $\theta_0 = 2\theta_L = 3.04 \times 10^{-4}$ with respect to the planes (1,0) of a superlattice of (10,10) nanotubes, is illustrated in Fig. 4. The maximum near 2 MeV corresponds to the first harmonic of coherent bremsstrahlung, and the intensity of the higher order harmonics (right up to the tenth harmonic) is comparable in magnitude to that of the first harmonic. This behavior of the coherent bremsstrahlung spectrum in nanotubes differs substantially from that of the bremsstrahlung spectrum in ordinary crystals. This is due to the specific nature of the structure factor $\Im(\mathbf{H})$ of a nanotube.

If radiative transitions occur between lower lying states ($\varepsilon \lesssim U_0$) of the transverse energy spectrum, then the perturbation theory with respect to the superlattice potential is inapplicable, the standard theory of coherent bremsstrahlung is inapplicable, and the wave function and energy eigenvalues satisfying Eq. (21) must first be found in order to calculate the radiation spectra. According to the Bloch theorem, the wave function of a particle in a periodic potential $U_s(\mathbf{r})$ has the form

$$\psi(\mathbf{r}) = \phi(\mathbf{r})\exp(i\boldsymbol{\kappa}\mathbf{r}), \qquad (26)$$

where $\phi(\mathbf{r})$ is a periodic function with period equal to the period of the superlattice and $\boldsymbol{\kappa}$ is the transverse quasimomentum of a particle, which can be assumed to be restricted by the first Brillouin zone. As is well known, the wave functions (26) correspond to transverse-energy bands $\varepsilon(\boldsymbol{\kappa})$, which in the limit of infinitely small tunneling through the potential barrier separating neighboring nanotubes degenerate into discrete levels. As estimates show, the number of quasidiscrete levels is relatively small when the energy of the electrons and positrons does not exceed 10 MeV. For such energies the electromagnetic radiation can be treated as being due to dipole radiative transitions between levels (bands), and the energy of an emitted photon is negligible compared to the energy of the particle ($\omega \ll E$). The spectral-angular probability density of emission due to spontaneous transition of a particle from a band n into a band n' with lower transverse energy can be represented in the form (see, e.g., [19])

$$\frac{d^2\omega}{d\omega d} = \frac{e^2\omega^3}{8\pi}$$

$$\times \sum_{n'} [(\theta^2 + E^{-2})^2 |\mathbf{n}_\perp \times \mathbf{r}_{nn'}|^2 - (\theta^2 - E^{-2})^2 |\mathbf{n}_\perp \mathbf{r}_{nn'}|^2] \qquad (27)$$

$$\times \delta\left[\frac{\omega}{2}(\theta^2 + E^{-2} - \chi_0'(\omega)) - \omega_{nn'}\right].$$

Here $d\Omega \simeq \theta d\theta d\varphi$ is the differential of the solid angle of emission, $\theta \ll 1$ is the polar angle of emission relative to the axis of a nanotube, $\mathbf{n}_\perp = \{\cos\varphi, \sin\varphi\}$ is the unit vector in the direction of the projection of the momentum of the emitted photon on a plane perpendicular to the axis of a nanotube, $\varepsilon_n(\boldsymbol{\kappa})$ is the transverse energy of a particle with quasimomentum $\boldsymbol{\kappa}$ in the nth

band, $\omega_{nn'} = \varepsilon_n(\boldsymbol{\kappa}) - \varepsilon_{n'}(\boldsymbol{\kappa})$ is the transition frequency, and δ is the Dirac delta function.

The matrix element of a dipole transition between bands has the form

$$\mathbf{r}_{nn'} = \int_S \phi^*_{n'\boldsymbol{\kappa}}(\mathbf{r})\mathbf{r}\phi_{n\boldsymbol{\kappa}}(\mathbf{r})d^2r. \qquad (28)$$

The quasimomentum of a particle is determined by the condition that the wave function is continuous at the boundary of the superlattice and is equal to (to within a reciprocal lattice vector) the transverse momentum $\mathbf{p}_{\perp 0}$ of the particle at the entrance into the superlattice; this value of the quasimomentum is conserved in radiative dipole transitions [19]. We note that, following [20], we have taken into account the possible influence of the polarization of the medium on the process of radiation by channeled particles by introducing the volume-averaged real part of the permittivity of the medium at x-ray frequencies $\chi'_0(\omega) \ll 1$. As will be shown below, for nanotubes, in contrast to ordinary crystals, the polarization can indeed considerably affect the spectrum of emitted frequencies because of the possibility of a complex Doppler effect.

To solve the wave equation (21) we represent the periodic part of the wave function $\phi(\mathbf{r})$ as an expansion in reciprocal-lattice vectors. As a result, Eq. (26) becomes

$$\psi(\mathbf{r}) = e^{i\boldsymbol{\kappa}\mathbf{r}}\sum_{\mathbf{G}} \psi_{\mathbf{G}} e^{i\mathbf{G}\mathbf{r}}.$$

Then the wave equation reduces to the following infinite system of algebraic linear equations for $\psi_{\mathbf{G}}$:

$$\left[\frac{1}{2E}(\boldsymbol{\kappa} + \mathbf{G})^2 - \varepsilon\right]\psi_{\mathbf{G}} + \sum_{\mathbf{H}} \psi_{\mathbf{G}-\mathbf{H}} U_{\mathbf{H}} = 0, \qquad (29)$$

where the formula (25) determines the coefficients of the system of equations $U_{\mathbf{H}}$. The condition for the existence of a nontrivial solution of the system (29) requires that the determinant of the matrix of this system vanish; this determines the possible transverse energy bands and the dependence (dispersion) of the transverse energy on the quasimomentum in each band. We note that the transition from the motion of particles along the axis of nanotubes to that along planes of a superlattice formally reduces to taking account of only that reciprocal-superlattice vectors **H** which are orthogonal to the corresponding planes.

Our numerical method for solving the system of equations consisted of step-by-step increasing the number of waves $2\pi/d_p$ taken into account and correspondingly increasing the rank of the system until the transverse-energy eigenvalues obtained no longer varied appreciably. Figures 5 and 6 show the resulting transverse-energy bands in the planar channel (1,0) of a superlattice of (10,10) nanotubes for positrons and electrons, respectively, with various energies (1, 3, and 9 MeV). The transverse quasimomentum in units of the Brillouin zone width δ is plotted along the abscissa. The same figures show the planar potentials for positrons and electrons. The number of allowed transverse-energy bands within a potential well increases with the total energy of the particles, and this number is greater for positrons than for electrons (with the same total energy E). For 1 and 3 MeV there are also relatively wide above-barrier bands, which ultimately pass into the continuous energy spectrum, corresponding to almost free particles. For 9 MeV positrons the number of quasidiscrete levels in a well is so large that, apparently, their motion can be investigated with adequate accuracy by methods of classical mechanics. As far as similar results for the motion of particles along nanotubes are concerned, the number of bands in this case is much larger than in the corresponding planar cases. For example, our calculations show that for 1 MeV positrons there are several

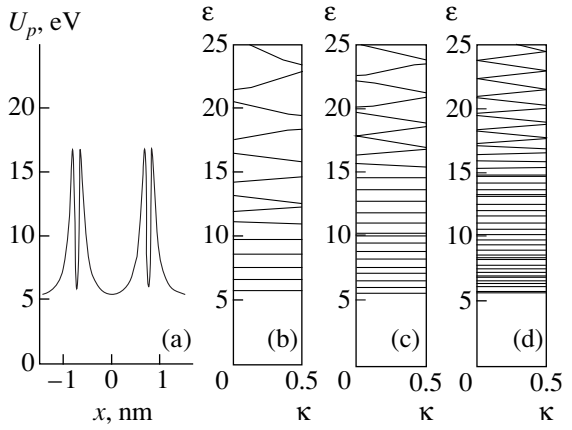

Fig. 5. (a) Potential well and transverse-energy bands of positrons with energy: (b) 1 MeV, (c) 3 MeV, (d) 9 MeV in a planar channel (1,0) of a superlattice of (10,10) nanotubes within the first Brillouin zone.

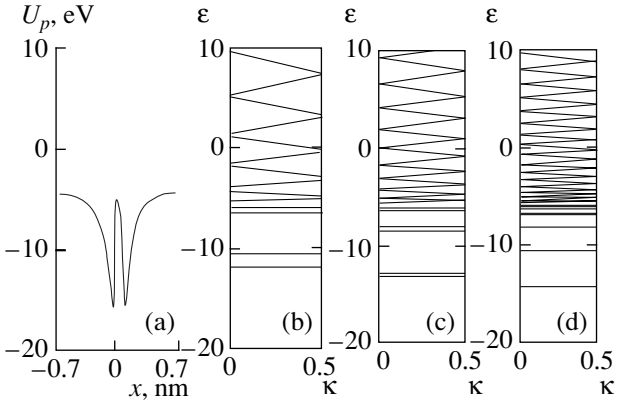

Fig. 6. (a) Potential well and transverse energy band of electrons with energy: (b) 1 MeV, (c) 3 MeV, (d) 9 MeV in a planar channel (1,0) of a superlattice of (10,10) nanotubes in the first Brillouin zone.

tens of bands within a potential well and the number of bands increases approximately linearly with the total energy of the particles.

According to Eq. (27) the frequencies of radiative transitions between bands are determined by the condition that the argument of the Dirac delta function, which contains the permittivity of the medium $\chi_0'(\omega)$, vanishes. Since, as we shall see below, the possible radiation frequencies can lie below the *K* edge of the photoeffect, the permittivity must be calculated taking account of the coupling of some electrons of the medium with the atomic nucleus and can be represented in the form

$$\chi_0'(\omega) = \frac{r_e\lambda^2 n_e}{\pi} f'(\omega),$$

where $\lambda = 2\pi/\omega$ is the wave length, $r_e = e^2/mc^2$ is the classical electron radius, $n_e = 2NZ/Sd_R$ is the average number of electrons per unit volume of the nanotube superlattice, and $f'(\omega)$ is the real part of the atomic scattering factor [21]. Moreover, since the potential well is symmetric and the states of transverse motion are characterized by a definite parity, dipole transitions are possible only for transitions between states with different parity.

Calculations show that 1 MeV electrons and positrons can emit only in the ultraviolet range, where radiation is completely absorbed in the substance itself and it is hardly possible to observe the radiation. Electrons with energy 3 MeV at zero angle of observation $\theta = 0$ can emit several lines in the range 290–311 eV, and as energy increases to 9 MeV radiation at even higher frequencies 1.3–2.2 keV is possible. The detailed spontaneous-emission spectra depend on the population of the energy bands, which is determined by the conditions of entry of the particle beam into the nanotubes.

4. SCATTERING OF X-RAYS BY A NANOTUBE SUPERLATTICE

Let an electromagnetic wave corresponding to x-ray frequencies ω propagate at a quite small angle with respect to the axes of the nanotubes. Then, similarly to the potential, we can introduce a permittivity $\chi(\omega, \mathbf{r})$ averaged along the axis of the nanotubes. The longitudinal component *F* of the electric (magnetic) field of the wave satisfies the equation

$$\Delta_r F + k_\perp^2 = 0. \tag{30}$$

Here $k_\perp^2 = \omega^2[1 + \chi(\omega, \mathbf{r})] - k_z^2$, k_z is the longitudinal component of the wave vector of the wave, and $\chi(\omega, \mathbf{r})$ is the complex permittivity. The transverse components of the electromagnetic field can be expressed in terms of the longitudinal components as follows:

$$\mathbf{E}_\perp = \frac{ik_z}{k_\perp^2}\nabla_r E_z, \quad \mathbf{H}_\perp = \frac{i\omega}{k_\perp^2}\left(-\frac{\partial E_z}{\partial y}, \frac{\partial E_z}{\partial x}\right)$$

for *TM* waves and

$$\mathbf{E}_\perp = \frac{i\omega}{k_\perp^2}\left(\frac{\partial H_z}{\partial y}, -\frac{\partial H}{\partial x}\right), \quad \mathbf{H}_\perp = \frac{ik_z}{k_\perp^2}\nabla_r H_z$$

for *TE* waves. Here ∇_r is a two-dimensional gradient. The dependence of all components of the electromagnetic field of the wave on the longitudinal coordinate *z* is determined by the factor $\exp(-ik_z z)$, where the possible values of the propagation constant k_z are determined by solving Eq. (30).

We shall represent the permittivity $\chi(\omega, \mathbf{r})$ as an expansion in terms of reciprocal-lattice vectors:

$$\chi(\omega, \mathbf{r}) = \sum_{\mathbf{H}} \chi_{\mathbf{H}}(\omega) e^{i\mathbf{H}\mathbf{r}}.$$

Let us assume that the energy ω of the photons is much higher than the binding energy of the *K* electrons in carbon ($\approx$283 eV). Then the real part of the permittivity $\chi'(\omega, \mathbf{r})$ is determined by the interaction of photons with the electrons in the medium, which can be assumed to be free. Then the explicit form of the coefficients $\chi_{\mathbf{H}}'(\omega)$ in the expansion of the real part of the susceptibility can be found as done above for the case of the electric potential. Since the distribution (19) of the electron density in a nanotube differs only by the coefficients from the corresponding potential distribution (2), under the same assumptions as used in the derivation of (25) we obtain

$$\chi_0'(\omega) = -\frac{r_e\lambda^2 n_e}{\pi},$$
$$\chi_{\mathbf{H}}'(\omega) = \chi_0'(\omega) a_j^{(e)} \exp\left(-\frac{H^2}{4b_j^{(e)2}}\right) J_0(RH). \tag{31}$$

Here $\lambda = 2\pi/\omega$ is the wavelength, $r_e = e^2/mc^2$ is the classical electron radius, and $n_e = 2NZ/Sd_R$ is the average number of electrons per unit volume of the nanotube superlattice. The imaginary part $\chi''(\omega)$ of the permittivity is due to absorption of photons (primarily as a result of the photoeffect on *K* electrons) and is related with the imaginary part of the atomic scattering factor $f''(\omega)$ by the relation

$$\chi''(\omega, \mathbf{r}) = \frac{r_e\lambda^2\rho}{\pi} f''(\omega).$$

Here $\rho(\mathbf{r})$ is the local density of atoms, which, since the *K* shell is relatively small, differs from zero only at the sites of the lattice of carbon atoms. Thus, we obtain the

following representation for the expansion coefficients of the imaginary part of the susceptibility:

$$\chi_0''(\omega) = \frac{r_e\lambda^2\rho_0}{\pi} f''(\omega), \quad \chi_{\mathbf{H}}''(\omega) \simeq \chi_0''(\omega) J_0(RH),$$

where $\rho_0 = 2N/Sd_R$ is the number of atoms per unit volume of the nanotube superlattice. The values of $f''(\omega)$ for carbon in the frequency range of interest to us are presented, for example, in [21].

We seek the solutions of Eq. (30) in the form

$$F(\mathbf{r}) = e^{i\boldsymbol{\kappa}\mathbf{r}} \sum_{\mathbf{G}} F_{\mathbf{G}} e^{i\mathbf{G}\mathbf{r}}.$$

As a result, we obtain a system of equations for the coefficients $F_{\mathbf{G}}$ that describes the multiwave diffraction of x-rays in a nanotube superlattice:

$$[-(\boldsymbol{\kappa} + \mathbf{G})^2 + (\omega^2 - k_z^2)]F_{\mathbf{G}} + \omega^2 \sum_{\mathbf{H}} F_{\mathbf{G}-\mathbf{H}} \chi_{\mathbf{H}} = 0. \tag{32}$$

The zeros of the determinant of the system (32) determine the modes of the electromagnetic field propagating freely along nanotubes or along the planes of the superlattice. By the order of magnitude, the number of modes localized inside a nanotube is determined by the squared ratio of the nanotube radius R to the transverse wavelength $\lambda_\perp = 1/k_\perp \approx 1/(\omega|\chi_0'|^{1/2})$. It is easy to see that, in contrast to relativistic electrons and positrons, the characteristic value of $\lambda_\perp$ does not depend on the particle (photon) energy ω and is about 85 Å for a superlattice of (10,10) nanotubes. Thus, localized modes can exist, irrespective of the value of ω, only for sufficiently large nanotube radii [10]. In this case, nanotubes can be treated as waveguides for x-rays similarly to the positron channeling examined above. Indeed, neglecting the vector character of the electromagnetic field, Eq. (32) differs from the similar Eq. (29) for positrons only by the meaning of the coefficients appearing in it.

In nanotube superlattices, in contrast to ordinary crystals, diffraction of relatively soft nanometer-range x-rays is possible. We shall consider the case of symmetric Bragg reflection under the conditions of the two-wave approximation of the dynamical theory of diffraction. Let a wave be incident at an angle close to the Bragg angle with respect to the planes of the superlattice, and, for the sake of simplicity, we shall assume that the entrance surface is parallel to these planes.

The differential coefficient of the Bragg reflection from superlattice planes can be represented in the form [22]

$$R(\Delta\theta) = \left|\eta \pm (\eta^2 - 1)^{1/2}\right|^2.$$

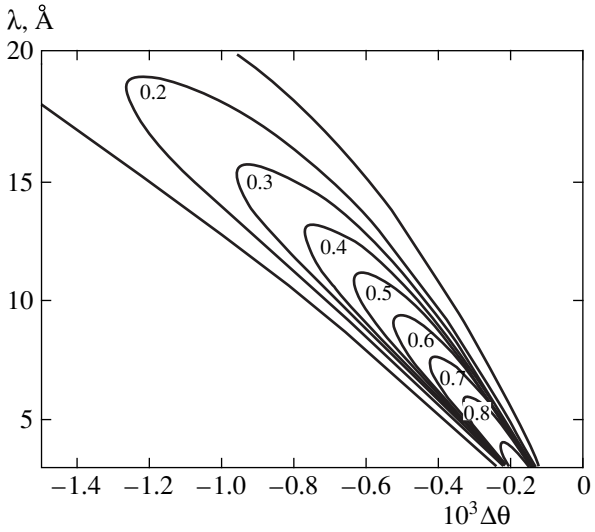

Fig. 7. Levels of constant differential Bragg reflection coefficient as a function of the wavelength of the incident radiation and angular deflection as a function of the kinematic direction for σ polarization.

The sign in front of the parentheses is chosen from the condition $R \le 1$. Here $\eta = \eta' + i\eta''$ is a complex parameter, related to $\chi_{\mathbf{H}}(\omega)$ as

$$\eta' = (-\Delta\theta \sin 2\theta_B + \chi_0')/|C|\chi_{\mathbf{H}}',$$

$$\eta'' = -\frac{\chi_{\mathbf{H}}''}{\chi_{\mathbf{H}}'}\left(\eta' - \frac{1}{|C|}\frac{\chi_0''}{\chi_{\mathbf{H}}''}\right),$$

where $C = 1$ for radiation polarized perpendicular to the plane of incidence (σ polarization), $C = \cos 2\theta_B$ for polarization in the plane of incidence (π polarization), the kinematic Bragg angle is determined by the equality $\sin\theta_B = 2\omega/H$, and $\Delta\theta = \theta - \theta_B$ is the angular deviation from kinetic Bragg direction. Another important characteristic is the integral coefficient

$$R_i = \int_{-\infty}^{\infty} R(\Delta\theta) d\Delta\theta.$$

The dependence of the differential reflection coefficient R on the deflection angle $\Delta\theta$ for various wavelengths λ for σ polarization and symmetric Bragg reflection by (1,0) planes of a superlattice of (10,10) nanotubes is displayed in Fig. 7. The curves correspond to constant values of the reflection coefficient, indicated on the curves. As the wavelength increases, the reflection maximum due to refraction at the boundary undergoes increasingly larger displacements relative to the kinematic direction and is broadened. On the whole, as the wavelength increases, the differential reflection coefficient $R(\Delta\theta)$ decreases as a result of an increase in the absorption in the medium, but it still remains substantial ($R \simeq 0.3$) even for very soft x-rays

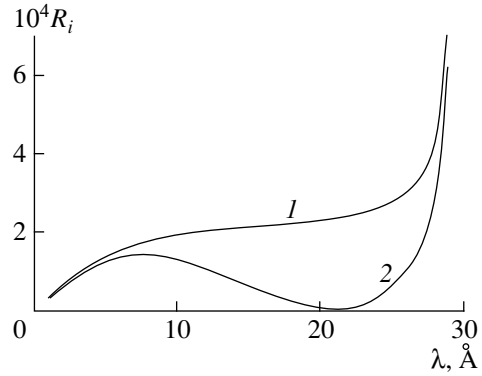

Fig. 8. Integral Bragg reflection coefficient as a function of wavelength for (*1*) σ polarization and (*2*) π polarization.

with wavelength 15 Å. The integral reflection coefficient is displayed in Fig. 8. It is interesting to note that near 20 Å the reflection of a π-polarized wave is substantially suppressed with respect to σ polarization, i.e., a nanotube superlattice acts as a polarizer in this frequency range.

ACKNOWLEDGMENTS

This work was supported by INTAS (project no. 97-30392).

REFERENCES

1. M. S. Dresselhaus, G. Dresselhaus, and P. C. Eklund, *Science of Fullerenes and Carbon Nanotubes* (Academic, San Diego, 1996).
2. M. M. J. Treacy, T. W. Ebbesen, and J. M. Gibson, Nature **381**, 678 (1996).
3. J. E. Fischer, H. Dai, A. Thess, *et al.*, Phys. Rev. B **55**, R4921 (1997).
4. S. Iijima and T. Ichihashi, Nature **361**, 603 (1993).
5. S. Iijima, Nature **354**, 56 (1991).
6. A. Thess, R. Lee, P. Nikolaev, *et al.*, Science **273**, 483 (1996).
7. S. J. Tans, M. H. Devoret, H. Dai, *et al.*, Nature **386**, 474 (1997).
8. C. Joachim and J. K. Gimzewski, Chem. Phys. Lett. **265**, 353 (1997).
9. L. G. Gevorgyan, K. A. Ispiryan, and R. K. Ispiryan, Pis'ma Zh. Éksp. Teor. Fiz. **66**, 304 (1997) [JETP Lett. **66**, 322 (1997)].
10. N. K. Zhevago and V. I. Glebov, Phys. Lett. A **250**, 360 (1998).
11. D. S. Gemmell, Rev. Mod. Phys. **129**, 46 (1974).
12. P. A. Doyle and P. S. Turner, Acta Crystallogr. A **24**, 390 (1968).
13. *Heigher Transcendental Functions (Bateman Manuscript Project)*, Ed. by A. Erdelyi (McGraw-Hill, New York, 1953; Nauka, Moscow, 1974), Vol. 2.
14. A. Magnus, Ann. Phys. **70**, 303 (1923).
15. Z. Su and P. Coppens, Acta Crystallogr. A **53**, 749 (1997).
16. N. K. Zhevago, Zh. Éksp. Teor. Fiz. **75**, 1389 (1978) [Sov. Phys. JETP **48**, 701 (1978)].
17. N. K. Zhevago, Zh. Éksp. Teor. Fiz. **94** (10), 23 (1988) [Sov. Phys. JETP **67**, 1967 (1988)].
18. A. I. Akhiezer and N. F. Shul'ga, Usp. Fiz. Nauk **137**, 561 (1982) [Sov. Phys. Usp. **25**, 541 (1982)].
19. V. A. Bazylev and N. K. Zhevago, *Radiation of Fast Particles in Substance and External Fields* (Nauka, Moscow, 1987).
20. V. A. Bazylev and N. K. Zhevago, Zh. Éksp. Teor. Fiz. **73**, 1697 (1977) [Sov. Phys. JETP **46**, 891 (1977)]; Phys. Lett. B **84**, 182 (1979).
21. B. L. Henke, At. Data **27**, 3 (1982).
22. B. W. Batterman and H. Cole, Rev. Mod. Phys. **36**, 681 (1964).

Translation was provided by AIP

Journal of Experimental and Theoretical Physics, Vol. 91, No. 3, 2000, pp. 515–523.
Translated from Zhurnal Éksperimental'noĭ i Teoreticheskoĭ Fiziki, Vol. 118, No. 3, 2000, pp. 592–601.

SOLIDS
Structure

Dynamics of Twist Point Defects with Stretching in a Polymer Crystal

E. A. Zubova[a, *], N. K. Balabaev[b], L. I. Manevich[a], and A. A. Tsygurov[a]

[a]*Semenov Institute of Chemical Physics, Russian Academy of Sciences, Moscow, 117977 Russia*
[b]*Institute of Mathematical Problems of Biology, Russian Academy of Sciences, Pushchino, Moscow oblast, 142292 Russia*
**e-mail: zubova@center.chph.ras.ru*

Received January 27, 2000

Abstract—A molecular-dynamics simulation of the behavior of a twist point defect with stretching in a chain of an equilibrium polymer crystal ("united" atoms approximation for polyethylene) is performed for immobile and mobile neighboring chains. It is shown that such a defect in a cold polymer crystal possesses soliton-type mobility. The upper limit of the spectrum of soliton velocities is found, and it is the same for both cases. The maximum possible velocity of defects is three times lower than the theoretical limit of the spectrum (which is equal to the velocity of "torsional" sound in an isolated chain). An explanation of the reason for this discrepancy is proposed: because of the interaction of two "degrees of freedom" of the defect (twisting and stretching) the energy of a nonlinear wave is dissipated in the linear modes of the system, which results in effective friction whose magnitude depends strongly on the velocity of the defect. The "boundary of the spectrum of soliton velocities" determines the transition between regimes of strong and weak braking of defects. © *2000 MAIK "Nauka/Interperiodica".*

1. INTRODUCTION

To predict the relaxation, plastic, and strength properties as well as the melting temperature of crystals it is necessary to study localized mobile defects—deviations from an ideal structure. Point structural defects (vacancies or inclusions) are responsible for the relaxation properties of solids. The impossibility of directed motion of such defects is the reason why relaxation processes are slow in low-molecular crystals. On account of the strong anisotropy of the properties of polymer crystals (the atoms in a polymer chain are bound with one another by chemical covalent bonds, and the intermolecular interaction is due to weak vander-Waals forces) vacancies with breaking of intrachain covalent bonds are virtually immobile. However, such crystals can contain different, specifically polymer, point defects, which are due to not the breaking of intrachain bonds but rather the deformation of a chain localized on a small section of the chain.

The concepts of such defects appeared in the physics of polymer crystals after the discovery of the anomalously rapid dielectric relaxation of oxidized polyethylene (see the review in [1]). Analysis of a number of possible molecular mechanisms of this process (see the review in [2]) made it possible to identify as the most likely mechanism the propagation of regions of twisting (by 180°) with stretching (by a half-period of the chain) that have a length of several tens of periods, in the absence of conformational changes, along the chains. It turned out [3, 4] that such defects are also created in the process of premelting of a crystal (their energy is much lower than the energy of purely stretching defects—vacancies of one unit of the chain without twisting).

The quasi-one-dimensional approximation with immobile neighboring chains (see, for example, [5] and the literature cited there) makes it possible to describe a point defect as a soliton-type topological excitation [6]—a localized nonlinear wave propagating with constant subsonic velocity along the chain, changing the state of the chain after it passes and therefore capable of causing rapid relaxation in the crystal.

An approximate analytical description of static point defects in polyethylene was proposed in [7], and a numerical investigation by molecular-mechanics methods was conducted in [8]. However, these works did not consider the question of the mobility of defects and correspondingly the role of defects in the relaxation of a crystal.

The problem of point defects in a polyethylene chain surrounded by immobile neighboring chains has recently been solved by a numerical-variational method [9, 4] (in the "united" atoms model, in which CH_2 groups are replaced by point particles). It was found that in this case point defects possess soliton-type mobility, and the upper limits of the velocity spectrum for different types of defects were calculated.

But, in a more previous work [10] a molecular-dynamics investigation of the behavior of torsional point defects with stretching in a polyethylene crystal was performed in the same "united-atoms" model but

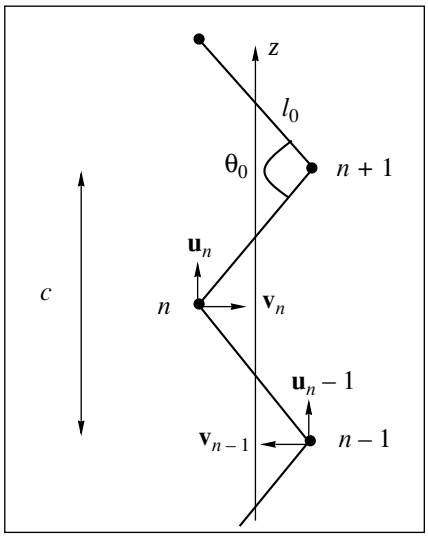

Fig. 1. Model of a polymer crystal (polyethylene with united atoms): parameters of the chain and the local coordinates of the atoms.

with mobile molecules of the first coordination sphere, and it was concluded from the results of the numerical experiment that the mobility of such defects is of nonsoliton type: they stopped rapidly.

However, the arrangement of this experiment was such that defects could stop rapidly for several reasons. In the first place, it was discovered in [11] that purely stretching defects in an equilibrium crystal configuration with all mobile chains in the crystal retain the initial velocities right up to values of the order of 0.6 times the velocity of sound, and for high velocities they slow down to this velocity but do not stop. In nonequilibrium or unstable structures (including the orthorhombic structure, which is nonequilibrium in the "united atoms" model of polyethylene used in [10]), requiring that the chains rotate in order for relaxation to occur, the vacancies stop rapidly. The clearly soliton character of the dynamics of pure vacancies led us to conjecture that the stopping of a defect in [10] was due precisely to the nonequilibrium nature of the initial crystal.

But, in addition, we saw in [11] that friction against the phonon modes of mobile neighboring chains substantially decreased the upper limit of the velocity spectrum even for pure vacancies (from the velocity of sound c (theoretical limit) down to ≈0.6 times this value). The interaction of a twist point defect with stretching with the same phonon modes can be much stronger and can even make the soliton mechanism of defect mobility impossible.

The purpose of the present work is to study the dynamics of twist point structural defects with stretching in a polymer crystal with all mobile chains on the basis of the same molecular-dynamics model as the one used in [11].

Our numerical model of a crystal for studying the dynamics of point defects is described in Section 2. An approximate analytical description of the defects is presented in Section 3 for immobile neighboring chains. The results of a molecular-dynamics simulation of the dynamics of defects with immobile and mobile neighboring chains are presented in Section 4. Finally, Section 5 is devoted to a discussion of these results and the conclusions.

2. NUMERICAL MODEL OF A POLYMER CRYSTAL (POLYETHYLENE WITH "UNITED" ATOMS); EQUILIBRIUM CRYSTAL CONFIGURATION

We adopted the following model of a polymer crystal [12] (polyethylene with united atoms; see Fig. 1): the chains are a planar trans-zigzags; the bonds between the atoms (point particles with mass m) are absolutely rigid and their length is l_0; the deformation energies of the valence (θ_n) and conformational (τ_n) angles are

$$U_3(\theta_n) = \frac{1}{2}K_\theta(\theta_n - \theta_0)^2, \quad (1)$$

$$U_4(\tau_n) = \alpha + \beta\cos\tau_n + \gamma\cos 3\tau_n, \quad (2)$$

the atoms separated by more than 2 neighbors or belonging to different chains interact according to the law

$$U(r) = \begin{cases} U_{LJ}(r) - U_{LJ}(R), & r \le R \\ 0, & r > R, \end{cases}$$

where $U_{LJ}(r) = 4\epsilon[(\sigma/r)^{12} - (\sigma/r)^6]$ is the Lennard–Jones potential with a minimum at the point $r_0 = 2^{1/6}\sigma$. The numerical values used for the constants are given in Table 1.

Table 1. Parameters of the model crystal

Parameter	Quantity	References	Parameter	Quantity	References
m	14 amu	–	β	1.675 kJ/mole	[13]
l_0	1.53 Å	[13]	γ	6.695 kJ/mole	[13]
θ_0	113°	[13]	ϵ	0.4937 kJ/mole	[14]
K_θ	331.37 kJ/mole	[13]	σ	3.8 Å	[14]
α	8.370 kJ/mole	[13]	R	$2r_0$	–

For the model of the crystal we adopted periodic boundary conditions in all three directions. A rectangular parallelepiped was chosen for the working cell. The leap-frog algorithm [15], taking account of the restrictions imposed by the rigid bonds [16], was used to solve the corresponding classical Lagrangian equations of the first kind numerically. Periodic boundary conditions along the axis of the molecules make it possible to follow the dynamics of a defect for an unbounded time and those in the transverse section of the crystal, to avoid introducing in the transverse section unphysical boundary conditions with a rigidly fixed second coordination sphere. To prevent a soliton from affecting itself the number of molecules in the working cell was chosen so that the image of each molecule was located no closer than in its fourth coordination sphere and the length of the molecule—for a defect of the order of 35 chain periods c long—was assumed to be $200c$ (one period contains two CH_2 groups).

Since the length of the projection of a molecule on its transverse section is $l_\perp = 0.843$ Å and the van-der-Waals radii of the united atoms r_0 are 4.265 Å $\approx 5l_\perp$, the packing of the zigzag planes in the crystal will be close to that of cylinders. Theoretically, two different mechanically equilibrium configurations are possible (see Fig. 2). Both possess a monoclinic cell and close energies. However, the second one is unstable and separates into two domains, each of which corresponds to the first configuration [11], as a result of relaxation. The parameters of a stable equilibrium structure a and b depend on the cutoff radius R. Table 2 presents data on the relaxation of samples for different values of R. The period along the axis of a molecule is always $c = 2.554$ Å. The density of the sample is $\rho = 1.155$ g/cm^3.

Molecular-dynamics simulation of a polyethylene crystal in the united-atoms model has shown [11] that the potential energy does not have a local minimum for an orthorhombic structure for any value of the cell parameters. The numerical experiment reveals the appearance of this minimum only in a model of polyethylene in which the CH_2 group is represented by three spatially separated force centers. Thus, the possibility of the existence of an orthorhombic structure in polyethylene is due to the presence of side groups and not the form of the backbone of the chains.

3. APPROXIMATE ANALYTICAL DESCRIPTION OF A TWIST POINT DEFECT WITH STRETCHING IN A CHAIN OF A POLYMER CRYSTAL

We shall consider the simplest description of the dynamics of a twist defect with stretching in the chain of a polymer crystal in the immobile-neighbors approximation using the continuum model.

The effective substrate potential generated for the nth atom by immobile neighbors in a stable equilibrium monoclinic lattice of a polymer crystal in the "united" atoms model should have two wells with different minima at the points (ϕ_n, ψ_n): (0, 0) and (π, π). Here ϕ_n is the angle in a cylindrical coordinate system whose axis is directed along the axis of a molecule, $\psi_n = (2\pi/c)u_n$ (u_n is the longitudinal displacement from the position of equilibrium).

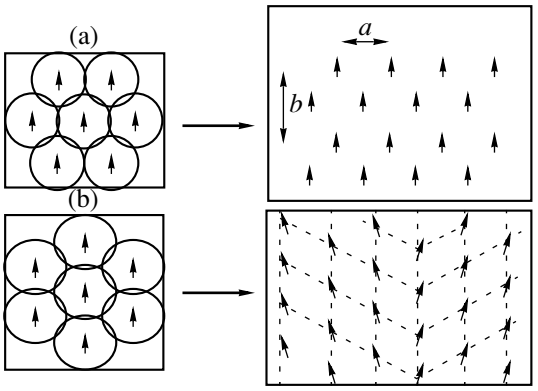

Fig. 2. Possible equilibrium configurations for planar zigzag packing: (a) stable, (b) unstable (the period of the transverse section of the molecules is shown; the arrows show the direction from the nearest atom of the molecule under the plane to the nearest above the plane).

Numerically, the substrate potentials obtained by calculating the energy of the crystal with all molecules secured except one, which can move and rotate (as a whole) along the axis. In the stable equilibrium configuration of the crystal the function

$$V(\phi_n, \psi_n) = A(1 - \cos\phi_n \cos\psi_n) + B(1 - \cos 2\phi_n) \quad (3)$$

with $A \approx 0.274$ kJ/mole and $B \approx 0.865$ kJ/mole approximates the substrate potential with an error of less than 10% of its maximum value. The level lines of the function (3) are shown in Fig. 3.

Adding a term $C(1 - \cos 4\phi_n)$ with $C = -0.067$ kJ/mole to the potential (3) decreases the error to 2%. We shall discuss in Section 5, after presenting the results of the molecular-dynamics simulation, how this addition can influence the dynamics of a defect.

In our analysis we neglect the dependence of the potential V on the third coordinate—transverse displacements v_n, since for long-wavelength waves (with characteristic scale much greater than the distance between the atoms of a chain) the transverse displacements of the atoms are much smaller than the longitudinal displacements (see Appendix A). The consequences (except for a small change in the form of the analytical solution) of taking account of the depen-

Table 2. The parameters a and b (in Å) of the equilibrium crystalline configuration for different cutoff radii R

Parameter	$R = 1.8r_0$	$R = 2r_0$
a	3.998	3.980
b	7.994	7.966

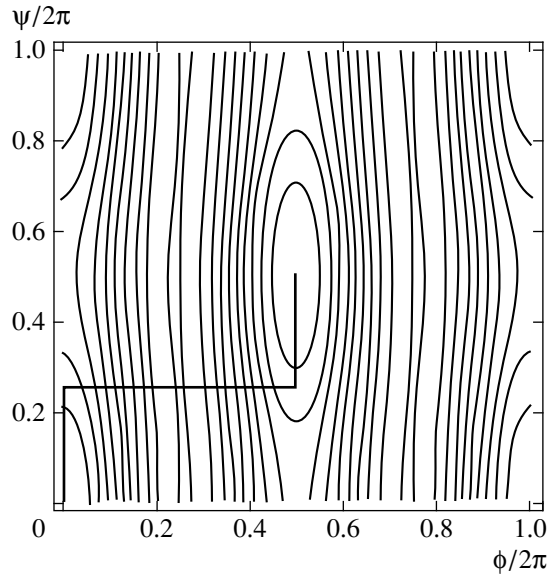

Fig. 3. Level lines of the potential surface (3) for a substrate, generated by immobile neighboring chains, in the (ϕ, ψ) plane. The broken line shows the "orbit" of the zeroth analytical approximation ($L_\phi/L_\psi \longrightarrow 0$) in the solution of the system of equations (8).

dence of the potential on this coordinate will be discussed in Section 5. Here we shall confine our attention to the main interaction of the two basic degrees of freedom of an atom.

The rigid-bond condition determines the relation between the longitudinal and transverse displacements u and v of the atoms (see Fig. 1). Switching to the continuum approximation (which is justified because of the relative weakness of the intermolecular interaction) and neglecting dispersion and nonlinearity, which are due to intramolecular interactions (see [17] for a more accurate continuum description), we obtain the Lagrangian system

$$L = \int \frac{dx}{l_0 s_0}\left[I_\phi \frac{\phi_t^2}{2} + I_\psi \frac{\psi_t^2}{2} - K_\phi \frac{\phi_x^2}{2} - K_\psi \frac{\psi_x^2}{2} - V(\phi, \psi)\right], \quad (4)$$

where the constants are related with the parameters of the numerical model of the crystal by the relations (see Appendix B)

$$I_\phi = m\left(\frac{l_0 c_0}{2}\right)^2, \quad I_\psi = m\left(\frac{l_0 s_0}{\pi}\right)^2, \quad (5)$$

$$(s_0 = \sin(\theta_0/2), \quad c_0 = \cos(\theta_0/2));$$

$$K_\phi = l_0^2(\beta + 9\gamma), \quad (6)$$

$$K_\psi = \left(\frac{2 l_0 s_0^2}{\pi c_0}\right)^2 K_\theta. \quad (7)$$

Here I_ϕ, I_ψ and K_ϕ, K_ψ are, respectively, the inertial and stiffness parameters of the chain—"torsional" (with respect to ϕ) and "longitudinal" (with respect to ψ). For our numerical values of the crystal parameters, $I_\phi/I_\psi \approx 1.08$ and $K_\phi/K_\psi \approx 0.29$.

Thus, we have obtained a very simple model for two coupled fields with different "stiffness characteristics" and "inertial properties," and the relation between them is due only to the anharmonicity of the external substrate potential.

The Lagrangian (4) with the potential function (3) corresponds to coupled equations describing the evolution of the fields ϕ and ψ:

$$\begin{aligned} I_\phi \phi_{tt} - K_\phi \phi_{xx} + A\sin\phi\cos\psi + 2B\sin 2\phi &= 0, \\ I_\psi \psi_{tt} - K_\psi \psi_{xx} + A\sin\psi\cos\phi &= 0. \end{aligned} \quad (8)$$

It is easy to see that the variables ϕ and ψ in the linearized equations of the system separate and give two branches of the dispersion curve with two sound velocities: torsional sound $v_\phi = \sqrt{K_\phi/I_\phi} \approx 7.63$ km/s and longitudinal sound $v_\psi = \sqrt{K_\psi/I_\psi} \approx 14.70$ km/s. In the continuum model of a chain without a substrate these are the maximum propagation velocities of small-amplitude disturbances along ϕ and ψ, respectively.

Let us assume that the system of equations (8) possesses a solution in the form of a topological solitary nonlinear wave moving with velocity $v \geq 0$ in a direction of positive values of x, i.e., in the form of "coupled" kinks

$$\phi = \phi(x - vt), \quad \psi = \psi(x - vt),$$

where the functions ϕ and ψ are such that as time t varies from $-\infty$ to $+\infty$ they vary from π to 0. Then this wave corresponds to a kink of twisting by 180° and stretching by a half-period of the chain. When such a wave passes along the chain all particles will move in turn into the position of the nearest neighbor in the direction of negative values of x.

If a solution of this form exists, then the problem consists simply of finding the correct "orbit" in the ϕψ plane. The exact solution of such a problem is known for identical stiffness and inertial parameters of the fields for a number of simple polynomial potentials $V(\phi, \psi)$ (see [18] and the references cited there, as well as [19]). In our case successive approximations to the solution can be found assuming that the kink width L_ϕ with respect to ϕ is much smaller than the kink width L_ψ with respect to ψ. Indeed, the "torsional" stiffness K_ϕ of the chain is approximately 3.4 times smaller than the longitudinal stiffness K_ψ, and the substrate is stiffer for ϕ than for ψ because of the presence of the large term $B(1 - \cos 2\phi_n)$ in the potential. In the limit $L_\phi/L_\psi \longrightarrow 0$ the orbit will approach the broken line shown in Fig. 3. For a finite value of the parameter the first approximation to the solution can be obtained by assuming that in the equation for ϕ the kink with respect to ψ is "infinitely wide" and setting $\psi \longrightarrow \pi/2$. The equation for ψ can be solved in two regions: for $x - vt < 0$, replacing ϕ by 0, and for $x - vt \geq 0$, making the

substitution $\phi \longrightarrow \pi$ and matching the solutions at the point $x - vt = 0$ (this method of solving a similar system of equations was proposed in [5]). In our case all three equations reduce to the sine-Gordon equations. After matching the solutions we obtain

$$\phi_1 = 2\arctan\exp\left(\frac{x - vt}{L_\phi\sqrt{1 - (v/v_\phi)^2}}\right), \quad (9)$$

$$L_\phi = \sqrt{\frac{K_\phi}{4B}} \approx 5.08\frac{c}{2};$$

$$\psi_1 = \begin{cases} 4\arctan\exp\left(\dfrac{x - vt}{L_\psi\sqrt{1 - (v/v_\psi)^2}} - \ln\tan\left(\dfrac{3\pi}{8}\right)\right), \\ x - vt < 0 \\ 4\arctan\exp\left(\dfrac{x - vt}{L_\psi\sqrt{1 - (v/v_\psi)^2}} + \ln\tan\left(\dfrac{3\pi}{8}\right)\right) - \pi, \\ x - vt \geq 0, \end{cases} \quad (10)$$

$$L_\psi = \sqrt{\frac{K_\psi}{A}} \approx 33.5\frac{c}{2},$$

and the kink velocity v must be less than the lowest of the velocities v_ϕ and v_ψ. We can see that the parameter $L_\phi/L_\psi \approx 0.152$ is indeed small.

When the term $C(1 - \cos 4\phi_n)$ is included in the potential the form of the kink with respect to ϕ will be the solution not of a simple but rather a double sine-Gordon equation.

In what follows we shall seek the solution in the form $\phi = \phi_1 + \phi_2 + \ldots$ $\psi = \psi_1 + \psi_2 + \ldots$, refining the form of the kinks. However, it is important to note that this procedure, even though it leads to the dependence of the form of the kink with respect to ϕ on the parameters L_ψ and v_ψ and the form of the kink with respect to ψ on the parameters L_ϕ and v_ϕ, it cannot change the maximum velocity of a kink. In the opposite approximation (the case of identical stiffness and inertial parameters of the field [19]) the exact solution also admits arbitrary subsonic velocities of coupled waves.

Thus, if the interaction $V(\phi, \psi)$ is such that there exists an "orbital" solution in the form of a coupled solitary wave $\phi - \psi$, then for the potential (3) with a small value of the parameter L_ϕ/L_ψ the form of this wave is given approximately by Eqs. (9) and (10) and the velocity $v < v_\phi < v_\psi$.

In summary, a twist defect with stretching (no breaking of covalent bonds (Fig. 4)) can move along the chain with subsonic velocity $v < v_\phi < v_\psi$, maintaining localization and not disrupting the crystal structure outside the region of the defect. This means that for velocities that are not too close to the sound velocity the dynamics of defects should be of a soliton character (as $v \longrightarrow v_\phi$ the discreteness and the intramolecular nonlinearity can no longer be neglected, since the solitons are too "narrow").

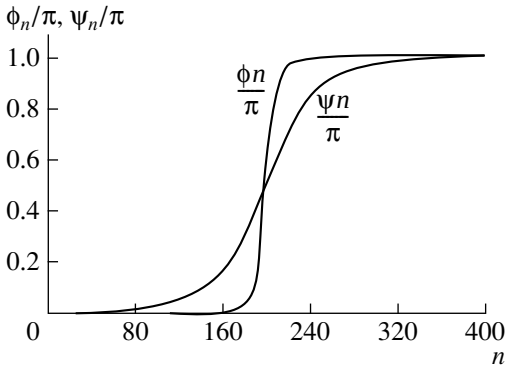

Fig. 4. Static twist point defect with stretching in a chain in a polymer crystal: the coordinates of the atoms in a chain with a defect in a cylindrical coordinate system with the axis directed along the axis of the molecule (see Fig. 1): the steeper curve is for the angle ϕ_n/π (ϕ in radians), the more gently sloping curve is for $\psi_n/\pi = (2/c)u_n$ (u_n is the longitudinal displacement from the position of equilibrium); the form of the defect after relaxation of the crystal in 15 ps is shown, and the difference of the curves shown from the analytical solution (9) and (10) does not exceed 0.005.

4. RESULTS OF MOLECULAR-DYNAMICS SIMULATION OF THE BEHAVIOR OF TWIST DEFECTS WITH STRETCHING IN A POLYMER CRYSTAL

The following time scales are characteristic for the system under study: the travel time of longitudinal sound over one period of the chain is approximately 1.7×10^{-2} ps, the width of a defect along ψ is about 0.58 ps, the travel time of "torsional" sound over the width of a defect along the ϕ coordinate is approximately 0.17 ps.

In the molecular-dynamics experiment, for one of the molecules of a crystal, which has relaxed and cooled down to 0.01 K, we set the atomic displacements and velocities according to the approximate analytic formulas (9) and (10) and observed the evolution of a defect for a prolonged time (of the order of hundreds of picoseconds).

Our analytical approximation (9) and (10) seems to be very rough: kinks "are not coupled." But it turns out that for a static kink the approximation fits the experimentally observed curves (Fig. 4) with an error of less than 0.5% of the magnitude of the jump.

In the numerical experiment we followed the position x_{cm} and velocity v_{cm} of the center of mass of a chain with a defect, which were easily converted to the displacement and velocity of a defect: $x_{def} = -N_1[x_{cm} - x_{cm}(t = 0)]$ and $v_{def} = -N_1 v_{cm}$ ($N_1 = 399$ is the number of atoms

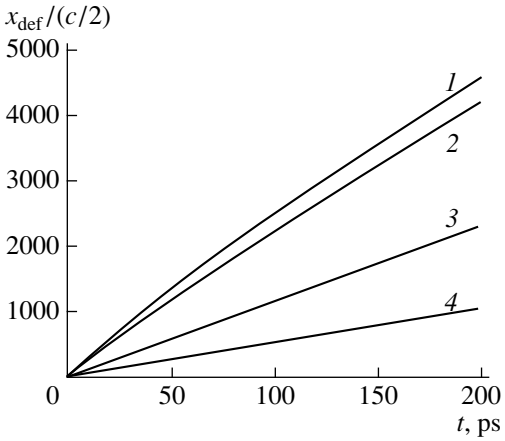

Fig. 5. Dynamics of defects in a chain on a substrate generated by stationary neighboring chains: displacement of defects (in half-periods of the chain) with velocities (*1*) 0.9, (*2*) 0.45, (*3*) 0.2, and (*4*) 0.1 of the velocity of the lower (torsional) sound as a function of time.

in a chain with a defect). With this conversion, because of the thermal vibrations of the atoms (for sample temperature of the order of several Kelvins) high-frequency oscillations are superposed on the true value of the velocity of a defect. These oscillations have no relation to the velocity, and we filtered them out to make the picture clearer.

4.1. Dynamics of Twist Defects with Stretching in a Chain on a Substrate, Generated by Stationary Neighboring Chains

We shall now describe the results of a simulation of the evolution of twist defects with stretching with initial velocities of approximately 0.9, 0.45, 0.2, and 0.1 times the velocity of the lower (torsional) sound (6.875, 3.437, 1.514, and 0.757 km/s) in a crystal in stable equilibrium (Fig. 2a) with clamped neighboring molecules.

As expected, soliton-type mobility of defects is clearly observed in the numerical experiment (see Fig. 5): they move along the chain with constant average velocity, traversing in 200 ps 4378, 4003, 2194, and 937 CH_2 groups (half-periods of the chain), respectively. By analogy with the dynamics of pure vacancies, one would expect that the defects would retain their initial velocities. However (Figs. 6 and 7), only two of the slowest defects retain their velocities. The velocity of fast defects decreases in the first 7 ps to 0.43 (from 0.9) and 0.39 (from 0.45) from the velocity of sound, and in the next 193 ps it slowly decreases to 0.34 and 0.33, respectively.

This picture strongly resembles the evolution of pure vacancies in the presence of mobile neighboring chains [11], when the energy of high-velocity vacancies was transferred to neighboring chains because of the stronger interaction, and the intensity of the interaction with phonon modes of the surrounding molecules decreased sharply with decreasing velocity of the vacancy, so that the (first) limit of the velocity spectrum of purely stretching solitons (approximately 0.6 times the upper (longitudinal) sound) in the presence of mobile neighboring chains arose.

For the propagation of a twist defect with stretching, the energy of a nonlinear wave in the coordinate ψ is transferred into phonon modes ϕ (or, possibly, other modes; see the discussion in Section 5) because of the interaction of the fields ϕ and ψ. Conversely, the energy of the nonlinear wave in the coordinate ϕ is transferred

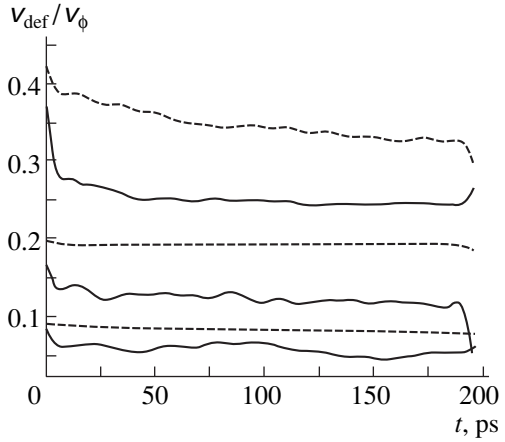

Fig. 6. Dynamics of defects in a polymer crystal: variation of the average velocity of defects with time with immobile (dashed lines) and mobile (solid lines) neighboring chains; the initial velocities of the defects are 0.45, 0.2, and 0.1 times the velocity of lower (torsional) sound (the sharp cutoffs at the end of the plots are not a physical effect; they are due to the averaging procedure).

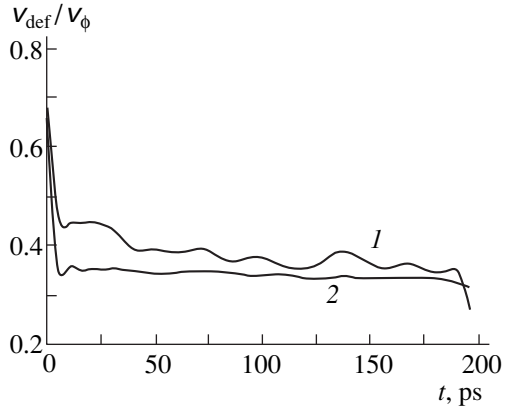

Fig. 7. Dynamics of a defect with initial velocity 0.9 times the velocity of lower (torsional) sound in a polymer crystal: variation of the average velocity with time with (*1*) immobile and (*2*) mobile neighboring chains (the sharp cutoffs at the end of the plots are not a physical effect; they are due to the averaging procedure).

into phonon (or other) modes ψ. The intensity of this interaction, just as the interaction of a purely stretching soliton with the phonon modes of neighboring mobile chains, decreases sharply with decreasing velocity of the wave, and the (second) limit of the velocity spectrum of twist solitons with stretching (~0.34 times the lower (torsional) sound) arises even with immobile neighboring chains. For initial defect velocities much higher than this limit the defect slows down (with large oscillations of even the average velocity; see Fig. 7, curve *1*). These oscillations are still noticeable with the initial velocity of the defect 0.45 times the "torsional" sound (above the limit where the velocity decreases; see Fig. 6, upper curve), but they are no longer noticeable for slow defects, whose initial velocities are remain unchanged.

4.2. Dynamics of Torsional Defects with Stretching in a Chain Surrounded by Mobile Neighboring Chains

The simulation results for the evolution of torsional defects with stretching with initial velocities of the order of 0.9, 0.45, 0.2, and 0.1 times the velocity of lower (torsional) sound in a crystal in stable equilibrium (Fig. 2a) with all mobile molecules are very close to the results obtained for a sample in which the neighboring chains are clamped.

Soliton-type mobility of defects is observed in the numerical experiment: they move steadily along the chain; the corresponding plot is very similar to Fig. 5, the only difference being that the tangents of the slope angles of the curves are smaller (see Figs. 6 and 7 for the velocities) and the average velocities fluctuate somewhat and decrease slightly even for the two slowest defects.

Deformations (along the axis of the molecule and along the angle) accompanying a defect during motion along a chain ("shadows") arise on the mobile neighboring chains near a defect. The form of the "shadows" can be determined analytically in perturbation theory for a purely stretching defect [20]. In the numerical experiment we took as the initial condition the atomic displacements and velocities using the analytical formulas (9) and (10) only on a chain with a defect. It is probably because the initial conditions are not completely "correct" that the velocities of the defects in the steady state with mobile neighbors turned out to be less than for immobile neighbors.

The oscillations of the magnitude and the hardly noticeable additional decrease of the velocities (Fig. 6) are the only direct consequences of the mobility of the neighboring chains in the crystal. It is obvious why this effect is so weak: the velocities of the defects are low, and the interaction of the defects with the phonon modes of neighboring chains become substantial, as we saw in [11], only for velocities higher than the first limit $\sim 0.6 v_\psi \approx 1.2 v_\phi$. It is interesting that near this first boundary (with initial defect velocity $0.9 v_\phi$) friction on the phonon modes of neighboring mobile chains is still so large that the velocity of a defect decreases much more rapidly than for stationary neighboring chains, and it decreases immediately down to the value of the lower second boundary, after which it stops decreasing and no longer fluctuates (Fig. 7, curve *2*).

5. DISCUSSION

In the present paper it was found by molecular-dynamics simulation that the dynamics of twist point defects with stretching in a (cold) polymer crystal is of a soliton character: defects retain their initial velocities if they lie below a certain value (~0.34 times the velocity of the lower (torsional) sound—the theoretical limit of the spectrum). The motion of a twist defect with stretching in a chain in a polymer crystal with velocities below torsional sound is essentially independent of whether the surrounding chains are mobile or immobile.

In [21] the upper limit of the velocity spectrum of twist solitons with stretching remained unnoticed in the molecular-dynamics study of defects in a close model of polyethylene with united atoms because the evolution of the solitons was followed for only a short time (in the time available the solitons could traverse only several tens of CH_2 groups). In [4, 9] it was found for other values of the parameters of the model crystal (with immobile neighboring chains) by a numerical-variational method (it was found that a numerical extremum of the Lagrangian of the system exists for solutions in the form of a solitary wave only for velocities below a certain limit), but the reason why this limit appeared was not discussed.

We believe that the reason is as follows: because the fields ϕ and ψ interact with one another energy is transferred from the nonlinear wave in the coordinate ψ into phonon modes ϕ (or modes localized near a kink) and, conversely, energy is transferred from the nonlinear wave in the coordinate ϕ into phonon modes ψ (or modes localized near a kink).

We note that, of course, for the sine-Gordon equation no localized linear modes other than a translational mode exist. However, in our analytical approximation we actually "separated" the kinks, and only because of this our equations all reduce to the sine-Gordon equations. In addition, in our theoretical analysis we confined our attention to the simplest potential for the interaction of the fields (3), which, however, approximates a real substrate to within about 10%. To increase the accuracy to 2%, a term $C(1 - \cos 4\phi_n)$ with $C = -0.067$ kJ/mole must be added; this leads to a double sine-Gordon equation in the variable ϕ. Finally, multiple harmonics can effectively be added to the main term in (3) when the transverse displacements of atoms in the zigzag plane are taken into account.

For the perturbed sine-Gordon equation, however, it is known that, just as in the nonintegrable φ^4 model, the

energy of the directed motion of a defect can be transferred not only to phonon but also to localized modes, which can be excited "around" a nonlinear wave (for certain types of perturbations see [22]; for the double sine-Gordon equation the form of these modes was obtained in [23]).

Thus, because the energy of a nonlinear wave is dissipated into linear modes of the system an effective friction arises, and the magnitude of this friction decreases sharply with decreasing velocity of the wave, so that a limit of the velocity spectrum of a twist soliton with stretching appears even with stationary neighboring chains. The friction against the phonon modes of neighboring mobile chains [11] is much smaller than the interaction of the torsional and longitudinal degrees of freedom of a defect, and consequently it is this interaction that limits the velocity of a defect in a cold polymer crystal.

Of course, the "orbital" analytical approximation in Section 3 cannot describe such a transfer of energy from a nonlinear wave into linear modes of the system. Further study of the evolution of the interacting fields is required (analytical "nonorbital" and numerical).

ACKNOWLEDGMENTS

This work was supported by the Russian Foundation for Basic Research (project no. 98-03-33366) and the Commission on Work with Young People of the Russian Academy of Sciences.

APPENDIX A

Possibility of Ruling Out Transverse Displacements of Atoms of a Flat Trans-Zigzag for Long Wavelengths in a Model with Absolutely Rigid Bonds between Atoms

The condition for the interatomic bonds to be rigid gives in the linear approximation

$$c_0(v_{n+1} + v_n) \approx s_0(u_{n+1} - u_n), \qquad \text{(A.1)}$$

where u_n and v_n are the local Cartesian coordinates of the atoms (see Fig. 1), and the third coordinate w_n forms a triplet with them.

If we confine our attention to waves with long wavelengths and switch from a discrete set $\{u_n(t)\}$ to a field $u(x, t)$ (and similarly for v and w), then Eq. (A.1) becomes

$$v(x, t) \approx \frac{s_0^2}{2c_0} l_0 u_x(x, t).$$

We can see that $v \sim (c/4L)u$ (c is the period of the chain and L is the characteristic scale of variation of u). Thus, for long wavelengths the coordinate v can be neglected in the zeroth approximation.

APPENDIX B

Relation between the Constants of the Analytical Model and the Parameters of the Numerical Model of the Crystal

The inertial constants (5) can be easily obtained by writing the kinetic energy of an atom in cylindrical coordinates (z, r, ϕ) with r = const.

To express the constant K_ψ in terms of K_θ, we shall write the increment to the angle θ_n in the linear approximation:

$$\theta_n - \theta_0 \approx \frac{c_0}{l_0}(u_{n+1} - u_{n-1}) + \frac{s_0}{l_0}(v_{n+1} + 2v_n + v_{n-1}).$$

It is independent of w (in the linear approximation). Using Eq. (A.1) it becomes

$$\theta_n - \theta_0 \approx \frac{1}{l_0 c_0}(u_{n+1} - u_{n-1}). \qquad \text{(B.1)}$$

Switching to the variable $\psi_n = (2\pi/c)u_n$, we obtain from the formula (1) for the deformation energy of the valence angle $U_3(\theta_n - \theta_0)$ and the relation (B.1) the expression (7) for K_ψ.

To obtain the constant K_ϕ we note that the conformational angle τ_n between the atomic planes [nth, (n + 1)st, (n +2)nd] and [(n – 1)st, nth, and (n + 1)st] with slow variation (for long-wavelength waves) is $\tau_n = \pi - \chi_n$, $\chi_n \ll 1$. The angle χ_n in the linear approximation does not depend on u_n and v_n and is equal to

$$\chi_n \approx \frac{1}{2l_0 c_0 s_0}(w_{n+2} + w_{n+1} - w_n - w_{n-1}).$$

In a cylindrical coordinate system with r = const this difference can be easily represented in terms of the difference of the angles ϕ_j:

$$\chi_n \approx \frac{1}{4s_0}(\phi_{n+2} + \phi_{n+1} - \phi_n - \phi_{n-1}). \qquad \text{(B.2)}$$

Using the harmonic (in the small angle χ_n) approximation of the potential $U_4(\tau_n)$ of the conformational angle and switching to the continuum approximation, we obtain from Eqs. (2) and (B.2) the relation (6) between K_ϕ and the constants β and γ for the numerical potential.

REFERENCES

1. R. H. Boyd, Polymer **26**, 323 (1985).
2. R. H. Boyd, Polymer **26**, 1123 (1985).
3. L. I. Manevich and A. V. Savin, Vysokomol. Soedin., Ser. A **40**, 931 (1998).
4. A. V. Savin and L. I. Manevitch, Phys. Rev. B **58**, 11386 (1998).

5. V. V. Ginzburg, L. I. Manevich, and N. G. Ryvkina, Mekh. Kompoz. Mater. **2**, 249 (1991).
6. *Solitons in Action*, Ed. by K. Lonngren and A. Scott (Academic, New York, 1978; Mir, Moscow, 1981).
7. V. V. Ginzburg and L. I. Manevich, Fiz. Tverd. Tela (Leningrad) **32**, 2414 (1990) [Sov. Phys. Solid State **32**, 1401 (1990)].
8. D. H. Reneker and J. Mazur, Polymer **29**, 3 (1988).
9. L. I. Manevich and A. V. Savin, Vysokomol. Soedin., Ser. A **40**, 788 (1998).
10. D. W. Noid, B. G. Sumpter, and B. Wunderlich, Macromolecules **24**, 4148 (1991).
11. E. A. Zubova, N. K. Balabaev, and L. I. Manevich, Zh. Éksp. Teor. Fiz. **115**, 1063 (1999) [JETP **88**, 586 (1999)].
12. N. K. Balabaev, O. V. Gendel'man, M. A. Mazo, and L. I. Manevich, Zh. Fiz. Khim. **69**, 24 (1995).
13. D. W. Noid, B. G. Sumpter, and B. Wunderlich, Macromolecules **23**, 664 (1990).
14. D. Rigby and R. J. Roe, Macromolecules **22**, 2259 (1989).
15. M. P. Allen and P. J. Tildesley, *Computer Simulation of Liquids* (Clarendon, Oxford, 1987).
16. P. G. Khalatur, N. K. Balabaev, and A. S. Pavlov, Mol. Phys. **59**, 753 (1986).
17. L. I. Manevitch and A. V. Savin, Phys. Rev. E **55**, 4713 (1997).
18. R. Rajaraman, *Solitons and instantons* (North-Holland, Amsterdam, 1982; Mir, Moscow, 1985).
19. D. Bazeia and E. Ventura, Chem. Phys. Lett. **303**, 341 (1999).
20. E. A. Zubova and L. I. Manevich, in *Proceedings of Scientific Conference of Semenov Institute of Chemical Physics, Russian Academy of Sciences, Moscow, 1999*, p. 46.
21. Fei Zhang and M. A. Collins, Phys. Rev. E **49**, 5804 (1994).
22. O. Braun, Y. Kivshar, and M. Peyrard, Phys. Rev. E **56**, 6050 (1997).
23. P. Sodano, M. El-Batanouny, and C. R. Willis, Phys. Rev. B **34**, 4936 (1986).

Translation was provided by AIP

Journal of Experimental and Theoretical Physics, Vol. 91, No. 3, 2000, pp. 524–530.
Translated from Zhurnal Éksperimental'noĭ i Teoreticheskoĭ Fiziki, Vol. 118, No. 3, 2000, pp. 602–609.

SOLIDS
Structure

Energy Level Crossing and Magnetic Anomalies in $PrVO_4$ in Strong Fields

Z. A. Kazeĭ[a, *], N. P. Kolmakov[b], V. V. Platonov[c], A. A. Sidorenko[b], and O. M. Tatsenko[c]

[a]*Moscow State University, Moscow, 119899 Russia*
[b]*Bryansk State Technical University, Bryansk, 241035 Russia*
[c]*Russian Federal Nuclear Center, Institute of Experimental Physics, Sarov, Nizhegorodsk oblast, 607190 Russia*
**e-mail: kazei@plms.phys.msu.su*

Received February 23, 2000

Abstract—The magnetic anomalies due to the interaction of the energy levels of a rare-earth ion in a strong magnetic field are studied experimentally and theoretically for the van-Vleck paramagnet $PrVO_4$. A maximum is discovered in the differential susceptibility dM/dH in a field $H_c \approx 45$ T, where the lower energy levels cross. The magnetocaloric effect in pulsed fields is calculated assuming the magnetization process to be adiabatic. This effect is characterized by the absence of initial heating of the sample when the field is turned on and strong cooling as H_c is approached. It is shown that in $PrVO_4$, which is an enhanced nuclear magnet, the hyperfine interaction plays an extremely important role for the magnetic anomalies associated with crossover. For another van-Vleck paramagnet, $HoVO_4$, it is shown that a second crossover occurs near 310 T and the magnetocaloric effect is calculated. © *2000 MAIK "Nauka/Interperiodica".*

1. INTRODUCTION

Recently, we began experimental and theoretical investigations of magnetic anomalies occurring in strong and superstrong magnetic fields as a result of the interaction of the energy levels of a rare-earth ion for paramagnets with the tetragonal structure of zircon RXO_4, X = P, V. In [1] magnetic susceptibility anomalies and the magnetocaloric effect for $YbPO_4$ in fields up to 400 T, obtained by an explosive method [2], were investigated. A wide peak in the differential susceptibility dM/dH in a field $H_c \approx 280$ T was observed. This peak is due to the crossing of energy levels (crossover) of magnetic ions in the field and a jump on the magnetization curve. Assuming the magnetization process to be adiabatic in pulsed fields, the magnetocaloric effect, characterized by a nonmonotonic field dependence and accompanied by substantial cooling of the crystal near H_c, was calculated. In [1] it was shown for $YbPO_4$ that crossover is very sensitive to the crystal field parameters and investigation of the anomalies due to them yields important information about the crystal field, which, as is well known, remains one of the main factors determining the physical properties of rare-earth compounds.

The quite low tetragonal symmetry of zircon (space group $D_{4h}^{19} = I4_1/amd$) and the absence of nonequivalent positions for the rare-earth ion give rise to the rich, weakly degenerate spectrum of the rare-earth ion and the large magnetic anisotropy along and perpendicular to the tetragonal axis of the crystal. Crossing or convergence of energy levels of the rare-earth ion in a magnetic field occurs, as our calculations in [3] showed, for all rare-earth vanadates and phosphates with zircon structure. For different rare-earth zircons with the field oriented along various crystallographic directions, one or two crossovers can occur in average fields—up to 40 T (e.g., in $DyVO_4$, $TmVO_4$, and $TbVO_4$), high fields—up to 100 T (e.g., in $PrVO_4$, $ErVO_4$), and superhigh fields—up to 300–500 T (e.g., in $NdVO_4$, $TmVO_4$). At lower temperatures crossover is accompanied by a jump on the magnetization curve. The character of this jump is substantially different in isothermal and adiabatic regimes for different rare-earth zircons and is determined by the specific nature of the interaction of the energy levels. The interaction of energy levels in a magnetic field is accompanied by anomalies not only of the magnetic but also magnetoelastic properties, as we showed in our study of the magnetostriction of $DyVO_4$ [4]. Several works study crossover near 12 T for field orientation along the tetragonal axis in the singlet paramagnet $HoVO_4$: the Zeeman effect is measured in [5] and calculated in [6], anomalies arising in the elastic constants as a result of crossover are investigated in [7], and special features of the magnetization curves, specifically, the strong character of the jumps at low temperatures down to 0.1 K, have been studied in detail in static fields [8] and at $T = 4.2$ K in pulsed fields [9].

2. EXPERIMENTAL RESULTS

In the present paper we continue the investigation of crossover phenomena for another zircon—praseodymium vanadate $PrVO_4$. The measurements were performed at 4.2 K by the induction method in pulsed

magnetic fields up to 400 T, obtained by an explosive method [2], on a $PrVO_4$ single crystal. The field rise time in the pulse was 15 μs. These measurements are single measurements and the measuring coils and samples are destroyed after each pulse. The signal in the measuring coils relative to the magnetic field cannot be completely compensated. Thus, the signal induced in the measuring coils can be written in the form

$$V_1(H) \sim \frac{dM}{dt} + K\frac{dH}{dt}, \tag{1}$$

where the first term corresponds to the signal from the sample and the second term corresponds to the decompensation signal of the coils. The signals V_1 and $V_2 \sim dH/dt$ from the measuring and "field" coils (about 8000 points) were recorded during the field pulse with an interval of 0.002 μs. These data make it possible to calculate the curves $V_1(H)/V_2(H)$, which, assuming the decompensation signal to depend on the magnitude of the field, are proportional, to within the constant K, to the differential susceptibility of the sample $dM/dH = (dM/dt)/(dH/dt)$, and to perform time-averaging of the signal in order to decrease the high-frequency fluctuations of the background. The discreteness of the signal measurements with step 0.002 μs can in principle result in cutoff of the sharp peaks of the signal dM/dH, which, as our analysis showed, does not exceed in our case 20% of the height of the peak. Sometimes, an oscillating signal, whose nature is not clear and which was removed when the data were processed, is induced in the measuring coils during a pulse.

The experimental and theoretical curves of dM/dH for a $PrVO_4$ single crystal with magnetic field along the [001] tetragonal axis are presented in Fig. 1. The sharp maximum of the susceptibility at $H_c \approx 45$ T is due to, as will be shown below, to the crossing of the low energy levels of the Pr^{3+} ion. The large magnitude of the signal in fields below the crossover field on the experimental curve $(dM/dH)_{\text{exp}}$, in contrast to the computed curve $(dM/dH)_{\text{cal}}$, is due to the decompensation of the coils and the large measurement error for weak fields.

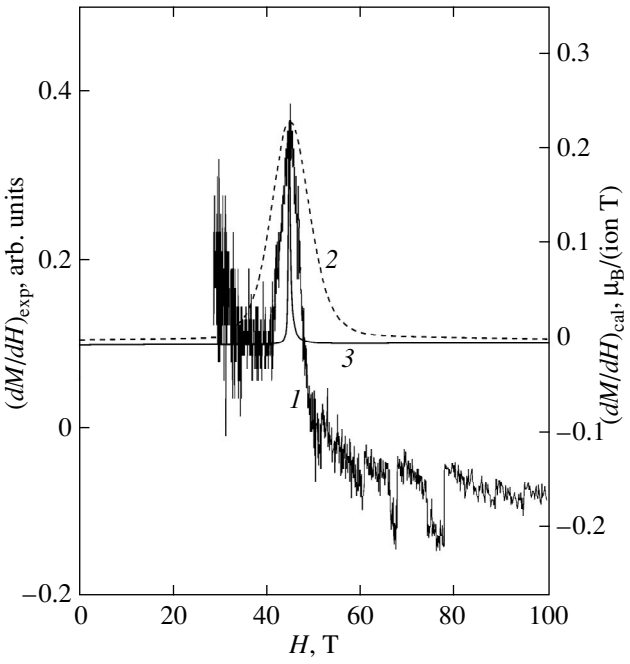

Fig. 1. (*1*) Experimental and theoretical (*2*) isothermal and (*3*, divided by 22) adiabatic differential susceptibilities dM/dH of a $PrVO_4$ single crystal for magnetic field along the [001]-axis at $T = 4.2$ K.

3. CALCULATION AND DISCUSSION

Hamiltonian. To calculate the Zeeman effect and the magnetic characteristics in a strong magnetic field, we used a Hamiltonian that includes the crystal field Hamiltonian H_{CF}, the Zeeman term H_z, and the hyperfine interaction Hamiltonian H_N:

$$\begin{gathered} H = H_{CF} + H_z + H_N, \\ H_{CF} = B_0^2C_0^2 + B_0^4C_0^4 + B_0^6C_0^6 \\ + B_4^4(C_4^4 + C_{-4}^4) + B_4^6(C_4^6 + C_{-4}^6), \\ H_z = g_J\mu_B\mathbf{HJ}, \quad H_N = H_Q + H_{HF} + H_{IZ}, \\ H_Q = (\mathbf{IPI}), \quad H_{HF} = (\mathbf{JAI}), \quad H_{IZ} = g_I\mu_B\mathbf{HI}. \end{gathered} \tag{2}$$

Here C_q^k are irreducible tensor operators, B_q^k are crystal-field parameters, g_J is the Landé factor, P is the quadrupole hyperfine interaction tensor, A is the magnetic hyperfine interaction tensor, and g_I is the nuclear g factor [10–12]. The hyperfine interaction for $PrVO_4$ was described in [13] using a spin Hamiltonian of the form

$$\begin{aligned} H_{NS} = &-\hbar[\gamma_\parallel H_zI_z + \gamma_\perp(H_xI_x + H_yI_y)] \\ &+ P\left[I_z^2 - \frac{1}{3}I(I+1)\right]. \end{aligned} \tag{3}$$

We used such a form for H_N for quantitative calculations with the values of the hyperfine parameters determined in [13].

In [13] the structure of the lower levels of $PrVO_4$ was established by means of fluorescence and Raman and infrared spectroscopy investigations. In the accordance with this structure, a singlet lies 35 cm^{-1} above the ground singlet state of the Pr^{3+} ion and a doublet lies ~84 cm^{-1} above the ground state. The positions of all other levels of the $^3H_4$ multiplet were determined in this work by an extrapolation procedure employing the symmetry properties of the Hamiltonian of a tetragonal crystal field and the experimental NMR data. The data obtained for the energies and wave functions of the Pr^{3+} ion enabled the authors to describe the measurements of the nuclear magnetic resonance spectrum and the

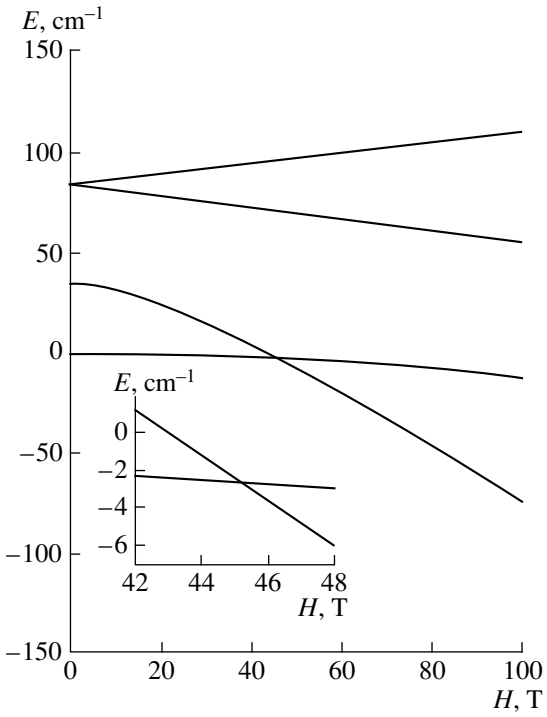

Fig. 2. Zeeman effect for $PrVO_4$ with $H \parallel [001]$ (the four lowest levels of the ground state multiplet are shown). Inset: Formation of a gap with crossover for field disorientation by 5° from the tetragonal axis.

low-temperature initial magnetic susceptibility. However, the values of the crystal field parameters are not presented in [13]. The crystal field parameters in $PrVO_4$ were proposed in [14]. They were determined from measurements of the magnetization in the temperature range 4–40 K and in magnetic fields up to 5 T with field orientation along and perpendicular to the tetragonal axis, using spectroscopic information from [13]. We note that the wave functions that they give for all multiplet levels and the positions of the upper levels differ substantially from those proposed in [13]. The crossover field with these parameters is found to be 86.5 T, which is much greater than our experimentally determined value $H_c \approx 45$ T.

Using the data in [13] for the splitting of the ground state multiplet of the Pr^{3+} ion, the magnetic susceptibility data [13, 15], and the magnetization curves [14] as well as our value of the crossover field T, we determined the crystal field parameters for the ion Pr^{3+} in $PrVO_4$. These parameters describe well all known experimental data for $PrVO_4$, and they reproduce the wave functions proposed in [13]. These parameters are not presented in the present work, since they differ from the parameters established reliably for other rare-earth vanadates (e.g., for $HoVO_4$ in [6]) by an amount greater than is considered to be admissible for measurements of parameters in the rare-earth series. The approximate character of the structure of the upper part of the multiplet of the Pr^{3+} ion in $PrVO_4$ [13] makes it impossible to give more accurate values for these parameters at the present time. It is interesting that quite small variations of the crystal-field parameters do not change the physical essence of the phenomena examined in the present work.

Zeeman effect and magnetization curves. Figure 2 shows the variation of the spectrum of the Pr^{3+} ion for field orientation along the tetragonal axis neglecting the hyperfine interaction (Zeeman effect). The total splitting of the multiplet is ~550 cm^{-1}; only the four lowest levels are presented in the figure. It is evident that crossing of the lower energy levels is observed in fields T. The symmetry of these levels is such (in the J, M_J representation the wave function of the lowest singlet level is the state $|4; 2^S\rangle$, while for the excited singlet the state is $|4; 4^S, 0\rangle$, i.e., a mixture of the states $|4^S\rangle$ and $|0\rangle$) that in a field $H \parallel [001]$ may cross without the formation of a gap. It is obvious that the calculation of the Zeeman effect in different bases—multiplet, term, or configuration—gives the same result: there is no gap at H_c, since the symmetry of the crossing levels, taking account of mixing with respect to the quantum numbers L, S, and J, does not change. Repulsion of these levels near crossover and gap formation are obtained when the hyperfine interaction in a magnetic field is taken into account. It will be shown below that when the hyperfine splitting of the electronic levels, which results in a large change of the electronic specific heat, and the nuclear specific heat are neglected, the magnetocaloric effect in pulsed fields cannot be calculated correctly. The large role of the hyperfine interaction in van-Vleck paramagnets, which are enhanced nuclear magnets and which $PrVO_4$, studied in the present work, is, is well known (see, e.g., [16]).

A disorientation of the field by several degrees results in an interaction of the first and second levels, the appearance of a small gap near H_c, and a shift of H_c to higher values. For disorientation angle 5° the gap is ~0.2 cm^{-1}, and the shift is $H_c \approx 1$ T (see inset in Fig. 2). A calculation of the Zeeman effect for $H \parallel [110]$ shows the presence of crossover in fields ≈290 T. The magnetization curves along the main crystallographic directions for $T = 4.2$ K are shown in Fig. 3. The magnetization jumps correspond to crossings of the lower energy levels. The magnitude of the jump (~$2.5\mu_B$) is much larger for $H \parallel [001]$ than for $H \parallel [110]$ (~$0.5\mu_B$). The figure also shows the adiabatic magnetization curves for initial temperature $T_0 = 4.2$ K, which were calculated taking account of the magnetocaloric effect. It is evident that for field directions characterized by crossover the adiabatic magnetization curves can be steeper than the isothermal curves. This attests to cooling of $PrVO_4$ as the crossover field is approached. For $H \parallel [100]$ the adiabatic and isothermal magnetization curves are identical; this is characteristic for the van-Vleck character of the magnetization in the "easy" direction.

Magnetocaloric effect. The adiabatic magnetization curves must be calculated in order to interpret magnetization processes in pulsed fields with quite short pulse durations. As investigations of magnetization processes in rare-earth dielectric paramagnets, performed in [17] on $Gd_3Ga_5O_{12}$, show (see references in [17]), if a strong magnetic field is turned on sufficiently rapidly (with rates much higher than 10^3 T/s) the magnetization process occurs without hysteresis and therefore without heat exchange with the surrounding medium through the surface of the sample. It is obvious that the presence of hysteresis will indicate that the process is irreversible and hence various temperature regimes exist as the field is switched on and off. In our experiment the field was switched on at a rate higher than the upper estimates, made in [17], for satisfaction of the adiabaticity condition. Consequently, we consider the process of magnetization of the $PrVO_4$ single crystal to be adiabatic. We note that if the field is switched on too rapidly, the magnetization can lag the field even in a paramagnet; this once again leads to the appearance of hysteresis.

To calculate the magnetic characteristics the Hamiltonian (2) was diagonalized numerically for each value of the field from 0 to 400 T with step $\Delta H = 0.01$ T in order to determine the spectrum and wave functions of the Pr^{3+} ion and the "elementary" magnetocaloric effect ΔT was calculated with the field varying from H to $H + \Delta H$:

$$\Delta T = -T\left(\frac{\partial M}{\partial T}\right)_H \frac{\Delta H}{C_H}. \qquad (4)$$

Here C_H is the total specific heat of a sample per formula unit, including the specific heat of the lattice

$$C_{\text{lat}} = \frac{12\pi^4}{5} k_B \nu \left(\frac{T}{T_D}\right)^3$$

(for a zirconium lattice the Debye temperature is T_D = 275 K [18], $\nu = 6$), and the electronic C_{el} and nuclear C_n magnetic specific heats, calculated for each value of the field and temperature on the basis of the electronic (taking account of the hyperfine interaction) and nuclear spectra of the rare-earth ion using the standard formula

$$C_{\text{el(n)}} = \frac{1}{k_B T^2}[\langle E^2_{\text{iel(n)}}\rangle - \langle E_{\text{iel(n)}}\rangle^2]. \qquad (5)$$

The parameters of the hyperfine interaction spin Hamiltonian, which were determined for $PrVO_4$ in [13], were used for the calculations. The computed dependences of the specific heat of the sample on the magnetic field for field orientation along the tetragonal axis are presented in Fig. 4 for $T_0 = 4.2$ K. It is evident that the nuclear specific heat plays a large role close to crossover, where the temperature becomes comparable to the nuclear multiplet splitting, i.e., as temperature decreases to tenths of a Kelvin (see Fig. 5). The lattice specific heat is negligibly small in this temperature range. The electronic subsystem makes the main contribution to the specific heat. The maxima in the curve $C_{\text{el}}(H)$ are Schottky anomalies. The dependence $C_{\text{el}}(H)$

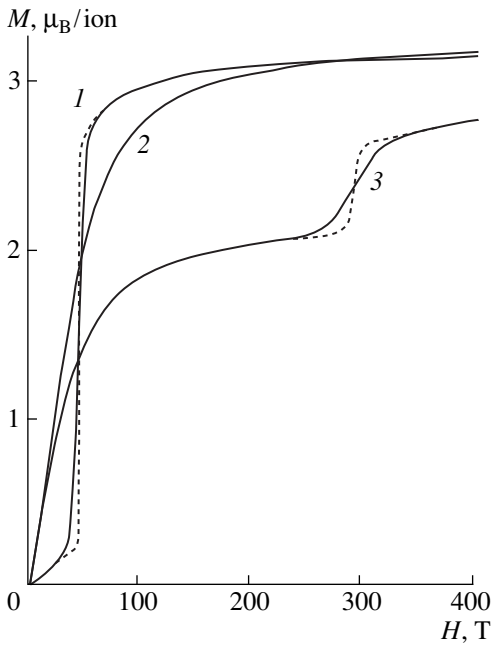

Fig. 3. Isothermal (solid lines) and adiabatic (dashed lines) magnetization curves for $PrVO_4$ at $T_0 = 4.2$ K: (*1*) $H \parallel [001]$, (*2*) $H \parallel [100]$, (*3*) $H \parallel [110]$.

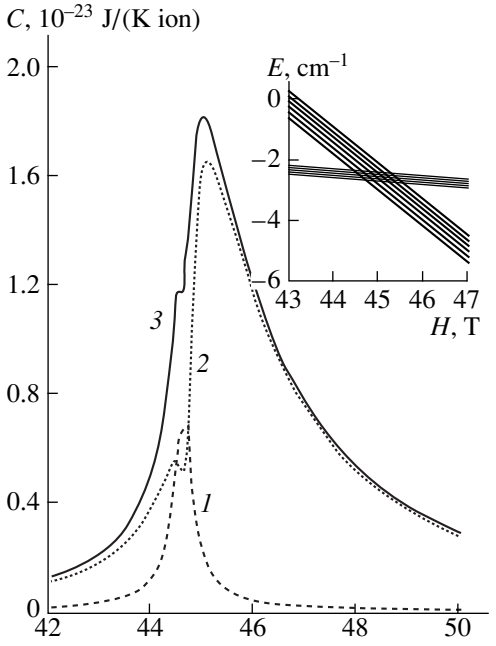

Fig. 4. Field dependences of the (*1*) nuclear, (*2*) electron, and (*3*) total specific heats of $PrVO_4$ with adiabatic magnetization along the tetragonal axis, $T_0 = 4.2$ K. Inset: Fragment of the Zeeman effect for the two lowest levels taking account of the hyperfine interaction.

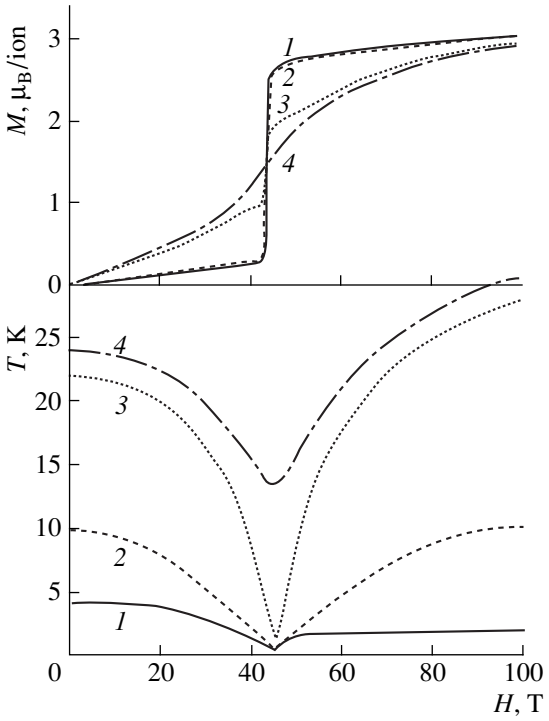

Fig. 5. Adiabatic magnetization curves and magnetocaloric effect for $PrVO_4$ with $H \parallel [001]$ for different initial temperatures T_0 = (*1*) 4.2, (*2*) 10, (*3*) 22, and (*4*) 24 K.

is not symmetric with respect to H_c. This asymmetry is due to the different magnitudes of the hyperfine splitting of the two lower levels of the Pr^{3+} ion and can be easily understood from the inset in Fig. 4.

Thus, with adiabatic magnetization the $PrVO_4$ sample cools from $T_0 = 4.2$ to $T \approx 0.17$ K near the crossover field H_c and then heats up to $T \approx 1.5$ K (Fig. 5). For $H > H_c$ the temperature does not increase to the initial value. This is due to the different rate of change of the magnetic part of the entropy of the system before and after crossover. The magnetocaloric effect close to H_c is symmetric with respect to H_c. However, a further increase of the field results in a large decrease of $(\partial M/\partial T)_H$, since the possibilities for decreasing the magnetic part of the entropy are exhausted (all electrons at these temperatures are in the bottom energy level and further lowering of this level with increasing field cannot change the entropy), and in a slower increase of the temperature. It is interesting to note that this asymmetric character of the magnetocaloric effect vanishes as the initial temperature increases above $T_{01} \approx 10$ K. Above this value of T_0 the minimum temperature which is reached with $H = H_c$ is not small enough in order for the mechanism described above to come into play. As T_0 increases further above $T_{02} \approx 22$ K the character of the magnetocaloric effect changes even more (see Fig. 5), since the character of the adiabatic magnetization curves changes substantially—the important section near the crossover field almost vanishes. We note that the height of the peak in the susceptibility decreases, but the width of the peak does not change—the peak remains narrow. The quantity T_{02} is determined by the energy gap between the ground and excited singlets for $H = 0$. Both characteristic temperatures T_{01} and T_{02} are determined by the electronic structure, formed by the crystal field, of the rare-earth ion. All this means that investigation of the temperature dependences of the magnetic anomalies with crossover should yield important information about the crystal field.

The computed magnetocaloric effect concerns the electronic subsystem of the compound under study. Whether or not the temperature of the sample decreases to the computed values depends on the spin-lattice relaxation time. According to estimates made for rare-earth van-Vleck dielectric paramagnets in a recent review [16, pp. 354, 378] (see also the citation in [16]), this time is 10^{-9}–10^{-10} s. However, the variance in the spin-lattice relaxation times is quite large for the compounds studied at low temperatures. According to some estimates made in [19] they can be 10^{-2}–10^{-3} s. If the spin-lattice relaxation times for $T \lesssim 1$ K are such, then the electronic subsystem is decoupled from the lattice under magnetization conditions in pulsed fields with rapid buildup of the field and it is impossible for the electronic subsystem to cool down, shedding heat into the lattice. In this case the larger width of the peak in the experiment than in theory can be interpreted as information about the spin-lattice relaxation times.

Magnetic susceptibility. We now return to the discussion of Fig. 1, where it is evident that the peak in the isothermal differential susceptibility, calculated at 4.2 K, is wider than the experimental peak. This shows that the $PrVO_4$ sample cools down below $T_0 = 4.2$ K. The susceptibility maximum calculated assuming the magnetization process to be adiabatic is narrower than the experimental peak. We note that if the hyperfine splitting of the electronic levels is neglected, the maximum becomes even narrower (approximately by a factor of 2). The hyperfine quadrupole interaction contributes to the width of the peak because of the lifting of the degeneracy of the energy levels even in zero field; the effect calculated with existing hyperfine interaction parameters [13] is three orders of magnitude weaker than the effect due to the hyperfine magnetic interaction. Taking into account a possible disorientation of the field in an experiment by several degrees relative to the tetragonal axis gives negligible broadening of the peak and a shift of the peak in the direction of high fields. The reasons why the width of the peak does not correspond to the differential susceptibility dM/dH, calculated assuming the magnetization process to be adiabatic, which for pulse duration ~15 μs is difficult to doubt, as discussed above, are still not clear. By analogy to $HoVO_4$, one reason could be, for example, the complicated structure of the peak, which in the situation where it is not

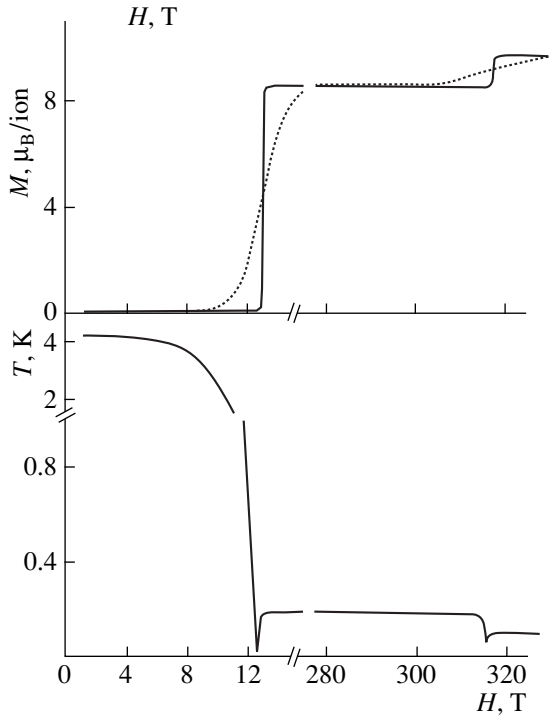

Fig. 6. Adiabatic (solid line) and isothermal (dotted line) magnetization curves and magnetocaloric effect in $PrVO_4$ with $T_0 = 4.2$ K and $H \parallel [001]$, calculated taking account of the hyperfine interaction.

resolved results in broadening. The fine structure of the peak at crossover in $HoVO_4$, consisting of the presence of two peaks and resulting in a total width of the peak of at least 1.5 T (in an isothermal regime at 0.1 K), has been observed in measurements in static [8] and pulsed [9] fields. The nature of such a profile of the peak has not been definitely established; the authors of [8, 9] conjecture either the influence of mechanical stresses in the sample or the existence of nonequivalent positions for the rare-earth ion. The fine structure of the peak in the differential susceptibility at crossover in $PrVO_4$ is attributed in [20] to the possibility of spin ordering in a magnetic field.

Magnetic anomalies at crossover in $HoVO_4$. As mentioned in the introduction, crossover in $HoVO_4$ in fields ≈11.8 T oriented along the tetragonal axis and the associated anomalies in various properties of the crystal have been investigated in a number of works. However, the magnetocaloric effect was not discussed in any of these works. The magnetocaloric effect in pulsed fields for $HoVO_4$ was calculated with the crystal-field and hyperfine interaction parameters from [6]. The Zeeman effects for $PrVO_4$ and $HoVO_4$ are very similar; the differences are only qualitative, for example, the gap separating the excited level from the ground state singlet is different (21 cm^{-1} for $HoVO_4$ and 35 cm^{-1} for $PrVO_4$) and, consequently, the crossover field is different (11.8 T for $HoVO_4$ and 45 T for $PrVO_4$). This suggests that the magnetocaloric effect, its evolution as a function of the initial temperature of the sample, and the behavior of the characteristics determining the effect (such as the specific heat) will be similar in $PrVO_4$ and $HoVO_4$. Our calculations confirm this. As an example, the isothermal and adiabatic magnetization curves (the latter being steeper than the isothermal curve) and the magnetocaloric effect for $T_0 = 4.2$ K and direction of the field along the tetragonal axis, which were calculated taking account of the hyperfine interaction, are displayed in Fig. 6. It is evident from Fig. 6 that in $HoVO_4$ with the field oriented along the tetragonal axis a second crossover occurs in a field ~310 T, in which the magnetization of the compound increases from the nominal value to $10\mu_B$/ion. For the hyperfine interaction parameters used, the temperature at the first crossover decreases to ~0.03 K and then increases to ~0.2 K; at the second crossover it decreases to ~0.08 K and then once again tends to a constant value ~0.13 K.

4. CONCLUSIONS

We call attention once again to the substantially different form of all anomalies associated with crossover in the singlet paramagnets $PrVO_4$ and $HoVO_4$ and in the paramagnet with the Kramers ion $YbPO_4$ [1]. In $PrVO_4$ and $HoVO_4$ the susceptibility peaks are much narrower than in $YbPO_4$, the magnetocaloric effect is anomalous—there is no initial increase of sample temperature when the field is switched on. The differences are due to the different character of the magnetization of these compounds. In $PrVO_4$ and $HoVO_4$ the magnetization before crossover is very small, and the magnetic susceptibility is of van-Vleck origin; in $YbPO_4$ the magnetization up to the moment of crossover is quite large, and the Curie term is the main term in the magnetic susceptibility. For singlet van-Vleck paramagnets $PrVO_4$ and $HoVO_4$ the hyperfine interaction, which modifies the electronic specific heat and forms a nuclear specific heat through the hyperfine splitting of the electronic and nuclear spectra, plays a very important role near the crossover field.

ACKNOWLEDGMENTS

This work was partially supported by the Russian Foundation for Basic Research (project no. 00-02-17756).

REFERENCES

1. Z. A. Kazeĭ, N. P. Kolmakova, R. Z. Levitin, *et al.*, Pis'ma Zh. Éksp. Teor. Fiz. **65**, 691 (1997) [JETP Lett. **65**, 720 (1997)].
2. A. I. Pavlovskiĭ, N. P. Kolokol'chikov, and O. M. Tatsenko, *Megagauss Physics and Techniques*, Ed. by P. Turchi (Plenum, New York, 1980), p. 1.

3. Z. A. Kazeĭ, N. P. Kolmakova, V. V. Platonov, *et al.*, in *Abstracts of 3rd International Conference on f-Elements, Paris, 1997*, p. 396.
4. Z. A. Kazeĭ, N. P. Kolmakova, A. A. Sidorenko, and V. V. Snegirev, in *Abstracts of 7th European Magnetic Materials and Applications Conference, Zaragoza, 1998*, p. 214.
5. J. E. Battison, A. Kasten, M. J. M. Leask, and J. B. Lowry, Phys. Lett. **55**, 173 (1975).
6. B. Bleaney, J. F. Gregg, P. Hansen, *et al.*, Proc. R. Soc. London **416**, 63 (1988).
7. T. Goto, A. Tamaki, T. Fujimura, and H. Unoki, J. Phys. Soc. Jpn. **55**, 1613 (1986).
8. P. Morin, J. Rouchy, and Z. Kazei, Phys. Rev. B **51**, 15103 (1995).
9. Z. A. Kazeĭ and Yu. F. Popov, Fiz. Tverd. Tela (St. Petersburg) **36**, 2099 (1994) [Phys. Solid State **36**, 1146 (1994)].
10. B. G. Wybourne, *Spectroscopic Properties of Rare-Earths* (Interscience, New York, 1965), p. 236.
11. B. Bleaney, in *Hyperfine Interactions*, Ed. by A. J. Freeman and R. B. Frankel (Academic, New York, 1967; Mir, Moscow, 1970).
12. A. Abragam and B. Bleaney, Proc. R. Soc. London **387**, 221 (1983).
13. B. Bleaney, R. T. Harley, J. F. Ryan, *et al.*, J. Phys. C **11**, 3059 (1978).
14. R. R. Andronenko, S. I. Andronenko, and A. N. Bazhan, Fiz. Tverd. Tela (St. Petersburg) **36**, 2396 (1994) [Phys. Solid State **36**, 1302 (1994)].
15. M.-D. Guo, A. T. Aldred, and S.-K. Chan, J. Phys. Chem. Solids **48**, 229 (1987).
16. L. K. Aminov, B. Z. Malkin, and M. A. Teplov, in *Handbook on the Physics and Chemistry of Rare Earths*, Ed. by K. A. Gschneidner, Jr. and L. Eyring (Elsevier, Amsterdam, 1996), vol. 22, p. 295.
17. R. Z. Levitin, V. V. Snegirev, A. V. Kopylov, *et al.*, J. Magn. Magn. Mater. **170**, 223 (1997).
18. A. Kasten, H. G. Kahle, P. Klofer, and D. Schafer-Siebert, Phys. Status Solidi B **144**, 423 (1987).
19. S. A. Altshuler and B. M. Kozyrev, *Electron Paramagnetic Resonance in Compounds of Transition Elements* (Nauka, Moscow, 1972; Halsted, New York, 1975).
20. K. Amaya, in *Recent Advances in Magnetism of Transition Metal Compounds* (World Scientific, Singapore, 1993), p. 327.

Translation was provided by AIP

Journal of Experimental and Theoretical Physics, Vol. 91, No. 3, 2000, pp. 531–536.
Translated from Zhurnal Éksperimental'noĭ i Teoreticheskoĭ Fiziki, Vol. 118, No. 3, 2000, pp. 610–616.

SOLIDS
Structure

Nonlinear Regimes of Resonance Precession of the Magnetization in a (111) Iron–Garnet Film

A. M. Shutyĭ* and D. I. Sementsov**

Ul'yanovsk State University, Ul'yanovsk, 432700 Russia
**e-mail: shut@mail.ru*
***e-mail: sements@quant.univ.simbirsk.su*

Received March 22, 2000

Abstract—The special features of precessional motion under ferromagnetic resonance conditions with perpendicular magnetization of the film are investigated on the basis of a numerical solution of the equations of motion of magnetization in a (111) type iron garnet film. Several nonlinear magnetization precession regimes exist for fixed values of the crystallographic and induced anisotropies. Depending on the value of the magnetizing field and the amplitude of the microwave field, precession occurs around the normal to the film with the third harmonic of the fundamental frequency making a small or large contribution to the nutation motion and with a small or large amplitude of the precession angle. Precession regimes around one of three symmetric directions different from the direction of the normal are possible. Narrow ranges of the static field, where dynamical bistability and regimes with a period which is a multiple of the period of the microwave field are realized, exist. © *2000 MAIK "Nauka/Interperiodica".*

1. INTRODUCTION

The nonlinear dynamics of the magnetization in magnetically ordered crystals is of interest because of the diversity of the nonlinear effects that arise when a dissipative spin system is exposed to a high-frequency field and also because of the possibility of attaining large precession angles and realizing dynamical chaos and various statistical and dynamical self-organizing structures [1–6]. One of the main manifestations of the nonlinear dynamics of magnetization for large angles of uniform precession is frequency doubling, which occurs with a linearly polarized high-frequency field. For precession in a transverse microwave field, as a rule, the analysis is confined to a given nonlinear effect. However, according to the analysis in this paper, as a result of a definite symmetry, associated with the crystallographic symmetry, of the magnetic anisotropy field in a material, primarily the higher-order harmonics of the fundamental precession frequency are manifested in the precessional motion of the magnetic moment. The interest in the behavior of the magnetization for large precession angles is also due to problems associated with the application of such precession for modulating laser radiation, whose efficiency is determined by the precession angle [7–9]. In the present paper the special features of the dynamics of the precessional motion of the magnetization in a nonlinear ferromagnetic resonance regime in a film with cubic symmetry and with the normal oriented along the crystallographic axis [111] and the direction of the magnetizing field are investigated.

Energy is transferred to spin waves from uniform precession with transverse magnetization by two mechanisms [3, 10]. The first mechanism involves a three-magnon process in which a magnon with the wave vector $\mathbf{k} = 0$ is annihilated and two magnons with wave vectors $\mathbf{k}$ and $-\mathbf{k}$ and frequency $\omega_{\mathbf{k}} = \omega$, where ω is the frequency of uniform precession, appear. The second mechanism involves a four-magnon process in which two magnons with $\mathbf{k} = 0$ vanish and two magnons with wave vectors $\mathbf{k}$ and $-\mathbf{k}$ and frequency $\omega_{\mathbf{k}} = \omega$ appear. Consequently, to attain large angles of uniform precession the precession frequency must equal the minimum frequency of the spin-wave spectrum, associated with the frequency of spin waves with $k = 0$ and direction along the magnetizing field, i.e., it is the ferromagnetic resonance frequency. Then neither mechanism for energy transfer from uniform precession to spin waves is realized. It is shown in [4, 11] that precession angles $\phi \approx 20$–$25°$ have been attained at the ferromagnetic resonance frequency in an iron-garnet film $Y_{2.9}La_{0.1}Fe_{3.9}Ga_{1.1}O_{12}$ grown on a gadolinium–gallium garnet substrate.

2. GENERAL EQUATIONS AND RELATIONS

Epitaxial iron garnet films are single-crystal layers with a cubic lattice. We shall assume that the crystallographic axis [111] is aligned with the x axis and the normal to the surface of the film and the axes $[11\bar{2}]$ and $[\bar{1}10]$ are aligned with the y and z axes; the polar angle θ and the azimuthal angle ψ of the magnetization vector

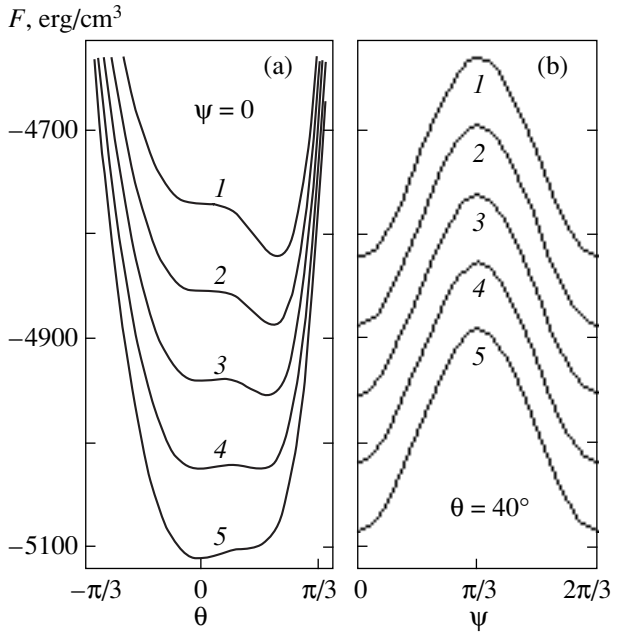

Fig. 1. Free energy density versus (a) the polar angle for the azimuthal direction $\psi = 0$ and (b) the azimuthal angle for the polar direction $\theta = 40°$ for the magnetizing field H = (*1*) 260, (*2*) 265, (*3*) 270, (*4*) 275, (*5*) 280 Oe.

M are measured from the *x* and *y* axes, respectively. We shall describe the dynamical behavior of the magnetization in external static **H** and varying **h** magnetic fields, which we shall assume are orthogonal to one another (**H** ⊥ **h**), by the equations of motion of the magnetization written in spherical coordinate systems [10]:

$$\begin{aligned} \dot{\psi} M \sin\theta &= \gamma \frac{\partial F}{\partial \theta} + \frac{\lambda}{M} \frac{1}{\sin\theta} \frac{\partial F}{\partial \psi}, \\ \dot{\theta} M &= \frac{\lambda}{M} \frac{\partial F}{\partial \theta} - \gamma \frac{1}{\sin\theta} \frac{\partial F}{\partial \psi}, \end{aligned} \tag{1}$$

where γ is the gyromagnetic ratio, λ is the damping parameter, and F is the free-energy density. Solving these equations makes it possible to find the precession frequency of the magnetic moment relative to its equilibrium orientation and the time dependence of the angles ψ and θ with a fixed geometry of the applied fields and time dependence of the external field. The resonance frequency ω_r is determined by the expression

$$\omega_r = \gamma H_{\text{eff}} = \frac{\gamma}{M \sin\theta} \sqrt{F_{\theta\theta} F_{\psi\psi} - F_{\theta\psi}^2}, \tag{2}$$

where the values of the second derivatives of the free-energy density are taken for the equilibrium values θ_0 and ψ_0, obtained from the conditions

$$\frac{\partial F}{\partial \theta} = 0, \quad \frac{\partial F}{\partial \psi} = 0.$$

For our orientation of the crystallographic axes the free-energy density is given by the expression

$$\begin{aligned} F &= -\mathbf{M}(\mathbf{H} + \mathbf{h}) + (K_u - 2\pi M^2) \sin^2\theta \\ &+ K_1 \left(\frac{1}{4} \sin^4\theta + \frac{1}{3} \cos^4\theta + \frac{\sqrt{2}}{3} \sin^3\theta \cos\theta \cos 3\psi \right), \end{aligned} \tag{3}$$

where K_u and K_1 are the constants characterizing the growth-induced and crystallographic anisotropies. Substituting Eq. (3) into Eq. (1) we can find the resonance frequency for an arbitrary orientation of the equilibrium magnetization.

As follows from the relations presented, together with the orientation and magnitude of the magnetizing field and the induced and crystallographic anisotropy fields, the polarization, amplitude, and initial phase of the microwave field also have a large effect on the dynamics of the magnetization in the precessional motion. For small amplitudes of the microwave field ($h \ll H$) linear ferromagnetic resonance, for which the precession angles are small and the time dependences $\theta(t)$ and $\psi(t)$ can be found from the linearized (with respect to the deviations of the magnetization from the equilibrium position) equations of motion, occurs at the frequency $\omega \approx \omega_r$. As the amplitude of the microwave field and, correspondingly, the precession angle increase, the contribution of the higher harmonics of the fundamental precession frequency to the indicated dependences increases and the nutation motion of the vector **M** becomes substantial. In this case the linear approximation is no longer adequate for solving Eqs. (1).

A detailed analysis of the special features of the precessional motion of magnetization taking account of all parameters determining the state of the magnetization in the film is possible only on the basis of numerical methods for solving Eqs. (1). As a simplification, the static field **H** is assumed to be perpendicular to the surface of the film, and the magnitude of the field is chosen to be such that for given values of the constants K_u and K_1 the equilibrium orientation of the vector **M** is in the direction of the normal ($\theta_0 = 0$). Then the resonance precession frequency is $\omega_r = \gamma H_{\text{eff}}(0)$, where the effective field is determined, according to Eq. (2), by the expression

$$H_{\text{eff}}(0) = H - 4\pi M + \frac{2}{M} \left(K_u - \frac{2}{3} K_1 \right). \tag{4}$$

The high-frequency field is assumed to be linearly polarized and lying in the *yz* plane, i.e., **h** ⊥ **H**. In this case several regimes of precessional motion of the magnetization occur, and their realization is determined by the magnitude of the static field H and, correspondingly, the frequency range of the ferromagnetic resonance.

To understand the special features of the precessional motion of the vector **M** it is necessary to know the energy relief given by the function $F(\theta, \psi)$. Figure 1 shows the free energy density as a function of the polar and azimuthal angles. The function $F(\theta)$ was con-

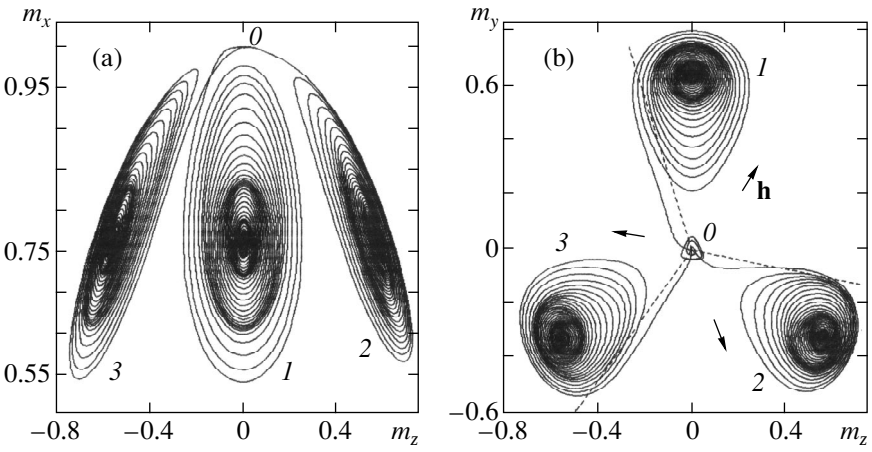

Fig. 2. Projections $m_\alpha = M_\alpha/M$ on the (a) *xz* and (b) *zy* planes of the magnetic moment, whose initial orientation is changed by the microwave field and which reaches one of the four stationary dynamical regimes characterized by a small precession amplitude. The amplitudes of the microwave field are h = (*0*) 0.04 and (*1–3*) 1.5 Oe; the orientation angles are ψ_h = (*1*) 0°, (*2*) 200°, and (*3*) 270°.

structed for the azimuthal direction $\psi = 0$, for which the function F characteristically has two local minima. The projection of the [001] crystallographic axis, which makes an angle $\theta = 54.7°$ with the normal to the film, is oriented in the direction $\psi = \pi$ (i.e., in the direction of the continuation of the indicated direction). The function $F(\psi)$ was constructed for the polar direction $\theta = 40°$, close to one of the vocal minima of the free energy with $\psi = 0$. The function $F(\psi)$ has a period equal to $2\pi/3$ in correspondence with the arrangement of the magnetizing field and the three {100} type crystallographic axes. The curves *1–5* correspond to various values of the magnetizing field. The calculation of these dependences, just as for the subsequent analysis, employed values of the parameters that are close to those of a real iron-garnet film with the indicated composition: $4\pi M = 214.6$ G, $\gamma = 1.755 \times 10^7$ (Oe s)$^{-1}$, $\lambda = 3 \times 10^6$ s^{-1}, $K_u = K_1 = -10^3$ ergs/cm^3 [11]. It follows from the curves presented that for the direction $\psi = \theta = 0$ (the vector **M** is parallel to the normal to the film) with magnetizing field $H > H_0$ the function $F(\theta)$possesses a local minimum which becomes more pronounced as H increases. The value of the field H_0 can be found from Eq. (4) with $\omega_r = 0$. For the structure chosen $H_0 = 254$ Oe. Besides the indicated minimum in the direction of the normal, there are also three local minima that correspond to the angles $\psi = 0°$, 120°, and 240° and $\theta \approx 35–41°$. As the magnetizing field increases, the minima slowly shift toward the normal and become less pronounced, vanishing for $H \approx 275$ Oe. Analysis shows that for sufficiently low resonance frequencies and weak magnetizing fields (for the experimental film $\omega_r < 4 \times 10^8$ s^{-1}, $H < 275$ Oe) the precession axis is aligned along the normal only for small amplitudes of the microwave field ($h \le 0.04$ Oe). Then the precession amplitude is several degrees ($\phi \approx 2°$). The form of the magnetization trajectory for $h \approx 0.04$ Oe is now substantially different from a circle because the vector **M** undergoes a nutation motion with the third harmonic of the precession frequency predominating. As the amplitude of the microwave field increases, the magnetization in the precessional motion tilts toward one of three directions (depending on the orientation of the field **h** in the *yz* plane and the initial phase of this field), which are determined by the corresponding local minima of the free-energy density with $\theta \neq 0$. Over times of the order of 200–500 ns precession of the magnetization along a stationary trajectory with average amplitude $\langle\phi\rangle \lesssim 3°$ is established near the indicated direction.

3. NUMERICAL ANALYSIS

The results of solving Eqs. (1) numerically that characterize the dynamics of the precessional motion in the experimental film are presented below. Figure 2 shows the projections of the magnetic moment, $m_\alpha = M_\alpha/M$, $\alpha = x, y, z$, on the *xz* and *zy* planes. The microwave field changes the initial orientation of the magnetic moment, as a result of which one of the four stationary dynamical regimes described above is established. The resonance precession frequency is chosen to be $\omega_r = 1.12 \times 10^8$ s^{-1}, which corresponds to the static field $H = 260$ Oe. The amplitude of the microwave field is $h = 0.04$ Oe (curve *0*), 1.5 Oe (curves *1–3*) and its initial phase is 0; the orientation angle, measured from the *y* axis, is $\psi_h = 0°$, 200°, and 270° (the curves *1–3*). For precession of the magnetic moment around the normal (the curve *0*) the initial phase and orientation of the high-frequency field are irrelevant. The dashed curves in the figure separate three sectors, corresponding to the initial orientational angles ψ_h of the high-frequency field **h** for which precession of the vector **M** around the corresponding direction is established. For an initial

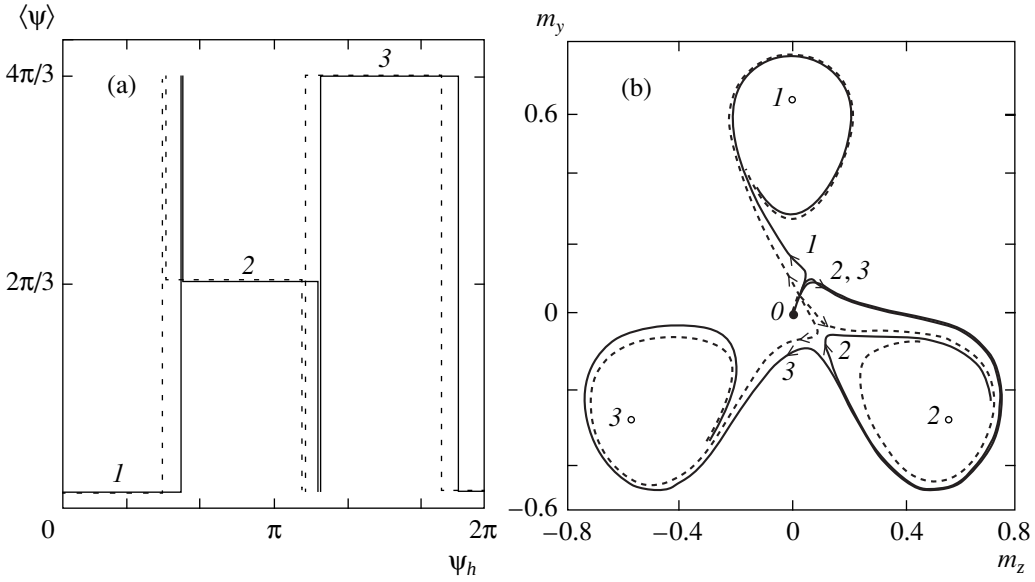

Fig. 3. (a) the angle $\langle\psi\rangle$ versus the orientation of a microwave field with amplitude $h = 0.5$ Oe (dashed curve) and 1.5 Oe (solid curve) in steady precessional motion and (b) projection of the trajectory of the vector M on the yz plane with a precession regime established near one of the three distinguished directions.

phase π, i.e., when the initial sign of the field **h** changes, the distinguished regions of the angles ψ_h shift by 180°.

Figure 3a shows the dependence of the average azimuthal angle $\langle\psi\rangle$ of the vector $\mathbf{M}(t)$ in stationary precessional motion near one of the three distinguished directions on the orientation ψ_h of a microwave field with amplitude $h = 0.5$ Oe (dashed curve) and 1.5 Oe (solid curve). The establishment of a magnetization precession regime with microwave field orientation close to the boundary of the sectors is shown in Fig. 3b (the curves *1*–*3*) for the angles $\psi_h = 85°, 87°, 88°$ ($h = 0.5$ Oe) and $\psi_h = 99°, 100°, 101°$ ($h = 1.5$ Oe). At first, a hop by 240° occurs at the boundary between the sectors, and precession in the position 3 replaces precession in the position 1. As the angle ψ_h increases further, a transition occurs to the missed position 2, where the dynamical regime is secured up to the next change in the angular sectors. When the amplitudes of the microwave field become sufficiently large (($h \gtrsim 1$ Oe), a hop by 240° occurs and the magnetization goes around the position 2 along the open trajectory and is drawn by the high-frequency fields to the next position 3. As the angle ψ_h increases, the envelope of the trajectory becomes spiral-shaped and magnetization precession in the position 2 is established. In weak fields ($h \lesssim$ Oe) a hop by 240° occurs without the magnetization going around the position 2. The precession regimes described above are due to the character of the arrangement of the four minima of the free energy $F(\theta, \psi)$ for weak static fields ($H > H_0$).

As the static field H increases further, the three minima of the function $F(\theta, \psi)$ which lie at an angle with respect to the normal vanish and precession occurs only with the axis oriented in the direction of the normal to the sample. Figure 4 shows for various values of the static field H and the high-frequency field $h = 1.5$ Oe the projections of the stationary trajectories of the precession of the magnetic moment on the yz plane. The orientation ψ_h and the initial phase of the microwave field in the cases considered do not affect the established precessional motion of the magnetization. In the dynamical regime arising immediately after the regime with four possible axes of precessional motion the magnetization goes around the three distinguished directions described above and resonance precession with maximum average amplitude $\langle\phi\rangle \approx 30°$ (curve *1*) is established as a result. The form of the trajectory of the vector **M** shows that for the prescribed ferromagnetic resonance geometry the third harmonic of the resonance frequency ω_r predominates in the nutation motion of the magnetic moment, and its contribution to the motion of the magnetization is greatest precisely in the dynamical regime characterized by the maximum precession angle $\langle\phi\rangle$. For static fields in the range 283.4 Oe $\lesssim H \lesssim$ 284.5 Oe bistability occurs, for which a second stationary orbit with a small amplitude $\langle\phi\rangle \approx 10°$ (curve *2*) is present together with the above-described stationary orbit of precession motion of the magnetization (curve *1*). Fluctuations of the parameters of the microwave field, the initial phase of the field, and the deviation of the initial orientation of the magnetization vector away from the direction along the normal influence the realization of one of the two dynamical states. For static fields in the range 284.5 Oe $< H <$ 285 Oe a small stationary orbit of precessional motion becomes the only orbit. Bifurcations leading to regimes with a complicated trajectory with period $T = 2\pi l/\omega_r$ arise in

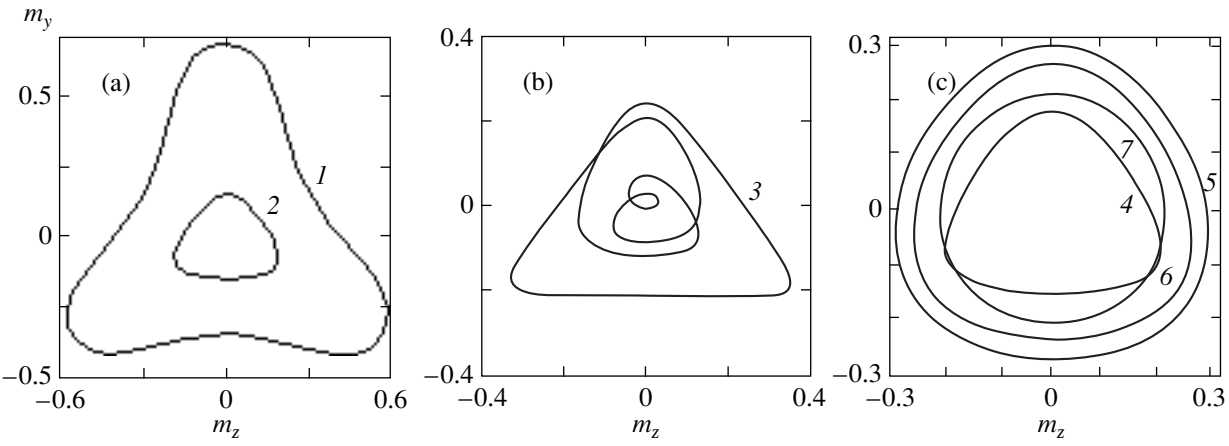

Fig. 4. Projections, on the *yz* plane, of the stationary precession trajectories of the magnetic moment for the static field H = (*1*) 283, (*2*) 284, (*3*) 285, (*4*) 290, (*5*) 350, (*6*) 400, (*7*) 600 Oe.

very narrow ranges of the fields H; specifically, $l = 4$ for trajectory 3. As H increases further (Fig. 3c) and for a fixed amplitude H of the microwave field the precession amplitude at first increases very little (curves 4 and 5) and then decreases (curves 6 and 7), and the form of the trajectory approaches a circle, i.e., the contribution of the higher harmonics of the resonance frequency to the nutation motion of the magnetization decreases.

For a quantitative estimate of the contribution of various harmonics of the resonance frequency to the nutation motion of the magnetization, we shall represent the time-dependent precession angle $\phi(t)$ in the form of the series

$$\phi(t) = \sum \phi_{n\omega} e^{i\omega n t}. \qquad (5)$$

The contributions of the first three harmonics for several values of the static field H are presented in table. It is evident that under ferromagnetic resonance conditions in a (111) type sample the third harmonic makes the largest contribution among the higher order harmonics ($n > 0$). For sufficiently strong static fields, $H \gg H_0$, and small precession amplitudes the angle $\phi(t)$ is determined, to a high degree of accuracy, by the constant angle $\phi_0 \equiv \langle\phi\rangle$ and the higher order harmonics can be neglected in the description of the motion of the magnetization. The average precession angle ϕ_0 in this case can be approximately found from the expression

$$\cos\phi_0 = 1 - \frac{h^2}{4M^2}(|\chi|^2 + |\chi_a|^2), \qquad (6)$$

where the complex diagonal $\chi = \chi' - i\chi''$ and off-diagonal $\chi_a = \chi_a' - i\chi_a''$ components of the high-frequency susceptibility tensor, determining the linear relation between the high-frequency fields and the magnetization, are, under resonance conditions ($\omega = \omega_r$),

$$\chi' = \frac{\gamma M(\lambda^2 - \gamma^2 M^2)}{\omega\Delta}, \quad \chi'' = \frac{\gamma^2 M^2(3\lambda^2 + 2\gamma^2 M^2)}{\omega\lambda\Delta},$$

$$\chi_a' = \frac{\gamma^3 M^3}{\omega\Delta}, \quad \chi_a'' = \frac{2\gamma^4 M^4}{\omega\lambda\Delta},$$

where $\Delta = \lambda^2 + 4\gamma^2 M^2$ [10]. Thus, the expression (6) for the precession amplitude with H = 500, 600, and 700 Oe gives, respectively, ϕ_0 = 17.510, 12.431, and 9.639°, which shows the high degree of accuracy of the results obtained from Eq. (6) for strong static fields.

The above analysis of resonance precessional motion of a magnetic moment in a (111) type iron-garnet film showed that there exist several nonlinear

Table

H, Oe	ϕ_0, deg	ϕ_ω/ϕ_0	$\phi_{2\omega}/\phi_0$	$\phi_{3\omega}/\phi_0$
283	30.108	2.50×10^{-2}	1.38×10^{-2}	3.62×10^{-1}
300	11.463	3.20×10^{-2}	4.42×10^{-2}	1.13×10^{-1}
400	16.904	1.98×10^{-3}	1.25×10^{-2}	5.03×10^{-2}
500	15.721	1.09×10^{-3}	6.29×10^{-3}	2.89×10^{-2}
600	12.228	7.24×10^{-4}	4.94×10^{-3}	1.65×10^{-2}
700	9.629	4.13×10^{-4}	4.94×10^{-3}	1.02×10^{-2}

regimes determined by the parameters of the magnetizing and high-frequency fields and by the crystallographic anisotropy field in the sample. It was shown that for sufficiently low resonance frequencies and weak magnetizing fields the precession axis is aligned in the direction of the normal only for small amplitudes of the microwave field; in the opposite case the magnetization tilts toward one of the three directions around which precessional motion with average amplitude $\langle\phi\rangle \lesssim 3^\circ$ is established. As the magnetization field and, correspondingly, the ferromagnetic resonance frequency increase, a precession regime, characterized by the maximum amplitude $\langle\phi\rangle \approx 30^\circ$ and a large contribution of the third harmonic of the fundamental precession frequency ω_r to the nutation motion of the magnetic moment, is realized around the normal to the field. In higher static fields there exists a range where dynamical bistability occurs; i.e., two stationary orbits of the precessing magnetic moment that differ strongly in amplitude ($\langle\phi\rangle_1 \approx 30^\circ$ and $\langle\phi\rangle_2 \approx 10^\circ$) exist. Bifurcations resulting in regimes with a complicated trajectory, whose period is a multiple of the period of the microwave field ($T = 2\pi l/\omega_r$), were observed.

REFERENCES

1. V. S. L'vov, *Nonlinear Spin Waves* (Nauka, Moscow, 1987).
2. P. Bryant, C. Jeffries, and K. Nakamura, Phys. Rev. A **38**, 4223 (1988).
3. S. M. Rezende and F. M. de Aguiar, Proc. IEEE **78**, 5 (1990).
4. B. Neite and H. Dötsch, J. Appl. Phys. **62**, 648 (1987).
5. G. S. Kandaurova and A. E. Sviderskiĭ, Pis'ma Zh. Éksp. Teor. Fiz. **47**, 410 (1988) [JETP Lett. **47**, 490 (1988)].
6. F. V. Lisovskiĭ, E. G. Mansvetova, E. P. Nikolaeva, and A. V. Nikolaev, Zh. Éksp. Teor. Fiz. **103**, 213 (1993) [JETP **76**, 116 (1993)].
7. A. M. Prokhorov, G. A. Smolenskiĭ, and A. N. Ageev, Usp. Fiz. Nauk **143**, 33 (1984) [Sov. Phys. Usp. **27**, 339 (1984)].
8. A. K. Zvezdin and V. A. Kotov, *Magneto-Optics of Thin Films* (Nauka, Moscow, 1988).
9. D. I. Sementsov and A. M. Shutyĭ, Opt. Spektrosk. **84**, 280 (1998) [Opt. Spectrosc. **84**, 238 (1998)].
10. A. G. Gurevich and G. A. Melkov, *Magnetic Oscillations and Waves* (Nauka, Moscow, 1994).
11. B. Neite and H. Dötsch, Proc. SPIE **1018**, 115 (1989).

Translation was provided by AIP

Journal of Experimental and Theoretical Physics, Vol. 91, No. 3, 2000, pp. 537–540.
Translated from Zhurnal Éksperimental'noĭ i Teoreticheskoĭ Fiziki, Vol. 118, No. 3, 2000, pp. 617–621.

SOLIDS
Structure

Electric and Magnetic Properties of the Perovskites $Ba_2(FeMo)O_x$ ($5.88 \le x \le 6.01$)

L. S. Lobanovskiĭ[a], I. O. Troyanchuk[a, *], and H. Szymczak[b]

[a]*Institute of Solid-State and Semiconductor Physics, Belarussian Academy of Sciences, Minsk,220072 Belarus*

[b]*Institute of Physics, Polish Academy of Sciences, 02-668 Warsaw, Poland*

**e-mail: troyan@ifttp.bas-net.by*

Received March 28, 2000

Abstract—The purpose of this work is to establish the relation between the magnetic, electric, and magnetotransport properties and the oxygen nonstoichiometry of the compounds $Ba_2(FeMo)O_x$ ($5.88 \le x \le 6.01$). The investigations established the behavior of the magnetization, resistance, and magnetoresistance of samples in this series. It is shown that the behavior of the magnetization can be described by assuming that the iron ions become divalent ($Fe^{3+} \longrightarrow Fe^{2+}$) as a result of the reduction of the samples and the molybdenum ions become hexavalent ($Mo^{5+} \longrightarrow Mo^{6+}$) as a result the oxidation of the samples. It is established that there are two contributions to the magnetoresistance which arise as result of magnetic ordering of the intragranular interlayer and intergranular transfer of spin-polarized charge carriers. It is inferred that electric transport in samples of this series is determined by percolation processes between granules with metallic conductivity separated by a dielectric interlayers. © *2000 MAIK "Nauka/Interperiodica".*

1. INTRODUCTION

The discovery of "colossal magnetoresistance" stimulated the search for new materials with a large magnetoresistance effect in weak magnetic fields. Such materials are required for information storage and processing devices. In manganites a large magnetoresistance effect is, as a rule, attained in comparatively large magnetic fields ($H \approx 5$ kOe). Recently there have appeared reports that perovskites of the type $Sr_2(FeMo)O_6$ manifest a large magnetoresistance effect in weak magnetic fields at room temperature [1–3]. This effect was explained by intergranular transfer of spin-polarized charge carriers. The results obtained on epitaxial thin films have confirmed this conclusion [4]. It has turned out that the magnetoresistance effect in epitaxial films is positive, while polycrystals manifest a negative effect, i.e., the electric resistance of samples decreases in a magnetic field. A magnetoresistance effect attributable to intergranular tunneling of spin-polarized electrons is also observed in manganites. Ordinarily, an effect of this type increases monotonically with decreasing temperature, while the intragranular magnetoresistance effect is strongest near phase-transition temperatures. The nature of the intergranular magnetoresistance effect is a subject of debate [5–8]. Consequently, further investigations are required.

The objective of the present work is to establish the relationship between the magnetic, electric, and magnetotransport properties and the oxygen nonstoichiometry of the compounds $Ba_2(FeMo)O_x$ ($5.88 \le x \le 6.01$).

2. EXPERIMENTAL PROCEDURE

The $Ba_2(FeMo)O_x$ sample was prepared from oxides and carbonates of the corresponding elements using the standard ceramic technology in an inert-gas atmosphere. The materials $BaCO_3$, Fe_2O_3, Mo_2O_3, and MoO_2 were taken in stoichiometric proportions and ground in an agate mortar. Next, the powder was compressed under a pressure of 6 kbar into tablets, which were synthesized in an argon flow at 1000°C for 3 h and then cooled at a rate of 100°C/h. X-Ray diffraction data, obtained in Co K_α radiation, for a sample after synthesis showed the presence of one phase with perovskite structure. The oxygen content in the composition obtained by this method was determined according to the mass loss after reduction in a hydrogen flow to barium oxide and metallic iron and molybdenum.

The samples were reduced at temperature 900°C in evacuated quartz ampullae in the presence of pulverized metallic tantalum, which served as an oxygen getter. To oxidize the samples, pulverized $LaMnO_{3.1}$ was placed into the ampulla. The oxygen content was determined according to the mass loss of the sample after reduction or according to the mass increase after oxidation. After each treatment x-ray diffraction analysis was performed on the samples to determine the unit-cell parameters and to monitor the single-phase composition of the samples.

The magnetic susceptibility was measured by the bridge method and the resistance was measured by the standard four-contact method. The magnetization was determined with a vibrating coil magnetometer.

1063-7761/00/9103-0537$20.00

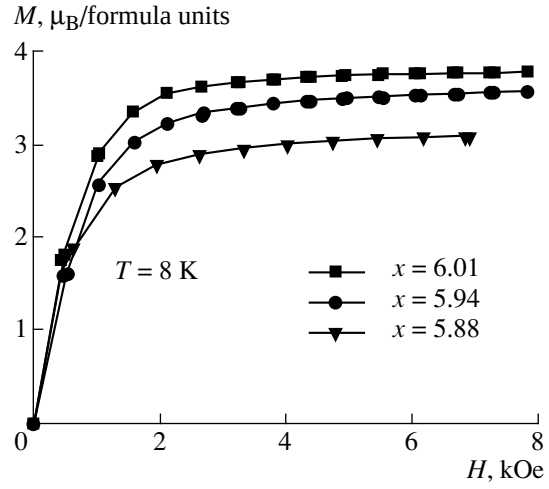

Fig. 1. Magnetization versus the external magnetic field for $Ba_2(FeMo)O_x$ ($5.88 \le x \le 6.01$) at 8 K.

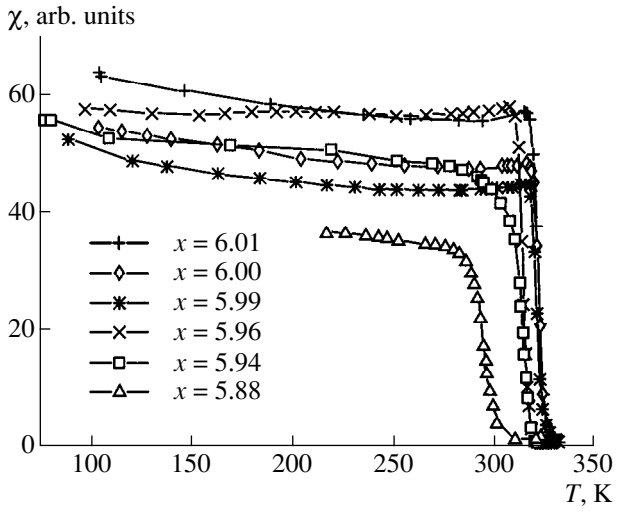

Fig. 2. Dynamic magnetic susceptibility versus temperature for $Ba_2(FeMo)O_x$ ($5.88 \le x \le 6.01$).

3. RESULTS AND DISCUSSION

The chemical formula for the sample after synthesis was determined to be $Ba_2(FeMo)O_{5.99}$. According to the x-ray diffraction data all samples were characterized by cubic symmetry of the unit cell. The x-ray diffraction patterns contained superstructural peaks due to ordering of the iron and molybdenum ions in a NaCl structure. The cell parameter a of the $Ba_2(FeMo)O_{5.99}$ sample was found to be 8.068 Å. Oxidation decreased the volume of the cubic unit cell. The parameter a for the $Ba_2(FeMo)O_{6.01}$ sample was 8.066 Å. The unit-cell volume was observed to increase with oxidation of the composition $Ba_2(FeMo)O_{5.99}$. The parameter a for the composition $Ba_2(FeMo)O_{5.88}$ was 8.086 Å.

Magnetization measurements (Fig. 1) showed that the $Ba_2(FeMo)O_{5.99}$ sample possesses a magnetic moment 3.5 Bohr magnetons per formula unit (μ_B/formula unit) at 8 K. The Curie temperature T_C, which for this sample is 324 K, was determined from the temperature dependence of the dynamic susceptibility (Fig. 2). When the $Ba_2(FeMo)O_{5.99}$ sample was reduced the magnetic moment and the Curie temperature decreased. For the composition $Ba_2(FeMo)O_{5.88}$ the magnetic moment was approximately $2.7\mu_B$/formula unit, and T_C decreased to 318 K. When the sample $Ba_2(FeMo)O_{5.99}$ was oxidized the magnetization increased but no appreciable changes were observed in T_C. The magnetic moment of the composition $Ba_2(FeMo)O_{6.01}$ was found to be $3.6\mu_B$/formula unit, while the Curie temperature remained unchanged—324 K.

The resistivity for the sample $Ba_2(FeMo)O_{5.99}$ at liquid nitrogen temperature was of the order of $10^{-3}\ \Omega$ cm (Fig. 3a). For all samples except $Ba_2(FeMo)O_{5.96}$ the resistivity decreased after reduction, and for the $Ba_2(FeMo)O_{5.88}$ sample the resistivity was of the order of $10^{-4}\ \Omega$ cm. The resistivity of the composition $Ba_2(FeMo)O_{5.96}$ was essentially identical to that of $Ba_2(FeMo)O_{5.99}$. For all reduced samples the resistivity increased negligibly with increasing temperature in the entire experimental temperature range. A kink was observed in the temperature dependence of the resistivity near the Curie temperature.

A strong increase in the resistivity was observed when the sample $Ba_2(FeMo)O_{5.99}$ was oxidized. In this case, the resistance increased by 4 orders of magnitude with a negligible change in the oxygen content from 5.99 to 6.01. The temperature dependence of the resistivity also changed. Up to temperature 325 K for $Ba_2(FeMo)O_{6.00}$ and 254 K for $Ba_2(FeMo)O_{6.01}$ the resistivity of the samples increased with temperature. However, under further heating the resistivity of the samples decreased.

The magnetoresistance was calculated as $MR = \{[\rho(H = 0) - \rho(H = 9\ \text{kOe})]/\rho(H = 0)\} \times 100\%$. The temperature dependence of the magnetoresistance is shown in Fig. 3b. For all samples except $Ba_2(FeMo)O_{5.96}$ the magnetoresistance at liquid-nitrogen temperature was 20–25% and decreased with increasing temperature to approximately 1%. Then, a peak was observed in the magnetoresistance, reaching 3–4%, at the Curie temperature. The composition $Ba_2(FeMo)O_{5.96}$ demonstrate a decrease of the magnetoresistance to 4% at 77 K, while the magnitude of the peak MR at T_C remained in the range 3–4%. With further reduction of this sample the magnetoresistance once again reached 20–25% at liquid-nitrogen temperature. An interesting feature in the behavior of the magnetoresistance can be seen by analyzing its temperature dependence. The magnetoresistance of the more highly oxidized sample $Ba_2(FeMo)O_{6.01}$ remained constant near 20% in the range from 77 K right up to

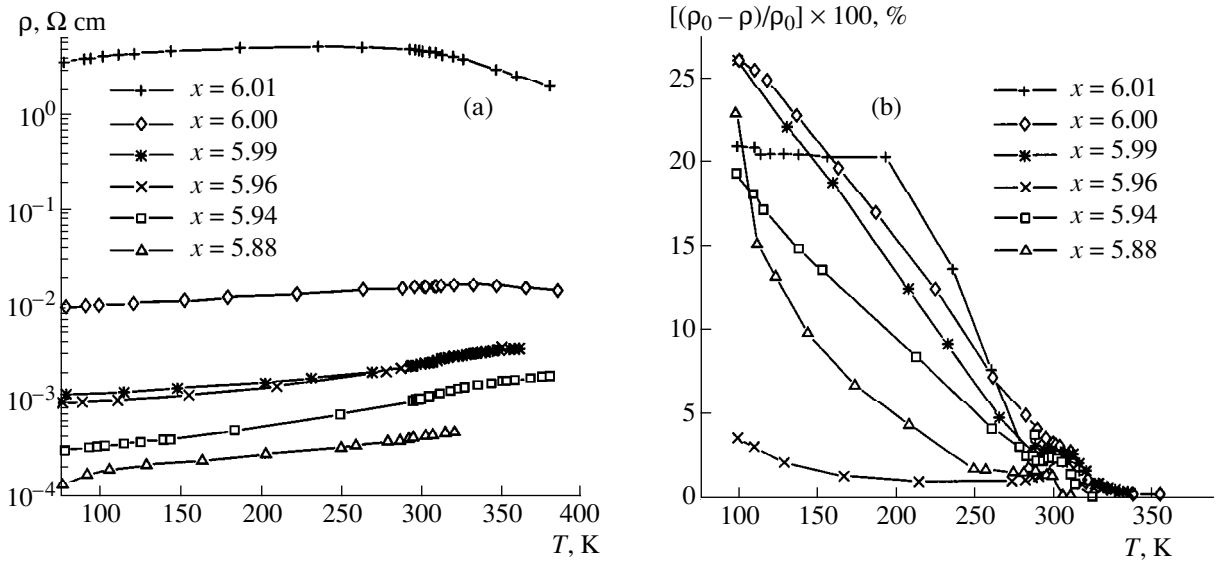

Fig. 3. (a) Electric conductivity and (b) magnetoresistance versus temperature for $Ba_2(FeMo)O_x$ ($5.88 \le x \le 6.01$).

185 K. However, as the oxygen content in the samples decreases, the low-temperature magnetoresistance effect decreases with increasing temperature more rapidly, even though the difference in the Curie temperature is very small.

Investigations of the field dependences of the magnetoresistance effect at liquid-nitrogen temperature established that the magnetoresistance saturates in fields up to 3 kOe (Fig. 4). Magnetization saturation was observed in approximately the same magnetic field.

Comparing the behavior of the compound $Ba_2(FeMo)O_x$ ($5.88 \le x \le 6.01$) with that of manganites with perovskite structure shows that the compositions $Ba_2(FeMo)O_x$ possess a smaller region of oxygen nonstoichiometry. When our experimental samples were reduced they retained a single-phase composition up to the oxygen parameter $x = 5.88$, while manganites retained a single-phase composition when the index x decreased from 3.0 to 2.66 [9]. Therefore, the $Ba_2(FeMo)O_x$ ($5.88 \le x \le 6.01$) retains a single-phase composition with variation of the oxygen content in only a 2% range, while for manganites with perovksite structure this range is approximately 11%.

Nuclear magnetic resonance investigations of $Ba_2(FeMo)O_6$ showed that the iron ions are in a trivalent state and the molybdenum ions are in a pentavalent state [10].

We assume that two opposing factors influence the change in the cell volume in our experimental compositions: when the compositions were reduced oxygen atoms were removed and iron ions became divalent. In addition, when one oxygen atom is removed two iron ions must become divalent in order for the sample to remain electrically neutral. The first process decreases the cell parameter, while the second process should increase the cell parameter, since the radius of the ion Fe^{2+} is much larger than that of Fe^{3+}. According to the experimental data, the cell volume increases on reduction. This shows that a decrease of the average oxidation state of the iron ions has a somewhat stronger effect on the change in the cell parameter.

It can be inferred that the oxidation of $Ba_2(FeMo)O_6$ is analogous to the oxidation of $LaMnO_3$ type perovskites. Then, when the samples are oxidized the valence state of the molybdenum ions changes ($Mo^{5+} \longrightarrow Mo^{6+}$) and cationic vacancies appear. Both processes decrease the cell volume. As mentioned earlier, for our samples the cell parameter decreased on oxidation.

The spontaneous magnetic moment for the stoichiometric composition $Ba_2(FeMo)O_6$ is determined by the antiparallel ordering of the magnetic moments of the iron ions $Fe^{3+}(3d^5)$ and the molybdenum ions $Mo^{5+}(4d^1)$. In

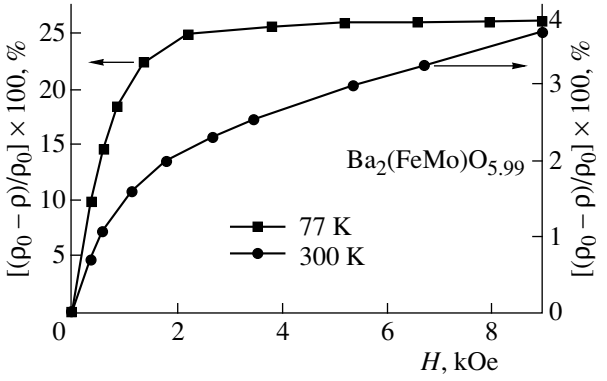

Fig. 4. Field dependences of the magnetoresistance of a $Ba_2(FeMo)O_{5.99}$ sample at various temperatures.

this case the spontaneous magnetic moment of a stoichiometric compound is $4\mu_B$/formula unit at 0 K. This value is somewhat higher than the value $3.6\mu_B$/formula unit at 8 K, which we obtained for the composition $Ba_2(FeMo)O_{6.01}$. Apparently, the strong hybridization of the orbitals of the Fe and Mo ions with the 2*p* orbitals of the oxygen ions decreases the effective magnetic moment of these ions.

Reduction of the sample $Ba_2(FeMo)O_{5.99}$ decreases the spontaneous magnetic moment per formula unit as result of a change in the electronic configuration of some iron ions from $Fe^{3+}(3d^5)$ to $Fe^{2+}(3d^6)$. The magnetic moment of $Fe^{2+}(3d^6)$ ions is smaller than that of $Fe^{3+}(3d^5)$ ions. Hexavalent molybdenum ions are diamagnetic. Consequently, oxidation of $Ba_2(FeMo)O_6$ should increase the spontaneous magnetic moment. The decrease in the Curie temperature on reduction of the samples is probably due to the fact that the exchange interaction Fe^{2+}–O–Mo^{5+} is weaker than the exchange interaction Fe^{3+}–O–Mo^{5+}.

Reduction increased the electric conductivity of the samples (Fig. 3a) despite the appearance of oxygen vacancies, which are structural defects and should limit the mobility of charge carriers. This tendency could be due to the disruption of the ordering of the Fe^{3+} and Mo^{5+} ions. In strongly reduced samples we observed a large decrease of the intensity of superstructural lines, which are due to ordering of iron and molybdenum ions. It is well-known that perovskites containing only Fe^{3+} ions are good dielectrics. When the oxygen content exceeded the value corresponding to the stoichiometric composition the resistance increased sharply. The form of the temperature dependence of the resistivity (Fig. 3a) can be interpreted assuming that the conductivity of strongly oxidized samples is due to percolation processes. Apparently, the samples consist of a main metallic phase and dielectric interlayers. At some locations the dielectric interlayers are weakened, which results in percolation conductivity along the metallic phase. When the samples are oxidized the surface composition of the granules probably changes radically first. This is confirmed indirectly by the fact that strongly oxidized samples became brittle.

Investigations of the magnetotransport properties shows that strongly oxidize samples are most promising for practical applications, since the magnetoresistance effect in such samples remains large at comparatively high temperatures. This tendency is probably due partially to the percolation character of the electrical conductivity of these samples. The magnetoresistance effect and the magnetic properties are correlated with one another. The more the exchange interactions in $Ba_2(FeMo)O_x$ are weakened, the more rapidly the magnetoresistance effect decreases with increasing temperature. The maximum of the magnetoresistance effect near the Curie temperature is probably due to the same processes as in metallic $Tl_2Mn_2O_7$ with pyrochlore structure [11]. In $Tl_2Mn_2O_7$ the conductivity type likewise remains unchanged at the point T_C, but the magnitude of the magnetoresistance effect is several times larger [11]. This could be due to the fact that $Ba_2(FeMo)O_6$ is a ferrimagnet, while $Tl_2Mn_2O_7$ is characterized by parallel ordering of the magnetic moments of all manganese ions. It is well-known that ferrimagnetic spinels of the type $MnFe_2O_4$ also manifest a maximum magnetoresistance effect near the Curie temperature, though its value is very small even compared with the analogous quantity for $Ba_2(FeMo)O_6$ type ferrimagnets.

The reasons for the sharp decrease of the magnetoresistance effect in the compositions $Ba_2(FeMo)O_6$ are not known. A complex of spectrometric investigations is now being conducted in order to determine the nature of the unusual behavior of this compound.

ACKNOWLEDGMENTS

This work was partially supported by the Belarussian Foundation for Basic Research (project no. F99M-033).

REFERENCES

1. K.-I. Kobayashi, T. Kimura, H Sawada, *et al.*, Nature **395**, 677 (1998).
2. T. H. Kim, M. Uehara, S.-W. Cheong, *et al.*, Appl. Phys. Lett. **74**, 1737 (1999).
3. K.-I. Kobayashi, T. Kimura, Y. Tomioka, *et al.*, Phys. Rev. B **59**, 11159 (1999).
4. H. Q. Yin, J.-S. Zhou, J.-P. Zhou, *et al.*, Appl. Phys. Lett. **75**, 2812 (1999).
5. H. Asano, S. B. Ogale, J. Garrison, *et al.*, Appl. Phys. Lett. **74**, 3696 (1999).
6. P. Raychaudhuri, K. Sheshadri, P. Taneja, *et al.*, Phys. Rev. B **59**, 13919 (1999).
7. M. Ziese, Phys. Rev. B **60**, R738 (1999).
8. S. Lee, H. Y. Hwang, B. I. Shraiman, *et al.*, Phys. Rev. Lett. **82**, 4508 (1999).
9. J. Briatico, B. Alascio, R. Allub, *et al.*, Phys. Rev. B **53**, 14 020 (1996).
10. H. Yokoyama and T. Nakagava, J. Phys. Soc. Jpn. **28**, 1197 (1970).
11. M. A. Subramanian, B. H. Toby, A. P. Ramírez, *et al.*, Science **273**, 81 (1996).

Translation was provided by AIP

Journal of Experimental and Theoretical Physics, Vol. 91, No. 3, 2000, pp. 541–546.
Translated from Zhurnal Éksperimental'noĭ i Teoreticheskoĭ Fiziki, Vol. 118, No. 3, 2000, pp. 622–628.

SOLIDS
Electronic Properties

Tunneling into a Size-Quantized Film

A. I. Khachaturov

Donetsk Physicotechnical Institute, Ukrainian National Academy of Sciences, Donetsk, 340114 Ukraine
e-mail: khach@sts.dipt.donetsk.ua

Received November 17, 1999

Abstract—A theoretical model making it possible to calculate the characteristics of metal–insulator–size-quantized film tunneling junctions in a wide range of voltages is proposed. The conditions for observing a geometric resonance in the differential tunneling conductivity are modeled, and the influence of temperature on the resonance oscillatory structure is investigated. It is shown that the geometric resonance is not the only possibility for manifestation of standing waves in real nonuniform films. For one polarity of the voltage resistance peaks which are stable with respect to temperature smearing can appear. Moreover, quantization of the spectrum as a whole changes the behavior of the curve $\sigma(V)$, shifting its minimum by a finite amount relative to zero voltage. It is inferred that this effect, which does not require any special conditions in order to appear, can serve as an indication of the presence of standing waves in one of the electrodes. *© 2000 MAIK "Nauka/Interperiodica".*

1. INTRODUCTION

The observation of size-quantization of electron waves in thin metallic films is one of the most remarkable achievements of electron tunneling spectroscopy [1, 2]. In the course of these investigations, the idea that effects of this kind are impossible in the characteristics of tunnel junctions changed. The problem was that even negligible changes in thickness (of the order of several angstroms, which are unavoidable in real films) should change the position of the energy levels by an amount comparable to the splitting between the levels and in consequence the effect under study should unavoidably be washed out. However, as shown in [1], because the thickness d of a nonuniform polycrystalline film can vary only discretely, $d = Na$ (a is the lattice constant and N is an integer), so-called commensurate levels, whose energies do not depend on the thickness, exist in thin films. Subsequent detailed experimental investigations made it possible not only to show, very reliably, the existence of commensurate levels in various materials (Au, Mg, Pb, Bi) but also to determine, on the basis of the data obtained, a variety of electronic characteristics of an experimental film [2–4]. For example, the position of certain singular points, located far from the Fermi level, was found in [4], and the slope of the dispersion curves near these points was also found. Size-quantization under conditions of high hydrostatic pressures (up to 10 kbar) was investigated in [5, 6]. Thus, by the beginning of the 1980s the effect under discussion was already a well-studied phenomenon and was described in detail in a number of monographs [7–9] as a promising method for investigating the electronic characteristics of solids. On the other hand, despite such a much-promising beginning, it must now be stated that the hopes were by no means fully justified. In our view, the effect has not found the expected applications in studying new materials, such as, for example, metal-oxide compounds. We believe that the main reason for this situation is that all preceding investigations were limited to the use of simple theoretical constructions, capable of predicting primarily only the locations of the experimental features on the voltage scale. In the present paper we present a theoretical model, which, in our view, gives the most complete possible description of the experimental situation. It makes it possible to calculate in a wide range of voltages (of the order of several volts) the characteristics of a tunneling contact, taking account of all two-dimensional bands of the quantized electrode that lie below the Fermi level and all bands of any importance above the Fermi level.

2. FORMULATION OF THE MODEL

Let us consider a tunneling contact in which one electrode is an ordinary Fermi metal. For simplicity, we assume that the second electrode is made of the same material, but because its thickness L is small (several hundreds of angstroms) size-quantization occurs in it, i.e., the wave number k_z in a direction perpendicular to the plane of the barrier and therefore the kinetic energy component E_z in this direction can assume only discrete values [7]:

$$E_{nz} = \frac{(\hbar k_{nz})^2}{2mL^2} = \frac{(\pi\hbar)^2}{2mL^2}n^2, \quad n = 1, 2, 3, \ldots \tag{1}$$

(the generally accepted notation is used in this expression). The electronic spectrum of the size-quantized electrode is quasicontinuous and consists of two-

1063-7761/00/9103-0541$20.00

dimensional bands—planes perpendicular in reciprocal space to the axis k_z:

$$\varepsilon_n(k_\parallel, E_{nz}) = E_{nz} + \frac{(\pi k_\parallel)^2}{2m}, \tag{2}$$

where $k_\parallel$ is the component of the wave vector parallel to the tunneling plane and varies continuously. We shall assume that the potential barrier separating the electrodes in the tunneling structure is a trapezoid with heights ϕ_1 and ϕ_2 and thickness d. When a bias voltage V is applied to one of the electrodes the form of the barrier changes as

$$\phi(z, V) = \phi_1 + (\phi_2 - eV - \phi_1)z/d. \tag{3}$$

Let us apply a bias voltage V to the size-quantized film and calculate the contribution $J_n(V)$ of the nth band to the total tunneling current $J(V)$. We note that the transmittance of the tunneling barrier depends only on the perpendicular component of the kinetic energy E_z and does not depend on the parallel component $E_\parallel$, so that all electrons in the nth band which have the same perpendicular energy component E_{nz} possess the same tunneling probability $P(E_{nz}, V)$. For a trapezoidal barrier [10]

$$P_1(E_{nz}, V) = \exp\left\{-\frac{A_d}{\phi_2 - eV - \phi_1} \times \left[\sqrt{(\phi_2 - eV - E_{nz})^3} - \sqrt{(\phi_1 - E_{nz})^3}\right]\right\}, \tag{4}$$

where $A_d = 4\sqrt{2m}\, d/3\hbar$ (the barrier height ϕ and the energy E_{nz} in this formula are measured from the conduction band bottom). We shall assume that tunneling is elastic and specular, i.e., the component $k_\parallel$ of the wave vector parallel to the barrier and the total energy E of the tunneling electron are conserved. We note that under these conditions the electrons lying in the reciprocal space of the initial electrode in the same plane k_{nz} = const will lie in the reciprocal space of the opposite electrode also in the same plane k'_{nz} = const (when the Fermi energies are the same $E_{F_1} = E_{F_2}$, their energies are related by the relation $E'_{nz} = E_{nz} + eV$). We note that although the spectrum of the opposite electrode is continuous, only the electrons lying in the plane E'_{nz} can participate in the tunneling process and contribute to the reverse tunneling current, since allowed states in the plane E_{nz} exist only for them in the first electrode. Thus, even though the spectrum of the opposite electrode is continuous on the whole, the spectrum of the tunneling electrons consists of two-dimensional bands. In contrast to "real" stationary bands of the quantized electrode, when a voltage is applied these "imaginary" bands will move along the k'_z-axis, satisfying the relation $k'_{nz} = \sqrt{2m(E_{nz} + eV)}/\hbar$. The number of electrons in the nth band is

$$N = 2\int_{-\infty}^{\infty}\int f(E_{nz}, E_\parallel, T)\frac{dk_x dk_y}{(2\pi)^2}, \tag{5}$$

where $f(E_{nz}, E_\parallel, T)$ is the Fermi distribution function (the factor of 2 takes account of the fact that each state is doubly spin-degenerate). Since all electrons belonging to the same band have the same group velocity v_{nz}, we find the contribution of the nth band to the current incident on the plane of the junction as

$$J_n^{(i)}(V) = ev_{nz}N = 4\pi e v_{nz}\int_0^\infty f(E_{nz}, E_\parallel, T)\frac{k_\parallel dk_\parallel}{(2\pi)^2}. \tag{6}$$

Multiplying this expression by the barrier transmittance $P(E_{nz}, V)$ and by the probability that the final state in the opposite electrode is not occupied, we find the contribution of the nth band to the direct tunneling current:

$$J_{1n}(V) = 4\pi e v_{nz}P_1(E_{nz}, V) \times \int_0^\infty f(E_{nz}, E_\parallel, T)[1 - f(E_{nz} + eV, E_\parallel, T)]\frac{k_\parallel dk_\parallel}{(2\pi)^2}. \tag{7}$$

Similar arguments for the reverse current lead to the expression

$$J'_{1n}(V) = 4\pi e v'_{nz}P_1(E_{nz}, V) \times \int_0^\infty f(E_{nz} + eV, E_\parallel, T)[1 - f(E_{nz}, E_\parallel, T)]\frac{k_\parallel dk_\parallel}{(2\pi)^2}. \tag{8}$$

For simplicity we shall assume that the group velocities

$$v = \sqrt{\frac{2E_{nz}}{m}}, \quad v' = \sqrt{\frac{2(E_{nz} + eV)}{m}}$$

are approximately the same (this assumption is obvious for small bias voltages, and for large eV, since the reverse current decreases rapidly with increasing voltage, it does not lead to a noticeable error). The tunneling current J_n from the nth band is the difference between the direct and reverse currents:

$$J_n(V) = J_{1n}(V) - J'_{1n}(V) = \sqrt{2m}\frac{e}{\pi\hbar^2}\sqrt{E_{nz}}P_1(E_{nz}, V) \times \int_0^\infty [f(E_{nz} + eV, E_\parallel, T) - f(E_{nz}, E_\parallel, T)]dE_\parallel \tag{9}$$

$$= C\sqrt{E_{nz}}kT\ln\left\{\frac{1+\exp\left[-\dfrac{E_{nz}-E_F}{kT}\right]}{1+\exp\left[-\dfrac{E_{nz}-E_F+eV}{kT}\right]}\right\}P_1(E_{nz}, V),$$

where $C = \sqrt{2m}\,e/\pi\hbar^2$. Differentiating the expression obtained with respect to the voltage, we find the contribution of the nth band to the differential tunneling conductivity

$$\sigma_n(V) = \sigma_{1n}(V) + \sigma_{2n}(V), \tag{10}$$

where

$$\sigma_{1n}(V) = C\sqrt{E_{nz}}kT \times \ln\left\{\frac{1+\exp\left[-\dfrac{E_{nz}-E_F}{kT}\right]}{1+\exp\left[-\dfrac{E_{nz}-E_F+eV}{kT}\right]}\right\}\frac{\partial P_1(E_{nz}, V)}{\partial V},$$

$$\sigma_{2n}(V) = eC\sqrt{E_{nz}} \times \frac{\exp\left(-\dfrac{E_{nz}-E_F+eV}{kT}\right)}{1+\exp\left(-\dfrac{E_{nz}-E_F+eV}{kT}\right)}P_1(E_{nz}, V),$$

$$\frac{\partial P_1(E_z, V)}{\partial V} = \frac{eA_d}{2}[(\phi_2 - 3\phi_1 - eV + 2E_z) \times \sqrt{\phi_2 - eV - E_z} + 2\sqrt{(\phi_1 - E_z)^3}]\frac{P_1(E_z, V)}{(\phi_2 - eV - \phi_1)^2}.$$

Figure 1 (curve *1*, negative bias voltages) displays the dependence $\sigma_n(V)$ for the band E_{nz} lying below the Fermi level E_F. Its behavior can be understood by taking account of the fact that the first term, $\sigma_{1n}(V)$, in Eq. (10) is determined by the change in the transmittance of the tunneling barrier and can grow monotonically when a bias voltage V is applied. The increase in the first term, $\sigma_{2n}(V)$, is due to the decrease in the number of reverse electrons tunneling into the band under study. This decrease is due to the fact that, as already noted, the plane E'_{nz} on which the reverse electrons are located moves, as voltage increases, along the k'_z-axis. A region populated with the largest number of reverse electrons—the section of this plane by the Fermi sphere—decreases in size. For voltage $eV \approx E_F - E_{nz}$, which corresponds to the plane indicated moving outside the Fermi sphere, the reverse current virtually vanishes, and the term due to this current in the conductivity decreases abruptly to zero. If the plane E_{nz} lies at a distance $E_{mz} - E_F \gg kT$ from the Fermi surface, then the contribution of this band to the conductivity can be assumed to be zero since the population of the band is small (Fig. 1, curve *2*, negative voltages).

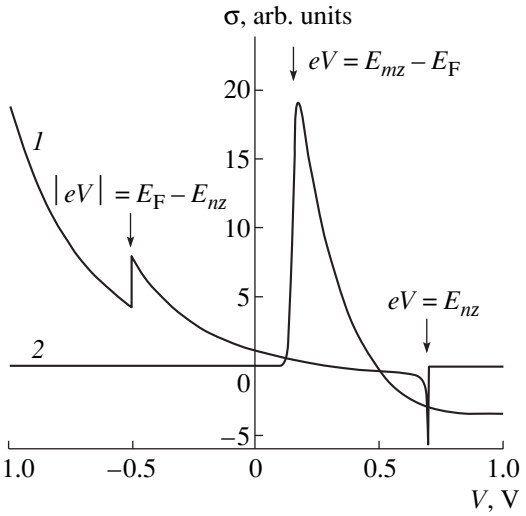

Fig. 1. Contribution of the two-dimensional band to the differential conductivity: *1*, $E_{nz} < E_F$ (E_{nz} = 0.7 meV, E_F = 1.2 meV); *2*, $E_{mz} > E_F$ (E_{nz} = 1.35 meV). Positive polarity corresponds to electron tunneling from the massive electrode into the band under study, negative polarity corresponds to tunneling from the band. The parameters of the square barrier are: d = 10 Å, $\varphi_1 = \varphi_2$ = 1 eV. The temperature T = 77 K. The computational step h = 1 meV.

When a negative voltage is applied to the massive electrode the "imaginary" bands of the massive electrode move toward the center of the Brillouin zone, $E'_{nz} = E_{nz} - eV$, while the planes from which the reverse electrons tunnel are stationary. As a result of this motion, the dependence of the barrier transmittance on the bias voltage $P_2(E_{nz},V)$ in this case differs somewhat from Eq. (4):

$$P_2(E_{nz}, V) = \exp\left\{-\frac{A_d}{\phi_1 - eV - \phi_2} \times [\sqrt{(\phi_1 - E_{nz})^3} - \sqrt{(\phi_2 - E_{nz} + eV)^3}]\right\}. \tag{11}$$

Arguments similar to those presented above yield an expression for the contribution of one band to the tunneling current:

$$J_{2n}(V) = C\sqrt{E_{nz} - eV}kT \times \ln\left\{\frac{1+\exp\left[-\dfrac{E_{nz}-E_F-eV}{kT}\right]}{1+\exp\left[-\dfrac{E_{nz}-E_F}{kT}\right]}\right\}P_2(E_{nz}, V). \tag{12}$$

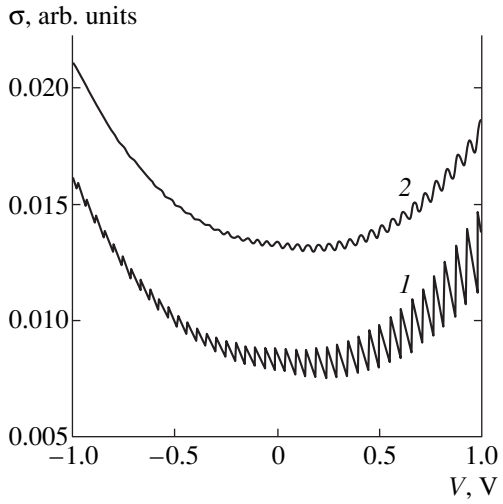

Fig. 2. Differential conductivity $\sigma(V)$ of a metal–insulator–quantized film, uniform over the thickness L = 500 Å, tunneling contact. It is assumed that the electrodes are made of a hypothetical metal with E_F = 4 eV and lattice constant a = 2 Å. The height of the square potential barrier $\varphi_1 = \varphi_2$ = 4 eV, the thickness d = 10 Å. Curve *1* corresponds to T = 4.2 K, curve *2* corresponds to T = 77 K (curve *2* is shifted upwards relative to curve *1* by 0.005 arb. units). The computational step h = 2 meV.

Differentiating Eq. (12) with respect to the voltage V we find an expression for the tunneling conductivity:

$$\sigma_n(V) = \sigma_{3n}(V) + \sigma_{4n}(V) + \sigma_{5n}(V), \qquad (13)$$

where

$$\sigma_{3n}(V) = -\frac{eC}{2\sqrt{E_{nz} - eV}}kT \times \ln\left\{\frac{1 + \exp\left[-\frac{E_{nz} - eV - E_F}{kT}\right]}{1 + \exp\left[-\frac{E_{nz} - E_F}{kT}\right]}\right\}P_2(E_{nz}, V),$$

$$\sigma_{4n}(V) = eC\sqrt{E_{nz} - eV} \times \frac{\exp\left(-\frac{E_{nz} - eV - E_F}{kT}\right)}{1 + \exp\left(-\frac{E_{nz} - eV - E_F}{kt}\right)}P_2(E_{nz}, V),$$

$$\sigma_{5n}(V) = -eCkT\sqrt{E_{nz} - eV} \times \ln\left\{\frac{1 + \exp\left[-\frac{E_{nz} - eV - E_F}{kT}\right]}{1 + \exp\left[-\frac{E_{nz} - E_F}{kT}\right]}\right\}\frac{\partial P_2(E_{nz}, V)}{\partial V}.$$

If the "imaginary" band of the massive electrode with V = 0 intersects the Fermi sphere, $E_{nz} < E_F$, then right up to voltages $eV \approx E_{nz}$ the dependence $\sigma_n(V)$ is a monotonically decreasing function (Fig. 1, curve *1*, positive polarity). This is a consequence of the fact that the indicated plane moves, when a voltage is applied, in the direction opposite to the k_z'-axis, and the barrier transmittance of electrons lying on it decreases. As the center of the Brillouin zone is approached, $eV = E_{nz}$, the differential conductivity tends towards minus infinity. It is obvious that for $eV > E_{nz}$the "imaginary" band vanishes and the contribution to the differential conductivity becomes zero. If the "imaginary" band does not cross the Fermi surface initially, $E_{mz} > E_F$, then it remains essentially empty right up to voltages $eV \approx E_{mz} - E_F$, and consequently its contribution to the conductivity is negligibly small. At the moment this band touches the Fermi surface, $eV = E_{mz} - E_F$, electrons start to occupy it, and this is reflected in the conductivity as a jump, whose sharpness depends on the temperature, after which the contribution to the conductivity, just as in the preceding case, becomes a monotonically decreasing function.

We note that Eqs. (10) and (13), though complicated, contain only elementary functions, i.e., we are dealing with a very unusual case where the calculation of the tunneling conductivity does not require any integration. This situation makes it possible to take into account the contributions from all bands lying below the Fermi level and all bands of any importance above the Fermi level, whose number n in real nonuniform metallic films can be very large:

$$\sigma(V) = \sum_n \sigma_n(V, E_{nz}). \qquad (14)$$

Figure 2 shows curves of the differential conductivity calculated for a square potential barrier $\varphi_1 = \varphi_2$ = 4 eV ($\varphi = \phi - E_F$) with thickness d = 10 Å (curve *1* corresponds to temperature T = 4.2 K, and curve *2* corresponds to T = 77 K). It is assumed that both electrodes are made of the same metal with E_F = 4 eV, but one electrode consists of a thin ideal uniform metallic film of thickness d = 500 Å, while the geometric dimensions of the second electrode make it possible to assume that its electronic spectrum is continuous. For definiteness, we assume that the positive polarity corresponds to electron tunneling from the massive electrode. As expected, features occur along the entire curve $\sigma(V)$ (curve *1*, Fig. 2). As temperature increases, these features wash out and an oscillatory structure is observed along the curve $\sigma(V)$, and the amplitude of the oscillations increases with the voltage (curve *2*, Fig. 2).

Figure 3 (curve *1*) shows a section of the differential conductivity calculated for positive polarity. A number of sharp conductivity dips, which appear at the moment the "imaginary" band approaches the center of the Bril-

louin zone, are also noticeable on this section together with a step structure. We note that this result in itself is very unusual, since it is well known that electrons located close to the Fermi surface make the overwhelming contribution to the tunneling current, but as eV approaches E_{nz} the contribution from the nth band, which is negligibly small compared with the contribution of the lower-lying bands to the tunneling current, decreases so rapidly that the corresponding term in the differential conductivity $\sigma(V)$ is larger than all other terms in Eq. (14). The fact that the contribution from the lower-lying bands is negligibly small is reflected in the fact that the corresponding dips in the conductivity are, as a result, extremely narrow in voltage even though their amplitudes are infinitely large. At first glance, the possibility of observing these features experimentally is very problematic, since any of three factors—temperature, modulation, and finite lifetime—can completely wash out the effects under discussion. However, as one can see from Fig. 3 (curve *2*), the resistance peaks are essentially unaffected by the temperature, while the structure which consists of wide steps and dominates at low temperatures completely vanishes at high temperatures. Thus, paradoxically, the temperature smearing should make it possible to observe sharp resistance peaks, since it washes out the structure which consists of wide conductivity steps and dominates at low temperatures. The finite lifetime of the quantized states, $\tau = \hbar/\Gamma$, was taken into account by replacing in Eq. (13) E_{nz} by $E_{nz} + i\Gamma$ and calculating the real part of the tunneling conductivity $\mathrm{Re}[\sigma(V)]$. The calculations showed that the resistance peaks gradually decrease as Γ increases, and they vanish completely for $\Gamma = 3$ meV (Fig. 3, curve *3*).

Figure 4 shows the computational results for the tunneling conductivity of a nonuniform film, whose thickness L varies from 470 to 530 Å around an average value $\bar{L}$, satisfying the Gaussian distribution,

$$W(L) \propto \exp\left\{-\left[\frac{\alpha(L-\bar{L})}{\bar{L}}\right]^2\right\}.$$

Following [3], it was assumed that $\alpha = 1/6$. We can see on curve 1 that, in complete agreement with the discussion presented above, the oscillator structure remained only in voltage ranges centered around the values 0.178 and 0.777 V, for which the geometric-resonance condition (1) is valid, respectively, for the values $S/Q = 2/3$ and $S/Q = 5/7$ (S/Q is an irreducible fraction) [3, 7]. As expected, the resonance structure corresponding to $S/Q = 5/7$ is much less pronounced than the structure corresponding to $S/Q = 2/3$. For other bias voltages on the tunnel junction the contributions from various values of the thickness L mutually cancel. It should be noted that the commensurate states are most strongly manifested when they lie outside the Fermi sphere. For positive voltages two resonance series appear in Fig. 4 (curve *1*); for negative voltages only one series appears.

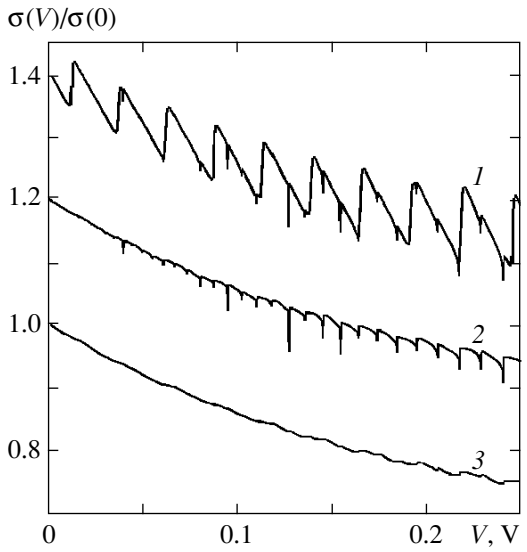

Fig. 3. A branch of the curve of the differential conductivity $\sigma(V)$, corresponding to electron tunneling from the ordinary electrode into a quantized film which is uniform over the thickness $L = 500$ Å. It is assumed that the electrodes are made of a hypothetical metal with $E_F = 1$ eV and lattice constant $a = 2$ Å. The height of the square potential barrier $\varphi_1 = \varphi_2 = 2$ eV, and the thickness $d = 10$ Å. Curve *1* corresponds to $T = 4.2$ K, $\Gamma = 0$; curve *2* corresponds to $T = 77$ K, $\Gamma = 0$; curve *3* corresponds to $T = 77$ K, $\Gamma = 3$ meV (curves *1* and *2* are shifted upwards by 0.04 and 0.02, respectively). The computational step $h = 0.5$ meV.

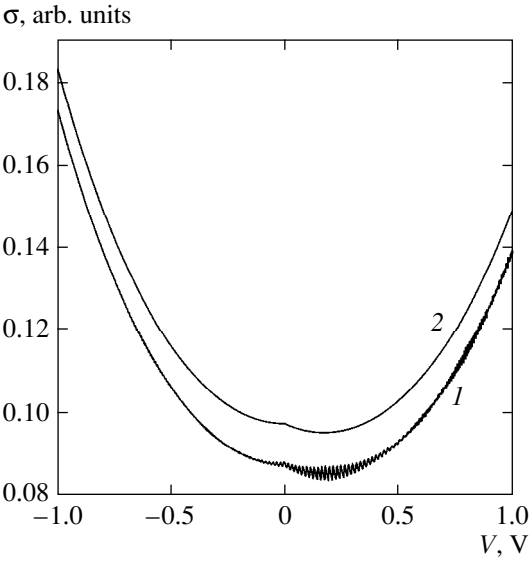

Fig. 4. Differential conductivity $\sigma(V)$ of a metal–insulator–quantized film (nonuniform over thickness) tunneling contact. It is assumed that the electrodes are made of a hypothetical metal with $E_F = 4$ eV and lattice constant $a = 2$ Å. The film thickness varies from 470 to 530 Å and satisfies the Gaussian distribution $W(L) \propto \exp\{-[\alpha(L - \bar{L})/\bar{L}]^2\}$, $L = 500$ Å, and $\alpha = 1/6$. The parameters of the square potential barrier are: $d = 10$ Å, $\varphi_1 = \varphi_2 = 4$ eV. Curve *1*, $T = 4.2$ K, curve *2*, $T = 77$ K (curve 2 is shifted upwards relative to curve *1* by 0.01 arb. units). The computational step $h = 2$ meV.

The high sensitivity of the computed curves to the temperature smearing, which makes it impossible to observe the resonance structure at nitrogen temperatures (Fig. 4, curve 2), is another unexpected result. We indicate one other feature in the behavior of $\sigma(V)$: calculated for a symmetric square barrier, assuming the Fermi energies are the same on both sides of the tunnel junction, this curve is strongly asymmetric. Its minimum is shifted along the voltage axis by the amount V_{sh}. Such shifts of the differential conductivity curve were investigated in detail in [10, 11] and were attributed exclusively to the asymmetry of the potential barrier. It was shown in [12] that together with the asymmetry of the potential barrier the difference in the Fermi energies of the materials of the edges of the tunneling junction can also result in an asymmetry in the differential conductivity. We underscore once again that in our case $E_{F_1} = E_{F_2}$, the potential barrier is square, $\varphi_1 = \varphi_2$, and therefore the effect under discussion is entirely due to the presence of size-quantization in one of the electrodes. This large-scale effect is stable with respect to an increase of temperature (see Fig. 4, curve *2*). In our view, of all the effects associated with the presence of the electronic standing waves, the latter promises to be the simplest to detect.

3. CONCLUSIONS

A theoretical model making it possible to calculate the voltage-dependence of the tunneling current and tunneling conductivity for a tunnel junction where size-quantization occurs in one of the electrodes was proposed. The minimal number of working parameters used in the model (the height φ and thickness d of the barrier, the Fermi energy E_F, and the lattice constant a of the size-quantized electrode) gives hope that it can be used successfully in experimental investigations. It was shown that together with the well-known oscillatory structure, which is due to two-dimensional resonance bands, additional effects can appear in the tunneling conductivity. For example, the presence of individual two-dimensional bands with tunneling of electrons from the quantized electrode can be manifested in the form of sharp isolated resistance peaks in the curve $\sigma(V)$. Moreover, the differential conductivity $\sigma(V)$ should be, on the whole, asymmetric relative to zero voltage and possess a minimum shifted along the voltage axis by the amount V_{sh}. A negative feature of these new effects is that they are insensitive to the temperature. As far as the last effect is concerned, being large-scale with respect to voltage it should be sensitive to modulation smearing and to smearing due to the finiteness of the lifetime. On this basis there is every reason to believe that the observation of the latter effect is a necessary indication of two-dimensional quantization in experimental samples.

ACKNOWLEDGMENTS

I thank V. M. Svistunov for encouragement in all respects.

REFERENCES

1. R. C. Jaklevic and J. Lambe, Phys. Rev. B **12**, 4146 (1975).
2. R. C. Jaklevic, J. Lambe, M. Mikkor, *et al.*, Phys. Rev. Lett. **26**, 88 (1971).
3. L. C. Davis, R. C. Jaklevic, and J. Lambe, Phys. Rev. B **12**, 798 (1975).
4. V. M. Svistunov, V. Yu. Tarenkov, and V. T. Vitchinkin, Fiz. Tverd. Tela (Leningrad) **21**, 3340 (1979) [Sov. Phys. Solid State **21**, 1929 (1979)].
5. V. M. Svistunov and V. Yu. Tarenkov, Pis'ma Zh. Éksp. Teor. Fiz. **26**, 34 (1977) [JETP Lett. **26**, 30 (1977)].
6. A. A. Galkin, V. M. Svistunov, A. I. D'yachenko, *et al.*, Pis'ma Zh. Éksp. Teor. Fiz. **21**, 259 (1975) [JETP Lett. **21**, 118 (1975)].
7. E. L. Wolf, *Principles of Electron Tunneling Spectroscopy* (Oxford Univ. Press, New York, 1985; Naukova Dumka, Kiev, 1990).
8. V. M. Svistunov and M. A. Belogolovskiĭ, *Tunnel Spactroscopy of Quasi-Particle Excitations in Metals* (Naukova Dumka, Kiev, 1986), p. 114.
9. Yu. F. Komnik, *The Physics of Metal Films* (Atomizdat, Moscow, 1979), p. 106.
10. R. B. Floyd and D. G. Walmsley, J. Phys. C **11**, 4601 (1978).
11. W. F. Brinkman, R. C. Dynes, and J. M. Rowell, J. Appl. Phys. **41**, 1915 (1970).
12. A. I. Khachaturov, V. M. Svistunov, and M. A. Belogolovskii, Czech. J. Phys. **46** (S2), 1031 (1996).

Translation was provided by AIP

Journal of Experimental and Theoretical Physics, Vol. 91, No. 3, 2000, pp. 547–552.
Translated from Zhurnal Éksperimental'noĭ i Teoreticheskoĭ Fiziki, Vol. 118, No. 3, 2000, pp. 629–636.

SOLIDS
Electronic Properties

Spin-Polarized Electron Tunneling in Lanthanum Manganite

V. M. Svistunov[a, *], Yu. V. Medvedev, V. Yu. Tarenkov, A. I. D'yachenko, E. Hatta[a], M. Mukasa[a], R. Aoki[b], H. Szymczak[c], S. Lewandowski[c], and J. Leszczynski[d]

Donetsk Physicotechnical Institute, Ukrainian National Academy of Sciences, Donetsk, 340114 Ukraine
[a]Nanoelectronics Laboratory, Faculty of Engineering, Hokkaido University, Sapporo 060, Japan
[b]College of Industrial Technology, Department of Electrical Engineering and Electronics, Amagasaki Nishi-koya 1-27-1, 661-0047 Japan
[c]Institute Fizyki, Polish Academy of Sciences, PLO2-668 Warsaw, Poland
[d]Technical University, 98-924 Lodz, Poland
**e-mail: svistuno@hsts.dipt.donetsk.ua*
Received December 31, 1999

Abstract—Tunneling experiments were performed on ceramic samples with the composition $La_{0.6}Sr_{0.4}MnO_3$, manifesting negative magnetoresistance. Two types of contacts were studied: symmetric (break junction type) and asymmetric ceramic–insulator–metal contact. A high magnetic-field sensitivity of the conductivity $\sigma(H)$ of the contacts was observed even when only one of the electrodes was magnetic. The effect was explained by the existence of spin-polarized localized states in the tunneling barrier. Their appearance was attributed to the formation of an oxygen-depleted, magnetically two-phase state of localized ferromagnetic nanoregions in an antiferromagnetic dielectric matrix in the near-contact region. © *2000 MAIK "Nauka/Interperiodica".*

1. INTRODUCTION

Manganese oxides $R_{1-x}A_xMnO_3$ (R = La, Pr, Nd ions and ions of other trivalent rare-earth elements, A = divalent Ca, Sr, Ba, Pb ions (see [1–4])) are the most widely studied materials whose resistive properties are highly sensitive to magnetic fields. The relationship between the magnetic and electric characteristics of these manganites is largely determined by the dopants A and the oxygen stoichiometry. For sufficiently high degrees of doping (x = 0.2–0.6), as temperature decreases, the compounds $R_{1-x}A_xMnO_3$ transform into the ferromagnetic phase, where they manifest metallic-type conductivity. The magnetoresistive effect in such compounds (up to 100% in a 6 T field) is observed near the ferromagnetic transition. It is usually determined as the relative change in the resistance $\Delta\rho/\rho(0) = [\rho(H) - \rho(0)]/\rho(0)$, where $\rho(0)$ and $\rho(H)$ are the resistivities, respectively, in a zero magnetic field and in a field H.

The properties of $R_{1-x}A_xMnO_3$ samples are also very sensitive to a change in the oxygen stoichiometry. Specifically, the composition $La_{0.8}Ca_{0.2}MnO_{3-\delta}$ is ferromagnetic with $\delta = 0$ but antiferromagnetic with $\delta = -0.2$ [5]. The magnetization of samples with low oxygen content exhibits, on cooling in zero field, a peak characteristic for the spin-glass state. As the oxygen deficiency increases, the samples demonstrate intensification of coercivity [6]. In systems with an oxygen deficiency (polycrystalline $La_{0.67}Ba_{0.33}MnO_{3-\delta}$ [7] or $La_{0.67}Pb_{0.4}MnO_{3-\delta}$ films [8]) the magnetoresistive effect intensifies as the oxygen content decreases.

The anomalies observed in [6] for the magnetic properties are similar to those observed in ferromagnetic granules embedded in a nonmagnetic matrix. This enabled the authors of [6] to assume that such manganites contain small ferromagnetic clusters (≤70 Å in size) in a nonferromagnetic matrix. As is well known [1], the model of a magnetically two-phase state of lanthanum manganites is an attempt to explain the nature of the colossal magnetoresistive effect near the Curie temperature. For oxygen-stoichiometric compound $La_{0.6}Sr_{0.4}MnO_3$ the magnetically two-phase state can arise as a result of electronic phase separation. The existence of this state is confirmed by the fact that the spontaneous magnetization at 4.2 K of the compound $La_{0.7}Sr_{0.3}MnO_3$ is, according to the data of [9], less than expected for total ferromagnetic ordering. Specifically, it is 95% of the latter. The investigations performed in [10, 11] also attest to the possible appearance of a magnetically two-phase state. This makes it possible to explain a variety of properties of the single crystal $La_{0.7}Sr_{0.3}MnO_3$ which are manifested near the Curie temperature [12]: sharp increase of the thermal expansion coefficient, maximum of the modulus of negative volume magnetostriction, and growth of the resistivity. The distinguishing feature of the behavior of such spin glasses—the presence of a susceptibility peak in weak fields ($H \leq 100$ Oe)—is observed in investigations of

the properties of the manganite series $R_{0.7}Ca_{0.3-x}Sr_xMnO_3$ (R = Pr, x = 0.10, 0.05, 0; R = Nd, x = 0) [13].

The weak binding of oxygen with the lattice and the ease with which oxygen diffuses have very important consequences for the formation of near-surface layers in manganite compounds as well as the formation of a transition contact layer between the material and the film deposited on it. Such an oxygen-deficient layer can also occur on intergrain boundaries of a polycrystalline material, which will also influence magnetoresistive effects in the samples. The most informative investigations of the electric properties of such near-surface layers can be performed by studying the tunneling effect in these materials. In the last few years electron tunneling has been used to study the low-field magnetoresistive effect in polycrystalline structures of doped manganites [14, 15] and junctions based on epitaxial films [16]. A spin valve has been realized in tunneling experiments [17] on contacts with lanthanum manganite electrodes with different coercive fields. Current flow in such magnetic structures is explained by the spin-dependent probability of tunneling of charge carriers [18] between ferromagnetic electrodes. However, such an approach neglects the possible influence of states inside the tunneling barrier on the character of the tunneling.

In the present work the magnetoresistive effect was observed on symmetric tunneling contacts $La_{0.6}Sr_{0.4}MnO_3$–$La_{0.6}Sr_{0.4}MnO_3$. The asymmetric contacts $La_{0.6}Sr_{0.4}MnO_3$–Pb showed a similar sensitivity to magnetic fields. Ceramic materials are used because high-quality tunneling contacts can be obtained by breaking thin ceramic plates [19]. The high degree of texture (i.e., the orientation of the ceramic microcrystals in one direction) is very important. It was achieved, as described in Section 2, by preparing compact ceramic plates with a small thickness (≤0.1 mm). As a rule, for a transverse breaking of such plates only one microcrystal–microcrystal tunneling contact (or Sharvin microcontact; in the latter case the resistance of the contact makes it possible to estimate its area) forms.

The simplicity of our technological technique made it possible to prepare many tens of tunneling contacts, making it possible to greatly expand the experimental possibilities. It was shown that a large spin-dependent tunneling magnetoresistive effect is possible even when one of the contacts is a nonmagnetic metal. The magnitude of the magnetoresistive effect in the contacts $La_{0.6}Sr_{0.4}MnO_3$–Pb reached 15%, while in symmetric contacts it was 10–30% (for $H \le 100$ Oe). The observed magnetoresistive effect can be explained by assuming that the spin-dependent electron tunneling occurs through states (or clusters of states) that possess a definite magnetic moment and are localized inside the barrier.

2. EXPERIMENT

Samples of lanthanum manganite $La_{0.6}Sr_{0.4}MnO_3$, which were prepared by solid-phase synthesis using chemically pure oxides La_2O_3 and MnO_2 and the carbonate $SrCO_3$, served as the experimental objects. Flat samples of a textured ceramic, whose grains are aligned in the direction of the transport current, are convenient for fabricating high-quality tunneling contacts (especially by the break junction method). Plates of such a ceramic with the dimensions $1 \times 0.1 \times 0.01$ cm^3 were obtained by compressing the $La_{0.6}Sr_{0.4}MnO_3$ powder between two flat steel anvils under 30–40 kbar pressure. Copper wires, 0.1–0.2 mm in diameter, glued parallel to one another on the surface of an anvil were used as a band support for the powder being compressed. As a result of pressure treatment, the powder was compressed into dense plane-parallel plates with thickness $d \le 0.1$ mm.

Samples with current and potential contacts consisting of silver paste were annealed at T = 1000°C for 6 h. The room-temperature resistance of the initial plates was 0.8–2 Ω. The transition resistance of the current and potential contacts was $R \approx 10^{-7}$ Ω cm^2. The temperature dependence of the resistance of the plates had a characteristic maximum near 360 K, associated with an insulator–metal phase transition. The measurements of the magnetization of the samples were performed in static magnetic fields up to 10 kOe by the weighing method. Under magnetization hysteresis was observed with coercivity field $H_c \approx 100$ Oe at 77 K (Fig. 1, inset (b)).

The tunneling junctions were prepared by two methods. Ceramic–nonmagnetic metal type junctions were produced by pressing a drop of lead into the ceramic surface. Symmetric junctions were produced by breaking a thin textured ceramic plate. A $La_{0.6}Sr_{0.4}MnO_3$ plate with current and potential contacts was placed on a textolite substrate and covered with a thick (thickness $d \sim$ 0.5 mm) layer of lacquer. After the lacquer polymerized, the substrate was bent until the ceramic broke. The moment of breaking was monitored according to the change in the resistance of the plate. A thick layer of lacquer eliminated possible shifts of the contact edges, and as a result its electric characteristics were highly stable. The resistance of the finished tunneling contacts was 100–300 Ω.

3. RESULTS AND DISCUSSION

Figure 1 shows the temperature dependence of the resistance of the experimental ceramic in magnetic fields H = 0 and H = 100 Oe at T = 77 K. Substantial hysteresis was observed (Fig. 1, inset a) in the magnetic field dependences of the resistance of the ceramic plate. In a H = 100 Oe field the resistance of a plate decreased by 3–4%.

At low temperatures the conductivity $\sigma(V)$ of both types of tunneling structures (symmetric and asymmet-

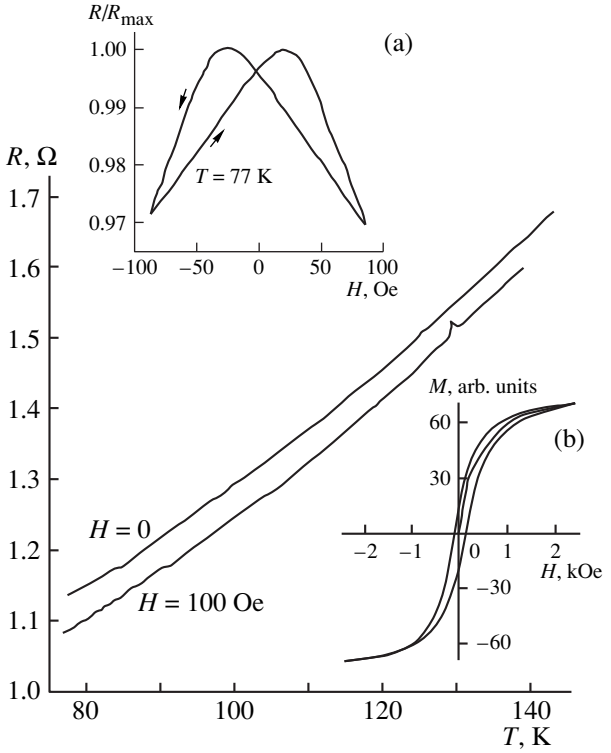

Fig. 1. Temperature dependence of the resistance of the ceramic $La_{0.6}Sr_{0.4}MnO_3$ in zero and finite magnetic fields. Inset: (a) Field dependence of the resistance of the sample, (b) field dependence of the magnetization of the sample.

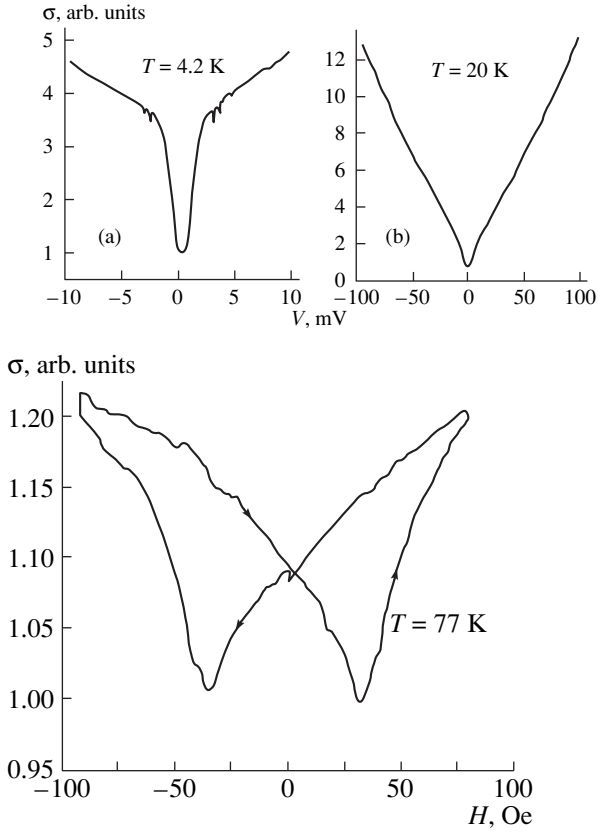

Fig. 2. Conductivity σ of a tunnel junction $La_{0.6}Sr_{0.4}MnO_3$–Pb versus the magnetic field. Inset: (a) Manifestation of an energy gap of lead and (b) $\sigma(V) = dI/dV$ at $T = 20$ K.

ric) was characterized by a linear dependence on the bias voltage V (see inset in Figs. 2 and 3). The appearance of such a linear background in the tunneling curves is attributed at present to resonance electron tunneling through localized states in the tunneling barrier. The Coulomb interaction of electrons (and holes) in localized levels results in the formation of a "soft" Coulomb pseudogap, when the density of localized levels in the barrier $g(\varepsilon) \sim |\varepsilon|$ [20]. In this case the differential conductivity of a tunneling contact, which is proportional to the density $g(\varepsilon)$, acquires the form $\sigma(V) = \sigma_0 + \gamma|V|$ [21, 22]. Here σ_0 is the conductivity of the contact at zero voltage and γ is a constant. The fact that such characteristics reflect electron tunneling processes is confirmed by the observation of an energy gap Δ_{Pb} of lead in the asymmetric junction $La_{0.6}Sr_{0.4}MnO$–Pb (see inset in Fig. 2). Even though the smearing of the gap features of Pb, which is associated in all probability with injection of spin-polarized electrons into the superconductor (see, for example, [23, 24]), is strong, the observed ratio $R(V = 0)/R(V > \Delta) \approx 4$ of the contact resistances is quite large, indicating that the tunneling mechanism of current flow through the manganite–lead structure predominates.

In the presence of resonance tunneling of electrons through localized states the conductivity of the contact $\sigma \sim \sigma_c N_{loc}$, where $\sigma_c = e^2/\pi\hbar$ is the conductivity quantum and $\sigma_c = 77.5 \times 10^{-6}\ \Omega^{-1}$ [25]. This relation makes it possible to estimate the number N_{loc} of localized states in the plane of the section of the experimental contacts according to the characteristic interval $\Delta\sigma$ in which a linear dependence of $\sigma(V)$ on the bias voltage V is observed (Figs. 2 and 3). For $\Delta\sigma \sim 0.1\ \Omega^{-1}$ we obtain the estimate $N_{loc} \geq 10^3$.

The conductivity of a $La_{0.6}Sr_{0.4}MnO$–Pb contact is highly sensitive to magnetic fields for $T > T_c$ of lead (Fig. 2) and with a lead electrode in the superconducting state (not shown in the figure). In the latter case a voltage $V > \Delta_{Pb}$ was applied to the contact. The experiments demonstrated that the character of the magnetoresistive effect in $La_{0.6}Sr_{0.4}MnO_3$–I–Pb contacts does not depend on the state (superconducting or normal) of the lead electrode.

Figure 3 shows the dependence $\sigma(H)$ for a symmetric contact $La_{0.6}Sr_{0.4}MnO_3$–$La_{0.6}Sr_{0.4}MnO_3$. It is evident that according to the form of the signal ("butterfly") and the magnetic-field sensitivity of the tunneling conductivity this characteristic is virtually identical to the curve $\sigma(H)$ for an asymmetric $La_{0.6}Sr_{0.4}MnO_3$–Pb contact (Fig. 2). There is a definite similarity between these curves and the dependence $R(H)$ for the ceramic

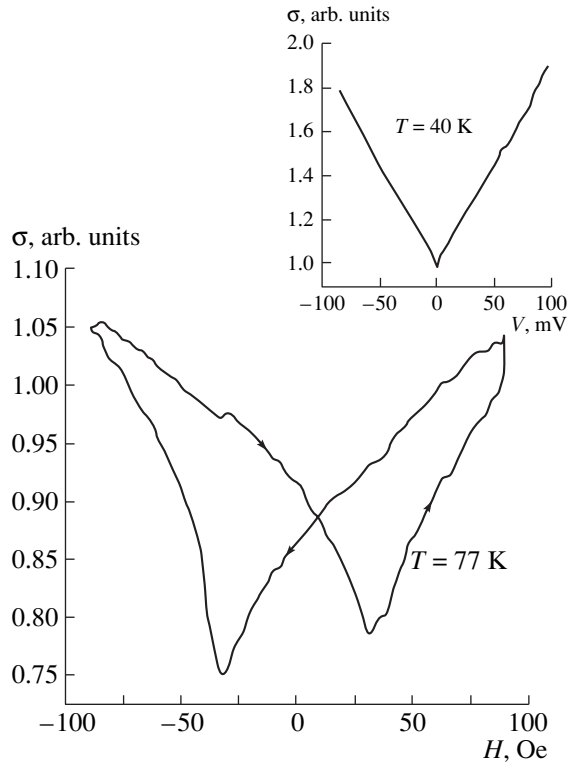

Fig. 3. $\sigma(H)$ for a $La_{0.6}Sr_{0.4}MnO_3$–$La_{0.6}Sr_{0.4}MnO_3$ contact. Inset: Tunneling conductivity $\sigma(V)$ at $T = 40$ K.

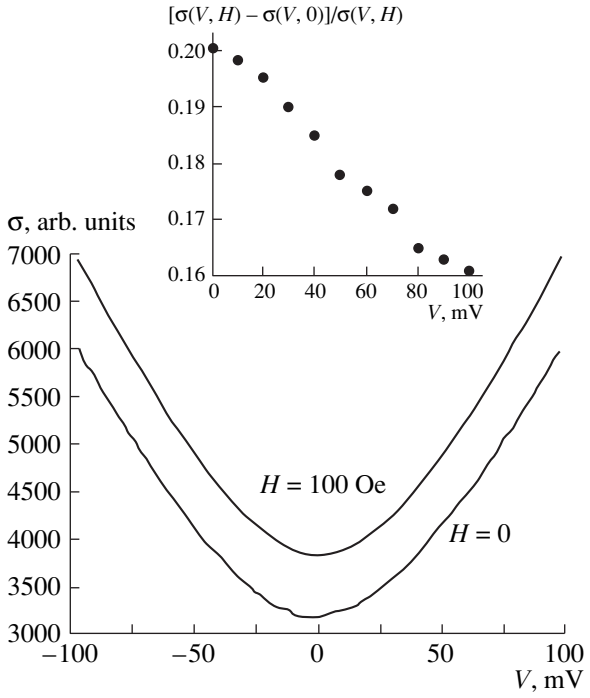

Fig. 4. Tunneling conductivity of a $La_{0.6}Sr_{0.4}MnO_3$–$La_{0.6}Sr_{0.4}MnO_3$ contact in the magnetic fields $H = 0$ and $H = 100$ Oe at $T = 77$ K. Inset: Normalized conductivity versus the bias voltage on a contact in a magnetic field $H = 100$ Oe.

$La_{0.6}Sr_{0.4}MnO_3$ (Fig. 1, inset a). However, the tunneling magnetoresistive effect is approximately an order of magnitude stronger.

The effect of a magnetic field on the conductivity $\sigma(V)$ of a symmetric $La_{0.6}Sr_{0.4}MnO_3$–$La_{0.6}Sr_{0.4}MnO_3$ contact is shown in Fig. 4. In a field $H = 100$ Oe the conductivity curve $\sigma(V)$ simply shifted upwards relative to the curve recorded in zero field. The maximum relative field sensitivity of the conductivity was observed for $V = 0$ and decreased with increasing voltage (see inset in Fig. 4).

The conductivity $\sigma(H)$ of the experimental tunneling junctions has a characteristic minimum (corresponding to the peak in the resistivity $\rho(H)$) in weak magnetic fields. For tunneling junctions with electrodes made of ferromagnetic materials (Fig. 3) the existence of such a peak is explained by the spin-valve effect [17, 26–28]. As is well known [23, 28], spin is conserved in elastic tunneling of electrons. Consequently, the strength of the tunneling current depends on the relative orientation of the magnetic moments $\mathbf{M}_1$ and $\mathbf{M}_2$ of the electrodes. The resistance of the contact is smallest when the magnetic moments $\mathbf{M}_1$ and $\mathbf{M}_2$ are parallel to one another and largest if the moments are antiparallel. In the latter case the channel for current flow with conservation of the orientation of the electron spin is closed.

Let the coercive fields of the electrodes H_{c1} and H_{c2} be different and let $H_{c1} > H_{c2}$. Then the antiparallel orientation of the magnetization vectors $\mathbf{M}_1$ and $\mathbf{M}_2$ of the electrodes is attained in magnetic fields $H_{c1} < |H| < H_{c2}$. In this range the resistance of a tunneling contact is maximum, $R = R_{\uparrow\downarrow}$. In the opposite case, for $|H| \gg H_{c2}$, H_{c1}, the magnetization vectors of the electrodes are oriented parallel and the contact resistance reaches its minimum value ($R = R_{\uparrow\uparrow}$).

However, a break junction type contact is, according to the method used to fabricate it, a symmetric (in the sense of the magnetic properties of the plates) tunnel contact, whereas in the asymmetric contact $La_{0.6}Sr_{0.4}MnO_3$–Pb lead is not a magnetic material. At the same time the behavior of the conductivity $\sigma(H)$ of these two, at first glance completely different, structures is completely identical (Figs. 2 and 3). This shows that the mechanism magnetoresistance in both cases is of the same nature.

In this connection we note that the data presented in Figs. 2 and 3 demonstrate a characteristic dependence of the magnetoresistance of tunneling junctions with impregnations of small ferromagnetic clusters in the insulating layer [20, 30]. Thus, in [29] a Co–Al_2O_3–Co tunneling contact contained a layer of inclusions of Co nanoparticles in aluminum oxide. In the experiment of [30] the tunneling structure was fabricated on the basis of a thin granular film, which consisted of a nonuni-

form magnetic medium consisting of magnetically hard ferromagnetic $Co_{80}Pt_{20}$ nanoparticles in a SiO_2 matrix.

On the basis of the behavior of the magnetoresistance in the cases of [29, 30] one can imagine that a phenomenon close to the spin-dependent tunneling transition of electrons through localized ferromagnetic centers in the barrier can also occur in our contacts. The most obvious reason for the appearance of ferromagnetic nanoparticles in the barrier layer, both a metal (Pb) in direct contact with $La_{0.6}Sr_{0.4}MnO_3$ and a symmetric contact, is an oxygen-depleted region in $La_{0.6}Sr_{0.4}MnO_3$ in the contact region. According to [6], for an average oxidation state of Mn ions of less than three in this region the magnetization M of the near-contact layer should decrease relative to the value of M in the interior volume, and the region itself will represent a structure consisting of ferromagnetic inclusions in a dielectric matrix. The coercivity field of such a magnetically two-phase near-barrier layer (in our case $H_{c1} \approx H_{c2} \sim 30$ Oe, Figs. 2 and 3) can differ appreciably from that of electrodes, which, as follows from measurements of the magnetization loop of $La_{0.6}Sr_{0.4}MnO_3$, is approximately 100 Oe (Fig. 1).

The conductivity $\sigma(H)$ of the contacts is lowest in an external field $H = H_{c1} \approx H_{c2}$ (Figs. 2 and 3). At this point the magnetization of the electrode sections next to the tunneling contact (and also of the ferromagnetic inclusions in the barrier) is zero, i.e., for tunneling electrons there is no fixed spin orientation. If $|H| > H_{c1} \approx H_{c2}$, then the magnetization vectors of the ferromagnetic inclusions in the barrier rotate together with the magnetization vectors of electrodes in the same direction as the field increases, so that the conductivity $\sigma(H)$ increases (Figs. 2 and 3). This picture is suitable for describing the symmetric and asymmetric contacts investigated in the present work.

We shall now examine a symmetric contact in greater detail (Fig. 3). If the coercive fields of the electrodes are essentially the same ($H_{c1} \approx H_{c2}$), then the tunneling magnetoresistive relation following from the spin-dependent electron tunneling model [26, 27] becomes $\Delta R/R_{\uparrow\uparrow} = (R_{\uparrow\downarrow} - R_{\uparrow\uparrow})/R_{\uparrow\uparrow} = P^2$, where P is the polarization of the electron spins at the boundaries of a tunneling contact. The maximum magnetoresistive effect (50%) for such a symmetric contact is attained with $P = 1$.

For the observed ratio $\Delta R/R_{\uparrow\uparrow} \approx 0.3$ (Fig. 3) this gives for $La_{0.6}Sr_{0.4}MnO_3$ the value $P = 0.55$ for the polarization of the electrons. We note that such an analysis of the tunneling curves of the junctions $La_{0.67}Sr_{0.33}MnO_3$–$SrTiO_3$–$La_{0.67}Sr_{0.33}MnO_3$ [16] and $La_{0.7}Sr_{0.3}MnO_3$–$SrTiO_3$–$La_{0.7}Sr_{0.3}MnO_3$ [17] at $T = 4.2$ K leads, respectively, to $P \approx 0.54$ and $P \approx 0.83$, which agrees with our results. The small discrepancy between the values of P [16, 17] and our value could be due to the state of the interface in the experimental tunneling contacts [33, 34], since the tunneling electrons "feel" only the near-surface electrode layer whose this is several tens of angstroms. In addition, there is a temperature dependence of the polarization [31], determined by the constant α, which depends on the material and is different for the interior volume or the surface ($P(T) = P_0(1 - \alpha T^{3/2})$). Ordinarily, α is larger for the latter because the exchange interaction decreases at the surface of the sample [32].

Returning to Fig. 4, we note that the observed dependence of the magnetoresistance on the bias voltage V on the contact (inset in Fig. 4) agrees with the results of [29] (see also [34]). The magnetoresistive ratio on the junctions CoFe–Al_2O_3–Co in the classical work [26] decreases almost by a factor of 20 over a range of 0.7 V, whereas for Co–Al_2O_3–Co samples [29] containing a layer of Co nanoparticle inclusions in aluminum oxide it decreases in the same voltage range by only a factor of 2, and on (Au–Cr)–thin granular $Co_{36}Al_{22}O_{42}$ film junctions its value is essentially independent of the voltage. This discrepancy is due to the complicated structure of the junction, because of which the voltage is redistributed between the granules and the electrodes. A possible explanation could lie in the fact (see [29]) that the strong dependence of the magnetoresistance on V in classical junctions is due to the fine structure in the spin-polarized density of states. However, in junctions with ferromagnetic microregions in an insulating interlayer this fine structure can wash out as a result of tunneling through resonance levels in the barrier.

In [26, 27] the analysis was based on the concept of direct (i.e., elastic) tunneling. In our case, however, a "linear background" $\gamma|V|$, $\sigma(V) \approx \sigma_0 + \gamma|V|$ (insets in Figs. 2 and 3) is observed in the conductivity σ of the contacts; this background indicates that resonance tunneling of electrons through a system of localized states in the barrier makes a large contribution to the current. In principle, both parameters, σ_0 and γ, could have been functions of the magnetic field H. But, experimentally (Fig. 4), the dependence of the slope γ of the tunneling curves on the field H is very weak, so that the result can be represented as the ratio $[\sigma_0(H = 0) - \sigma_0(H)]/[\sigma_0(H) + \gamma V]$, i.e., the quantity $\Delta\sigma/\sigma$ should decrease as the voltage V on the contact increases. The inset in Fig. 4 demonstrates this dependence.

The parameter γ, determining the slope of the linear background directly, is due to resonance tunneling of electrons through localized states [21, 22]. Its independence on the magnetic field shows for our experimental contacts the resonance processes make virtually no contribution to the spin-polarized electron tunneling.

It is well known that localized states located at the center of the tunneling barrier make the dominant contribution to resonance tunneling of electrons [21, 25]. According to what has been said above, such states do not contribute to the observed spin-polarized tunneling of electrons. Consequently, the dependence $\sigma(H)$ is entirely determined by electron tunneling through

localized states located at the edges of the tunneling barrier close to the contact electrodes.

4. CONCLUSIONS

The experimental results show that spatially nonuniform magnetic states consisting of ferromagnetic microregions in an antiferromagnetic matrix can occur in the near-surface layers of lanthanum manganite. The explanation of this magnetoresistive effect in such tunneling contacts could also involve the formation of a band of localized states in the insulating interlayer. Localized levels located at the center of the tunneling barrier contribute to resonance tunneling of electrons and lead to a "linear background" $\gamma|V|$ in the conductivity $\sigma(V)$ of a contact. Tunneling through such states makes no contribution to the magnetoresistive effect (in any case, in weak magnetic fields $H \leq 100$ Oe). However, localized states at the junction edges possess a definite magnetic moment, so as to ensure functioning of the tunneling spin valve.

ACKNOWLEDGMENTS

One of us (V. M. S.) is grateful to professor N. Tambo, the President of the University of Hokkaido, for the opportunity to perform joint experiments and the Ministry of Education, Science and Culture of Japan for supporting a series of lectures at the University of Hokkaido. This work was also supported in part by the Telecommunications Advancement Organization of Japan and by the Polish Committee of Scientific Research (KBN), grant no. 2 P03B 148 14.

REFERENCES

1. É. L. Nagaev, Usp. Fiz. Nauk **166**, 833 (1996) [Phys. Usp. **39**, 781 (1996)].
2. A. P. Ramírez, J. Phys.: Condens. Matter **9**, 8171 (1997).
3. C. N. R. Rao, R. Mahesh, A. K. Raychaudhuri, and R. Mahendiran, J. Phys. Chem. Solids **59**, 487 (1998).
4. Y. Tokura and Y. Tomioka, J. Magn. Magn. Mater. **200**, 1 (1999).
5. S. Tamura, Phys. Lett. A **73**, 401 (1980).
6. H. L. Ju and H. Sohn, J. Magn. Magn. Mater. **167**, 200 (1997).
7. H. L. Ju, J. Gopalakrishnan, J. L. Peng, *et al.*, Phys. Rev. B **51**, 6143 (1995).
8. A. Milner, A. Gerber, B. Groisman, *et al.*, Phys. Rev. Lett. **76**, 475 (1996).
9. A. Urushibara, Y. Moritomo, T. Arima, *et al.*, Phys. Rev. B **51**, 14103 (1995).
10. R. von Helmolt, J. Wecker, K. Samwer, and K. Barner, J. Magn. Magn. Mater. **151**, 411 (1995).
11. O. Yu. Gorbenko, R. V. Demin, F. R. Kaul', *et al.*, Fiz. Tverd. Tela (St. Petersburg) **40**, 290 (1998) [Phys. Solid State **40**, 263 (1998)].
12. L. I. Koroleva, R. V. Demin, and A. M. Balbashov, Pis'ma Zh. Éksp. Teor. Fiz. **65**, 449 (1997) [JETP Lett. **65**, 474 (1997)].
13. A. Maiqnan, U. V. Varadaraju, F. Millange, and B. Raveau, J. Magn. Magn. Mater. **168**, L237 (1997).
14. H. Y. Hwang, S-W. Cheong, N. P. Ong, and B. Batlogg, Phys. Rev. Lett. **77**, 2041 (1996).
15. A. Gupta and J. Z. Sun, J. Magn. Magn. Mater. **200**, 24 (1999).
16. Yu Lu, X. W. Li, G. Q. Gong, *et al.*, Phys. Rev. B **54**, R8357 (1996).
17. M. Viret, M. Drouet, J. Nassar, *et al.*, Europhys. Lett. **39**, 545 (1997).
18. J. S. Helman and B. Abeles, Phys. Rev. Lett. **37**, 1429 (1976).
19. V. M. Svistunov, V. Yu. Tarenkov, A. I. D'yachenko, *et al.*, Zh. Éksp. Teor. Fiz. **113**, 1397 (1998) [JETP **86**, 763 (1998)].
20. A. L. Efros and B. I. Shklovski, J. Phys. C **8**, L49 (1975).
21. A. I. D'yachenko, V. Yu. Tarenkov, A. V. Abaleshev, and V. M. Svistunov, Fiz. Nizk. Temp. **22**, 267 (1996) [Low Temp. Phys. **22**, 208 (1996)].
22. A. I. D'yachenko, Fiz. Tekh. Vys. Davleniĭ **5**, 5 (1995).
23. P. M. Tedrow and R. Meservey, Phys. Rev. Lett. **26**, 192 (1971); Phys. Rev. B **7**, 318 (1973).
24. V. A. Vas'ko, V. A. Larkin, P. A. Kraus, *et al.*, Phys. Rev. Lett. **78**, 1134 (1997).
25. A. I. Larkin and K. A. Matveev, Zh. Éksp. Teor. Fiz. **93**, 1030 (1987) [Sov. Phys. JETP **66**, 580 (1987)].
26. M. Julliere, Phys. Lett. A **54**, 225 (1975).
27. J. S. Moodera, L. R. Kinder, T. M. Wong, and R. Meservey, Phys. Rev. Lett. **74**, 3273 (1995).
28. R. Meservey and P. M. Tedrow, Phys. Rep. **239**, 174 (1994).
29. L. P. Schelp, A. Fert, F. Fettar, *et al.*, Phys. Rev. B **56**, R5747 (1997).
30. K. Inomata, H. Ogiwara, Y. Saito, *et al.*, Jpn. J. Appl. Phys. **36**, L1380 (1997).
31. Chang He Shang, J. Nowak, R. Jansen, and J. S. Moodera, Phys. Rev. B **58**, R2917 (1998).
32. D. L. Mills and A. A. Maradudin, J. Phys. Chem. Solids **28**, 1855 (1967); J. Mathon and S. B. Ahmad, Phys. Rev. B **37**, 660 (1988).
33. D. T. Pierce, R. J. Celotta, J. Unguris, and H. C. Siegmann, Phys. Rev. B **26**, 2566 (1982); D. Mauri, D. Scholl, H. C. Siegmann, and E. Kay, Phys. Rev. Lett. **61**, 758 (1988).
34. S. Mitani, S. Takahasi, K. Takanashi, *et al.*, Phys. Rev. Lett. **81**, 2799 (1998).

Translation was provided by AIP

Journal of Experimental and Theoretical Physics, Vol. 91, No. 3, 2000, pp. 553–561.
Translated from Zhurnal Éksperimental'noĭ i Teoreticheskoĭ Fiziki, Vol. 118, No. 3, 2000, pp. 637–646.

SOLIDS
Electronic Properties

Temperature Dependence of the Conductivity of Cu : SiO_2 Composite Films. Experiment and Numerical Simulation

D. A. Zakheĭm*, I. V. Rozhansky, I. P. Smirnova, and S. A. Gurevich
Ioffe Physicotechnical Institute, Russian Academy of Sciences, St. Petersburg, 194021 Russia
**e-mail: mitya@quantum.ioffe.rssi.ru*
Received January 15, 2000

Abstract—Experimental data are presented for the temperature dependence of the conductivity of Cu : SiO_2 metal-insulator composite films containing 3-nm Cu granules. At low temperatures in the concentration range 17–33 vol % Cu, all of the conductivity curves have a temperature dependence of the form $\sigma \propto \exp\{-(T_0/T)^{1/2}\}$, while at higher temperatures a transition is observed to an activational dependence. A numerical simulation of the conduction in a composite material shows that an explanation of the observed temperature dependence must include the Coulomb interaction and the presence of a rather large random potential. The simulation also yields the size dependence and temperature dependence of the mesoscopic scatter of the conductivities of composite conductors. It is shown that a self-selecting percolation channel of current flow is formed in the region of strong mesoscopic scatter. *© 2000 MAIK "Nauka/Interperiodica".*

1. INTRODUCTION

Recently there has been a considerable heightening of interest in the study of the electrical properties of granular nanostructures—island metal films and metal-insulator structures containing conducting granules several nanometers in size, separated by thin insulating gaps. If the density of granules in such a structure is not too high, a regime of tunneling conduction is realized in it. The most important circumstance is that for metallic granules of the order of several nanometers in size, their charge energy $E_c = e^2/2C$ (C is the capacitance of a granule) amounts to around 100 meV, which is considerably higher than the energy of thermal noise, even at room temperature. For this reason, macroscopic charge quantization effects are observed at high temperatures [1]. For comparison, the structures fabricated for single-electron electronics by electron lithography contain metallic islands with minimum dimensions of the order of tens of nanometers, which corresponds to a working temperature range of the order of a few Kelvins [2]. For this reason, granular nanostructures are being considered as a basis for the creation of new nanoelectronic devices capable of operating at room temperature [3].

Electron transport processes in granular metallic films and metal-insulator nanostructures are largely determined by the random character of these media. Indeed, it is in principle impossible to control precisely the size and position of the individual metallic granules, and the charge energies and tunneling resistances between granules therefore have a certain scatter. Furthermore, in such structures there are fluctuations of the potential of the individual granules [4]. For this reason the electrical properties of such structures must be described using statistical approaches. Statistical approaches have been developed in a number of studies [4–6]. It has been shown [6] that the value of the random potential in the structure determines the equilibrium density of positively and negatively charged particles. The density of single-particle excitations in a random granular medium was considered in [7], and it was shown that it has a dip (the Coulomb gap) near the Fermi level, the shape of this dip depending on the value of the random potential. In [8] the charge energies of the granules were calculated with allowance for the random environment and finite size of the particle. Although some progress has been made, the theory of transport phenomena in granular media is not yet complete. In particular, there is as yet no generally accepted explanation for the experimental temperature dependence of the conductivity. For example, some authors attribute the behavior $\ln\sigma \sim T^{-1/2}$ that is widely observed in experiment [9] to structural features of the composite material [5, 10], while others point to an important role of the Coulomb interaction between charged granules [6, 11]. There have also been indications that it is necessary to take into account the multiparticle excitations in the theoretical treatment of the conduction of granular structures at low temperatures [12]. Moreover, it remains an open question as to how much the statistical character of the media affects the reproducibility of the electrical characteristics of granular conductors, the sizes of the conductors and the temperatures for which one can observe a strong mesoscopic scatter of the electrical characteristics.

In this paper we report an experimental study of the temperature dependence of the conductivity of Cu : SiO_2 composite films obtained by the method of simultaneous magnetron sputtering of two sources. The exper-

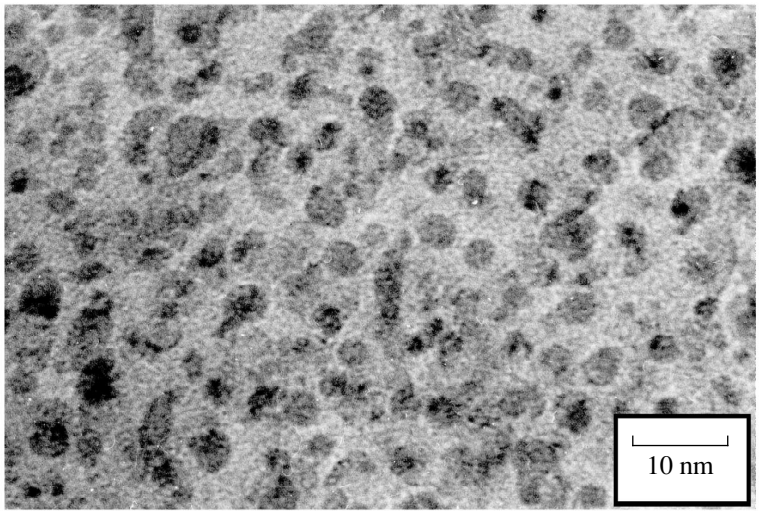

Fig. 1. Transmission electron micrograph of a composite film with a copper concentration of 27 vol %.

imental data are analyzed on the basis of a numerical model of hopping conduction. The model is based on a representation of the electrical conduction by means of single-electron hops; the single-electron density of states is constructed with allowance for the Coulomb interactions between charged granules. We believe that neglecting the multiparticle excitations is justified, since we are not considering the case of very low temperatures here and, as was shown in [12], taking multiparticle excitations into account does not lead to a new type of temperature dependence of the conductivity but only modifies the parameter T_0 of the law $\sigma \propto \exp\{-(T_0/T)^{1/2}\}$ at low temperatures.

The results of the simulation are in good agreement with the experimental data; moreover, approximation of the experimental temperature dependence of the conductivity by the calculated dependence made it possible to determine the values of several microscopic parameters of the composite material and their dependence on the concentration of the metallic phase. The model constructed here can also be used to study the dependence of the amplitude of the mesoscopic scatter of the conductivity of composite conductors as a function of the dimensions of the conductors and the temperature.

2. TECHNOLOGY OF FABRICATION OF COMPOSITE CONDUCTORS

To obtain amorphous SiO_2 films containing copper clusters we used the method of joint magnetron sputtering. The sputtering was done on an Alcatel SCM-450 apparatus with the chamber pumped down beforehand to 10^{-5} Pa. As a source of SiO_2 we used a planar magnetron with a quartz target, which was sputtered in an rf discharge. A dc voltage was applied to the second magnetron, which had a copper target. Both targets were 100 mm in diameter and were made of high-purity materials (99.999%). The sputtering was done in an atmosphere of high-purity argon at a pressure of 0.3 Pa. As in [13], the substrate holders were rotated during deposition of the films, so that the substrates were alternately found above the Cu and SiO_2 targets. The necessary copper concentration in the film was set by changing the rate of deposition of the SiO_2 in the interval 0.42–3.3 nm/min at a constant copper deposition rate of 0.53 nm/min. The rotational velocity of the substrate holders was 8 rpm, so that less than one monolayer of material was deposited in each revolution of the substrate above the sources. The volume concentration of copper in the samples studied was 17–33 vol %.

Figure 1 shows a typical transmission electron micrograph of a Cu : SiO_2 composite film with a copper concentration of 27 vol %. The presence of metallic clusters with a characteristic size of 3 nm is evident. We note that the films grown were not subjected to additional annealing, and cluster formation occurred in them directly during deposition. Analysis of the resulting films by x-ray photoelectron spectroscopy [14] showed that the copper in the film was found in the unoxidized state. Furthermore, studies of the fine structure of the x-ray absorption edge [15] show that the clusters have close to a crystalline structure, but a small fraction of the Cu atoms remain dispersed in the insulating matrix. It has been shown that the structure of the clusters become less dense as the volume concentration of Cu decreases. A study of the films by the small-angle x-ray scattering technique yielded a more precise value of the average size of the granules. For example, for films with a Cu concentration of 27 vol % the average granule size determined by that technique was 1.6 nm. The scatter of the granule size in these films, according to the transmission electron microscope and small-angle x-ray scattering data, is 20–30%.

To study the temperature dependence of the conductivity of composite films we prepared a series of samples with different volume concentrations of Cu. The films were 200 nm thick; Au/Cr contacts 400×400 μm in area were deposited on top, with the gap between contacts varied over the interval 5–200 μm.

3. TEMPERATURE DEPENDENCE OF THE CONDUCTIVITY

The temperature dependence of the conductivity was measured in the temperature range 30–300 K in an Oxford 450 closed-cycle helium cryostat by means of a V7-49 precision electrometer. The current-voltage characteristics of the samples with different gaps between contacts had a linear initial segment at all temperatures in the range studied, and the value of the differential conductivity obtained from the slope of this linear segment did not depend on the size of the gap. This confirms the Ohmic properties of the contacts between the metal and the composite film. The temperature dependence of the conductivity was measured at a voltage corresponding to the linear part of the current-voltage characteristics.

Figure 2 shows the experimental temperature dependence of the conductivity for four samples with different Cu concentrations, plotted on a scale of $\ln\sigma$ versus $1/T^{1/2}$. As we see in this plot, at low temperatures the conductivity of all the samples studied follows a $\sigma \propto \exp\{-(T_0/T)^{1/2}\}$ law, in good agreement with the published data available [5, 11]. At the same time, at higher temperatures a smooth increase of the exponent in the power law occurs, in the direction of an activational dependence. Here the boundary temperature at which the transition from one temperature dependence to the other occurs increases with decreasing concentration of the metallic phase. The dashed lines in Fig. 2 show the linear (on a scale of $\ln\sigma$ versus $1/T^{1/2}$) approximations of the experimental temperature dependence of the conductivity. The parameters T_0 of these approximations are listed in table.

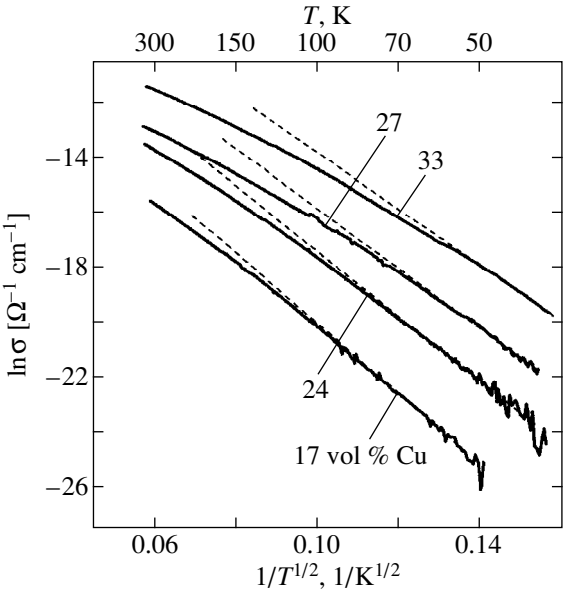

Fig. 2. Experimental temperature dependence of the conductivity of composite films on the scale of $\ln\sigma$ versus $1/T^{1/2}$.

4. MODELING OF THE TEMPERATURE DEPENDENCE OF THE CONDUCTIVITY

4.1. Construction of the Model

For analysis of the experimental data obtained, we constructed a numerical model of the conductivity of a composite material. In this model a composite material is represented as a set of metallic spheres randomly situated in a cube of a specified volume; the radius r of the spheres is Gaussian-distributed about its mean value r_0:

$$\rho(r) = \frac{1}{\sqrt{2\pi}\xi}\exp\left(-\frac{(r-r_0)^2}{2\xi^2}\right), \qquad (1)$$

where ξ is the width of the distribution.

Another parameter of the random distribution of spheres is the minimum admissible distance between them, $\Delta_{\min}$. For a numerical realization of the quasi-random arrangement we used the following method: first the spheres were analyzed in a regular cubic lattice, and then each sphere was randomly assigned an initial velocity, and its motion was calculated with allowance for collisions, with the remaining spheres being assumed immobile. The effective radius of each sphere was taken as $r_i^{\text{eff}} = r_i + \Delta_{\min}/2$, which guaranteed that the distance between spheres was not smaller than $\Delta_{\min}$. In addition, all of the collisions were assumed to be completely inelastic, i.e., after each collision the velocity of the sphere was again assumed random. The calculation was terminated when each sphere had undergone a large number of collisions (several hundred), and then the resulting arrangement was used as a random realization.

Of course, the arrangement of the granules in a real composite material is not completely random; a certain correlation can be present in their relative position on account of the way these granules were grown. In the

Results of a fitting of the experimental temperature dependence of the conductivity of composite films by the calculated dependence

c_{Cu}, vol %	kT_0, meV	ϵ	$\Delta_{\min}$, nm	Δ_ϕ, eV
17	1464	6.6	0.6	0.12
24	1248	6	0.54	0.11
27	1048	6	0.52	0.1
33	702	6	0.5	0.09

proposed model the parameter $\Delta_{\min}$ is introduced for taking such correlations into account, and the dependence of the electronic properties of the material on the value of this parameter will be demonstrated below.

For describing the electrical properties of the resulting system of metallic spheres the Hamiltonian of the system is written as

$$H = \sum_i \phi_i Q_i + \frac{1}{2}\sum_{i,j} C_{ij}^{-1} Q_i Q_j. \quad (2)$$

Here Q_i is the charge on the ith sphere, which can take one of three values: $-e$, 0, and e, where e is the charge of the electron (for simplicity we are neglecting multiply charged states of the spheres); (ϕ_i is the so-called random potential [4, 6] on granule i, and C_{ij}^{-1} is the inverse matrix of the capacitance coefficients and the coefficients of the electrostatic induction of the system [16]. In the simplest approximation, when the distance between spheres is much larger than the size of the spheres, the diagonal elements of this matrix are equal to the reciprocal of the capacitance of a metallic sphere of radius r embedded in an insulator with dielectric constant ϵ: $C_{ii}^{-1} = 1/\epsilon r_i$, and the off-diagonal elements can be written as $C_{ij}^{-1} = 1/\epsilon d_{ij}$, where d_{ij} is the distance between the centers of the ith and jth spheres. However, this extremely simple approximation is poorly suited to the case of a composite material with a relatively high concentration of the metallic phase, since the characteristic distances between granules are comparable to the dimensions of these granules, and one must take into account the corrections due to the polarization of the granules. The matrix C_{ij}^{-1} is calculated using the dipole approximation described in [8]. It follows from [8] that taking the polarization of the spheres into account leads to a decrease in their charge energy E_c in comparison with the value for an isolated sphere of radius r: $E_{c0} = Q^2/2\epsilon r$; taking the polarization into account also modifies the dependence of the Coulomb interaction of the charged spheres (the off-diagonal elements of the matrix C_{ij}^{-1}) on the distance between spheres. Furthermore, the values of the charge energies of the spheres have some scatter due to the random distribution of the spheres surrounding them.

The random potential ϕ_i that appears in Eq. (2) and serves as a source of diagonal disorder in the system is ordinarily attributed to the electric fields of defects and charged states in the matrix and at the granule/insulator interfaces. We note that, unlike [6], we do not attribute the random potential to fluctuations of the size and shape of the granules. The random potential here is understood to mean the potential of an external (with respect to the system of metallic spheres) electric field; only such a potential can give rise to charged granules in the ground state of the system. In the given model the random potential is assumed to be uniformly distributed over a specified range $\Delta\phi$: $-\Delta\phi/2 < \phi_i < \Delta\phi/2$.

4.2. Ground State and Spectrum of Single-Particle Excitations

The first step in the modeling of the conductivity of a composite material is to find the ground state of the system described by Hamiltonian (2). Because of the presence of the random potential, some of the metallic particles in this state may be charged. In our model the ground state of the system was determined by minimizing Hamiltonian (2) with respect to all possible pairwise permutations of electrons between the metallic spheres. The procedure was repeated many times until pairwise permutations that lowered the total energy of the system were no longer found.

Let us construct the spectrum of single-particle excitations of the system, taking the energy of the ground state as the zero of energy. The energy of a single-particle excitation is defined as the energy required to transport an electron or hole from infinity to one of the spheres. This energy is the value of Hamiltonian (2) calculated for a system with one "extra" electron or hole in comparison with the ground state. Here a sphere having charge 0 in the ground state supplies one level to each of the electron and hole densities of states, a sphere having charge $+e$ supplies a twofold degenerate level to the electron density of states, and a sphere having charge $-e$ supplies a twofold degenerate level to the hole density of states. Furthermore, it must be taken into account that each metallic sphere actually represents not an isolated energy level but a spectrum of levels, with a density of the form $\rho(E) = \theta(E - E_i)$ [11], where E_i is an energy level obtained by the method described above. The total electron or hole density of states will thus have the form

$$\rho^{e,h}(E) = \sum_i \theta(E - E_i^{e,h}) g_i^{e,h}, \quad (3)$$

where $g_i^{e,h} = \{1, 2\}$ is the degeneracy of the electron or hole level mentioned above.

Figure 3 shows the results of a calculation of the density of states (3) for a monodisperse system containing 10^3 spheres, for different values of the scatter $\Delta\phi$ of the random potential. The inset in this figure gives the dependence of the fraction of charged granules in the ground state as a function of the scatter $\Delta\phi$ of the random potential. It is seen that charged granules arise in the ground state when the random potential becomes comparable to the charge energy of the particles. At such values of the random potential a parabolic Coulomb gap is formed in the density of states, after which the form of the density states changes little as the scatter of the random potential is increased further.

4.3. Modeling the Temperature Dependence of the Conductivity

In the single-particle approximation the probability of a tunneling transition of the electron (or hole) between any two granules i, j is determined by the distance d_{ij} between them, the single-particle energy levels E_i, E_j, and the temperature. For this probability we used the expression [17]

$$\gamma_{ij} = \gamma^0 \exp(-\chi d_{ij}) N(E_j - E_i), \tag{4}$$

where γ^0 is a constant, χ is the tunneling transparency of the insulator (we have used the value $\chi = 10^8$ cm^{-1} for SiO_2), and $N(\Delta_{ij})$ is the equilibrium Planck distribution of phonons with energy $\Delta_{ij} = E_j - E_i$:

$$N(\Delta_{ij}) = \frac{\Delta_{ij}^2}{\exp(\Delta_{ij}/kT) - 1}, \tag{5}$$

where k is Boltzmann's constant and T is the temperature. To find the total current between granules i and j we must sum the rates of all possible electron and hole transitions $i \longrightarrow j$ and $j \longrightarrow i$ (with the opposite sign) with the density of states (3) taken into account. For example, the rate of the electron transition $i \longrightarrow j$ is

$$\Gamma^e_{i\to j} = \gamma^0 \exp(-\chi d_{ij}) \int\limits_{E_i = E_i^e}^{\infty} \int\limits_{E_j = E_j^e}^{\infty} f(E_i) \times (1 - f(E_j)) |N(E_j - E_i)| dE_i dE_j. \tag{6}$$

Here $E^e_{i,j}$ are the single-particle energy levels of electrons on granules i, j, and $f(E)$ is the Fermi distribution: $f(E) = 1/(\exp((E - E_F)/kT) + 1)$ (the Fermi level E_F is zero in our case, since the numbers of electrons and holes are always equal, and the density of states is symmetric about the zero of energy). For holes the rate of the transition $i \longrightarrow j$ is written in an analogous way, but with the electron energy levels $E^e_{i,j}$ in (6) replaced by the hole levels $E^h_{i,j}$. The total current between granules i and j is the sum of the electron and hole currents (here we are neglecting the generation-recombination current that arises in the presence of creation and annihilation of electron-hole pairs, since this current is small in the weak-field regime):

$$I_{ij} = -e[\Gamma^e_{i\to j} - \Gamma^h_{i\to j} - \Gamma^e_{i\leftarrow j} + \Gamma^h_{i\leftarrow j}]. \tag{7}$$

It is easily seen that this current is zero in a state of equilibrium. However, an external electric field disturbs this equilibrium, and in the linear approximation the current between granules can be written in the form [17]

$$I_{ij} = R_{ij}^{-1}(U_j - U_i), \tag{8}$$

where U_i, U_j are the potentials of granules i, j in the external field, and the coefficient R_{ij}, which has the meaning of the electrical resistance of the junction, is given by

$$R_{ij} = \frac{kT}{e[\Gamma^{e0}_{i\to j} + \Gamma^{h0}_{i\to j}]}, \tag{9}$$

where Γ^{e0} and Γ^{h0} are the electron and hole transition rates in the absence of an electric field. Thus, in a weak electric field the problem reduces to one of calculating the conductivity of a network of resistances. We note that the above scheme for calculating the conductivity is a modification of the approach used for describing impurity conductivity in semiconductors [17].

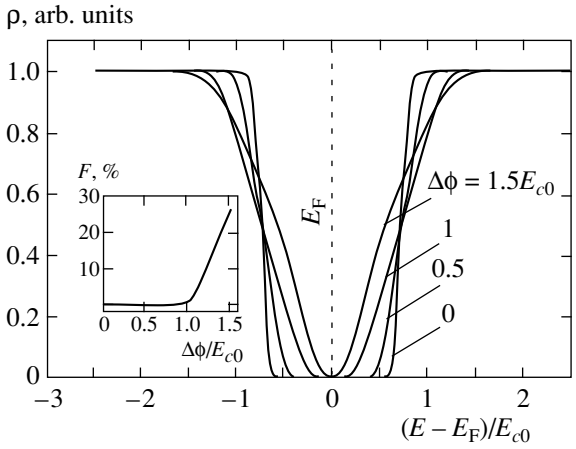

Fig. 3. Calculated density of single-particle excitations for various values of $\Delta\phi$; $\Delta_{min} = 0.5r$, $\epsilon = 5$, metal concentration 20 vol %. Inset: Dependence on $\Delta\phi$ of the fraction of charged spheres in the ground state.

The Ohmic contacts to the composite material were modeled as follows: all the spheres found in a certain layer near one of the faces of the cube were assigned an external applied potential U, while the spheres in a layer near the opposite face were assigned zero potential. The potentials of all the rest of the spheres and the currents between them were found by numerical solution of the system of Kirchhoff's equations, and the total current through the sample was found by summing the elementary currents in one cross section. In that way we have calculated the total conductivity of the system at the given temperature.

Figure 4 shows the temperature dependence of the conductivity of a system of 10^3 spheres; the curves were calculated for the same values of the amplitude of the random potential as for the density of states in Fig. 3. We see that at $\Delta\phi \sim E_{c0}$ the temperature dependence of the conductivity at low temperatures has the form $\ln\sigma \propto T^{-1/2}$. At small $\Delta\phi$ the temperature dependence has an activational character.

Figure 5 illustrates how the temperature dependence of the conductivity of a composite material depends on the variance of the size of the spheres in the absence of a random potential and at an amplitude of the random

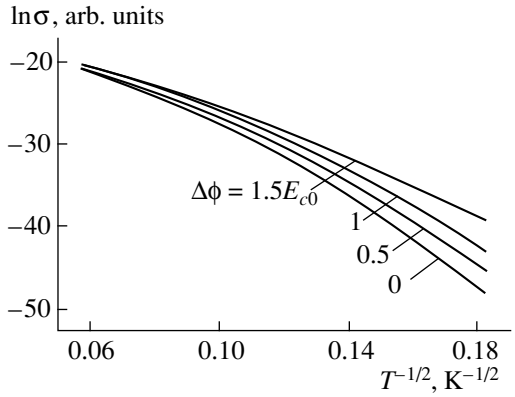

Fig. 4. Calculated temperature dependence of the conductivity for various values of $\Delta\phi$; $\Delta_{min} = 0.5r$, $\epsilon = 5$, metal concentration 20 vol %.

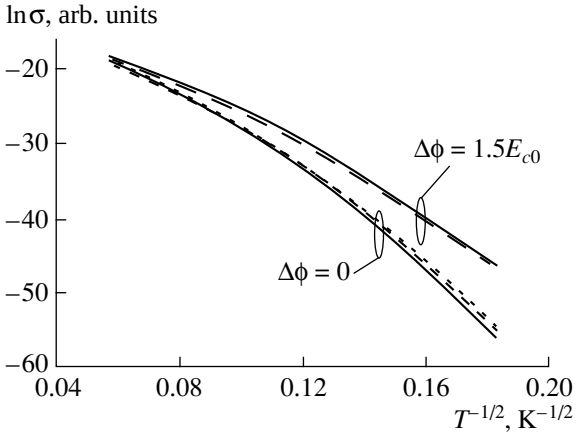

Fig. 5. Calculated temperature dependence of the conductivity for various values of the variance of the size of the spheres (the solid curve is for 0%, the dashed curve for 20%, and the dotted curve for 30%) in the absence of a random potential and for an amplitude of the random potential equal to $1.5E_{c0}$; $\Delta_{min} = 0.5r$, $\epsilon = 5$.

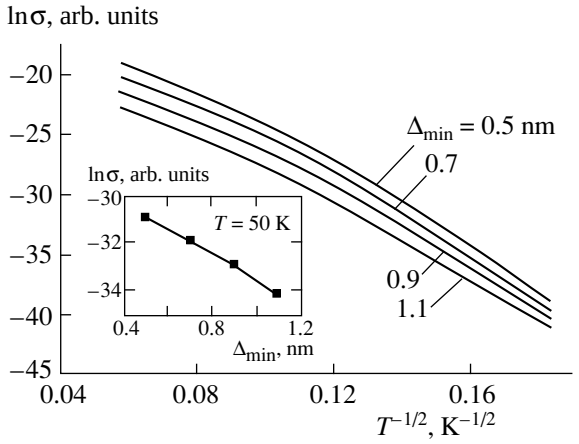

Fig. 6. Calculated temperature dependence of the conductivity for various values of Δ_{min}; $\Delta\phi = 1.5E_{c0}$, $\epsilon = 5$, metal concentration 20 vol %. Inset: plot of $\ln\sigma$ versus Δ_{min} at a fixed temperature.

potential of $1.5E_{c0}$ (here E_{c0} is understood to mean the energy of an isolated sphere of average radius r_0). We see in the figure that an rms deviation of the sphere size of the order of a few tens of percent has a smaller influence on the form of the temperature dependence of the conductivity than does the random potential, and cannot by itself account for the $T^{-1/2}$ law observed experimentally. Since the real scatter of the particle size in the investigated composite films, according to microscopy data, is not over 20–30%, we shall henceforth assume that the system of spheres is monodisperse and devote our attention mainly to the effect of the random potential on the conductivity. We note that in calculating the temperature dependence shown in Fig. 5 we used the approximation of point charges in calculating the off-diagonal elements of the matrix C_{ij}, i.e., the scatter in the size of the spheres was taken into account only in the values of their charge energies E_{ci}. This is because of the great computational difficulties that arise when the polarization is taken into account in a system of metallic particles of different sizes, but one expects that taking the polarization into account will not alter our main conclusion that the variance of granule size has only a weak effect on the form of the temperature dependence of the conductivity.

Figure 6 shows a series of curves of the conductivity versus temperature as calculated for different values of the spacing parameter Δ_{min} at an amplitude of the random potential $\Delta\phi = 1.5E_{c0}$ in the case of zero variance in the size of the spheres. The inset to this figure shows the dependence of the conductivity on Δ_{min} at a fixed temperature. It is seen in the figure that the conductivity depends exponentially on Δ_{min}, while this parameter has only a weak effect on the form of the temperature dependence of the conductivity. The exponential character of the dependence of the conductivity on Δ_{min} is expected according to Eq. (4), where the distance between granules appears in the argument of the exponential function.

4.4. Mesoscopic Scatter of the Conductivity

Since the position of the granules in the structure has a random character and the potential on them is a random quantity, the electrical properties of the structure as a whole can, in principle, fluctuate as one goes from one random realization of the sample to another. To estimate the scale of the mesoscopic fluctuations of the conductivity due to the random character of the medium and to assess its dependence on the macroscopic parameters of the system as a whole, we did a series of calculations of the temperature dependence of the conductivity for samples containing different numbers of spheres. The results of these calculations are presented in Fig. 7 in the form of a plot of the rms deviation of the common logarithm of the conductivity versus the inverse temperature and the size of the sample (the number of spheres). As we see in the figure, for

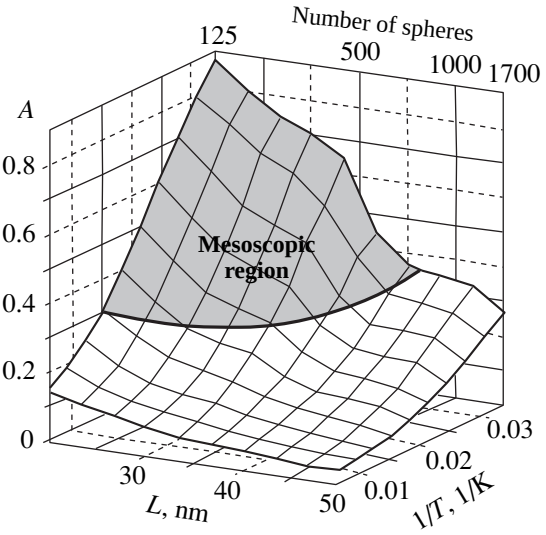

Fig. 7. Amplitude of the mesoscopic scatter of the conductivity as a function of the size of the sample and the temperature.

samples containing ~100 spheres the mesoscopic scatter of the conductivity relative to its value is an order of magnitude at a temperature of 30 K, whereas for samples containing ~1000 spheres this scatter on average is not more than a factor of 1.5–2. At room temperature the mesoscopic scatter of the conductivity is a factor of 1.5–2 for samples containing ~100 spheres and not over 10–20% for samples containing ~1000 spheres.

Figure 8 shows the distribution of the current over the sample, as calculated in the region of large mesoscopic scatter (T = 30 K, 125 spheres) and for the same structure at T = 300 K (the region of small scatter). It is seen that in the first case practically all of the current is concentrated in one chain of several spheres, and, consequently, the electrical conductivity of the sample as a whole is determined by the parameters of this chain. As the temperature increases, current flow over the entire cross section of the sample occurs, and the conductivity of the structure is the result of an averaging of the parameters of all the spheres contained in it.

We note that if the potentials of the spheres in the system are changed in some way, then the position of the conducting chain described above can also change, and that can cause the conductivity of the system as a whole to jump to a new value. Such a redistribution of the potential can be caused, e.g., by an external electric field applied by means of a third (gate) electrode added to the system. It could also be caused by an electric field produced by granules not belonging to the chain (floating gate) if their charge state is changed. This might be the explanation for the memory effect observed in composite materials [18].

5. COMPARISON OF THE CALCULATIONS WITH EXPERIMENTAL DATA

To compare the results of the numerical simulation with the experimental data, model calculations were done in the region of macroscopic conduction, where the conductivity fluctuations are small (see Fig. 7). Furthermore, the calculated temperature dependence was averaged over 10 random realizations. The minimum distance Δ_{min} between granules, the amplitude $\Delta\phi$ of the random potential, and the dielectric constant ϵ of the film were used in these calculations as adjustable parameters, and the variance of the granule size was not taken into account. With these parameters a simulta-

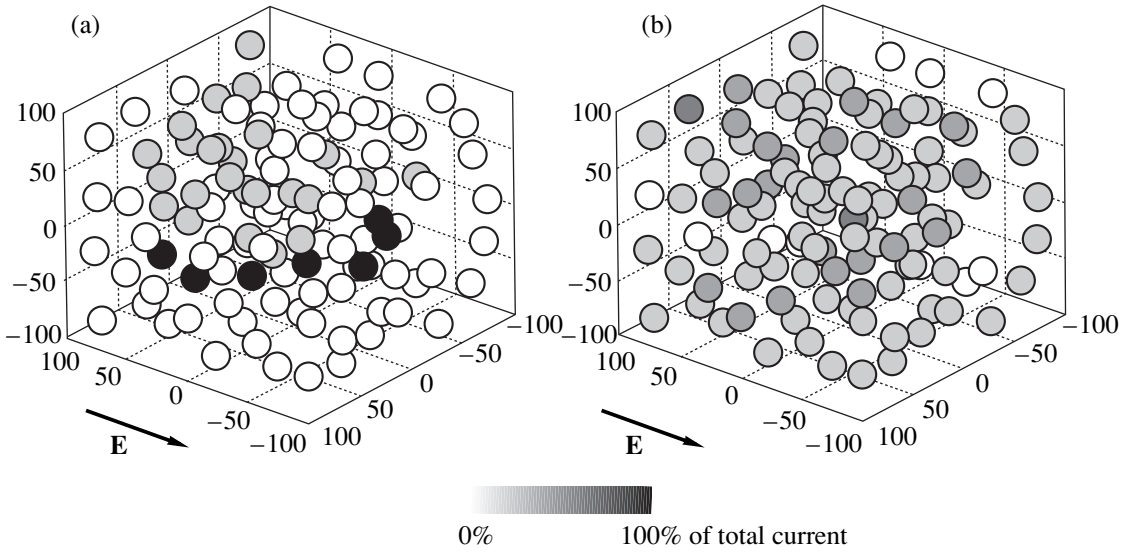

Fig. 8. Distribution of the current over a sample consisting of 125 spheres at temperatures of (a) 30 and (b) 300 K; the field direction is indicated by arrows.

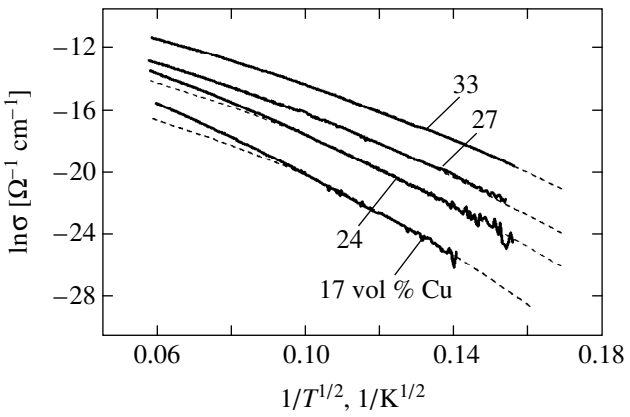

Fig. 9. Results of a fitting of the experimental temperature dependence of the conductivity by the calculated dependence: the solid curves are experimental, the dotted curves theoretical.

neous fitting of both the temperature dependence of the conductivity and the dependence of the conductivity on the volume concentration of Cu was done. The fitting of the temperature dependence of the conductivity yielded the values of $\Delta\phi$ and ϵ, and an independent fitting of the dependence of the conductivity on the concentration of the metal yielded the value of Δ_{min} in each case.

The curves of the experimental and calculated temperature dependence of the conductivity are shown in Fig. 9, and the values of the parameters giving their best fit are listed in the table. As we see in the figure, at high concentrations of the metallic phase in the films there is practically total agreement between the calculated and experimental curves of the temperature dependence of the conductivity, whereas for films with Cu concentrations of 17 and 24 vol % there is a divergence in the high-temperature region. As we see from the table, the amplitude of the random potential $\Delta\phi$ depends weakly on the concentration and is equal in order of magnitude to E_{c0} ($E_{c0} \approx 100$ meV for a sphere of radius 1.5 nm). The minimum distance between granules, Δ_{min}, which is a measure of the "anticorrelation" in the positions of the granules, falls off monotonically as the concentration increases. The values obtained for ϵ are somewhat greater than the characteristic value for pure SiO_2 ($\epsilon_{SiO_2} = 4$). This disagreement is indicative of the presence of a small amount of dissolved metal in the matrix. This explanation is confirmed by the results of separate measurements of the capacitance of capacitors filled with a dielectric in the form of a SiO_2 film having a copper content of 0–5% (at these concentrations the copper does not form clusters). The measured values of the dielectric constant lay in the range 4–5.

The observed divergence of the calculated and experimental curves at high temperatures in the case of low concentrations may be due to the circumstance that in such films the clusters do not have a dense structure and possibly have a broader size distribution. In addition, it follows from the results presented above that as the concentration decreases, there is an increase in the relative fraction of the metal dissolved in the matrix, which may cause the real structure of the material to deviate from the model structure.

6. CONCLUSION

The numerical model proposed in this paper for the conduction of a composite material, based on the concept of single-electron hops between metallic granules of nanometer size, has enabled us to describe the experimentally observed temperature dependence of the conductivity of Cu : SiO_2 films with copper concentrations of 17–33 vol %. Approximation of the experimental data by the calculated results has yielded an estimate of such parameters of the material as the value of the random potential of the granules, the average distance between Cu granules, and the fraction of dissolved metal in the matrix. We have shown that in order to explain the experimentally observed universal low-temperature law $\sigma \propto \exp\{-(T_0/T)^{1/2}\}$ it is necessary to take into account the Coulomb interaction between charged granules and also the presence of a rather large random potential, which brings about charge exchange between the initially neutral granules in the ground state of the system. The deviation from this law at higher temperatures is also satisfactorily described in the framework of the proposed model. We have analyzed the influence of the variance of the granule size on the conductivity of a composite material and found that the existing scatter in the granule size in the experimental samples has a weak effect on the form of the temperature dependence of the conductivity.

We have also carried out a numerical investigation of the dependence of the amplitude of the mesoscopic scatter of the conductivity of composite conductors on their size and temperature. These studies showed that a strong (severalfold) scatter of the conductivity should be expected for conductors less than ~40 nm in size at temperatures ~30 K and for conductors smaller than ~20 nm at room temperature. We have also shown that in the region of strong mesoscopic scatter of the conductivity a percolation channel for the current is formed, i.e., a chain of granules along which more than 90% of the total current flows. The position of this channel can in principle be controlled by applying an external electric field (e.g., using an additional, third (gate) electrode), and that opens up a way of creating devices for nanoelectronics (such as, e.g., a transistor or memory cell) based on composite materials.

ACKNOWLEDGMENTS

This study was supported by the Russian Foundation for Basic Research (project no. 98-02-18210), the program "Physics of Solid-State Nanostructures" of the Ministry of Science of the Russian Federation (projects

no. 97-2014 and 97-1035), and the European Research Office of the US Army (grant no. 68171-98-M-5654).

REFERENCES

1. W. Chen, H. Ahmed, and K. Nakazoto, Appl. Phys. Lett. **66**, 3383 (1995).
2. K. K. Likharev, IEEE Trans. Magn. **23**, 1142 (1987).
3. S. V. Vyshenski, Pis'ma Zh. Éksp. Teor. Fiz. **61**, 105 (1995) [JETP Lett. **61**, 111 (1995)].
4. P. Sheng, Philos. Mag. B **65**, 357 (1992).
5. B. Abeles, P. Sheng, M. D. Coutts, and Y. Arie, Adv. Phys. **24**, 407 (1975).
6. E. Cuevas, M. Ortuño, and J. Ruíz, Phys. Rev. Lett. **71**, 1871 (1993).
7. A. Möbius, M. Richter, and B. Drittler, Phys. Rev. B **45**, 11568 (1992).
8. D. A. Zakgeĭm, I. V. Rozhanskiĭ, and S. A. Gurevich, Pis'ma Zh. Éksp. Teor. Fiz. **70**, 100 (1999) [JETP Lett. **70**, 105 (1999)].
9. A. Möbius, H. Vinzelberg, C. Gladun, *et al.*, J. Phys. C **18**, 3337 (1985).
10. E. Z. Meĭlikhov, Zh. Éksp. Teor. Fiz. **115**, 1484 (1999) [JETP **88**, 819 (1999)].
11. J. Klafter and P. Sheng, J. Phys. C **17**, L93 (1984).
12. A. Pérez-Garrido, M. Ortuño, E. Cuevas, *et al.*, Phys. Rev. B **55**, R8630 (1997).
13. S. A. Gurevich, A. I. Ekimov, I. A. Kudryavtsev, *et al.*, Fiz. Tekh. Poluprovodn. (St. Petersburg) **26**, 102 (1992) [Sov. Phys. Semicond. **26**, 57 (1992)].
14. S. A. Gurevich, T. A. Zaraĭskaya, C. G. Konnikov, *et al.*, Fiz. Tverd. Tela (St. Petersburg) **39**, 1889 (1997) [Phys. Solid State **39**, 1691 (1997)].
15. A. V. Kolobov, H. Oyanagi, S. A. Gurevich, *et al.*, J. Surf. Anal. **3**, 486 (1997).
16. L. D. Landau and E. M. Lifshitz, *Course of Theoretical Physics*, Vol. 8: *Electrodynamics of Continuous Media* (Nauka, Moscow, 1982; Pergamon, New York, 1984).
17. B. I. Shklovskiĭ and A. L. Efros, *Electronic Properties of Doped Semiconductors* (Nauka, Moscow, 1979; Springer-Verlag, New York, 1984).
18. S. A. Gurevich, D. A. Zakheim, V. V. Horenko, *et al.*, in *Proceedings of International Symposium "Nanostructures: Physics and Technology", St.-Petersburg, 1997,* p. 474.

Translation was provided by AIP

Journal of Experimental and Theoretical Physics, Vol. 91, No. 3, 2000, pp. 562–567.
Translated from Zhurnal Éksperimental'noĭ i Teoreticheskoĭ Fiziki, Vol. 118, No. 3, 2000, pp. 647–653.

SOLIDS
Electronic Properties

Thermopower and Resistivity in the Normal State of Mercury Cuprate HTSCs under Pressure

E. S. Itskevich*, V. F. Kraĭdenov, and I. G. Kuzemskaya

Institute of High-Pressure Physics, Russian Academy of Sciences, Troitsk, Moscow oblast, 142092 Russia
**e-mail: itskev@ns.hppi.troitsk.ru*

Received February 16, 2000

Abstract—Nonmonotonic temperature dependence of the thermopower and resistivity of the multilayer mercury cuprate $HgBa_2Ca_4Cu_5O_{12}$ are observed in the temperature range 77–300 K under hydrostatic pressures of up to 10 kbar. The data obtained agree qualitatively with our results for $HgBa_2Ca_2Cu_3O_8$ [V. F. Kraidenov *et al.*, Fiz. Nizk. Temp. 16, 1016 (1990), 20, 76 (1994), 19, 835 (1993) [Sov. J. Low Temp. Phys. 16, 593 (1990), 20, 64 (1994), 19, 597 (1993)]]. Experiments are carried out to check the applicability of the new "correlated polaron" model to the description of the normal state of HTSCs. It is shown that the model proposed by J. B. Goodenough and J. S. Zhou [Phys. Rev. B 49, 4251 (1994), 51, 3104 (1995)] is suitable for use as a working hypothesis for describing research on the temperature dependence of the thermopower, including its behavior under pressure. *© 2000 MAIK "Nauka/Interperiodica".*

1. INTRODUCTION

Two of us (E.S.I. and V.F.K.) have previously investigated the temperature dependence of the thermopower $\alpha(T)$ and resistivity $\rho(T)$ in the Y-123 and Y-124 systems in the temperature range 77–300 K under hydrostatic pressures of up to 15 kbar [1–3]. For Y-123 the measurements were made for different oxygen concentrations ($YBa_2Cu_3O_x$). We have analyzed our results in the framework of the model proposed in [4], which allows one to determine a number of the parameters of the normal state. However, that model cannot answer many questions and, in particular, cannot explain the main feature of $\alpha(T)$—a broad maximum at $T \approx 140$ K or the linear dependence $\rho(T) \propto T$ at $T > T_c$. Most important, the model used in [4] did not take into account the possible mechanisms of high-T_c superconductivity and their relation to the normal state of superconductors with a high superconducting transition temperature T_c.

In papers [5, 6] published in 1994–1995 the phenomenological concept of "correlated polaron" was proposed in order to explain the features of the transport properties of high-T_c superconducting copper oxides. This is a nonuniform charge state that takes place in thermodynamically distinguishable phases with localized and mobile carriers in a variable-valence system. The use of this new concept makes it possible to understand, from a unified point of view, many properties of the normal state of HTSCs, including the existence of stripes, and the transport properties, including the maximum on the $\alpha(T)$ curve at $T = 140$ K. This concept has aroused great interest and is under active development. As was pointed out in [7], "Recently a growing number of experiments have provided indications of the key role of polarons … in doped perovskites…"

In the model proposed in [5] it is assumed that the occurrence of high-T_c superconductivity in copper cuprates can be understood in the framework of the hypothesis that a "thermodynamically distinguishable phase" arises in which the HTSC state is realized. This phase is stable near the crossover of two types of behavior of the carriers. The authors of the model mention three characteristic features of the normal state of this phase: (1) an anomalously large compressibility of the Cu–O bonds in the CuO_2 planes; (2) the existence of mobile stripes with a high concentration of holes; (3) appreciable rise of the thermopower, with a maximum of $\alpha(T)$ in the region $100\ \text{K} < T_{\max} < 150$ K [8].

It is stated in [5, 6] that the existing theories of high-T_c superconductivity, including the Fermi-liquid, Hubbard, weak electron-phonon interaction, and standard polaron models, do not give a complete description of the normal state of a HTSC (see the references cited in [5]); the normal state of the superconducting phase is characterized by a strong dependence of α_0 on doping and by the presence of a maximum of $\delta\alpha$.

The temperature at which the maximum is observed (140 K) suggests a coupling of the holes with optical vibrations of the lattice. There arises a vibronic resonance of two types of bonding (ionic and covalent), which can be characterized as a polaron. There is reason to suppose that on cooling ($T < 300$ K), the polarons in the normal state condense into a polaron liquid. Superconductivity is brought about by a strong electron-lattice interaction, the so-called vibronic stabilization, which is associated with the dynamic Jahn–Teller effect.

1063-7761/00/9103-0562$20.00

From an analysis of the well-known Mott formula describing the thermopower one can say that: (1) the value of the thermopower depends on the details of the carrier dispersion relation $E(k)$ at the Fermi energy E_F; (2) the interaction of the carriers with other excitations changes the sign of the dispersion relation near E_F and causes the values of the width $\delta\alpha$ and temperature T_{max} of the maximum to depend on the interaction energies. It follows, in view of the behavior of $\alpha(T)$ in $La_{2-x}Sr_xCuO_4$ for $x \geq 0.15$, that the superconducting phase should have a dispersion relation $E(k)$ with a substantial nonmonotonicity near E_F. Such a picture for $E(k)$ was discussed in [5] in terms of the polaron-liquid concept. The formation of a gap leads to growth of the effective mass of the carriers.

For $T_c < T < 300$ K in the interval of concentrations x corresponding to the superconducting composition, $\alpha_0(T)$ is strongly dependent on doping and weakly dependent on T. The temperature dependence is mainly determined by the maximum $\delta\alpha$. In the underdoped case there is only the monotonic dependence $\alpha_0(T)$, and for overdoping there is only the maximum $\delta\alpha(T)$.

In the case of heavy doping, perturbations of the periodic potential in the CuO_2 layer due to defects in the apical oxygen sites suppress the thermopower peak and superconductivity [6].

The authors of [5, 6] performed a series of experimental studies expressly to test their theory. They mainly studied model systems based on single-layer La_2CuO_4 with variable concentrations of oxygen and the strontium dopant and also the systems Y-123 and Y-124. They measured $\alpha(T)$ and $\rho(T)$ at temperatures from 4 to 300 K and above. Experiments were also done under hydrostatic pressure in a fixed-pressure low-temperature chamber [9]. The experimental results [6, 10–12] were analyzed in the light of the theory described above, and rather good agreement was demonstrated. Thus another model featuring polarons of a special type (or mobile nonuniform charge states on a mesoscopic scale) has come into being; it explains many known experimental results and to a certain extent ties in with the generally accepted knowledge base for the other models.

Our previous papers [1–3] on yttrium systems showed the same features in $\alpha(T)$ as were later observed in [10, 12]: rather large absolute values of α, a broad maximum on the $\alpha(T)$ curve, dependence of α and ρ on the pressure and the degree of oxidation. It should be noted that there is a detailed quantitative agreement of the $\alpha(T)$ curves obtained under pressure for Y-124 samples not containing nonstoichiometric oxygen. Of course, the data of our studies [1–3] have been refined by the detailed and careful measurements made for the purpose of testing the concept of anomalies in the transport properties [5, 6].

We wish to point out that the value of studying the thermopower over a wide temperature interval and under hydrostatic pressure was correctly foreseen by us at that time. It was clear that the maximum in the temperature dependence of the thermopower, which we observed in all the HTSCs that we studied, derives from some specific properties, and the effect of pressure on this maximum might clarify considerably the interrelationships involved.

Recently we have published the results of our measurements of $\alpha(T)$ and $\rho(T)$ under hydrostatic pressure for samples of the phase Hg-1223 [13] with the optimum oxygen concentration and the highest values of T_c known at the present time. These data were confirmed in [14]. In the present paper we report analogous measurements for the phase Hg-1245, which differs from Hg-1223 by the presence of three types of CuO_2 layers and which has certain features in the pressure dependence $T_c(P)$ [15]. The main features in $\alpha(T)$ are manifested in the region of the normal state, permitting a comparison with the correlated polaron theory for the system of layered mercury cuprate HTSCs. The results are discussed in relation to the experimental data on yttrium HTSCs. The theory of [5, 6] is itself subject to divided opinion, and a detailed comparison with it would be extremely desirable.

2. MEASUREMENTS AND RESULTS

Samples. Ceramic samples of mercury HTSCs were prepared by the technology described in [16]. Our sample consisted 90% of the $HgBa_2Ca_4Cu_5O_{12+\delta}$ phase and 10% of the $HgBa_2Ca_3Cu_4O_{8+\delta}$ phase. The tetragonal lattice parameters were a = 3.849(4) Å and c = 22.151(5) Å.

The sample was in the form of a half cylinder 3 mm in diameter and 3 mm in height, with a transverse cross-sectional area of 3.5 mm^2. The heater and heat sink were mounted to the sample using silver paste, which was then dried for several hours at 60°C.

Thermopower. The measurements were made in the temperature interval 77–300 K under hydrostatic pressure up to 10 kbar. The method of longitudinal heat flow with a constant power delivered to the heater [17] was used. The temperature drop across the sample ranged from 0.1 K at atmospheric pressure to 0.04 K at the highest pressure. The temperature was measured by a (Cu + 0.1%Fe + 0.017%Li)–Cu thermocouple. The thermopower of the sample was determined relative to that of copper. The correction to the thermopower for copper was taken from published data and from our measurements of α_{Cu} relative to the Y-123 superconducting sample. The pressure dependence of α_{Cu} was not taken into account as it would introduce a negligible correction.

The overall trend of the $\alpha(T)$ curve at all pressures follows the typical dependence for the majority of HTSCs (see, e.g., [1, 10]). As the temperature is lowered from 300 K the thermopower α increases linearly at a rate of 7.7×10^{-2} $\mu V/K^2$, passes through a maxi-

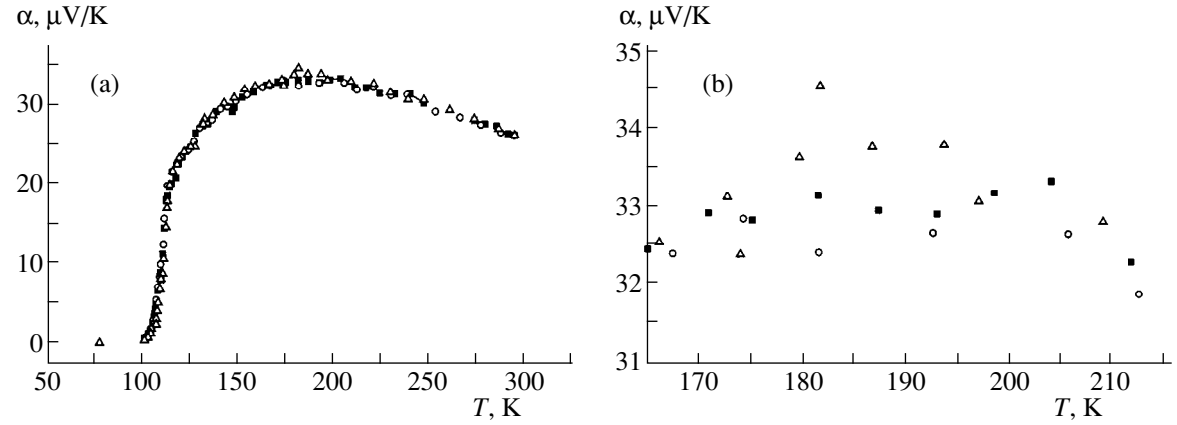

Fig. 1. (a) Temperature dependence of the thermopower for the Hg-1245 sample at different pressures: (○) 4.3 kbar, (■) 6.4 kbar, (△) 9.5 kbar; (b) enlargement of the part of these curves in the interval from 165 to 215 K.

mum at $T \approx 190$ K, and falls off sharply near T_c (Fig. 1). A comparison with the data for single crystals shows that the thermopower of the ceramic corresponds to the contribution from the CuO_2 layers [18].

The insufficient single-phaseness of our samples does not alter the qualitative picture observed in the $\alpha(T)$ curve, and from a quantitative standpoint we estimated the extraneous contribution to be 3–4% of the measured value in the Hg-1245 phase.

The first kink in the region of the superconducting transition occurs at $T \approx 132$–134 K, which indicates that traces of the Hg-1223 phase may be present. From this kink to 113 K the thermopower decreases linearly at a low rate, passing through the superconducting transition of the Hg-1234 phase. At 113 K the sharp superconducting transition of the Hg-1245 phase begins, ending at $T \leq 100$ K. The width of the transition according to the linear part of the drop is approximately 6 K. A linear extrapolation to pressure $P = 0$ by the least-squares method gives a value $T_{c0} = 110$ K. As the pressure is increased, T_c increases at a rate of $dT_c/dP = 0.19$ K/kbar.

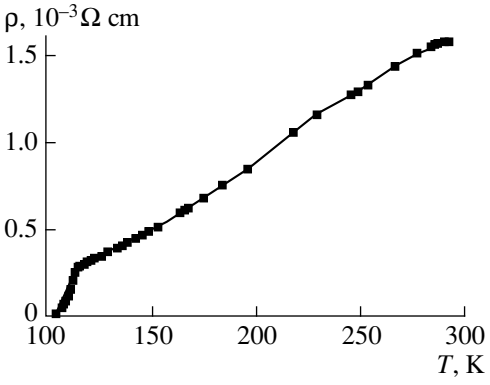

Fig. 2. Plot of the resistivity of the Hg-1245 sample versus temperature at $P = 0$.

In the range from T_c to 300 K the value of α increases very weakly with increasing pressure. For cuprate superconductors the thermopower is usually written in the form

$$\alpha(T) = \alpha_0 + \delta\alpha(T),$$

where $\alpha_0 \approx \alpha_{300}$ is the part of the thermopower that depends weakly on temperature, and $\delta\alpha(T)$ is the part that includes the broad maximum.

In the region of the monotonic behavior of $\alpha(T)$ the thermopower at 300 K is equal to 25 μV/K, and

$$d\alpha_{300}/dP = 2 \times 10^{-3}\ \mu\text{V/K kbar},$$

and near the maximum

$$\frac{d\ln\delta\alpha}{dP} = (4 \pm 1) \times 10^{-3}\ \text{kbar}^{-1}.$$

The behavior of T_{max} under pressure is difficult to estimate. For all the samples from the papers cited above, T_{max} shifts to lower temperatures. Such behavior is most clearly observed in Y-123 and La–Sr–Cu–O [1, 10].

Resistivity. The resistivity ρ was measured by the usual four-contact scheme. The potential leads were made of copper wire 50 μm in diameter, attached by means of silver paste at a distance of 1.2 mm apart. At room temperature and atmospheric pressure $\rho_{300} = 1.55 \times 10^{-3}$ Ω cm. Measurements of ρ under pressure were not made.

The $\rho(T)$ curve is shown in Fig. 2. The resistivity falls off linearly in the temperature interval 290–230 K. Extrapolation to $\rho = 0$ gives a value of 70 K for the T-axis intercept. As the temperature is decreased further to $T = 113$ K the resistivity decreases, the $\rho(T)$ curve being convex downward and close to the extrapolated straight line. The superconducting transition

begins at 113 K and ends at approximately 100 K. The temperature width of the steep linear drop of ρ is ΔT_c = 6 K. The midpoint temperature of the superconducting transition is 110 K, in agreement with the value obtained from the thermopower measurements. The form of the $\rho(T)$ curve is not entirely typical for cuprate HTSCs. A similar dependence has been observed in Hg-1223 [13, 19], although the curve had an upward convexity.

It may be that the two different slopes of the two parts of the $\rho(T)$ curves observed in the Hg-1223 phase [13] are analogous to the results of [1] for Y-124 samples, and the temperature at which the change in slope occurs is correlated with T_{max} for $\delta\alpha$.

3. DISCUSSION OF THE RESULTS

The main difference between our experimental results on the thermopower of mercury HTSCs [13] and those in the present study for the optimally oxidized samples with a high value of T_c is that the maximum of $\delta\alpha$ is not very high and is strongly broadened, and the positive values of $d\alpha/dP$ are an order of magnitude smaller than in the yttrium samples. The maximum $\delta\alpha$ for the Hg-1245 phase has a relative value $\delta\alpha/\alpha_0$ = 0.27, while for Hg-1223 it is $\delta\alpha/\alpha_0$ = 0.55. The position of the maximum, T_{max}, was shifted by approximately 50 K to higher temperatures, and the absolute value of α_{300} increased substantially. For the Hg-1223 phase at T_c^{max} we have $d\ln\delta\alpha/dP = 3.8 \times 10^{-3}$ kbar^{-1}, while for Hg-1245 the derivative $d\ln\delta\alpha/dP = 4.0 \times 10^{-3}$ kbar^{-1} and is almost independent of temperature.

The absolute value of α_0(300 K) for Hg-1223 is approximately the same as for $YBa_2Cu_3O_{6.9}$. The value of T_c agrees with the experimental results [20]. The signs of the derivatives dT_c/dP and $d\delta\alpha/dP$ are positive and correspond to the theory of [5, 6], according to which the value of T_c is determined by the curvature of the dispersion curve $E(k)$ at $E = E_F$, which increases under pressure [10]. The changing gap for $E > E_F$ increases the value of $\delta\alpha(T)$, so that it increases under pressure.

The presence of a maximum of $\delta\alpha$ at $T_{max} \approx 140$ K is attributed in the theory [5, 6] to optical phonons participating in the formation of vibrons. The published values of the vibrational frequencies in CuO_2 correspond approximately to such a binding energy [21]. For Y-124 there is a 431 cm^{-1} mode. In Hg-1223 the value of T_{max} is 158 K [13], while in Hg-1245 this value rises to 195 K, and it is not yet clear which vibrations in the given material can correspond to this temperature. All of the modes investigated in [21] are shifted to higher frequencies under pressure. Granted, the shift is small and, on account of the broadening of the maximum, is hard to observe at low pressures. However, experiments on lightly and optimally doped Y-123 and La–Sr–Cu–O samples have shown that the shift of T_{max} is quite noticeable and is in the opposite direction: $d\ln T_{max}/dP = -1.2 \times 10^{-2}$ kbar^{-1} in Y-123 [1].

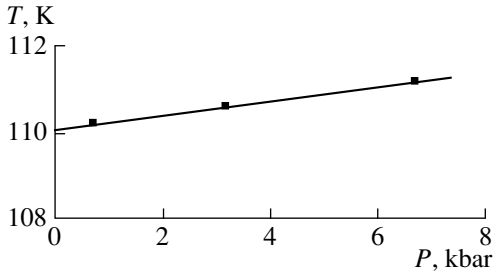

Fig. 3. Pressure dependence of the superconducting transition temperature T_c of the Hg-1245 sample, as estimated from the center of the linear drop of the thermopower $\alpha(T)$.

According to [15, 22], T_c^{max} = 111–113 K for Hg-1245. If these numbers refer, as in our case, to the center and not to the start of the superconducting transition, then our sample is very close to optimally doped (Fig. 3). The empirical relation $\alpha_{300} = f(T_c/T_c^{max})$ for cuprate HTSCs from [22] gives $T_c/T_c^{max} = 1$ and α_{300} = 3–5 μV/K. Our measurements gave a value α_{300} = 25 μV/K, which is six times greater than the expected result. This is all the more surprising because in Hg-1223 one has α_{300} = 5 μV/K [13], which corresponds to the universal dependence. The maximum of $\delta\alpha$ in Hg-1245 is also strongly shifted to higher temperatures. The absolute value of α in its temperature dependence for Hg-1245 is reminiscent of the analogous dependence for the "yttrium" HTSCs with a carrier concentration considerably below δ_{opt}, such as, e.g., in $YBa_2Cu_3O_{7-\delta}$ with $\delta = 0.4$ (T_c = 65.9 K, T_c^{max} = 129 K, α_{max} = 31.7 μV/K, α_{300} = 22 μV/K). In Hg-1245 this may be due to a sharp decrease in the hole concentration in the CuO_2 layer [15]. Such a decrease can give rise to nonuniformity of the potential in the middle CuO_2 layer and is very similar to the data for $La_{2-x}Sr_xCuO_4$ (α_{300} = 25 μV/K) from [6].

Data on the pressure derivative for yttrium samples are given in [10–12]. For both $YBa_2Cu_3O_x$ and $YBa_2Cu_4O_8$ samples there are no major discrepancies with our data [1–3]. The main difference between these samples and purely layered ones lies in the influence of the orthorhombicity (the Cu–O chains), which has been used to explain the behavior of $\alpha(T)$ in $YBa_2Cu_4O_8$ [12].

The shift of T_c^{max} on the $\alpha(T)$ curve for the Hg-1245 phase is very likely due to the vibronic spectrum, which determines T_c^{max} for $\delta\alpha(T)$. The Hg-1245 phase has five CuO_2 layers of three different types, each type having different carrier concentrations [15]. This can complicate the picture set forth in [12] and can lead to a

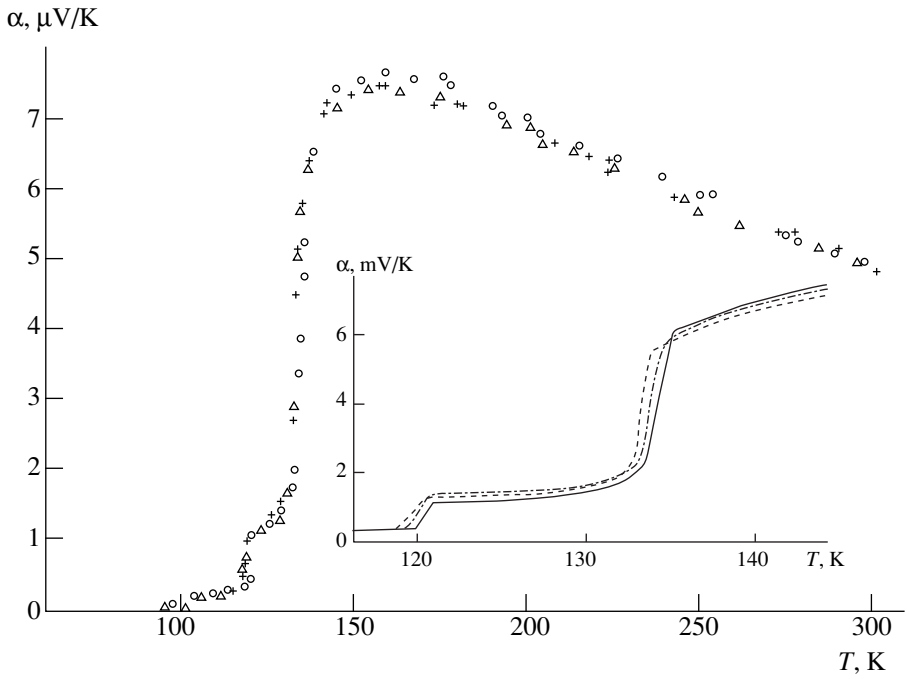

Fig. 4. Temperature dependence of the thermopower $\alpha(T)$ of the Hg-1223 sample at different pressures P: △ and dashed curve, 4.3 kbar; + and the dot-and-dash curve, 6.4 kbar; ○ and the solid curve, 10 kbar.

growth of the frequency ω_{vibr}. This is evidenced by the curves of T_c versus the tetragonal lattice parameter a for samples with different numbers of layers n: the Hg-1245 layer has a smaller parameter a than do phases with smaller n [23]. In mercury HTSCs the parameter a is intimately related to the amount of nonstoichiometric oxygen in the samples and, hence, to the carrier concentration. There can be an additional effect from the defects that exist in all phases of mercury HTSCs and can disrupt the periodic potential of the CuO_2 layer. This last, we believe, has more of an influence on the absolute value of $\delta\alpha(T)$ and its change under pressure. All of the above reasoning in the comparison with theory can also be applied to the behavior of $\rho(T)$ obtained for all our samples at temperatures between T_c and 300 K, where we see a growth $\rho \propto T$.

It should be noted that the absolute values of $\rho(T)$ for Y-HTSCs [1, 9] decrease under pressure (with increase of T_c). For $YBa_2Cu_3O_{6.6}$ samples the same values of the derivative $d\ln\rho/dP = -0.01$ kbar^{-1}were found in [1] and [6].

The signs of the derivatives dT_c/dP and $d\delta\alpha/dP$ should be the same for each sample [5, 6]. However, in experiments on yttrium HTSCs [10] and in our samples this correlation of the signs was not observed ($dT_c/dP > 0$, while $d\delta\alpha/dP < 0$). The signs were the same only for the yttrium sample with concentration $x = 6.7$ at a pressure of 17 kbar in the interval from $T_c = 70$ K to approximately 120 K [10].

It is apropos to mention some specifics of the approach of [5, 6] to the interpretation of the transport properties. It was assumed that the HTSC state arises in a so-called thermodynamically distinguishable phase that can be stable near the crossover at which the character of the carrier behavior changes from localized to collectivized. This occurs through the trapping of holes into clusters of "short" Cu–O bonds: the clusters are thereby stabilized, so that, according to [5], one can speak of a "second electronic phase." In the $Cu–O_2$ planes, clusters consisting of five or six copper ions, which are called "correlated polarons," are stabilized on account of the fact that besides the local lattice deformation there occurs a sharp decrease in the Hubbard effective correlation U in a certain small region of real (coordinate) space. The question of whether or not this formation corresponds to a renormalized quasiparticle cannot be answered in the framework of the phenomenological approach [5, 6]. However, the phenomenological picture is partially confirmed by a proposed [24] microscopic approach to the problem. Mobile polarons in the normal state condense into mobile stripes. In this picture, the presence of stripes brings about a superconducting pairing.

4. CONCLUSION

We have measured the thermopower and resistivity of the Hg-1245 phase as functions of temperature and pressure in the region of the normal state. The qualitative picture of $\alpha(T)$ agrees with the results of our measurements of the Hg-1223 phase. The absolute value of $\alpha(T)$ in the Hg-1245 phase in the normal state is six times as large as in the Hg-1223 phase.

The results described here on the thermopower and resistivity of two phases of the mercury cuprate HTSCs $HgBa_2Ca_{x-1}Cu_xO_{2x+2-\delta}$ (x = 3.5) under hydrostatic pressure over a wide temperature interval of the normal state are in agreement with the model of [5, 6] and allow one to generalize that model to multilayer HTSC systems with high T_c.

We made a comparative analysis of the results which we obtained previously for yttrium HTSCs [1–3] with the implications of the correlated polaron model and the experimental results of the authors of that model for the same HTSCs. It was found that the majority of our results are in rather good agreement with both the model and the experiment done by the authors of that model.

At the same time, there are a number of experimental results that disagree in the two groups of studies and that have not been analyzed in terms of the model of [5, 6]. These include the direction of pressure effects on Y-123 samples, the downward shift of the temperature T_{max} of the broad maximum under pressure, and the high value of α_{300} for Hg-1245.

ACKNOWLEDGMENTS

In closing, we thank A.L. Kuzemsky for a discussion of the theoretical questions touched on in this paper.

This study was supported by the Government Program on HTSC and the Russian Foundation for Basic Research (project no. 00-02-16019).

REFERENCES

1. V. F. Kraĭdenov, E. S. Itskevich, and N. M. Kotov, Fiz. Nizk. Temp. **16**, 1016 (1990) [Sov. J. Low Temp. Phys. **16**, 593 (1990)].
2. V. F. Kraĭdenov and E. S. Itskevich, Fiz. Nizk. Temp. **20**, 76 (1994) [Low Temp. Phys. **20**, 64 (1994)].
3. V. F. Kraĭdenov and E. S. Itskevich, Fiz. Nizk. Temp. **19**, 835 (1993) [Low Temp. Phys. **19**, 597 (1993)].
4. V. É. Gasumyants, G. A. Kaz'min, A. N. Kaĭdanov, *et al.*, Sverkhprovodimost: Fiz., Khim., Tekh. **4**, 1289 (1991).
5. J. B. Goodenough and J. S. Zhou, Phys. Rev. B **49**, 4251 (1994).
6. J. S. Zhou and J. B. Goodenough, Phys. Rev. B **51**, 3104 (1995).
7. A. Lanzara, N. L. Saini, A. Bianconi, *et al.*, cond-mat/9812425.
8. J. S. Zhou and J. B. Goodenough, in *Abstracts of the Second International Conference "Stripes and HTS," Rome, 1998*, WE 1-2.
9. E. S. Itskevich, Prib. Tekh. Éksp., No. 4, 148 (1963); No. 6, 161 (1966).
10. J. S. Zhou and J. B. Goodenough, Phys. Rev. Lett. **77**, 151 (1996).
11. J. B. Goodenough and J. S. Zhou, Phys. Rev. B **53**, R 11 976 (1996).
12. J. S. Zhou and J. B. Goodenough, Phys. Rev. Lett. **77**, 4253 (1996).
13. V. F. Kraĭdenov and E. S. Itskevich, Fiz. Nizk. Temp. **22**, 1028 (1996) [Low Temp. Phys. **22**, 784 (1996)].
14. A. Fukuoka, A. Tokiwa-Yamamoto, M. Itoh, *et al.*, Phys. Rev. B **55**, 6612 (1997).
15. E. S. Itskevich, Fiz. Nizk. Temp. **25**, 685 (1999) [Low Temp. Phys. **25**, 509 (1999)].
16. K. A. Lokshin, I. G. Kuzemskaya, L. F. Kulikova, *et al.*, Physica C (Amsterdam) **279**, 11 (1997).
17. S. L. Bud'ko, A. G. Gapotchenko, E. S. Itskevich, *et al.*, Prib. Tekh. Éksp., No. 5, 189 (1986).
18. Y. Nakamura and S. Uchida, Phys. Rev. B **47**, 8369 (1993).
19. A. Carrington, D. Colson, J. Dumont, *et al.*, Physica C (Amsterdam) **234**, 1 (1994).
20. B. A. Scott, E. Y. Suard, C. C. Tsuel, *et al.*, Physica C (Amsterdam) **230**, 239 (1994).
21. N. Watanabe, M. Kosuge, N. Koshizuka, *et al.*, Phys. Rev. B **49**, 9226 (1994).
22. S. D. Obertelli, J. R. Cooper, and J. L. Tallon, Phys. Rev. B **46**, 14 928 (1992).
23. K. A. Lokshin, D. A. Pavlov, M. L. Kovba, *et al.*, Physica C (Amsterdam) **300**, 71 (1998).
24. G. I. Bersuker and J. B. Goodenough, Physica C (Amsterdam) **274**, 267 (1997).

Translation was provided by AIP

Journal of Experimental and Theoretical Physics, Vol. 91, No. 3, 2000, pp. 568–578.
Translated from Zhurnal Éksperimental'noĭ i Teoreticheskoĭ Fiziki, Vol. 118, No. 3, 2000, pp. 654–665.

SOLIDS
Electronic Properties

Influence of the Electron-Electron Coupling Parameters on the Temperature of the Superconducting Transition in Compounds with the A-15 Structure

R. O. Zaĭtsev[a] and Yu. V. Mikhaĭlova[b]

[a]*Russian Research Centre Kurchatov Institute, Moscow, 123182 Russia*

[b]*State Science Center, Research Institute of Thermal Devices, Moscow, 129085 Russia*
e-mail: zaitsev@mbslab.kiae.ru

Received February 25, 2000

Abstract—The electronic structure of the A-15 compounds is investigated with allowance for the Hubbard energy as the largest energy parameter. The conditions for the onset of Cooper instability are obtained. The dependence of the superconducting transition temperature on the degree of filling of the *d* and *p* shells of the transition elements and nontransition elements is determined. The relation between the transition temperature and the energy of a hop along the chains is established, permitting a comparison with experiment. A qualitative explanation is given for the empirical results of Matthias. © *2000 MAIK "Nauka/Interperiodica".*

1. STATEMENT OF THE PROBLEM

According to BCS theory, which is based on the usual phonon mechanism, the various properties of superconductors are represented by universal relations containing T_c as a scale parameter. There are several properties of superconducting compounds with the A-15 structure which disagree with a phonon mechanism or which are at least atypical for it. These include the everywhere positive curvature of $H_{c2}(T)$, the low-degree power-law behavior of the resistance (T or T^2), and the anomalous values of the dimensionless ratios $2\Delta_0/T_c$ and $2\Delta C/\gamma T_c$ [1]. The impossibility of explaining these facts without using a rather strong electron-electron interaction indicates that the role of the Coulomb interactions is substantial in comparison with the usual electron-phonon interaction [2].

It is even more difficult to explain the dependence of the superconducting transition temperature T_c on the position of the Fermi level. Table 1 gives the values of the superconducting transition temperature for all of the known superconductors with the A-15 structure. Listed for each compound in Table 1 are the average number n_d of *d* electrons per A cation of the transition element, and the average number n_p of *p* electrons per B anion in the investigated compounds A_3B. In the lower part of the table is the dependence of T_c on $(3n_d + n_d')$, where n_d' is the average number of *d* electrons per transition element D in the configurations A_3D. The experimentally observed value of T_c in Kelvin is written in brackets after the corresponding element. Membership in a particular subgroup, corresponding to a given value of $3n_d + n_p$ (or $3n_d + n_d'$) is determined for each individual compound A_3B (or A_3D) on the assumption that the s shell of the A cation is empty and the *s* shell of the B anion is completely filled. As to the number n_d' for the elements D of one of the transition groups, its *s* shell is assumed to be unfilled.

All superconducting compounds of the type A_3B are naturally divided into two groups of eight subgroups, each of which corresponds to a certain number $\mathcal{R} = 2n_d + n_p$, with $14 \le \mathcal{R} \le 28$. For a specified number of electrons per unit cell, $(3n_d + n_p)$ or $(3n_d + n_d')$, one observes anomalies due to the possibility of an increase in T_c on going to a B element with a larger number of nucleons. The change in T_c as a function of the number of electrons per unit cell obeys the relations discovered by Matthias [3]. The superconducting transition temperature has two sharp peaks: at $3n_d + n_p \simeq 16.8$ ($T_c = 20.1$) and at $3n_d + n_d' \simeq 25.7$ ($T_c = 12.7$).

These relations cannot be explained from the standpoint of a purely electron-phonon interaction. However, they can be understood on the basis of the strong-coupling approximation with allowance for the strong electron-electron repulsion within the same atom—the so-called generalized Hubbard–Emery model [4, 5].

Our problem is to investigate the influence of the electron-electron coupling parameters on the superconducting transition temperature in compounds with the A-15 structure for A_3B compounds (with a nontransi-

tion element). In the compounds studied, the distance between atoms of the metal and nonmetal (≈2 Å) is the same as in the CuO_2 layers of high-T_c superconductors.

2. GENERAL RELATIONS

Let us assume that the Hubbard energy is large for both the *p* and *d* electrons. In the ladder approximation the condition for the onset of the Cooper instability is the same as the requirement that there be a nontrivial solution of the system of homogeneous equations for the two-particle vertex part $\Gamma_{\alpha,\beta}$, which is calculated for zero total frequency, momentum, and spin:

$$\Gamma_{\alpha,\beta} = -T\sum_{\omega,\mathbf{p}} g_{\alpha,\beta;\gamma,\delta}(\mathbf{p})G^{\gamma,\lambda}_{-\omega}(-\mathbf{p})G^{\delta,\nu}_{\omega}(\mathbf{p})\Gamma_{\lambda,\nu}. \quad (1)$$

Direct calculations show that the Hubbard energy is the largest energy parameter for both the *d* and *p* electrons, and we therefore treat it as infinite below. The zeroth vertex part $g_{\alpha,\beta;\gamma,\delta}(\mathbf{p})$ is calculated by Dyson's method.

In the simplest case, when the lower Hubbard subband for both the *p* and *d* electrons is filled, we have

$$g_{\alpha,\beta;\lambda,\nu}(\mathbf{p}) = -\delta_{\alpha,\lambda}g_\alpha t^{\beta,\nu}(\mathbf{p}) - \delta_{\beta,\lambda}g_\beta t^{\alpha,\lambda}(\mathbf{p}). \quad (2)$$

The indices α, λ in the odd positions differ only in the sign of the spin projection from the indices β, ν in the even positions. The coefficients g_α and g_β are equal to ±1 or 0 and are determined by the product of the structure constants corresponding to the given transitions [6]. The vertex part turns out to be proportional to the matrix of the hopping integrals $t^{\beta,\nu}(\mathbf{p})$, in terms of which the inverse single-particle Green's function is expressed:

$$[G^{-1}_{\omega}(\mathbf{p})]_{\gamma,\lambda} = \delta_{\gamma,\lambda}(i\omega - \epsilon_\gamma) - f_\gamma b^2_\gamma t^{\gamma,\lambda}(\mathbf{p}). \quad (3)$$

Here ϵ_γ is the energy of the single-particle transitions γ, and f_γ is the so-called end factor, equal to the sum of the initial and final occupation numbers corresponding to the given transition. This last equation corresponds to the zero-loop approximation and will be used in writing the equations of state.

This same equation is conveniently used for calculating the products $t^{\gamma,\lambda}(\mathbf{p})G^{\lambda,\nu}_{\omega}(\mathbf{p})$:

$$\sum_\lambda t^{\gamma,\lambda}(\mathbf{p})G^{\lambda,\nu}_{\omega}(\mathbf{p}) = \frac{1}{b^2_\gamma f_\gamma}(i\omega - \epsilon_\gamma)G^{\gamma,\nu}_{\omega}(\mathbf{p}) - \frac{1}{f_\gamma b^2_\gamma}\delta_{\gamma,\nu}. \quad (4)$$

Table 1

$3n_d + n_p$	A_3B (A is a transition element, B is a nontransition element)
14	Zr_3Sn(0.94), Zr_3Pb(0.76)
15	Ti_3Sb(5.8), Zr_3Sb(0)
16	V_3Al(11.65), V_3Ga(16.5), V_3In(13.9)
16	Nb_3Al(18.55), Nb_3Ga(20.3), Nb_3In(9.2)
16.75	$Nb_3Al_{0.75}Ge_{0.25}$(20.10)
17	V_3Si(17.1), V_3Ge(6.1), V_3Sn(3.8), Nb_3Si(16.5)
17	Nb_3Ge(6.9), Nb_3Sn(18.0), Nb_3Bi(9.6), Ta_3Sn(5.8)
18	V_3As(0), V_3Sb(0.8), V_3Bi(0)
18	Nb_3Sb(1.95), Nb_3Bi(3.05), Ta_3Sb(0.66)
19	Nb_3Te(2.5), Cr_3Ga(0), Mo_3Al(0.58), Mo_3Ga(0.76)
20	Cr_3Si(0), Cr_3Ge(0), Mo_3Si(1.3)
20	Mo_3Ge(1.43), Mo_3Sn(0), W_3Si(0)
22	Mo_3O(4.5), W_3O(0.4–3.35)
24	Cr_3O(0)
$3n_d + n'_d$	A_3D (A and D are transition elements)
21	Ti_3Ir(4.3)
22	Ti_3Pt(0.49), Zr_3Au(0.92)
23	Ti_3Au(<0.015), Zr_3Au(0.92), V_3Os(5), Nb_3Os(0.95)
24	V_3Rh(1.075), V_3Ir(1.7), Nb_3Ir(1.71)
25	V_3Ni(0), V_3Pb(0.082), V_3Pt(3.2), Nb_3Pt(9.8), Ta_3Pt(0.4)
25	Mo_3Tc(14), Mo_3Re(15)
25.7	$Nb_3Pt_{0.3}Au_{0.7}$(12.7)
26	V_3Au(3.14), Nb_3Au(10.8), Ta_3Au(0.51–10)
26	Cr_3Ru(3.3), Cr_3Os(4.24), Mo_3Os(11.76)
27	Cr_3Rh(0.072), Cr_3Ir(0.17), Mo_3Ir(8.5)
28	Cr_3Pt(<0.3), Mo_3Pt(4.5)

Substituting (4) and (2) into the basic Eq. (1) and neglecting the nonlogarithmic terms, we find

$$\Gamma_{\alpha,\beta} = -Tg_\alpha\frac{\epsilon_\alpha}{b^2_\alpha f_\alpha}\sum_{\omega,\mathbf{p}} G^{\alpha,\lambda}_{-\omega}(-\mathbf{p})G^{\beta,\nu}_{-\omega}(\mathbf{p})\Gamma_{\lambda,\nu} - Tg_\beta\frac{\epsilon_\beta}{b^2_\beta f_\beta}\sum_{\omega,\mathbf{p}} G^{\alpha,\lambda}_{-\omega}(-\mathbf{p})G^{\beta,\nu}_{\omega}(\mathbf{p})\Gamma_{\lambda,\nu}. \quad (5)$$

The indices α, λ in the odd positions in the definition of the vertex parts Γ refer to final states with a positive spin projection. The indices β, ν in the even positions in the definition of the vertex parts Γ refer to final states with a negative spin projection. As to the crystal indices (*p*, *d*), the vertex parts $\Gamma_{\alpha,\beta}$ and $\Gamma_{\lambda,\nu}$ in Eqs. (1) and (5) should be assumed diagonal in these indices.

Since the single-particle Green's functions do not depend on the spin indices, we rewrite Eq. (5) in the form

$$\Gamma_n = -2\gamma_n\frac{\epsilon_n}{b^2_n f_n}T\sum_{\omega,\mathbf{p},m} G^{n,m}_{-\omega}(-\mathbf{p})G^{n,m}_{\omega}(\mathbf{p})\Gamma_m, \quad (6)$$

Table 2

Interval	γ_π	f_π	b_π^2	g_π	Interval	γ_p	f_p	b_p^2	g_p
$0 < n_\pi < 1$	1	$1 - 3n_\pi/4$	1	4	$0 < n_p < 1$	1	$1 - 5n_p/6$	1	6
$1 < n_\pi < 2$	3/4	$(2 + n_\pi)/12$	3/2	3	$1 < n_p < 2$	3/2	$(4 - n_p)/18$	3	9
$2 < n_\pi < 3$	–3/4	$(6 - n_\pi)/12$	3/2	4	$2 < n_p < 3$	2/3	$(5n_p - 6)/36$	2	4
$3 < n_\pi < 4$	–1	$(3n_\pi - 8)/12$	1	1	$3 < n_p < 4$	–2/3	$(24 - 5n_p)/36$	2	9
Interval	γ_σ	f_σ	b_σ^2	g_σ	Interval	γ_p	f_p	b_p^2	g_p
$0 < n_\sigma < 1$	1	$1 - n_\sigma/2$	1	2	$4 < n_p < 5$	–3/2	$(n_p - 2)/18$	3	6
$1 < n_\sigma < 2$	–1	$n_\sigma/2$	1	1	$5 < n_p < 6$	–1	$(5n_p - 24)/6$	1	1

where the dimensionless coefficients γ_n and the sums of the squares of the genealogical coefficients b_n^2 are evaluated for each particular group of transitions (see Table 2). In order to integrate over the momentum variable **p**, we decompose all of the matrix elements of each Green's function $G_\pm(\pm\mathbf{p})$ with respect to the normal coordinates:

$$G_\omega^{(dd)}(\mathbf{p}) = \sum_{\lambda=\pm} A^{-\lambda}(\mathbf{p})\frac{1}{i\omega - \xi_\mathbf{p}^\lambda},$$
$$G_\omega^{(pp)}(\mathbf{p}) = \sum_{\lambda=\pm} A^{\lambda}(\mathbf{p})\frac{1}{i\omega - \xi_\mathbf{p}^\lambda}, \quad (7)$$

$$G_\omega^{(dp)}(\mathbf{p}) = \sum_{\lambda=\pm} B^{-\lambda}(\mathbf{p})\frac{1}{i\omega - \xi_\mathbf{p}^\lambda},$$
$$G_\omega^{(pd)}(\mathbf{p}) = \sum_{\lambda=\pm} C^{\lambda}(\mathbf{p})\frac{1}{i\omega - \xi_\mathbf{p}^\lambda}, \quad (8)$$

where the normal coordinates are defined in terms of the matrix elements appearing in the definition of the single-particle Green's function (3):

$$A^\lambda(\mathbf{p}) = \frac{1}{2}\times\left[1 + \mathrm{sgn}(\lambda)\frac{\xi_\mathbf{p}^p - \xi_\mathbf{p}^d}{\sqrt{(\xi_\mathbf{p}^p - \xi_\mathbf{p}^d)^2 + 4f_p b_p^2 f_d b_d^2 |t_\mathbf{p}|^2}}\right], \quad (9)$$

$$B^\lambda(\mathbf{p}) = \mathrm{sgn}(\lambda)\frac{f_d b_d^2 t_\mathbf{p}}{\sqrt{(\xi_\mathbf{p}^p - \xi_\mathbf{p}^d)^2 + 4f_p f_d b_p^2 b_d^2 |t_\mathbf{p}|^2}},$$
$$C^\lambda(\mathbf{p}) = \frac{f_p b_p^2 t_\mathbf{p}^*}{f_d b_d^2 t_\mathbf{p}} B^\lambda(\mathbf{p}). \quad (10)$$

Here we have introduced the notation

$$\xi_\mathbf{p}^{p,d} = \epsilon_{p,d} + f_{p,d} b_{p,d}^2 t_\mathbf{p}^{p,d},$$
$$\xi_\mathbf{p}^\pm = \frac{1}{2}[\xi_\mathbf{p}^p + \xi_\mathbf{p}^d \pm \sqrt{(\xi_\mathbf{p}^p - \xi_\mathbf{p}^d)^2 + 4f_p f_d b_p^2 b_d^2 |t_\mathbf{p}|^2}]. \quad (11)$$

One notices that for specified branch number λ and quasimomentum **p** the coefficients $A^\lambda(\mathbf{p})$, $B^\lambda(\mathbf{p})$, and $C^\lambda(\mathbf{p})$ of the decomposition are related by

$$A^\lambda(\mathbf{p})A^{-\lambda}(\mathbf{p}) = B^\lambda(\mathbf{p})C^\lambda(\mathbf{p}). \quad (12)$$

After substituting decompositions (10) and (11) into Eq. (6) and using relation (12), one can show that the products of four Green's functions vanish from the final condition. The final form of the solvability condition contains only a sum of the diagonal matrix elements:

$$-\sum_{\mathbf{p},\lambda}\left[\gamma_d\frac{\epsilon_d}{b_d^2 f_d}(A^{-\lambda}(\mathbf{p}))^2 + \gamma_p\frac{\epsilon_p}{b_p^2 f_p}(A^\lambda(\mathbf{p}))^2\right] \times\frac{1}{\xi_\mathbf{p}^\lambda}\tanh\left(\frac{\xi_\mathbf{p}^\lambda}{2T}\right) = 1. \quad (13)$$

This equation must be supplemented by the equations of state, which relate the average occupation numbers n_p and n_d with the energy difference $r = \epsilon_p - \epsilon_d$ and the chemical potential $\mu = -(\epsilon_p + \epsilon_d)/2$:

$$n_d = [n_d] + g_{[n_d]+1} b_d^2 f_d \sum_{\mathbf{p},\lambda} A^{-\lambda}(\mathbf{p}) n_F(\xi_\mathbf{p}^{(\lambda)}), \quad (14)$$

$$n_p = [n_p] + g_{[n_p]+1} b_p^2 f_p \sum_{\mathbf{p},\lambda} A^{\lambda}(\mathbf{p}) n_F(\xi_\mathbf{p}^{(\lambda)}), \quad (15)$$

where $[n_{p,d}]$ is the integer part of $n_{p,d}$, $n_F(\xi)$ is the Fermi function, and g_n is the degree of degeneracy of the n-particle states. All of the coefficients specified for each whole-number interval of the n_p and n_d values are given in Table 2.

3. EQUATIONS OF STATE

Let us discuss the phase diagram of the superconducting state in the variables n_d, n_p. According to band calculations [2] the *d* states of the A cations are split by the crystalline field into four sublevels. The lower *xy* state is separated from the (*xz*, *yz*) states by at least 3/2 eV (the π band). In turn, the next, $(3z^2 - r^2)$, state is separated from the $(x^2 - y^2)$ state by 1 eV (the σ band). Then the distance between the (*xz*, *yz*) and $(3z^2 - r^2)$ states is of the same order of magnitude. Thus it is necessary to consider the sequential filling of the *xy*, (*xz*, *yz*), $(3z^2 - r^2)$, and $(x^2 - y^2)$ states with the simultaneous filling of the *p* shells of the B anions of the nontransition element. In the compounds under study the B atoms are found at the corners of the cube, and it is therefore necessary to take into account the threefold degeneracy of the *p* or *t* states and also the twofold degeneracy of the e_g states of the B atoms.

Our problem is to study the influence of the electron-electron coupling parameters on the superconducting transition temperature in compounds with the A-15 structure for compounds A_3B (with a nontransition element). Below we use a simple model in which only the hops to a neighboring site are taken into account. Accordingly, for the description of the compounds under study we have a Hamiltonian of the Emery–Hubbard type:

$$\begin{aligned}\hat{H} = &\sum_{\mathbf{r},\mathbf{r}',\mathbf{r}\neq\mathbf{r}'} [t^{\lambda,\lambda'}_{\mathbf{r},\mathbf{r}'}\hat{d}^{+}_{\mathbf{r},\sigma}(\lambda)\hat{d}_{\mathbf{r}',\sigma}(\lambda') + \text{H.c.}]\\ &+ \sum_{\mathbf{r},\mathbf{r}',\mathbf{r}\neq\mathbf{r}'} [t^{\nu,\nu'}_{\mathbf{r},\mathbf{r}'}\hat{p}^{+}_{\mathbf{r},\sigma}(\nu)\hat{p}_{\mathbf{r}',\sigma}(\nu') + \text{H.c.}]\\ &+ \sum_{\mathbf{r},\mathbf{r}',\mathbf{r}\neq\mathbf{r}'} [t^{\lambda,\nu}_{\mathbf{r},\mathbf{r}'}\hat{d}^{+}_{\mathbf{r},\sigma}(\lambda)\hat{p}_{\mathbf{r}',\sigma}(\nu) + \text{H.c.}]\\ &+ \epsilon_d \sum_{\mathbf{r},\sigma,\lambda} \hat{d}^{+}_{\mathbf{r},\sigma}(\lambda)\hat{d}_{\mathbf{r},\sigma}(\lambda) + \epsilon_p \sum_{\mathbf{r},\sigma,\nu} \hat{p}^{+}_{\mathbf{r},\sigma}(\nu)\hat{p}_{\mathbf{r},\sigma}(\nu).\end{aligned} \tag{16}$$

Here $\hat{d}^{+}_{\mathbf{r},\sigma}(\lambda)$, $\hat{p}^{+}_{\mathbf{r},\sigma}(\nu)$ and $\hat{d}_{\mathbf{r},\sigma}(\lambda)$, $\hat{p}_{\mathbf{r},\sigma}(\nu)$ are creation and annihilation operators for electrons of the *d* and *p* shells, respectively, and $\mu = \{x, y, z\}$ and λ denote three types of *p* states and two t_{2g} states, respectively.

As we see from Table 1, a finite value of the superconducting transition temperature T_c is observed under the condition $n_p + 3n_d \le 14$. Therefore, the filling of the lower *xy* subband, for which $n_d \le 2$ and $n_p \le 6$, is not of interest.

In considering the filling of the degenerate (*xz*, *yz*) subbands we make use of the circumstance that the hopping integral $J \approx 1$ eV along the chain is considerably greater than the energy of hybridization between the *p* and *d* electrons while, of course, remaining smaller than the Hubbard energy. If it is assumed that the Hubbard energy is the largest energy parameter, then the energy of the excitations has the usual form, but with a hopping integral τ that depends on the density of π electrons [4]:

$$\xi_{\mathbf{p}} = 2\tau\cos p_z + \epsilon_d. \tag{17}$$

Here $\tau = J_\pi b_\pi^2 f_\pi$ is the product of the hopping integral and the sum of the squares of the genealogical coefficients b_π^2 and the end factor f_π (b_π^2 and f_π are given in Table 2 for each whole-number interval of n_π).

Let us consider the start of the filling of the π-(*xz*, *yz*) subband: $0 < n_p < 2$. If the Hubbard energy is assumed infinite, then the equation of state can be written for each whole-number interval in $[n_d] < n_d < [n_d + 1]$ as follows:

$$n_d = [n_d] + g_d f_d K_d, \quad K_d = \sum n_{\mathrm{F}}(\xi_{\mathbf{p}}). \tag{18}$$

Here g_d is the degree of degeneracy of the lowest $([n_d] + 1)$-particle state and is also given in Table 2.

One notices that the *p* electrons are strongly hybridized only with the *e* electrons. The hopping energy itself, for hopping between nearest anions, is of the order of $|t| \approx 2 \times 10^{-2}$ Ry = 0.3 eV. The Hubbard energy of the *p* electrons is at least five times as large as the hopping energy $|t|$, so that we can again use the "Hubbard I" approximation [4]:

$$n_p = [n_p] + 3g_p f_p K_p, \quad K_p = \frac{1}{3}\sum n_{\mathrm{F}}(\xi_{\mathbf{p}}). \tag{19}$$

Here the three branches of the spectrum $\xi_{\mathbf{p}}$ depend on an additional parameter, the tetragonal anisotropy parameter *b*:

$$\xi^{(k)}_{\mathbf{p}} = -|t| g_p^2 f_p \left[(1-b)\cos p_k + b\sum_{s=x,y,z}\cos p_s\right]. \tag{20}$$

Here and below $n_{\mathrm{F}}(\xi)$ is the Fermi distribution, and $b < 1$.

At a fixed energy difference $\epsilon_p - \epsilon_d$ the system of equations (18) and (20) determine the average occupation numbers n_d and n_p for the region $0 < n_\pi < 2$, $0 < n_p < 3$. The equations in the remaining regions, $0 < n_\pi < 2$, $3 < n_p < 6$, $2 < n_\pi < 4$, and $0 < n_p < 6$ can be obtained from (14), (15) by the particle-hole symmetry transformation

$$n_\pi \longrightarrow 4 - n_\pi, \quad n_p \longrightarrow 6 - n_p, \quad \epsilon_{p,d} \longrightarrow -\epsilon_{p,d}.$$

According to Eqs. (14) and (15), the equation of state can be written for each whole-number interval n_π as follows:

$$n_\pi = g_\pi f_\pi^{(1)} \sum_{\mathbf{p}} n_{\mathrm{F}}(\xi^{(1)}_{\mathbf{p}}), \quad f_\pi^{(1)} = 1 - \frac{3n_\pi}{4},$$

$$\xi_{\mathbf{p}}^{(1)} = 2f_\pi^{(1)} J_\pi b_\pi^2 \cos p_z + \epsilon_d, \quad 0 < n_\pi < 1,$$
$$n_\pi = 1 + g_\pi f_\pi^{(2)} \sum_{\mathbf{p}} n_{\mathrm{F}}(\xi_{\mathbf{p}}^{(2)}), \quad f_\pi^{(2)} = \frac{2+n_\pi}{12}, \quad (21)$$
$$\xi_{\mathbf{p}}^{(2)} = 2f_\pi^{(2)} J_\pi b_\pi^2 \cos p_z + \epsilon_d, \quad 1 < n_\pi < 2.$$

After the substitution $\epsilon = -\cos\theta$ we obtain the following equation of state for $T = 0$:

$$n_\pi = [n_\pi] + g_\pi f_\pi \frac{\theta}{\pi}. \quad (22)$$

For the p electrons we have

$$n_p = g_p f_p^{(1)} \sum_{\mathbf{p},\lambda} n_{\mathrm{F}}(\xi_{\mathbf{p}}^{(1,\lambda)}),$$
$$f_p^{(1)} = 1 - \frac{5n_p}{6}, \quad 0 < n_p < 1,$$
$$n_p = 1 + g_p f_p^{(2)} \sum_{\mathbf{p},\lambda} n_{\mathrm{F}}(\xi_{\mathbf{p}}^{(2,\lambda)}),$$
$$f_p^{(2)} = \frac{4-n_p}{18}, \quad 1 < n_p < 2,$$
$$n_p = 2 + g_p \times \frac{4}{3} f_p^{(3)} \sum_{\mathbf{p},\lambda} n_{\mathrm{F}}(\xi_{\mathbf{p}}^{(3,\lambda)}), \quad (23)$$
$$f_p^{(3)} = \frac{5n_p - 6}{36}, \quad 2 < n_p < 3,$$
$$\xi_{\mathbf{p}}^{(1)} = b_p^2 f_p \times 2t[\cos p_x + b(\cos p_y + \cos p_z)] + \epsilon_p,$$
$$\xi_{\mathbf{p}}^{(2)} = b_p^2 f_p \times 2t[\cos p_y + b(\cos p_z + \cos p_x)] + \epsilon_p,$$
$$\xi_{\mathbf{p}}^{(3)} = b_p^2 f_p \times 2t[\cos p_z + b(\cos p_y + \cos p_x)] + \epsilon_p.$$

We further assume that $b = 1$, and after the substitution $\epsilon = -\cos(\beta/2)$ we obtain the following equation of state for $T = 0$:

$$n_p = [n_p] + g_p f_p \frac{\beta - \sin\beta}{2\pi}. \quad (24)$$

At a fixed energy difference $\epsilon_p - \epsilon_d$ the system of Eqs. (21) and (22) determines the average occupation numbers n_π and n_p for the region $\{0 < n_\pi < 2, 0 < n_p < 3\}$. The equations in the remaining regions $\{0 < n_\pi < 2, 3 < n_p < 6\}$, $\{2 < n_\pi < 4, 0 < n_p < 6\}$ can be obtained from (21), (22) by means of the particle-hole symmetry transformation

$$n_\pi \longrightarrow 4 - n_\pi, \quad n_p \longrightarrow 6 - n_p, \quad \epsilon_{p,d} \longrightarrow -\epsilon_{p,d}.$$

According to Eq. (13), the contribution to the equation for T_c for each type of excitation is determined independently [6–9]. The equation for T_c has the form

$$1 = S_p\left(1 - g_p f_p \frac{\beta - \sin\beta}{2\pi}\right) + S_d\left(1 - g_\pi f_\pi \frac{\alpha}{\pi}\right), \quad (25)$$

where S_p and S_d are the contributions for the p and d types of excitations, respectively.

In Eq. (25) we have introduced the factors $[1 - g_p f_p(\beta - \sin\beta)/2\pi]$ and $(1 - g_\pi f_\pi \alpha/\pi)$, which take into account the vanishing of the effective value of the hopping upon the complete filling of the band ($\alpha = \pi$ for the π band, and $\beta = 2\pi$ for the p electrons), which corresponds to localization of the electronic excitations upon the complete filling of the lower Hubbard band. We note that $f_\pi(\alpha = \pi) = 1/g_\pi$, $f_p(\beta = 2\pi) = 1/g_p$ (see Table 2).

For the π excitations we have

$$S_d = \frac{\gamma_\pi}{\pi} T \sum_{\omega} \int_{-\pi}^{\pi} d\varphi \frac{J_\pi \cos\varphi}{\omega^2 + (2J_\pi f_\pi b_\pi^2 t \cos\varphi + \epsilon_\pi)^2}, \quad (26)$$

where $\omega = \pi T(2k + 1)$, $k = 0, \pm 1, \pm 2, \pm 3, \ldots$.

For the p excitations we calculate S_p in an analogous way:

$$S_p = \frac{\gamma_p}{3} T \sum_{\omega} \int_{-1}^{1} \rho(\epsilon) d\epsilon \frac{t\epsilon}{\omega^2 + (2J_p f_p b_p^2 \epsilon + \epsilon_p)^2}, \quad (27)$$

where the density of states $\rho(\epsilon)$ in an isotropic model has the form

$$\rho(\epsilon) = \frac{2}{\pi}\sqrt{1 - \epsilon^2}. \quad (28)$$

The dimensionless scattering amplitudes γ_π and γ_p for the excitations in expressions (26) and (27) were calculated by Dyson's method [7], and the genealogical coefficients b_π^2, b_p^2 and the degrees of degeneracy g_π, g_p are given in Table 2.

4. INFLUENCE OF THE p ELECTRONS

It follows from Eq. (25) that the transition temperature is determined by two terms, each of which depends on either the p or the d electrons. To describe the influence of the p electrons, let us consider the variation of S_p as a function of temperature and the degree of filling β of the band. One can calculate S_p as follows. After summation over n and the substitution

$$\mu = -2f_p t b_p^2 \cos\frac{\beta}{2}$$

we obtain

$$S_p = \frac{\gamma_p}{6 f_p b_p^2} \times \int_{-1}^{1} \rho(\epsilon) d\epsilon \frac{\epsilon}{\epsilon + \cos(\beta/2)} \tanh\left[\frac{\epsilon + \cos(\beta/2)}{T/t f_p b_p^2}\right]. \quad (29)$$

Let us assume that $T \ll f_p t b_p^2$. We write the expression for S_p as a sum of two terms:

$$S_p = S_{p0} + S_{p1},$$

where

$$S_{p0} = \frac{\gamma_p}{6 f_p b_p^2} \int_{-1}^{1} \epsilon \rho(\epsilon) d\epsilon \times \left[\frac{\tanh\left[\frac{\epsilon + \cos(\beta/2)}{T/t f_p b_p^2}\right]}{\epsilon + \cos(\beta/2)} - \frac{1}{|\epsilon + \cos(\beta/2)| + T/f_p t b_p^2}\right], \quad (30)$$

$$S_{p1} = \frac{\gamma_p}{6 f_p b_p^2} \int_{-1}^{1} \epsilon \rho(\epsilon) d\epsilon \frac{1}{|\epsilon + \cos(\beta/2)| + T/f_p t b_p^2}. \quad (31)$$

As a function of T the quantity S_{p0} tends toward a finite limit as $T \longrightarrow 0$. In fact, making the substitution

$$x = \frac{\epsilon + \cos(\beta/2)}{T/t f_p b_p^2},$$

we obtain

$$S_{p0} \simeq \lim_{T \to 0} S_{p0}(T) = -\frac{\gamma_p \sin\beta}{3\pi f_p b_p^2} \int_0^\infty dx \left[\frac{\tanh x}{x} - \frac{1}{1+x}\right] = -\frac{\gamma_p \sin\beta}{3\pi f_p b_p^2} \ln\frac{4\pi}{\pi} \simeq -0.819 \frac{\gamma_p \sin\beta}{3\pi f_p b_p^2}, \quad (32)$$

where $\ln\gamma = C \approx 0.577$ (C is Euler's constant).

We also write the expression for S_{p1} for low temperatures to an accuracy up to terms of order $O(1)$. Then

$$S_{p1} \simeq \frac{\gamma_p}{3\pi f_p b_p^2}\left[-\frac{\pi - \beta}{2}\cos\beta + \frac{3}{2}\sin\beta - \sin\beta \ln\frac{2 f_p t \sin^2(\beta/2)}{T}\right], \quad (33)$$

and, ultimately,

$$S_p \simeq \frac{\gamma_p}{6\pi f_p b_p^2}\left[-(\pi - \beta)\cos\beta + 3\sin\beta - 2\sin\beta \ln\frac{8\gamma f_p t \sin^2(\beta/2)}{\pi T}\right]. \quad (34)$$

We see that S_p consists of two parts, one of which (S_p^1) is the factor multiplying the logarithm, while the other (S_p^c) plays the role of a correction to unity in the BCS formula for the transition temperature:

$$S_p = S_p^1 + S_p^c, \quad (35)$$

where

$$S_p^c = F(\beta)\Psi_p,$$

$$F(\beta) \simeq -\frac{\pi - \beta}{2}\cos\beta + 0.681\sin\beta, \quad \Psi_p = \frac{\gamma_p}{3\pi f_p b_p^2},$$

$$S_p^1 = -\Psi_p \sin\beta \ln\frac{2 f_p t \sin^2(\beta/2)}{T}.$$

We note that $F(\beta)$ is an odd function of the argument $\pi - \beta$, $F(0) = -\pi/2$, $F(\pi) = \pi/2$, and for $\beta \approx 0.124$ it has a minimum equal to $F(0.124) \approx -1.5815$. The function $F(\beta)$ does not depend on the number of the subband, i.e., on the character of the dependence of the end factors f_p on α. The value of Ψ_p, on the other hand, is largely determined by the number of the subband. The maximum value of S_p^c depends on the number of the subband and for the first band has the value $x = 0.007$ (for $n_p = 0.97$), while for the second ($1 < n_p < 2$) and third ($2 < n_p < 3$) bands it has the value $x = 0.021$ (for $n_p = 1.899$ and $n_p = 2.725$, respectively).

The coefficient multiplying the logarithm of the temperature is given by

$$S_p^2 = -\frac{\gamma_p \sin\beta}{3\pi f_p b_p^2}.$$

The maximum of S_p^2 in the first band ($0 < n_p < 1$) has the value $x = 0.019$ at $n_p = 0.97$, while in the second ($1 < n_p < 2$) and third ($2 < n_p < 3$) bands the maximum of S_p^2 has the values $x = 0.0589$ at $n_p = 1.806$ and at $n_p = 2.56$, respectively. Therefore, if the d electrons are not taken into account, one would expect that the transition temperature in the first band ($0 < n_p < 1$) is of the order of

$$T_c \approx t \exp\left(-\frac{1}{0.02}\right) \approx 10^{-22} t,$$

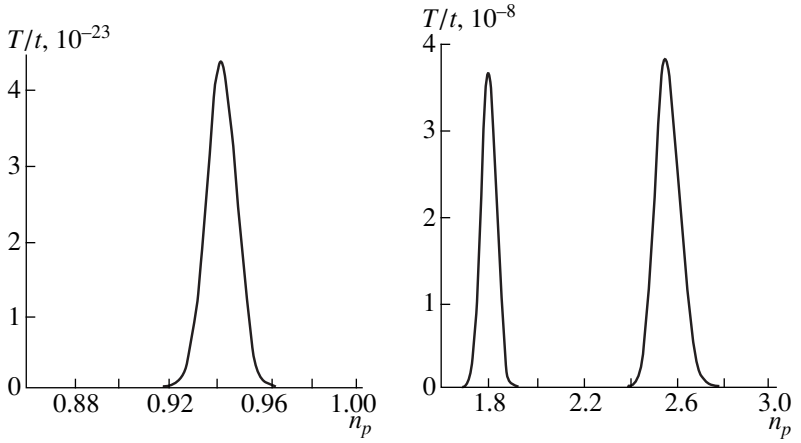

Fig. 1. Dependence of the transition temperature on the concentration for *p* electrons.

while in the second ($1 < n_p < 2$) and third ($2 < n_p < 3$) bands the transition temperature is of the order of

$$T_c \approx t\exp\left(-\frac{1}{0.06}\right) \approx 10^{-7}t.$$

Thus when only the *p* excitations are taken into account, the transition temperature T_c does not exceed $0.5 \times 10^{-7}t$. The dependence of T_c on the *p*-electron concentration n_p is shown in Fig. 1.

5. INFLUENCE OF THE *d* ELECTRONS

The value of S_d can be calculated in an analogous way. After summation over $\omega_n = \pi T(1 + 2n)$ in (6) and the substitution $\mu_\pi = -2f_\pi J_\pi b_\pi^2 \cos\theta$ we have

$$S_d = \frac{\gamma_\pi}{4\pi f_\pi b_\pi^2}\int_{-\pi}^{\pi} d\varphi \frac{\cos\varphi}{\cos\varphi + \cos\theta} \times \tanh\left[\frac{\cos\varphi + \cos\theta}{T/J_\pi f_\pi b_\pi^2}\right]. \tag{36}$$

We shall assume that $T \ll J_\pi f_\pi b_\pi^2$. We write the expression for S_d as a sum of two terms:

$$S_d = S_{d0} + S_{d1},$$

where

$$S_{d0} = \frac{\gamma_\pi}{4\pi f_\pi b_\pi^2}\int_{-\pi}^{\pi} \cos\varphi d\varphi \left[\frac{\tanh\left(\dfrac{\cos\varphi + \cos\theta}{T/f_\pi J_\pi b_\pi^2}\right)}{\cos\varphi + \cos\theta} - \frac{1}{|\cos\varphi + \cos\theta| + T/J_\pi f_\pi b_\pi p^2}\right], \tag{37}$$

$$S_{d1} = \frac{\gamma_\pi}{4\pi f_\pi b_\pi^2} \times \int_{-\pi}^{\pi} \cos\varphi d\varphi \frac{1}{|\cos\varphi + \cos\theta| + T/J_\pi f_\pi b_\pi p^2}. \tag{38}$$

As a function of *T* the quantity S_{d0} tends toward a finite limit as $T \longrightarrow 0$:

$$S_{d0} \simeq \lim_{T\to 0} S_{d0}(T) = -\frac{\gamma_\pi \cot\theta}{\pi f_\pi b_\pi^2}\int_0^\infty dx\left(\frac{\tanh x}{x} - \frac{1}{1+x}\right) = -\frac{\gamma_\pi \cot\theta}{\pi f_\pi b_\pi^2}\ln\frac{4\gamma}{\pi} \simeq -0.819\frac{\gamma_\pi \cot\theta}{4\pi f_\pi b_\pi^2}. \tag{39}$$

For S_{d1} we obtain in a similar way

$$S_{d1} \simeq -\frac{\gamma_\pi \cot\theta}{\pi f_\pi b_\pi^2}\ln\left(\frac{2J_\pi b_\pi^2 f_\pi \sin^2\theta}{T}\right), \tag{40}$$

and, ultimately,

$$S_d = S_{d0} + S_{d1} = \frac{\gamma_\pi}{\pi f_\pi b_\pi^2} \times \left[(\pi - 2\theta) - \cot\theta \ln\left(\frac{8\gamma J_\pi b_\pi^2 f_\pi \sin^2\theta}{\pi T}\right)\right]. \tag{41}$$

We see that S_d consists of two terms, one of which (S_{d1}) determines the factor multiplying the logarithm, while the other (S_{d0}) plays the role of a correction to unity in the BCS formula for the transition temperature. It is easy to see that the factor

$$S_1 \simeq -\frac{\gamma_\pi \cot\theta}{\pi f_\pi b_\pi^2}$$

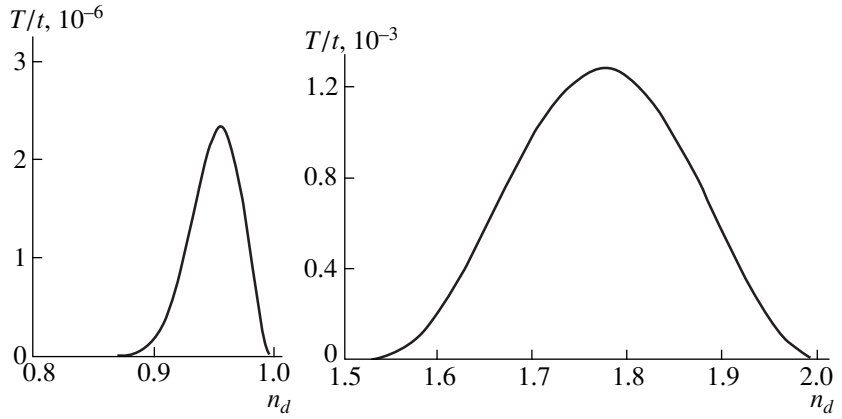

Fig. 2. Dependence of the transition temperature on the concentration for $d(\pi)$ electrons.

multiplying the logarithm is positive for $\theta > \pi/2$, increases monotonically with increasing θ for $\theta > \pi/2$, and at $\theta = \pi$ takes the value $S_1 \approx 0.1$ for $n_d = 1$ and $S_1 \approx 0.2$ for $n_d = 2$. In absolute value S_{d0} is not greater than 0.06 for all values of n_d in the region where $S_1 > 0$. Therefore the transition temperature in the region $n_d > 1$ can exceed the maximum value of the transition temperature for the region $n_d < 1$ by an order of magnitude. According to a numerical calculation, the maximum transition temperature in the region $n_d < 1$ does not exceed $10^{-6}t$, while in the region $n_d > 1$ the maximum transition temperature does not exceed $10^{-3}t$ (Fig. 2).

6. SUPERCONDUCTING TRANSITION TEMPERATURE AS A FUNCTION OF THE OCCUPATION NUMBERS n_p AND n_d

Let us write the resulting equation for T_c. We have

$$1 = \frac{\gamma_p}{6\pi f_p b_p^2}\left[-(\pi-\beta)\cos\beta + 3\sin\beta - 2\sin\beta\ln\frac{8\gamma f_p t\sin^2(\beta/2)}{\pi T}\right]\left(1 - g_p f_p\frac{\beta-\sin\beta}{2\pi}\right) + \frac{\gamma_\pi}{\pi f_\pi b_\pi^2}\left[(\pi-2\theta) - \cot\theta\ln\left(\frac{8\gamma J_\pi b_\pi^2 f_\pi \sin^2\theta}{\pi T}\right)\right] \times\left(1 - g_\pi f_\pi\frac{\theta}{\pi}\right). \quad (42)$$

The superconducting region at $T = 0$ is determined by the condition that the coefficient multiplying the logarithm of the dimensionless temperature vanishes, i.e., by the condition

$$0 = \frac{\gamma_p\sin\beta}{3\pi f_p b_p^2}\left(1 - g_p f_p\frac{\beta-\sin\beta}{2\pi}\right) + \frac{\gamma_\pi\cot\theta}{\pi f_\pi b_\pi^2}\left(1 - g_\pi f_\pi\frac{\theta}{\pi}\right), \quad (43)$$

which is sketched in Fig. 3.

As is seen in Fig. 3, in the regions where $\mu_\pi > 0$ and $\mu_p > 0$ superconductivity always exists for $T = 0$. We note that, although superconductivity is mainly determined by the influence of the π excitations, there exist regions of n_p values in which the superconducting state at $T = 0$ does not arise at any values of n_d. Figure 3 shows the largest possible existence regions of the superconductivity but does not give any idea of the values of the transition temperature. It follows from Eq. (42) and the discussion in Sections 2 and 3 that the superconducting transition temperature T_c is mainly determined by the π excitations, and it is clear that for $3 < n_d < 5$ the values of T_c can reach higher values than

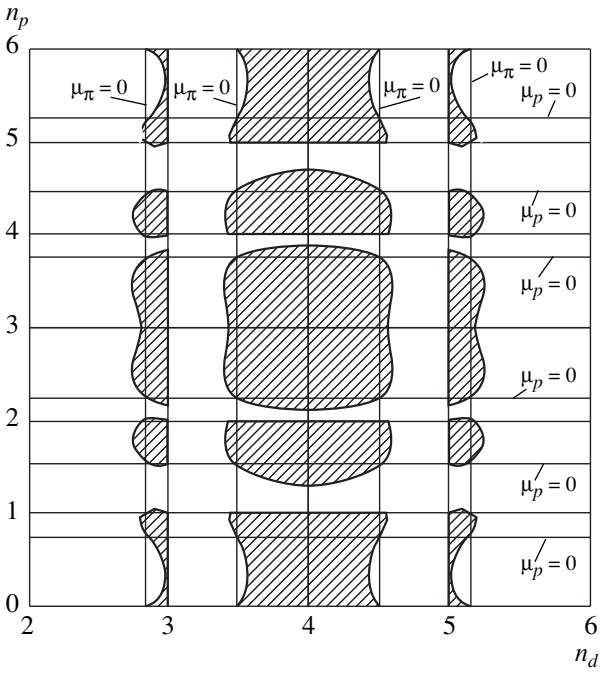

Fig. 3. Phase diagram for the existence of superconductivity at $T = 0$. The superconducting regions are shaded.

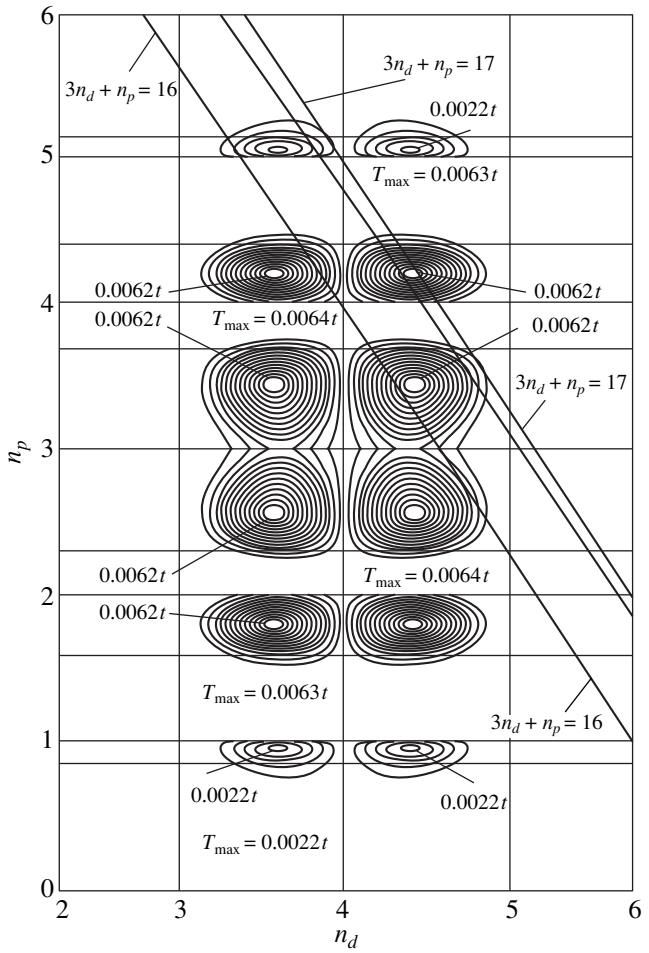

Fig. 4. Transition temperature as a function of the concentration of p and d electrons. Shown are the level lines of the transition temperature, at intervals of $\Delta T_c = 0.0005t$ starting at $0.0002t$. The straight lines are the lines of electrical neutrality of the compounds with the highest values of T_c.

for $2 \le n_d \le 3$ and $5 \le n_d \le 6$. The combined effect of the p excitations and π excitations can lead to an increase in the transition temperature by approximately an order of magnitude. The results of a numerical calculation of the transition temperature as a function of the degree of filling of the bands, i.e., of n_p and n_d, are presented in Fig. 4. The calculation was done under the assumption that the hopping integrals for the p and d electrons are the same. Figure 4 shows level lines of the transition temperature, at intervals of $\Delta T_c = 0.0005t$ starting at $\Delta T_c = 0.0002t$. We note that these transition temperatures are observed only if $3 < n_d < 5$. The highest possible transition temperature obtained in our calculation is $0.0064t$. Figure 4 shows the lines of electrical neutrality corresponding to compounds with the highest experimentally observed transition temperatures: V_3Ga, V_3In, Nb_3Al, Nb_3Ga, $Nb_3Al_{0.2}Ge_{0.8}$, V_3Si, and Nb_3Sn. As one moves along the electrical-neutrality lines $3n_d + n_p = \rho$ on the n_d, n_p phase diagram, the chemical potential is fixed, while the parameter $r = \epsilon_p - \epsilon_d$ formally varies from $-\infty$ to $+\infty$. For all of these compounds the lines of electrical neutrality pass through regions with values of T_c close to the maximum possible value.

7. FEATURES OF THE ONE-DIMENSIONAL MOTION ALONG THE CHAINS

It can be noted that in addition to the Cooper singularity our model also has the Peierls instability, which arises on account of the logarithmic integration of the diagram of the zero-sound type (see Fig. 5c).

At zero temperature and zero energy transfer the polarization operator depends only on the momentum

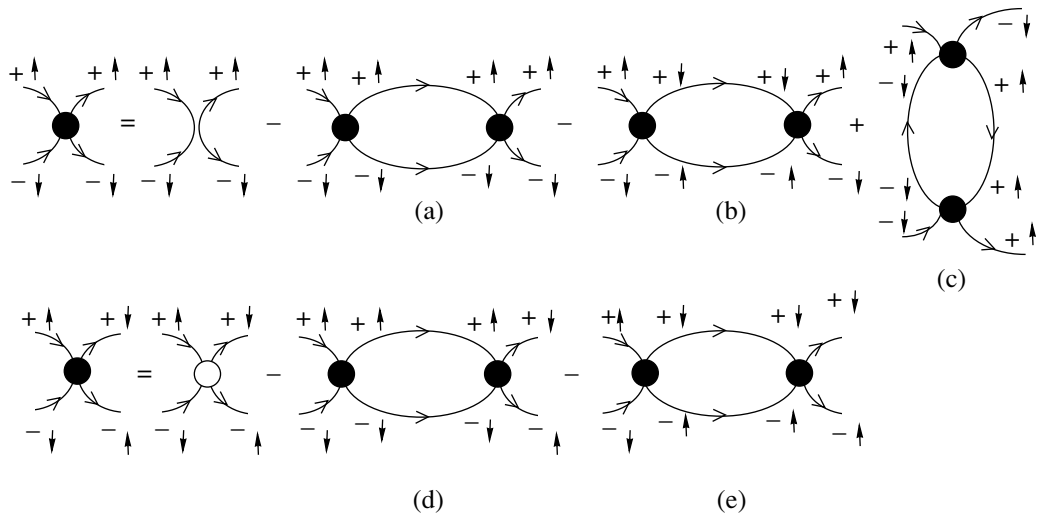

Fig. 5. Sudakov equations. The plus and minus signs denote motion parallel and antiparallel to the x-axis. The arrows ↑ and ↓ indicate the directions of the spin projection.

transfer q and the position of the Fermi level $\mu_\pi = 2f_\pi J_\pi b_\pi^2 \cos\theta$:

$$\Pi_0(q) = -\frac{1}{4\pi f_d J_\pi b_\pi^2 \sin(q/2)} \ln\left|\frac{\sin(\theta + q/2)}{\sin(\theta - q/2)}\right|. \quad (44)$$

It follows that the logarithmic divergence arises under the condition

$$q = 2\arccos\frac{\mu_\pi}{2 f_d J_\pi b_\pi^2},$$

i.e., at a momentum transfer equal to twice the Fermi momentum.

If we ignore the additional features due to umklapp processes at $q = \pi$ or $\mu_\pi = 0$, then taking the parquet diagrams into account will lead to the well-known Sudakov equations [10], which for one-dimensional systems and low Hubbard energies were obtained by Dzyaloshinskiĭ and Larkin [11].

In our case, "backscattering" occurs only for excitations with opposite spin projections. For this reason it is sufficient to consider two amplitudes: one of them, Γ_1, corresponds to any collision with no change in the spin direction, while the other, Γ_2, is the amplitude for "backscattering" with a change in the sign of the spin projection:

$$\begin{aligned}\Gamma_1(\tau) &= g_1 - \int_0^\tau [\Gamma_1^2(\eta) + \Gamma_2^2(\eta)]d\eta + \kappa\int_0^\tau \Gamma_1^2(\eta)d\eta,\\ \Gamma_2(\tau) &= g_2 - 2\int_0^\tau \Gamma_1(\eta)\Gamma_2(\eta)d\eta.\end{aligned} \quad (45)$$

Here

$$\tau = \ln\frac{\epsilon_F}{\max(T, pv_F)}$$

is a logarithmic variable that depends on the Fermi energy ϵ_F and on the temperature T or the product of the momentum p times the Fermi velocity v_F. Under the condition of completely nesting $\kappa = 1$, so that the contribution of the Cooper and zero-sound diagram cancel each other halfway. For $\kappa = 0$ we obtain the ladder equations that were used in the present study. Solving the equations in the two limiting cases, we establish the connection between the transition temperature in the case of complete nesting and the value which was obtained in the previous sections.

Differentiating the equations with respect to the logarithmic variable τ, we obtain the following system:

$$\begin{aligned}\dot\Gamma_1(\tau) &= -\Gamma_1^2(\tau) - \Gamma_2^2(\tau) + \kappa\Gamma_1^2(\tau),\\ \dot\Gamma_2(\tau) &= -2\Gamma_1(\tau)\Gamma_2(\tau).\end{aligned} \quad (46)$$

For $\kappa = 1$ the system has a solution that satisfies the condition $2g_1^2 - g_2^2 = g^2 > 0$:

$$\begin{aligned}\Gamma_1(\tau) &= \frac{|g|}{\sqrt{2}}\coth[(\tau - \tau_0)|g|\sqrt{2}],\\ \Gamma_2(\tau) &= \pm\frac{|g|}{\sinh[(\tau - \tau_0)|g|\sqrt{2}]}.\end{aligned} \quad (47)$$

It can be noted that the zeroth vertex part does not depend on the momentum direction: $g_1 = g_2 = g$. From this we find the relation between the bare vertex part $|g|$ and the arbitrary constant τ_0:

$$|g|\tau_0\sqrt{2} = \operatorname{arctanh}\sqrt{2} \approx 0.8815. \quad (48)$$

The transition temperature in the case of complete nesting is found from the condition $\tau = \tau_0$:

$$T_{c2} = \bar\epsilon e^{-\tau_0} = \bar\epsilon\exp\left(-\frac{\operatorname{arctanh}\sqrt{2}}{|g|\sqrt{a}}\right). \quad (49)$$

To solve the system of differential equations for $\kappa = 0$ it is sufficient to multiply these equations together. The solution of the resulting equation has the familiar ladder form:

$$\Gamma = \Gamma_1 + \Gamma_2 = \frac{2|g|}{1 - 2|g|\tau}. \quad (50)$$

From this equation we find the transition temperature as the condition for a singularity to appear at $\tau = 1/(2|g|)$:

$$T_{c1} = \bar\epsilon\exp\left(-\frac{1}{2|g|}\right). \quad (51)$$

Thus the transition temperature in the presence of nesting is related to the superconducting transition temperature in the absence of nesting as

$$\begin{aligned}T_{c2} &= \bar\epsilon\left(\frac{T_{c1}}{\bar\epsilon}\right)^{g_1/g_2},\\ \frac{g_1}{g_2} &= \frac{2\operatorname{arctanh}\sqrt{2}}{\sqrt{2}} \approx 1.246.\end{aligned} \quad (52)$$

As expected, the presence of nesting, which leads to the Peierls instability, lowers the temperature of the superconducting transition. If the system contains small amounts of impurities, then the zero-sound anomaly due to nesting vanishes, while the Cooper instability does not vanish, since here $T_{c2} \approx T_{c1}$. The value of $\bar\epsilon$ can be obtained formally by proceeding from the ladder approximation, where the calculations in the absence of nesting are done to a higher accuracy than in the parquet approximation. Thus the transition temperatures found in the present study for compounds with the A-15 structure are in good agreement with the experimentally observed values.

8. CONCLUSIONS

In summary, the proposed mechanism gives a qualitatively correct explanation for the empirical Matthias rules, which are reflected in the existence of a sharp and nonmonotonic dependence of the transition temperature on the average number of electrons in incompletely filled shells. The onset of the Cooper instability at a fixed position of the Fermi level is due to the possibility of a change in sign of the scattering amplitude for the entire Fermi surface at once. A calculation of the scattering amplitude for all values of the Fermi energy allows one to find the dependence of the superconducting transition temperature for the entire concentration region within which the Cooper instability can exist. The change in sign of the scattering amplitude occurs once and only once between the center and the boundaries of the Brillouin zone. Accordingly, boundaries of the superconducting region arises within each whole-number interval of variation of n_p and n_d. Thus one can explain the cause of the nonmonotonic dependence of the superconducting transition temperature on the number of conduction electrons. The numerical values obtained for the highest possible transition temperature are somewhat higher than those observed in experiment. However, taking into account the features due to the motion along the one-dimensional chains lowers the transition temperature and results in quantitative agreement with experiment.

REFERENCES

1. B. Brandow, Phys. Rep. **296**, 1 (1998).
2. M. Veger and I. Goldberg, Solid State Phys. **28**, 2 (1973).
3. B. T. Matthias, T. H. Geballe, and V. B. Compton, Rev. Mod. Phys. **35**, 1 (1963).
4. D. J. Hubbard, Proc. R. Soc. London, Ser. A **276**, 238 (1963).
5. V. J. Emery, Phys. Rev. Lett. **58**, 2794 (1987).
6. R. O. Zaĭtsev and Yu. V. Mikhaĭlova, Zh. Éksp. Teor. Fiz. **109**, 1859 (1996) [JETP **82**, 1001 (1996)].
7. R. O. Zaĭtsev, Phys. Lett. A **134**, 199 (1988).
8. R. O. Zaĭtsev, Fiz. Tverd. Tela (Leningrad) **29**, 1631 (1987).
9. R. O. Zaĭtsev, Solid State Commun. **76**, 795 (1990).
10. V. V. Sudakov, I. T. Dyatlov, and K. A. Ter-Martirosyan, Zh. Éksp. Teor. Fiz. **32**, 677 (1957) [Sov. Phys. JETP **5**, 560 (1957)].
11. I. E. Dzyaloshinskiĭ and A. I. Larkin, Zh. Éksp. Teor. Fiz. **61**, 791 (1971) [Sov. Phys. JETP **34**, 422 (1971)].

Translation was provided by AIP

Journal of Experimental and Theoretical Physics, Vol. 91, No. 3, 2000, pp. 579–587.
Translated from Zhurnal Éksperimental'noĭ i Teoreticheskoĭ Fiziki, Vol. 118, No. 3, 2000, pp. 666–675.

SOLIDS
Electronic Properties

Negative *U* Centers, Percolation, and the Insulator-Metal Transition in High-T_c Superconductors

K. V. Mitsen* and O. M. Ivanenko

Lebedev Physical Institute, Russian Academy of Sciences, Moscow, 117924 Russia
**e-mail: mitsen@csi.lebedev.ru*

Received March 1, 2000

Abstract—The mechanism of formation of $-U$ centers in high-T_c superconductors (HTSCs) is considered. It is shown that the transition from the insulator to the metallic state on doping passes through a certain range of dopant concentrations in which it becomes possible for local transitions of singlet electron pairs to occur from oxygen ions to two neighboring cations (a $-U$ center), while single-electron transitions are still forbidden. Conduction arises in such systems at a concentration of $-U$ centers exceeding the percolation threshold for the orbitals of singlet hole pairs. A phase diagram constructed on the basis of the proposed model for the HTSC compounds of the Ln-214 class is in complete agreement with experiment. The mechanisms of formation and relaxation of free hole carriers are considered. It is shown that a distinctive feature of the normal state of HTSCs is the dominant contribution of electron–electron scattering to the charge carrier relaxation processes. It is concluded from the analysis presented that HTSCs comprise a special class of solids in which a nonstandard mechanism of superconductivity, different from the BCS mechanism, is realized. © *2000 MAIK "Nauka/Interperiodica".*

1. INTRODUCTION

In the 13 years since the discovery of high-T_c superconductivity [1] there have been many models proposed (see the review [2]) that might explain the nature of the ground state and the anomalous properties of these compounds. However, the lack of any decisive experiment makes it impossible to decide in favor of any one of them.

In the present paper we show that the mechanism responsible for the many anomalous properties of these compounds (including the high-T_c superconductivity) is apparently the interaction of the electrons with the so-called $-U$ centers [3]. To do this, we consider how the insulator-metal transition occurs in HTSCs upon doping. We show in the framework of a simple ionic model that this transition must pass through a certain range of dopant concentrations in which it becomes possible, within individual microclusters consisting of several unit cells, for local transitions of singlet electron pairs to occur from oxygen ions to a pair of neighboring cations (a $-U$ center), while the single-electron transitions are still forbidden. We believe that it is this range of concentrations that corresponds to the HTSC region, in which the electron–electron attraction is due to the interaction of the electron pairs with $-U$ centers [4–11]. Conduction in such a system arises at a concentration of $-U$ centers exceeding the percolation threshold for the orbitals of the singlet hole pairs. We consider the questions of which parts of the crystal structure form the $-U$ centers and in what range of concentrations does an infinite percolating cluster connecting the $-U$ centers exist, and on these considerations we construct the phase diagram of Ln-214 compounds. Comparison of the resulting phase diagram with the phase diagrams of the Ln-214 compounds that have been studied in detail should, we believe, be the decisive experiment for choosing the mechanism responsible for the unusual properties of HTSCs. We conclude with a discussion of the formation and relaxation of hole carriers in HTSCs.

2. MECHANISM OF FORMATION OF $-U$ CENTERS IN HTSCs

There are some rather weighty grounds for assuming that the electron spectrum of the insulator phases of the various HTSC compounds in the neighborhood of E_F can be best approximated by the model of a charge-transfer (CT) insulator [12], i.e., an insulator with a gap due to charge transfer. In such a model (Fig. 1a) the upper, empty band, formed by the unfilled orbitals of the cations, is separated by a gap from the $2p$ valence band, formed mainly by oxygen states. The gap Δ_{CT} in the spectrum is due to the transfer of an electron from the oxygen to a neighboring cation and has a value of 1.5–2 eV for all HTSCs [13].

What is the mechanism for the insulator–metal phase transition in doped HTSCs? As an example, let us consider the HTSC compounds of the Ln-214 type. For them the value of Δ_{CT} in the framework of a simple ionic model is given by a relation among three quantities [14]:

$$\Delta_{CT} \approx |\Delta E_M| + A_p - I_d,$$

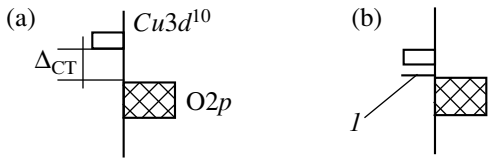

Fig. 1. (a) Electron spectrum of a CT insulator in the vicinity of E_F; (b) modification of the electron spectrum of a CT insulator upon doping; *1* is pair level of a $-U$ center.

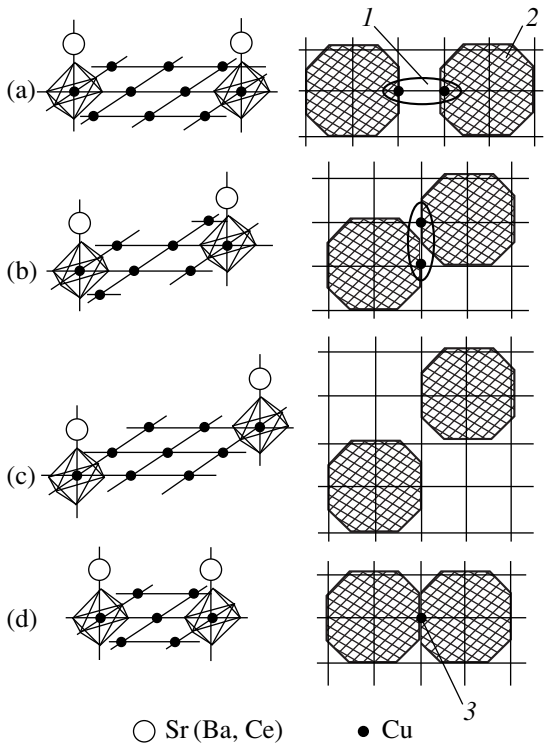

Fig. 2. Different types of $M_2Cu_2O_n$ clusters in cuprates. On the left are fragments of the crystal structure, and on the right are the corresponding projections on the CuO_2 plane. The Sr (Ba, Ce) ions can be located on either side of the CuO_2 plane. In $Nd_{2-x}Ce_xCuO_4$ the apical oxygen ions are absent. (a, b) Two types of clusters forming $-U$ centers at the inner Cu ions (M = Cu for Ln-214); (c) a cluster "intermediate" between a and b, for which a $-U$ center does not form; it corresponds to an insulating phase; (d) an M_2CuO_n cluster on which a normal-phase nucleus forms: (*1*) $-U$ centers; (*2*) localization regions of the doped carriers; (*3*) a "nucleus" of the metallic phase.

where I_d is the second ionization potential of copper, A_p is the electronegativity of oxygen in relation to the formation of O^{2-}, and $|\Delta E_M|$ is the difference between the electrostatic Madelung energies of the configuration in which the oxygen and copper ions are found in the states Cu^{2+} and O^{2-} and the configuration in which they are in the states Cu^{1+} and O^-. Since $I_d \sim 20$ eV and $\Delta_{CT} \sim 1.5$–2 eV, the balance between these three quantities is rather delicate. It can be altered by heterovalent doping, e.g., the doping of La_2CuO_4 by divalent Sr or of Nd_2CuO_4 by tetravalent Ce. The doped carriers (in any case at small x) are localized [15–17] near the dopant ion: either on the oxygen orbitals (as holes in $La_{2-x}Sr_xCuO_4$) or on the copper orbitals (as electrons in $Nd_{2-x}Ce_xCuO_4$). Here it important to note that the admixture of both electrons (on the Cu orbitals) and holes (on the O orbitals) will lead to the same result: a decrease of $|\Delta E_M|$ and, hence, a lowering of Δ_{CT} for the other pairs of copper and oxygen ions found in the neighborhood of the doped carrier. At a certain critical concentration x_c the gap Δ_{CT} vanishes throughout the entire crystal. Thus it becomes possible for electron transitions to occur from the oxygen to the copper, and the substance is transformed into an ordinary metal.

This is the general picture of the transition from a CT insulator to a metallic state on doping in the framework of the ionic model. However, we suppose that in HTSC compounds the transition from the insulator to metal on increasing x first passes through a special stage or, more precisely, through a concentration region $x_0 < x < x_c$ in which two-electron transitions can occur from the oxygen ions to certain pairs of neighboring cations, while the single-electron transitions are still forbidden.

Let us consider a cluster $Cu_2M_2O_n$, where the Cu ions are "built into" the CuO_2 plane and where M = Cu in the CuO_2 plane for Ln-214, M = Cu (in chains) for $YBa_2Cu_3O_7$, and M = Bi for Bi-2212 and Bi-2223. The condition for the formation of $-U$ centers at the Cu ions in the CuO_2 plane is the presence of a localized doped carrier in the neighborhood of each M ion (in YBCO and BSCCO the doped carriers are localized in the CuO_3 chains and in the BiO planes, respectively). In Ln-214 there are two possible types of such clusters (Fig. 2), in which the projections of the dopant ion on the CuO_2 plane are at distances of either $3a$ or $a\sqrt{5}$, where a is the lattice constant in the CuO_2 plane. In both of these cases the presence of a doped carrier in the neighborhood of each M ion decreases Δ_{CT} for the neighboring Cu ion and creates conditions (i.e., forms a local minimum of the potential energy) for the simultaneous transition of two electrons to the internal Cu ions from the oxygen ions surrounding this pair. We note immediately that in the intermediate case, when the M ions are found at a distance $a\sqrt{8}$, pairs of neighboring Cu ions such that doped carriers would be localized in the neighborhoods of the adjacent cations do not appear, and a $-U$ center is not formed.

The lowering of Δ_{CT} for a given Cu ion owing to the presence of a single hole around the neighboring Cu ion can be estimated by taking into account the interaction between nearest neighbors only and the fact that this hole is "distributed" (Fig. 3) over twelve nearest oxygen ions (the first and second coordination spheres).[1]

[1] This assumption agrees with the experimentally determined solubility limit of the dopant (see below).

Then the lowering of the energy of the $3d^{10}$ state for the given Cu ion is $\Delta E \approx (1/4)e^2/r \sim 1.8$ eV. Here it is assumed that on three of the twelve oxygen ions the hole "sees" the unscreened Cu ion located at a distance $r = a/2 \approx 2$ Å from them (e is the charge of an electron and a is the lattice constant in the CuO_2 plane). Thus, on account of the doping, the energy of these states is below the bottom of the conduction band by $\Delta E \approx 1.8$ eV, which is approximately 0.1–0.2 eV smaller than $\Delta_{CT} \approx 1.9$–2.0 eV for La_2CuO_4.

An additional lowering of the energy of the $3d^{10}$ state of Cu is achieved on account of the formation of a bound state of two electrons on neighboring Cu ions in the presence of two holes in the nearest-neighbor environment of this pair. Such a lowering of the energy can occur for the bonding orbital of a singlet hole pair, as takes place in the H_2 molecule. Here this analogy is more appropriate, since the distance between electrons on Cu ions is approximately 3.8 Å and is close to the quantity $R_0\varepsilon_\infty \sim 3.6$ Å, where $R_0 \approx 0.8$ Å is the distance between nuclei in the H_2 molecule, and $\varepsilon_\infty \approx 4.5$ [18] is the high-frequency dielectric constant. Therefore the additional lowering of the density ΔE_U on account of the transition of two electrons to neighboring copper ions in our case can be estimated from the relation $\Delta E_U \sim \Delta E_{H_2}/\varepsilon_\infty^2 \approx 0.23$ eV, where $\Delta E_{H_2} = 4.75$ eV is the binding energy in the H_2 molecule. This estimate, however, is too low, since the oxygen ions effectively screen the repulsive interaction of the electrons on Cu and weakly screen the electron-hole attraction.

Thus it can be assumed that Δ_{CT}, which has a value of to 1.5–2.0 eV for doped cuprates, is depressed for two-electron transitions to neighboring Cu ions. In this case it appears that the holes occupy mainly the $\pi p_{x,y}$ orbital [19], bringing about in a natural way a bonding character of the orbital of the hole pair on account of the geometry of the bonds in the CuO_2 plane, so that holes can be found in the space between Cu ions (Fig. 4).

It follows from the foregoing analysis that upon the creation of a local minimum of Δ_{CT} (on account of doping) it is possible for electronic bound states to arise on pairs of neighboring Cu ions, i.e., the formation of a $-U$ center can occur.[2] In that case the singlet hole pair will be localized in the vicinity of the $-U$ center at a distance $\sim a/2$. The localization region of the hole pair will be limited by the condition such that the position of the pair level coincides with the top of the valence band (the energy of the pair level becomes higher as the localization region of the hole pair increases).

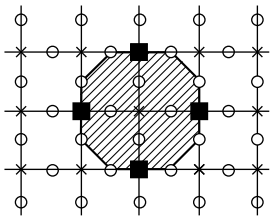

Fig. 3. The localization region of a doped hole in the CuO_2 plane (shaded) includes 12 oxygen ions: (×) copper ions, (○) oxygen ions, (■) Cu ions for which Δ_{CT} is lowered by 1.8 eV by the presence of a hole at three nearest-neighbor oxygen ions.

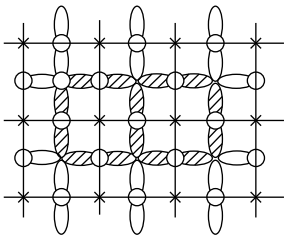

Fig. 4. The formation of bonding orbitals of a singlet hole pair of a $-U$ center from $\pi p_{x,y}$ orbitals of oxygen.

Conduction can occur in such a system if these localization regions of the hole pairs form a percolating cluster, or, in other words, when the threshold of classical percolation is exceeded for the orbitals of the hole pairs of the $-U$ centers. The delocalization region of the hole carriers in a cluster of $-U$ centers is also limited by the condition such that the position of the pair level coincides with the top of the valence band. This mechanism will cause the pair level to be pinned precisely at the top of the valence band (Fig. 1b) for $x < x_{over}$, where x_{over} corresponds to the transition to the "overdoped" region. In that case the localization regions of the carriers will overlap substantially, and the pair level will sink below the top of the valence band.

On the other hand, if the M ions are located at a distance of 2a (Fig. 2d), for the inner Cu ion the gap Δ_{CT} vanishes for single-electron transitions as well.[3] Such a fragment is a nucleus of the metallic (nonsuperconducting) phase. At the corresponding concentrations the entire crystal undergoes a transition to the normal metal state. For this state there is a single-band electronic spectrum in the vicinity of E_F. In $La_{2-x}Sr_xCuO_4$ the charge carriers in the normal metal phase will be electrons, since the filling of the band $\rho < 1/2$ on account of the doping by divalent Sr, while in $Nd_{2-x}Ce_xCuO_4$ they will be holes, since $\rho > 1/2$ on account of the doping by tetravalent Ce.

[2] In $Ba_{1-x}K_xBiO_3$ the $-U$ centers are formed on neighboring Bi cations in the presence of three K ions in the eight cells surrounding each of them, i.e., each hole depresses $\Delta_{CT} \approx 2$ eV by approximately 0.6 eV. The smaller influence of a hole in $Ba_{1-x}K_xCuO_3$ in comparison with La-214 is explained by the proximity of the Bi ion to the K ion, which is negatively charged (in relation to Ba).

[3] In $Ba_{1-x}K_xBiO_3$ this occurs in the presence of four K ions in the eight cells surrounding a Bi ion.

Concentration values determining the boundaries of the percolation regions for various L

L^2	x_p	x_m	State
16	0.0371	0.0625	Insulator
13	0.0456	0.0769	"
10	0.0593	0.100	"
9	0.0659	0.111	HTSC
8	0.0742	0.125	Insulator
5	0.118	0.200	HTSC
4	0.148	0.250	Metal

3. PHASE DIAGRAM OF HTSC COMPOUNDS OF THE Ln-214 TYPES

Let us now construct the phase diagram of the HTSC compounds Ln-214, starting from the following postulates:

(1) The $-U$ centers are formed on pairs of neighboring Cu ions belonging only to clusters with $L = 3$ and $L = \sqrt{5}$.

(2) The orbitals of hole pairs are found in the immediate vicinity of these ion pairs at a distance $\sim a/2$.

(3) The conduction in the system arises upon percolation of the $-U$ centers.

(4) The localized doped carriers cannot be found at a distance of less than $2a$.

This last postulate follows from the existence of solubility limits for dopants in the Ln-214 lattice, $x_{\text{lim}} = 0.2–0.25$ [20–22], the exceeding of which results in decomposition of the single-phase state and/or disruption of the oxygen stoichiometry. The existence of a solubility limit, we believe, is due to the presence of a repulsive interaction of the localized doped carriers. This repulsion will in turn affect the distribution of the dopant ions if they have a high enough mobility at the heat-treatment temperature. We therefore assume that the dopant ions (more precisely, their projection on the CuO_2 plane) cannot be located at distances $L < 2a$.

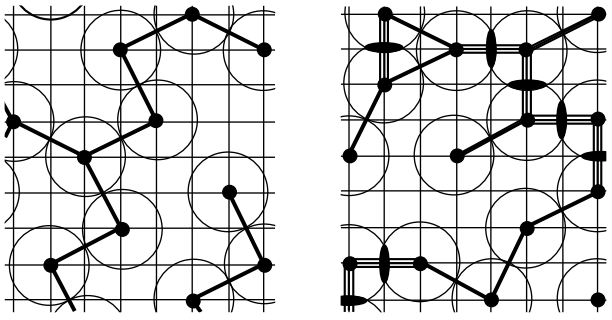

Fig. 5. A construction illustrating the method of determining the percolation thresholds of units of length $L = \sqrt{5}$ in a square lattice: (a) correct determination of the percolation threshold of units of length L when the distance between atoms is not less than L; (b) determination of the percolation threshold for $L = \sqrt{5}$ in the presence of units with $L = 2$ gives a lowered value of x_p because of the superposition of circles. The triple lines indicate units with $L = 2$. The regions of overlap of the circles are shaded.

Under this assumption the threshold of two-dimensional ($2D$) percolation for $-U$ centers in a square lattice can be determined in the following way. Suppose we have a square mesh with a cell parameter $a = 1$, with a fraction x of the mesh points (sites) occupied by atoms, and we are to determine the percolation threshold for units of length L (i.e., pairs of atoms located at a distance L). For this we place each occupied site at the center of a circle of radius $L/2$ (Fig. 5). The sum of the areas of the circles constructed around these atoms is $S = x\pi L^2/4$. For a square lattice, percolation sets in when $S \geq 0.466$ [23]. It follows that the concentration corresponding to the percolation threshold is

$$x_p = 0.593/L^2.$$

Here we are assuming that the distribution of the atoms over sites is random and that L is the shortest distance between atoms at the given concentration. Otherwise there will be superpositions of circles (Fig. 5b) and x_p will be larger than that obtained from the above relation. Then the percolating cluster will include units of length L and smaller (connecting sites found at shorter distances). The maximum number of units of length L, x_m, can be achieved for an ordered arrangement of the atoms in the square lattice with a constant L, i.e., $x_m = 1/L^2$. The values of x_p and x_m for different L are given in Table 1. Also indicated, in the right-hand column, is the state (insulator, ordinary metal, or HTSC) that would correspond to the given value of L in the case of an ordered distribution of the dopant.

Figure 6a shows the percolation region for units of various lengths L (the corresponding value of L^2 is given to the left of each rectangle in the figure). The left-hand side of each rectangle corresponds to the threshold of $2D$ percolation for units of length L in the case of a random distribution of the atoms in the absence of units with lengths shorter than L. The right-hand side corresponds to the points x_m. The heavy lines indicate the percolation regions for units with $L = 3$ and $L = \sqrt{5}$ (i.e., clusters containing $-U$ centers). It is in these regions that high-T_c superconductivity is realized, according to the proposed model.

As we shall see below, the relationship between order and disorder in the distribution of the atoms over sites plays a very large role and determines all of the features of the phase diagrams of HTSC compounds of the Ln-214 type. The tendency toward ordering is due, we think, to the difference of the ionic radii of Ln and

the dopant. It should be largest for the pair La/Ba and smallest for Nd/Ce. The degree of ordering should also increase with increasing x (decreasing L). When we talk about ordering of the dopant, we will mean the formation of a large number of ordered clusters (less than 100 Å in size), separated by narrow regions (domain walls) with a random distribution of the dopant.[4] Thus only short-range order is present in the system. It is assumed that when x_m is exceeded there is a transition to a random distribution of atoms over sites, with the formation of an infinite percolating cluster with a smaller L.

Let us now consider the different concentration intervals in Fig. 6a, assuming that not more than two types of units coexist at each point $x > 0.1$.

(1) $0.20 < x < 0.25$. Here 2D percolation occurs via units with $L = 2$ (i.e., via clusters of normal (nonsuperconducting) metal).

(2) $0.148 < x < 0.20$. In this interval the 2D percolation threshold for units with $L = 2$ depends on the degree of ordering of the dopant atoms with $L = \sqrt{5}$. In the presence of ordering of the atoms with $L = \sqrt{5}$ the $L = 2$ percolation sets in at $x = 0.2$, whereas in the case of a random distribution, percolation for $L = 2$ is achieved at $x = 0.148$. Therefore, this region corresponds either to high-T_c superconductivity or to a mixed state of a HTSC and a normal (nonsuperconducting) metal.

(3) $0.125 < x < 0.148$. Here there is "pure" 2D percolation via units with $L = \sqrt{5}$. This region corresponds to high-T_c superconductivity.

(4) $0.118 < x < 0.125$. In the presence of ordering of the atoms with $L = \sqrt{8}$ the onset of percolation with $L = \sqrt{5}$ occurs at $x = 0.125$, while in the case of a random distribution of the atoms with $L = \sqrt{8}$ and $L = \sqrt{5}$, percolation for $L = \sqrt{5}$ is achieved at $x = 0.118$. This region corresponds to an insulator in the first case and to a HTSC in the second.

(5) $0.111 < x < 0.118$. This is the region of "pure" 2D percolation with $L = \sqrt{8}$, and here an insulating state is realized.

(6) $0.10 < x < 0.111$. In the case of ordering of the atoms with $L = 3$ (more precisely, in the absence of pairs with $L = \sqrt{8}$), 2D percolation with $L = 3$ sets in at $x = 0.10$, while in the case of a random distribution of the atoms with $L = 3$ and $L = \sqrt{8}$ there is no 2D percolation. In this case the percolating cluster will include regions with $L = 3$ (HTSC) and with $L = \sqrt{8}$ (insulator), and conduction is possible only in cases of 3D percolation (with Josephson tunneling between CuO_2 planes and/or with tunneling between clusters with $L = 3$).

[4] Subject to the restriction $L \geq 2$.

(7) $0.066 < x < 0.10$. In this interval there is no 2D percolation with $L = 3$. 3D percolation (and superconductivity) can still occur for $x > 0.077$, while for $x < 0.077$ there is no 3D superconductivity.

For comparison, Figs. 6b–6d shows the experimental phase diagrams of $La_{2-x}Ba_xCuO_4$ [24], $La_{2-x}Sr_xCuO_4$ [25], and $Nd_{2-x}Ce_xCuO_4$ [26]. The difference between the phase diagrams of $La_{2-x}Ba_xCuO_4$ and $La_{2-x}Sr_xCuO_4$ is that the dip in T_c occurs at $x_k = 0.125$ in the first case and at $x_k = 0.115$ in the second. Comparing Figs. 6a and 6b–6d, we can easily see that all of the extremal points on the

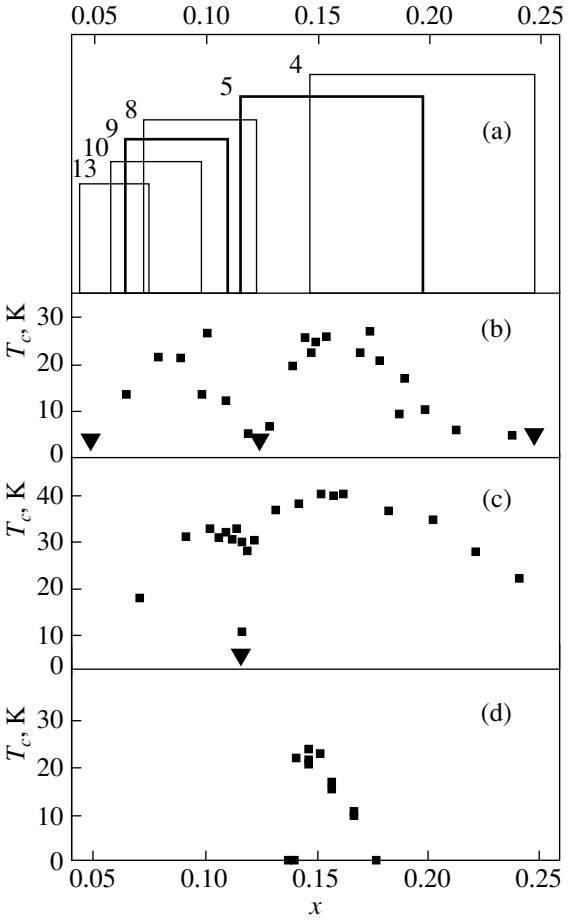

Fig. 6. (a) Boundaries of the percolation regions for units with different L. The left-hand side of each rectangle corresponds to the threshold of 2D percolation for units of length L (in the case of a random distribution of circles of radius $L/2$ over sites and in the absence of units with lengths shorter than L). The right-hand side of each rectangle is the point of the maximum number of units of length L, corresponding to an ordered distribution of atoms in a square lattice with constant L. To the left of each rectangle is the corresponding value of L^2. The height of the rectangles qualitatively reflects the growth of the number of units with as L decreases; (b)–(d) experimental phase diagrams $T_c(x)$ for Ln-214 HTSCs. The triangles on the diagrams indicate compositions for which superconductivity is not observed down to 4.2 K: (b) $La_{2-x}Ba_xCuO_4$ [24], (c) $La_{2-x}Sr_xCuO_4$ [25], (d) $Nd_{2-x}Ce_xCuO_4$ [26].

experimental phase diagrams practically coincide with the boundaries of the percolation regions for units with different L. The difference in the values of x_k for $La_{2-x}Ba_xCuO_4$ and $La_{2-x}Sr_xCuO_4$ is naturally attributed to the high degree of ordering in the La/Ba sublattice in comparison with La/Sr, as a result of which the percolation threshold for units with $L = \sqrt{5}$ is shifted to $x = 0.125$—the point of maximum ordering for $L = \sqrt{8}$.

The maximum of T_c at $x \approx 0.15$ is due to the onset (for $x > 0.148$) of clusters of the normal metal inside the superconducting phase. Therefore, the superconductivity in this region is "weakened" in proportion to the ratio of the superconducting and normal phases. This is confirmed [27] by measurements of the volume of the "Meissner phase" as a function of magnetic field. Those measurements show that whereas for $x = 0.15$ this quantity is practically independent of field, for $x > 0.15$ the volume of the superconducting phase is noticeably decreased by a magnetic field. At the same time, in low fields the volume of the Meissner phase increases up to $x \approx 0.2$ [20, 27], which attests to ordering of the dopants with $L = \sqrt{5}$.

In $Nd_{2-x}Ce_xCuO_4$ (Fig. 5b), because of the small difference in the sizes of the Nd and Ce ions, there is practically no ordering of Ce. Therefore 2D percolation can occur only for $x > 0.118$ (with $L = \sqrt{5}$). However, because of the absence of ordering the percolating cluster will also include regions with $L = 2$, and the percolation threshold for $L = \sqrt{5}$ is shifted to the percolation threshold for $L = 2$. This agrees with the experimental phase diagram.

Let us now discuss the region $x < 0.12$ for $La_{2-x}Ba_xCuO_4$ and $La_{2-x}Sr_xCuO_4$ (the underdoped region) in special detail. It follows from Fig. 6 that for $x < 0.077$ even 3D percolation does not occur, and only "traces" of superconductivity can be observed. This conclusion agrees with the results of [28], it was observed that 3D superconductivity did not occur in $La_{2-x}Sr_xCuO_4$ for $x \leq 0.08$. As we have said, it would be hard to expect 2D percolation (in any case for $La_{2-x}Sr_xCuO_4$) in the interval $0.08 < x < 0.12$ either, on account of the proximity of the percolation regions for three different types of units with $L^2 = 8$, 9, and 10. Here there will most likely be 3D percolation and/or tunneling between clusters with $L = 3$. This conclusion explains the results of [29], where in $La_{2-x}Sr_xCuO_4$ for $x < 0.15$ a logarithmic divergence of the resistivity was observed as $T \longrightarrow 0$, with suppression of the superconductivity by a magnetic field.

The proposed model offers another way of interpreting the results of experiments on the observation of a pseudogap in HTSCs in the underdoping and optimal doping regions [30–32]. It follows from the experiment that the pseudogap has the value and symmetry of the superconducting gap but collapses at $T^* > T_c$ (where T^* increases with decreasing x).

We assume that the observed pseudogap is actually the superconducting gap but that it arises at $T > T_c$ as a result of large fluctuations of the number of particles due to transitions of electrons between pair levels and the band. The point is that, unlike an ordinary superconductor with an electron-phonon interaction, where the superconducting gap closes on account of thermal excitations which decrease the number of states to which the electron pairs can be scattered (the interaction with real phonons), in our case the mechanism leading to suppression of the gap is the filling of $-U$ centers by electrons (the interaction with real bosons). Therefore the fluctuational decrease of the population of electron pair levels will tend to increase the "superconducting" interaction and can lead to a fluctuational "turning on" of superconductivity (a second-order transition) at $T^* > T_c > T_{c0}$ (here T_{c0} is the equilibrium value of T_c) [33]. Such large fluctuations can occur in underdoped and optimally doped samples when an appreciable fraction of the $-U$ centers belong to small isolated clusters. As x decreases, an ever-larger fraction of the $-U$ centers are found in small clusters, and the relative fluctuations of the number of particles increase (i.e., T^* increases). On the other hand, in the "overdoped" samples, when practically all of the copper ions belong to the infinite percolating cluster, such large fluctuations become impossible.

We can thus conclude that all of the features observed on the phase diagrams of Ln-214 HTSCs are only a reflection of the geometrical relationships in a square lattice and the competition between order and disorder in the distribution of the dopant ions. The agreement of the calculated phase diagram with the experimental diagrams confirms the conclusion that the superconductivity in Ln-214 is due specifically to the investigated fragments containing pairs of neighboring Cu ions in the CuO_2 plane and is convincing evidence in favor of the proposed model of high-T_c superconductivity.

4. MECHANISM OF RELAXATION OF HOLE CARRIERS IN HTSCs

It follows from the above discussion that for $x > x_0$, two-electron transitions can occur to certain pairs of neighboring copper ions, while single-electron transitions are still forbidden. The pair level of the $-U$ center formed on a pair of neighboring Cu ions is located at the top of the valence band and is essentially a pair resonance state. The transition of electrons from the oxygen ions to these $-U$ centers will give rise to additional hole carriers in a certain neighborhood of the $-U$ center and will lead to the onset of hole conduction of the percolation type at a dopant concentration $x > x_p$ ($x_0 < x_p < x_c$). We stress that for $x_p < x < x_c$ the state is not the ordinary metallic state (which arises only for $x > x_c$) but is an intermediate state (the $-U$ phase) in which doping gives

rise to $-U$ centers, and the hole carriers that arise can effectively interact with one another at these centers. We assume that it is this concentration range $x_p < x < x_c$ that corresponds to the HTSC region on the phase diagram. Let us now consider the kinetics of the hole carriers in the $-U$ phase.

Since the mechanism described leads to the appearance of a pair electron level located near the top of the oxygen valence band in the electron spectrum of the initial insulator phase, its hybridization with the states of the oxygen band causes both the itinerant and localized states to be broadened. As the two-particle hybridization increases, the broadening of the pair level is [9, 10]

$$\Gamma \approx 4\pi kT(DV)^2$$

(here V is the hybridization constant, D is the density of states in the valence band, and T is the temperature). The corresponding broadening of the itinerant states is

$$\gamma \propto \Gamma \propto T.$$

The broadening of the itinerant states causes a smearing of the features in the density of states of the oxygen valence band and leads to its energy independence in the interaction region (an analog of the "marginal" Fermi liquid [34]).

The occupation η of the pair states and the hole concentration n in the valence band are determined by the equality of the rates of transitions from the band to the pair level and back. If N is the concentration of $-U$ centers, then $2N\eta = n$. The frequency of transitions from the pair level to the band is equal to $N\eta\Gamma \propto T\eta$. The rate of the reverse process is determined by the frequency of electron–electron scattering and is proportional to $\gamma^2(1-\eta) \propto T^2(1-\eta)$. Hence

$$\eta = \frac{T}{T_0 + T},$$

where the constant T_0 is independent of temperature. Hence $n \propto T$ in the low-temperature region, while at high T it approaches a constant (equal to $2N$). This result agrees with the data of Hall measurements [35, 36] on $YBa_2Cu_3O_7$.

Thus, on account of the interaction of the electrons with $-U$ centers, the distribution of hole carriers is nondegenerate in the sense that the chemical potential $\mu = 0$ for holes is zero for all T, whereas the condition of degeneracy is that $\mu > 0$. When one takes into account the nondegenerate character of the distribution (the absence of Pauli exclusion) and the high concentration (10^{21}–10^{22} cm^{-3}), it is expected that the predominant contribution to the relaxation process will be from electron–electron scattering (in this case, the scattering of hole carriers on one another). However, since the interaction of two holes in a system with $-U$ centers corresponds to an effective attraction, this is not ordinary Coulomb scattering. In the present case the main mechanism for electron–electron scattering is analogous to that [37] which takes place in metals and alloys with a strong electron–phonon interaction. In those substances, for the electrons found in a layer with a thickness of about $k\Theta_D$ (Θ_D is the Debye temperature) on the Fermi surface, the effective electron–electron interaction responsible for the attraction, involving the exchange of virtual phonons, is much greater than the screening of the Coulomb repulsion. Therefore the main channel of electron–electron scattering in the present case will also be due to the exchange of virtual phonons. The contribution from these processes [37] becomes substantial at $T < \Theta_D$. Here the electron–electron scattering amplitude is independent of the energy E of the particles being scattered for $E \ll k\Theta_D$ and falls off sharply for $E \sim k\Theta_D$. For $E > k\Theta_D$ only the Coulomb interaction contributes to the scattering amplitude. A contribution of electron–electron scattering to the resistivity ρ ($\rho = AT^2$) greater than the electron–phonon contribution has been observed experimentally in Al [38] at $T < 4$ K and in superconductors with the A15 lattice [39] at $T < 50$ K. Here the amplitude A was more than an order of magnitude larger than the value calculated on the assumption of a Coulomb scattering mechanism.

Thus the main contribution to the hole carrier relaxation processes in HTSCs is from hole–hole scattering at a $-U$ center due to the exchange of a virtual boson with energy Ω. Since $\Omega \sim 0.1$–1 eV, the temperature interval in which there is a substantial contribution from scattering processes involving a virtual boson is broadened to $T \sim 10^3$ K.

The temperature dependence of the resistivity $\rho(T)$ in this model can be obtained from the Drude formula: $\rho = m^*(\nu/n)e^2$ (where m^* is the effective mass of the holes, and ν is the hole carrier scattering frequency). For $\Omega \gg E$ the scattering amplitude is independent of the particle energy E. The scattering frequency ν will therefore be determined by the hole concentration and the statistical factor in the scattering cross section, i.e., the volume of phase space accessible to the particles being scattered, which is proportional to $E_1 + E_2$ (here E_1 and E_2 are the energies of the scattering particles measured from the top of the oxygen band), i.e.,

$$\nu \propto n(E_1 + E_2).$$

For steady conduction one has $E_1 \sim E_2 \sim \gamma \propto T$, $\nu \propto nT \propto T^2$, and thus $\rho(T) \propto T$. A dependence of this form is observed experimentally for optimally doped samples of $YBa_2Cu_3O_7$, $La_{2-x}Sr_xCuO_4$, $Bi_2Sr_2CaCu_2O_y$, etc.

In overdoped HTSCs the additional carriers cause the pair level to sink below the top of the valence band by an amount δE. Then, as a result of the degeneracy of the hole carriers at low temperatures, n ceases to depend on T for $\gamma \ll \delta E$. The temperature-dependent part of the resistivity in this case takes the form

$$\rho(T) \propto \gamma^2 \propto T^2.$$

A dependence of this kind is observed experimentally in a number of HTSCs in the "overdoped" regime. As T increases, when γ becomes of the same order as δE, a transition to a linear $\rho(T)$ dependence is observed.

The predominant contribution of the electron–electron interaction to the scattering processes will affect both the frequency dependence and temperature dependence of the optical conductivity σ_{opt}:

$$\sigma_{opt} = (en/m^*)[\nu/(\omega^2 + \nu^2)]$$

(here ω is the frequency of the radiation and ν is the "optical" relaxation frequency). For electron–electron scattering (at a concentration $n \sim 10^{22}$ cm^{-3}) the collision frequency $\nu \geq 10^{15}$ s^{-1}. Therefore, in the IR range we have $\nu \gg \omega$, and the formula for the optical conductivity takes an even simpler form:

$$\sigma_{opt} = en/m^*\nu.$$

For optical relaxation we have $E_1 \sim \omega$, $E_2 \sim \gamma \propto T$, and $\nu \propto n\omega$ for $\omega \gg \Gamma$ and $\nu \propto nT$ for $\omega \ll \Gamma$. Hence $\sigma_{opt} \propto \omega^{-1}$ for $\omega \gg \Gamma$ and $\sigma_{opt} \propto T^{-1}$ for $\omega \ll \Gamma$. These results are in complete agreement with the data of various experiments [40, 41].

5. CONCLUSION

We have presented an elementary model for the modification of the electron spectrum of a HTSC upon doping, wherein the transition from an insulator to a metallic state passes through a certain range of dopant concentrations in which transitions of singlet electron pairs can occur from the oxygen ions to two neighboring cations (a $-U$ center), while the single-electron transitions are still forbidden. We believe that it is this range of concentrations that corresponds to the region of high-T_c superconductivity, when the electron–electron attraction is due to the scattering of electron pairs on $-U$ centers. Conduction arises in the system as a result of percolation via $-U$ centers,[5] or, more precisely, via $\pi p_{x,y}$ orbitals of singlet hole pairs. The aforementioned features of the electron spectrum give rise to qualitatively new properties of the system (a nondegenerate distribution of hole carriers, a predominant contribution of electron–electron scattering to energy relaxation processes), which are responsible for the unusual behavior of HTSCs in the normal state. Starting from the proposed model we have constructed the phase diagram of HTSC compounds of the Ln-214 class and obtained complete agreement with experiment. Based on the analysis presented here, one can conclude that HTSCs are indeed a special class of solids in which a nonstandard mechanism of superconductivity, different from the BCS mechanism, is realized.

[5] Interestingly, the percolating cluster in the case of doped HTSCs of the Ln-214 class is reminiscent of a Little "polymer" [42], while in the case of the HTSCs $YBa_2Cu_3O_7$ and $Bi_2Sr_2Ca_nCu_{n+1}O_y$ (where the doping holes are found in a plane parallel to the CuO_2 plane) it resembles a Ginzburg "sandwich" [43].

ACKNOWLEDGMENTS

The authors are grateful to L.V. Keldysh, B.A. Volkov, and all the participants in the Solid State Theory Seminar of the Department of Theoretical Physics of the Lebedev Physical Institute, Russian Academy of Sciences for a discussion of this study.

REFERENCES

1. J. G. Bednorz and K. A. Muller, Z. Phys. B **64**, 189 (1986).
2. V. M. Loktev, Fiz. Nizk. Temp. **22**, 3 (1996) [Low Temp. Phys. **22**, 1 (1996)].
3. P. W. Anderson, Phys. Rev. Lett. **34**, 953 (1975).
4. E. Simanek, Solid State Commun. **32**, 731 (1979).
5. C. S. Ting, D. N. Talwar, and K. L. Ngai, Phys. Rev. Lett. **45**, 1213 (1980).
6. H.-B. Schuttler, M. Jarrell, and D. J. Scalapino, Phys. Rev. Lett. **58**, 1147 (1987).
7. J. Yu, S. Massida, A. J. Freeman, and D. D. Koelling, Phys. Lett. A **122**, 203 (1987).
8. B. A. Volkov and V. V. Tugushev, Pis'ma Zh. Éksp. Teor. Fiz. **46**, 193 (1987) [JETP Lett. **46**, 245 (1987)]; Pis'ma Zh. Éksp. Teor. Fiz. **49**, 337 (1989) [JETP Lett. **49**, 384 (1989)].
9. G. M. Éliashberg, Pis'ma Zh. Éksp. Teor. Fiz. **46** (12), 94 (1987) [JETP Lett. **46**, S81 (1987)].
10. I. O. Kulik, Fiz. Nizk. Temp. **13**, 879 (1987) [Sov. J. Low Temp. Phys. **13**, 505 (1987)].
11. P. I. Arseev, Zh. Éksp. Teor. Fiz. **101**, 1246 (1992) [Sov. Phys. JETP **74**, 667 (1992)].
12. J. Zaanen, G. A. Sawatzky, and J. W. Allen, Phys. Rev. Lett. **55**, 418 (1985).
13. Y. Ohta, T. Tohyama, and S. Maekava, Phys. Rev. Lett. **66**, 1228 (1991).
14. S. Mazumdar, Solid State Commun. **69**, 527 (1989).
15. R. L. Martin, Phys. Rev. B **53**, 15501 (1996).
16. P. C. Hammel, R. L. Martin, B. W. Statt, *et al.*, Phys. Rev. B **57**, 712 (1998).
17. H. Romberg, M. Alexander, N. Nucker, *et al.*, Phys. Rev. B **42**, 8768 (1990).
18. D. R. Harshman and A. P. Mills, Jr., Phys. Rev. B **45**, 10 684 (1992).
19. J. E. Hirsh and S. Tang, Solid State Commun. **69**, 987 (1989).
20. P. G. Radaelli, D. G. Hinks, A. W. Mitchell, *et al.*, Phys. Rev. B **49**, 4163 (1994).
21. K. Yoshimura, H. Kubota, H. Tanaka, *et al.*, J. Phys. Soc. Jpn. **62**, 1114 (1993).
22. E. F. Paulus, I. Yehia, H. Fuess, *et al.*, Solid State Commun. **73**, 791 (1990).

23. J. M. Ziman, *Models of Disorder: The Theoretical Physics of Homogeneously Disordered Systems* (Cambridge Univ. Press, Cambridge, 1979).
24. A. R. Moodenbaugh, Y. Xu, M. Suenaga, *et al.*, Phys. Rev. B **38**, 4596 (1988).
25. K. Kumagai, K. Kawano, I. Watanabe, *et al.*, J. Supercond. **7**, 63 (1994).
26. H. Takagi, S. Uchida, and Y. Tokura, Phys. Rev. Lett. **62**, 1197 (1989).
27. K. Kitazawa, Y. Tomioka, T. Nagano, and K. Kishio, J. Supercond. **7**, 27 (1994).
28. C. Marin, T. Charvolin, D. Braithwaite, and R. Calemczuk, Physica C (Amsterdam) **320**, 197 (1999).
29. Y. Ando, G. S. Boebinger, A. Passner, *et al.*, Phys. Rev. Lett. **75**, 4662 (1995).
30. H. Ding, T. Yokoya, J. C. Campuzano, *et al.*, Nature **382**, 51 (1996).
31. S. Marshall, D. S. Dessau, A. G. Loeser, *et al.*, Phys. Rev. Lett. **76**, 4841 (1996).
32. H. Ding, M. R. Norman, T. Yokoya, *et al.*, Phys. Rev. Lett. **78**, 2628 (1997).
33. O. M. Ivanenko and K. V. Mitsen, Physica C (Amsterdam) **235–240**, 2361 (1994).
34. C. M. Varma, P. B. Littlewood, S. Schmitt-Rink, *et al.*, Phys. Rev. Lett. **63**, 1996 (1989).
35. N. G. Ong, T. W. Jing, T. R. Chien, *et al.*, Physica C (Amsterdam) **185–189**, 34 (1991).
36. A. Kapitulnik, Physica C (Amsterdam) **153–155**, 520 (1988).
37. A. H. MacDonald, Phys. Rev. Lett. **44**, 489 (1980).
38. J. C. Garland and R. Bowers, Phys. Rev. Lett. **21**, 1007 (1968).
39. M. Gurvitch, A. K. Ghosh, B. L. Gyorffy, *et al.*, Phys. Rev. Lett. **41**, 1616 (1978).
40. D. B. Tanner, M. A. Quijada, D. N. Basov, *et al.*, J. Supercond. **8**, 563 (1995).
41. Z. Schlesinger, R. T. Collins, F. Holtzberg, *et al.*, Phys. Rev. B **41**, 11237 (1990).
42. W. A. Little, Phys. Rev. **134**, A1416 (1964).
43. V. L. Ginzburg, Zh. Éksp. Teor. Fiz. **47**, 2318 (1964) [Sov. Phys. JETP **20**, 1549 (1964)].

Translation was provided by AIP

Journal of Experimental and Theoretical Physics, Vol. 91, No. 3, 2000, pp. 588–596.
Translated from Zhurnal Éksperimental'noĭ i Teoreticheskoĭ Fiziki, Vol. 118, No. 3, 2000, pp. 676–686.

SOLIDS
Electronic Properties

Magnetization of Type-II Superconductors in the Range of Fields $H_{c1} \le H \le H_{c2}$: Variational Method

V. V. Pogosov[a, *], A. L. Rakhmanov[b], and K. I. Kugel'[b]

[a] *Moscow Institute of Physics and Technology, Dolgoprudnyi, Moscow oblast, 141700 Russia*
[b] *Institute of Theoretical and Applied Electrodynamics, Russian Academy of Sciences, Moscow, 127412 Russia*
** e-mail: wpogosov@mail.ru*

Received April 4, 2000

Abstract—A variational method is proposed to find the magnetic field dependence of the magnetization of type-II superconductors in the mixed state by a self-consistent technique. This model allows for suppression of the order parameter to zero at the centers of Abrikosov vortices and also for the magnetic field dependence of the order parameter. The results can be applied to the entire range of fields $H_{c1} \le H \le H_{c2}$ for any values of the Ginzburg–Landau parameter $\kappa > 1/\sqrt{2}$. It is shown that in weak fields where $\kappa \gg 1$ the behavior of the magnetization can be described exactly in the London approximation provided that the correct value of H_{c1} is used. Near the second critical field this dependence shows good agreement with the well-known Abrikosov result. It is also shown that using the concept of isolated vortices and applying the principle of superposition of the fields and currents generated by these vortices to calculate the magnetization gives inaccurate quantitative results even in fairly weak fields. By going beyond these concepts, it was possible to allow more accurately for the influence of the vortex cores on the magnetization behavior in the intermediate range of fields $H_{c1} \ll H \ll H_{c2}$ and to identify the range of validity of various approximations used widely in the literature. © *2000 MAIK "Nauka/Interperiodica".*

1. INTRODUCTION

The magnetization of type-II superconductors is a fundamental electromagnetic characteristic. It can be used to find various important parameters of the superconductor such as the lower H_{c1} and upper H_{c2} critical fields, and the Ginzburg–Landau parameter κ [1–3]. An enormous number of experimental and theoretical studies have been devoted to magnetization (see, for example, the reviews [4, 5]). In this context it is important to obtain formulas for the magnetization of superconductors which would be suitable for quantitative calculations over a wide range of external magnetic fields. This problem has been discussed in the literature for some time (see, for example, [2–15]). Until recently, however, there was no convenient and reliable approach which could be applied to calculate the magnetization of a type-II superconductor analytically over the entire range of external fields $H_{c1} \le H \le H_{c2}$.

The problem of calculating the magnetic moment M of a superconductor can be solved most easily in weak fields $H \ll H_{c2}$. Here the cores of the Abrikosov vortices occupy only a small part of the volume and $M(H)$ is obtained for $\kappa \gg 1$ using the London approximation where the modulus of the order parameter is assumed to be constant to calculate the local fields and currents outside the core [1–3]. In the London model the dependence of the magnetization M of an ideal isotropic superconductor on the magnetic field H in the range $H_{c1} \ll H \ll H_{c2}$ can be described using the Fetter formula [6]:

$$-4\pi M = H_{c1} - \frac{1}{4\kappa}\{\ln[2\kappa(H - H_{c1})] + 1.34\}. \quad (1)$$

In this formula and subsequently we use a system of units [3] in which all the distances are normalized to the London depth of penetration of the magnetic field λ, the magnetic field is normalized to $H_c\sqrt{2}$ (where H_c is the thermodynamic critical field), the order parameter is normalized to its equilibrium value, and the vector potential is normalized to $\hbar c/2e\xi$, where $\hbar$ is Planck's constant, c is the velocity of light, e is the electron charge, and ξ is the coherence length. The dimensionless values of the local magnetic field, the vector potential, and the order parameter are denoted by **h**, **a**, and f. Note that in this system of units the flux quantum is $\Phi_0 = 2\pi/\kappa$ and $H_{c2} = \kappa$. The lower critical field H_{c1} cannot be calculated self-consistently in the London model. For this reason, H_{c1} appears in Eq. (1) as a parameter and for $\kappa \gg 1$ may be written in the form [3]

$$H_{c1} = \frac{1}{2\kappa}(\ln\kappa + \varepsilon). \quad (2)$$

The constant ε is determined by the structure of the order parameter at the vortex core and its value $\varepsilon \approx 0.50$ was determined by Hu [7] by means of a numerical solution of the complete Ginzburg–Landau system of

equations (see also [8]). The Fetter dependence (1) differs for $H \longrightarrow H_{c1}$. In the immediate vicinity of H_{c1} the magnetization in the London model can be obtained numerically [14] or analytically using an approximation which only allows for vortex interaction with nearest neighbors in the vortex lattice [3]. In order to extend the validity of the London approximation, various approaches have been developed which make partial allowance for the contribution of the vortex cores to the free energy of the superconductor (see [5, 13]).

The London model cannot be applied in strong fields because the vortex density is high in this case. The behavior of the magnetization near the second critical field is described by the well-known Abrikosov expression [3]:

$$M = \frac{H - H_{c2}}{4\pi\beta_A(2\kappa^2 - 1)}, \quad H_{c2} - H \ll H_{c2}, \qquad (3)$$

where, for a triangular vortex lattice, we have $\beta_A = 1.16$.

In [10, 11] Clem proposed a fairly simple variational model which allows for the structure of the order parameter near the center of the vortex. The following trial function was used for the modulus of the order parameter:

$$f = \frac{f_\infty r}{\sqrt{r^2 + \xi_\nu^2}}, \qquad (4)$$

where r is the distance from the center of the vortex, ξ_ν and f_∞ are variational parameters characterizing the spatial distribution of the order parameter. This model was used to obtain a formula for H_{c1} [10,11] which for $\kappa \gg 1$ may be expressed in the form (2) where $\varepsilon \approx 0.52$ which shows good agreement with the results of [7, 8].

Hao and Clem then generalized this variational model to the case of a regular vortex lattice and obtained a unified formula for $M(H)$ which can be applied over the entire range of fields $H_{c1} \le H \le H_{c2}$ [11]. One of the most important conclusions of this study is that, even in weak fields, the influence of the vortex cores cannot be neglected and consequently the London model cannot generally give an exact result [11,12]. In the range of fields near H_{c2} the dependence $M(H)$ obtained in [11] is almost the same as the Abrikosov result (3). This theory was subsequently generalized to the case of anisotropic superconductors [16]. The model proposed in [11] has been widely used in the literature. The formula for the magnetization has been actively used to analyze experimental data from measurements of the magnetic moment of various superconductors such as: $YBa_2Cu_3O_{7-\delta}$ [17], $YBa_2Cu_4O_8$ [18], $Bi_2Sr_2Ca_2Cu_3O_{10}$ [19], $(Tl,Pb)Sr_2Ca_2Cu_3O_9$ and $Tl_2Ba_2Ca_2Cu_3O_{10}$ [20], $HgBa_2Ca_2Cu_3O_{8+\delta}$ [21, 22], $Hg_{0.8}Pb_{0.2}Ba_{1.5}Sr_2Cu_3O_{8-\delta}$ [23], and $Nd_{1.85}Ce_{0.15}CuO_{4-\delta}$ [24].

In the present paper we show that several errors were made in the derivation of the formula for $M(H)$ in [11]. For example, the expression for the free energy of the vortex lattice F was formulated using the principle of superposition of the fields and currents of isolated vortices, which can only be applied in weak fields. At the same time, also in the calculations of F the transition was made from summation over the reciprocal vortex lattice to integration. The error associated with this transition and also using an inaccurate value of H_{c1} in (1) led these authors [11] to the erroneous conclusion that the London model is incorrectly formulated for $\kappa \gg 1$ even in weak fields. Unfortunately, this statement is now accepted by a whole range of researchers. Additionally, the dependences of the variational parameters ξ_ν and f_∞ on the magnetic induction given in [11] (which determine the behavior of the magnetization to a considerable extent) were not obtained self-consistently and do not follow from the expression used for the free energy, but are simply convenient approximations.

Here we use a variational model to obtain a self-consistent derivation of the expression for $M(H)$. In this case, the spatial distribution of the order parameter was simulated using the trial function (4) and the unit cell of the regular vortex lattice was replaced by a circular one (Wigner–Seitz approximation). The formula obtained for $M(H)$ can be applied over the entire range of fields $H_{c1} \le H \le H_{c2}$ for any values of the Ginzburg–Landau parameter $\kappa > 1/\sqrt{2}$. The result for the magnetization in weak fields for $\kappa \gg 1$ agrees with the London dependence (allowing for the exact value of H_{c1}) whereas in strong fields it shows good agreement with the Abrikosov result. The formula obtained for the magnetization can easily be generalized to the case of anisotropic superconductors where the vortices are oriented along one of the principal axes of the crystal. For this orientation a scaling transformation exists which can be used to calculate the magnetization of an anisotropic superconductor from an isotropic one simply by changing the notation of κ [11]. This aspect is considered in Section 2.

We also discuss the correctness of the approximation of isolated vortices in the mixed state of a superconductor. It is shown that even in weak fields, when the density of vortex filaments is still low, using the principle of superposition of the fields created by separate vortices leads to appreciable quantitative errors in calculations of the magnetization.

2. WIGNER–SEITZ APPROXIMATION

In weak fields the distances between the neighboring vortex filaments are many times the dimensions of the vortex cores. This means that the vortices can be considered as independent interacting objects (see, for example, [1–3, 13, 25, 26]). Thus, in weak fields the principle of field superposition is satisfied: the self-induced field of each filament is assumed to be the same as that of an isolated filament and the local field at an

arbitrary point in the superconductor is the sum of the fields of all the filaments. The energy of the vortex lattice is expressed as the sum of the self-energies of the filaments and the energies of their pairwise interaction [3, 25]. A particular case of this approach is the London approximation which neglects the influence of the spatial variation of the order parameter in the core of each filament on its field which is valid when $\kappa \gg 1$.

In strong fields the vortex concentration is high and for this reason the concept of independent filament interactions becomes meaningless (see, e.g., [3]). However, as was shown in [27], the local magnetic field in the regular vortex lattice can still be represented as the sum of terms interpreted as contributions from isolated unit cells. In strong fields however, calculation of these contributions is a nontrivial problem. Moreover, this approach is artificial since the vortices are no longer isolated objects and their properties are determined by the lattice as a whole. In this case, it is far simpler to calculate the local magnetic field distribution and the order parameter in an isolated lattice unit cell. The area of the cell is uniquely related to the magnetic induction and the existence of translational invariance in the system yields the boundary condition that the current density at the cell boundary is zero. This method can also be used to obtain the well-known Abrikosov result for the magnetization in fields near H_{c2}. This approach is also convenient for numerical solutions of the Ginzburg–Landau equation over the entire range of external fields $H_{c1} < H < H_{c2}$ [28, 29].

An important simplification in this case is the Wigner–Seitz approximation, i.e., replacing the hexagonal vortex cell with a circle of the same area. In [9] the Wigner–Seitz approximation was applied to find the magnetization in weak fields when $\kappa \gg 1$ and the results show good agreement with the London model. This approximation has frequently been used in numerical calculations of vortex structures [30–34]. It has been found that in Ginzburg–Landau theory [30] and in microscopic superconductivity theory [34] the approximation of a circular cell yields good results not only in weak fields but also near H_{c2}.

In the present paper we propose a variational model to obtain analytic expressions for the magnetization in the Wigner–Seitz approximation. Instead of solving the complete system of Ginzburg–Landau equations, we use the trial function (4) to model the distribution of the order parameter in a Wigner–Seitz cell and the corresponding local magnetic field is calculated from the second Ginzburg–Landau equation. The fact that expression (4) contains two variational parameters means that the vortex shape at the center of the cell can be varied widely for an arbitrary induction.

We shall calculate the magnetic field distribution in a Wigner–Seitz cell. For the case of cylindrical symmetry the second Ginzburg–Landau equation for the magnetic field can be expressed in the form [3]

$$\frac{1}{r}\frac{d}{dr}\left(\frac{r}{f^2}\frac{dh}{dr}\right) = h. \tag{5}$$

Equation (5) allowing for (4) has the solution

$$h = \alpha K_0(f_\infty\sqrt{r^2+\xi_v^2}) + \beta I_0(f_\infty\sqrt{r^2+\xi_v^2}), \tag{6}$$

where I_n is an nth-order Bessel function of an imaginary argument, K_n is an nth-order Macdonald function, and α and β are constant coefficients. The values of the constants α and β can be determined from the conditions for quantization of the magnetic field flux through the Wigner–Seitz cell $\Phi = 2\pi/\kappa$ and zero superconducting current $\mathbf{j} = \mathrm{rot}\mathbf{h}$ at its interface. This gives:

$$\alpha = \frac{f_\infty}{\kappa\xi_v}\frac{I_1(f_\infty\rho)}{K_1(f_\infty\xi_v)I_1(f_\infty\rho) - I_1(f_\infty\xi_v)K_1(f_\infty\rho)}, \tag{7}$$

$$\beta = \frac{f_\infty}{\kappa\xi_v}\frac{K_1(f_\infty\rho)}{K_1(f_\infty\xi_v)I_1(f_\infty\rho) - I_1(f_\infty\xi_v)K_1(f_\infty\rho)}, \tag{8}$$

where we introduce the notation $\rho = \sqrt{R^2+\xi_v^2}$, $R = \sqrt{2/B\kappa}$ is the cell radius, $B = 2\pi/\kappa A_{\mathrm{cell}}$ is the magnetic induction, and A_{cell} is the cell area. We stress that this result can be applied for any κ and in particular for $\kappa \sim 1$ when $H_{c1} \sim H_{c2}$ and the concept of independent filaments is only valid in a narrow range of fields near H_{c1}.

The free energy density of the vortex lattice may be expressed in the form

$$F = F_{\mathrm{core}} + F_{\mathrm{em}},$$

where F_{core} is the energy density associated with the change in the order parameter near the centers of the vortices, and F_{em} is the electromagnetic energy density [3, 11]. In Ginzburg–Landau theory F_{core} and F_{em} are given by the expressions [3]

$$F_{\mathrm{core}} = \frac{1}{A_{\mathrm{cell}}}\int\left[\frac{1}{2}(1-f^2)^2 + \frac{1}{\kappa^2}(\mathrm{grad}\, f)^2\right]d^2r, \tag{9}$$

$$F_{\mathrm{em}} = \frac{1}{A_{\mathrm{cell}}}\int\left[\mathbf{h}^2 + f^2\left(\mathbf{a} + \frac{1}{\kappa}\mathrm{grad}\gamma\right)^2\right]d^2r, \tag{10}$$

where γ is the phase of the order parameter and integration is performed over the cell area. We shall find the dependence of F on the variational parameters and the magnetic induction. For the electromagnetic energy density using the second Ginzburg–Landau equation from (10) we can easily obtain: $F_{\mathrm{em}} = Bh(0)$, where $h(0)$ is the magnetic field at the center of the vortex filament [11]. If we substitute Eqs. (4) and (6)–(8) into this formula, we have

$$F_{\mathrm{em}} = \frac{Bf_\infty}{\kappa\xi_v}\frac{K_0(f_\infty\xi_v)I_1(f_\infty\rho) + I_0(f_\infty\xi_v)K_1(f_\infty\rho)}{K_1(f_\infty\xi_v)I_1(f_\infty\xi_v) - I_1(f_\infty\xi_v)K_1(f_\infty\rho)}. \tag{11}$$

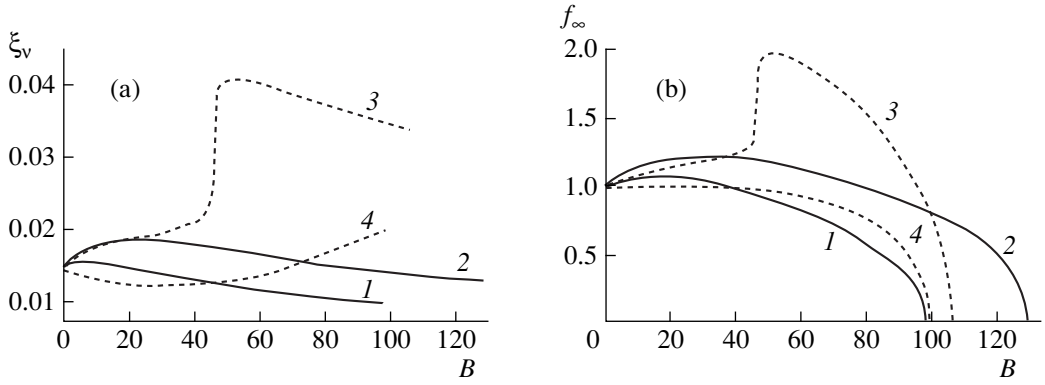

Fig. 1. Curves of $\xi_\nu(B)$ (a) and $f_\infty(B)$ (b) for $\kappa = 100$; (*1*) obtained using the Wigner–Seitz approximation, (*2*) obtained using the isolated vortex approximation, (*3*) using the Hao–Clem continuous approximation, and (*4*) using formulas (24) given in [11].

The expression for F_{core} is derived from Eqs. (4) and (9) and was calculated by Hao and Clem [11]:

$$F_{\text{core}} = \frac{1}{2}(1-f_\infty^2)^2 + \frac{1}{2}B\kappa\xi_\nu^2 f_\infty^2\left[(1-f_\infty^2)\ln\left(1+\frac{2}{B\kappa\xi_\nu^2}\right)\right] + \frac{f_\infty^4}{2} - \frac{f_\infty^4}{2+B\kappa\xi_\nu^2} + \frac{Bf_\infty^2(1+B\kappa\xi_\nu^2)}{\kappa(2+B\kappa\xi_\nu^2)^2}. \quad (12)$$

Thus, we have obtained the dependence of the free energy density of the vortex lattice on the magnetic induction and the variational parameters ξ_ν and f_∞. In order to achieve self-consistency in the theory the dependences $\xi_\nu(\kappa, B)$ and $f_\infty(\kappa, B)$ should be obtained by numerically minimizing the function $F(\kappa, B, \xi_\nu, f_\infty)$ with respect to ξ_ν and f_∞. Figures 1a and 1b gives the curves $\xi_\nu(B)$ and $f_\infty(B)$ plotted for the case $\kappa = 100$ (curves *1*). The numerical calculations show that for arbitrary values of κ they can be approximated by the following formulas:

$$\xi_\nu(B,\kappa) = \xi_{\nu 0}\left[\left(1-4.3\left(1.01-\frac{B}{1.05\kappa}\right)^{6.3}\left(\frac{B}{\kappa}\right)^{0.98}\right)\times\left(1-0.56\left(\frac{B}{\kappa}\right)^{0.9}\right)\right]^{1/2}, \quad (13)$$

$$f_\infty(B,\kappa) = \left(1-\frac{B^2}{2.8\kappa^2}\right)\times\left[\left(1+\frac{1.7B}{\kappa}\left(1-\frac{1.4B}{\kappa}\right)^2\right)\left(1-\left(\frac{B}{s\kappa}\right)^4\right)\right]^{1/2}, \quad (14)$$

where the constant is $s = 0.985$ and $\xi_{\nu 0}$ is the value of the parameter ξ_ν at $B = 0$. This value is obtained from the condition $\partial F/\partial\xi_\nu = 0$ for $B = 0$:

$$\kappa\xi_{\nu 0} = \sqrt{2}\left[1-\frac{K_0^2(\xi_{\nu 0})}{K_1^2(\xi_{\nu 0})}\right], \quad (15)$$

from which it follows that $\xi_{\nu 0} \approx \sqrt{2}/\kappa$ for $\kappa \gg 1$.

The magnetic field H is determined from the condition for minimum Gibbs thermodynamic potential $G = F - 2BH$:

$$H = \frac{1}{2}\frac{\partial F}{\partial B}. \quad (16)$$

For the magnetization we then have

$$M = \frac{B-H}{4\pi} = -\frac{1}{8\pi}\frac{\partial}{\partial B}(F-B^2). \quad (17)$$

The magnetization can be conveniently expressed in the form $M = M_{\text{core}} + M_{\text{em}}$ where the terms

$$M_{\text{core}} = -\frac{\partial F_{\text{core}}}{\partial B}, \quad M_{\text{em}} = -\frac{1}{8\pi}\frac{\partial}{\partial B}(F_{\text{em}}-B^2) \quad (18)$$

are the contributions made to the total magnetization by the energy associated with the change in the order parameter at the vortex core and the electromagnetic energy.

For M_{em} using Eq. (11) we then obtain the following expression:

$$-4\pi M_{\text{em}} = \frac{f_\infty}{2\kappa\xi_\nu}\times\frac{K_0(f_\infty\xi_\nu)I_1(f_\infty\rho)+I_0(f_\infty\xi_\nu)K_1(f_\infty\rho)}{K_1(f_\infty\xi_\nu)I_1(f_\infty\rho)-I_1(f_\infty\xi_\nu)K_1(f_\infty\rho)} + \frac{1}{2B\kappa^2\xi_\nu^2\rho^2}\times[K_1(f_\infty\xi_\nu)I_1(f_\infty\rho)-I_1(f_\infty\xi_\nu)K_1(f_\infty\rho)]^{-2} - B. \quad (19)$$

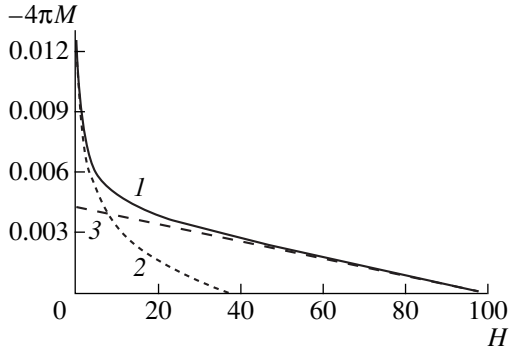

Fig. 2. Curves of $-4\pi M(H)$ over the entire range of fields $H_{c1} \le H \le H_{c2}$ for $\kappa = 100$ obtained using different approximations: (*1*) Wigner–Seitz approximation; (*2*) London approximation [using the Fetter formula (1)] allowing for the exact value of H_{c1}, and (*3*) Abrikosov approximation for strong fields.

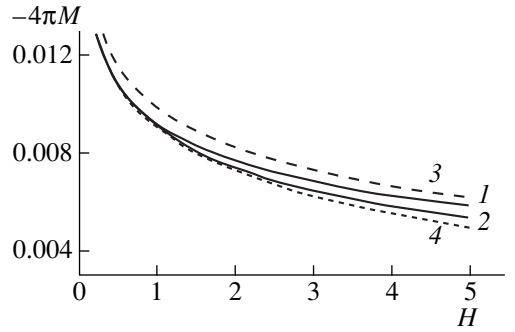

Fig. 3. Curves of $-4\pi M(H)$ in weak fields $H_{c1} \le H \le 0.05H_{c2}$ for $\kappa = 100$ obtained using different approximations: (*1*) Wigner–Seitz approximation, (*2*) isolated vortex approximation, (*3*) Hao–Clem continuous approximation, (*4*) London approximation [using the Fetter formula (1)] allowing for the exact value of H_{c1}.

The relationship for $-4\pi M_{\text{core}}$ from [11] still holds and, in accordance with (9), has the form

$$-4\pi M_{\text{core}} = \frac{\kappa f_\infty^2 \xi_\nu^2}{2}\left[\frac{1-f_\infty^2}{2}\ln\left(\frac{2}{B\kappa\xi_\nu^2}+1\right) - \frac{1-f_\infty^2}{2+B\kappa\xi_\nu^2} + \frac{f_\infty^2}{(2+B\kappa\xi_\nu^2)^2}\right] + \frac{f_\infty^2(2+3B\kappa\xi_\nu^2)}{2\kappa(2+B\kappa\xi_\nu^2)^3}. \quad (20)$$

The dependences $\xi_\nu(\kappa, B)$ and $f_\infty(\kappa, B)$ are determined by Eqs. (13) and (14). According to Eq. (17), we have $H(B) = B - 4\pi M$. Thus we obtained the implicit dependence $M(H)$.

We shall consider the limiting case $\kappa \gg 1$. Then, in the range of fields $H \ll H_{c2}$ the variational parameters can be assumed to be constant: $\xi_\nu = \xi_{\nu 0} = \sqrt{2}/\kappa \ll 1$, $f_\infty = 1$ and it also follows from (20) that M_{core} can be assumed to be constant. In this case, Eqs. (19) and (20) can be expanded in powers of ξ_ν and we obtain the well-known expression [9]:

$$-4\pi M(B) = H_{c1} + \frac{1}{2\kappa}\left[\frac{K_1(R)}{I_1(R)} + \frac{1}{2I_1^2(R)}\right] - B + O(\xi_\nu^2). \quad (21)$$

This expression holds as far as H_{c1} where the Fetter dependence (1) diverges. If $H \gg H_{c1}$, we find $R \ll 1$. In this case, the expression for $-4\pi M$ can be expanded not only in powers of ξ_ν but also in powers of R and we obtain expression (1) with H_{c1} in the form (2) with $\varepsilon \approx 0.52$.

Figure 2 gives the curve of $-4\pi M(H)$ calculated using Eqs. (19) and (20) for $\kappa = 100$ (curve *1*). Figure 3 gives this curve in weak fields (curve *1*). In weak fields the dependence is the same as the Fetter curve (Fig. 2, curve *2*, Fig. 3, curve *4*). In strong fields it shows good agreement with the Abrikosov result (3) (Fig. 2, curve *3*). Note that the Abrikosov expression fairly accurately describes the behavior of the magnetization as far as fields of around $0.4H_{c2}$. In fields close to H_{c1} where the Fetter formula cannot be applied, our dependence agrees with the calculations [14] for the London model. Thus, in order to calculate the magnetization in weak fields for $\kappa \gg 1$ we can use the London approximation provided that we allow for the correct value of H_{c1}. In fact, in weak fields in the London approximation, the influence of the structure of the order parameter inside the cores of vortex filaments on the self-energy of each filament can be taken into account by introducing the exact value of H_{c1}. At the same time when $\kappa \gg 1$ the structure of the order parameter has a negligible influence on the filament interaction energy because the distances between neighboring filaments are many times greater than their core dimensions. Our result agrees with the conclusion reached by Hao and Clem [11, 12] that the London approximation is inaccurate even in weak fields for $\kappa \gg 1$. The authors of [11] used an inexact value of H_{c1} in the Fetter expression (1) (which is equivalent to using an inexact value of the self-energy of an isolated vortex). In addition, an approximation for the electromagnetic energy was used in [11]. We shall show that the error associated with this approximation is also significant.

Figure 4 gives the curve $M(H)$ for $\kappa = 5$ (curve *1*). Note that in this case the magnetization can only be described using the approximation of independent vortices near H_{c1}. As for large κ, the Abrikosov dependence (3) (Fig. 4, curve *2*) remains valid as far as fields around $0.4H_{c2}$.

The upper critical field in the variational model is defined as the field at which the order parameter in the entire superconductor becomes zero. According to the approximation (14), f_∞ (and thus the order parameter) is zero for $H \approx 0.985\kappa$. This value is fairly close to the true

value $H_{c2} = \kappa$. The small difference between H_{c2} and κ can be attributed to the approximate nature of the variational model. As a result, the values of the first and second critical fields calculated using this model cannot be identically equal to the true values H_{c1} and H_{c2}. Thus, although the value of the lower critical field is fairly close to the exact value H_{c1}, it still differs from it. The same applies to the upper critical field.

For practical application of the formula for the magnetization, we can set $s = 1$ in (14). This leads to better agreement between the dependence obtained and the Abrikosov expression (3) for $H \longrightarrow H_{c2}$. In fields below H_{c2} our result remains the same. Small differences from the Abrikosov result for $H \longrightarrow H_{c2}$ can be explained by the fact that the variational model uses a circular vortex cell.

Thus, in the asymptotic limits of weak and strong fields this dependence of the magnetization on the magnetic field agrees with the well-known results: the London dependence (1) (for $\kappa \gg 1$) and the Abrikosov result (3) (in a wide range of values $\kappa > 1/\sqrt{2}$). In this model the values of the first and second critical fields are fairly close to the true values of H_{c1} and H_{c2}. This suggests that this dependence accurately describes the behavior of the magnetization of type-II superconductors in the mixed state.

The magnetic properties of anisotropic superconductors are described by Ginzburg–Landau equations with an effective mass tensor. The variational model can easily be generalized to this case if the vortices are oriented along one of the principal axes of the crystal x_i, $i = 1, 2, 3$. Note that these directions of the external magnetic field are usually used in experimental studies. It was shown in [35] (see also [11]) that in this case, the anisotropic Ginzburg–Landau equations can be transformed to the isotropic form by means of a simple scaling transformation. In order to obtain the dependence of the magnetization on the external field in the anisotropic case from the known dependence for an isotropic superconductor, we need to replace the Ginzburg–Landau parameter κ with $\tilde{\kappa}_\alpha = \kappa u_\alpha^{-1/2}$, where $u_\alpha = m_\alpha/\sqrt{m_1 m_2 m_3}$, and m_i are the effective masses in the direction of the x_i axis (here the vortices are directed along the x_α axis). The case of arbitrary orientation of the vortex filaments relative to the principal axes is studied in [24].

3. APPROXIMATION OF ISOLATED VORTICES

The field h_0 generated by an isolated vortex is a decreasing solution of the Ginzburg–Landau equation (5) over large distances. If the order parameter in the entire superconductor is distributed according to Eq. (4), taking into account the flux quantization condition, we can find the dependence of the field h_0 on the distance from the vortex axis r from (5) [11]:

$$h_0(r) = \frac{f_\infty K_0(f_\infty \sqrt{r^2 + \xi_v^2})}{\kappa \xi_v K_1(f_\infty \xi_v)}. \qquad (22)$$

Note that this formula is a particular case of Eq. (6) for $B \longrightarrow 0$.

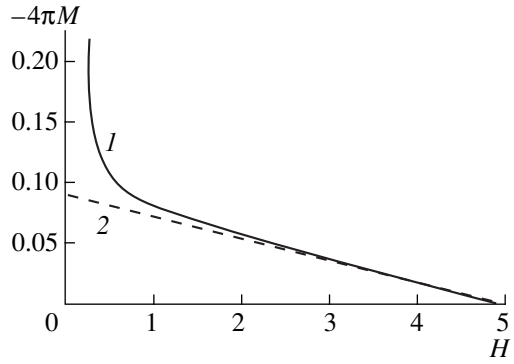

Fig. 4. Curves of $-4\pi M(H)$ in fields $H_{c1} \le H \le H_{c2}$ for $\kappa = 5$ obtained using various approximations: (*1*) Wigner–Seitz approximation; (*2*) Abrikosov result (3).

The trial function (4) was used in [11, 12] to model the distribution of the order parameter in each unit cell of a regular vortex lattice. In these studies the local magnetic field over the entire range of fields $H_{c1} < H < H_{c2}$ was obtained from the sum of the contributions of isolated cells. These contributions should be calculated from the second Ginzburg–Landau equation for a given periodic distribution of the order parameter. Instead it was assumed in [11] that each contribution at an arbitrary point in the superconductor is given by Eq. (22) which is valid for an isolated vortex. In this approximation the local magnetic field $h_l(\mathbf{r})$ may be obtained by the simple superposition:

$$h_l(\mathbf{r}) = \sum_i h_0(|\mathbf{r} - \mathbf{r}_i|). \qquad (23)$$

This approach retains the concept of vortices as isolated objects. We shall therefore call it the "isolated vortex approximation."

As in the case of the London approximation, the approach described above should only remain exact in weak fields. Unlike the Wigner–Seitz model, this approximation can be used to study various vortex lattice configurations and also to study the vortex state near the surface. We shall analyze the range of validity of this approach. To do this we shall calculate the superconductor magnetization in this approximation over the entire range of external fields $H_{c1} \le H \le H_{c2}$ and we shall compare this with known results.

As before, the electromagnetic energy density in this case is determined by the formula $F_{em} = Bh(0)$ and the magnetic field at the center of the vortex $h(0)$ is

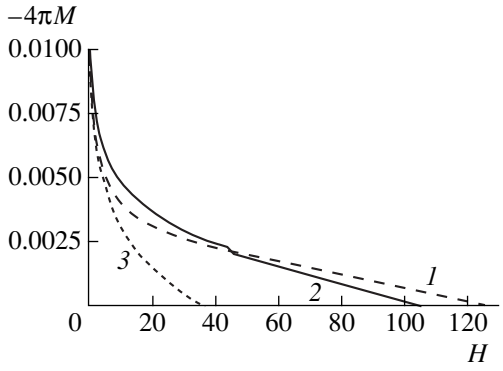

Fig. 5. Curves of $-4\pi M(H)$ in fields $H_{c1} \le H \le H_{c2}$ for $\kappa = 100$ calculated using various approximations: (*1*) isolated vortex approximation; (*2*) Hao–Clem continuous approximation, (*3*) London approximation [using the Fetter formula (1)] allowing for the exact value of H_{c1}.

made up of the self-induced field of the vortex and the fields generated by all the other vortices. In this case, for $F_{\rm em}$ for a triangular lattice we have

$$F_{\rm em} = \frac{B f_\infty}{\kappa \xi_\nu K_1(f_\infty \xi_\nu)} \times \sum_{m=-\infty}^{\infty} \sum_{n=-\infty}^{\infty} [K_0(f_\infty \sqrt{d_f^2(3m^2+n^2)+\xi_\nu^2}) + K_0(f_\infty \sqrt{d_f^2(m+1/2)^2 + d_f^2(n+1/2)^2 + \xi_\nu^2})], \quad (24)$$

where $d_f = (4\pi/B\kappa\sqrt{3})^{1/2}$ is the vortex lattice constant. Using Eq. (24), we find the dependence of the magnetization on the magnetic field. The values of the variational parameters are determined by numerical minimization of the free energy density. However, this procedure using Eq. (24) directly is difficult since in strong fields we need to have thousands of terms in each of the single sums to achieve the required accuracy. In the Appendix, Eq. (24) is transformed to a more suitable form for the numerical calculations.

A numerical minimization of the free energy gives the dependences $\xi_\nu(B)$ and $f_\infty(B)$ plotted in Figs. 1a and 1b for $\kappa = 100$ (curves *2*). The corresponding curve $-4\pi M(H)$ is plotted in Fig. 3 (curve *2*) and Fig. 5 (curve *1*) for $\kappa = 100$. In this case, the value of the second critical field H_{c2} is higher than the correct value, being approximately $1.29H_{c2}$. Thus, in this approximation the behavior of the magnetization near H_{c2} does not agree with the Abrikosov result because the concept of isolated vortices becomes meaningless here as a result of the substantial overlap of the vortex cores.

For comparison the Fetter curve (*1*) is plotted in Fig. 3 (curve *4*) and Fig. 5 (curve *3*). In weak fields the dependence $M(H)$ in the isolated vortex approximation is almost the same as the corresponding curve in the London approximation (for $\kappa = 100$ as far as fields $H \approx 0.02H_{c2}$ which is approximately $80H_{c1}$). At the same time, a comparison of the magnetization curves (Fig. 3, curves *1* and *2*) shows that even in fairly weak fields $H \approx 0.05H_{c2}$ for $\kappa = 100$, when the spacings between the vortices are still large, the error associated with using the isolated vortex approximation is quite appreciable and is around 10%. This is because, in accordance with formula (17), the magnetization is determined by the difference between two numbers, each many times greater than the magnetization itself. Thus, even small corrections to F_{em} may be significant.

To conclude this section we note that for $H \ll H_{c2}$ we have calculated the free energy of a square vortex lattice in the isolated vortex approximation. As was to be predicted, this was higher than the free energy of a triangular lattice. Thus, in the London approximation a triangular lattice is thermodynamically more favorable than a square one.

4. CONTINUOUS APPROXIMATION

In [11] a transition was made to summation over the reciprocal lattice to calculate the free energy F_{em}, followed by a transition to the continuous limit, i.e., the sum was approximated by an integral. For $\kappa \gg 1$ the following approximations were given in [11] for $\xi_\nu(B)$ and $f_\infty(B)$:

$$\xi_\nu = \xi_{\nu 0}\sqrt{\left(1-2\left(1-\frac{B}{\kappa}\right)^2\frac{B}{\kappa}\right)\left(1+\left(\frac{B}{\kappa}\right)^4\right)}, \qquad f_\infty = \sqrt{1-\left(\frac{B}{\kappa}\right)^4}. \quad (25)$$

It can be seen from (25) that $f_\infty = 0$ for $B = H_{c2} = \kappa$. A numerical check shows that the formulas (25) do not follow from the expression for the free energy density in the continuous approximation [11]. Minimizing F in terms of the parameters ξ_ν and f_∞ gives the curves $\xi_\nu(B)$ and $f_\infty(B)$ plotted in Figs. 1a and 1b for the case $\kappa = 100$ (curves *3*). The dependences (25) are also plotted in these figures (curves *4*).

It can be seen from Figs. 1a and 1b that in the field $H_{cr} \approx H_{c2}/2$ the dependences $\xi_\nu(B)$ and $f_\infty(B)$ exhibit an abrupt jump. This jump leads to a small jump in the magnetization $M(H)$ which is nevertheless incorrect from the physical point of view (Fig. 5, curve *2*). Such an abrupt change in $\xi_\nu(B)$ and $f_\infty(B)$ occurs because in addition to the absolute minimum of the free energy, a local minimum occurs near the field H_{cr}. As B varies, the absolute and local extrema abruptly change places. It can be seen from Figs. 1b (curve *3*) and Fig. 5 (curve *2*) that the field for which the order parameter in the superconductor becomes zero differs from the true value $H_{c2} = \kappa$. For $\kappa = 100$ this difference is approximately 6.6%.

More significantly, using this approximation for F_{em} gives an error in the behavior of $M(H)$ in weak fields.

Thus, the difference between the values of the magnetization obtained using the continuous approximation (Fig. 3, curve *3*) and the self-consistent (Wigner–Seitz) and London approaches (Fig. 3, curves *1* and *4*) is around 10% for $H \approx 0.01H_{c2}$, $\kappa = 100$. It can be seen from Fig. 3 that a similar difference does not occur when the isolated vortex approximation is used systematically (curve *2*). Using the approximations (25) barely alters the behavior of the magnetization in weak fields although it eliminates the jump in $M(H)$ and more accurately describes the behavior of the magnetic moment near H_{c2}.

Consequently, the difference between the magnetization in weak fields in the Hao–Clem model and the magnetization in the London approximation for $\kappa \gg 1$ is a result of the inaccuracy of the approximation [11] for the electromagnetic energy.

5. CONCLUSIONS

Thus, we have used a variational model which allows for the structure of the order parameter inside the vortex core and the dependence of the modulus of the order parameter on the magnetic field to determine the magnetization of a homogeneous isotropic type-II superconductor in the mixed state over the entire range of magnetic fields $H_{c1} \le H \le H_{c2}$. The model has been generalized to the case of anisotropic superconductors when the filaments are oriented along one of the principal axes of the crystal. In weak fields when $\kappa \gg 1$ our results agree with the predictions of the London model, while in fields near the second critical value they agree with the well-known results obtained by Abrikosov. The proposed model is self-consistent and can be applied for a quantitative description of the magnetization of type-II superconductors. We have also analyzed the accuracy of representing the mixed state as a set of isolated vortices for $\kappa \gg 1$. We have shown that calculating the magnetization using these representations yields appreciable quantitative errors in weak fields.

ACKNOWLEDGMENTS

We thank L.G. Mamsurova, K.S. Pigal'skiĭ, and N.G. Trusevich for useful discussions, and also V.K. Manukyan for assistance with the numerical calculations. This work was supported financially by the scientific-technical program "Topical Problems in Physics of the Condensed State ("Superconductivity"), the Russian Foundation for Basic Research (projects nos. 00-15-96570, 00-02-18032), and also by the joint program INTAS–RFBR (project no. IR-97-1394).

APPENDIX

Electromagnetic Energy Density of a Vortex Lattice in the Isolated Vortex Approximation

We shall express the field b created by a single vortex array in a convenient form for numerical calculations. We introduce coordinates centered on the axis of one of the vortices with the x-axis perpendicular to the plane in which the vortex array lies, the y-axis lying in this plane and orthogonal to the vortices, and the z-axis directed along the vortex axes. In this case, using (22) we obtain

$$b(x, y) = \frac{f_\infty}{\kappa\xi_\nu K_1(f_\infty\xi_\nu)} \times \sum_{m=-\infty}^{\infty} K_0\left(f_\infty\sqrt{(d_f m + y)^2 + x^2 + \xi_\nu^2}\right). \tag{A.1}$$

Let us perform Fourier transformation of a summand in Eq. (A.1) with respect to the coordinate y:

$$b(x, y) = \frac{f_\infty}{\kappa\xi_\nu K_1(f_\infty\xi_\nu)} \sum_{m=-\infty}^{\infty} \int \frac{dq\,dy}{2\pi} \times \exp(-iqy) K_0\left(f_\infty\sqrt{(d_f m + y)^2 + x^2 + \xi_\nu^2}\right).$$

Having integrated this expression with respect to y, using known formulas for the definite integrals of the Bessel functions [36] and the relationships

$$\sum_{m=-\infty}^{\infty} \exp(-iqm) = 2\pi \sum_{m=-\infty}^{\infty} \delta(q - 2\pi m),$$

we can obtain

$$b(x, y) = \frac{f_\infty}{\kappa\xi_\nu K_1(f_\infty\xi_\nu)} \frac{\pi}{d_f} \sum_{m=-\infty}^{\infty} \left(f_\infty^2 + \frac{4\pi^2 m^2}{d_f^2}\right)^{-1/2} \times \exp\left(-\sqrt{x^2 + \xi_\nu^2}\sqrt{f_\infty^2 + \frac{4\pi^2 m^2}{d_f^2}}\right) \cos\left(\frac{2\pi m y}{d_f}\right). \tag{A.2}$$

Taking into account Eqs. (24) and (A.2), we have

$$F_{\text{em}} = \frac{B\pi}{\kappa\xi_\nu K_1(f_\infty\xi_\nu) d_f \sqrt{3}} \times \left\{ 2\sum_{n=1}^{\infty} \exp\left(-f_\infty\sqrt{3d_f^2 n^2 + \xi_\nu^2}\right) + \exp(-f_\infty\xi_\nu) + 2\sum_{m=1}^{\infty} \left(1 + \frac{4\pi^2 m^2}{3d_f^2 f_\infty^2}\right)^{-1/2} \exp\left(-\xi_\nu\sqrt{f_\infty^2 + \frac{4\pi^2 m^2}{3d_f^2}}\right) + 4\sum_{m=1}^{\infty}\sum_{n=1}^{\infty} \left(1 + \frac{4\pi^2 m^2}{3d_f^2 f_\infty^2}\right)^{-1/2} \times \exp\left(-\sqrt{f_\infty^2 + \frac{4\pi^2 m^2}{3d_f^2}}\sqrt{d_f^2 n^2 + \xi_\nu^2}\right)$$

$$+4\sum_{m=1}^{\infty}\sum_{n=0}^{\infty}(-1)^m\left(1+\frac{4\pi^2m^2}{3d_f^2f_\infty^2}\right)^{-1/2}$$

$$\times\exp\left(-\sqrt{f_\infty^2+\frac{4\pi^2m^2}{3d_f^2}}\sqrt{d_f^2(n+1/2)^2+\xi_\nu^2}\right)$$

$$+2\sum_{m=0}^{\infty}\exp(-f_\infty\sqrt{d_f^2(n+1/2)^2+\xi_\nu^2})\Bigg\}. \qquad \text{(A.3)}$$

Although this formula is fairly cumbersome, it is more convenient for numerical calculations than Eq. (24) since only a few terms (several tens) need be taken into account in the double sums on the right-hand side of Eq. (A.3).

REFERENCES

1. A. A. Abrikosov, Zh. Éksp. Teor. Fiz. **32**, 1442 (1957) [Sov. Phys. JETP **5**, 1174 (1957)].
2. P. G. de Gennes, *Superconductivity of Metals and Alloys* (Benjamin, New York, 1966; Mir, Moscow, 1968).
3. D. St. James, G. Sarma, and E. J. Thomas, *Type II Superconductivity* (Pergamon, Oxford, 1969; Mir, Moscow, 1970).
4. E. H. Brandt and U. Essman, Phys. Status Solidi B **144**, 13 (1987).
5. E. H. Brandt, Rep. Prog. Phys. **58**, 1465 (1995).
6. A. L. Fetter, Phys. Rev. **147**, 153 (1966).
7. C.-R. Hu, Phys. Rev. B **6**, 1756 (1972).
8. E. A. Shapoval, Pis'ma Zh. Éksp. Teor. Fiz. **69**, 532 (1999) [JETP Lett. **69**, 577 (1999)].
9. D. Ihle, Phys. Status Solidi B **47**, 423 (1971).
10. J. R. Clem, J. Low Temp. Phys. **18**, 427 (1975).
11. Z. Hao, J. R. Clem, M. W. Elfresh, *et al.*, Phys. Rev. B **43**, 2844 (1991).
12. Z. Hao and J. R. Clem, Phys. Rev. Lett. **67**, 2371 (1991).
13. I. G. de Oliveira and A. M. Thompson, Phys. Rev. B **57**, 7477 (1998).
14. A. S. Krasilnikov, L. G. Mamsurova, N. G. Trusevich, *et al.*, Supercond. Sci. Technol. **8**, 1 (1995).
15. H. Koppe and J. Willebrand, J. Low Temp. Phys. **2**, 499 (1970).
16. Z. Hao and J. R. Clem, Phys. Rev. B **43**, 7622 (1991).
17. J. H. Gohng and D. K. Finnemore, Phys. Rev. B **46**, 398 (1992).
18. W. Chen, J. Gohng, D. K. Finnemore, *et al.*, Phys. Rev. B **51**, 6035 (1995).
19. Q. Li, M. Suenaga, D. K. Finnemore, *et al.*, Phys. Rev. B **46**, 3195 (1992).
20. D. N. Zheng, A. M. Campbell, and R. S. Liu, Phys. Rev. B **48**, 6519 (1993).
21. V. C. Kim, J. R. Thompson, J. G. Ossandon, *et al.*, Phys. Rev. B **51**, 11767 (1995).
22. M.-S. Kim, M. K. Bae, W. C. Lee, *et al.*, Phys. Rev. B **51**, 3261 (1995).
23. Y. Zhuo, J.-H. Choi, M.-S. Kim, *et al.*, Phys. Rev. B **55**, 12719 (1997).
24. A. Nugroho, I. M. Sutjahja, A. Rusydi, *et al.*, Phys. Rev. B **60**, 15384 (1999).
25. V. I. Shmidt and G. S. Mkrtchyan, Usp. Fiz. Nauk **112**, 459 (1974) [Sov. Phys. Usp. **17**, 170 (1974)].
26. V. G. Kogan, A. Gurevich, J. H. Cho, *et al.*, Phys. Rev. B **54**, 12386 (1996).
27. J. R. Clem, in *Superconducting Electronics*, Ed. by H. Weinstock and M. Nisenoff (Springer-Verlag, Berlin, 1989), p. 1.
28. E. H. Brandt, Phys. Status Solidi B **51**, 345 (1972).
29. E. H. Brandt, Phys. Rev. Lett. **78**, 2208 (1997).
30. D. Ihle, Phys. Status Solidi B **47**, 429 (1971).
31. R. J. Watts-Tobin, L. Kramer, and W. Pesch, J. Low Temp. Phys. **17**, 71 (1974).
32. J. Rammer, W. Pesch, and L. Kramer, Z. Phys. B **68**, 49 (1987).
33. J. Rammer, J. Low Temp. Phys. **71**, 323 (1988).
34. A. I. Larkin and Yu. N. Ovchinnikov, Phys. Rev. B **51**, 5965 (1995).
35. V. G. Kogan and J. R. Clem, Phys. Rev. B **24**, 2497 (1981).
36. I. S. Gradshteyn and I. M. Ryzhik, *Table of Integrals, Series, and Products* (Nauka, Moscow, 1971; Academic, New York, 1980).

Translation was provided by AIP

Journal of Experimental and Theoretical Physics, Vol. 91, No. 3, 2000, pp. 597–609.
Translated from Zhurnal Éksperimental'noĭ i Teoreticheskoĭ Fiziki, Vol. 118, No. 3, 2000, pp. 687–700.

SOLIDS
Electronic Properties

Multiple-Quantum Dynamics of One-Dimensional Nuclear Spin Systems in Solids

S. I. Doronin*, I. I. Maksimov, and E. B. Fel'dman*****

Institute of Chemical Physics, Russian Academy of Sciences, Chernogolovka, Moscow oblast, 142432 Russia
**e-mail: serge@pro.icp.ac.ru*
***e-mail: iimxv@camille.icp.ac.ru*
****e-mail: feldman@feb.uupc.chg.ru*

Received April 27, 2000

Abstract—Multiple-quantum spin dynamics is studied using analytic and numerical methods for one-dimensional finite linear chains and rings of nuclear spins 1/2 coupled by dipole–dipole interactions. An approximation of dipole–dipole interaction between nearest neighbors having the same constants is used to obtain exact expressions for the intensities of the multiple-quantum coherences in the spin systems studied, which are initially in thermal equilibrium and whose evolution is described by a two-spin two-quantum Hamiltonian. An approximation of nearest neighbors with arbitrary dipole–dipole interaction constants is used to establish a simple relationship between the multiple-quantum dynamics and the dynamics of spin systems with an *XY* Hamiltonian. Numerical methods are developed to calculate the intensities of multiple-quantum coherences in multiple-quantum NMR spectroscopy. The integral of motion is obtained to expand the matrix of the two-spin two-quantum Hamiltonian into two independent blocks. Using the nearest-neighbor approximation the Hamiltonian is factorized according to different values of the projection operator of the total spin momentum on the direction of the external magnetic field. Results of calculations of the multiple-quantum dynamics in linear chains of seven and eight nuclear spins and a six-spin ring are presented. It is shown that the evolution of the intensities of the lowest-order multiple-quantum coherences in linear chains is accurately described allowing for dipole–dipole interaction of nearest and next-nearest neighbors only. Numerical calculations are used to compare the contributions of nearest and remote spins to the intensities of the multiple-quantum coherences. © *2000 MAIK "Nauka/Interperiodica".*

1. INTRODUCTION

As a result of the local nature of the main spin–spin interactions, NMR spectroscopy in solids can only be used to obtain structural information on the nearest neighborhoods of the nuclei being studied which restricts the scope for studying structural characteristics at distances greater than 6 Å [1]. In multiple-quantum NMR spectroscopy, various multiple-spin transitions are excited between the Zeeman levels of a system of interacting spins in an external magnetic field which in principle can be used to obtain information on the structure at considerably greater distances. In this case, information on the nuclear spin distribution in the solid is extracted by determining the number $N(\tau)$ of spins correlated as a result of spin–spin interactions at the end of a preparation period of duration τ when the system is exposed to a specially selected sequence of rf pulses which ensure the appearance and evolution of various multiple-quantum coherences [2]. Multiple-quantum NMR has been successfully used to study proton distributions in liquid crystals [3], in simple organic systems [2], and in light-sensitive polymer mixtures [4]. The time evolution of the multiple-quantum coherences has proved to be very sensitive to the spatial dimension of the system of interacting nuclear spins [5], which opens up new prospects for structural investigations.

The main theoretical method for interpreting multiple-quantum NMR experiments is the statistical theory [2]. According to [2], for a fairly long irradiation time the probability of the excitation of all possible multiple-spin transitions is the same. Determining the intensities of the multiple-quantum coherences then simply reduces to a combinatorial problem. In a group of N interacting nuclear spins ($S = 1/2$) the number of transitions responsible for multiple-quantum coherence of order n is C_{2N}^{N-n}. According to the Stirling formula [6] this expression for $N > 6$ is accurately approximated by the Gaussian function $2^{2N}(\pi N)^{-1/2}\exp(-n^2/N)$. By approximating the experimentally observed dependence of the multiple-quantum coherence intensities on the time τ [2] (the multiple-quantum coherence profile) by a Gaussian curve, we can find the number of spins $N(\tau)$ coupled by spin–spin interactions. In a small cluster the value of $N(\tau)$ corresponds to the number of nuclear spins contained in it. However, in solids with a macroscopically large number of spins this definition becomes meaningless. The initial hypothesis of the statistical theory [2] that the probabilities of excitation of various multiple-quantum coherences are the same for

long times of excitation of the system by a series of pulses (compared with the time determined by the spin–spin interactions) remains unclear. Although the experimental data for spin systems with a small number of spins ($6 < N < 8$) agree qualitatively with the predictions of the statistical theory [7], a more detailed analysis of multiple-quantum NMR experiments revealed substantial discrepancies [8]. An experimental investigation [9] of the profiles of multiquantum coherences showed that in some cases, these profiles are in fact described by Gaussian curves and in other cases, the multiple-quantum coherences decay exponentially, as was noted in [8]. It should be stressed that it is very difficult to uniquely determine the correspondence between the profiles of multiple-quantum coherences and the predictions of the statistical theory since we are dealing with experimental signals obtained for long excitation times when the imperfections of the pulse sequences used and relaxation processes are important factors.

The phenomenological theory [10, 11] of multiple-quantum NMR spectroscopy, like the theory of the NMR line shape in systems with chemical exchange [12], has played an important role in understanding various aspects of multiple-quantum dynamics. In this theory, multiple-spin, multiple-quantum dynamics is considered as multipositional exchange of coherences in Liouville space. A hopping model [10, 11] was used to describe some features of multiple-quantum NMR experiments. However, an explanation of the oscillating nature of multiple-quantum dynamics was outside the scope of phenomenological theory.

Multiple-quantum coherences form at times $t > \omega_{loc}^{-1}$ (ω_{loc} is the characteristic frequency of the spin–spin interactions). Hence, perturbation theory methods cannot be used effectively to study multiple-spin, multiple-quantum dynamics. This leaves us with numerical methods of multiple-quantum NMR and exactly solvable models.

A numerical analysis of the evolution (growth) of multiple-quantum coherences is extremely difficult since for a system of N spins the density matrix contains 4^N elements. Numerical solutions describing the growth of multiple-quantum coherences have been obtained for systems containing up to six spins by diagonalizing the Hamiltonian [13]. An iterative method [14] was used for a numerical study of the growth of multiple-quantum coherences in a nine-spin system. However, the accuracy of the results obtained in [14] for long times requires additional verification.

Multiple-spin multiple-quantum dynamics is simplified appreciably in one-dimensional spin systems. Thus, an experimental investigation [1, 15] of multiple-quantum NMR in calcium hydroxyapatite $Ca_5(OH)(PO_4)_3$ containing quasi-one-dimensional chains of hydroxyl-group protons opened up a new direction for the evolution of theoretical methods of analyzing multiple-spin multiple-quantum dynamics. Since the constant of the dipole–dipole interactions of neighboring nuclear spins in a chain is 20 times the maximum constant of dipole–dipole interactions of neighboring chains, we can assume that in calcium hydroxyapatite the main contribution to the intensity of the multiple-quantum coherences is made by the nuclear spins of one-dimensional linear chains. It is also significant that the dipole–dipole interaction of the spins decreases with distance r as r^{-3} and thus the dipole–dipole interaction of the next nearest neighbors is eight times weaker than the interactions of the nearest neighbors. This gives rise to the problem of multiple-quantum NMR of one-dimensional linear chains of nuclear spins in which only nearest neighbors interact. This problem was solved exactly in the thermodynamic limit when the number of spins in the chain is $N \longrightarrow \infty$ at high [16, 17] and low [18] temperatures.

Problems involving multiple-quantum NMR of finite linear chains and rings of nuclear spins when allowance is only made for spin–spin interactions of nearest neighbors belong to the exactly solvable one-dimensional models [19] and their solution can be used to study multiple-quantum dynamics of various clusters of nuclear spins.

In the present paper we use analytic and numerical methods to study multiple-quantum spin dynamics for one-dimensional linear chains and rings of nuclear spins. For these systems which are initially in thermal equilibrium, whose evolution is described by a two-spin, two-quantum Hamiltonian [2], we obtain exact formulas for the intensities of multiple-quantum coherences when allowance is only made for nearest-neighbor interactions and the constants of these interactions are assumed to be the same. It is shown that in the nearest-neighbor approximation (for arbitrary dipole–dipole interaction constants) the multiple-quantum dynamics can be reduced by means of a simple transformation to spin dynamics described by the flip-flop Hamiltonian [20]. Thus, an example which can be achieved experimentally was constructed for the first time when as a result of dipole–dipole interaction of nuclear spins in a solid, only the part responsible for flip-flop processes remains. This implies that multiple-quantum dynamics is closely related to the dynamics of systems of nuclear spins with an *XY* Hamiltonian [19, 21, 22]. Numerical methods have also been developed to calculate multiple-quantum coherences in multiple-quantum NMR spectroscopy in solids. These are implemented using an integral of motion which can expand the matrix of the dipole–dipole interactions into two independent blocks. In order to obtain a numerical solution of the problem of multiple-quantum NMR in the nearest-neighbor approximation, the two-spin two-quantum Hamiltonian is factorized in terms of different values of the operator of the total spin momentum on the direction of the external magnetic field. Numerical solutions of the problem of multiple-quantum NMR are obtained for linear chains consisting of seven and eight spins, and a ring of six spins. It is shown that in these

linear chains the time evolution of the zeroth- and second-order multiple-quantum coherences is described allowing only for dipole–dipole interaction of nearest and next-nearest neighbors. The contributions of nearest and remote spins to the intensities of multiple-quantum coherences of various orders are compared on the basis of the numerical calculations.

2. INTENSITIES OF MULTIPLE-QUANTUM COHERENCES IN SOLIDS

Multiple-quantum NMR spectroscopy in solids uses a two-dimensional NMR experiment in which a preparation period of duration τ, free evolution over the time interval t_1, a mixing period τ, and detection are systematically carried out [2]. In the preparation period the system is exposed to a sequence of pulses which leads to the appearance and evolution of multiple-quantum coherences. We shall assume that the exciting sequence is periodic and that one period contains eight rf pulses [2]:

$$\frac{\Delta}{2} - X - \Delta' - X - \Delta - X - \Delta' - X - \Delta - \bar{X} - \Delta' - \bar{X} - \Delta - \bar{X} - \Delta' - \bar{X} - \frac{\Delta}{2}, \quad (1)$$

where Δ, $\Delta' = 2\Delta + t_p$ are the time intervals between pulses (t_p is the pulse duration), X and $\bar{X}$ are resonant pulses having the phase difference π which flip the spins by 90° about the x axis of coordinates rotating at the pulse carrier frequency [23]. Then the average Hamiltonian $\mathcal{H}$ determining the dynamics of the nuclear spin system may be written in the form [17]

$$\mathcal{H} = \mathcal{H}^{+2} + \mathcal{H}^{-2}, \quad (2)$$

where

$$\mathcal{H}^{\pm 2} = -\frac{1}{2}\sum_{j<k} D_{jk} I_j^{\pm} I_k^{\pm}. \quad (3)$$

In (3) $I_j^{\pm}$ are the raising and lowering spin angular momentum operators of spin j. The dipolar coupling constant D_{jk} between spins j and k is given by

$$D_{jk} = \frac{\gamma^2\hbar}{2r_{jk}^3}(1 - 3\cos^2\theta_{jk}), \quad (4)$$

where r_{jk} is the distance between spins j and k, θ_{jk} is the angle between the internuclear vector $\mathbf{r}_{jk}$ and the external magnetic field $\mathbf{H}_0$, and γ is the gyromagnetic ratio. In one-dimensional linear chains, the angle θ_{jk} is the same for all spin pairs. We shall subsequently assume that for these chains $\theta_{jk} \equiv 0$. The distance between nearest neighbors r_{ii+1} can generally differ for different spins. However, in order to find an exact solution, we shall assume that they are the same.

Assuming that a system of nuclear spins is initially in thermal equilibrium with the lattice, at the end of the preparation period, apart from terms which do not depend on τ, in the high-temperature approximation the density matrix $\rho(\tau)$ is given by:

$$\rho(\tau) = \exp(-i\mathcal{H}\tau) I_z \exp(i\mathcal{H}\tau), \quad (5)$$

where I_z is the projection of the total spin angular momentum on the z direction of the external field $\mathbf{H}_0$. Following [17], we expand the density matrix $\rho(\tau)$ as a series:

$$\rho(\tau) = \sum_n \rho_n(\tau), \quad (6)$$

where $\rho_n(\tau)$ is the contribution to $\rho(\tau)$ made by multiple-quantum coherences of order n. Special experimental methods are used to separate the signals from the multiple-quantum coherences of various orders [2, 24, 25]. These methods can be described by introducing the offset field $\Delta\omega$ over the evolution period of the system. Assuming that $\omega_{loc} \ll |\Delta\omega|$ (ω_{loc} is the local dipole frequency determined by the dipole–dipole interaction), we find that the density matrix $\rho(\tau + t_1)$ at the end of the free evolution period has the form

$$\rho(\tau + t_1) = \exp(i\Delta\omega t_1 I_z)\rho(\tau)\exp(-i\Delta\omega t_1 I_z). \quad (7)$$

Taking into account Eq. (7), we rewrite Eq. (6) as follows:

$$\tau(\tau + t_1) = \sum_n \exp(in\Delta\omega t_1)\rho_n(\tau). \quad (8)$$

During the mixing period the multiple-quantum coherences are transformed into longitudinal magnetization. Over this period we use a multipulse sequence which changes the sign of the Hamiltonian (2) as in time-reversal experiments in a system of interacting nuclear spins [26, 27]. This sequence ensures that different contributions to the multiple-quantum coherence of a given order n have the same phases. At the end of the mixing period the density matrix has the form

$$\rho(2\tau + t_1) = \exp(i\mathcal{H}\tau)\exp(i\Delta\omega t_1 I_z)\exp(-i\mathcal{H}\tau) \times I_z \exp(i\mathcal{H}\tau)\exp(-i\Delta\omega t_1 I_z)\exp(-i\mathcal{H}\tau). \quad (9)$$

In this case, the longitudinal polarization $I_z(\tau, t_1)$ is defined as follows [17]:

$$\begin{aligned} I_z(\tau, t_1) &= \mathrm{Tr}\{I_z\rho(2\tau + t_1)\} \\ &= \mathrm{Tr}\{\rho(\tau)\exp(i\Delta\omega t_1 I_z)\rho(\tau)\exp(-i\Delta\omega t_1 I_z)\} \\ &= \sum_{n,m} e^{in\phi}\mathrm{Tr}\{\rho_m(\tau)\rho_n(\tau)\} \\ &= \sum_{n,m} e^{in\phi}\delta_{-m,n}\mathrm{Tr}\{\rho_{-n}(\tau)\rho_n(\tau)\} \end{aligned} \quad (10)$$

$$= \sum_n e^{in\phi}\mathrm{Tr}\{\rho_{-n}(\tau)\rho_n(\tau)\},$$

where $\delta_{-m,n}$ is the Kronecker delta and $\phi = \Delta\omega t_1$. Thus, the intensity of the multiple-quantum coherence of order n is determined using the formula

$$J_n(\tau) = \mathrm{Tr}\{\rho_n(\tau)\rho_{-n}(\tau)\}. \quad (11)$$

3. MULTIPLE-QUANTUM NUCLEAR MAGNETIC RESONANCE IN NUCLEAR SPIN RING SYSTEMS

We shall consider a ring of N nuclear spins in an external magnetic field $\mathbf{H}_0$ perpendicular to the plane of the ring. It follows from Eq. (2) that the Hamiltonian of this system is given by

$$\mathcal{H} = b\sum_{j=1}^{N}(I_j^+ I_{j+1}^+ + I_j^- I_{j+1}^-), \quad (12)$$

where $b = -D_{jj+1}/2$. It is assumed that the cyclicity condition, i.e., $I_j^\pm = I_{j+N}^\pm$ ($j = 1, 2, \ldots, N$) is satisfied in this system. In order to diagonalize the Hamiltonian (12) we need to convert from the spin operators $I_j^\pm$ to the fermion operators Ψ_k^+, Ψ_k which is achieved by means of the Jourdan–Wigner transformation [19]:

$$I_n^+ = \Psi_n \exp\left(-i\pi\sum_{m>n}\Psi_m^+\Psi_m\right),$$
$$I_n^z = \frac{1}{2} - \Psi_n^+\Psi_n. \quad (13)$$

Here, I_n^z is the projection operator of the spin angular momentum of nucleus n on the z axis ($I_z = \sum_{n=1}^N I_n^z$). In the fermion representation, the Hamiltonian $\mathcal{H}$ (12) may be rewritten as follows:

$$\mathcal{H} = b\sum_{n=1}^{N-1}(\Psi_n\Psi_{n+1} + \Psi_{n+1}^+\Psi_n^+)$$
$$- b\prod_{n=2}^{N-1}(1 - 2\Psi_n^+\Psi_n)\Psi_N\Psi_1 \quad (14)$$
$$+ b\prod_{n=2}^{N-1}(1 - 2\Psi_n^+\Psi_n)\Psi_N^+\Psi_1^+.$$

The Hamiltonian (12) does not commute with I_z, but it is easy to check that

$$[\mathcal{H}, \exp(i\pi I_\alpha)] = 0, \quad \alpha = x, y, z. \quad (15)$$

The integral of motion $\exp(i\pi I_z)$ may be used to represent the Hamiltonian $\mathcal{H}$ as two blocks having dimensions of $2^{N-1} \times 2^{N-1}$.[1] It is convenient to introduce the operator

$$M = \frac{N}{2} - I_z, \quad (16)$$

which has an integer-value spectrum (0, 1, 2, …, N). It is easy to check that in the fermion representation we have

$$M = \sum_{k=1}^{N}\Psi_k^+\Psi_k, \quad (17)$$

and using Eq. (17), the formula for $\mathcal{H}$ (14) can be transformed to give

$$\mathcal{H} = b\sum_{n=1}^{N-1}(\Psi_n\Psi_{n+1} + \Psi_{n+1}^+\Psi_n^+)$$
$$- be^{-i\pi M}\Psi_N\Psi_1 + be^{i\pi M}\Psi_N^+\Psi_1^+. \quad (18)$$

In accordance with Eq. (15), $\mathcal{H}$ is divided into two blocks, $\mathcal{H}_{-1}$ and $\mathcal{H}_1$ for $e^{i\pi M} = -1$ and $e^{i\pi M} = 1$ (for odd and even eigenvalues of the operator M). The block $\mathcal{H}_{-1}$ has the standard form [19]:

$$\mathcal{H}_{-1} = b\sum_{n=1}^{N}(\Psi_n\Psi_{n+1} + \Psi_{n+1}^+\Psi_n^+),$$
$$\Psi_{N+1} = \Psi_1, \quad (19)$$

needed for its diagonalization. At the same time, the block $\mathcal{H}_1$ in accordance with (18) does not have the standard form [19] since the interaction of fermions 1 and N has the opposite sign relative to the interaction of the remaining fermions. For an odd number of spins N we can reduce $\mathcal{H}_1$ to the standard form. Transforming to the new fermion operators $\bar{\Psi}_n$:

$$\Psi_n = (-1)^{n-1}i\bar{\Psi}_n, \quad n = 1, 2, \ldots, N, \quad (20)$$

we obtain from (18) for odd N

$$\mathcal{H}_1 = b\sum_{n=1}^{N}(\bar{\Psi}_n\bar{\Psi}_{n+1} + \bar{\Psi}_{n+1}^+\bar{\Psi}_n^+). \quad (21)$$

Multiple-quantum NMR in systems with even N requires a different approach and will be analyzed below.

3.1. Multiple-Quantum Nuclear Magnetic Resonance in Rings with an Odd Number of Spins

For odd N the diagonalization of $\mathcal{H}$ is performed in two stages which we shall illustrate using the example

[1] This division of the Hamiltonian was used in [28] to derive analytic formulas for multiple-quantum coherences for $N = 2$ and 3.

of the diagonalization of $\mathcal{H}_{-1}$. Performing the Fourier transformation

$$\Psi_n = \frac{1}{\sqrt{N}}\sum_k a_k e^{ikn}, \quad \Psi_n^+ = \frac{1}{\sqrt{N}}\sum_k a_k^+ e^{-ikn}, \quad (22)$$

we rewrite $\mathcal{H}_{-1}$ using the fermion creation and annihilation operators a_k^+ and a_k in the form

$$\mathcal{H}_{-1} = b\sum_n (a_k a_{-k} + a_k^+ a_{-k}^+) e^{-ik}. \quad (23)$$

At the second stage we need to perform the Bogolyubov transformation [29]:

$$\begin{aligned} a_k &= \frac{\sin k}{\sqrt{2}|\sin k|} D_k + \frac{i}{\sqrt{2}} D_{-k}^+, \\ a_{-k} &= -\frac{\sin k}{\sqrt{2}|\sin k|} D_{-k} + \frac{i}{\sqrt{2}} D_k^+. \end{aligned} \quad (24)$$

The diagonal representation of $\mathcal{H}_{-1}$ in terms of the fermion operators D_k and D_k^+ is written as follows:

$$\begin{aligned} \mathcal{H}_{-1} &= -2b\sum_k |\sin k|\left(D_k^+ D_k - \frac{1}{2}\right), \\ k &= \frac{2\pi}{N} l, \quad l = 0, 1, \ldots, N-1. \end{aligned} \quad (25)$$

The Hamiltonian block $\mathcal{H}_{+1}$ can be diagonalized in exactly the same way. For $e^{i\pi M} = -1$ the projection operator of the total spin angular momentum I_z^{-1} on the z axis in the fermion representation has the form

$$I_z^{-1} = \frac{N}{2} - \sum_k a_k^+ a_k. \quad (26)$$

Equation (26) determines the initial condition for the equation for the density matrix $\sigma_{-1}(t)$ ($\hbar = 1$)

$$i\frac{d\sigma_{-1}(t)}{dt} = [\mathcal{H}_{-1}, \sigma_{-1}(t)], \quad (27)$$

whose solution may be expressed as

$$\sigma_{-1}(t) = \sigma_{-1}^0(t) + \sigma_{-1}^2(t) + \sigma_{-1}^{-2}(t), \quad (28)$$

where

$$\sigma_{-1}^0(t) = \frac{1}{2}\sum_k \cos[4bt\sin k](1 - 2a_k^+ a_k), \quad (29)$$

$$\sigma_{-1}^2(t) = -\frac{1}{2}\sum_k \sin[4bt\sin k] a_k a_{-k}, \quad (30)$$

$$\sigma_{-1}^{-2}(t) = \frac{1}{2}\sum_k \sin[4bt\sin k] a_k^+ a_{-k}^+. \quad (31)$$

The structure of Eqs. (29)–(31) suggests that $\sigma_{-1}^0(t)$ describes a zeroth-order multiple-quantum coherence and $\sigma_{-1}^2(t)$ and $\sigma_{-1}^{-2}(t)$ describe multiple-quantum coherences of the second and "minus" second orders. The intensities of these coherences are determined by Eq. (11) and are given by

$$\begin{aligned} J_0^{-1}(\tau) &= 2^{N-3}\sum_k \cos^2[4b\tau\sin k], \\ J_{\pm 2}^{-1}(\tau) &= 2^{N-4}\sum_k \sin^2[4b\tau\sin k]. \end{aligned} \quad (32)$$

Similar formulas are obtained for the block $\mathcal{H}_1$ of the Hamiltonian $\mathcal{H}$ as a result of which the intensities of the multiple-quantum coherences (32) are doubled. The sum of the intensities of the multiple-quantum coherences does not depend on time [1,30] which follows directly from Eqs. (6) and (11). Normalizing this sum to unity, we finally obtain

$$\begin{aligned} J_0(\tau) &= \frac{1}{N}\sum_k \cos^2[4b\tau\sin k], \\ J_{\pm 2}(\tau) &= \frac{1}{2N}\sum_k \sin^2[4b\tau\sin k]. \end{aligned} \quad (33)$$

As in the thermodynamic limit $N \longrightarrow \infty$ in linear chains [16, 17], the profile of the multiple-quantum coherences in the rings only consists of lines corresponding to multiple-quantum coherences of the zeroth and plus/minus second orders. The intensities of the multiple-quantum coherences averaged over the irradiation time T are determined using the formulas

$$\begin{aligned} \bar{J}_0 &= \lim_{T\to\infty}\frac{1}{T}\int_0^T J_0(\tau)d\tau = \frac{1}{2} + \frac{1}{2N}, \\ \bar{J}_{\pm 2} &= \lim_{T\to\infty}\frac{1}{T}\int_0^T J_{\pm 2}(\tau)d\tau = \frac{1}{4} - \frac{1}{4N}. \end{aligned} \quad (34)$$

3.2. Multiple-quantum Nuclear Magnetic Resonance in Rings with an Even Number of Spins

In order to analyze multiple-quantum coherences in rings with an even number of nuclear spins N we again consider the Hamiltonian $\mathcal{H}$ in the spin space (12). Let us assume that 1, 2, …, N is the numbers of spins in the ring. We perform a unitary transformation U of the Hamiltonian $\mathcal{H}$ which is a composition of π pulses

which flip even-numbered spins through 180° about the *x* axis:

$$U = \exp(-i\pi I_2^x)\exp(-i\pi I_4^x)\ldots\exp(-i\pi I_N^x), \quad (35)$$

where I_k^x is the projection operator of the spin angular momentum of nucleus *k* on the *x* axis. As a result, the transformed Hamiltonian $\overline{\mathcal{H}}$ has the following structure:

$$\overline{\mathcal{H}} = U\mathcal{H}U^+ = b\sum_{j=1}^{N}(I_j^+I_{j+1}^- + I_j^-I_{j+1}^+), \quad (36)$$

and the initial density matrix has the form

$$\bar{\sigma}(0) = UI_zU^+ = \sum_{j=1}^{N}(-1)^{j-1}I_j^z. \quad (37)$$

Equations (36) and (37) show that the problem of multiple-quantum dynamics is reduced to the dynamics of spin systems with an *XY* Hamiltonian solved in [21]. Using [21], we can obtain a solution for the density matrix $\bar{\sigma}(t)$ of a system with the Hamiltonian (36) and the initial condition (37). The density matrix $\bar{\sigma}(t)$ consists of blocks $\bar{\sigma}_\alpha(t)$ corresponding to even and odd eigenvalues of the operator *M* [21]:

$$\begin{aligned}\bar{\sigma}_\alpha(t) = -\frac{2^{2N}}{N^2}\sum_{k,k'}\exp[-i(\epsilon_k^\alpha - \epsilon_{k'}^\alpha)\tau]\\ \times\sum_{j=1}^{N}(-1)^{j-1}\exp[-i(k-k')j]\\ \times\sum_{n,n'}\frac{\exp[i(kn-k'n')]}{2^{n+n'}}\\ \times I_{n+1}^zI_{n+2}^z\ldots I_N^zI_n^-I_{n'+1}^zI_{n'+2}^z\ldots I_N^zI_{n'}^+,\end{aligned} \quad (38)$$

where the single-fermion energies (in frequency units) ϵ_k^α for odd ($\alpha = o$) and even ($\alpha = e$) eigenvalues *M* are respectively given by [21]:

$$\begin{aligned}\epsilon_k^o = 2b\cos k, \quad \epsilon_k^e = 2b\cos\left(k+\frac{\pi}{N}\right),\\ k = \frac{2\pi m}{N}, \quad m = 1, 2, \ldots, N.\end{aligned} \quad (39)$$

It is easy to check that the sum over *j* in Eq. (38) is nonzero (and equal to *N*) only when $k - k' = \pm\pi$. Also bearing in mind that $\epsilon_k^\alpha = -\epsilon_{k\pm\pi}^\alpha$ ($\alpha = o, e$) we write the density matrix $\bar{\sigma}_\alpha(\tau)$ in the form

$$\begin{aligned}\bar{\sigma}_\alpha(\tau) = \frac{2^{2N}}{N}\sum_k e^{-2i\epsilon_k^\alpha t}\sum_{n,n'}(-1)^{n'}\frac{\exp[ik(n-n')]}{2^{n+n'}}\\ \times I_{n+1}^zI_{n+2}^z\ldots I_N^zI_n^-I_{n'+1}^zI_{n'+2}^z\ldots I_N^zI_{n'}^+.\end{aligned} \quad (40)$$

Performing a transformation inverse of (37),

$$\sigma_\alpha(\tau) = U^+\bar{\sigma}_\alpha(\tau)U, \quad (41)$$

we express the density matrix $\sigma_\alpha(\tau)$ describing the multiple-quantum dynamics of a ring having an even number of nuclear spins, in the form

$$\sigma_\alpha(\tau) = \sigma_\alpha^0(\tau) + \sigma_\alpha^2(\tau) + \sigma_\alpha^{-2}(\tau), \quad (42)$$

where

$$\begin{aligned}\sigma_\alpha^0(\tau) = \frac{2^{2N}}{N}\sum_k\exp(-2i\epsilon_k^\alpha\tau)\\ \times\left(\sum_{n=1,3,\ldots}\sum_{n'=1,3,\ldots}(-1)^{(n+n')/2}\frac{\exp[ik(n-n')]}{2^{n+n'}}\right.\\ \times I_{n+1}^zI_{n+2}^z\ldots I_N^zI_n^-I_{n'+1}^zI_{n'+2}^z\ldots I_N^zI_{n'}^+\\ +\sum_{n=2,4,\ldots}\sum_{n'=2,4,\ldots}(-1)^{(n+n')/2}\frac{\exp[ik(n-n')]}{2^{n+n'}}\\ \left.\times I_{n+1}^zI_{n+2}^z\ldots I_N^zI_n^+I_{n'+1}^zI_{n'+2}^z\ldots I_N^zI_{n'}^-\right),\end{aligned} \quad (43)$$

$$\begin{aligned}\sigma_\alpha^2(\tau) = \frac{2^{2N}}{N}\sum_k\exp(-2i\epsilon_k^\alpha\tau)\\ \times\sum_{n=2,4,\ldots}\sum_{n'=1,3,\ldots}(-1)^{(n+n'+1)/2}\frac{\exp[ik(n-n')]}{2^{n+n'}}\\ \times I_{n+1}^zI_{n+2}^z\ldots I_N^zI_n^+I_{n'+1}^zI_{n'+2}^z\ldots I_N^zI_{n'}^+,\end{aligned} \quad (44)$$

$$\begin{aligned}\sigma_\alpha^{-2}(\tau) = -\frac{2^{2N}}{N}\sum_k\exp(-2i\epsilon_k^\alpha\tau)\\ \times\sum_{n=1,3,\ldots}\sum_{n'=2,4,\ldots}(-1)^{(n+n'+1)/2}\frac{\exp[ik(n-n')]}{2^{n+n'}}\\ \times I_{n+1}^zI_{n+2}^z\ldots I_N^zI_n^-I_{n'+1}^zI_{n'+2}^z\ldots I_N^zI_{n'}^-.\end{aligned} \quad (45)$$

Here $\sigma_\alpha^0(\tau)$ describes transitions between Zeeman levels of the spin system in an external magnetic field which are responsible for zeroth-order multiple-quan-

tum coherence and $\sigma_\alpha^2(\tau)$ and $\sigma_\alpha^{-2}(\tau)$ describe multiple-quantum coherences of plus/minus second order. The intensities of the multiple-quantum coherences are determined by Eq. (11) as before and in normalized form are given by

$$J_0(\tau) = \frac{1}{2N}\left(\sum_k \cos^2[4b\tau\cos k] + \sum_k \cos^2\left[4b\tau\cos\left(k+\frac{\pi}{N}\right)\right]\right), \quad (46)$$

$$J_{\pm 2}(\tau) = \frac{1}{4N}\left(\sum_k \sin^2[4b\tau\cos k] + \sum_k \sin^2\left[4b\tau\cos\left(k+\frac{\pi}{N}\right)\right]\right). \quad (47)$$

The profile of the multiple-quantum coherences again consisted of only three lines corresponding to multiple-quantum coherences of zeroth and plus/minus second orders. It is interesting to note that in all arguments in the Eqs. (46) and (47) the cosine may be replaced by a sine without changing the intensities of the multiple-quantum coherences. Despite the difference between Eqs. (33) and (46), (47) for the intensities of multiple-quantum coherences in rings with even and odd numbers of nuclear spins, the intensities of the multiple-quantum coherences averaged over the excitation time are the same for both cases and are determined by Eqs. (34).

4. MULTIPLE-QUANTUM NUCLEAR MAGNETIC RESONANCE IN LINEAR CHAINS OF NUCLEAR SPINS

The Hamiltonian of the system of nuclear spins in linear chains only differs from the Hamiltonian of ring systems in that there is no interaction of nuclear spins 1 and N at the ends of the chains:

$$\mathcal{H} = b\sum_{i=1}^{N-1}(I_i^+ I_{i+1}^+ + I_i^- I_{i+1}^-). \quad (48)$$

We perform the transformation U of the Hamiltonian (48)

$$U = \exp(-i\pi I_2^x)\exp(-i\pi I_4^x)\ldots, \quad (49)$$

applying selective π pulses which flip the spins in even positions through 180° about the x-axis of the rotating reference frame. As a result, we obtain

$$\bar{\mathcal{H}} = U\mathcal{H}U^+ = b\sum_{i=1}^{N-1}(I_i^+ I_{i+1}^- + I_i^- I_{i+1}^+). \quad (50)$$

The initial density matrix $\sigma(0) = I_z$ is then transformed

$$\bar{\sigma}(0) = UI_zI^+ = \sum_{j=1}^{N}(-1)^{j-1}I_j^z. \quad (51)$$

A solution of the equation for the density matrix $\sigma(t)$ of a spin system with the Hamiltonian (50) subject to the initial condition (51) may be obtained by generalizing the analytic solution obtained in [22]. Thus, we find

$$\bar{\sigma}(t) = \frac{2}{N+1}\sum_{k,k'}\exp[-i(\epsilon_k-\epsilon_{k'})t] \times\left(\sum_{j=1}^{N}(-1)^{j-1}\sin(kj)\sin(k'j)\right)\beta_k^+\beta_{k'} - \frac{1}{4}[1-(-1)^N], \quad (52)$$

$$k = \frac{\pi n}{N+1}, \quad n = 1, 2, \ldots, N,$$

where the fermion operators β_k are related to the spin operators by

$$\beta_k = \sqrt{\frac{2}{N+1}}\sum_{l=1}^{N}(-2)^{l-1}\sin(kl)I_1^zI_2^z\ldots I_{l-1}^zI_l^-, \quad (53)$$

and the single-fermion energies (in frequency units) are determined using the formula

$$\epsilon_k = 2b\cos k, \quad (54)$$

It is convenient to replace the wave vector k with $k = \pi - \tilde{k}$. We then have

$$\sum_{j=1}^{N}(-1)^{j-1}\sin(kj)\sin(k'j) = \sum_{j=1}^{N}(-1)^{j-1}\sin[(\pi-\tilde{k})j]\sin(k'j) = \sum_{j=1}^{N}\sin(\tilde{k}j)\sin(k'j) = \frac{N+1}{2}\delta_{\tilde{k}k'}. \quad (55)$$

Taking into account Eq. (55), we can rewrite Eq. (52) for the density matrix $\bar{\sigma}(t)$ as follows:

$$\bar{\sigma}(t) = -\frac{2}{N+1}\sum_k\exp(2i\epsilon_k t) \times\sum_{l,l'=1}^{N}(-1)^{l'}2^{l+l'-2}\sin(kl)\sin(kl') \times I_1^zI_2^z\ldots I_{l-1}^zI_l^+I_1^zI_2^z\ldots I_{l'-1}^zI_{l'}^- - \frac{1}{4}[1-(-1)^N], \quad (56)$$

$$k = \frac{\pi n}{N+1} \quad n = 1, 2, \ldots, N.$$

After performing a transformation the inverse of Eq. (49), we obtain the density matrix $\sigma(t)$ describing the multiple-quantum dynamics of a one-dimensional linear chain of nuclear spins in the form

$$\sigma(t) = \sigma_0(t) + \sigma_2(t) + \sigma_{-2}(t), \tag{57}$$

where

$$\sigma_0(t) = -\frac{2}{N+1}\sum_k \exp(2i\epsilon_k t)$$

$$\times\Bigg(\sum_{l=1,3,\ldots}\sum_{l'=1,3,\ldots}(-1)^{(l+l')/2}2^{l+l'-2}\sin(kl)\sin(kl')$$

$$\times I_1^z I_2^z \ldots I_{l-1}^z I_l^+ I_1^z I_2^z \ldots I_{l'-1}^z I_{l'}^- \tag{58}$$

$$+\sum_{l=2,4,\ldots}\sum_{l'=2,4,\ldots}(-1)^{(l+l')/2}2^{l+l'-2}\sin(kl)\sin(kl')$$

$$\times I_1^z I_2^z \ldots I_{l-1}^z I_l^- I_1^z I_2^z \ldots I_{l'-1}^z I_{l'}^+\Bigg),$$

$$\sigma_2(t) = -\frac{2}{N+1}\sum_k \exp(2i\epsilon_k t)$$

$$\times\sum_{l=1,3,\ldots}\sum_{l'=2,4,\ldots}(-1)^{(l+l'+1)/2}2^{l+l'-2}\sin(kl) \tag{59}$$

$$\times\sin(kl')I_1^z I_2^z \ldots I_{l-1}^z I_l^+ I_1^z I_2^z \ldots I_{l'-1}^z I_{l'}^+,$$

$$\sigma_{-2}(t) = -\frac{2}{N+1}\sum_k \exp(2i\epsilon_k t)$$

$$\times\sum_{l=2,4,\ldots}\sum_{l'=1,3,\ldots}(-1)^{(l+l'+1)/2}2^{l+l'-2}\sin(kl) \tag{60}$$

$$\times\sin(kl')I_1^z I_2^z \ldots I_{l-1}^z I_l^- I_1^z I_2^z \ldots I_{l'-1}^z I_{l'}^-.$$

According to Eq. (11), the normalized intensities of the zeroth- and plus/minus second-order multiple-quantum coherences, which only occur in a linear spin chain when allowance is made for nearest-neighbor interactions, are given by

$$J_0(\tau) = \frac{1}{N}\sum_k \cos^2(4b\tau\cos k),$$
$$J_{\pm 2}(\tau) = \frac{1}{2N}\sum_k \sin^2(4b\tau\cos k). \tag{61}$$

In the thermodynamic limit when $N \longrightarrow \infty$, by replacing summation with integration in Eqs. (61), we arrive at the formulas

$$J_0(\tau) = \frac{1}{2} + \frac{1}{2}\tilde{\mathcal{J}}_0(8b\tau),$$
$$J_{\pm 2}(\tau) = \frac{1}{4} - \frac{1}{4}\tilde{\mathcal{J}}_0(8b\tau), \tag{62}$$

where $\tilde{\mathcal{J}}_0(x)$ is the zeroth-order Bessel function. The formulas (62) are the same as those obtained in [16, 17] by a method valid only for $N \gg 1$. The intensities of the multiple-quantum coherences averaged over the excitation time T for a linear chain having an odd number of nuclear spins are given by

$$\bar{J}_0 = \lim_{T\to\infty}\frac{1}{T}\int_0^T J_0(\tau)d\tau = \frac{1}{2} + \frac{1}{2N},$$
$$\bar{J}_{\pm 2} = \lim_{T\to\infty}\frac{1}{T}\int_0^T J_{\pm 2}(\tau)d\tau = \frac{1}{4} - \frac{1}{4N}. \tag{63}$$

Similar formulas for a linear chain having an even number of nuclear spins have the form

$$\bar{J}_0 = \frac{1}{2}, \quad \bar{J}_{\pm 2} = \frac{1}{4} \tag{64}$$

and agree with the expressions obtained for $N \longrightarrow \infty$ in [16, 17]. Unlike rings in linear chains the average intensities of the multiple-quantum coherences differ for even and odd N.

5. RELATIONSHIP BETWEEN MULTIPLE-QUANTUM DYNAMICS OF SPIN SYSTEMS AND DYNAMICS DESCRIBED BY AN *XY* HAMILTONIAN

A study of the multiple-quantum dynamics of spin systems in rings having an even number of nuclear spins N and in finite linear chains carried out in Sections 3.2 and 4 demonstrates a clear relationship between multiple-quantum dynamics and the dynamics of systems with an *XY* Hamiltonian [31]. Although in multiple-quantum NMR all possible transitions are excited between the Zeeman levels of a multiple-spin system in an external magnetic field [2], in the approximation of nearest-neighbor interaction multiple-quantum dynamics is closely related to the dynamics determined by the *XY* Hamiltonian when all the transitions are merely attributable to flip-flop processes [20]. Let us assume that the initial density matrix of a linear chain consisting of N spins is

$$\sigma(0) = \sum_k I_k^x = I_x.$$

Then, as a result of the time evolution described by the two-spin, two-quantum Hamiltonian [see (48)], at the

end of the preparation period the density matrix $\sigma(t)$ is given by

$$\sigma(t) = \exp(-i\mathcal{H}t)I_x\exp(i\mathcal{H}t), \quad (65)$$

and all multiple-quantum coherences of odd orders evolve in the system [24]. At the same time, by performing the transformation (49) we obtain

$$\bar{\sigma}(t) = U\sigma(t)U^+ = \exp(-i\overline{\mathcal{H}}t)I_x\exp(-i\overline{\mathcal{H}}t), \quad (66)$$

where the Hamiltonian $\overline{\mathcal{H}}$ is given by Eq. (50). In this case, at the end of the preparation period only multiple-quantum coherences of plus/minus first order appear. The exact solutions obtained above can explain this factor. If the term in the expression for the density matrix (56) contains raising I_l^+ and lowering $I_{l'}^-$ operators with odd or even l and l', it is responsible for zeroth-order multiple-quantum coherence. If the parity of l and l' differs, the term obtained after transforming (49) becomes responsible for multiple-quantum coherence of plus/minus second order.

The approximation of nearest-neighbor interactions does not necessarily lead to an exactly solvable model. If the spacings between the spins and thus the dipole–dipole interaction constants differ, the methods proposed above do not lead to exactly solvable problems of multiple-quantum dynamics. In this case, in principle multiple-quantum coherences of all even orders may occur.

We propose a simple experiment to demonstrate the relationship between multiple-quantum dynamics and the dynamics determined by the *XY* Hamiltonian. For this we need to study the multiple-quantum dynamics of benzene molecules dissolved in a liquid-crystal matrix [32], by exposing the system to a sequence of pulses (1). The initial condition $\sigma(0) = I_1^z$ is created by cross polarization [33] between the ^{13}C nucleus and a proton associated with it in the benzene. A similar process can be used to measure the polarization of the first proton at time τ which is not too large so that interactions with next-nearest neighbors can be neglected. In this multiple-quantum experiment the polarization of the first proton $\langle I_1^z\rangle(t)$ is determined by

$$\begin{aligned}\langle I_1^z\rangle(t) &= \mathrm{Tr}[\sigma(t)I_1^z] \\ &= \mathrm{Tr}[\exp(-i\mathcal{H}t)I_1^z\exp(i\mathcal{H}t)I_1^z].\end{aligned} \quad (67)$$

Performing the transformation (35) in (67) we find

$$\langle I_1^z\rangle(t) = \mathrm{Tr}[\exp(-i\overline{\mathcal{H}}t)I_1^z\exp(i\overline{\mathcal{H}}t)I_1^z], \quad (68)$$

where $\overline{\mathcal{H}}$ is an *XY* Hamiltonian. The experimentally observed signal $\langle I_1^z\rangle(t)$ can be compared with the similar theoretical result for the dynamics in a nuclear spin ring with an *XY* Hamiltonian, obtained in [21]. This relationship between multiple-quantum dynamics and the dynamics of a system with an *XY* Hamiltonian allows us to create an effective *XY* Hamiltonian in a solid. A completely different method was used to solve the analogous problem for liquids [34].

6. NUMERICAL ANALYSIS OF MULTIPLE-QUANTUM NUCLEAR MAGNETIC RESONANCE IN RINGS AND LINEAR CHAINS

In order to make numerical calculations for the intensities of multiple-quantum coherences it is convenient to use the formula

$$J_n(\tau) = \sum_{i,j}{}'|\rho_{ij}(\tau)|^2 \quad (69)$$

derived directly from Eq. (11). The prime after the summation sign in Eq. (69) implies that summation is merely limited to states i, j for which the projections M_i, M_j of the total spin momentum on the z axis satisfy

$$M_i - M_j = n. \quad (70)$$

In order to make calculations using Eq. (69) we need to find the density matrix

$$\rho(\tau) = e^{-i\mathcal{H}\tau}I_z e^{i\mathcal{H}\tau} = Ce^{-i\Lambda\tau}(C^TI_zC)e^{i\Lambda\tau}C^T, \quad (71)$$

where the orthogonal matrix C ($CC^T = E$) diagonalizes the Hamiltonian of the system:

$$\Lambda = C^T\mathcal{H}C. \quad (72)$$

In order to diagonalize the Hamiltonian $\mathcal{H}$ we used the integral of motion $\exp(-i\pi I_z)$ (15) which divides the matrix $\mathcal{H}$ into two blocks of dimensions $2^{N-1}\times 2^{N-1}$. More effective factorization can be achieved in linear chains in the approximation of nearest-neighbor interactions when, however, the dipole–dipole interaction constants of the pairs differ. For this we need to perform the transformation (49) and factorize the matrix of the Hamiltonian $\overline{\mathcal{H}}$ (50) into blocks of dimensions $C_N^k \times C_N^k$ ($k = 0, 1, 2, \ldots, N$) in terms of eigenvalues of the operator I_z which commutes with $\overline{\mathcal{H}}$. For the following calculations we need to perform a transformation the inverse of Eq. (49). The accuracy of the calculations was monitored by checking the normalization condition $\sum_n J_n(\tau) = 1$ and comparing the numerical solutions in the approximation of nearest-neighbor interactions with the analytic solutions presented in Sections 3 and 4.

Figure 1 gives time dependences of the multiple-quantum coherences of a linear chain of seven nuclear spins coupled by dipole–dipole interaction. For $\tau = 0$ only zeroth-order multiple-quantum transitions occurred in the system. In the initial period of evolution multiple-

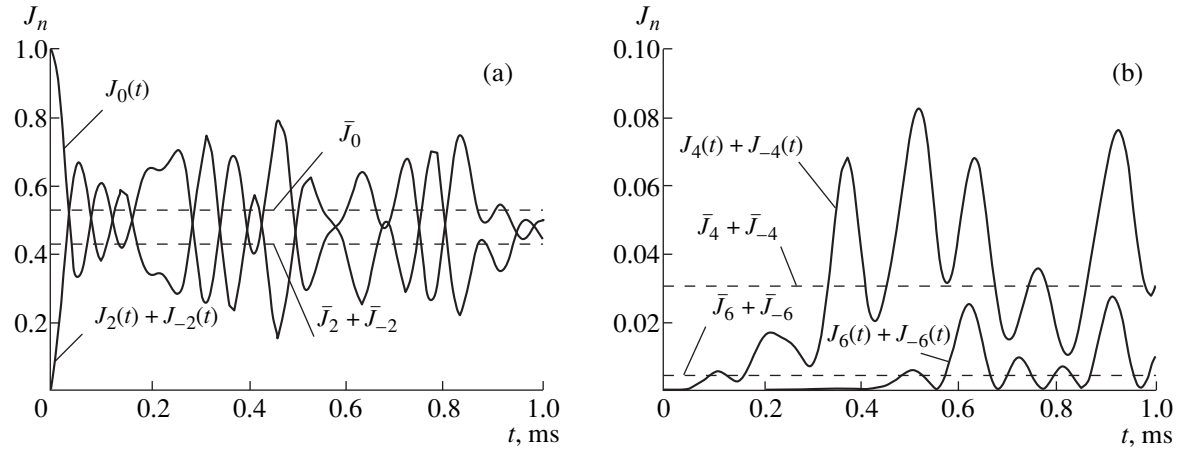

Fig. 1. Time dependences of the intensities of multiple-quantum coherences in a linear chain of seven spins coupled by dipole–dipole interaction, $b = -D_{ii+1}/2 = 2\pi \times 1475\ \text{s}^{-1}$: (a) intensities of multiple-quantum coherences of zeroth $J_0(t)$ and second $J_2(t) + J_{-2}(t)$ orders; (b) intensities of multiple-quantum coherences of fourth $J_4(t) + J_{-4}(t)$ and sixth $J_6(t) + J_{-6}(t)$ orders. The horizontal straight lines give the time-averaged multiple-quantum coherence intensities $\bar{J}_n = T^{-1}\int_0^T J_n(\tau)d\tau$, $T = 1$ ms.

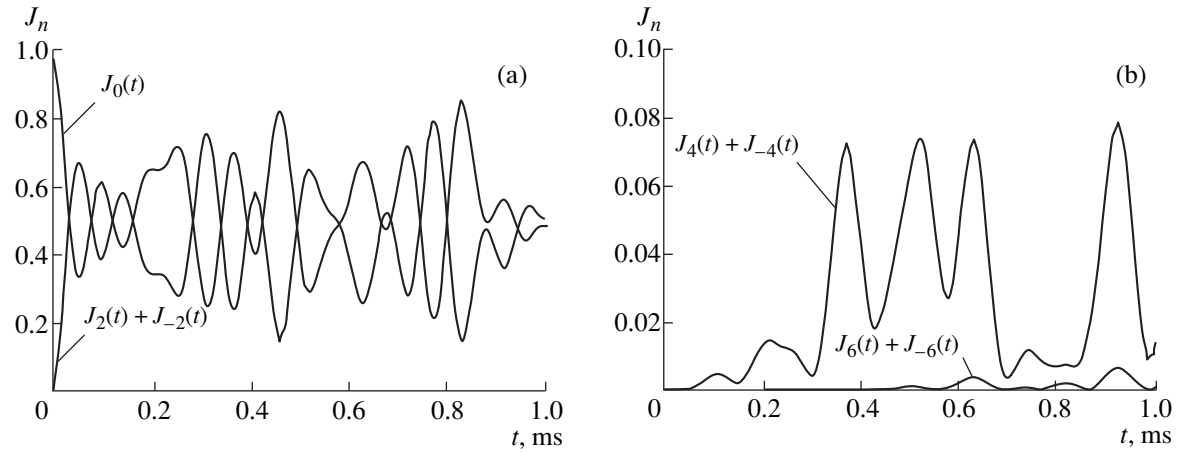

Fig. 2. Time dependences of the intensities of multiple-quantum coherences in a linear chain of seven spins coupled by dipole–dipole interaction of nearest and next-nearest neighbors, $b = -D_{ii+1}/2 = 2\pi \times 1475\ \text{s}^{-1}$: (a) intensities of multiple-quantum coherences of zeroth $J_0(t)$ and second $J_2(t) + J_{-2}(t)$ orders; (b) intensities of multiple-quantum coherences of fourth $J_4(t) + J_{-4}(t)$ and sixth $J_6(t) + J_{-6}(t)$ orders.

quantum coherences of plus/minus second order appear and the multiple-quantum dynamics at this stage can be described as exchange between multiple-quantum coherences of zeroth and plus/minus second orders. For $0 < \tau < 0.2$ ms the multiple-quantum dynamics of a seven-spin system is fairly accurately described by the analytic solution (61) obtained in the nearest-neighbor approximation. For $\tau > 0.1$ ms fourth-order multiple-quantum transitions occur and for $\tau > 0.4$ ms sixth-order transitions. Figure 2 shows the multiple-quantum dynamics of a seven-spin system for the case when only nearest- and next-nearest-neighbor interactions are taken into account. A comparison of Figs. 1a and 2a shows that allowance for dipole–dipole interaction of nearest and next-nearest neighbors is quite sufficient for quantitative agreement between the multiple-quantum dynamics of zeroth- and second-order coherences and the similar dynamics allowing for all dipole–dipole interactions in the system, at least as far as time $\tau \approx 1$ ms. Small differences in the dynamics of fourth- and sixth-order multiple-quantum coherences (Figs. 1b and 2b) are insufficient to obtain structural information outside the limits of two "coordination spheres" in quasi-one-dimensional systems. In two-dimensional and three-dimensional systems the small contribution made by various remote nuclear spins to the intensity of the multiple-quantum coherences is compensated by the large number of these spins. Thus, multiple-quantum NMR

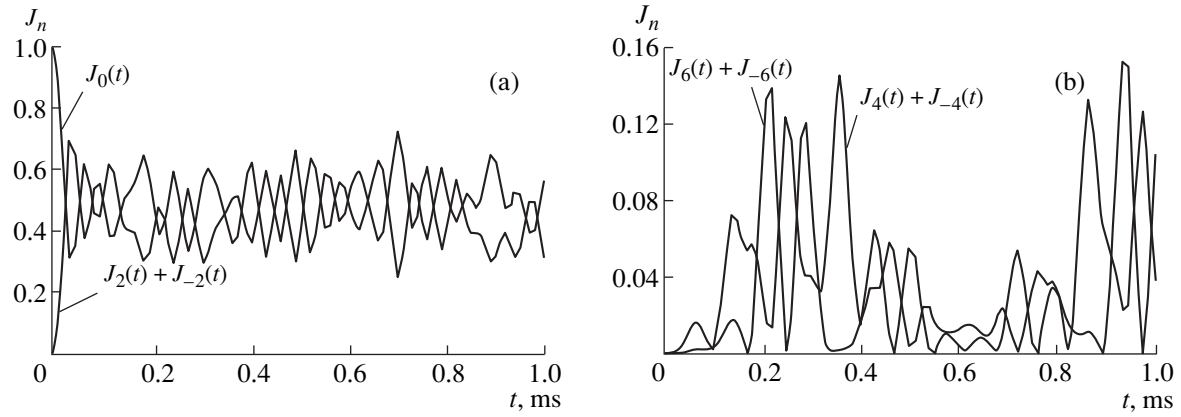

Fig. 3. Time dependences of the intensities of multiple-quantum coherences in a ring of six spins coupled by dipole–dipole interaction, $b = -D_{ii+1}/2 = 2\pi \times 2253.6\ \text{s}^{-1}$: (a) intensities of multiple-quantum coherences of zeroth $J_0(t)$ and second $J_2(t) + J_{-2}(t)$ orders; (b) intensities of multiple-quantum coherence of fourth $J_4(t) + J_{-4}(t)$ and sixth $J_6(t) + J_{-6}(t)$ orders.

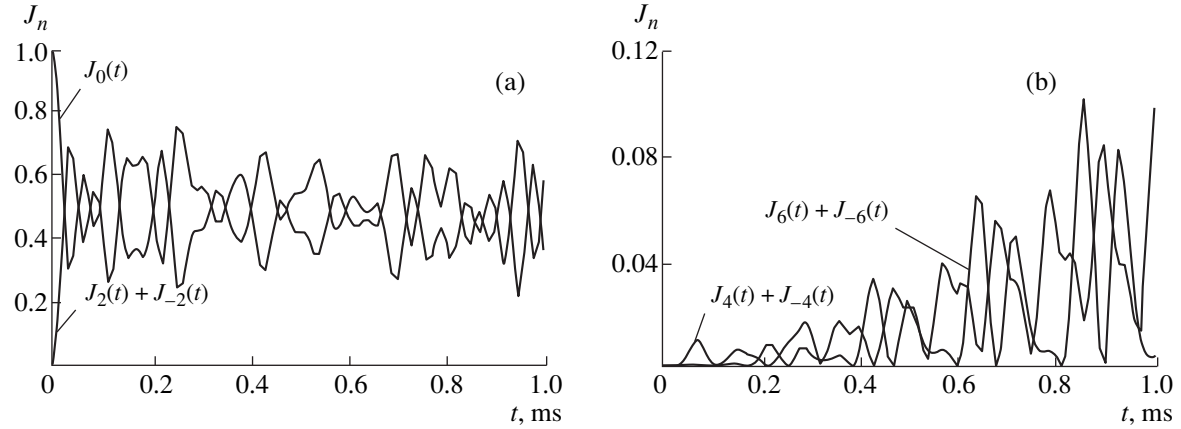

Fig. 4. Time dependences of the intensities of multiple-quantum coherences in a ring of six spins coupled by dipole–dipole interaction of nearest and next-nearest neighbors, $b = -D_{ii+1}/2 = 2\pi \times 2253.6\ \text{s}^{-1}$: (a) intensities of multiple-quantum coherences of zeroth $J_0(t)$ and second $J_2(t) + J_{-2}(t)$ orders; (b) intensities of multiple-quantum coherences of fourth $J_4(t) + J_{-4}(t)$ and sixth $J_6(t) + J_{-6}(t)$ orders.

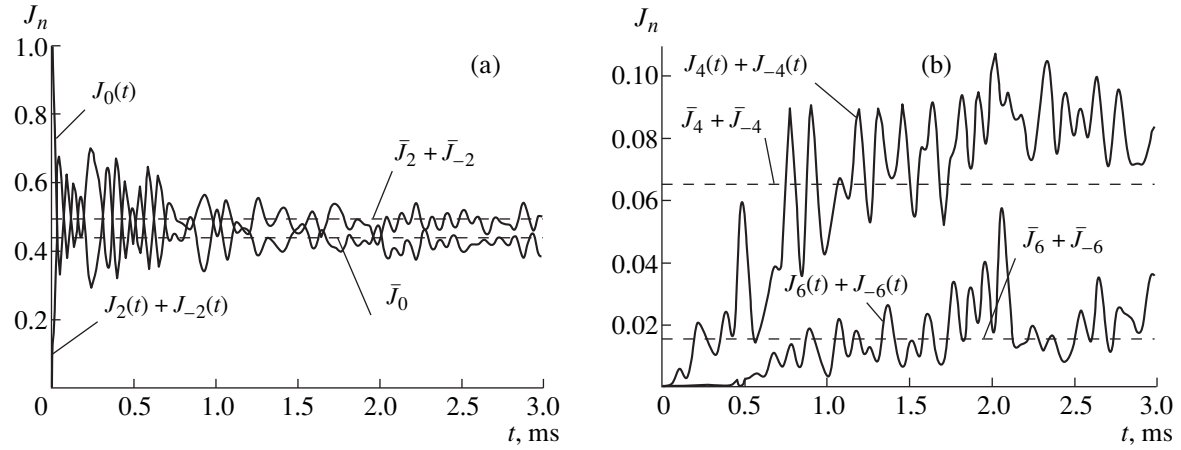

Fig. 5. Time dependences of the intensities of multiple-quantum coherences in a linear chain of eight spins coupled by dipole–dipole interaction, $b = -D_{ii+1}/2 = 2\pi \times 1475\ \text{s}^{-1}$: (a) intensities of multiple-quantum coherences of zeroth $J_0(t)$ and second $J_2(t) + J_{-2}(t)$ orders; (b) intensities of multiple-quantum coherence of fourth $J_4(t) + J_{-4}(t)$ and sixth $J_6(t) + J_{-6}(t)$ orders. The horizontal straight lines give the multiple-quantum coherence intensities averaged over irradiation time $\bar{J}_n = T^{-1}\int_0^T J_n(\tau)d\tau$, $T = 3$ ms.

Table

Number of multiple-quantum coherence	0	2	4	6
Statistical theory [2]	0.393	0.244	0.056	0.004
Results of numerical calculations	0.436	0.243	0.032	0.007

may be potentially extremely useful for obtaining structural information in solids.

Unlike one-dimensional linear chains in ring systems information on the contributions of remote spins to the intensity of multiple-quantum coherences can be obtained from the profiles of the multiple-quantum coherences. Figure 3 shows the evolution of multiple-quantum coherences in a six-spin ring. Similar multiple-quantum dynamics allowing for dipole–dipole interaction of nearest and next-nearest neighbors is shown in Fig. 4. Here there are clear differences even between the multiple-quantum dynamics of the zeroth- and second-order coherences plotted in Figs. 3a and 4a. The reason for this is whereas in a linear chain the dipole–dipole interaction of the first and fourth spins is 3.375 times weaker than that of the first and third spins, in a six-spin ring the ratio of the similar dipole–dipole interactions is only 1.5. Hence the contribution of remote spins in a ring is more significant than that in chains. At the same time, in rings it is possible in principle to extract fine structural information associated with the interactions of remote spins. Figure 5 gives dependences of the multiple-quantum coherence intensities in a linear eight-spin system. Qualitatively the multiple-quantum dynamics of an eight-spin system does not differ from the similar dynamics of a seven-spin system. In the table the intensities of the multiple-quantum coherences averaged over the excitation time are compared with the predictions of the statistical theory [2] for an eight-spin linear chain.

It can be seen from the table that the predictions of the statistical theory [2] agree qualitatively with the result of a numerical analysis. The intensity of the eighth-order multiple-quantum coherence is negligible and has an insignificant influence on the multiple-quantum dynamics in an eight-spin chain.

7. CONCLUSIONS

We have studied analytic and numerical methods of analyzing the multiple-quantum dynamics of one-dimensional systems of nuclear spins in solids. In the approximation of nearest-neighbor interactions the multiple-quantum dynamics of linear chains and rings with an even number of spins was directly related to the dynamics of systems with an *XY* Hamiltonian. This factor leads to new experimental methods of studying multiple-quantum dynamics and new numerical methods for interpreting the experimental results. In an experimental study of multiple-quantum dynamics in quasi-one-dimensional systems of hydroxyl protons of calcium hydroxyapatite [15] it was found that the observed one-dimensional cluster consists of approximately twelve spins [15]. It is therefore extremely relevant to make a numerical study of multiple-quantum dynamics in linear chains consisting of between twelve and twenty spins. The solution of this problem requires significant development of new approaches to the numerical methods used for calculations in multiple-quantum dynamics.

ACKNOWLEDGMENTS

We thank V.A. Atsarkin, F.S. Dzheparov, M.Yu. Loshkevich, T.N. Khazanovich, participants at the seminar on "Problems of Magnetic Resonance" for useful discussions, and A.N. Ivanova and Yu.D. Kobyzev for assistance in organizing and carrying out the numerical calculations.

This work was supported by the Russian Foundation for Basic Research (project no. 98-03-33151).

REFERENCES

1. G. Cho and J. P. Yesinowski, J. Phys. Chem. **100**, 15716 (1996).
2. J. Baum, M. Munowitz, A. N. Garroway, and A. Pines, J. Chem. Phys. **83**, 2015 (1985).
3. J. Baum and A. Pines, J. Am. Chem. Soc. **108**, 7447 (1986).
4. B. E. Scruggs and K. K. Gleason, Macromolecules **25**, 1864 (1992).
5. D. H. Levy and K. K. Gleason, J. Phys. Chem. **96**, 8125 (1992).
6. M. A. Lavrent'ev and B. V. Shabat, *Methods of the Theory of Functions of a Complex Variable* (Nauka, Moscow, 1973).
7. D. P. Weitekamp, Adv. Magn. Reson. **11**, 111 (1983).
8. S. Lacelle, S. J. Hwang, and B. G. Gerstein, J. Chem. Phys. **99**, 8407 (1993).
9. M. Tomaselli, S. Hediger, D. Suter, and R. R. Ernst, J. Chem. Phys. **105**, 10672 (1996).
10. M. Munowitz, A. Pines, and M. Mehring, J. Chem. Phys. **86**, 3172 (1987).
11. M. Munowitz and M. Mehring, Solid State Commun. **64**, 605 (1987).
12. C. S. Johnson, J. Chem. Phys. **41**, 3277 (1964).
13. M. Munowitz, Mol. Phys. **71**, 959 (1990).
14. B. E. Scruggs and K. K. Gleason, Chem. Phys. **166**, 367 (1992).
15. G. Cho and J. P. Yesinowski, Chem. Phys. Lett. **205**, 1 (1993).
16. E. B. Fel'dman and S. Lacelle, Chem. Phys. Lett. **253**, 27 (1996).
17. E. B. Fel'dman and S. Lacelle, J. Chem. Phys. **107**, 7067 (1997).
18. E. B. Fel'dman and S. Lacelle, J. Chem. Phys. **106**, 6768 (1997).

19. M. Gaudin, *La fonction d'onde de Bethe* (Masson, Paris, 1983; Mir, Moscow, 1984).
20. A. Abragam, *The Principles of Nuclear Magnetism* (Clarendon Press, Oxford, 1961; Inostrannaya Literatura, Moscow, 1963).
21. E. B. Fel'dman and M. G. Rudavets, Chem. Phys. Lett. **311**, 453 (1999).
22. E. B. Fel'dman, R. Bruschweiler, and R. R. Ernst, Chem. Phys. Lett. **294**, 297 (1999).
23. M. Goldman, *Spin Temperature and Nuclear Magnetic Resonance in Solids* (Clarendon Press, Oxford, 1970; Mir, Moscow, 1972).
24. D. Suter, S. B. Liu, J. Baum, and A. Pines, Chem. Phys. **114**, 103 (1987).
25. D. Shykind, J. Baum, S. B. Liu, *et al.*, J. Magn. Reson. **60**, 205 (1984).
26. H. Schneider and H. Schmiedel, Phys. Lett. A **30**, 298 (1969).
27. W.-K. Rhim, A. Pines, and J. S. Waugh, Phys. Rev. Lett. **25**, 218 (1970).
28. A. K. Roy and K. K. Gleason, J. Magn. Reson. A **120**, 139 (1996).
29. L. D. Landau and E. M. Lifshitz, *Statistical Physics* (Nauka, Moscow, 1976; Pergamon, Oxford, 1980), Part 2.
30. D. A. Lathrop, E. S. Handy, and K. K. Gleason, J. Magn. Reson. A **111**, 161 (1994).
31. E. H. Lieb, T. Schultz, and D. C. Mattis, Ann. Phys. **16**, 407 (1961).
32. A. K. Khitrin and B. M. Fung, J. Chem. Phys. **111**, 7480 (1999).
33. S. Zhang, B. H. Meier, and R. R. Ernst, Phys. Rev. Lett. **69**, 2149 (1992).
34. Z. L. Madi, B. Brutscher, T. Schulte-Herbruggen, *et al.*, Chem. Phys. Lett. **268**, 300 (1997).

Translation was provided by AIP

Journal of Experimental and Theoretical Physics, Vol. 91, No. 3, 2000, pp. 610–614.
Translated from Zhurnal Éksperimental'noĭ i Teoreticheskoĭ Fiziki, Vol. 118, No. 3, 2000, pp. 701–706.

SOLIDS
Electronic Properties

Forced Charge Oscillations in a Double Quantum Dot

V. A. Burdov
Nizhni Novgorod State University, Nizhni Novgorod, 603600 Russia
e-mail: burdov@phys.unn.runnet.ru
Received April 28, 2000

Abstract—The two-electron wave function and charge distribution are obtained in a symmetric double quantum dot in a weak variable electric field. It is shown that the action of a variable field under resonance conditions when the perturbation frequency is close to the frequency of the quantum transition leads to the appearance of electron density oscillations between the dots having the characteristic form of beats. However, the Coulomb repulsion between the electrons strongly "quenches" the amplitude of the beats even in a resonant variable field.

1. INTRODUCTION

Various low-dimensional nanostructures, which have recently attracted increasing attention among researchers, are attractive mainly from the point of using their functional possibilities. The question of controllably influencing the electron dynamics in quantum structures using various external actions then arises.

The possibility of achieving so-called controlled tunneling in double quantum-well heterostructures by exposing the electron to a strong variable electric field has been reported [1–5] and the electron distribution in a double quantum dot and electron transport in the presence of a variable electric field were studied in [6]. Charge oscillations in tunnel-coupled quantum wells and dots and terahertz emission accompanying these oscillations were observed experimentally in [7–10].

It should be noted that the theoretical studies cited above [1–6] only considered the single-electron dynamics which may appreciably distort the true pattern, especially for systems of quantum dots where the Coulomb interaction between the electrons is already significant, and generally leads to a Coulomb blockade effect. The aim of the present paper is to study the influence of Coulomb interaction on the electron dynamics in a double quantum dot in the presence of an external variable field.

2. BASIC APPROXIMATIONS

We shall consider a symmetric double quantum dot with two electrons in an external variable electric field whose intensity varies with time as $\mathscr{E}(t) = \mathscr{E}\sin(\omega t)$. To simplify the calculations to some extent we shall assume that the size of each quantum dot is so small that the characteristic Coulomb interaction energy is much lower than the characteristic energy of the quantum transition in an isolated quantum dot:

$$\frac{e^2}{\epsilon R} \ll \frac{\hbar^2\pi^2}{mR^2},$$

where R is the size of the quantum dot, m is the effective mass, and ϵ is the permittivity of the medium. Estimates show that this inequality is satisfied for quantum dots having dimensions not exceeding tens of nanometers. Such quantum dots are obtained, for example, by ion implantation followed by deposition, and are generally close to spherical [11–13].

Tunnel coupling between quantum dots leads to splitting of the electron energy levels in isolated dots by the amount Δ which is small compared with the transition energies in an isolated quantum dot. According to estimates, the ratio of these energies may be lower than or of the order of 10^{-2}, giving a value of approximately 10^{-2} eV or lower for the splitting energy Δ which is of the same order of magnitude as the thermal energy at room temperatures.

We shall also assume that the probability of electron "dumping" to upper levels of the isolated quantum dot is low, for which we impose the constraint that the amplitude of the variable electric field and its frequency satisfy the inequalities

$$e\mathscr{E}R \ll \frac{\hbar^2\pi^2}{mR^2}, \quad \omega \ll \frac{\hbar\pi^2}{mR^2}.$$

Taking into account all the relationships given above between the energy parameters of the problem, we conclude that only two levels obtained as a result of splitting of the ground-state energy level in an isolated dot are "involved" in the electron dynamics. We take the energy origin midway between these levels, and then the level energies will be $\pm\Delta/2$. The single-electron wave functions of the steady states corresponding to these levels are denoted by $\chi_{0,1}(\mathbf{r})$.

The functions $\chi_{0,1}(\mathbf{r})$ are respectively even and odd with respect to the coordinate z [we take the axis of the double quantum point as the z coordinate axis and assume that the function $\chi_0(\mathbf{r})$ is everywhere positive while $\chi_1(\mathbf{r})$ is only positive in the region $z > 0$]. Using $\chi_0(\mathbf{r})$ and $\chi_1(\mathbf{r})$ we can then construct two other orthonormalized functions:

$$\Psi_L(\mathbf{r}) = \frac{\chi_0(\mathbf{r}) - \chi_1(\mathbf{r})}{\sqrt{2}}, \quad \Psi_R(\mathbf{r}) = \frac{\chi_0(\mathbf{r}) + \chi_1(\mathbf{r})}{\sqrt{2}},$$

which will be almost completely localized at the left and right quantum dots, respectively. For the subsequent calculations we shall neglect the product of the functions $\Psi_L(\mathbf{r})\Psi_R(\mathbf{r})$, assuming that this is small.

3. SELECTION RULES FOR QUANTUM TRANSITIONS

We shall begin with calculations of the spectrum and wave functions of the steady states of a two-electron system for $\mathscr{E} = 0$ for which we find the eigenfunctions and eigenvalues of the Hamilton operator

$$\hat{H}(\mathbf{r}_1, \mathbf{r}_2) = \hat{H}_0(\mathbf{r}_1) + \hat{H}_0(\mathbf{r}_2) + \frac{e^2}{\epsilon|\mathbf{r}_1 - \mathbf{r}_2|}, \qquad (1)$$

where $\hat{H}_0(\mathbf{r})$ is the single-electron Hamilton operator in a double quantum dot (its eigenfunctions $\chi_{0,1}(\mathbf{r})$ and eigenvalues $\pm\Delta/2$). As a result of using the two-level approximation in the single-electron problem the eigenfunction $\Psi(\mathbf{r}_1, \mathbf{r}_2)$ of the operator $\hat{H}(\mathbf{r}_1, \mathbf{r}_2)$ will be a superposition of four orthonormalized vectors $\chi_i(\mathbf{r}_1)\chi_j(\mathbf{r}_2)$ $(i, j = 0, 1)$ with various expansion coefficients.

As a result of solving the equation

$$\hat{H}\Psi_j(\mathbf{r}_1, \mathbf{r}_2) = E_j\Psi_j(\mathbf{r}_1, \mathbf{r}_2), \qquad (2)$$

we arrive at the following expressions for the energies E_j and wave functions of the steady states:

$$E_{0,3} = U \mp \sqrt{V^2 + \Delta^2}, \quad E_{1,2} = U \mp V; \qquad (3)$$

$$\Psi_{0,3}(\mathbf{r}_1, \mathbf{r}_2) = \frac{\varepsilon_{0,3} + V + \Delta}{2\sqrt{(\varepsilon_{0,3} + \Delta)^2 + V^2}} \times (\Psi_L(\mathbf{r}_1)\Psi_L(\mathbf{r}_2) + \Psi_R(\mathbf{r}_1)\Psi_R(\mathbf{r}_2)) - \frac{\varepsilon_{0,3} - V + \Delta}{2\sqrt{(\varepsilon_{0,3} + \Delta)^2 + V^2}} \times (\Psi_L(\mathbf{r}_1)\Psi_R(\mathbf{r}_2) + \Psi_R(\mathbf{r}_1)\Psi_L(\mathbf{r}_2)), \qquad (4)$$

$$\Psi_{1,2}(\mathbf{r}_1, \mathbf{r}_2) = \frac{\Psi_{L,R}(\mathbf{r}_1)\Psi_R(\mathbf{r}_2) - \Psi_{R,L}(\mathbf{r}_1)\Psi_L(\mathbf{r}_2)}{\sqrt{2}}.$$

Here $\varepsilon_{0,3} = E_{0,3} - U$, and U and V are the two-particle (pair) matrix elements of the Coulomb interaction operator:

$$V = \frac{e^2}{2\epsilon}\int d\mathbf{r}_1 d\mathbf{r}_2 \times \frac{(\Psi_R^2(\mathbf{r}_1) - \Psi_L^2(\mathbf{r}_1))(\Psi_R^2(\mathbf{r}_2) - \Psi_L^2(\mathbf{r}_2))}{|\mathbf{r}_1 - \mathbf{r}_2|},$$

$$U = \frac{e^2}{2\epsilon}\int d\mathbf{r}_1 d\mathbf{r}_2 \times \frac{(\Psi_R^2(\mathbf{r}_1) + \Psi_L^2(\mathbf{r}_1))(\Psi_R^2(\mathbf{r}_2) + \Psi_L^2(\mathrm{r}_2))}{|\mathbf{r}_1 - \mathbf{r}_2|}.$$

It follows from Eqs. (4) that the wave functions of the zeroth, second, and third steady states are symmetric with respect to exchanges of the particles, whereas the wave function of the first steady state is antisymmetric. Here we do not write the spin part of the wave function because Coulomb interaction like the external electric field does not lead to spin flipping. The spin component, being symmetric for the first state and antisymmetric for the remaining states, conserves its form and can thus be omitted.

We shall now introduce the external electric field with the result that an additional term appears in the Hamiltonian (1)

$$\hat{W}(\mathbf{r}_1, \mathbf{r}_2) = e\mathscr{E}(\mathbf{r}_1 + \mathbf{r}_2)\sin(\omega t), \qquad (5)$$

this being the energy of the interaction of the electrons with the external field.

We express the wave function of the system in a variable field as a linear combination of the wave functions obtained for the steady states (4) with time-dependent coefficients:

$$\Psi(\mathbf{r}_1, \mathbf{r}_2, t) = \sum_{j=0}^{3} C_j(t)\Psi_j(\mathbf{r}_1, \mathbf{r}_2)\exp\left\{-i\frac{E_j}{\hbar}t\right\}. \qquad (6)$$

Substituting the Eq. (6) into the Schrödinger equation yields a system of equations for the four expansion coefficients:

$$i\hbar\frac{dC_0}{dt} = W_{02}(t)C_2\exp(-i\omega_{20}t), \quad i\hbar\frac{dC_1}{dt} = 0,$$

$$i\hbar\frac{dC_2}{dt} = W_{20}(t)C_0\exp(i\omega_{20}t) + W_{23}(t)C_3\exp(-i\omega_{32}t), \qquad (7)$$

$$i\hbar\frac{dC_3}{dt} = W_{32}(t)C_2\exp(i\omega_{32}t),$$

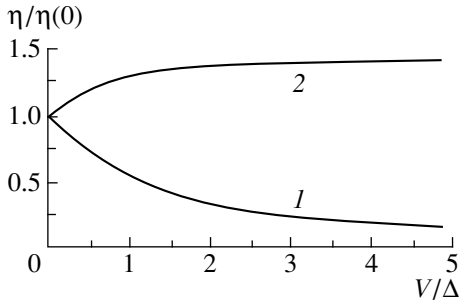

Fig. 1. Dependences of the resonant Rabi frequency for the (*1*) 0 ⟷ 2 and (*2*) 2 ⟷ 3 transitions on the Coulomb energy.

where $\omega_{ij} = (E_i - E_j)/\hbar$ and $W_{ij}(t)$are the matrix elements of the operator $\hat{W}$ where

$$W_{n2} = W_{2n} = \frac{\sqrt{2}(\varepsilon_n + V + \Delta)}{\sqrt{(\varepsilon_n + \Delta)^2 + V^2}} e\mathscr{E} \cdot \mathbf{L} \sin(\omega t) \quad (8)$$

for $n = 0, 3$ and all the other matrix elements W_{ij} go to zero. The vector **L** connects the origin positioned midway between two identical quantum dots and the center of the "right" quantum dot (to be specific we assume that the dots are spherical).

The matrix elements W_{ij} determine the so-called selection rules for transitions in a four-level spectrum (3). The expressions (8) show that in our system under the action of a variable electric field of frequency ω: it is only possible to have two transitions: between the ground and second excited states (0 ⟷ 2) and between the second and third excited states (2 ⟷ 3).

Transitions from or to the first excited state are impossible. The reason for this is the different symmetry of the wave functions of the first state and all the other states relative to particle exchanges and since the perturbation operator $\hat{W}(\mathbf{r}_1, \mathbf{r}_2)$ is symmetric with respect to exchanges, it cannot lead to transitions accompanied by a change in the symmetry of the wave function. Thus, the coefficient C_1 remains constant, as follows from the second equation in the system (7), and is defined only by the initial condition.

In addition to the selection rules determined by the specific structure of the wave functions of the steady states, there are also additional selection rules for the direction of the electric field vector $\mathscr{E}$. It can be seen from expressions (8) that in order to achieve 0 ⟷ 2 and 2 ⟷ 3 transitions the vector $\mathscr{E}$ must have a nonzero projection on the axis of the double quantum dot. In the opposite case when $\mathscr{E} \perp \mathbf{L}$, all the transitions are forbidden.

4. ELECTRON DENSITY OSCILLATIONS

We obtain solutions of Eqs. (7) in the so-called resonance approximation, when the frequency of the external field is close to the frequency of one of the transitions (see, for example, [14]). We shall assume that the external field is sufficiently weak for the inequality $|W_{ij}| \ll \hbar\omega$ to be satisfied. We shall begin with the case $\omega \approx \omega_{20}$ and we shall introduce the deviation from resonance $\delta = \omega - \omega_{20}$ which we shall assume to be small compared with the frequency ω. In Eqs. (7) we can only retain resonant harmonics as a result of which we find $C_3 \approx$ const and the remaining pair of equations for C_0 and C_2 are the standard evolution equations for a two-level system in a weak resonant external field.

As is well-known [14], the solution of the equations (7) in this case describes periodic transitions between two states coupled by the variable field (in our case, between the ground and second excited states) at the Rabi frequency

$$\nu = \sqrt{\delta^2/4 + \eta^2}, \quad (9)$$

where the parameter η defines the Rabi frequency in strict resonance and is given by

$$\eta = \frac{e\mathscr{E} \cdot \mathbf{L}}{\hbar} \sqrt{\frac{\sqrt{1 + V^2/\Delta^2} - V/\Delta}{2\sqrt{1 + V^2/\Delta^2}}}. \quad (10)$$

For the initial condition corresponding to a 100% ground-state population we have

$$C_0(t) = \left[\cos(\nu t) - i\frac{\delta}{2\nu}\sin(\nu t)\right]\exp\left(i\frac{\delta}{2}t\right),$$
$$C_2 = \frac{\eta}{\nu}\exp\left(i\frac{\delta}{2}t\right)\sin(\nu t). \quad (11)$$

These expressions are also well known from the theory of resonant excitation of a two-level system [14].

New here is the expression for the frequency η, which depends explicitly on the parameters characterizing the interelectron Coulomb interaction and the steady-state energy spectrum of the system. In the single-frequency case the Rabi frequency in strict resonance is determined only by the perturbation amplitude and does not depend on the structure of the single-electron spectrum. Allowance for Coulomb interaction abruptly changes the situation and the dependence of the Rabi frequency on the pair interaction parameters is fairly significant, as can be seen from Fig. 1 (curve *1*).

In particular, for noninteracting particles in the limit for $V \longrightarrow 0$ the frequency $\eta(0) = e\mathscr{E} \cdot \mathbf{L}/\hbar\sqrt{2}$ has a maximum, decreasing with increasing characteristic Coulomb energy *V*. The decrease in η compared with η(0) as a function of *V* is determined by the level splitting energy Δ (on scales $V \sim \Delta$ the frequency decreases appreciably). For quantum dots having dimensions of several nanometers [11–13] and distances between dots

no exceeding their size, the ratio Δ/V becomes fairly small (of the order of 0.1) which leads to a reduction in the Rabi frequency by more than order of magnitude.

For the case of the other possible transition $2 \longleftrightarrow 3$ the two equations for the coefficients C_2 and C_3 can be retained in Eqs. (7) and the coefficient C_0 can be considered to be almost constant. The solution of these equations is exactly the same as the previous case for the $0 \longleftrightarrow 2$ transition and yields exactly the same structure of the expression for the Rabi frequency (9) and the same solutions (11) in which we now merely need to substitute $C_0 \longrightarrow C_2$ and $C_2 \longrightarrow C_3$ (assuming that the second level is initially completely filled).

The only difference will be the specific dependence of the resonant Rabi frequency η on the parameters of the energy spectrum. Instead of Eq. (10) we will now have

$$\eta = \frac{e\mathscr{E}\cdot\mathbf{L}}{\hbar}\sqrt{\frac{\sqrt{1+V^2/\Delta^2}+V/\Delta}{2\sqrt{1+V^2/\Delta^2}}}. \qquad (12)$$

In accordance with (12) the resonant Rabi frequency for the $2 \longleftrightarrow 3$ transition of an interacting two-electron system is always higher than the Rabi frequency for a system of noninteracting electrons (see Fig. 1, curve *2*) and as V increases, the graph of the ratio $\eta/\eta(0)$ reaches a constant whose value is $\sqrt{2}$.

Unlike the $0 \longleftrightarrow 2$transition for which the Rabi frequency decreases monotonically as the characteristic Coulomb energy V increases and may decrease arbitrarily strongly as far as zero, for the $2 \longleftrightarrow 3$ transition the Rabi frequency does not change substantially, remaining close to its value in the complete absence of interaction. As has been noted, the maximum increase in the Rabi frequency relative to this value may be a factor of $\sqrt{2}$.

However, it should be noted that under real conditions at temperatures below or around room temperature, and for the sizes of quantum dots noted above the thermal energy is considerably lower than the Coulomb energy V and consequently the probability of $2 \longleftrightarrow 3$ transitions will be almost zero because of the negligible population of these levels. Conversely, the $0 \longleftrightarrow 2$ transition will be quite feasible so that the following discussion of the electron dynamics will be made for this case.

We shall now analyze the spatial distribution of the charge in a double quantum dot and its variation with time. For this we shall use the wave function (6) in which we set $C_1 = C_3 = 0$ and for C_0 and C_2 we use Eqs. (11). We shall calculate the quantity of charge concentrated, for example, in the left quantum dot $Q_L(t)$ as a function of time. Bearing in mind that the wave function will be strictly symmetric with respect to particle exchanges and also that the functions $\Psi_L(\mathbf{r})$ and $\Psi_R(\mathbf{r})$ will be almost completely localized in "their"

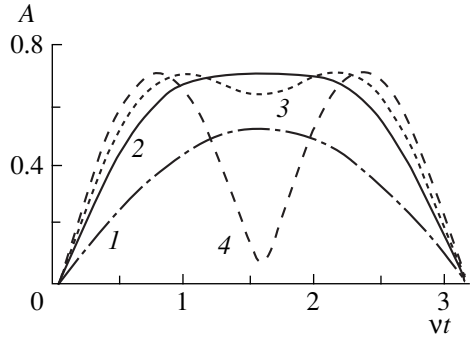

Fig. 2. Beat amplitude in the absence of interelectron interaction for various values of the ratio η/ν: η/ν = (*1*) 0.1, (*2*) $1/\sqrt{2}$, (*3*) 0.85, and (*4*) 0.999.

quantum dots, we shall determine the charge $Q_L(t)$ as an integral over the region $z < 0$ from the single-particle distribution density multiplied by $-2e$.

As a result, we obtain the following expression for $Q_L(t)$ [naturally the charge $Q_R(t)$ in the right quantum dot can be obtained as $-2e - Q_L(t)$]:

$$Q_L(t) = -e\left[1 - \sqrt{\frac{\sqrt{1+V^2/\Delta^2}-V/\Delta}{\sqrt{1+V^2/\Delta^2}}}A(t)\sin[\varphi(t)+\omega t]\right], \qquad (13)$$

where $A(t)$ and $\varphi(t)$ are slowly varying functions of time given by

$$A(t) = \sqrt{2}\frac{\eta}{\nu}|\sin(\nu t)|\sqrt{1-\frac{\eta^2}{\nu^2}\sin^2(\nu t)},$$
$$\cot\varphi(t) = \frac{\delta}{2\nu}\tan(\nu t). \qquad (14)$$

The dependence $Q_L(t)$ exhibits a characteristic beat form. Fast oscillations of charge between the quantum dots which take place at a frequency close to that of the external electric field are modulated in amplitude by a slower function $A(t)$ of frequency ν.

The functions $A(t)$ and $\varphi(t)$ have the simplest form in the two limiting cases: $\delta = 0$ and $\delta \gg \eta$. For $\delta = 0$ Eq. (14) gives: $A(t) = |\sin 2\nu t|/\sqrt{2}$, $\varphi(t) = \pi/2$ and for $\delta \gg \eta$ we have $A(t) = (\sqrt{2}\,\eta/\nu)|\sin\nu t|$, $\varphi(t) = \pi/2 - \nu t$. For all other intermediate cases the dependences $A(t)$ and $\varphi(t)$ are more complex and have the form shown in Fig. 2. It can be seen that as δ decreases and the ratio η/ν increases, the envelope of $A(t)$ differs increasingly from sinusoidal. For $\eta/\nu > 1/\sqrt{2}$ the curve of $A(t)$ at the point $\nu t = \pi/2$ has a minimum instead of a maximum, which progresses with increasing δ and in the limit $\eta/\nu \longrightarrow 0$ falls to zero, doubling the envelope frequency.

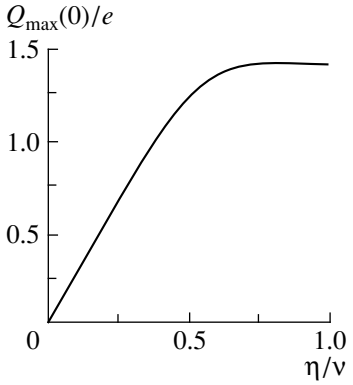

Fig. 3. Maximum quantity of charge oscillating between quantum dots in the absence of interaction. From $\eta/\nu = 1/\sqrt{2}$ the curve reaches a constant value of $\sqrt{2}$.

The function $eA(t)$ describes the amplitude of the charge oscillations for the case of noninteracting particles. The maximum value of the function $A(t)$ is $\eta\delta/(\sqrt{2}\,\nu^2)$ for $0 < \eta/\nu \le 1/\sqrt{2}$ and $1/\sqrt{2}$ for $1/\sqrt{2} \le \eta/\nu \le 1$. The maximum quantity of charge flowing from one quantum dot to another in the absence of interparticle interaction $Q_{max}(0)$ may be defined as the product of the maximum of the function $A(t)$ and twice the electron charge $2e$. The dependence of $Q_{max}(0)$ on the ratio η/ν is plotted in Fig. 3. Far from resonance ($\eta \ll \delta$) very little charge tunnels from one dot to another, keeping the equal populations of both quantum dots constant (in terms of $-e$). As resonance is approached, the amplitude of the charge oscillations increases and from $\eta/\nu = 1/\sqrt{2}$ these oscillations are the most intensive, at certain times the charge in one quantum dot is $-e(1 + 1/\sqrt{2})$.

Coulomb interaction between the electrons changes the amplitude of the envelope: now, in accordance with Eq. (13) the amplitude acquires an additional factor which is equal to the ratio of the Rabi frequency for the interacting electrons (10) to the Rabi frequency in the absence of interaction $\eta(0)$. The maximum quantity of charge flowing from one quantum dot to another for $V \neq 0$ is now also determined by this factor:

$$Q_{max}(V) = Q_{max}(0)\frac{\eta}{\eta(0)}. \qquad (15)$$

The frequency ratio $\eta/\eta(0)$ whose dependence on the Coulomb energy V was discussed above (see Fig. 1) thus accurately determines the ratio of the maximum charges $Q_{max}(V)/Q_{max}(0)$.

Consequently, on the basis of (15) and Fig. 1 we can confirm that the quantity of charge tunneling between quantum dots will always be smaller for interacting particles than in the absence of interaction. Moreover, as soon as the typical Coulomb energy V exceeds the level splitting energy Δ, the amount of flowing charge falls rapidly and for $V \gg \Delta$ we can assume that no tunneling takes place between the quantum dots. As has been noted for quantum dots having dimensions of a few nanometers, the ratio $\eta/\eta(0)$ is less than or of the order of 0.1 which reduces the quantity of tunneling charge by more than an order of magnitude.

The sharp drop in the amplitude of the charge oscillations and their slowing in a weak variable electric field is another manifestation of the so-called Coulomb tunneling blockade. A reduction in the sizes of the quantum dots accompanied by convergence of the electrons leads to fairly strong interelectron repulsion forces, which cannot be overcome in weak electric fields even in a resonance regime.

REFERENCES

1. F. Grossmann, T. Dittrich, P. Jung, and P. Hanggi, Phys. Rev. Lett. **67**, 516 (1991).
2. R. Bavli and H. Metiu, Phys. Rev. A **47**, 3299 (1993).
3. A. A. Gorbatsevich, V. V. Kapaev, and Yu. V. Kopaev, Zh. Éksp. Teor. Fiz. **107**, 1321 (1995) [JETP **80**, 734 (1995)].
4. Y. Dakhnovskii, R. Bavli, and H. Metiu, Phys. Rev. B **53**, 4657 (1996).
5. V. A. Burdov, Zh. Éksp. Teor. Fiz. **116**, 217 (1999) [JETP **89**, 119 (1999)].
6. C. A. Stafford and N. S. Wingreen, Phys. Rev. Lett. **76**, 1916 (1996).
7. K. Leo, J. Shah, E. O. Gobel, *et al.*, Phys. Rev. Lett. **66**, 201 (1991).
8. I. Brener, P. C. M. Planken, M. C. Nuss, *et al.*, Appl. Phys. Lett. **63**, 2213 (1993).
9. M. S. C. Luo, S. L. Chuang, P. C. M. Planken, *et al.*, Phys. Rev. B **48**, 11043 (1993).
10. R. H. Blick, D. W. van der Weide, R. J. Haug, and K. Eberl, Phys. Rev. Lett. **81**, 689 (1998).
11. M. A. Olshavsky, A. N. Goldstein, and A. P. Alivisatos, J. Am. Chem. Soc. **112**, 9438 (1990).
12. Q. Ye, R. Tsu, and E. H. Nicolian, Phys. Rev. B **44**, 1806 (1991).
13. A. D. Andreev and A. A. Lipovskii, Phys. Rev. B **59**, 15402 (1999).
14. L. D. Landau and E. M. Lifshitz, *Quantum Mechanics: Non-Relativistic Theory* (Nauka, Moscow, 1989, 4th ed.; Pergamon, Oxford, 1977, 3rd ed.).

Translation was provided by AIP

Journal of Experimental and Theoretical Physics, Vol. 91, No. 3, 2000, pp. 615–625.
Translated from Zhurnal Éksperimental'noĭ i Teoreticheskoĭ Fiziki, Vol. 118, No. 3, 2000, pp. 707–719.

NONLINEAR PHYSICS

Geometry of a Chaotic Layer

I. I. Shevchenko

Pulkovo Observatory, Russian Academy of Sciences, St. Petersburg, 196140 Russia
e-mail: iis@gao.spb.ru

Received March 13, 2000

Abstract—An analysis is made of the dependence of the geometric shape of the chaotic layer near the separatrix of a nonlinear resonance of a Hamiltonian system on the parameters of this system. A separatrix algorithmic mapping, which describes the motion near the separatrix in the presence of an asymmetric perturbation having an arbitrary degree of asymmetry. The separatrix algorithmic mapping is an algorithm containing conditional transfer instructions, is considered. An analytic procedure is derived to reduce the separatrix algorithmic mapping to the unified surface of the cross section of the initial Hamiltonian system (mapping synchronization procedure). It is observed that in the case of the high-frequency perturbation $\lambda \longrightarrow +\infty$ (where λ is the ratio of the perturbation frequency to the frequency of small phase oscillations at resonance), the chaotic layer is subjected to strong bending in the sense that during motion near the separatrix the amplitude of the energy deviations relative to the unperturbed separatrix value is much larger than the layer width. However, the synchronized separatrix algorithmic mapping ensures an accurate representation of the phase portrait of the layer for both low and high values of the parameter λ provided that the amplitude of the perturbation is fairly small. This is demonstrated by comparing the phase portraits obtained using the synchronized separatrix algorithmic mapping with the results of direct numerical integrations of the initial Hamiltonian system. *© 2000 MAIK "Nauka/Interperiodica".*

1. INTRODUCTION

The theory of the chaotic layer [1–5] in Hamiltonian dynamics has applications in the most diverse fields of physics and mechanics (see, for example, [5, 6] and the literature cited there). The most important characteristic of the chaotic layer is its width. Problems associated with estimating this have been studied in detail in [4, 5, 7, 8]. However, the geometric profile of the layer and how it is influenced by the parameters of the dynamic system have been very little studied (except for the role of marginal resonances [4, 5, 9]). The procedure for synchronization of the separatrix mapping [9] (reducing it to the unified surface of the cross section) can be used to obtain real phase portraits of the motion near the separatrix and thus to assess the real profile of the chaotic layer. In the present study the synchronization procedure is derived for the general case of an asymmetric perturbation and is applied to analyze the layer geometry.

1.1. Hamiltonian of the Problem

The equations for a nonlinear pendulum with periodic perturbations are an important paradigm in various fields of modern physics and mechanics, ranging from plasma physics [10, 11] to the dynamics of orbital [12] and spin-orbit [13,14] resonances in the motion of celestial bodies.

We shall consider the Hamiltonian

$$H = \frac{\mathcal{G}p^2}{2} - \mathcal{F}\cos\varphi + \cos(k\varphi - \tau) + b\cos(k\varphi + \tau), \quad (1)$$

where $\tau = \Omega t + \tau_0$. The first two terms are the Hamiltonian of the pendulum while the last two are the periodic perturbations. The variable φ is the deviation of the pendulum from the equilibrium position, p is its conjugate momentum, τ is the phase angle of the perturbation, and τ_0 is its initial value. The value of Ω is the perturbation frequency, and $\mathcal{F}$, $\mathcal{G}$, a, b, and k are constants. The Hamiltonian of the unperturbed pendulum is then denoted by H_0, i.e.,

$$H_0 = \frac{\mathcal{G}p^2}{2} - \mathcal{F}\cos\varphi.$$

The number k is assumed to be integer or half-integer.

The Hamiltonian (1) is present in many problems in mechanics and physics. The case of a symmetric perturbation $a = b$ and $k = 1$ has been studied in the greatest detail. In this case, the Hamiltonian is reduced to the form of the Hamiltonian of a pendulum with a modulated frequency of small oscillations. In particular, it corresponds to a pendulum with a vertically oscillating suspension point [15].

The case with an arbitrary nonzero value of k and zero a or b has been studied as the problem of particle

motion in the field of two plane waves [16–19]. The case of $k = 1/2$ and special values of a, b defines an approximate description of motion in orbital 3/1 resonance in satellite systems [12, 20]. The case $k = 1$, $b = -a/7$ corresponds to the problem of the rotational motion of a nonspherical satellite in synchronous spin–orbit resonance in an elliptic orbit [13, 14].

The reason for the universal nature of the Hamiltonian (1) is that a nonlinear pendulum is a model of nonlinear resonance under very general conditions [4–6].

1.2. Separatrix Mapping

Motion near a pendulum separatrix or a nonlinear resonance separatrix in the symmetric case $a = b$ (and $k = 1$) was considered by Chirikov in [4, 5]. He showed that this motion is effectively described by some mapping known as a "whisker mapping." Now the term "separatrix mappings" is more commonly used for this type of mapping. The description of motion using a mapping affords certain advantages both in qualitative analyses of motion and in computer analyses.

The usual separatrix mapping (in the form [4, 5]) has the form

$$\begin{aligned} w_{n+1} &= w_n - W\sin\tau_n, \\ \tau_{n+1} &= \tau_n + \lambda\ln\frac{32}{|w_{n+1}|} \pmod{2\pi}, \end{aligned} \tag{2}$$

where w denotes the relative (to the value on the separatrix) pendulum energy $w = H_0/\mathcal{F} - 1$ and τ is the phase of the perturbation as above. The constants λ and W are parameters: λ is the ratio of the perturbation frequency Ω to the frequency $\omega_0 = (\mathcal{F}\mathcal{G})^{1/2}$ of the small pendulum oscillations, and

$$W = \frac{a}{\mathcal{F}}\lambda\alpha_2^c(\lambda), \tag{3}$$

where $\alpha_2^c(\lambda) = A_2(\lambda) + A_2(-\lambda)$,

$$A_2(\lambda) = 4\pi\lambda\frac{\exp(\pi\lambda/2)}{\sinh(\pi\lambda)}$$

is the Melnikov–Arnold integral [4–6]; the functions $\alpha_k^c(\lambda)$ ($k = 0, 1, 2, \ldots$) are analyzed in the Appendix. The notation ε is subsequently used for $a/\mathcal{F}$.

A single iteration of the separatrix mapping (2) corresponds to a single period of the pendulum rotation or a half-period of its oscillation.

An important property of the separatrix mapping in the form (2) is that it represents the motion asynchronously [9–11]: the action variable is mapped for $\varphi = \pm\pi$ whereas the phase angle of the perturbation is mapped for $\varphi = 0$. This property follows from the derivation procedure [4–6] for the separatrix mapping (2). Abdullaev and Zaslavsky [10, 11] derived a "shifted" separatrix mapping in which the phase angle of the perturbation can be taken at an arbitrary cross section. In this case, the action variable is mapped for $\varphi = \pm\pi$ as in the usual separatrix representation.

In order to completely describe the motion around a separatrix we need to find a procedure for synchronizing the mapping to the surface of the cross section $\varphi = 0$, i.e., for the position of the pendulum in stable equilibrium. In fact, the cross section $\varphi = \pm\pi$ (for the position of the pendulum in unstable equilibrium) does not completely represent the dynamic behavior since the angle of the pendulum during oscillations does not reach $\pm\pi$.

The procedure for synchronizing an ordinary separatrix mapping (case of a symmetric perturbation) to the surface of the cross section $\varphi = 0$ was described in [9]. As was noted in that article, desynchronization of the separatrix mapping leads, in particular, to asymmetry of the phase portraits relative to the lines $\tau = 0$ or $\tau = \pi$ (see for example Fig. 1 in [21], Fig. 6 in [22], or Fig. 3b in [23]).

In the present study the synchronization procedure is derived for the general case of an asymmetric perturbation (1). Synchronization confers meaning to comparisons of cross-section surfaces obtained by numerical integration with the corresponding phase portraits of a separatrix mapping. Such comparisons provide direct information on the range of validity of the separatrix mapping in parameter space. By means of such comparisons we confirm that the separatrix mapping can be applied over a wide range of values of the parameter λ, both for low-frequency and for high-frequency perturbations, provided that their amplitudes are fairly small.

2. SEPARATRIX ALGORITHMIC MAPPING

The separatrix mapping for an asymmetric perturbation differs from that for the symmetric case since the energy increments differ for the forward and backward motion of the pendulum. Following [20, 24], we write the separatrix mapping in the asymmetric case in the form of an algorithm which we shall call a separatrix algorithmic mapping:

$$\begin{aligned} &\text{if } w_n < 0 \text{ and } W = W^{\pm}, \text{ then } W := W^{\mp}; \\ &w_{n+1} = w_n - W\sin\tau_n, \\ &\tau_{n+1} = \tau_n + \Delta_{n+1}\tau \pmod{2\pi}. \end{aligned} \tag{4}$$

In this mapping the sign of the upper index of W alternates on each iteration if $w_n < 0$ (i.e., during oscillations); $W^{\pm}$ implies W^+ or W^- and $W^{\mp}$ implies W^- or W^+, respectively. The value of $\Delta_{n+1}\tau$, approximately equal to $\lambda\ln(32/|w_{n+1}|)$ [as is assumed for the ordinary separatrix mapping (2)], is determined more accurately below, while W^+ and W^- are the values of the parameter W for the forward and backward motion, respectively. Here the motion is called forward (or backward respectively) if the change of φ with time is positive (nega-

tive). Any functions having the indices + or – subsequently denote forward and backward motion, respectively.

The separatrix algorithmic mapping may be written in a more detailed "expanded" form if we introduce numbering of W as for the variables τ and w; however we stress that W is a parameter which can only have two values W^+ and W^-. The separatrix algorithmic mapping in expanded form is as follows:

$$\begin{aligned} &\text{if } w_n < 0 \text{ and } W_n = W^-, \text{ then } W_{n+1} := W^+, \\ &\text{if } w_n < 0 \text{ and } W_n = W^+, \text{ then } W_{n+1} := W^-; \\ &w_{n+1} = w_n - W_{n+1}\sin\tau_n, \\ &\tau_{n+1} = \tau_n + \Delta_{n+1}\tau \pmod{2\pi}. \end{aligned} \tag{5}$$

The essence of the algorithm (4), (5) is that it takes into account the alternation of the values of the parameter W. These alternate when the direction of motion changes. The algorithm (4), (5) does not contain the condition $w_n > 0$ (case of rotation) because in this case the direction of motion remains unchanged.

The logarithmic expression for the phase increment in the ordinary separatrix mapping (2) is a fairly rough approximation. Thus, using (2) we can only predict an exact representation of the phase portraits of the motion for very small values of the perturbation amplitude W. If the perturbation is not weak, in order to improve the accuracy of the separatrix algorithmic mapping (4), (5), the logarithmic approximation of the phase increment should be replaced by its exact value. The analytic expressions for this differ in regions of oscillation and rotation and have the form [9]

$$\Delta_{n+1}\tau = \begin{cases} 2\lambda K\left(\left(1+\dfrac{w_{n+1}}{2}\right)^{1/2}\right), & w_{n+1} < 0 \\ 2\lambda\left(1+\dfrac{w_{n+1}}{2}\right)^{-1/2} K\left(\left(1+\dfrac{w_{n+1}}{2}\right)^{-1/2}\right), & w_{n+1} > 0, \end{cases} \tag{6}$$

where $K(k)$ is a complete elliptic integral of the first kind. The first line in Eq. (6) corresponds to the oscillations of a model pendulum while the second corresponds to its rotations.

By using this separatrix algorithmic mapping we can rapidly reproduce the cross sections of the phase space of the system near the separatrix. Using this mapping gives a relative advantage of two or three orders of magnitude in terms of calculation speed. However, its use not only involves an increase in calculation speed. More importantly, this theory gives a direct analytic description of the phase space of a system: an analysis of the separatrix mapping allows us to precalculate the locations of resonance and chaos boundaries [24], and the appearance of marginal resonances [9] by means of compact analytic expressions. For instance, when the values of the parameters of the separatrix algorithmic mapping are known, it is easy to find the conditions of Hamiltonian intermittency [9] which are the same as the conditions for the appearance of marginal resonances.

3. SYNCHRONIZATION PROCEDURE

In order to find the cross section of phase space for $\varphi = 0$ we need to calculate the energy increment not in the time limits $(-\infty, +\infty)$ as are used to derive the ordinary separatrix mapping but in the limit $(-\infty, 0)$ or $(0, +\infty)$. We first give various formulas required for the motion at the unperturbed separatrix:

$$\varphi^\pm(t) = \pm[4\arctan\exp(t) - \pi]. \tag{7}$$

$$p^\pm(t) = \pm\frac{2\omega_0}{\mathcal{G}}\cos\frac{\varphi^\pm(t)}{2}. \tag{8}$$

The plus sign in these expressions corresponds to the forward motion of the pendulum while the minus sign corresponds to the backward motion.

We shall calculate the required energy increment:

$$\Delta H_0 = \int_{-\infty}^{0}\frac{dH_0}{dt}dt.$$

The derivative is given by the Poisson bracket:

$$\begin{aligned} \frac{dH_0}{dt} &= \{H_0, H\} = \frac{\partial H_0}{\partial\varphi}\frac{\partial H}{\partial p} - \frac{\partial H}{\partial\varphi}\frac{\partial H_0}{\partial p} \\ &= ak\mathcal{G}p[\sin(k\varphi - \tau) + \eta\sin(k\varphi + \tau)], \end{aligned} \tag{9}$$

where $\tau = \Omega t + \tau_0$, $\eta = b/a$. The value of ΔH_0 is expressed in terms of the Melnikov–Arnold (MA) integrals (see Appendix). For the case of forward motion we have

$$\begin{aligned} \Delta^+ H_0 &= -a\frac{\lambda}{2}[A_{2k}(\lambda) + \eta A_{2k}(-\lambda)]\sin\tau_0 \\ &- a\{\lambda[B_{2k}(\lambda) - \eta B_{2k}(-\lambda)] + 1 + \eta\}\cos\tau_0. \end{aligned} \tag{10}$$

The derivation of the expression for backward motion is similar. We shall subsequently calculate the change in the pendulum energy H_0 relative to the line of the unperturbed separatrix $H_0 = \mathcal{F} = \text{const}$, in accordance with the definition of the relative energy $w = H_0/\mathcal{F} - 1$ given above. In general for the increment of the relative energy w we have

$$\Delta^\pm w = -\frac{W^\pm}{2}\sin\tau_0 - \delta^\pm W^\pm\cos\tau_0, \tag{11}$$

where

$$W^+(k, \lambda, \eta) = \frac{a}{\mathcal{F}}\lambda[A_{2k}(\lambda) + \eta A_{2k}(-\lambda)], \tag{12}$$

$$W^-(k, \lambda, \eta) = \frac{a}{\mathcal{F}}\lambda[\eta A_{2k}(\lambda) + A_{2k}(-\lambda)], \tag{13}$$

$$\delta^{\pm}(k, \lambda, \eta) = \frac{\lambda[B_{2k}(\lambda) - \eta B_{2k}(-\lambda)] + 1 + \eta}{\lambda[A_{2k}(\lambda) + \eta A_{2k}(-\lambda)]}, \quad (14)$$

$$\delta^{-}(k, \lambda, \eta) = \frac{\lambda[\eta B_{2k}(\lambda) - B_{2k}(-\lambda)] + 1 + \eta}{\lambda[\eta A_{2k}(\lambda) + A_{2k}(-\lambda)]}. \quad (15)$$

Formulas (14) and (15) for δ are valid for calculating the changes in H_0; i.e., the relative energy is defined as $w = H_0/\mathcal{F} - 1$ (see above). They will be different for the changes in the total energy H; i.e., if $w = H_0/\mathcal{F} - 1$. In this last case, the derivation is slightly simpler since

$$\frac{dH}{dt} = \frac{\partial H}{\partial t}.$$

The expressions for $W^{\pm}$ remain the same as for the H_0 case but the expressions for $\delta^{\pm}$ change and have the form

$$\delta^{+}(k, \lambda, \eta) = \frac{B_{2k}(\lambda) - \eta B_{2k}(-\lambda)}{A_{2k}(\lambda) + \eta A_{2k}(-\lambda)}, \quad (16)$$

$$\delta^{-}(k, \lambda, \eta) = \frac{\eta B_{2k}(\lambda) - B_{2k}(-\lambda)}{\eta A_{2k}(\lambda) + A_{2k}(-\lambda)}. \quad (17)$$

Everywhere subsequently we take the definition of w in terms of H_0; i.e., Eqs. (16) and (17) are not used for the calculations in this article. We also note that all the formulas are given here for arbitrary integer or half-integer values of the parameter k, but all the computer experiments described below are given for $k = 1$. Construction of the cross sections for $k \neq 1$ requires separate analysis; the case of a separatrix algorithmic mapping with $k = 1/2$ was considered in [20].

The value of the relative energy on the surface $\varphi = 0$ when its value w_n on the surface $\varphi = \pm\pi$ is known, is given by the algebraic summation of w_n and the calculated increment $\Delta^{\pm}w_n$:

$$\begin{aligned}\tilde{w}_n &= \tilde{w}_n - \frac{W^{\pm}}{2}\sin\tau_n - \delta^{\pm}W^{\pm}\cos\tau_n \\ &= \frac{w_n + w_{n+1}}{2} - \delta^{\pm}W^{\pm}\cos\tau_n,\end{aligned} \quad (18)$$

where w_{n+1} is the next iteration of the nonsynchronized separatrix algorithmic mapping. The phase portrait of the separatrix mapping is synchronized by replacing the pair (w_n, τ_n) with $(\tilde{w}_n, \tau_n)$.

In other words, the procedure (18) synchronizes the separatrix algorithmic mapping (4) where the index of δ ("+" or "–") is the same as for the instantaneous value of W. In the specific case of the ordinary separatrix mapping (symmetric perturbation), quantities having the indices "+" and "–" are equal to each other. The algorithm (4) together with the synchronization procedure (18) can be used to construct synchronized phase portraits of the motion near the separatrix of the Hamiltonian (1).

The new variable $\tilde{w}$ is dynamic. The separatrix algorithmic mapping can be written in terms of the new ("synchronized") variables $\tilde{w}, \tau$. The change of variables $w, \tau \longrightarrow \tilde{w}, \tau$ is canonical. However, the new mapping is implicit. Since it is merely of theoretical interest and not very suitable for practical calculations, we shall confine ourselves to writing the mapping in terms of the new variables merely in the symmetric case $a = b$:

$$\begin{aligned}\tilde{w}_{n+1} &= \tilde{w}_n - W\sin\frac{\tau_n + \tau_{n+1}}{2} \\ &\times\left(\cos\frac{\tau_{n+1} - \tau_n}{2} - 2\delta\sin\frac{\tau_{n+1} - \tau_n}{2}\right), \\ \tau_{n+1} &= \tau_n + \lambda\ln 32 - \lambda\ln\left|\frac{\tilde{w}_n + \tilde{w}_{n+1}}{2}\right. \\ &+ \frac{W}{2}\cos\frac{\tau_n + \tau_{n+1}}{2} \\ &\left.\times\left(\sin\frac{\tau_{n+1} - \tau_n}{2} + 2\delta\cos\frac{\tau_{n+1} - \tau_n}{2}\right)\right| \pmod{2\pi}.\end{aligned} \quad (19)$$

In practice, it is clearly easier to iterate a nonsynchronized mapping and by thus calculating w_n, w_{n+1}, and τ_n, go over to the variables $\tilde{w}_n$ and τ_n using Eq. (18).

To conclude this section, we note that the synchronization parameter δ can be determined not only analytically but also by means of a direct numerical estimate of the MA integrals in Eqs. (14), (15) or (16), (17). We shall consider the case $k = 1$. The primitives of the integrand expressions in the formulas for $\alpha_2^c(\lambda)$ and $\beta_2^s(\lambda)$ (quantities in terms of which δ is expressed, see their integral representations in the Appendix) as a function of the independent variable t oscillate at $t \longrightarrow +\infty$ and the amplitude of the oscillations is nonzero in the limit. In the Appendix the analytic representations of $\alpha_2^c(\lambda)$ and $\beta_2^s(\lambda)$ are given in accordance with the agreement that in the limit when the upper limit of integration tends to infinity ($t \longrightarrow +\infty$) the mean is taken for the primitive. This implies that for the numerical integration the upper limits of integration should be taken as certain nonarbitrary discrete values given in Table 1 in the Appendix.

4. SYNCHRONIZATION FOR A SYMMETRIC PERTURBATION

We shall consider the case of a symmetric perturbation (we also assume $k = 1$) in greater detail. We recall that this corresponds to the ordinary separatrix mapping (2).

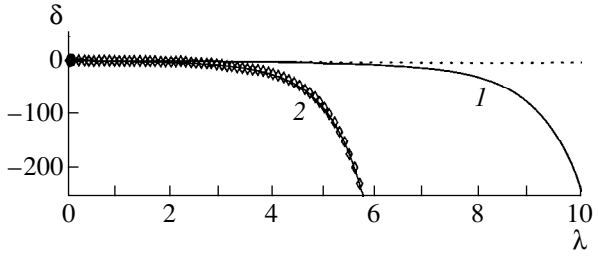

Fig. 1. Curve *1*, synchronization parameter $\delta(\lambda)$ (21) for a symmetric perturbation and $k = 1$ in the Hamiltonian (1). Curve *2*, synchronization parameter $\delta(\lambda)$ (23) for the Hamiltonian (22).

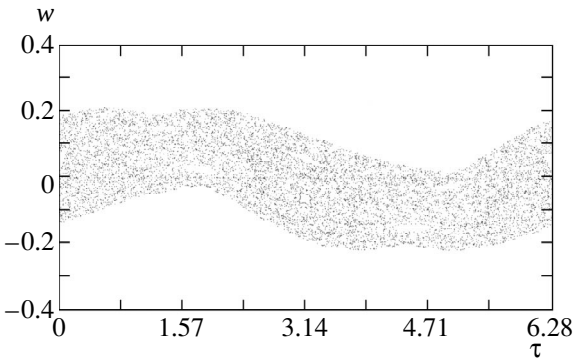

Fig. 2. Phase plane of separatrix mapping (2) with $\lambda = 0.5$, $W = 0.181$.

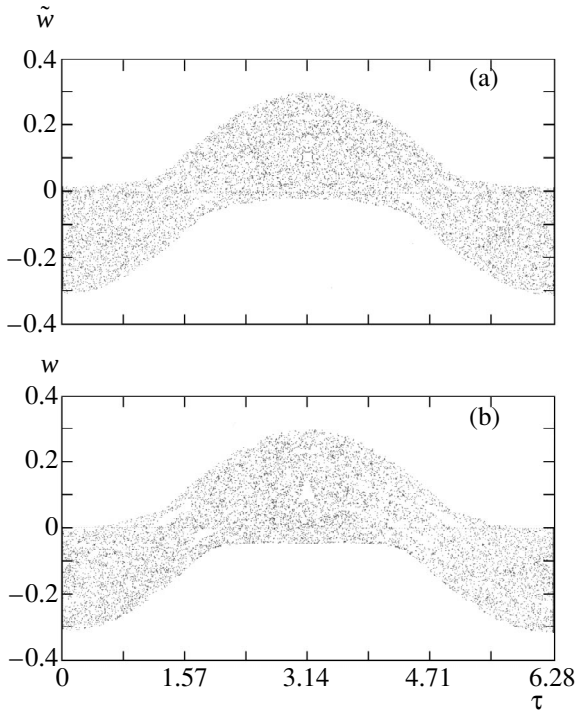

Fig. 3. (a) As Fig. 2 but synchronized to the unified plane of the cross section $\varphi = 0$; $\delta = 0.929$. (b) Cross section obtained by direct numerical integration. In Fig. 3b the synchronization sign (tilde) in the notation of the *w* axis is omitted since no analytic synchronization procedure is used in integration and the values of *w* and τ are obtained directly at the unified surface of the cross section $\varphi = 0$.

In this case, the expression for the synchronization parameter δ has the form

$$\delta(\lambda) = \frac{\beta_2^s(\lambda) + 2/\lambda}{\alpha_2^c(\lambda)}. \tag{20}$$

This is the same for the forward and backward motion. The functions $\alpha_2^c(\lambda)$ and $\beta_2^s(\lambda)$ are determined in the Appendix. Using the recurrence formulas given in the Appendix we find

$$\delta(\lambda) = \frac{1}{2\pi}\left(\beta_1^c(\lambda) + \frac{2}{\lambda^2}\right)\sinh\frac{\pi\lambda}{2} = \frac{1}{\pi} \times \left\{\mathrm{Re}\left[\psi\left(i\frac{\lambda}{2}\right) - \psi\left(i\frac{\lambda}{4}\right)\right] + \frac{1}{\lambda^2} - \ln 2\right\}\sinh\frac{\pi\lambda}{2}, \tag{21}$$

where $\psi(z) = \Gamma'(z)/\Gamma(z)$ is the digamma function (see Appendix), and *i* is the imaginary unit. This expression was given earlier in [9] without derivation. The dependence of the synchronization parameter δ on λ, Eq. (21), is plotted in Fig. 1 (curve *1*). We note the existence of a central plateau and the steep slopes to $\pm\infty$ in the low- and high-frequency limits.

We shall compare the phase portraits of the separatrix algorithmic mapping with the cross sections of the phase space of the initial Hamiltonian calculated directly (by numerical integration). By way of example we consider the case of a relatively low-frequency perturbation: $\lambda = 0.5$. In Figs. 2 and 3a the behavior of the same system is shown represented by nonsynchronized and synchronized separatrix mappings. Here and everywhere subsequently the figures only show the principal chaotic layer. Note that the asymmetry relative to the line $\tau = \pi$ which is quite appreciable in Fig. 2 had disappeared in Fig. 3a, i.e., this asymmetry is an artifact of the desynchronization of the initial separatrix mapping.

Figure 3b shows the same phase plane as in Fig. 3a but obtained by direct numerical integration of the system with the Hamiltonian (1). The values of the parameters in the Hamiltonian (1) correspond to the values of the separatrix mapping parameters $\lambda = 0.5$ and $W = 0.181$ and specifically $\mathcal{F} = 1$, $\mathcal{G} = 2$, $a = b = 0.05$, and $\Omega = 2^{-1/2}$. Here and subsequently integration is performed by the eighth-order Dormand–Prince method [25] with a controlled step size. The local accuracy of the integrator is set to 10^{-10}.

The calculated values of the variables *w* and τ were taken directly at the unified cross section surface $\varphi = 0$.

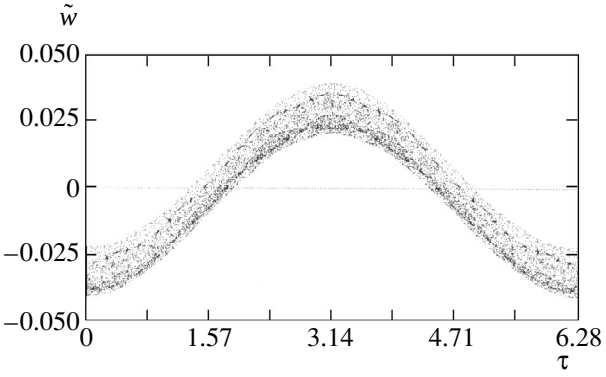

Fig. 4. Phase plane of synchronized separatrix mapping (2), (18) with $\lambda = 0.1$, $W = 0.00797$, $\delta = 4.99$.

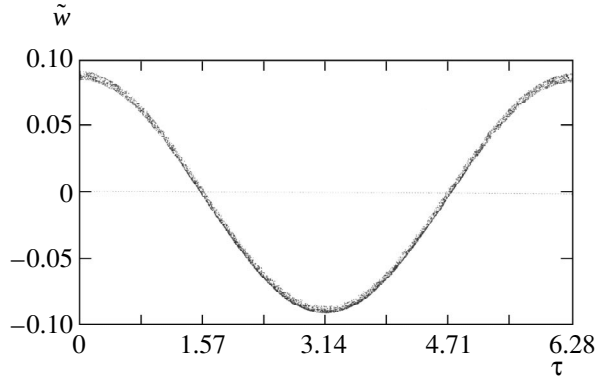

Fig. 5. Phase plane of synchronized separatrix mapping (2), (18) with $\lambda = 10$, $W = 0.000379$, $\delta = -232.5$.

The phase plane in Fig. 3b can be compared directly with the phase plane of the synchronized separatrix mapping in Fig. 3a. It can be seen that the behavior of the real system (1) is well described by the phase portrait of the synchronized separatrix mapping. The accuracy of the agreement along the $\tilde{w}$-axis is better than 0.03. The small differences disappear as the amplitude $a = b$ of the perturbation decreases.

Figures 4 and 5 give another two examples of the phase portraits of a synchronized separatrix mapping for cases of low- and high-frequency perturbations (λ being 0.1 and 10, respectively, where the amplitude of the perturbation ε in the initial Hamiltonian is 0.01 and 1, respectively). Numerical integration gives almost identical cross sections so that the figures giving the results of integration are not included here. The accuracy of the agreement along the $\tilde{w}$-axis is better than 0.001 in the first case and better than 0.002 in the second. The accuracy of the agreement between the amplitudes of the layer bending is better than 0.001 in both cases.

The case of an asymmetric perturbation is considered in Section 6.

5. SYNCHRONIZATION FOR A DIFFERENT TYPE OF PERTURBATION

As an example of a nonlinear resonance with a perturbation different to that considered earlier [see the Hamiltonian (1)] we consider the following Hamiltonian:

$$H = \frac{\mathcal{G}p^2}{2} - \mathcal{F}\cos\varphi + a\varphi\sin\tau. \tag{22}$$

Holmes and coworkers analyzed this model in a study of separatrix splitting [26].

An important difference from the case of the Hamiltonian (1) is that the synchronization formula here is the same for variations of H_0 and H since the perturbing term degenerates to zero for $\varphi = 0$. We obtain (the derivation is similar to that given in Section 3)

$$\begin{aligned} W(\lambda) \equiv W^+(\lambda) &= -W^-(\lambda) \\ &= \frac{a}{\mathcal{F}}\alpha_1^c(\lambda) = \frac{a}{\mathcal{F}}\frac{2\pi}{\cosh(\pi\lambda/2)}, \\ \delta(\lambda) \equiv \delta^+(\lambda) &= \delta^-(\lambda) = \frac{\beta_1^s(\lambda)}{\alpha_1^c(\lambda)} \\ &= \frac{1}{2}\sinh\frac{\pi\lambda}{2} - \frac{1}{\pi}\cosh\frac{\pi\lambda}{2}\operatorname{Im}\psi\left(\frac{1+i\lambda}{4}\right), \end{aligned} \tag{23}$$

and the separatrix algorithmic mapping has the universal form (4). In this case, it reduces to the simple form:

$$\begin{aligned} &\text{if } w_n < 0, \text{ then } W := -W; \\ &w_{n+1} = w_n - W\sin\tau_n, \\ &\tau_{n+1} = \tau_n + \Delta_{n+1}\tau \pmod{2\pi}. \end{aligned} \tag{24}$$

The dependence of the synchronization parameter δ on λ, Eq. (23), is shown in Fig. 1 (curve *2*). The qualitative difference from curve *1*, Eq. (21), is that for $\lambda \longrightarrow 0$ the value of $\delta(\lambda)$ does not go to $+\infty$ but tends to zero.

Figure 6 shows the behavior of the system (22) with the parameters $\lambda = 0.5$, $\varepsilon = 0.02$ represented using the synchronized separatrix algorithmic mapping (24), (18) with the corresponding parameters $\lambda = 0.5$, $W = 0.0949$, $\delta = -0.303$, Eq. (23). The phase portrait for the backward direction of motion (Fig. 6b) is simply a shift of the phase portrait for the forward direction (Fig. 6a) along the τ-axis by π. The results of finding the cross section $\varphi = 0$ by direct numerical integration are visually the same as Fig. 6 (the accuracy of the agreement along the $\tilde{w}$-axis is better than 0.01) and are not given here.

6. BENDING OF THE CHAOTIC LAYER IN THE HIGH-FREQUENCY PERTURBATION LIMIT

Figures 4 and 5 (examples corresponding to low and high perturbation frequencies) clearly show that the

chaotic layer is strongly bent. This bending is only observed for the synchronized mapping; the chaotic layer of the nonsynchronized separatrix mapping always completely encloses the line $w = 0$. Figure 1 (curve *1*) shows that the synchronization parameter $\delta(\lambda)$, Eq. (21), goes steeply to $\pm\infty$ in the low- and high-frequency perturbation limits. However, this does not yield any conclusions on the geometry of the chaotic layer. The energy amplitude of the layer bending is measured by the product δW [see Eq. (18)]. From Eqs. (12)–(15) for $k = 1$ we have: for a symmetric perturbation ($\eta = 1$) if $\lambda = 0$, then $\delta^+W^+ = \delta^-W^- = 4\varepsilon$, if $\lambda = +\infty$, then $\delta^+W^+ = \delta^-W^- = 0$; for a perturbation having maximum asymmetry ($\eta = 0$) if $\lambda = 0$, then $\delta^+W^+ = \delta^-W^- = 2\varepsilon$ and if $\lambda = +\infty$, $\delta^+W^+ = \delta^-W^- = 0$. Thus, in the high-frequency perturbation limit the energy amplitude of the layer bending always tends to zero.

Nevertheless, qualitatively the layer bending effect cannot be considered to be weak. We shall analyze the ratio of the amplitude of the layer bending and its half-width. In the high-frequency perturbation limit the latter is given by the product λW [4, 5]. Thus, for energies $w > 0$, which correspond to rotations of the model pendulum, the unknown ratio is δ^+/λ for forward motion and δ^-/λ for backward motion. For any value of the asymmetry parameter the absolute values of these quantities increase without bound if $\lambda \longrightarrow +\infty$. For energies $w < 0$, which correspond to oscillations of the model pendulum, a rougher estimate (since the forward and backward motion alternates) of the unknown ratio is δ^+/λ for forward motion and δ^-W^-/δ^+W^+ for backward motion (without loss of generality it is implied here that the parameter η is contained within the limits $0 \le \eta \le 1$).

Thus, in the case of a high-frequency perturbation the chaotic layer is subjected to strong bending in the sense that during motion near the separatrix the amplitude of the energy deviations relative to the unperturbed separatrix value is much greater than the layer width.

In accordance with Eq. (18), as a result of synchronization the line of the unperturbed separatrix $w = 0$ is transformed into the curve $\tilde{w} = \delta^\pm W^\pm \cos\tau$. If the chaotic layer is relatively thin ($\lambda W^\pm \ll \delta^\pm W^\pm$), in natural polar coordinates ρ, τ (where $\rho \equiv \tilde{w} + 2$, i.e., the zero point of the relative pendulum energy is taken to be its rest state in the lower equilibrium position) the line of the layer is a Pascal limaçon: $\rho = 2 + \delta^\pm W^\pm \cos\tau$.

For the Hamiltonian (22) the conclusions on the nature of the layer bending in the high-frequency perturbation limit are qualitatively similar.

The case of a low-frequency perturbation is not considered here because no reliable estimates of the layer width are available for it.

From these approximate formulas for the ratio of the bending amplitude of the chaotic layer to its half-width, it follows that when the perturbation has the maximum possible asymmetry ($\eta = 0$) the layer bending effect

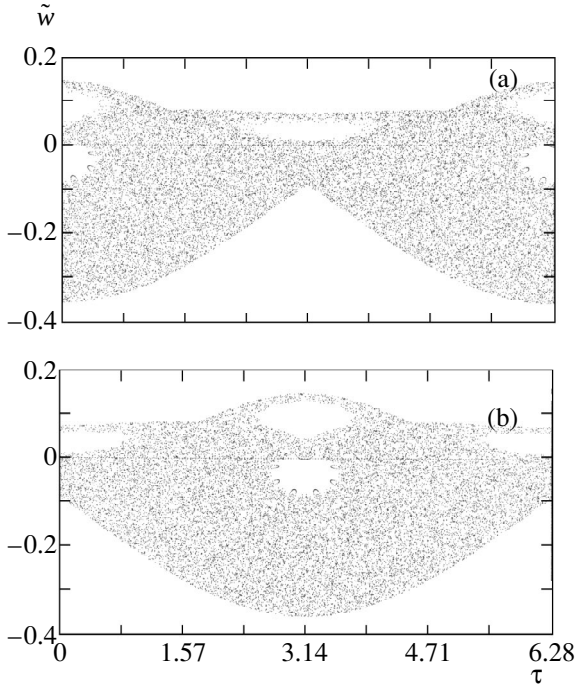

Fig. 6. Phase plane of synchronized separatrix algorithmic mapping (24), (18) for the Hamiltonian (22) with the parameters $\lambda = 0.5$, $\varepsilon = 0.02$; (a) forward motion, (b) backward motion as defined in text.

begins to become appreciable (the amplitude becomes greater than the width) for the forward and backward motion if $\lambda \approx 8$. Thus, this last value was selected to construct the phase portraits of the separatrix algorithmic mapping.

Figure 7 shows the behavior of the system (1) with the parameters $\lambda \approx 8$, $\varepsilon = 1$, $\eta = 0$, $k = 1$ using the representation of the synchronized separatrix algorithmic mapping (4), (18) with the corresponding parameters $\lambda = 8$, $W^+ = 0.00561$, $W^- = 6.82 \times 10^{-14}$, $\delta^+ = -62.3$, $\delta^- = 2.96 \times 10^{12}$ [Eqs. (12)–(15)]. The results of finding the cross section $\varphi = 0$ by direct numerical integration are similar to those obtained using the synchronized separatrix algorithmic mapping (the accuracy of the agreement along the $\tilde{w}$ axis is better than 0.04, and the accuracy of the agreement between the layer bending amplitudes is better than 0.01) and are not given here.

Thus, the bending effect is in fact present and corresponds to the predicted value for the forward and backward motion. The theoretical bending amplitudes for a thin chaotic layer ($\delta^+W^+ = -0.35$, $\delta^-W^- = 0.20$) show good agreement with those observed (with an accuracy better than 0.01).

In this last example the perturbation parameter is large $\varepsilon = 1$, but nevertheless the synchronized separatrix algorithmic mapping very accurately represents the real motion. We shall now explain why this synchronized separatrix algorithmic mapping is so efficient.

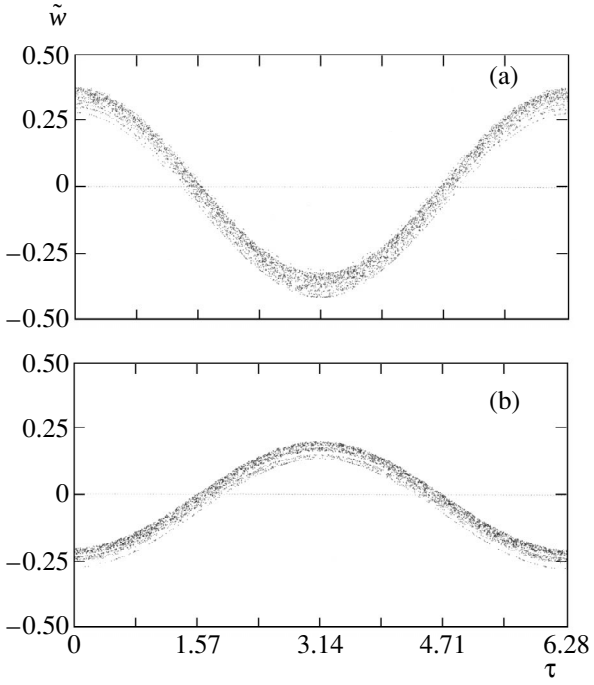

Fig. 7. Phase plane of synchronized separatrix algorithmic mapping (4), (18) for the Hamiltonian (1) with the parameters $\lambda = 8$, $\varepsilon = 1$, $\eta = 0$, $k = 1$; (a) forward motion, (b) backward motion.

The first of the two terms in the final expression (18) for the procedure for reducing the phase pattern of the motion to the unified surface of the cross section $\varphi = 0$ is the average of two successive values of w. The structure of the formula thus resembles taking the average (first term) with a certain correction (second term describing the layer bending effect). This division is formal. No approximation is used here. Both terms appear in the course of the exact derivation of the synchronization procedure.

The synchronized separatrix algorithmic mapping is free from any averaging effects. It may seem that these may appear and influence the accuracy of the mapping and the synchronization procedure since oscillating terms of amplitude of order ε are neglected in the numerical determination of the MA integrals (see Appendix). In fact, the primitives of the MA integrals A_n and B_n for any n oscillate at infinity with nonzero amplitude. We shall first consider the case of the Hamiltonian (1) and the definition of the relative energy in terms of H_0: $w = H_0/\mathcal{F} - 1$. As we can easily see, the initial expression for the increment H_0 [see (9)] contains the MA integrals in the form of sums $A_{2k-1} + A_{2k+1}$ and $B_{2k-1} + B_{2k+1}$. The primitives of the MA integrals in these sums oscillate at infinity in antiphase and therefore compensate for each other. (The absence of oscillations can be clearly seen from the presence of the cofactor $p(t)$ in the initial integrand expression; this suppresses oscillations at infinity.) The final expressions for the parameters of the separatrix algorithmic mapping and the synchronization (12)–(15) were obtained using recurrence formulas (A.10) and (A.11) by means of which these sums are transformed in the MA integrals A_{2k} and B_{2k}. The primitives of the last integrals oscillate at infinity. This "acquisition" of oscillations in the final expression is essentially a mathematical artifact which occurs as a result of using the recurrence formulas (A.10) and (A.11) which are derived using averaging. These oscillations have no relation to the accuracy of the nonsynchronized separatrix algorithmic mapping nor to the accuracy of the synchronization procedure.

When the relative energy is defined in terms of H ($w = H/\mathcal{F} - 1$) the oscillations of the energy increment in the nonsynchronized separatrix algorithmic mapping are real [since there is no cofactor $p(t)$ in the integrand expression in the initial formula for the increment]. However, in the synchronized separatrix algorithmic mapping these oscillations compensate for each other, as we can easily see, when the increments of H are summed on intervals between τ_{n-1} and $+\infty$ and between $-\infty$ and τ_n.

Finally, for the Hamiltonian (22) the situation with the oscillations of the energy increment is quite similar since it is determined by the presence or absence of the cofactor $p(t)$ in the initial integrand expressions. The apparent difference is that the integrals in the final formulas for the parameters of the separatrix algorithmic mapping and the synchronization (23) (specifically, the integrals α_1^c and β_1^s) are proper although in reality, when the relative energy is defined in terms of H, oscillations of the energy increments do occur in the nonsynchronized separatrix algorithmic mapping. The "loss" of oscillations here like their "acquisition" in the case of the Hamiltonian (1) in the previous example can be explained by the fact that the MA integrals were subjected to transformations using averaging.

On the whole, no real oscillations of the energy increments occur in the case of a synchronized separatrix algorithmic mapping and this mapping is free from averaging effects.

For perturbations $\varepsilon \sim \lambda^{-1}$ and above the standard Poincaré–Melnikov method of calculating exponentially small effects associated with separatrix splitting generally requires corrections [27, 28]. In the two examples given above (Figs. 5 and 7) the perturbation is large ($\varepsilon = 1$). In the last case, however, the perturbation is completely asymmetric ($\eta = 0$) and for this reason (see [27]) the correction is zero. For a perturbation of arbitrary asymmetry the correction factor for the system (1) according to the Simó [27] hypothetical formula is

$$|f(x)| = \left|\frac{\sinh x}{x}\right|,$$

where

$$x \equiv \sqrt{2\varepsilon_1\varepsilon_2} = \frac{\sqrt{2ab}}{\mathcal{F}_*}$$

(here x may be either a real or an imaginary number depending on the signs of a and b). In the case of Fig. 5, the correction factor to the separatrix mapping parameter W is $f(\sqrt{2}) = 1.368$.

This correction merely influences the layer width but not its bending amplitude since the bending has no relation to the separatrix splitting effect (in fact, the product δW contains no MA integrals of the first kind), i.e., this correction should be made to W and not to δW. In the case shown in Fig. 5 the layer width is very small compared with the bending amplitude and allowance for the correction negligibly influences the general form of the phase portrait.

Thus, our comparisons of the cross sections of the phase space obtained using the separatrix algorithmic mapping and direct numerical integration show that the synchronized separatrix algorithmic mapping correctly represents the near-separatrix motion both at high and low values of the relative perturbation frequency λ.

7. CONCLUSIONS

In the present article we have derived a procedure for synchronizing a separatrix algorithmic mapping describing motion near the separatrix of a nonlinear resonance in the presence of an asymmetric perturbation [20, 24] to the unified cross section surface of an initial Hamiltonian system. This procedure is valid for the particular case of a symmetric perturbation, i.e., for an ordinary separatrix mapping. This choice of cross-section surface gives a complete description of the phase space of the motion near the separatrix. It is found that in the limit of a high perturbation frequency the principal chaotic layer undergoes strong bending: during motion near the separatrix the amplitude of the energy deviations relative to the unperturbed separatrix value is greater than the layer width (their ratio tends to infinity for $\lambda \longrightarrow +\infty$). However, a comparison with the results of direct numerical integration shows that the synchronized separatrix algorithmic mapping gives (both at high and low perturbation frequencies) a correct representation of the phase portrait of the motion near the separatrix if the perturbation amplitude is fairly small.

To conclude we note that the synchronized separatrix algorithmic mapping describes motion of the system on the "phase perturbation angle—energy deviation" plane for a fixed resonant phase angle $\varphi = 0$. When the Poincaré cross sections are constructed numerically in applied problems, a different plane is usually used, i.e., the "resonant phase angle—momentum" plane for a fixed perturbation phase angle, for example $0 \pmod{2\pi}$. Can we construct a second type of cross section using separatrix mapping? To quite a good approximation, this is in fact possible. The procedure is again an algorithm containing conditional transfer instructions. This algorithm was described in [20, 24].

ACKNOWLEDGMENTS

I am deeply grateful to B.V. Chirikov and V.V. Vecheslavov for useful discussions.

APPENDIX

Melnikov–Arnold Integrals

In the Appendices to [4, 5] a detailed description is given of functions of the type

$$A_n(\lambda) = \int_{-\infty}^{\infty} \cos\left(\frac{n}{2}\phi(t) - \lambda t\right) dt, \qquad \text{(A.1)}$$

where

$$\phi(t) = 4\arctan\exp(t) - \pi \qquad \text{(A.2)}$$

(separatrix solution of the unperturbed equation for a nonlinear pendulum), the parameter λ is any real number, and n is a real parameter $n \geq 0$. Everywhere subsequently we confine ourselves to the case of integer values of $n \geq 0$. Following the terminology adopted in [4, 5] the functions (A.1) will be called Melnikov–Arnold integrals (MA integrals). They can be written in the following form:

$$A_n(\lambda) = 2\int_0^{\infty} \cos[n\arctan\sinh(t) - \lambda t]\, dt. \qquad \text{(A.3)}$$

In order to synchronize the separatrix mapping we need to introduce the related functions

$$B_n(\lambda) = \int_0^{\infty} \sin[n\arctan\sinh(t) - \lambda t]\, dt, \qquad \text{(A.4)}$$

which we shall call MA integrals of the second kind. (The functions A_n are therefore MA integrals of the first kind.) We also introduce the auxiliary functions

$$\begin{aligned} \alpha_n^c(\lambda) &= A_n(\lambda) + A_n(-\lambda), \\ \alpha_n^s(\lambda) &= A_n(\lambda) - A_n(-\lambda), \\ \beta_n^c(\lambda) &= B_n(\lambda) + B_n(-\lambda), \\ \beta_n^s(\lambda) &= B_n(\lambda) - B_n(-\lambda). \end{aligned} \qquad \text{(A.5)}$$

The upper indices c and s indicate that the expression in the integrand contains the cosine or sine of λt respectively according to the following equivalent formulas:

$$\alpha_n^c(\lambda) = 4\int_0^\infty \cos[n\,\mathrm{arctan\,sinh}(t)]\cos(\lambda t)\,dt, \quad \text{(A.6)}$$

$$\alpha_n^s(\lambda) = 4\int_0^\infty \sin[n\,\mathrm{arctan\,sinh}(t)]\sin(\lambda t)\,dt, \quad \text{(A.7)}$$

$$\beta_n^c(\lambda) = 2\int_0^\infty \sin[n\,\mathrm{arctan\,sinh}(t)]\cos(\lambda t)\,dt, \quad \text{(A.8)}$$

$$\beta_n^s(\lambda) = -2\int_0^\infty \cos[n\,\mathrm{arctan\,sinh}(t)]\sin(\lambda t)\,dt. \quad \text{(A.9)}$$

The primitives in the integrands in the integral representations for A_n, B_n, α_n^c, α_n^s, β_n^c, β_n^s as functions of the independent variable t oscillate for $t \longrightarrow +\infty$ and the amplitude of the oscillations may be nonzero in the limit depending on the parity of the parameter n. Thus, determining the numerical values of the MA integrals should follow a specific formal agreement. According to the traditional natural agreement for A_n [4, 5] the numerical values of all these MA integrals are taken to be averages in the limit when the upper limit of integration for their integral representations tends to infinity. For example, we therefore assume here that $A_0 = 0$.

In [4, 5] the MA integrals of the first kind with any real values of n were estimated analytically by means of residue theory. If we confine ourselves to natural values of n, the MA integrals of the second kind of interest to us and also those of the first kind are most easily determined by induction. We shall first consider functions of the first kind. The induction for A_n in terms of A_{n-1} and A_{n-2} for any natural $n \geq 2$ is described by the Zhirov recurrence formula (see Appendices in [4, 5]). Thus, if A_0 and A_1 are known, the expression for A_n is obtained by induction:

$$\begin{aligned} A_0(\lambda) &= 0, \\ A_1(\lambda) &= \frac{\alpha_1^c(\lambda) + \alpha_1^s(\lambda)}{2}, \\ &\cdots \\ A_n(\lambda) &= \frac{2\lambda}{n-1}A_{n-1}(\lambda) - A_{n-2}(\lambda). \end{aligned} \quad \text{(A.10)}$$

Deriving the recurrence formula for B_n similar to the derivation for A_n (by integrating by parts) for the functions of the second kind we have similar formulas:

$$\begin{aligned} B_0(\lambda) &= -\frac{1}{\lambda}, \\ B_1(\lambda) &= \frac{\beta_1^c(\lambda) + \beta_1^s(\lambda)}{2}, \\ &\cdots \\ B_n(\lambda) &= \frac{2\lambda}{n-1}B_{n-1}(\lambda) - B_{n-2}(\lambda) + \frac{2}{n-1}. \end{aligned} \quad \text{(A.11)}$$

The induction bases in Eqs. (A.10), (A.11) are A_0, B_0, A_1, and B_1. The last two functions are expressed in terms of α_1^c, α_1^s, β_1^c, and β_1^s. Using tabular integrals from [29] or [30], we have the following expressions for these four functions:

$$\alpha_1^c(\lambda) = \frac{2\pi}{\cosh(\pi\lambda/2)}, \quad \text{(A.12)}$$

$$\alpha_1^s(\lambda) = \frac{2\pi}{\sinh(\pi\lambda/2)}, \quad \text{(A.13)}$$

$$\beta_1^c(\lambda) = 2\mathrm{Re}\left[\psi\left(i\frac{\lambda}{2}\right) - \psi\left(i\frac{\lambda}{4}\right)\right] - 2\ln 2, \quad \text{(A.14)}$$

$$\beta_1^s(\lambda) = \pi\tanh\frac{\pi\lambda}{2} - 2\mathrm{Im}\,\psi\left(\frac{1+i\lambda}{4}\right). \quad \text{(A.15)}$$

In these expressions $\psi(z) = \Gamma'(z)/\Gamma(z)$ is the digamma function and i is the imaginary unit. The real part of the digamma function of imaginary argument is given by the series [31]:

$$\mathrm{Re}\,\psi(iy) = -C + y^2\sum_{n=1}^\infty \frac{1}{n(n^2+y^2)}, \quad \text{(A.16)}$$

where y is any real number, $C \approx 0.577216$ is the Euler constant.

Formulas (A.11) together with Eqs. (A.14) and (A.15) express the MA integrals of the second kind of any natural order n in terms of special functions.

How can the MA integrals be estimated with prescribed accuracy by numerical integration, for example, to check the formulas given above? In addition to selecting a suitable quadrature scheme to improve the accuracy we also need to increase the upper limit of numerical integration. A direct monotonic increase would lead to oscillations of the results in some cases because of the nature of the MA integrals described above. This can be avoided by selecting the upper limit of integration successively equal to the values of t for which the oscillating component of the primitive (selected assuming $\lambda t \gg 2\pi$) goes to zero. This reasoning gives values for the upper limits of numerical integration and these are given in the table.

In Table 1 the number M implies any sufficiently large natural number (the values of M =5–6 are usually

Upper limits of numerical integration to estimate Melnikov–Arnold integrals

n	$A_n(\lambda)$	$\alpha_n^c(\lambda)$	$\alpha_n^s(\lambda)$	$B_n(\lambda)$	$\beta_n^c(\lambda)$	$\beta_n^s(\lambda)$
Odd	$\frac{\pi(M+1/2)}{\lvert\lambda\rvert}$	*	$\frac{\pi(M+1/2)}{\lvert\lambda\rvert}$	$\frac{\pi M}{\lvert\lambda\rvert}$	$\frac{\pi M}{\lvert\lambda\rvert}$	*
Even	$\frac{\pi M}{\lvert\lambda\rvert}$	$\frac{\pi M}{\lvert\lambda\rvert}$	*	$\frac{\pi(M+1/2)}{\lvert\lambda\rvert}$	*	$\frac{\pi(M+1/2)}{\lvert\lambda\rvert}$

sufficient to ensure four significant digits in the result); the asterisk indicates that the upper limit of integration can be any fairly large positive real number (since oscillations of nonzero amplitude at infinity do not occur in the corresponding cases).

REFERENCES

1. N. N. Filonenko, R. Z. Sagdeev, and G. M. Zaslavsky, Nucl. Fusion **7**, 253 (1967).
2. G. M. Zaslavskiĭ and N. N. Filonenko, Zh. Éksp. Teor. Fiz. **54**, 1590 (1968) [Sov. Phys. JETP **27**, 851 (1968)].
3. G. M. Zaslavskiĭ and B. V. Chirikov, Usp. Fiz. Nauk **105**, 3 (1972) [Sov. Phys. Usp. **14**, 549 (1972)].
4. B. V. Chirikov, *Nonlinear Resonance* (Novosib. Gos. Univ., Novosibirsk, 1977).
5. B. V. Chirikov, Phys. Rep. **52**, 263 (1979).
6. A. J. Lichtenberg and M. A. Lieberman, *Regular and Chaotic Dynamics* (Springer-Verlag, New York, 1992; Mir, Moscow, 1984).
7. V. V. Vecheslavov and B. V. Chirikov, Zh. Éksp. Teor. Fiz. **114**, 1516 (1998) [JETP **87**, 823 (1998)].
8. V. V. Vecheslavov, Zh. Éksp. Teor. Fiz. **116**, 336 (1999) [JETP **89**, 182 (1999)].
9. I. I. Shevchenko, Phys. Scr. **57**, 185 (1998).
10. S. S. Abdullaev and G. M. Zaslavsky, Phys. Plasmas **2**, 4533 (1995).
11. S. S. Abdullaev and G. M. Zaslavsky, Phys. Plasmas **3**, 516 (1996).
12. R. Malhotra, Icarus **87**, 249 (1990).
13. J. Wisdom, S. J. Peale, and F. Mignard, Icarus **58**, 137 (1984).
14. A. Celletti, Z. Angew. Math. Phys. **41**, 174 (1990).
15. B. S. Bardin and A. P. Markeev, Prikl. Mat. Mekh. **59**, 922 (1995).
16. D. F. Escande and F. Doveil, J. Stat. Phys. **26**, 257 (1981).
17. D. F. Escande, Phys. Rep. **121**, 165 (1985).
18. G. M. Zaslavskiĭ, R. Z. Sagdeev, D. A. Usikov, *et al.*, *Weak Chaos and Quisi-Regular Structures* (Nauka, Moscow, 1991).
19. G. M. Zaslavsky and S. S. Abdullaev, Phys. Rev. E **51**, 3901 (1995).
20. I. I. Shevchenko, in *The Chaotic Universe*, Ed. by V. G. Gurzadyan and R. Ruffini (World Scientific, London, 2000), p. 599.
21. B. V. Chirikov, Preprint No. 90-109, INF (Institute of Nuclear Physics, Novosibirsk, 1990).
22. T. Ahn, G. Kim, and S. Kim, Physica D (Amsterdam) **89**, 315 (1996).
23. D. Jeon, M. Bai, C. M. Chu, *et al.*, Phys. Rev. E **54**, 4192 (1996).
24. I. I. Shevchenko, Celest. Mech. Dyn. Astron. **73**, 259 (1999).
25. E. Hairer, S. Nørsett, and G. Wanner, *Solving Ordinary Differential Equations*, I: *Nonstiff Problems* (Springer-Verlag, Heidelberg, 1987; Mir, Moscow, 1990).
26. P. Holmes, J. Marsden, and J. Scheurle, Contemp. Math. **81**, 213 (1988).
27. V. G. Gelfreich, Nonlinearity **10**, 175 (1997).
28. D. V. Treshchev, *An Introduction to the Perturbation Theory of Hamiltonian Systems* (FAZIS, Moscow, 1998).
29. I. S. Gradshteyn and I. M. Ryzhik, *Table of Integrals, Series, and Products* (Fizmatgiz, Moscow, 1962; Academic, New York, 1980).
30. *Tables of Integral Transforms (Bateman Manuscript Project)*, Ed. by A. Erdelyi (McGraw-Hill, New York, 1954; Nauka, Moscow, 1969), Vol. 1.
31. *Handbook of Mathematical Functions*, Ed. by M. Abramowitz and I. A. Stegun (Dover, New York, 1971; Nauka, Moscow, 1979).

Translation was provided by AIP

Journal of Experimental and Theoretical Physics, Vol. 91, No. 3, 2000, pp. 626–634.
Translated from Zhurnal Éksperimental'noĭ i Teoreticheskoĭ Fiziki, Vol. 118, No. 3, 2000, pp. 720–729.

NONLINEAR PHYSICS

Stability of Solitons in Nonlinear Composite Media

A. T. Il'ichev

Steklov Institute of Mathematics, Russian Academy of Sciences, Moscow, 117966 Russia
e-mail: ilichev@orc.ru, ilichev@mi.ras.ru

Received April 4, 2000

Abstract—An analysis is made of the dynamic stability of soliton solutions of the Hamilton equations describing plane waves in nonlinear elastic composite media in the presence and absence of anisotropy. In the anisotropic case two two-parameter soliton families, fast and slow, are obtained in analytic form; in the absence of anisotropy there is a single three-parameter soliton family. It is shown that solitons from the slow family in an anisotropic composite and solitons in an isotropic composite are dynamically stable if their velocities lie in a certain range known as the range of stability. The analysis of stability is based on the spectral properties of the "linearized Hamiltonian" $\mathscr{H}$. It is shown that the operator $\mathscr{H}$ is positively semidefinite on some linear subspace of the main solution space from which stability follows. Problems of instability of the fast soliton family in the anisotropic case and representatives of soliton families whose velocities lie outside the range of stability in the presence and absence of anisotropy are discussed. © *2000 MAIK "Nauka/Interperiodica".*

1. INTRODUCTION

Recently so-called composite materials have continued to attract interest in the theory of elasticity and its applications. As is indicated by their name, these materials consist of various substances capable of mixing at the macrolevel. A method of averaging the fundamental equations of the model (see, for example, [1]) is used to study large-scale processes having a characteristic scale substantially exceeding the inhomogeneity scale. In this case, it is found that the properties of a composite described by averaged equations reveal substantial differences compared with those of the component materials. In particular, a typical case is the appearance of dispersion despite the fact that no dispersion occurs in any of the elastic materials forming the composite [1]. Consequently, a composite consisting of elastic materials with a nonlinear equation of state is a dispersive medium in which waves resulting from the interaction of nonlinear and dispersion effects, including solitons, can propagate. The possibility of observing solitons in practice is naturally related to the dynamic stability of these waves.

The literature relating to the study of soliton stability by modern methods in various problems in nonlinear physics is fairly extensive. Without claiming to give any kind of complete review of this topic, we shall confine ourselves to mentioning various studies in this field. An analysis of orbital stability based on a geometric approach to the study of Hamiltonian systems was first developed in [2, 3] where the authors analyzed the nonlinear stability of solitons of the Korteweg–de Vries equation and the alternative Benjamin–Bona–Mahony equation. The application of similar methods to study soliton solutions of various other model equations can be found in [4] (Benjamin–Ono equation), [5] (nonlinear Klein–Gordon equation and nonlinear Schrödinger equation), [6] (so-called equation for moderate wavelengths), [7] (generalized Boussinesq equation), [8, 9] (Kawahara equation), [10] (family of generalized Korteweg–de Vries equations), and [11] (asymptotic stability of solitons of Korteweg–de Vries equations). A discussion of general aspects of the stability and instability of solitons in hydromechanics and plasma is reported in [12, 13].

Soliton structures in elastic media are found, for example, in geometrically nonlinear media where dispersion occurs as a result of stretching and bending effects (see, for example, [14]). The only results on the dynamic stability of solitons in elastic media known to the author are results on the stability of solitons in nonlinear elastic strings [7] and plane loop solitons in elastic inextensible thin rods [15,16].

In the present study we investigate the dynamic stability of solitons in nonlinear elastic composites in the presence and absence of anisotropy. For our study of stability we follow a method which was definitively developed in [17]. Dynamic (orbital) stability of boundary states (solitons) in Hamiltonian systems having symmetries automatically occurs if the "linearized Hamiltonian" [defined as $\mathscr{H}$ by Eq. (2.7) in the present article] is nonnegative over the entire functional space of the solutions of the system. However, for the general position, the operator $\mathscr{H}$ has at least one negative eigenvalue. If this value is unique and simple, soliton stability will also occur for specific properties of the soliton solution in translationally invariant Hamiltonian systems [17].

The present article is organized as follows. In Section 2 we give the basic equations describing plane

waves in nonlinear anisotropic elastic composites, we present the Hamiltonian form of the basic equations, we consider symmetries in the anisotropic and isotropic cases, give the conserved quantities, and obtain soliton solutions. Two families of solitons are obtained in the anisotropic case. These families are parametrized by the propagation velocity and shear. In the absence of anisotropy there is a single family which depends on an additional parameter, the angle of rotation in the wave plane. Section 3 is devoted to the spectral properties of the operator $\mathcal{H}$. It is shown that for certain ranges of velocities of the slow soliton family in an anisotropic medium and for the soliton family in the absence of anisotropy (the so-called range of stability), the operator $\mathcal{H}$ is positive definite on a certain closed linear subspace of the entire functional space of the solutions of the basic equations. This implies orbital stability (dynamic shape stability) of these soliton families for these ranges of velocities. Evidence of stability is provided in Section 4. In Section 5 we give a brief formulation of the results and discuss problems of instability of the fast soliton family in the anisotropic case and representatives of soliton families whose velocities lie outside the range of stability in the presence and absence of anisotropy.

2. FORMULATION OF THE PROBLEM

2.1. Basic Equations

We shall consider plane wave motion in an inhomogeneous nonlinear elastic medium (composite) when the displacements w_α, strains $u_\alpha = \partial w_\alpha/\partial x$, and particle velocities v_α, $\alpha = 1, 2, 3$ depend on a single spatial variable, the Cartesian coordinate $x = x_3$, and the time t. We are interested in incompressible elastic media when u_3 and v_3 are constant. These constants may be set equal to zero without limiting the generality.

Although the motion of a nonlinear elastic body is described by a hyperbolic system of equations [18], the existence of an internal inhomogeneous material structure at the macrolevel leads to wave dispersion [1]. For an elastic medium we shall assume that the nonlinearity, anisotropy, and dispersion are small and represented by terms of the same order of magnitude. The system of basic equations can then be written in the form [19]

$$\frac{\partial u_i}{\partial t} - \frac{\partial v_i}{\partial x} = 0,$$
$$\rho_0 \frac{\partial v_i}{\partial t} - \frac{\partial}{\partial x}\left(\frac{\partial \Phi}{\partial u_i}\right) + m\frac{\partial^2 u_i}{\partial x^3} = 0, \qquad (2.1)$$
$$i = 1, 2.$$

Here ρ_0 is the average density of the material, Φ is the elastic potential which is given by

$$\Phi = \frac{1}{2} f(u_1^2 + u_2^2)$$
$$+ \frac{1}{2} g(u_2^2 - u_1^2) - \frac{1}{4}\kappa(u_1^2 + u_2^2)^2.$$

The constants $g > 0$ and κ characterize the anisotropy and nonlinearity, respectively. An expression for the constants f, g, and κ is given in [18]. A dispersion term with $m > 0$ appears in the equations of motion [second pair of equations in (2.1)], for example in cases when a homogeneous, elastic, easily deformable medium contains uniformly distributed rods having sufficient bending rigidity and positioned parallel to the x axis [19].

The equations (2.1) may be written in the Hamiltonian form:

$$\frac{\partial}{\partial t}\mathbf{w} = \mathcal{J}E'(\mathbf{w}),$$
$$E = \frac{1}{2}\int_{-\infty}^{\infty}\Big[v_1^2 + v_2^2 + \mu_1 u_1^2 + \mu_2 u_2^2 \qquad (2.2)$$
$$-\frac{\kappa}{2\rho_0}(u_1^2 + u_2^2)^2 + \frac{m}{\rho_0}(\partial_x u_1)^2 + \frac{m}{\rho_0}(\partial_x u_2)^2\Big]dx,$$

where

$$\mu_1 = (f - g)/\rho_0, \quad \mu_2 = (f + g)/\rho_0,$$

$\mathbf{w} = \{u_1, u_2, v_1, v_2\}^t$ is an unknown vector function, the prime denotes the variational derivative $\delta/\delta w = \{\delta/\delta u_1, \delta/\delta u_2, \delta/\delta v_1, \delta/\delta v_2\}^t$, and $\mathcal{J}$ is a skew-symmetric operator:

$$\mathcal{J} = \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix}\frac{\partial}{\partial x}.$$

2.2. Conserved Quantities and Symmetries

Quite clearly, the Hamiltonian E is constant on account of the system (2.2). In addition, it is easy to see that the functional

$$Q = \int_{-\infty}^{\infty}[u_1 v_1 + u_2 v_2]dx$$

is also an invariant. Another formally conserved quantity is the vector functional

$$\mathbf{A} = \int_{-\infty}^{\infty}\mathbf{w}dx.$$

The system of Eqs. (2.2) has additional conserved quantities in the particular case of degeneracy of the anisotropy. We shall subsequently make separate analyses of the anisotropic $g \neq 0$ and isotropic $g = 0$ cases.

2.2.1. Anisotropic case. Equations (2.1) and (2.2), and also the functions E, Q, and $\mathbf{A}$ are invariant with respect to the group of translations

$$T(s)\mathbf{w} = \mathbf{w}(x+s) = \exp(s\partial_x)\mathbf{w}(x), \quad s \in \mathbb{R}.$$

The functional Q is a conserved quantity as a result of the translational invariance (2.1).

2.2.2. Isotropic case. In this case we have the additional rotational symmetry

$$G(\varphi)\mathbf{w} = \exp(\mathscr{A}\varphi)\mathbf{w}, \quad \varphi \in \mathbb{S}^1,$$

where $\mathbb{S}^1$ denotes a circle and the matrix $\mathscr{A} = \mathrm{diag}(a, a)$ is a partitioned diagonal matrix with the blocks

$$a = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}.$$

As a result of the rotational symmetry, the following quantity is formally conserved

$$U = \frac{1}{2}\int_{-\infty}^{\infty} [y_1 v_2 - y_2 v_1 + u_2 w_1 - u_1 w_2]\,dx,$$
$$\partial_x y_i = u_i, \quad \partial_x w_i = v_i.$$

The Hamiltonian E and functional Q are also invariant with respect to the rotation group.

2.3. Soliton Solutions

The soliton solutions of the system (2.1) are traveling waves which decay rapidly at infinity. After substituting into (2.1) $\mathbf{w} = \mathbf{w}(\xi)$, where $\xi = \xi - Vt$ (V is the constant wave propagation velocity) and integrating once, using the conditions of decay at infinity we obtain

$$v_i = -Vu_i,$$
$$\frac{m}{\rho_0}\ddot{u}_i = (\mu_i - V^2)u_i - \frac{\kappa}{\rho_0}u_i(u_1^2 + u_2^2). \tag{2.3}$$

The dots here denote differentiation with respect to the variable ξ. The Eq. (2.3) for isolated waves are written in the equivalent form:

$$E'(\phi_V) + VQ'(\phi_V) = 0. \tag{2.4}$$

In equation (2.4) the vector function $\phi_V = \{u_1^c, u_2^c, v_1^c, v_2^c\}^T$, $v_i^c = -Vu_i^c$ denotes soliton solutions (2.3).

For $\mu_i > 0$, $\kappa > 0$ Eqs. (2.3) have soliton solutions. As in the previous paragraph we consider the anisotropic and isotropic cases separately.

2.3.1. Solitons in the presence of anisotropy. In this case, Eqs. (2.3) have two families of soliton solutions:

$$u_1^c = \pm\sqrt{2\rho_0\kappa^{-1}(\mu_1 - V^2)}\cosh^{-1}\sqrt{2\rho_0 m^{-1}(\mu_1 - V^2)}\xi,$$
$$u_2^c = 0, \quad \text{at} \quad 0 < V^2 < \mu_1,$$
$$u_1^c = 0, \tag{2.5}$$
$$u_2^c = \pm\sqrt{2\rho_0\kappa^{-1}(\mu_2 - V^2)}\cosh^{-1}\sqrt{2\rho_0 m^{-1}(\mu_2 - V^2)}\xi,$$
$$\text{at} \quad 0 < V^2 < \mu_2.$$

Each of the families (2.5) is a two-parameter one, the parameters being the velocity V and the shear s along the ξ coordinate. We shall subsequently call the first family of solitons in (2.5) slow since these exist in the lower velocity range and we shall call the second fast.

2.3.2. Solitons in an isotropic material. In the absence of anisotropy $\mu_1 = \mu_2 = \mu$ we have additional rotational symmetry: if $\phi_V = \{u_1^c, u_2^c, v_1^c, v_2^c\}^T$ is a soliton solution of Eqs. (2.3), then $\exp(\mathscr{A}\varphi)\phi_V$, $\varphi \in \mathbb{S}^1$ is also a soliton solution. Thus, it is sufficient to consider only the specific case with fixed φ:

$$u_1^c = \pm\sqrt{2\rho_0\kappa^{-1}(\mu - V^2)}\cosh^{-1}\sqrt{2\rho_0 m^{-1}(\mu - V^2)}\xi,$$
$$u_2^c = 0, \quad \text{at} \quad 0 < V^2 < \mu. \tag{2.6}$$

The family of solitons represented by Eq. (2.6) is a three-parameter one, the third parameter being the angle of rotation φ.

By virtue of Eq. (2.4), the behavior of the functional $E(\mathbf{w}) = VQ(\mathbf{w})$ near $\mathbf{w} = \phi_V$ is completely determined by the spectral properties of the self-conjugate operator

$$\mathscr{H} = E''(\phi_V) + VQ''(\phi_V). \tag{2.7}$$

We shall discuss these properties in the following section.

3. SPECTRAL PROPERTIES OF THE OPERATOR $\mathscr{H}$

The operator $\mathscr{H}$ has the form

$$\mathscr{H} = \begin{pmatrix} H_1 & 0 & V & 0 \\ 0 & H_2 & 0 & V \\ V & 0 & 1 & 0 \\ 0 & V & 0 & 1 \end{pmatrix},$$
$$H_1 = \mu_1 - \frac{3\kappa}{\rho_0}u_1^{c2} - \frac{\kappa}{\rho_0}u_2^{c2} - \frac{m}{\rho_0}\frac{d^2}{d\xi^2}, \tag{3.1}$$
$$H_2 = \mu_2 - \frac{3\kappa}{\rho_0}u_2^{c2} - \frac{\kappa}{\rho_0}u_1^{c2} - \frac{m}{\rho_0}\frac{d^2}{d\xi^2}.$$

It follows from Eq. (3.1) that the eigenvalue problem

$$\mathcal{H}\bar{\chi} = \lambda\bar{\chi}, \quad \bar{\chi} = \{\chi_1, \chi_2, \chi_3, \chi_4\}^T,$$

reduces to two independent problems:

$$\begin{aligned}(H_1 - V^2)\chi_1 &= \left(\lambda - \frac{V^2\lambda}{\lambda - 1}\right)\chi_1,\\ (H_2 - V^2)\chi_2 &= \left(\lambda - \frac{V^2\lambda}{\lambda - 1}\right)\chi_2,\end{aligned} \tag{3.2}$$

and the third and fourth components $\bar{\chi}$ are determined using the formulas

$$\chi_{3,4} = V\chi_{1,2}/(\lambda - 1).$$

As before, we shall discuss the anisotropic and isotropic cases separately.

3.1. Spectrum in Anisotropic Case

For the slow soliton family ($u_2^c = 0$) the spectral problem (3.2) has the form

$$\begin{aligned}&\left(\mu_1 - V^2 - \frac{3\kappa u_1^{c2}}{\rho_0} - \frac{m}{\rho_0}\frac{d^2}{d\xi^2}\right)\chi_1 = \nu\chi_1,\\ &\left(\mu_1 - V^2 - \frac{\kappa u_1^{c2}}{\rho_0} - \frac{m}{\rho_0}\frac{d^2}{d\xi^2}\right)\chi_2 = \nu_1\chi_2,\\ &\nu = \left(\lambda - \frac{V^2\lambda}{\lambda - 1}\right), \quad \nu_1 = \left(\lambda - \frac{V^2\lambda}{\lambda - 1} - 2\frac{g}{\rho_0}\right).\end{aligned} \tag{3.3}$$

The differential operators on the left-hand sides of Eqs. (3.3) are Sturm–Liouville operators and the corresponding eigenvalue problems are investigated using general theory. In the first equation in (3.3) we have for $\nu = 0$ the antisymmetric eigenfunction $\chi_1 = \partial_\xi u_1^c$. This means that there is a single simple eigenvalue $\nu_- < 0$ and the positive spectrum ν is isolated from zero:

$$\nu \geq \nu_+ > 0.$$

For the lowest state ν_- we have

$$\lambda - \frac{V^2\lambda}{\lambda - 1} = \nu_-. \tag{3.4}$$

It follows from (3.4) that

$$\lambda_\pm = \frac{1 + V^2 + \nu_-}{2} \pm \frac{1}{2}\sqrt{(1 + V^2 + \nu)^2 - 4\nu_-},$$

whence $\lambda_- < 0$, $\lambda_+ > 0$. For the continuous spectrum λ of the first problem (3.3) we have

$$\lambda - \frac{V^2\lambda}{\lambda - 1} \geq \mu_1 - V^2 > 0.$$

From this we obtain

$$\lambda > \frac{1 + \mu_1}{2} - \sqrt{(1 - \mu_1)^2 + 4V^2} > 0,$$

i.e., the continuous spectrum λ is positive and nonzero.

The lowest state of the second spectral problem (3.3) is a simple eigenvalue $\nu_1 = 0$ with the symmetric eigenfunction u_1^c. The positive spectrum of ν_1 is nonzero. For the eigenvalues λ corresponding to $\nu_1 = 0$ we have

$$\begin{aligned}&\lambda - \frac{V^2\lambda}{\lambda - 1} - \frac{2g}{\rho_0} = 0,\\ &\lambda_\pm = \frac{1 + V^2 + 2g/\rho_0}{2}\\ &\pm \frac{1}{2}\sqrt{\left(\frac{1 + V^2 + 2g}{\rho_0}\right)^2 - \frac{8g}{\rho_0}} > 0.\end{aligned}$$

For the continuous spectrum similar to the previous case we obtain

$$\lambda > \frac{1 + \mu_2}{2} - \sqrt{(1 - \mu_2)^2 + 4V^2} > 0.$$

Thus, for the slow soliton family in the anisotropic case the spectrum of the operator $\mathcal{H}$ is organized as follows:

(i) there is exactly one simple negative eigenvalue;

(ii) the kernel of the operator $\mathcal{H}$ is one-dimensional and pulled onto the eigenvector $\partial_\xi\phi_V$;

(iii) the positive spectrum of $\mathcal{H}$ is nonzero.

We shall subsequently consider the spectrum of the operator $\mathcal{H}$ for the fast soliton family ($u_1^c = 0$). For this family the spectral problem (3.2) has the form

$$\begin{aligned}&\left(\mu_2 - V^2 - \frac{\kappa u_2^{c2}}{\rho_0} - \frac{m}{\rho_0}\frac{d^2}{d\xi^2}\right)\chi_1 = \nu_1\chi_1,\\ &\left(\mu_2 - V^2 - \frac{3\kappa u_2^{c2}}{\rho_0} - \frac{m}{\rho_0}\frac{d^2}{d\xi^2}\right)\chi_2 = \nu\chi_2,\\ &\nu = \left(\lambda - \frac{V^2\lambda}{\lambda - 1}\right), \quad \nu_1 = \left(\lambda - \frac{V^2\lambda}{\lambda - 1} + 2\frac{g}{\rho_0}\right).\end{aligned} \tag{3.5}$$

The lowest state of the first spectral problem in (3.5) is $\nu_1 = 0$ with the eigenfunction $\chi_1 = u_2^c$. The positive spectrum of ν_1 is isolated from zero. For the eigenvalues λ corresponding to $\nu_1 = 0$ we have

$$\lambda - \frac{V^2\lambda}{\lambda - 1} + 2\frac{g}{\rho_0} = 0,$$

$$\lambda_{\pm} = \frac{1 + V^2 - 2g/\rho_0}{2} \pm \frac{1}{2}\sqrt{\left(\frac{1 + V^2 - 2g}{\rho_0}\right)^2 + \frac{8g}{\rho_0}},$$

from which it follows that $\lambda_- < 0$, $\lambda_+ > 0$. The positive and nonzero nature of the continuous spectrum for the first spectral problem in (3.5) is demonstrated as in the previous cases.

The analysis of the second spectral problem is exactly the same as the first spectral problem for the slow family of solitons. As a result for the spectrum of the operator $\mathcal{H}$ for the fast family we have:

(i) two simple negative eigenvalues λ where the eigenvalue λ_- corresponds to the eigenfunction $\bar{\chi}_-^0 = \{u_2^c, 0, V^{-1}(1 + 2g\lambda_-^{-1})u_2^c, 0\}^T$;

(ii) a one-dimensional kernel pulled to the zero eigenvector $\partial_\xi \phi_V$;

(iii) a positive spectrum isolated from zero.

Thus, for the fast family there is an additional unstable direction $\bar{\chi}_-^0$ compared with the slow family. The existence of this unstable direction presumably leads to instability of the fast soliton family. The question of instability is discussed in greater detail in Section 5.

3.2. Spectrum in the Isotropic Case

In the isotropic case $\mu_1 = \mu_2 = \mu$ the spectral problem (3.2) degenerates:

$$\left(\mu - V^2 - \frac{3\kappa u_1^{c2}}{\rho_0} - \frac{m}{\rho_0}\frac{d^2}{d\xi^2}\right)\chi_1 = \nu\chi_1,$$

$$\left(\mu - V^2 - \frac{\kappa u_1^{c2}}{\rho_0} - \frac{m}{\rho_0}\frac{d^2}{d\xi^2}\right)\chi_2 = \nu\chi_2, \qquad (3.6)$$

$$\nu = \left(\lambda - \frac{V^2\lambda}{\lambda - 1}\right).$$

Studying (3.6) as in the previous case, we find that the first eigenvalue problem has a simple negative eigenvalue, a simple zero eigenvalue with the eigenfunction $\chi_1 = \partial_\xi u_1^c$ and a positive nonzero spectrum. The second problem in (3.6) gives a simple zero eigenvalue with the eigenfunction $\chi_2 = u_1^c$ and a positive nonzero spectrum. Summarizing we find that in the isotropic case the spectrum of the operator $\mathcal{H}$ consists of:

(i) a simple negative eigenvalue;

(ii) a double zero eigenvalue with the eigenfunctions $\partial_\xi \phi_V$ and $\mathcal{A}\phi_V$;

(iii) a positive part of the spectrum isolated from zero.

The existence of an additional eigenvalue compared with the anisotropic case is caused by the appearance of additional rotational symmetry.

We then introduce the notation: for $\mathbf{x} = \{x_1, x_2, x_3, x_4\}^T$, $\mathbf{z} = \{z_1, z_2, z_3, z_4\}^T$,

$$\langle \mathbf{x}, \mathbf{z} \rangle = \int_{-\infty}^{\infty} (x_1 z_1 + x_2 z_2 + x_3 z_3 + x_4 z_4) d\xi$$

and we consider the properties of the bilinear form $\langle \mathcal{H}\mathbf{y}, \mathbf{y} \rangle$ for the slow soliton family in the anisotropic case and for the isotropic case.

3.3. Properties of the Bilinear Form $\langle \mathcal{H}\mathbf{y}, \mathbf{y} \rangle$ for a Slow Soliton Family in an Anisotropic Medium and for Solitons in an Isotropic Medium

We shall analyze the bilinear form $\langle \mathcal{H}\mathbf{y}, \mathbf{y} \rangle$ in the functional space X which is defined as follows:

$$\mathbf{y} = \{y_1, y_2, y_3, y_4\}^T \in X = H^1(\mathbb{R}) \times H^1(\mathbb{R}) \times L^2(\mathbb{R}) \times L^2(\mathbb{R}),$$

where $H^1(\mathbb{R})$ is the Sobolev space of quadratically integrable functions together with the derivatives, and $L_2(\mathbb{R})$ is the space of quadratically integrable functions on the real $\mathbb{R}$-axis. The space X is chosen because it is the space of least smoothness where the Hamiltonian E is continuous.

In the following analysis a key role will be played by the positivity of

$$d(V) = \partial Q(\phi_V)/\partial V,$$

which has the form

$$d(V) = (2V^2 - \mu_1)\frac{2\sqrt{\rho_0 m}}{\sqrt{\mu_1 - V^2}}$$

for the slow soliton family in the anisotropic case and

$$d(V) = (2V^2 - \mu)\frac{2\sqrt{\rho_0 m}}{\sqrt{\mu - V^2}}$$

for solitons in the isotropic case. Thus, we have $d(V) > 0$ when $V^2 > \mu_1/2$ for the anisotropic case and $V^2 > \mu/2$ for the isotropic case. For conciseness we shall subsequently call the range of the parameter V where $d(V) > 0$ the stability range. It follows from the spectral properties of the operator $\mathcal{H}$ that the bilinear form $\langle \mathcal{H}\mathbf{y}, \mathbf{y} \rangle$ is not positively semidefinite in the entire space X because the negative spectrum of the operator $\mathcal{H}$ is not empty. Nevertheless, for $d(V) > 0$ this bilinear form is nonnegative definite on the linear subspace

$$L = \{\mathbf{y} \in X, \langle Q'(\phi_V), \mathbf{y} \rangle = 0\},$$

tangential to the manifold $Q(\mathbf{w}) = Q(\phi_V)$ at the point $\mathbf{w} = \phi_V$. Moreover we have

$$\langle \mathscr{H}\mathbf{y}, \mathbf{y}\rangle \geq c\langle \mathbf{y}, \mathbf{y}\rangle, \quad \mathbf{y} \in L_1, \tag{3.7}$$

where c is a certain positive constant which depends on the velocity V,

$$L_1 = L \cap \langle \mathbf{y}, \partial_\xi \phi_V\rangle = 0$$

for the slow family in the anisotropic case and

$$L_1 = L \cap \{\langle \mathbf{y}, \partial_\xi \phi_V\rangle = 0, \langle \mathbf{y}, \mathscr{A}\phi_V\rangle = 0\}$$

for the isotropic case. In order to prove (3.7) we note that the condition

$$E'(\phi_V) + VQ'(\phi_V) = 0$$

after differentiation once with respect to velocity yields

$$\begin{aligned}\langle \mathscr{H}\partial_V\phi_V, \partial_V\phi_V\rangle &= -\langle Q'(\phi_V), \partial_V\phi_V\rangle \\ &= -\partial_V Q(\phi_V) = -d(V) < -\epsilon,\end{aligned} \tag{3.8}$$

where $\epsilon > 0$ is a fairly small number which depends on V [we recall that the case $d(V) > 0$ is being considered]. We then express the vector function $\partial_V\phi_V$ in the form

$$\begin{aligned}\partial_V\phi_V = a_0\bar{\chi}_- + b_0\partial_\xi\phi_V + b_1\mathscr{A}\phi_V + \mathbf{p}_0,\\ \mathbf{p}_0 \in P,\end{aligned} \tag{3.9}$$

where $\bar{\chi}_-$ is a single negative eigenvector of $\mathscr{H}$:

$$\mathscr{H}\bar{\chi}_- = \lambda_-\bar{\chi}_-, \quad \langle \bar{\chi}_-, \bar{\chi}_-\rangle = 1,$$

P is a positive half-space of $\mathscr{H}$; for $\mathbf{p}_0 \in P$ we have $\langle \mathscr{H}\mathbf{p}_0, \mathbf{p}_0\rangle \geq \delta\langle \mathbf{p}_0, \mathbf{p}_0\rangle$, δ is a certain positive constant, and $b_1 = 0$ in the anisotropic case. From Eqs. (3.8) and (3.9) we have

$$\langle \mathscr{H}\mathbf{p}_0, \mathbf{p}_0\rangle < a_0^2|\lambda_-| - \epsilon. \tag{3.10}$$

We then represent the vector $\mathbf{y} \in L_1$ as

$$\mathbf{y} = a\bar{\chi}_- + \mathbf{p}, \quad \mathbf{p} \in P. \tag{3.11}$$

It follows from (3.9) and (3.11) for $\mathbf{y} \in L_1$ that

$$\begin{aligned}0 = \langle Q'(\phi_V), \mathbf{y}\rangle &= \langle \mathscr{H}\partial_V\phi, \mathbf{y}\rangle \\ &= a_0 a\lambda_- + \langle \mathscr{H}\mathbf{p}_0, \mathbf{p}\rangle.\end{aligned} \tag{3.12}$$

Then using (3.10), (3.11), (3.12), and the Cauchy–Bunyakovsky inequality, we obtain for $\mathbf{y} \in L_1$

$$\begin{aligned}\langle \mathscr{H}\mathbf{y}, \mathbf{y}\rangle &= -a^2|\lambda_-| + \langle \mathscr{H}\mathbf{p}, \mathbf{p}\rangle \\ &\geq -a^2|\lambda_-| + \delta\langle \mathbf{p}, \mathbf{p}\rangle,\end{aligned} \tag{3.13}$$

and also

$$\begin{aligned}\langle \mathscr{H}\mathbf{y}, \mathbf{y}\rangle &= -a^2|\lambda_-| + \langle \mathscr{H}\mathbf{p}, \mathbf{p}\rangle \\ &\geq -a^2|\lambda_-| + \frac{\langle \mathscr{H}\mathbf{p}, \mathbf{p}_0\rangle}{\langle \mathscr{H}\mathbf{p}_0, \mathbf{p}_0\rangle} \\ &> -a^2|\lambda_-| + \frac{a_0^2 a^2\lambda^2}{a_0^2|\lambda_-| - \epsilon} > \frac{a^2\epsilon}{a_0^2}.\end{aligned} \tag{3.14}$$

We then multiply (3.13) by $\epsilon/2|\lambda_-|a_0^2$ and add to (3.14). As a result we obtain

$$\begin{aligned}\langle \mathscr{H}\mathbf{y}, \mathbf{y}\rangle &\geq \frac{|\lambda_-|\epsilon}{2|\lambda_-|a_0^2 + \epsilon}\left(a^2 + \frac{\delta}{|\lambda_-|}\langle \mathbf{p}, \mathbf{p}\rangle\right) \\ &\geq c(a^2 + \langle \mathbf{p}, \mathbf{p}\rangle),\end{aligned} \tag{3.15}$$

where

$$c = \frac{|\lambda_-|\epsilon}{2|\lambda_-|a_0^2 + \epsilon}\min\{1, \delta/|\lambda_-|\}.$$

From (3.15) and the orthogonality condition $\langle \bar{\chi}_-, \mathbf{p}\rangle = 0$ we then directly obtain the estimate (3.7).

Inequality (3.7) can be made even stronger and specifically we have the inequality

$$\langle \mathscr{H}\mathbf{y}, \mathbf{y}\rangle \geq c_1\|\mathbf{y}\|^2, \quad \mathbf{y} \in L_1, \tag{3.16}$$

where $\|\cdot\|$ denotes the norm in X space:

$$\|\mathbf{w}\| = \left(\int_{-\infty}^{\infty}(u_1^2 + u_2^2 + \partial_\xi u_1^2 + \delta_\xi u_2^2 + v_1^2 + v_2^2)d\xi\right)^{1/2}.$$

To prove (3.16) we note that the definition of $\langle \mathscr{H}\mathbf{y}, \mathbf{y}\rangle$ yields

$$\begin{aligned}\langle \mathscr{H}\mathbf{y}, \mathbf{y}\rangle = \int_{-\infty}^{\infty}\Bigg[&y_3^2 + y_4^2 + \left(\mu_1 - \frac{3\kappa}{\rho_0}u_1^{c2}\right) \\ &+ \left(\mu_2 - \frac{\kappa}{\rho_0}u_1^{c2}\right)y_2^2 + \frac{m}{\rho_0}(\partial_\xi y_1)^2 + \frac{m}{\rho_0}(\partial_\xi y_2)^2 \\ &+ Vy_1y_3 + Vy_2y_4\Bigg]d\xi \geq \int_{-\infty}^{\infty}\Bigg[5(V^2 - \mu_1)y_1^2 \\ &+ (V^2 - \mu_2)y_2^2 + \frac{m}{\rho_0}(\partial_\xi y_1^2)^2 + \frac{m}{\rho_0}(\partial_\xi y_2)^2 \\ &+ \frac{3}{4}y_3^2 + \frac{3}{4}y_4^2\Bigg]d\xi.\end{aligned} \tag{3.17}$$

Here and subsequently in the isotropic case we should write μ instead of μ_1 and μ_2. Multiplying both sides

of (3.7) by $1 - \varepsilon$, where ε is fairly small and using (3.17), we obtain

$$\begin{aligned}\langle \mathcal{H}\mathbf{y}, \mathbf{y}\rangle &\geq \varepsilon\langle \mathcal{H}\mathbf{y}, \mathbf{y}\rangle + (1-\varepsilon)c\langle \mathbf{y}, \mathbf{y}\rangle \\ &\geq \int_{-\infty}^{\infty}\Big\{[(1-\varepsilon)c - 5\varepsilon(\mu_1 - V^2)]y_1^2 \\ &\quad + [(1-\varepsilon)c - \varepsilon(\mu_2 - V^2)]y_2^2 \\ &\quad + \varepsilon m/\rho_0(\partial_\xi y_1)^2 + \frac{\varepsilon m}{\rho_0}(\partial_\xi y_2)^2 \\ &\quad + \Big[(1-\varepsilon)c + \frac{3\varepsilon}{4}\Big](y_3^2 + y_4^2)\Big\}d\xi.\end{aligned} \tag{3.18}$$

From (3.18) we directly obtain (3.16) with the choice

$$c_1 = \min[(1-\varepsilon)c - 5\varepsilon(\mu_1 - V^2),$$
$$(1-\varepsilon)c - \varepsilon(\mu_2 - V^2), (1-\varepsilon)c + 3\varepsilon/4, \varepsilon m/\rho_0].$$

4. ANALYSIS OF STABILITY

We shall determine soliton orbits, i.e., families of solutions, which depend on the parameters of the symmetry group. In the anisotropic case we shall call the soliton orbit ϕ_V a single-parameter family of all shears $T(s)\phi_V$. Similarly, in the isotropic case the soliton orbit is defined as the two-parameter family $T(s)G(\varphi)\phi_V$. A small perturbation of the soliton may lead to the appearance of a soliton at a different velocity and in the isotropic case a soliton formed as a result of a perturbation may correspond to a different angle φ. Thus, it is quite clear that the physical stability of the solitons will be orbital stability where one of the elements of the orbit is transferred to an element of the same orbit. In other words, a soliton is called stable if for any $\epsilon_1 > 0$, $\delta_1 > 0$ exists such that if $\|\mathbf{w}(0) - \phi_V\| < \delta_1$, then

$$\sup_{t>0}\inf_{s\in\mathbb{R}}\|\mathbf{w}(t) - T(s)\phi_V\| < \epsilon_1$$

for the slow soliton family (2.5) in the anisotropic case and

$$\sup_{t>0}\inf_{s\in\mathbb{R}}\inf_{\varphi\in\mathbb{S}}\|\mathbf{w}(t) - T(s)G(\varphi)\phi_V\| < \epsilon_1$$

for the soliton family (2.6) in the isotropic case. Here $\mathbf{w}(t)$ denotes a continuous solution of the basic equations (2.2) in time on an arbitrary segment $t \in [0, T)$ with the initial value $\mathbf{w}(0)$.[1]

We note that for the groups of shifts $T(s)$ and rotations $G(\varphi)$ there are optimal shifts $s(\mathbf{w})$ and rotations $\varphi(\mathbf{w})$ which minimize the distances from the soliton orbits to the vector functions $\mathbf{w}$. These shifts and rotations satisfy [15, 17]

$$\langle T(s(\mathbf{w}))\mathbf{w}, \partial_\xi\phi_V\rangle = 0 \tag{4.1}$$

in the anisotropic case and

$$\begin{aligned}\langle T(s(\mathbf{w}))G(\varphi(\mathbf{w}))\mathbf{w}, \partial_\xi\phi_V\rangle &= 0, \\ \langle T(s(\mathbf{w}))G(\varphi(\mathbf{w}))\mathbf{w}, \mathcal{A}\phi_V\rangle &= 0\end{aligned} \tag{4.2}$$

in the isotropic case.

Orbital stability of these soliton families occurs if the Hamiltonian $E(\mathbf{w})$ has a local minimum near $\mathbf{w} = \phi_V$ subject to the condition $Q(\mathbf{w}) = Q(\phi_V)$ [17], more accurately if the following inequality is satisfied

$$\begin{aligned}E(\mathbf{w}) - E(\phi_V) &\geq c_2\|T(s(\mathbf{w}))\mathbf{w} - \phi_V\|^2, \\ Q(\mathbf{w}) &= Q(\phi_V)\end{aligned} \tag{4.3}$$

for the anisotropic case with $s(\mathbf{w})$ from (4.1) and

$$\begin{aligned}E(\mathbf{w}) - E(\phi_V) &\geq c_2\|T(s(\mathbf{w}))G(\varphi(\mathbf{w}))\mathbf{w} - \phi_V\|^2, \\ Q(\mathbf{w}) &= Q(\phi_V)\end{aligned} \tag{4.4}$$

for the isotropic case with $s(\mathbf{w})$, $\varphi(\mathbf{w})$ from (4.2) for low values of the norms on the right-hand sides of (4.3), (4.4), and $c_2 = \text{const} > 0$. We shall prove this statement. We shall subsequently confine ourselves to the isotropic case, in the anisotropic case all the reasoning is simplified because of the absence of additional rotational symmetry.

Following [17], we shall assume that the family ϕ_V is orbitally unstable. There then exists a series of initial data $\mathbf{w}_n(0)$ and $\delta_2 > 0$ such that

$$\inf_{s\in\mathbb{R}}\inf_{\varphi\in\mathbb{S}}\|\mathbf{w}_n(0) - T(s)G(\varphi)\phi_V\| \longrightarrow 0,$$
$$\sup_{t>0}\inf_{s\in\mathbb{R}}\inf_{\varphi\in\mathbb{S}}\|\mathbf{w}_n(t) - T(s)G(\varphi)\phi_V\| \geq \delta_2,$$

where $\mathbf{w}_n(t)$ are solutions of (2.2) with the initial data $\mathbf{w}_n(0)$. We select the series of instants t_n when

$$\inf_{s\in\mathbb{R}}\inf_{\varphi\in\mathbb{S}}\|\mathbf{w}_n(t_n) - T(s)G(\varphi)\phi_V\| = \delta_2. \tag{4.5}$$

As a result of the continuity of the conserved functionals E and Q, we have

$$E(\mathbf{w}_n(t_n)) = E(\mathbf{w}_n(0)) \longrightarrow E(\phi_V),$$
$$Q(\mathbf{w}_n(t_n)) = Q(\mathbf{w}_n(0)) \longrightarrow Q(\phi_V), \quad n \longrightarrow \infty.$$

We then select the auxiliary series $\mathbf{v}_n$ such that $Q(\mathbf{v}_n) = Q(\phi_V)$ and

$$\lim_{n\to\infty}\|\mathbf{v}_n - \mathbf{w}_n(t_n)\| = 0.$$

It then follows from (4.4) that for fairly small δ_2

$$\begin{aligned}0 \longleftarrow E(\mathbf{v}_n) - E(\phi_V) &\geq c_2\|T(s(\mathbf{v}_n))G(\varphi(\mathbf{v}_n))\mathbf{v}_n - \phi_V\| \\ &= c_2\|\mathbf{v}_n - T(-s(\mathbf{v}_n))G(-\varphi(\mathbf{v}_n))\phi_V\|.\end{aligned}$$

[1] The existence and uniqueness of the Cauchy problem for (2.2) occurs in classes of smooth functions. For the present stability analysis it is sufficient to confine ourselves to continuous functions with respect to time having values in X.

From this it follows that

$$\|\mathbf{v}_n - T(-s(\mathbf{v}_n))G(-\varphi(\mathbf{v}_n))\phi_V\| \longrightarrow 0,$$

which contradicts (4.5).

Thus, in order to determine the orbital stability of these soliton families, it remains for us to establish the validity of inequalities (4.3) and (4.4) which are obtained for $V^2 > \mu_1/2$ and $V^2 > \mu/2$, respectively. As before, we shall confine our analysis to the isotropic case: inequality (4.3) is a particular case of (4.4).

We perform the expansion

$$\mathbf{v} = T(s(\mathbf{w}))G(\varphi(\mathbf{w}))\mathbf{w} - \phi_V = a_1 Q'(\phi_V) + \mathbf{y}, \quad (4.6)$$

where $a_1 = \text{const}$ and $\langle Q'(\phi_V), \mathbf{y}\rangle = 0$. We then have

$$Q(\phi_V) = Q(\mathbf{w}) = Q(T(s(\mathbf{w}))G(\varphi(\mathbf{w}))\mathbf{w})$$
$$= Q(\phi_V) + \langle Q'(\phi_V), \mathbf{v}\rangle + O(\|\mathbf{v}\|^2)$$
$$= Q(\phi_V) + a_1\langle Q'(\phi_V), Q'(\phi_V)\rangle + O(\|\mathbf{v}\|^2).$$

From this it follows that

$$a_1 = O(\|\mathbf{v}\|^2).$$

We then introduce the notation

$$R(\mathbf{w}) = E(\mathbf{w}) + VQ(\mathbf{w})$$

and perform the expansion

$$R(\mathbf{w}) = R(T(s(\mathbf{w}))G(\varphi(\mathbf{w}))\mathbf{w})$$
$$= R(\phi_V) + \langle R'(\phi_V), \mathbf{v}\rangle + \frac{1}{2}\langle \mathscr{H}\mathbf{v}, \mathbf{v}\rangle + o(\|\mathbf{v}\|^2).$$

Since $R'(\phi_V) = 0$ and $Q(\mathbf{w}) = Q(\phi_V)$, we obtain

$$E(\mathbf{w}) - E(\phi_V) = \frac{1}{2}\langle \mathscr{H}\mathbf{v}, \mathbf{v}\rangle + o(\|\mathbf{v}\|^2)$$
$$= \frac{1}{2}\langle \mathscr{H}\mathbf{y}, \mathbf{y}\rangle + O(a_1^2) + O(a_1\|\mathbf{v}\|) + o(\|\mathbf{v}^2\|)$$
$$= \frac{1}{2}\langle \mathscr{H}\mathbf{y}, \mathbf{y}\rangle + o(\|\mathbf{v}\|^2).$$

It is easy to see from (4.6) and (4.2) that

$$\langle \mathbf{y}, \partial_\xi \phi_V\rangle = 0, \quad \langle \mathbf{y}, \partial_\xi \mathscr{A}\phi_V\rangle = 0,$$

i.e., $\mathbf{y} \in L_1$. As a result of (3.16) we have

$$E(\mathbf{w}) - E(\phi_V) \geq \frac{1}{2}c_1\|\mathbf{y}\|^2 + o(\|\mathbf{v}\|^2).$$

Having then noted that

$$\|\mathbf{y}\| = \|\mathbf{v} - a_1 Q'(\phi_V)\|$$
$$\geq \|\mathbf{v}\| - |a_1|\|Q'(\phi_V)\| \geq \|\mathbf{v}\| - O(\|\mathbf{v}\|^2),$$

we arrive at (4.4).

5. DISCUSSION OF RESULTS

In the present paper we have obtained the following results on the nonlinear dynamic stability of the soliton families (2.5) and (2.6):

(i) in an anisotropic material solitons in the slow family (2.5) are stable for velocities V in the range $\mu_1/2 < V^2 < \mu_1$;

(ii) in the absence of anisotropy solitons in the family (2.6) are stable in the velocity range $\mu/2 < V^2 < \mu$.

The operator $\mathscr{H}$ for the fast soliton family in the anisotropic case has an additional unstable direction $\bar{\chi}_-^0$ which corresponds to perturbations of the zeroth component u_1^c of the fast soliton family; we can assume that these perturbations will destroy the soliton in the fast family over the entire velocity range of its existence.

Unlike the situation typical of hydromechanics when the solitons are arranged such that amplitude growth takes place together with an increase in velocity and a reduction in soliton "width", in an elastic medium we have a decrease in amplitude and increase in soliton "width" with increasing velocity. This is responsible for the different behavior of the quantity Q conserved as a result of the translational invariance in hydrodynamics and elasticity theory; in hydrodynamic models it is an increasing function of velocity whereas in elasticity theory models it either decreases (see, for example [15]) or has different behavior on various velocity intervals (as in the case considered in the present study). This factor ensure that the global and local stabilities of the soliton families are equivalent in hydromechanics in the sense that if the elements of a family from a small velocity interval are stable, the entire family is stable. For the case of an elastic medium considered in the present study the situation is different: stability of solitons on these velocity intervals (range of stability) does not imply stability on additional intervals. Moreover, it seems that solitons in the slow family in the anisotropic case for $0 < V^2 < \mu_1/2$ and solitons in the absence of anisotropy for $0 < V^2 < \mu/2$ will be unstable. The physical substantiation for this assumption is that a soliton propagating at lower velocity (and thus having a larger amplitude) is destroyed by the resistance formed as a result of the elastic response of the medium. To overcome this resistance the soliton must have a fairly high velocity, in this case a velocity whose square exceeds $\mu_1/2$ in the anisotropic case or $\mu/2$ in isotropic materials.

The mathematical substantiation of the assumption that solitons having velocities outside the stability range are unstable is as follows. In [17] sufficient instability conditions were formulated for Hamiltonian systems of the form (2.2). The only violation of these conditions in this case is the purely formal fact that the operator $\mathscr{J}$ is not reversible on the functional space X. A similar situation occurs for generalized Korteweg–de Vries equations with higher-order nonlinearity. In [10]

the difficulty associated with the irreversibility of $\mathcal{J}$ in proving the instability of solitons of the generalized Korteweg–de Vries equation was overcome using the properties of an invariant similar to **A** in Section 2.2 in the present study. Strict proof of the instability of solitons in an elastic composite having velocities outside the stability range will be the subject of forthcoming investigations.

ACKNOWLEDGMENTS

I am grateful to A.G. Kulikovskiĭ for useful discussions. This work was supported by the Russian Foundation for Basic Research (project no. 99-01-01150).

REFERENCES

1. N. S. Bakhvalov and M. É. Églit, Tr. Mat. Inst. im. V.A. Steklova, Akad. Nauk SSSR **192**, 5 (1990).
2. T. B. Benjamin, Proc. R. Soc. London, Ser. A **328**, 153 (1972).
3. J. L. Bona, Proc. R. Soc. London, Ser. A **344**, 363 (1975).
4. D. P. Bennet, R. W. Brown, S. E. Stansfield, *et al.*, Math. Proc. Cambridge Philos. Soc. **94**, 351 (1983).
5. J. Shatah and W. A. Strauss, Commun. Math. Phys. **100**, 173 (1985).
6. J. P. Albert, J. L. Bona, and D. B. Henry, Physica D (Amsterdam) **24**, 343 (1987).
7. J. L. Bona and R. L. Sachs, Commun. Math. Phys. **118**, 15 (1988).
8. A. Il'ichev and A. Semenov, Theor. Comput. Fluid Dyn. **3**, 307 (1992).
9. F. Dias and E. Kuznetsov, Phys. Lett. A **263**, 98 (1999).
10. J. L. Bona, P. E. Souganidis, and W. A. Strauss, Proc. R. Soc. London, Ser. A **411**, 395 (1987).
11. R. L. Pego and M. I. Weinstein, Commun. Math. Phys. **164**, 305 (1994).
12. E. A. Kuznetsov, A. M. Rubenchik, and V. E. Zakharov, Phys. Rep. **142**, 103 (1986).
13. R. L. Pego and M. I. Weinstein, Philos. Trans. R. Soc. London, Ser. A **340**, 47 (1992).
14. A. E. H. Love, *A Treatise on the Mathematical Theory of Elasticity* (Dover, New York, 1944).
15. A. Beliaev and A. Il'ichev, Physica D (Amsterdam) **90**, 107 (1996).
16. D. J. Dichmann, J. N. Maddocs, and R. L. Pego, Arch. Ration. Mech. Anal. **135**, 357 (1996).
17. M. Grillakis, J. Shatah, and W. Strauss, J. Funct. Anal. **74**, 160 (1987).
18. A. G. Kulikovskiĭ and E. I. Sveshnikova, *Nonlinear Waves in Elastic Medium* (Moskovs. Litseĭ, Moscow, 1998).
19. N. I. Gvozdovskaya and A. G. Kulikovskiĭ, Prikl. Mekh. Tekh. Fiz. **40**, 174 (1999).

Translation was provided by AIP

Journal of Experimental and Theoretical Physics, Vol. 91, No. 3, 2000, pp. 635–651.
From Zhurnal Éksperimental'noĭ i Teoreticheskoĭ Fiziki, Vol. 118, No. 3, 2000, pp. 730–748.

MISCELLANEOUS

Linear Relaxation Processes Governed by Fractional Symmetric Kinetic Equations¶

A. V. Chechkin and V. Yu. Gonchar

Kharkov Institute for Theoretical Physics, National Science Center Kharkov Institute of Physics and Technology, Kharkov, 310108 Ukraine

Institute of Single Crystals, National Academy of Sciences of Ukraine, Kharkov, 310001 Ukraine
**e-mail: chechkin@ipp.kharkov.ua*
**e-mail: achechkin@kipt.kharkov.ua*

Received November 4, 1999

Abstract—The fractional symmetric Fokker–Planck and Einstein–Smoluchowski kinetic equations that describe the evolution of systems influenced by stochastic forces distributed with stable probability laws are derived. These equations generalize the known kinetic equations of the Brownian motion theory and involve symmetric fractional derivatives with respect to velocity and space variables. With the help of these equations, the linear relaxation processes in the force-free case and for the linear oscillator is analytically studied. For a weakly damped oscillator, a kinetic equation for the distribution in slow variables is obtained. Linear relaxation processes are also studied numerically by solving the corresponding Langevin equations with the source given by a discrete-time approximation to white Levy noise. Numerical and analytical results agree quantitatively. © *2000 MAIK "Nauka/Interperiodica".*

1. INTRODUCTION

The study of the evolution of various systems under the influence of external stochastic forces constitute a large section of statistical physics. It has a great variety of applications in physics, chemistry, biology, economy, and sociology (see, e.g., [1–3]). The best known problem is the description of Brownian motion. In the classical formulation, Brownian motion manifests itself as an unceasing chaotic motion of small macroscopic particles in a liquid or a gas. This is explained by atoms nudging one another and, hence, reveals the atomic structure of continuous medium in which the motion occurs.

The achievements of the theory of probability serve as a mathematical basis for the kinetic theory of Brownian motion. They are as follows:

(i) the Central Limit Theorem, which justifies that the stochastic force acting on a Brownian particle is Gaussian;

(ii) the theory of Markovian stochastic processes; an important assumption used in the kinetic description of Brownian motion is that the behavior of the particle at a given instant depends only on the instantaneous values of the physical parameters, but not on their previous history;

(iii) studies of stochastic Gaussian processes, and primarily, the work of Bachelier (1900), who was the first to study a continuous stochastic Gaussian process with independent increments, and the work of Wiener (1927), who gave a rigorous mathematical formulation of this process and studied the properties of its sample paths.

The basic equations of the kinetic theory of Brownian motion are the Fokker–Planck equation for the probability density function (PDF) $f(x, v, t)$ in the phase space of coordinates and velocities, and the Einstein–Smoluchowski equation for the PDF $f(x, t)$ in real space. The relaxation in the phase space can occur in two steps: the first is the "fast" stage, at which relaxation over velocities occurs and a Maxwellian PDF is established, and the second is the "slow" diffusion stage, at which relaxation in the real (coordinate) space occurs. If the physical conditions in the system are such that the two relaxation stages can be separated, it is possible to pass from the Fokker–Planck equation to the Einstein–Smoluchowski equation and describe the system at the diffusion stage with a simpler equation. The transition to the diffusion stage is discussed in more detail by Chandrasekhar in [4], where a brilliant presentation of the classical Brownian motion theory is given, and also in monograph [3], which contains a modern presentation of the Brownian motion theory including the motion in nonlinear open systems.

In the second half of the 1980s, the term "Levy motion" started to become widely used in statistical physics, in particular, in anomalous diffusion problems, where the characteristic displacement of diffusive particle grows as t^{μ} $\mu \neq 1/2$ (the case where $\mu = 1/2$ corresponds to classical Brownian motion). The Levy anomalous diffusion appears in different areas of physics, including turbulence [5], solid and amorphous state physics [6], chaotic dynamics [7], plasma physics [8], etc. It is also worthwhile mentioning nonphysical areas, e.g., biology and physiology [9], and the theory of

¶This article was submitted by the authors in English.

finance [10]. Levy motion can be considered as a generalization of Brownian motion. The mathematical foundation of this generalization is provided by remarkable properties of stable probability laws, whose theory was begun in the works by Levy and Khintchine [11]. From the limit theorem point of view, the stable probability laws are a generalization of the widely used Gaussian law. Namely, the stable laws are the limiting the probability laws of (properly normalized) sums of independent identically distributed (i.i.d.) random variables [12]. Therefore, these laws (similarly to the Gaussian one) occur when the evolution of a physical system or the result of an experiment is determined by a sum of a large number of identical independent random factors. An important distinction of stable PDFs is a power-law tail decreasing as $|x|^{-1-\alpha}$, $x \longrightarrow \infty$, where α is the Levy index, $0 < \alpha < 2$. Hence, the PDF moments of the order $q \geq \alpha$ diverge.

The above-mentioned properties of stable PDFs allow one to obtain a simple intuitive basis for anomalous diffusion in the framework of the model of independent random "jumps" [6]. However, in order to construct a consistent theory of Levy motion, it is necessary to obtain kinetic equations that generalize those of Brownian motion, namely, the Fokker–Planck and the Einstein–Smoluchowski equations. It is clear from very simple arguments that these equations must contain fractional space and/or time derivatives. During the last two decades, several monographs solely devoted to the theory of fractional calculus have appeared, see, e.g., [13], and an extensive treatment of fractional-order differential equations applied to heat and mass transfer has been given [14]. Different forms of diffusion (for example, equations with fractional derivatives) were proposed [15–20]. These equations were used, in particular, in the description of anomalous diffusion on random fractals [21, 22] and in chaotic Hamiltonian systems, for which the orders of fractional space and time derivatives are determined by delicate properties of the phase space [7, 23]. We also refer to [24], where the general description of the fractional relaxation-oscillation and fractional diffusion-wave phenomena was provided using a simple adaptation of a mathematical approach to the fractional calculus.

Our paper deals with fractional generalizations of the classical Fokker–Planck and Einstein–Smoluchowski equations describing the respective relaxation in the phase space and in real space [44]. We follow the classical approach [4] in deriving kinetic equations for Brownian motion and also the approach of [25], where, as far as we know, the fractional kinetic equation for the phase space PDF was proposed for the first time. Throughout the paper, we restrict ourselves to a one-dimensional coordinate space and two-dimensional (coordinate plus velocity) phase space. In addition, we restrict ourselves to symmetric fractional kinetic equations, that is, those involving symmetric fractional derivatives (see below).

First, we derive kinetic equations with symmetric fractional derivatives, which generalize the Fokker–Planck and Einstein–Smoluchowski equations in the case of Levy motion. These equations are called the fractional symmetric Fokker–Planck equation (FSFPE) and the fractional symmetric Einstein–Smoluchowski equation (FSESE), respectively. Second, we use these equations in studying classical problems of linear relaxation, namely, relaxation in a force-free case and relaxation of a linear Levy oscillator.

It is worthwhile to note that the force-free relaxation in a spatially homogeneous case and the relaxation of a linear oscillator were first studied in [26]. In this paper, the equations for characteristic functions were obtained by solving the corresponding Langevin equations and by subsequently averaging the Liouville equation over the phase space density. The linear oscillator was also considered in [27] with the help of a fractional kinetic equation for the diffusion stage of relaxation. Using the FSFPE and FSESE, we study in detail both the "fast" and the "slow" stages of a linear relaxation and demonstrate a transition from the first level of description (the use of FSFPE) to the second one (the use of FSESE).

We next consider two limiting cases for the oscillator, namely, an overdamped and a weakly damped oscillator. Both these cases are very important in studying the Levy motion in nonlinear open systems. We propose a new kinetic equation for the weakly damped linear oscillator and study its solutions. We also solve numerically the Langevin equations that correspond to fractional kinetic equations. We demonstrate numerical results for the linear relaxation problems that are solved analytically, and show a close agreement between the kinetic theory results and the numerical modeling.

The paper is organized as follows. In Section 2, we derive fractional generalizations of the Fokker–Planck and Einstein–Smoluchowski equations following the approaches of [4] and [25]. In Section 3, we investigate relaxation in real space for force-free Levy motion and for the Levy linear oscillator. In Section 4, we investigate both relaxation problems in the phase space. In Section 5, we consider relaxation for the overdamped and weakly damped limits of the Levy oscillator. Conclusions and a discussion are presented in Section 6.

2. FRACTIONAL FOKKER–PLANCK AND FRACTIONAL EINSTEIN–SMOLUCHOWSKI EQUATIONS

The derivation of the Fokker–Planck equation is usually based on the finiteness assumption for the second moment of the PDF. Since this assumption breaks down for stable PDFs, we find that the classical approach used by Chandrasekhar [4] can be adopted for a derivation not using the finiteness of the second moment. A similar treatment was undertaken in [28], where the discussion of [29] was adopted for the pur-

pose of obtaining a kinetic equation in coordinate space.

2.1. Fractional Fokker–Planck Equation

Similar to the Brownian motion theory, the initial equations in our approach are as follows:

(1) the integral equation for the PDF $f(x, v, t)$ of the Markovian stochastic process in the phase space,

$$f(x, v, t+\Delta t) = \iint d(\Delta x)d(\Delta v) \times f(x-\Delta x, v-\Delta v, t)\Psi(x-\Delta x, v-\Delta v; \Delta x, \Delta v, \Delta t), \quad (2.1)$$

where $\Psi(x, v; \Delta x, \Delta v, \Delta t)$ is the transition probability, that is, the probability for the coordinate x to acquire the increment Δx and for the velocity v the increment Δv during the time interval Δt;

(2) the Langevin equations

$$\frac{dx}{dt} = v, \quad \frac{dv}{dt} = -\nu v + F + A(t), \quad (2.2)$$

where ν is the friction coefficient (which is assumed to be independent of v), F is the regular external force, and $A(t)$ is the fluctuation component of the external force.

Following traditional assumptions in the theory of Brownian motion [4], one then obtains the expressions for the coordinate and velocity increments during the time interval Δt that is larger than the characteristic time intervals of $A(t)$ but smaller than the time intervals during which physical parameters change appreciably:

$$\Delta x = v\Delta t, \quad \Delta v = -(\nu v - F)\Delta t + B(\Delta t). \quad (2.3)$$

Here

$$B(\Delta t) = \int_t^{t+\Delta t} A(t')dt'$$

is a nonstationary stochastic process that is assumed to be a homogeneous Gaussian process with independent increments possessing a PDF

$$w(B(\Delta t)) = \frac{1}{\sqrt{4\pi D\Delta t}}\exp\left(-\frac{(B(\Delta t))^2}{4D\Delta t}\right). \quad (2.4)$$

The Central Limit Theorem serves as a mathematical justification of this assumption. In accordance with the above, we generalize the Chandrasekhar approach by generalizing the Central Limit Theorem to i.i.d. variables with infinite variances. We recall that Levy and Khintchine [11] discovered a class of stable probability laws. These are limiting laws for the probability laws of normalized sums of i.i.d. random variables. Each stable law with a characteristic Levy index ($0 < q < 2$) possesses finite moments of the orders q, $0 < q < \alpha$, but infinite moments of higher orders. The Gaussian law is also a stable one with the characteristic index $\alpha = 2$, and it possesses moments of all orders. Returning to Eq. (2.4), we note that in the theory of stochastic processes, the corresponding generalizations of a homogeneous Gaussian process $B(t)$ with independent increments are stable Levy processes $L(t)$ with the characteristic function (we restrict ourselves to symmetric stable laws) [30]

$$\hat{w}_L(k, \Delta t) = \langle e^{ikL}\rangle = \exp(-D|k|^2\Delta t), \quad (2.5)$$

where $D > 0$ and $(D\Delta t)^{1/\alpha}$ is called the scale parameter. At $\alpha = 2$, one recovers the Gaussian process $B(t)$. The above statements justify the expediency of the generalization $B(\Delta t) \longrightarrow L(\Delta t)$ in Eqs. (2.3).

With Eqs. (2.3) and (2.5), the transition probability in Eq. (2.1) is

$$\Psi(x, v; \Delta x, \Delta v) = \psi(x, v; \Delta v)\delta(\Delta x - v\Delta t),$$

where δ is the Dirac delta-function and

$$\psi(x, v, \Delta v) = \int_{-\infty}^{\infty}\frac{dk}{2\pi} \times \exp[-ik(\Delta v + \nu v\Delta t - F\Delta t) - D|k|^\alpha\Delta t] \quad (2.6)$$

is the transition probability in the velocity space.

We insert Eq. (2.6) in Eq. (2.1), expand the left- and the right-hand sides into series in Δt, and let $\Delta t \longrightarrow 0$. As a result, we obtain

$$\frac{\partial f}{\partial t} + v\frac{\partial f}{\partial x} = -\int d(\Delta v) f(x, v-\Delta v, t) \times \int_{-\infty}^{\infty}\frac{dk}{2\pi}\exp(-ik\Delta v)[-ikF + ikv(v-\Delta v) + D|k|^\alpha]. \quad (2.7)$$

We now turn to the physical space by making the inverse Fourier transform with respect to the velocity on the right-hand side of Eq. (2.7). We treat each term in the square brackets separately. The first and the second terms, being "classical," are transformed trivially, yielding $-F(x, t)\partial f/\partial v$ and $\nu\partial(vf)/\partial v$, respectively. The last term can be written as

$$-D\int_{-\infty}^{\infty}\frac{d\kappa}{2\pi}e^{-i\kappa x}\int_{-\infty}^{\infty}\frac{dk}{2\pi}e^{-ikv}|k|^\alpha\hat{f}(\kappa, k, t), \quad (2.8)$$

where $\hat{f}(\kappa, k, t)$ is the characteristic function.

We use the symmetric fractional derivative of an arbitrary order $\alpha > 0$ that can be defined, for a "sufficiently well-behaved" function $\phi(v)$, $v \in \mathbf{R}$, as the pseudo-differential operator characterized by its Fourier representation,

$$\frac{d^\alpha}{d|v|^\alpha}\phi(v) \div -|k|^\alpha\hat{\phi}(k), \quad v, k \in \mathbf{R}, \quad \alpha > 0.$$

On the left-hand side, we adopt the notation introduced in [19].

To treat this kind of fractional derivatives properly, we recall the definition of the left and right Liouville derivatives on the infinite axis [13],

$$\mathcal{D}_+^\alpha \phi(v) = \frac{1}{\Gamma(1-\alpha)}\frac{d}{dv}\int_{-\infty}^{v}\frac{\phi(\xi)d\xi}{(x-\xi)^\alpha},$$
$$\mathcal{D}_-^\alpha \phi(v) = -\frac{1}{\Gamma(1-\alpha)}\frac{d}{dv}\int_{v}^{\infty}\frac{\phi(\xi)d\xi}{(\xi-x)^\alpha}, \quad (2.9)$$

where $0 < \alpha < 1$. For $\alpha \geq 1$,

$$\mathcal{D}_\pm^\alpha \phi(v) = \frac{(\pm 1)^n}{\Gamma(n-\alpha)}\frac{d^n}{dv^n}\int_0^\infty \xi^{n-\alpha-1}\phi(v \mp \xi)d\xi, \quad (2.10)$$

$n = [\alpha] + 1$, where the square brackets denote the integer part. Derivatives (2.9)–(2.10) are characterized by their Fourier representation as

$$\mathcal{D}_\pm^\alpha \phi(v) \div (\mp ik)^\alpha \hat{\phi}(k), \quad \alpha \geq 0,$$

where

$$(\mp ik)^\alpha = |k|^\alpha \exp\left(\mp \frac{i\alpha\pi}{2}\operatorname{sgn} k\right).$$

Thus, the symmetric fractional space derivative can be written as

$$\frac{d^\alpha}{d|v|^\alpha}\phi(v) = -\frac{1}{2\cos(\pi\alpha/2)}[\mathcal{D}_+^\alpha \phi(v) + \mathcal{D}_-^\alpha \phi(v)], \quad (2.11)$$

where $\alpha \neq 1, 3, \ldots$.

We now return to Eq. (2.7) and write the kinetic equation for the PDF $f(x, v, t)$ in the phase space as

$$\frac{\partial f}{\partial t} + v\frac{\partial f}{\partial x} + F\frac{\partial f}{\partial v} = \nu\frac{\partial}{\partial v}(vf) + D\frac{\partial^\alpha}{\partial |v|^\alpha} f, \quad (2.12)$$

where the last term is defined through Eqs. (2.9)–(2.11). This is a fractional Fokker–Planck equation for Levy motion. For $\alpha = 2$, this is the standard Fokker–Planck equation for Brownian motion.

We note that Eq. (2.12) becomes meaningless when α is an odd integer. That is why the particular case $\alpha = 1$ must be treated separately in our range of interest $0 < \alpha \leq 2$. However, if one uses the Fourier transform over velocity in solving a particular problem, this case is not singled out.

2.2. Fractional Einstein–Smoluchowski Equation

Along with the relaxation parameter $1/\nu$, the description of Brownian motion involves another relaxation parameter that characterizes diffusion in real space. If the characteristic time of this process is much greater than $1/\nu$, it is possible to pass from the Fokker–Planck equation for the PDF $f(x, v, t)$ to the Einstein–Smoluchowski equation for a simpler PDF $f(x, t)$.

As in deriving the fractional symmetric Einstein–Smoluchowski equation in the previous subsection, we follow the reasoning used in the theory of Brownian motion.

Instead of Eq. (2.1), an integral equation in the coordinate space serves as an initial one,

$$f(x, t + \Delta t) = \int d(\Delta x) f(x - \Delta x, t)\psi(x - \Delta x; \Delta x, \Delta t),$$

where $\psi(x; \Delta x, \Delta t)$ is the transition probability, that is, the probability for the coordinate x to square the increment Δx during the interval Δt.

In the kinetic theory of Brownian motion, the transition to the Einstein–Smoluchowski equation corresponds to neglecting the term dv/dt in Langevin Eq. (2.2) [2]. Thus, instead of two equations, we have a single Langevin equation,

$$\frac{dx}{dt} = \frac{F}{\nu} + \frac{1}{\nu}A(t), \quad (2.13)$$

and instead of Eqs. (2.3), we obtain

$$\Delta x = F\Delta t/\nu + L(\Delta t)/\nu,$$

where $L(t)$ is a stable process with a symmetric PDF and characteristic function (2.5), as before. We now obtain the fractional symmetric Einstein–Smoluchowski equation

$$\frac{\partial f}{\partial t} = -\frac{\partial}{\partial x}\left(\frac{F}{\nu}f\right) + \frac{D}{\nu^\alpha}\frac{\partial^\alpha}{\partial |x|^\alpha} f. \quad (2.14)$$

In what follows, we give examples of relaxation processes governed by Eqs. (2.12) and (2.14). In all cases, in order to obtain the solutions, we pass to the characteristic functions $\hat{f}(\kappa, t)$ with respect to the coordinate of the FSESE and to the characteristic function $\hat{f}(\kappa, k, t)$ with respect to the coordinate and the velocity of the FSFPE.

3. SOLUTIONS TO THE FRACTIONAL SYMMETRIC EINSTEIN–SMOLUCHOWSKI EQUATION

In this section, we consider two simple examples of relaxation processes governed by FSESE, namely, relaxation in a force-free case and relaxation of the Levy linear oscillator.

3.1. Force-Free Relaxation

Setting $F = 0$ in Eq. (2.14), we seek the solution with the initial condition

$$f(x, t = 0) = \delta(x - x_0).$$

The characteristic function of the solution is

$$\hat{f}(\kappa, t) = \exp\left\{ i\kappa x_0 - \frac{D}{\nu^\alpha}|\kappa|^\alpha t \right\}. \quad (3.1)$$

In the force-free case, the random process $x(t)$ is therefore a non-stationary Levy stable process with independent increments and with the scale parameter $(Dt)^{1/\alpha}/\nu$. In the real space, the Levy stable PDFs are expressed in terms of the Fox' H functions [31]. This representation of all stable PDFs was achieved in [32]. Mathematical details on H functions are presented in [33, 34]. In the present paper, we do not consider the real space representation for an arbitrary α, however.

Because the variance and higher moments of integer orders diverge for stable PDFs, the moments of fractional orders can be used as statistical means characterizing the properties of these processes [26, 35]. To guarantee the reality, the moments must be defined for the modulus of the stochastic variable. For the force-free relaxation, the fractional-order moments are therefore given by

$$M_x(t; q, \alpha) = \int_{-\infty}^{\infty} dx|x|^q f(x, t) = \begin{cases} ((Dt)^{1/\alpha}/\nu)^q C(q; \alpha), & 0 < q < \alpha \\ \infty, & q \geq \alpha, \end{cases} \quad (3.2)$$

for $0 < \alpha < 2$, and by

$$M_x(t; q, 2) = \frac{(Dt)^{q/2}}{\nu^2} C(q; 2) \quad (3.3)$$

for $\alpha = 2$ and an arbitrary q, where

$$C(q; \alpha) = \int_{-\infty}^{\infty} dx_2|x_2|^q \int \frac{dx_1}{2\pi} \exp(-ix_1x_2 - |x_1|^\alpha).$$

The coefficient $C(q; \alpha)$ can be evaluated using the generalized function theory [26] as

$$C(q; \alpha) = \frac{2}{\pi q} \sin\left(\frac{\pi q}{2}\right) \Gamma(1 + q) \Gamma\left(1 - \frac{q}{\alpha}\right), \quad 0 < q < \alpha. \quad (3.4)$$

Equations (3.2)–(3.4) have a direct physical consequence for the description of anomalous diffusion. For ordinary Brownian motion, the characteristic displacement $\delta x(t)$ of a particle can be written through the second moment as

$$\delta x(t) = M_x^{1/2}(t; 2, 2) \propto t^{1/2}.$$

One can note from Eqs. (3.3) that for the normal diffusion, $M_x^{1/q}(t; q, 2) \propto t^{1/2}$ at any q, and therefore, any order of the moment can serve as a measure of the normal diffusion rate,

$$\delta x(t) \approx M_x^{1/q}(t; q, 2) \propto t^{1/2},$$

if one is interested in the time dependence of the characteristic displacement, but not in the value of the prefactor. We remind the reader that the time dependence alone, without the prefactor, usually serves as an indicator of normal or anomalous diffusion [6]. In analogy with the Brownian case, it follows from Eqs. (3.2) that the quantity $M_x^{1/q}(t; q, \alpha)$ for $0 < \alpha < 2$ and any $q < \alpha$ can serve as a measure of the anomalous diffusion rate,

$$\delta x(t) \approx M_x^{1/q}(t; q, \alpha) \propto t^{1/\alpha}, \quad 0 < q < \alpha < 2. \quad (3.5)$$

This describes a fast anomalous diffusion, or superdiffusion, as the second moment $M_x(t; 2, \alpha)$ diverges. We note that in order to extract the scaling form for the second moment, the "walker" was enclosed in an "imaginary growing box" in [27]. This formal procedure leads to the diffusion scaling $t^{2/\alpha}$ for the variance, which is consistent with Eq. (3.5), and it was implemented numerically. However, it seems that a more physically relevant procedure must take the finite velocity of a diffusive particle into account. This problem is beyond the scope of our paper. We only mention an extensive recent discussion on this interesting and important theme [36, 37].

The numerical simulation results for the Langevin equation (2.13) are shown in Figs. 1 and 2. Here and in what follows, the stochastic source $A(t)$ is represented in numerical simulation as a discrete approximation of a "white Levy noise", that is, as a stationary sequence of independent identically distributed variables having a symmetric stable PDF with the Levy index α and with the scale parameter equal to 1. To obtain the sequence, we use the generator based on the inversion method [38] and the Gnedenko limit theorem [12]. This generator was described in our recent papers [39] and [40] in more detail. The time interval between the terms of the sequence is equal to unity. In the force-free problem, we estimate the moments $M_x(t; q, \alpha)$ numerically by averaging realizations of $x(t)$. The total number of realizations is equal to 500, each of a length of 512.

In Fig. 1, we show $M_x(t; q, 1)$ versus t at a different q in a log-log scale. For $q < \alpha = 1$, the dependence is well fitted by the straight line whose slope allows us to define the diffusion exponent. For $q \geq \alpha$, the theoretical value of the moment is infinite, and in numerical simulation the moment strongly fluctuates, and it is therefore impossible to obtain the diffusion exponent.

In Fig. 2, we show the exponent μ involved in the relation $M_x(t; q, \alpha) \propto t^\mu$ versus the Levy index α of the discrete approximation of the white noise. The order q of the moment is equal to 0.25, which is smaller than the smallest value $\alpha = 1$ used in numerical simulation. The theoretical dependence $0.25/\alpha$ is shown by the dotted line. The values of μ obtained in simulations are shown by black dots. Close agreement between the theory and numerical simulations is obvious.

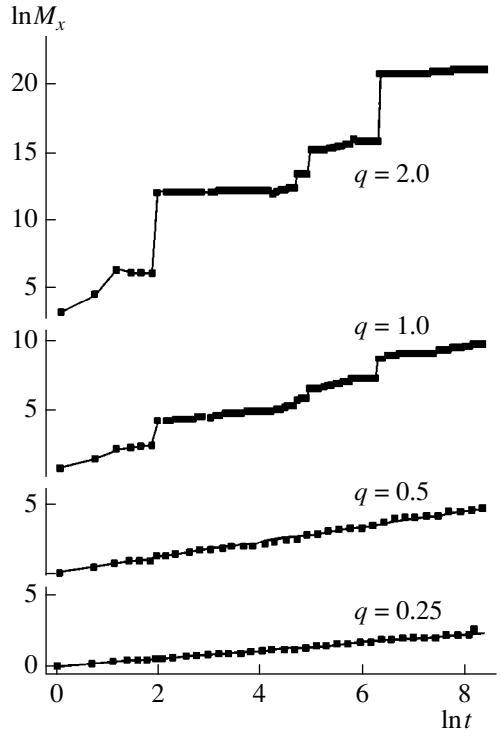

Fig. 1. Force-free relaxation in the FSESE framework. The moment M_x versus t at different values of the moment exponent q. The Levy index α is 1.

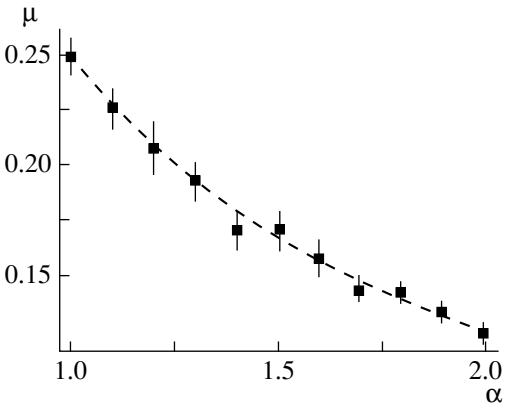

Fig. 2. Force-free relaxation in the FSESE framework. The exponent μ of the time dependence of the moment $M_x(t; q, \alpha) \sim t^{\mu}$ versus the Levy index α. The moment order is $q = 0.25$. The theoretical dependence $0.25/\alpha$ is depicted by the dotted line. Numerical results are depicted by black dots.

3.2. Relaxation of the Linear Levy Oscillator

Setting $F = -\omega^2 x$ in Eq. (2.14), we seek the solution with the initial condition

$$f(x, t = 0) = \delta(x - x_0).$$

The characteristic function of the solution is

$$\hat{f}(\kappa, t) = \exp\{i\kappa x_0 e^{-\omega^2 t/\nu} - D_{\rm osc}(t)|\kappa|^{\alpha}\}, \qquad (3.6)$$

where

$$D_{\rm osc}(t) = \frac{D}{\alpha\omega^2\nu^{\alpha-1}}(1 - e^{-\alpha\omega^2 t/\nu}).$$

This result was recently obtained in [27].

It can be seen that the relaxation of the oscillator, contrary to the force-free case, is not a Levy stable process with independent increments. It can be named a "stable-like" or a "Levy-like" process, because the exponent of the characteristic function involves $|\kappa|^{\alpha}$, but the scale parameter for the oscillator, $(D_{osc}(t))^{1/\alpha}$, does not increase as $t^{1/\alpha}$, which is a manifestation of the Levy stable process with independent increments, see Eq. (2.5). The process $x(t)$ behaves as a Levy stable one only asymptotically for short times,

$$t \ll \tau_x = \nu/\alpha\omega^2.$$

On the other hand, for $t \gg \tau_x$, the process $x(t)$ becomes asymptotically stationary with a stable PDF that does not depend on t and with the Levy index α and the scale parameter

$$D_{\rm osc}^{1/\alpha}(t = \infty) = \left(\frac{D}{\alpha\omega^2\nu^{\alpha-1}}\right)^{1/\alpha}. \qquad (3.7)$$

We also note that for the Brownian oscillator only, the stationary solution has the Boltzmann form

$$f_{\rm st}(x;\ \alpha = 2) = \sqrt{\frac{\nu\omega^2}{2\pi D}}\exp\left\{-\frac{\nu\omega^2}{2D}x^2\right\}. \qquad (3.8)$$

In what follows, we return to the problems related to stationary solutions of fractional kinetic equations.

It is convenient to define fractional moments after subtracting from a stochastic quantity x its regular part containing the initial condition, that is, $x_0\exp(-\omega^2 t/\nu)$. Thus, the moments are

$$\begin{aligned} M_x(t;\ q, \alpha) &= \left\langle \left| x - x_0\exp(-\omega^2 t/\nu)\right|^q \right\rangle \\ &= D_{\rm osc}^{q/\alpha}(t)C(q;\ \alpha), \end{aligned} \qquad (3.9)$$

where $C(q; \alpha)$ is the same as in (3.4).

Numerical simulations of the linear oscillator relaxation involve solution of Langevin equation (2.13) with the external force $F = -\omega^2 x$ and the calculation of the qth order moments. In Fig. 3, we present the results obtained for various values of ω by averaging over 300 realizations, each of length 4096. The Levy index α is equal to 1, and the order of the moment is 0.25. The values obtained in the numerical simulation are depicted by black dots, whereas the solid line indicates the values estimated from Eq. (3.9). The vertical mark indicates the value τ_x, after which the random process $x(t)$

becomes stationary. For $t > \tau_x$, the moment tends to a constant value that can be estimated from Eqs. (3.7) and (3.9). Numerical results demonstrate close agreement with theoretical estimates on both the non-stationary and stationary stages of the evolution.

4. SOLUTIONS TO THE FRACTIONAL SYMMETRIC FOKKER–PLANCK EQUATION

In this section, as in the previous one, we consider the same examples of relaxation processes, but governed by FSFPE.

4.1. Force-Free Relaxation

Setting $F = 0$ in Eq. (2.12), we seek the solution of FSFPE with the initial condition

$$f(x, v, t = 0) = \delta(x - x_0)\delta(v - v_0).$$

For clarity, it is expedient to consider the space-homogeneous relaxation first.

4.1.1. Space-homogeneous force-free relaxation. The solution for the characteristic function with respect to the velocity is

$$\hat{f}(k, t) = \exp\{ikv_0 e^{-\nu t} - |k|^\alpha D_{ff}^{(\nu)}(t)\}, \qquad (4.1)$$

where

$$D_{ff}^{(\nu)}(t) = (D/\alpha\nu)(1 - e^{-\alpha\nu t})$$

and "*ff*," means "force-free." The space homogeneous relaxation in a force-free case was first considered in [26]. The relaxation process is not a Levy stable process with independent increments, but, following the terminology used in the previous section, it can be called a stable-like, or Levy-like process, since its characteristic function (4.1) involves $|k|^\alpha$ in the exponent, but $D_{ff}^{(\nu)}(t)$ is not a linear function of t, and hence, the scale parameter $(D_{ff}^{(\nu)}(t))^{1/\alpha}$ does not scale as $t^{1/\alpha}$. A stable process with independent increments arises for short times, $t \ll \tau_\nu = 1/\alpha\nu$, when the exponent in (4.1) can be expanded in a power series. With the accuracy up to linear terms in t inclusively, we obtain the Levy stable process. On the other hand, for $t \gg \tau_\nu$, the stochastic process $v(t)$ becomes asymptotically stationary with a stable PDF that is independent of t and with the Levy index α and the scale parameter

$$(D_{ff}^{(\nu)}(t = \infty))^{1/\alpha} = \left(\frac{D}{\alpha\nu}\right)^{1/\alpha}. \qquad (4.2)$$

Here it seems expedient to discuss the problems related to stationary solutions of fractional kinetic equations. In the classical theory of Brownian motion, the equilibrium Maxwell PDF over velocity is reached as $t \longrightarrow \infty$. It is characterized by the temperature T of the surrounding medium. There exists a relation between the parameter D of the PDF of the random source in the Langevin equations, see Eq. (2.2), and the friction coefficient ν such that

$$D = \nu k_B T/m,$$

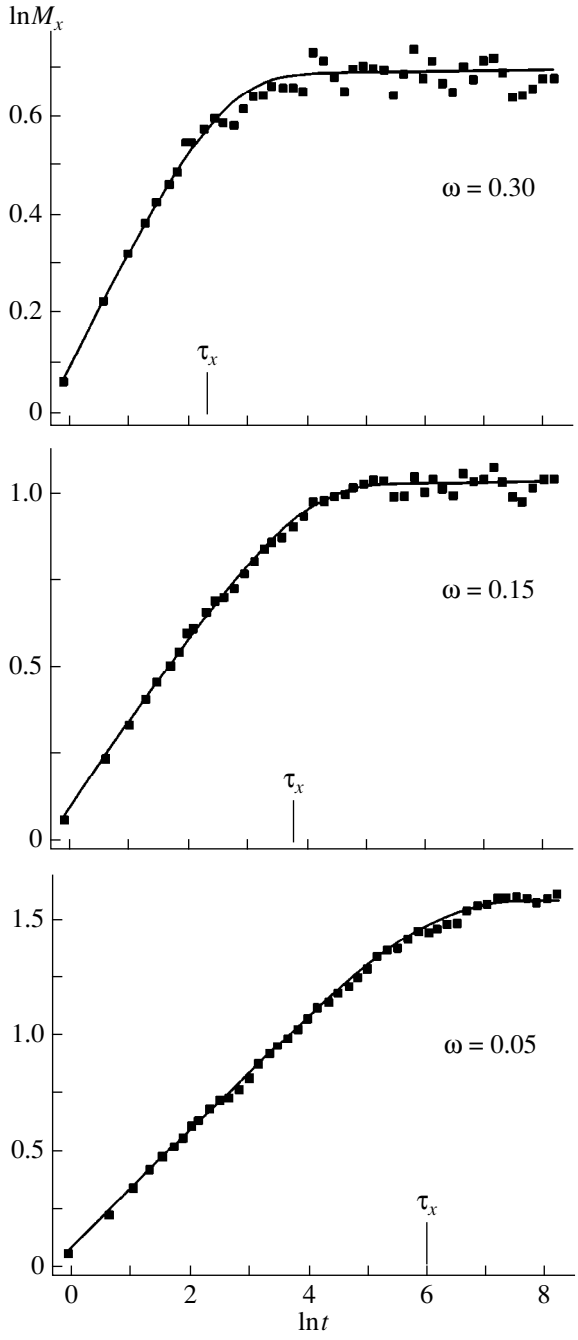

Fig. 3. Relaxation of a linear oscillator in the FSESE framework. The qth order coordinate moment versus time in a twice logarithmic scale. The numerical simulation results are depicted by black dots, the moment obtained from the FSESE is shown by the solid line. Vertical marks indicate the coordinate relaxation time.

where m is the mass of the Brownian particle and k_B is the Boltzmann constant. The temperature T is a measure of the mean kinetic energy of the Brownian particle,

$$\langle E_{\rm kin}\rangle = \frac{m\langle v^2\rangle}{2} = \frac{k_B T}{2}. \tag{4.3}$$

Equation (4.3) is an example of fluctuation-dissipation relations. In this case, the source $A(t)$ in the Langevin equation is called the source of internal fluctuations. Relation (4.3) may not be satisfied, as it takes place, e.g., in auto-oscillation systems. One then says that $A(t)$ is a source of external (relative to the system) fluctuations in Eq. (2.2). However, the Maxwell–Boltzmann exponential form of stationary solutions is preserved [2]. As for the Levy motion, the fluctuation-dissipation relations cannot be satisfied, because of the infinite squared velocity $\langle v^2\rangle = \infty$ for $0 < \alpha < 2$. Therefore, we can only speak of $A(t)$ as of a source of external fluctuations. Moreover, it follows from the example considered in this subsection and from the linear oscillator example considered above that the stationary solutions do not possess the Maxwell–Boltzmann form but rather a more general form of stable distributions. At present, there is no theory of equilibrium state based on stable PDFs. Perhaps, it can be constructed with the help of Tsallis' statistics and his generalizations, see recent review [41] and references therein.

We also write the qth order fractional moment of the velocity,

$$\begin{aligned} M_v(t; q, \alpha) &= \left\langle \left| v - v_0 e^{-\nu t}\right|^q \right\rangle \\ &= ((D_{ff}^{(v)}(t))^{q/\alpha} C(q; \alpha)), \end{aligned} \tag{4.4}$$

where $D_{ff}^{(v)}(t)$ is determined above. This formula is compared with the results of numerical simulations at the end of Subsection 4.1.

4.1.2. Space-inhomogeneous force-free relaxation. We turn to the force-free relaxation in the general case, which is governed by Eq. (2.12) with $F = 0$. The solution for the characteristic function $\hat{f}(\kappa, k, t)$ can be obtained by the method of characteristics. For Brownian motion, where $\alpha = 2$, one can make an analytical Fourier transformation [4]. For an arbitrary α, $0 < \alpha < 2$, the analysis becomes rather complicated. However, since we already have some information about the velocity relaxation, we study the evolution of a simpler function,

$$f(x, t|x_0, v_0) = \int dv\, f(x, v, t|x_0, v_0),$$

whose characteristic function is

$$\begin{aligned} &\hat{f}(\kappa, k = 0, t|x_0, v_0) \\ &= \exp\left\{ i\kappa x_0 + i\kappa \frac{v_0}{\nu}(1 - e^{-\nu t}) - D_{ff}^{(x)}(t)|\kappa|^\alpha \right\}, \end{aligned} \tag{4.5}$$

where

$$D_{ff}^{(x)}(t) = \frac{D}{\nu^\alpha}\int_0^t d\tau (1 - e^{-\nu\tau})^\alpha.$$

The random coordinate is a stable-like process. For prolonged times $t \gg \tau_v$, we obtain

$$D_{ff}^{(x)}(t \longrightarrow \infty) \longrightarrow \frac{Dt}{\nu^\alpha} \tag{4.6}$$

and the characteristic function (4.5) coincides with the solution of the Einstein–Smoluchowski equation in the force-free case. In the prolonged time limit, therefore, the random process $x(t)$ becomes an α-stable process with independent increments and with the Levy index α and the scale parameter $(Dt)^{1/\alpha}/\nu$. One can see that the space-inhomogeneous relaxation occurs in two stages, namely, the "fast" stage, at which a stationary stable velocity PDF is established after the time period τ_v, and the "slow" diffusion stage, whose characteristic relaxation time τ_x can be defined as

$$\tau_x \approx (\nu L)^\alpha / D,$$

where L is an external scale of the system. For sufficiently large systems, $\tau_x \gg \tau_v$.

For the coordinate, we also write the qth order fractional moment that is estimated in the numerical simulations given in what follows,

$$\begin{aligned} M_x(t; q; \alpha) &= \left\langle \left| x - x_0 - \frac{v_0}{\nu}(1 - e^{-\nu t})\right|^q \right\rangle \\ &= (D_{ff}^{(x)}(t))^{q/\alpha} C(q; \alpha). \end{aligned} \tag{4.7}$$

Numerical simulation of the force-free relaxation process described by FSFPE involves solving Langevin equations (2.2) with $F = 0$ and estimating the moments $M_v(t; q, \alpha)$ and $M_x(t; q, \alpha)$. The results are shown in Figs. 4, 5.

Figure 4 has an illustrative purpose. It shows typical velocity trajectories (at the left) and coordinate trajectories (at the right) for various values of the Levy index. We set $\nu = 0.03$ and, thus, the velocity relaxation time τ_v is equal to 20 for $\alpha = 1.7$, 26 for $\alpha = 1.3$, and 33 for $\alpha = 1.0$, respectively. In most of the realizations presented here, the process $v(t)$ is stationary. Large outliers that are clearly seen in the velocity realizations appear to be due to the power law asymptotic behavior of the stable PDFs of the velocity. With the Levy index decreasing (from top to bottom), the asymptotic behavior becomes flatter, leading to the growth of the outlier amplitudes. Large outliers of the velocity, in their turn, lead to large jumps in the trajectories $x(t)$, that is, the Levy flights, see illustrations on the right.

In Fig. 5, we depict the velocity moments M_v (at the left) and coordinate moments M_x (at the right) versus time at different values of the friction coefficient v. The

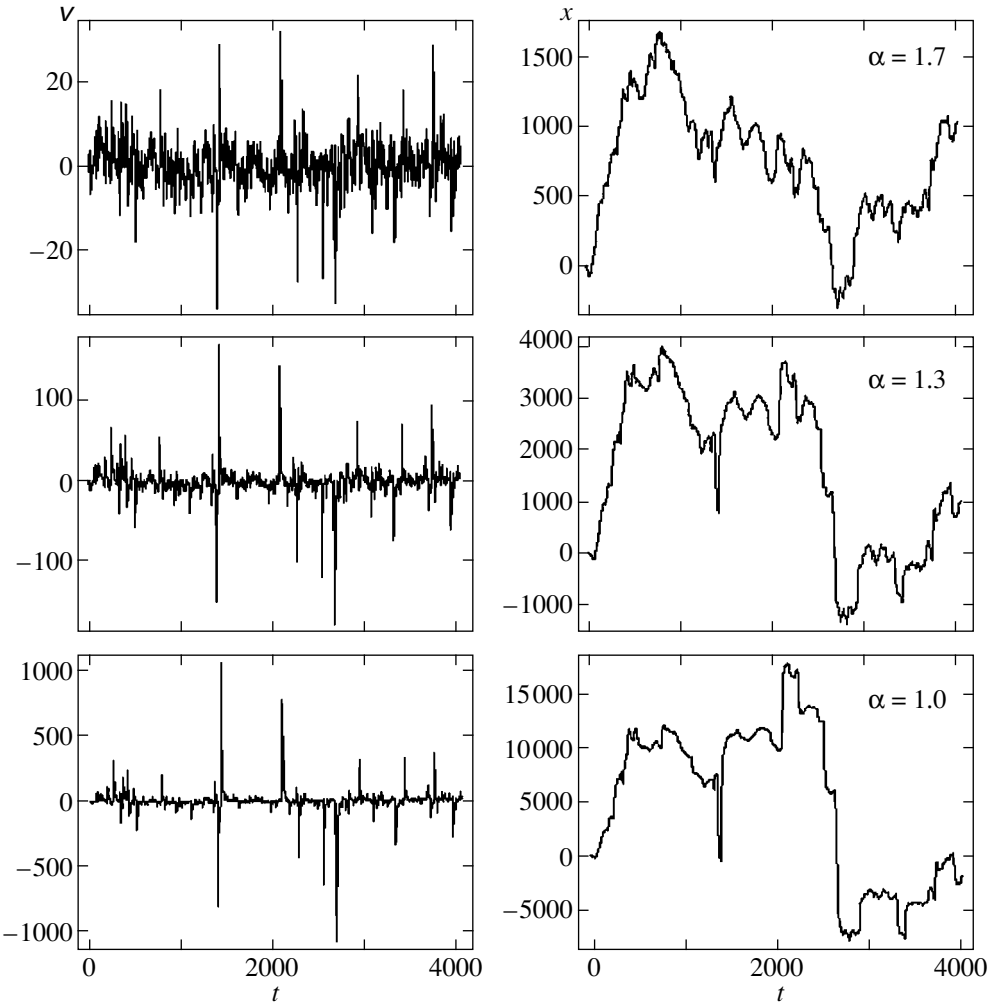

Fig. 4. Force-free relaxation in the FSFPE framework. The typical trajectories of the velocity (at the left) and of the coordinates (at the right) for various values of the Levy index α. The friction coefficient ν is 0.03.

moments are obtained by averaging over 50 realizations, each of length 4096. The moment exponent q is equal to 0.25 for the velocity and of the coordinate, the Levy index is equal to 1.3. The moments estimated by numerically solving the Langevin equations are shown by black dots, whereas the moments estimated with Eqs. (5.5) and (4.7) are shown by solid lines. The vertical mark indicates the relaxation time τ_v. At the intervals greater than τ_v, the random process $v(t)$ becomes stationary and the velocity moment tends to the constant value $D/\alpha\nu$. At the same time, it follows from the right-hand figures that the process $x(t)$ remains non-stationary, and the moment of the coordinate tends to a linear (in a twice logarithmic scale) asymptotic regime, which has a slope q/α and is shown by the dotted line in the right-hand figures. From Fig. 5, we can make conclusions about the agreement between the theoretical results obtained for the force-free relaxation by solving FSFPE and the results obtained by numerically solving the Langevin equations.

4.2. Relaxation of the Linear Levy Oscillator

Setting $F = -\omega^2 x$ in Eq. (2.12), we seek the solution of FSFPE with the initial condition $f(x, v, t = 0) = \delta(x - x_0)\delta(v - v_0)$.

Making the Fourier transform and using the method of characteristics, we obtain the characteristic function

$$\hat{f}(\kappa, k, t) = \exp\left\{ i\omega^2 x_0 e^{-\nu t/2}\left[-\frac{2k}{\nu_1}\sinh\frac{\nu_1 t}{2} + \frac{\kappa}{\nu_1\omega^2}\left(\nu\sinh\frac{\nu_1 t}{2} + \nu_1\cosh\frac{\nu_1 t}{2}\right)\right]\right.$$

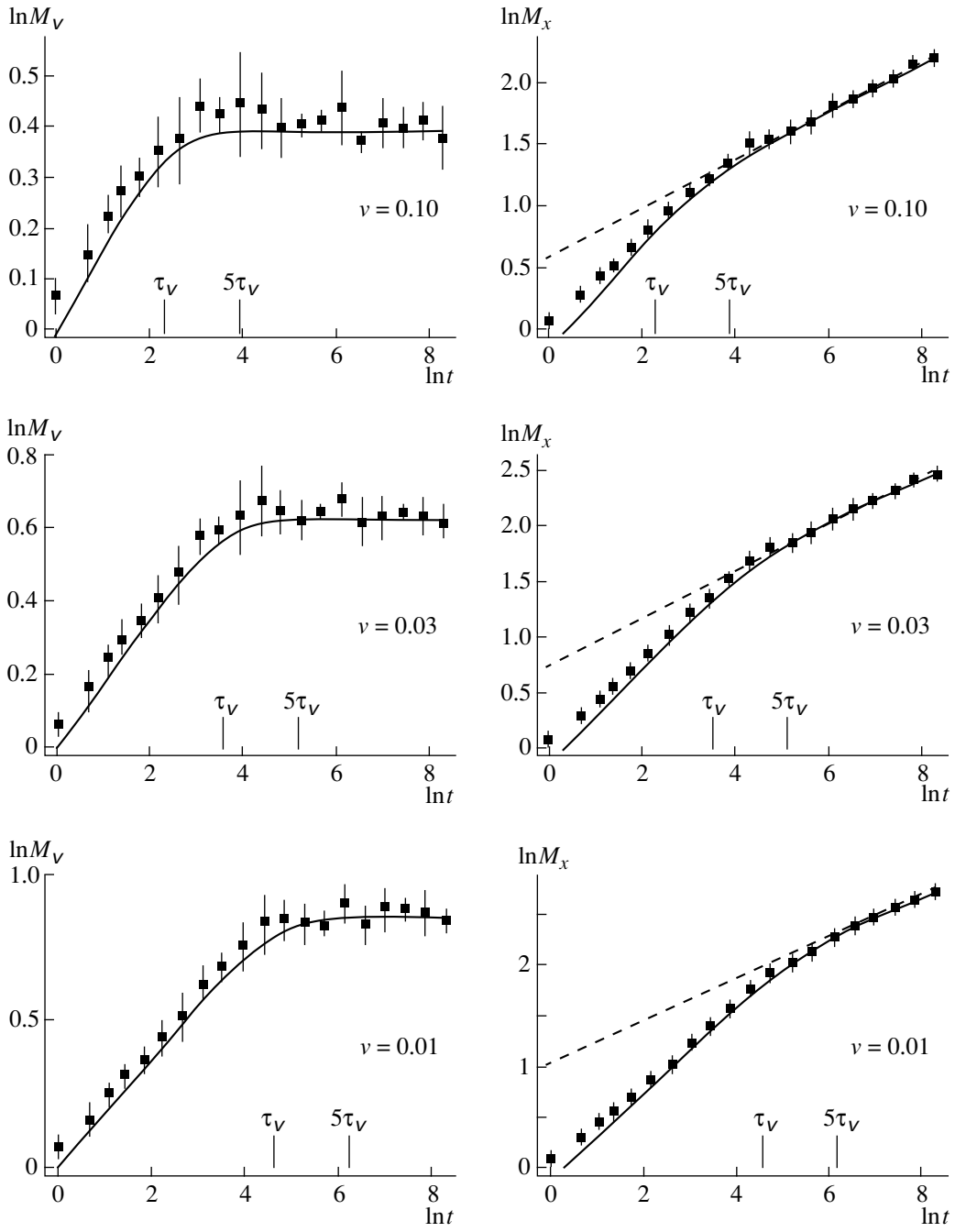

Fig. 5. Force-free relaxation in the FSFPE framework. The velocity moments (at the left) and the coordinate moments (at the right) versus time in a twice logarithmic scale. Black dots indicate the numerical simulation results for the Langevin equations, the solid line indicates the moments obtained from the FSFPE. The moment order q is 0.25. The Levy index α is 1.3. Vertical marks show the velocity relaxation time τ_v and the time $5\tau_v$. The dotted lines in the right-hand figures indicate the theoretical values q/α of the straight-line asymptotic regime (in a log-log scale) for the moment of the coordinate.

$$+ iv_0 e^{-\nu t/2}\left[\frac{2\kappa}{\nu_1}\sinh\frac{\nu_1 t}{2} - \frac{k}{\nu_1}\left(\nu\sinh\frac{\nu_1 t}{2} - \nu_1\cosh\frac{\nu_1 t}{2}\right)\right] - D\int_0^t d\tau e^{-\alpha\nu\tau/2} \times \left|\frac{2\kappa}{\nu_1}\sinh\frac{\nu_1\tau}{2} + k\left(\cosh\frac{\nu_1\tau}{2} - \frac{\nu}{\nu_1}\sinh\frac{\nu_1\tau}{2}\right)\right|^{\alpha}\Bigg\}, \quad (4.8)$$

where $\nu_1 = (\nu^2 - 4\omega^2)^{1/2}$. This expression is valid for $\nu^2 - 4\omega^2 > 0$. If $\nu^2 - 4\omega^2 < 0$, we introduce $\omega_1 = \sqrt{\omega^2 - \nu^2/4}$

and make the following changes:

$$\cosh(\nu_1 t/2) \longrightarrow \cos(\omega_1 t),$$

$$\sinh(\nu_1 t/2)/\nu_1 \longrightarrow \sin(\omega_1 t)/2\omega_1,$$

$$\sinh(\nu_1 t)/\nu_1 \longrightarrow \sin(2\omega_1 t)/2\omega_1.$$

In the aperiodic case, $\nu = 4\omega^2$, evidently,

$$\cosh(\nu_1 t/2) \longrightarrow 1, \quad \sinh(\nu_1 t/2)/\nu_1 \longrightarrow t/2,$$

$$\sinh(\nu_1 t)/\nu_1 \longrightarrow t.$$

Setting $\alpha = 2$ and using the first and the second derivatives of the characteristic function at $\kappa = k = 0$, we can obtain the means and the variances of the velocity and coordinates for the Brownian oscillator. These formulas are given in [4]. The expressions for the means are also valid for all α greater than unity.

We now turn to more complicated fractional moments,

$$\begin{aligned} M_v(q) &= \langle |v|^q \rangle = (D_{\mathrm{osc}}^{(v)}(t))^{q/\alpha} C(q;\alpha), \\ M_x(q) &= \langle |x|^q \rangle = (D_{\mathrm{osc}}^{(x)}(t))^{q/\alpha} C(q;\alpha), \end{aligned} \tag{4.9}$$

where

$$D_{\mathrm{osc}}^{(v)}(t) = D\int_0^t d\tau \exp\left(-\frac{\alpha\nu\tau}{2}\right)\left|\cosh\frac{\nu_1\tau}{2} - \frac{\nu}{\nu_1}\sinh\frac{\nu_1\tau}{2}\right|^\alpha,$$

$$D_{\mathrm{osc}}^{(x)}(t) = \frac{D}{\nu^\alpha}\int_0^t d\tau \exp\left(-\frac{\alpha\nu\tau}{2}\right) \times \left[\exp\left(\frac{\nu_1\tau}{2}\right) - \exp\left(-\frac{\nu_1\tau}{2}\right)\right]^\alpha,$$

and $C(q;\alpha)$ is the same as in previous sections. Equations (4.9) are compared with the results of numerical simulations at the end of Section 5.

Equation (4.8), in principle, allows one to study the stationary solution, which is defined by

$$f_{st}(x, v) = \int_{-\infty}^{\infty}\frac{d\kappa}{2\pi} \times \int_{-\infty}^{\infty}\frac{dk}{2\pi}\exp(-i\kappa x - ikv)\hat{f}_{st}(\kappa, k), \tag{4.10}$$

where

$$\hat{f}_{\mathrm{st}}(\kappa, k) = \exp\left\{-D\int_0^\infty d\tau e^{-\alpha\nu\tau/2} \times \left|\frac{2\kappa}{\nu_1}\sinh\left(\frac{\nu_1\tau}{2}\right) + k\left[\cosh\left(\frac{\nu_1\tau}{2}\right) - \frac{\nu}{\nu_1}\sinh\left(\frac{\nu_1\tau}{2}\right)\right]\right|^\alpha\right\}.$$

A simple analytic expression can be obtained for the stationary solution in the case where $\alpha = 2$ only. However, Eq. (4.10) allows one to derive some conclusions for simpler stationary PDFs, namely,

$$f_{\mathrm{st}}(x) = \int_{-\infty}^{\infty} dv\, f_{\mathrm{st}}(x, v) = \int_{-\infty}^{\infty}\frac{d\kappa}{2\pi}e^{-i\kappa x}\hat{f}_{\mathrm{st}}(\kappa, k = 0)$$

and

$$f_{\mathrm{st}}(v) = \int_{-\infty}^{\infty} dx\, f_{\mathrm{st}}(x, v) = \int_{-\infty}^{\infty}\frac{d\kappa}{2\pi}e^{-ikv}\hat{f}_{\mathrm{st}}(\kappa = 0, k).$$

Both stationary PDFs are stable with the Levy index α and with the scale parameters expressed as integrals over τ, see Eq. (4.9).

Although the linear oscillator, as we see, admits an exact solution, the general formulas are not easy to analyze analytically. Therefore, it is instructive to consider two limiting cases, namely, an overdamped and a weakly damped oscillator. Both cases are also very important in problems related to the nonlinear oscillator.

5. LIMIT CASES OF THE LEVY OSCILLATOR

5.1. The Overdamped Levy Oscillator, $\omega/\nu \ll 1$

We consider the relaxation of the moments for an overdamped oscillator. We first turn to the velocity relaxation. It follows from Eq. (4.9) that we can restrict ourselves by the zero-order approximation in ω/ν. We then obtain $D_{osc}^{(v)}(t)$ and τ_v that obviously coincide with $D_{ff}^{(v)}(t)$ and τ_v or the force-free case. The conclusion is that the velocity relaxation for the overdamped oscillator in the leading order in the small parameter ω/ν is the same as in the force-free case.

We next consider space relaxation, which differs from the force-free case. We obtain from Eq.(4.9) in the first order in ω/ν,

$$D_{\mathrm{osc}}^{(x)}(t) \approx \frac{D}{\nu^\alpha}\int_0^t d\tau (e^{-\omega^2\tau/\nu} - e^{-\nu\tau})^\alpha. \tag{5.1}$$

It follows from Eq. (5.1) that for $t \gg \tau_v = 1/\alpha\nu$, the second term in the brackets gives a negligibly small contribution, and therefore,

$$D_{\mathrm{osc}}^{(x)}(t \gg \tau_v) \approx \frac{D}{\alpha\omega^2\nu^{\alpha-1}}(1 - e^{-\alpha\omega^2\tau/\nu}), \tag{5.2}$$

which coincides with the result obtained in the FSESE framework. Thus, we conclude that at time intervals greater than the velocity relaxation τ_v, the overdamped oscillator can be described with the help of a simpler kinetic equation, namely, the FSESE. For the overdamped oscillator, the relaxation process occurs in two stages: the fast velocity relaxation stage, at which a sta-

tionary stable velocity PDF is established during the time interval τ_v, and the slow diffusion stage, at which the stable PDF is established in real space during the time interval $\tau_x = \nu/\alpha\omega^2$.

5.2. The Weakly Damped Levy Oscillator, $\omega/\nu \gg 1$

In this case, the theory of the Brownian oscillator provides us with the method of simplifying the kinetic description [2]. It is based on the method of slowly varying amplitudes, or the van-der-Pol method, which is frequently used, e.g., in radiophysics [42, 43]. In this approach, the solution of the Langevin equations

$$\frac{dx}{dt} = v, \quad \frac{dv}{dt} = -\omega^2 x - \nu v + A(t) \tag{5.3}$$

is sought in the form

$$x = \tilde{x}\cos(\omega t) + \frac{\tilde{v}}{\omega}\sin(\omega t),$$
$$v = \tilde{v}\cos(\omega t) - \omega\tilde{x}\sin(\omega t),$$

where the amplitudes $\tilde{x}$ and $\tilde{v}$ are slowly varying during the period $2\pi/\omega$. The evolution equation for the PDF $f(\tilde{x}, \tilde{v}, t)$ is given by

$$\frac{\partial f}{\partial t} = \frac{\nu}{2}\frac{\partial}{\partial\tilde{x}}(\tilde{x} f) + \frac{\nu}{2}\frac{\partial}{\partial\tilde{v}}(\tilde{v} f) + D_{\tilde{x}}\frac{\partial^\alpha f}{\partial|\tilde{x}|^\alpha} + D_{\tilde{v}}\frac{\partial^\alpha f}{\partial|\tilde{v}|^\alpha}, \tag{5.4}$$

where

$$D_{\tilde{x}} = D/(2\omega)^\alpha, \quad D_{\tilde{v}} = D/2^\alpha.$$

The detailed derivation of Eq. (5.4) is presented in the Appendix and is analogous to that of FSFPE. If, as usual, the initial condition is

$$f(\tilde{x}, \tilde{v}, t = 0) = \delta(\tilde{x} - x_0)\delta(\tilde{v} - v_0),$$

the solution for the characteristic function is

$$\hat{f}(\kappa, k, t) = \exp\left\{ i\kappa x_0 e^{-\nu t/2} + ikv_0 e^{-\nu t/2} - \frac{2D_{\tilde{x}}}{\alpha\nu}|\kappa|^\alpha(1 - e^{-\alpha\nu t/2}) - \frac{2D_{\tilde{v}}}{\alpha\nu}|k|^\alpha(1 - e^{-\alpha\nu t/2})\right\}. \tag{5.5}$$

We also obtain the fractional moments

$$M_{\tilde{v}}(t; q, \alpha) = (D_{\rm osc}^{(\tilde{v})}(t))^{q/\alpha} C(q; \alpha),$$
$$M_{\tilde{x}}(t; q, \alpha) = (D_{\rm osc}^{(\tilde{x})}(t))^{q/\alpha} C(q; \alpha), \tag{5.6}$$

where

$$D_{osc}^{(\tilde{v})} = \frac{2D_{\tilde{v}}}{\alpha\nu}(1 - e^{-\alpha\nu t/2}),$$

$$D_{\rm osc}^{(\tilde{x})} = \frac{2D_{\tilde{x}}}{\alpha\nu}(1 - e^{-\alpha\nu t/2}).$$

It follows from the last equations that unlike in the case of the overdamped Levy oscillator, the coordinate and the velocity of the weakly damped oscillator possess equal relaxation times,

$$\tau_v = \tau_x = \tau = 2/\alpha\nu.$$

For the weakly damped oscillator, it is therefore impossible to divide the relaxation process into the fast and slow stages. At time intervals greater than τ, the random processes $\tilde{x}(t)$ and $\tilde{v}(t)$ become stationary with stable PDFs. The characteristic function of the stationary state is determined from Eq. (5.5) as

$$\hat{f}_{\rm st}(\kappa, k) = \exp\left(-\frac{2D_{\tilde{x}}}{\alpha\nu}|\kappa|^\alpha - \frac{2D_{\tilde{v}}}{\alpha\nu}|k|^\alpha\right),$$

and the PDF retains the Maxwell–Boltzmann form for $\alpha = 2$ only.

Numerical simulations of the linear oscillator relaxation involve solving Langevin equations (2.2) with the external force $F = -\omega^2 x$ and subsequently calculating the velocity and coordinate moments. The results are shown in Figs. 6–8.

Figure 6 has an illustrative character. It shows the typical trajectories of the velocity (at the top) and of the coordinate (at the-bottom) for the overdamped oscillator (at the left) and for the weakly damped oscillator (at the right), respectively. The frequency value is equal to 0.003 and 0.3 for the overdamped and weakly damped oscillators, respectively. The friction coefficient is equal to 0.03 and the Levy index is 1.3. The trajectories shown in the figures have a single large outlier. This allows us to visually demonstrate the difference in the behavior of the two kinds of oscillators: the relaxation process for the overdamped oscillator (at the left) resembles the force-free relaxation and is radically different from the rapidly oscillating behavior of the velocity and coordinates of the weakly damped oscillator (at the right).

In Fig. 7, we show the velocity moments M_v (at the top) and coordinate moments M_x (at the bottom) versus t for the overdamped oscillator (at the left) and for the weakly damped oscillator (at the right), respectively. The oscillation frequencies ω are equal to 0.01 and 0.1 for the overdamped and the weakly damped oscillators, respectively, the friction coefficient is 0.03. The order of the moment is 0.25 and the Levy index is 1.3. The moments obtained by numerical simulation are shown by black dots, whereas the theoretical values, see Eq. (4.9), are shown by a solid line. The numerical values are obtained by averaging over 200 realizations, each of length 1024. The vertical marks indicate the velocity relaxation time $\tau_v = 1/\alpha\nu$ and the coordinate relaxation time $\tau_x = \nu/\alpha\omega^2$ for the overdamped oscillator, and also the relaxation time $\tau_v = \tau_x = 2/\alpha\nu$ for the

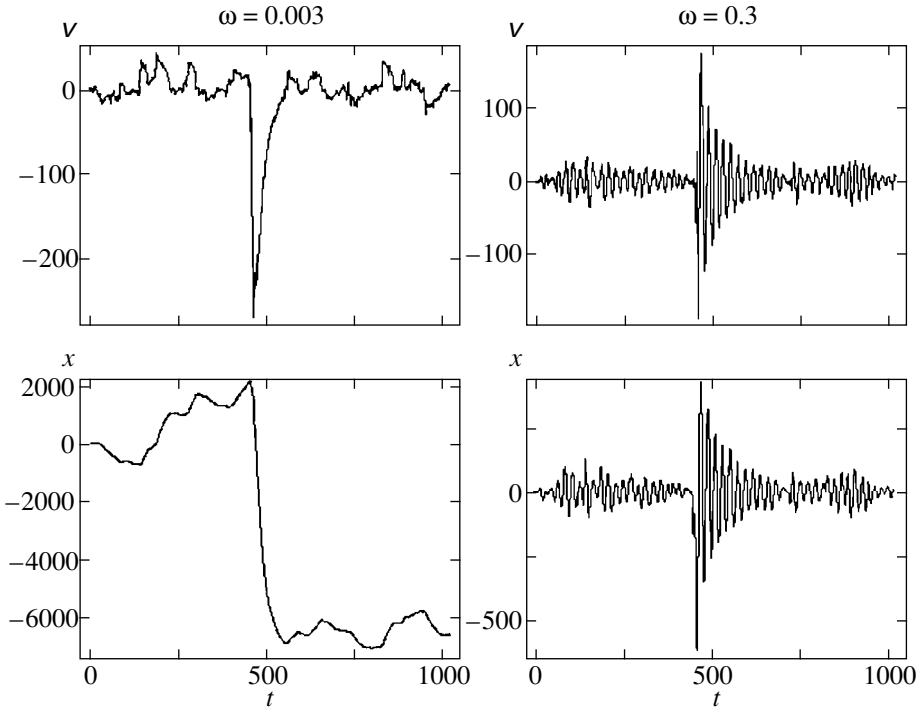

Fig. 6. Relaxation of the linear oscillator in the FSFPE framework. Typical trajectories of the velocity (top) and the coordinate (bottom) for the overdamped (left) and the weakly damped (right) oscillators. The frequencies are 0.003 (overdamped) and 0.3 (weakly damped), the friction coefficient is 0.03, and the Levy index is 1.3.

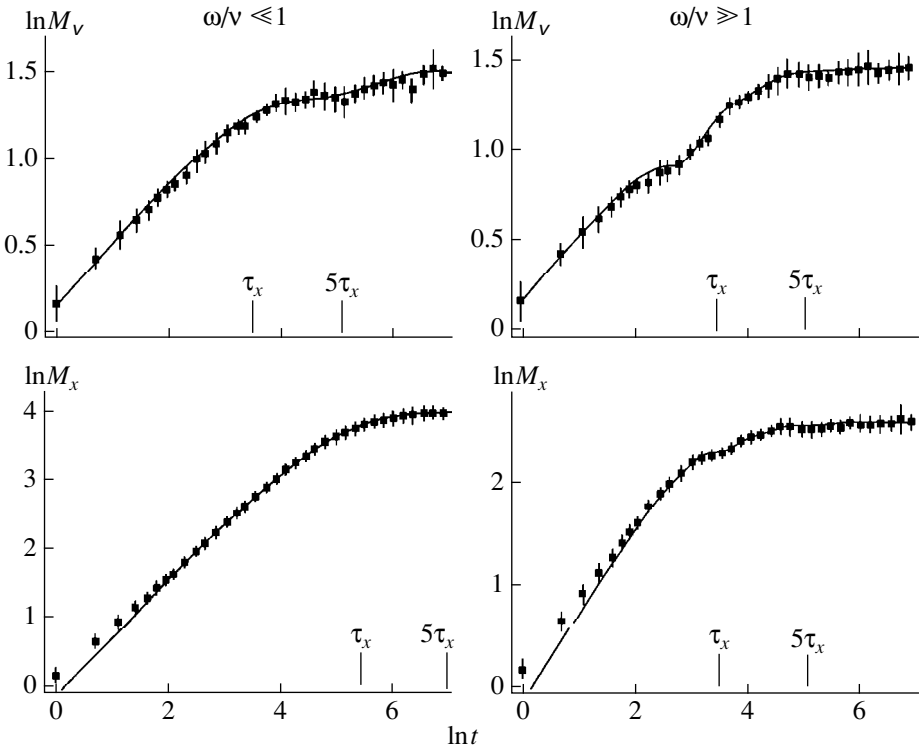

Fig. 7. Relaxation of the linear oscillator in the FSFPE framework. The velocity moments (at the top) and the coordinate moments (at the bottom) versus time in a twice logarithmic scale. At the left, the results are for the overdamped oscillator are presented, with ω = 0.01, and at the right, the results are for the weakly damped oscillator, with ω = 0.1. The frequency coefficient ν is 0.03, the order q of the moments is 0.25, and the Levy index α is 1.3. Black dots indicate the numerical simulation results of the Langevin equations, the solid lines indicate the moments obtained from the FSFPE, see Eq. (4.9). The arrows show velocity and coordinate relaxation times.

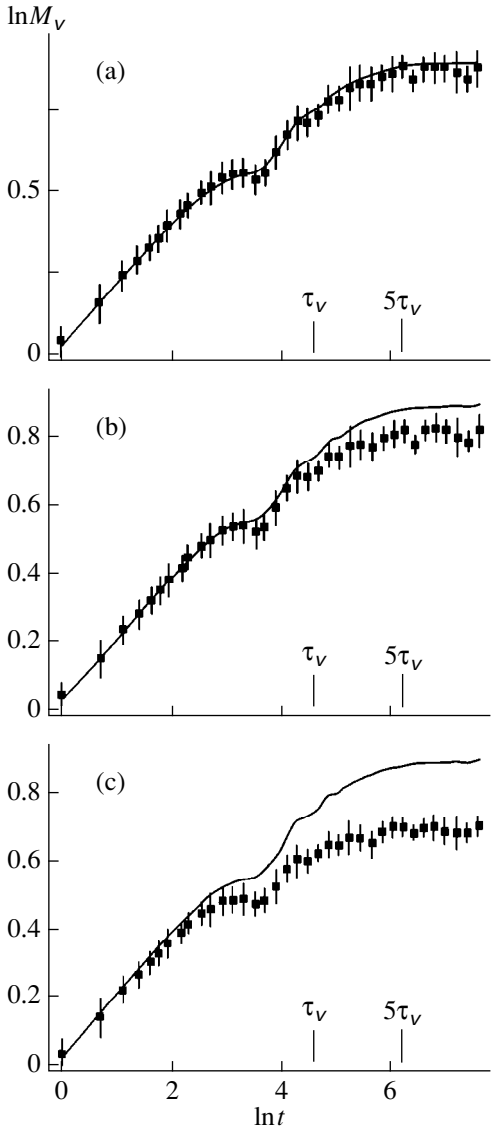

Fig. 8. Relaxation of the linear oscillator in the FSFPE framework. The velocity moments versus time in a twice logarithmic scale for different truncation parameters $A_{\max}$ for the stable PDFs of the noise in the Langevin equations. The parameters used in simulations are as follows: $\omega = 0.05$, $\nu = 0.01$, $q = 0.25$, and $\alpha = 1.3$. $A_{\max}$ is 1600, 300, and 50 for the cases (a), (b), and (c), respectively.

weakly damped oscillator. Both theoretical and numerical curves reach a "plateau" at intervals greater than the relaxation times. This implies that the processes $v(t)$ and $x(t)$ become stationary. The figures demonstrate an important difference between the overdamped and the weakly damped oscillators: for the overdamped oscillator, the coordinate relaxation time is much greater than the velocity relaxation time, whereas for the weakly damped oscillator both times are the same. From Fig. 7, we also conclude that there is close quantitative agreement between theoretical results obtained by solving FSFPE and the numerical solution of the corresponding Langevin equations. In this respect, it is worthwhile to draw one's attention to the coincidence of the theoretical curve with numerical dependence in the nonstationary parts.

To estimate the influence of the power-law asymptotic behavior on the evolution of the moments, we replace the stable PDFs in the Langevin equation by the "truncated" ones, in which large values of random quantities are cut off. In Fig. 8, the velocity moments versus time are depicted for the weakly damped oscillator in a log-log scale. The oscillator frequency is 0.05, the friction coefficient is 0.01, the order of the moment is 0.25, and the Levy index is 1.3. The moments are obtained by averaging over 1500 realizations, each of length 2048, thus the total number of points is 3×10^6. The mode with the maximum value (that is, the most probable value) is of the order $N^{1/\alpha}$. In the figure, the moments obtained numerically are shown by black dots. The Langevin source $A(t)$ is modeled as the effect of independent random variables possessing a truncated stable PDF, that is $|A(t)| < A_{\max} = 1600, 300, 50$ for the respective variants (a), (b), and (c). The solid lines indicate the moments estimated analytically from the FSFPE. It is seen that the role of large outliers increases as time increases, because large values become more and more probable. Therefore, the discrepancy between theoretical results and numerical simulations using the truncated PDF grows with time. As the truncation parameter decreases, the discrepancy increases. Thus, it is clearly seen that the discrepancies are most essential at the stationary stage of the evolution.

We have already mentioned that our studies demonstrate a good quantitative agreement between the theory based on the FSFPE and the numerical simulations based on the Langevin equations. To show this fact more precisely, we estimated the velocity and coordinate moments by averaging over 50×10^3 realizations with the total number of points 10^8, which is much larger than in simulations presented in Fig. 7. It is seen that the numerical results strictly repeat all the bends of the theoretical curves at the nonstationary stage of the evolution.

6. RESULTS

The main results are as follows.

(1) We have obtained the fractional symmetric Fokker–Planck equation and the fractional symmetric Einstein–Smoluchowski equation. These equations generalize the Fokker–Planck and Einstein–Smoluchowski equations for Brownian motion. The FSFPE describes a linear relaxation in the phase space of systems influenced by stochastic forces distributed with symmetric stable laws. The FSFPE contains a fractional velocity derivative instead of a second-order derivative. The FSESE describes relaxation in real space. It contains a fractional space derivative.

(2) Using the kinetic equations obtained, we considered the linear relaxation processes for two problems:

the force-free relaxation and the relaxation of the linear oscillator. We have obtained general analytic solutions of the FSFPE and FSESE and the expressions for the fractional velocity and coordinate moments.

(3) For both problems, we solved numerically the Langevin equations with random sources that are discrete approximations to white Levy noises. After averaging many realizations, we estimated the fractional moments and compared the numerical results with the results of analytical solutions to the kinetic equations. The analytical and numerical results appear to be in a quantitative agreement.

(4) In studying the relaxation of the linear oscillator, it is expedient to distinguish between the two variants: the overdamped oscillator with $\omega/\nu \ll 1$ and the weakly damped oscillator with $\omega/\nu \gg 1$. Both cases are of special importance in nonlinear generalizations of the theory presented. We study in detail, both analytically and numerically, each of the limiting cases, and point out substantially different properties of the relaxation processes in the two cases.

(5) The relaxation of an overdamped oscillator occurs in two stages, which are described in the FSFPE framework: the "fast" stage, at which a stationary stable PDF over velocity is established during the time interval $\tau_v = 1/\alpha\nu$ and the "slow" diffusion stage, at which a stationary stable PDF in the real space is established during the time interval $\tau_x = \nu/\alpha\omega^2$. At the diffusion stage, the relaxation of the overdamped oscillator can be described in the FSESE framework.

(6) For the weakly damped Levy oscillator, we have derived a kinetic equation for the PDF depending on slowly varying (at the period of oscillations) variables. This equation contains fractional derivatives over the velocity and the coordinate. The relaxation process cannot be divided into two stages. On the contrary, the velocity and the coordinate relax during the same time interval $\tau = 2/\alpha\nu$.

ACKNOWLEDGMENTS

The authors thank F. Mainardi (Bologna University) for valuable comments on problems related to fractional calculus and diffusion-like equations. This paper was supported by the Project "Chaos-2" of National Academy of Science of the Ukraine and by the INTAS, project no. 98-01.

APPENDIX

Kinetic Equation for the Weakly Damped Levy Oscillator

We start from Eqs. (5.3) and look for the solution in the form

$$x = \tilde{x}\cos(\omega t) + \frac{\tilde{v}}{\omega}\sin(\omega t), \qquad v = \tilde{v}\cos(\omega t) + \omega\tilde{x}\sin(\omega t), \tag{A.1}$$

where the amplitudes $\tilde{x}$ and $\tilde{v}$ are slowly varying over the period $2\pi/\omega$. The choice of solution (A.1) is equivalent to the condition

$$\cos(\omega t)\frac{d\tilde{x}}{dt} + \frac{\sin(\omega t)}{\omega}\frac{d\tilde{v}}{dt} = 0.$$

We insert (A.1) in (5.3) and, after averaging the period, obtain the Langevin equations for $\tilde{x}$ and $\tilde{v}$,

$$\frac{d\tilde{x}}{dt} + \frac{\nu}{2}\tilde{x} = A_{\tilde{x}}(t), \quad \frac{d\tilde{v}}{dt} + \frac{\nu}{2}\tilde{v} = A_{\tilde{v}}(t), \tag{A.2}$$

where

$$A_{\tilde{x}}(t) = -\frac{1}{2\pi}\int_{t-2\pi/\omega}^{t} dt' A(t')\sin(\omega t'), \qquad A_{\tilde{v}}(t) = \frac{\omega}{2\pi}\int_{t-2\pi/\omega}^{t} dt' A(t')\cos(\omega t'). \tag{A.3}$$

It follows from Eqs. (A.2) that the Langevin sources $A_{\tilde{x}}(t)$ and $A_{\tilde{v}}(t)$ do not contain the "fast" time $2\pi/\omega$. Therefore, it follows from Eqs. (A.3) that $A(t)$ can be represented as

$$A(t) \approx a(t)\cos(\omega t) - b(t)\sin(\omega t), \tag{A.4}$$

where $a(t)$ and $b(t)$ are random stationary functions that are related to $A_{\tilde{x}}(t)$ and $A_{\tilde{v}}(t)$ as

$$a(t) = 2A_{\tilde{v}}(t), \quad b(t) = -2\omega A_{\tilde{x}}(t). \tag{A.5}$$

Equations (A.4) and (A.5) have the following meaning [42]. According to Eqs. (A.2) and (A.3), the random force influences the oscillator by means of slowly varying components $A_{\tilde{x}}(t)$ and $A_{\tilde{v}}(t)$ [or $a(t)$ and $b(t)$] only. Therefore, if one considers the random influence on the weakly damped oscillator, the main components of the random force are singled out by Eq. (A.4). Furthermore, if $A(t)$ is a stationary Gaussian process, then the expression for the correlation function of this process shows that the one-point PDFs of $A(t)$, $a(t)$, and $b(t)$ coincide [43] with each other. We assume that the conclusion about identical PDFs of $A(t)$, $a(t)$, and $b(t)$ is also valid for (symmetric) stable PDFs, even though the proof of this statement is not so trivial as in the Gaussian case. It follows from this coincidence that the processes

$$L(\Delta t) = \int_{t}^{t+\Delta t} dt' A(t'), \qquad L_a(\Delta t) = \int_{t}^{t+\Delta t} dt' a(t'), \qquad L_b(\Delta t) = \int_{t}^{t+\Delta t} dt' b(t') \tag{A.6}$$

also have identical PDFs, and thus, with Eq. (2.5) taken into account, the PDFs of L_a and L_b are given by

$$w(L_{a,b}) = \int_{-\infty}^{\infty} \frac{dk}{2\pi} \exp(-ikL_{a,b} - D|k|^{\alpha}\Delta t). \quad \text{(A.7)}$$

We also define the processes

$$\begin{aligned} L_{\tilde{v}}(\Delta t) &= \int_{t}^{t+\Delta t} dt' A_{\tilde{v}}(t'), \\ L_{\tilde{x}}(\Delta t) &= \int_{t}^{t+\Delta t} dt' A_{\tilde{x}}(t') \end{aligned} \quad \text{(A.8)}$$

which in accordance with (A.5) are related to L_a and L_b as

$$L_a(\Delta t) = 2L_{\tilde{v}}(\Delta t), \quad L_b(\Delta t) = -2\omega L_{\tilde{x}}(\Delta t). \quad \text{(A.9)}$$

Now, with the help of Eqs. (A.7) and (A.9), we are able to obtain the characteristic functions $\hat{w}(L_{\tilde{x}})$ and $\hat{w}(L_{\tilde{v}})$ and their PDFs $w(L_{\tilde{x}})$ and $w(L_{\tilde{v}})$,

$$\begin{aligned} \hat{w}(L_{\tilde{x}}) &= \exp(-D_{\tilde{x}}|k|^{\alpha}\Delta t), \\ \hat{w}(L_{\tilde{v}}) &= \exp(-D_{\tilde{v}}|k|^{\alpha}\Delta t), \end{aligned} \quad \text{(A.10)}$$

where

$$D_{\tilde{x}} = \frac{D}{(2\omega)^{\alpha}}, \quad D_{\tilde{v}} = \frac{D}{2^{\alpha}}. \quad \text{(A.11)}$$

The equation for $f(\tilde{x}, \tilde{v}, t)$ is derived similarly to the FSESE. The initial equation is

$$\begin{aligned} &f(\tilde{x}, \tilde{v}, t+\Delta t) \\ &= \iint d(\Delta\tilde{x})d(\Delta\tilde{v})\, f(\tilde{x}-\Delta\tilde{x}, v-\Delta\tilde{v}, t) \\ &\quad \times \Psi(\tilde{x}-\Delta\tilde{x}, v-\Delta\tilde{v}; \Delta x, \Delta\tilde{v}, \Delta t), \end{aligned} \quad \text{(A.12)}$$

where Ψ is the transition probability. For the increments $\Delta\tilde{x}$ and $\Delta\tilde{v}$, we obtain from Eqs. (A.2)

$$\begin{aligned} \Delta\tilde{x} + \frac{\nu}{2}\tilde{x}\Delta t &= L_{\tilde{x}}(\Delta t), \\ \Delta\tilde{v} + \frac{\nu}{2}\tilde{v}\Delta t &= L_{\tilde{v}}(\Delta t), \end{aligned} \quad \text{(A.13)}$$

where the PDFs for $L_{\tilde{x}}$ and $L_{\tilde{v}}$ are given by Eqs. (A.10) and (A.11). We now construct Ψ. From the structure of Eqs. (A.13), it follows that

$$\begin{aligned} &\Psi(\tilde{x}-\Delta\tilde{x}, v-\Delta\tilde{v}; \Delta\tilde{x}, \Delta\tilde{v}, \Delta t) \\ &= \Psi_{\tilde{x}}(\tilde{x}-\Delta\tilde{x}; \Delta\tilde{x}, \Delta t)\Psi_{\tilde{v}}(\tilde{v}-\Delta\tilde{v}; \Delta\tilde{v}, \Delta t), \end{aligned} \quad \text{(A.14)}$$

where

$$\begin{aligned} \Psi_{\tilde{x}}(\tilde{x}; \Delta\tilde{x}, \Delta t) &= \int_{-\infty}^{\infty} \frac{d\kappa}{2\pi} \\ &\times \exp\left(-i\kappa\left(\Delta\tilde{x} + \frac{\nu}{2}\tilde{x}\Delta t\right) - D_{\tilde{x}}|\kappa|^{\alpha}\Delta t\right), \\ \Psi_{\tilde{v}}(\tilde{v}; \Delta\tilde{v}, \Delta t) &= \int_{-\infty}^{\infty} \frac{dk}{2\pi} \\ &\times \exp\left(-ik\left(\Delta\tilde{v} + \frac{\nu}{2}\tilde{v}\Delta t\right) - D_{\tilde{v}}|k|^{\alpha}\Delta t\right). \end{aligned} \quad \text{(A.15)}$$

Inserting Eqs. (A.14) and (A.15) in Eq. (A.12), expanding into a power series in Δt and letting Δt go to zero, we arrive at

$$\begin{aligned} \frac{\partial f}{\partial t} &= -\iint d(\Delta\tilde{x})d(\Delta\tilde{v})\, f(\tilde{x}-\Delta\tilde{x}, \tilde{v}-\Delta\tilde{v}, t) \\ &\quad \times \int_{-\infty}^{\infty} \frac{d\kappa}{2\pi} \int_{-\infty}^{\infty} \frac{dk}{2\pi} \exp(-i\kappa\Delta\tilde{x} - ik\Delta\tilde{v}) \\ &\quad \times \left[i\kappa\frac{\nu}{2}(\tilde{x}-\Delta\tilde{x}) + ik\frac{\nu}{2}(\tilde{v}-\Delta\tilde{v}) + D_{\tilde{x}}|\kappa|^{\alpha} + D_{\tilde{v}}|k|^{\alpha} \right]. \end{aligned} \quad \text{(A.16)}$$

Transforming the terms in the right-hand side as described in Section 2, we arrive at the differential equation for $f(\tilde{x}, \tilde{v}, t)$, see (5.4).

REFERENCES

1. N. G. van Kampen, *Stochastic Processes in Physics and Chemistry* (North Holland, Amsterdam, 1981; Vyschaya Shkola, Moscow, 1986).
2. Yu. L. Klimontovich, *Statistical Physics* (Nauka, Moscow, 1982; Harwood Academic, New York, 1986).
3. Yu. L. Klimontovich, *Statistical Physics in Open Systems* (Moscow, Yanus, 1995; Kluwer, Dordrecht, 1994).
4. S. Chandrasekhar, Rev. Mod. Phys. **15**, 1 (1943); *Stochastic Problems in Physics and Astronomy* (Inostrannaya Literatura, Moscow, 1947).
5. J. Klafter, A. Blumen, and M. F. Shlesinger, Phys. Rev. A **35**, 3081 (1987).
6. J.-P. Bouchaud and A. Georges, Phys. Rep. **195**, 127 (1990).
7. G. M. Zaslavsky, Physica D (Amsterdam) **76**, 110 (1994).
8. G. Zimbardo, P. Veltri, G. Basile, and S. Principato, Phys. Plasmas **2**, 2653 (1995).
9. B. J. West and W. Deering, Phys. Rep. **246**, 1 (1994).
10. J.-P. Bouchaud and M. Potters, *Theorie des Risques Financiers* (Alea-Saclay-Eyrolles, Paris, 1997); http://www.science-finance.fr/.
11. P. Levy, *Theorie de l'Addition des Variables* (Gauthier-Villiers, Paris, 1937); A. Y. Khintchine and P. Levy, C. R. Hebd. Seances Acad. Sci. **202**, 374 (1936).

12. B. V. Gnedenko and A. N. Kolmogorov, *Limit Distributions for Sums of Independent Random Variables* (Tekhniko-Teor. Literatura, Moscow, 1949; Addison-Wesley, Reading, 1954).
13. S. G. Samko, A. A. Kilbas, and O. I. Marichev, *Fractional Integrals and Derivatives, Theory and Applications* (Nauka i Technika, Minsk, 1987; Gordon and Breach, Amsterdam, 1993).
14. Yu. I. Babenko, *Heat and Mass Transfer* (Khimiya, Leningrad, 1986).
15. W. Wyss, J. Math. Phys. **27**, 2782 (1986).
16. W. R. Schneider and W. Wyss, J. Math. Phys. **30**, 134 (1989).
17. T. F. Nonnenmacher, J. Phys. A **23**, L697 (1990).
18. G. Jumarie, J. Math. Phys. **33**, 3536 (1992).
19. A. I. Saichev and G. M. Zaslavsky, Chaos **7**, 753 (1997).
20. R. Gorenffo, G. De Fabritiis, and F. Mainardi, Physica A (Amsterdam) **269**, 79 (1999).
21. M. Giona and H. E. Roman, J. Phys. A **25**, 2093 (1992).
22. H. E. Roman and M. Giona, J. Phys. A **25**, 2107 (1992).
23. G. M. Zaslavsky, M. Edelman, and B. A. Niyazov, Chaos **7**, 159 (1997).
24. F. Mainardi, Chaos, Solitons and Fractals **7**, 1461 (1996).
25. F. E. Peseckis, Phys. Rev. A **36**, 892 (1987).
26. B. J. West and V. Seshadri, Physica A (Amsterdam) **113**, 203 (1982).
27. S. Jespersen, R. Metzler, and H. C. Fogedby, Phys. Rev. E **59**, 2736 (1999).
28. H. C. Fogedby, Phys. Rev. E **58**, 1690 (1998).
29. R. L. Stratonovich, *Topics on the Theory of Random Noise* (Gordon and Breach, New York, 1963), Vol. 1.
30. A. V. Skorokhod, *Random Processes with Independent Increments* (Nauka, Moscow, 1964; Kluwer, Dordrecht, 1991).
31. C. Fox, Trans. Am. Math. Soc. **98**, 395 (1961).
32. W. R. Schneider, in *Stochastic Processes in Classical and Quantum Systems*, Ed. by S. Albeverio, G. Casati, and D. Merlini (Springer-Verlag, Berlin, 1986), Lecture Notes in Physics, p. 497.
33. A. M. Mathai and R. K. Saxena, *The H-functions with Applications in Statistics and Other Disciplines* (Wiley, New Delhi, 1978).
34. H. M. Srivastava, K. G. Gupta, and S. P. Goyal, *The H-functions of One and Two Variables with Applications* (South Asian, New Delhi, 1982).
35. A. V. Chechkin, D. Schertzer, A. V. Tur, and V. V. Yanovsky, Ukr. Phys. J. **40**, 434 (1995).
36. V. M. Zolotarev, V. V. Uchaikin, and V. V. Sayenko, Zh. Éksp. Teor. Fiz. **114**, 1411 (1999) [JETP **88**, 780 (1999)].
37. E. Barkai and J. Klafter, in *Chaos, Kinetics and Nonlinear Dynamics in Fluids and Plasmas: Proceedings of Workshop, Carry-Le Rouet, 1997*, Ed. by S. Benkadda and G. M. Zaslavsky, p. 373.
38. M. G. Kendall and S. B. Babington, *Random Sampling Numbers* (Cambridge Univ. Press, Cambridge, 1939).
39. A. V. Chechkin and V. Yu. Gonchar, cond-mat/9903400; Physica A (Amsterdam) **277**, 312 (2000).
40. A. V. Chechkin and V. Yu. Gonchar, cond-mat/9907234; submitted to Chaos, Solitons and Fractals.
41. C. Tsallis, cond-mat/9903356.
42. A. N. Malakhov, *Fluctuations in Auto-Oscillation Systems* (Nauka, Moscow, 1968).
43. S. A. Akhmanov, Yu. E. Dyakov, and A. S. Chirkin, *Introduction to Statistical Radiophysics and Optics* (Nauka, Moscow, 1981).
44. A. V. Chechkin and V. Yu. Gonchar, cond-mat/9910091.