Journal of Experimental and Theoretical Physics, Vol. 95, No. 4, 2002, pp. 581–600.
Translated from Zhurnal Éksperimental'noĭ i Teoreticheskoĭ Fiziki, Vol. 122, No. 4, 2002, pp. 675–695.

NUCLEI, PARTICLES, AND THEIR INTERACTION

On the Summation of Divergent Perturbation Series in Quantum Mechanics and Field Theory

D. I. Kazakov[a, b, *] and V. S. Popov[a]

[a] *Institute of Theoretical and Experimental Physics, Russian Academy of Sciences, ul. Bol'shaya Cheremushkinskaya 25, Moscow, 117218 Russia*
[b] *Joint Institute for Nuclear Research, Dubna, Moscow oblast, 141980 Russia*
**e-mail: kazakovd@thsun1.jinr.ru*
Received April 17, 2002

Abstract—The possibility of recovering the Gell-Mann–Low function in the asymptotic strong-coupling regime by known first-order perturbation-theory (PT) terms β_n and their asymptotics $\tilde{\beta}_n$ as $n \longrightarrow \infty$ is investigated. Conditions are formulated that are necessary for recovering the required function at the physical level of rigor: (1) a large number of PT coefficients are known whose asymptotics has already been established, and (2) there is no intermediate asymptotics. Higher orders of PT, their asymptotic behavior, and power corrections are calculated in quantum mechanical problems that involve divergent PT series (including series for a funnel potential, the $\varphi^4_{(0)}$ model, and the Stark effect in a strong field). The scalar field theory $\varphi^4_{(4)}$ is considered in the $\overline{\text{MS}}$ and MOM regularization schemes. It is shown that one cannot make any definite conclusion about the asymptotics of the Gell-Mann–Low function as $g \longrightarrow \infty$ on the basis of information available for the above theory. © *2002 MAIK "Nauka/Interperiodica".*

1. INTRODUCTION AND STATEMENT OF THE PROBLEM

It is well known that perturbation-theory (PT) series in quantum theory are, as a rule, divergent, i.e., are asymptotic series in the Poincaré sense [1, 2] and have zero radius of convergence. Nevertheless, for small values of the coupling constant g, PT gives valuable and quite reliable results whose accuracy in the case of asymptotic series is determined by the last term retained in a series expansion. The divergence of PT series is attributed to the factorial growth of higher orders of PT and is directly associated with the essential singularity of the required function at zero (with respect to the coupling constant g) and its discontinuity across a cut when the sign of the coupling constant is changed, $g \longrightarrow -0$ (the Dyson phenomenon [3]).

To proceed into the domain $g \gtrsim 1$, one has to apply summation methods for divergent series.[1] Here, in addition to information from PT, one also uses information (or assumptions) on the analyticity of the required function. The Lipatov method [5] allowed one to calculate the asymptotics of higher orders of PT in a number of problems in quantum field theory (see also [4, 6, 7]). At first sight, one could expect that matching the Lipatov asymptotics to the first few PT series terms calculated from the Feynman diagrams allows one to find important quantities of the theory (such as the Gell-Mann–Low function $\beta(g)$, which determines the behavior of invariant charge at small distances) for any g. However, even the first attempts in this direction have shown that such a program can only be realized in a certain finite interval[2] of g but not in the limit of strong coupling $g \longrightarrow \infty$. In particular, such a conclusion was made in [4, 8–12].

Recently, Suslov [13, 14] has tried to revise these results. In [14], he states that he "developed a summation procedure for divergent perturbation-theory series for arbitrary values of the coupling constant, where information on all the terms of a series is obtained by interpolating the known first terms by the Lipatov asymptotics." In [15] (see also [16]), he developed a method for calculating power corrections to the leading term of the Lipatov asymptotics. Applying these procedures, Suslov concluded [13, 14] that the asymptotics of the Gell-Mann–Low function in the scalar $\varphi^4_{(4)}$ theory with the Euclidean action

[1] Monograph [1] provides a detailed account of various summation methods from a mathematical point of view. A survey of certain methods that have been successfully used in theoretical physics can be found in [4]. The summation methods used for calculating the Gell-Mann–Low function are discussed below in Section 4.

[2] Naturally, this interval may increase either when one adds higher orders of PT or corrections to the Lipatov asymptotics or when one applies more efficient methods for summing divergent series. Here, it is useful to compare the results obtained for the Gell-Mann–Low function in the case of the $\varphi^4_{(4)}$ theory in [8, 9] and in the subsequent studies [10, 11].

$$S[\varphi] = \int d^4x \left\{ \frac{1}{2}(\partial\varphi)^2 + \frac{16\pi^2}{4!} g\varphi^4 \right\} \quad (1)$$

and in quantum electrodynamics (QED) are represented as $\beta(g) \propto g^\alpha$ with the exponent $\alpha < 1$, which implies the absence of zero charge behavior [17, 18] and that these theories are internally consistent at arbitrarily small distances [19]. This conclusion differs significantly from the results of all earlier authors; therefore, it needs to be carefully analyzed.

Below in this paper, we discuss the following question. Is it possible (and, if yes, under what conditions) to recover sufficiently reliably the function $\beta(g)$ in the strong-coupling regime, provided that the first several orders of PT β_n and their asymptotics $\tilde\beta_n$ as $n \longrightarrow \infty$, including a finite number of power corrections to it, are known. Let us introduce the following definitions, which will be needed below:

$$\beta(g) = \sum_{n=2}^{\infty} \beta_n (-g)^n, \quad (2)$$

$$\begin{aligned} \beta_n &= \tilde\beta_n \left(1 + \frac{c_1}{n} + \frac{c_2}{n^2} + \ldots\right) \\ &\equiv \tilde{\tilde\beta}_n \left(1 + \frac{c_1'}{n} + \frac{c_2'}{n^2} + \ldots\right), \end{aligned} \quad (3)$$

$$\tilde\beta_n = n! a^n n^b c_0, \quad \tilde{\tilde\beta}_n = \Gamma(n + b + 1) a^n c_0, \quad (4)$$

where n is the order of PT; a, b, and c_0 are parameters of the asymptotics; and the relation between the coefficients of the power corrections c_j and c_j' is given by

$$\begin{aligned} c_1 &= c_1' + \frac{1}{2} b(b+1), \\ c_2 &= c_2' + \frac{1}{2} b(b+1)\left[c_1' + \frac{1}{12}(3b^2 - b - 2)\right], \end{aligned} \quad (5)$$

etc.[3] Asymptotic series of type (2) arise not only for the Gell-Mann–Low functions but also in many other physics problems,[4] including quantum mechanics (see Section 3 below). Of course, our analysis also applies to these cases.

Let us briefly outline the contents of this paper. In Section 2 (see also Appendix A), we show that, in the general case, the answer to the question posed is negative; i.e., basically, the behavior of the function $\beta(g)$ as $g \longrightarrow \infty$ is not determined by the first PT coefficients β_n and their asymptotics $\tilde\beta_n$. In this section, we also introduce and illustrate the concept of intermediate asymptotics, which is important for further analysis. Next (in Section 3), we consider certain quantum mechanical problems that involve divergent PT series. We show that, in many simple problems, the asymptotics of the ground-state energy $E^{(0)}(g)$ can be recovered to a sufficient degree of accuracy by the above-listed data. However, this requires that the following additional conditions should be satisfied:

(1) It is necessary (but not sufficient!) that a large number ($n_0 \gg 1$) of PT coefficients should be known that reach their own asymptotics for $n \sim n_0$.

(2) The problem considered should not have an intermediate asymptotics.

The study of the scalar field theory (1) shows (Section 4) that the above conditions are not satisfied in this case; therefore, it is impossible to uniquely recover the asymptotics of the Gell-Mann–Low function by the data available to date. In the Section 5, we formulate the main conclusions of the present paper and criticize the assertions made in [13, 14]. In the appendices, we present certain details of calculations, additional examples, the proof of the divergence of expansions (3) for the $\varphi^4_{(0)}$ model, as well as the discussion of the calculation methods for the Stark effect in a very strong electric field for a short-range potential and for a hydrogen atom.

2. ANALYTIC EXAMPLES

First, let us consider several model examples. For the function

$$\varphi_1(g) = cg^{N+1}/(1 + \mu g), \quad \mu > 0, \quad (6)$$

the first N coefficients of the Taylor series expansion vanish, while the others can easily be made arbitrarily small by an appropriate choice of the constants c and μ. Therefore, the first N PT terms of the function $\beta_1(g) = \beta(g) + \varphi_1(g)$, as well as the Lipatov asymptotics (4), are the same as those of the original Gell-Mann–Low function $\beta(g)$. However, $\varphi_1(g) \propto g^N$ as $g \longrightarrow \infty$, which "beats" any power asymptotics of $\beta(g)$ for sufficiently large N.

The only singularity of $\varphi_1(g)$ is a simple pole at the point $g = -1/\mu < 0$. This fact does not contradict the analytic properties of the Gell-Mann–Low function because there does not exist any field theory with action (1) for $g < 0$.[5]

[3] Note that $\tilde\beta_n = \tilde{\tilde\beta}_n$ if $b = 0$ or -1. In particular, examples (20), (24), and (35) considered below refer to this case.

[4] Sometimes, this asymptotics has the form $\tilde\beta_n \propto (n\nu)!$ with $\nu \neq 1$. For example, for a quantum oscillator with anharmonicity of the type gr^{2N}, $\nu = N - 1$; for QED, $\nu = 1/2$; etc. A generalization of the formulas for this case is straightforward.

[5] The potential $U(\varphi) \propto g\varphi^4 \longrightarrow -\infty$ for $\varphi \longrightarrow \infty$ and $g < 0$; therefore, such a theory is unstable: it lacks the ground (vacuum) state.

The simple generalization

$$\varphi_2(g) = cg^{N+1}\frac{(1+\lambda_1 g)\ldots(1+\lambda_L g)}{(1+\mu_1 g)\ldots(1+\mu_M g)} \quad (7)$$

shows that,[6] as g increases, this function can successively pass through several intermediate asymptotics (in the sense of [20]) before its final asymptotics as $g \longrightarrow \infty$ is established. For instance, setting $\lambda_j \equiv 0$ and $1 \gg \mu_1 \gg \mu_2 \gg \ldots \gg \mu_n > 0$, we have

$$\varphi_2(g) \approx \begin{cases} cg^{N+1}, & g \ll \mu_1^{-1}, \\ c_1 g^N, & \mu_1^{-1} \ll g \ll \mu_2^{-1}, \\ \ldots\ldots\ldots\ldots\ldots\ldots \\ c_M g^{N-M+1}, & g \gg \mu_M^{-1}, \end{cases} \quad (8)$$

where $c_1 = c/\mu_1$, $c_2 = c/\mu_1\mu_2$, etc. Here, only the last line in (8) gives a true asymptotics of $\varphi_2(g)$ at infinity.

Note that the powers of intermediate asymptotics by no means should successively decrease. By an appropriate choice of the constants λ_j and μ_i, the order of variation of these powers can be specified at will; if some of the λ_j are negative, then an intermediate asymptotics changes its sign.

Next, let

$$\varphi_3(g) = \sum_{n=0}^{\infty}\frac{a_n}{\Gamma(n-N)}(-g)^n, \quad (9)$$

where a_n is an analytic function of number n. For example, if $a_n = (-\mu)^n$, then

$$\varphi_3(g) = (\mu g)^{N+1}e^{\mu g} \quad (10)$$

exponentially grows as $g \longrightarrow \infty$, while the PT coefficients decrease as $n^{N+1}/n!$ and cannot change the Lipatov asymptotics $\tilde{\beta}_n$.

One can readily construct other examples of this kind. For instance, for

$$\varphi_4(g) = \sum_{n=0}^{\infty}\frac{(\mu g)^n}{[\Gamma(n-N)]^2} = (\mu g)^{N+1}I_0(2\sqrt{\mu g}) \quad (11)$$

with $g \longrightarrow \infty$, we have

$$\varphi_4(g) \propto g^{N+3/4}\begin{cases} \exp(2\sqrt{\mu g}), & \mu > 0, \\ \cos(2\sqrt{-\mu g} - \pi/4), & \mu < 0; \end{cases} \quad (12)$$

therefore, the asymptotics of the function $\beta(g) + c\varphi_4(g)$ at infinity can be determined by the second term, which makes a negligible contribution to the PT coefficients for small c. Other examples will be discussed in Appendix *A*.

[6] Here, μ_i and λ_j are constants; all $\mu_i > 0$, while λ_j may be of either sign.

We conclude with an example that relates directly to the $\varphi^4_{(4)}$ theory. Let us modify this theory by adding the term $\mu g^{2N+1}\varphi^4$ with integer $N \geq 1$ and arbitrarily small $\mu > 0$ to action (1); this is equivalent to the following renormalization of the coupling constant: $g \longrightarrow \hat{g} = g(1 + \mu g^{2N})$. In this case, we will construct a PT series in powers of g, considering $\mu g^{2N+1}\varphi^4$ as a perturbation. A difference between the PT coefficients β_n and $\hat{\beta}_n$ manifests itself only starting from the terms of order g^{2N+2}. On the other hand, suppose that a discontinuity of the Gell-Mann–Low function across the cut for $g < 0$, which determines the Lipatov asymptotics (4), is given by

$$\begin{aligned} \mathrm{disc}\beta(g) &\equiv \frac{1}{2i}[\beta(g+i0) - \beta(g-i0)] \\ &= \mathrm{const}e^{a/g}(-g)^{-(\alpha+1)}(1 + b_1 g + b_2 g^2 + \ldots), \\ &g \longrightarrow -0 \end{aligned} \quad (13)$$

(specifically for (1), we have [5] $a = 1$ and $\alpha = 7/2$). The discontinuity of the action $\hat{S}[\varphi]$ in the variable $\hat{g}$ has the same form as (13); hence,

$$\begin{aligned} \mathrm{disc}\hat{\beta}(g) &= \mathrm{const}e^{a/g}(-g)^{-(\alpha+1)} \\ &\times [1 + b_1 g + \ldots + (b_{2N-1} - \mu a)g^{2N-1} + \ldots]. \end{aligned} \quad (14)$$

The difference between (13) and (14) for small g only manifests itself in the terms of order g^{2N-1} and higher, while, for the coefficients β_n and $\hat{\beta}_n$, it manifests itself in power corrections of order $n^{-(2N-1)}$. Thus, the two theories with the action functions $S[\varphi]$ and $\hat{S}[\varphi]$ have $2N + 1$ identical PT coefficients and the same asymptotics $\tilde{\beta}_n$, including $2N - 2$ identical power corrections. However, while, for (1), the Gell-Mann–Low function is

$$\beta(g) \approx c_\infty g^\alpha(\ln g)^\rho, \quad g \longrightarrow \infty,$$

in the second case we have

$$\hat{\beta}(g) \equiv \beta(\hat{g}) \approx \begin{cases} \beta(g), & 1 \ll g \ll g_0, \\ (2N+1)^\rho g^{2N\alpha}\beta(g), & g \gg g_0, \end{cases} \quad (15)$$

where $g_0 = \mu^{-1/2N}$. Therefore, if the parameter μ is sufficiently small, then the Gell-Mann–Low function $\hat{\beta}(g)$ for the theory with the action $\hat{S}[\varphi]$ first passes through an intermediate asymptotics of $\beta(g)$, and its final asymptotics at infinity, which is substantially different from $\beta(g)$, is established only for $g \gg g_0 \gg 1$.

Table 1. Higher orders of PT for the ground-state energy in a funnel potential

n	$-E_n$	$\tilde{E}_n/E_n$	n	$-E_n$	$\tilde{E}_n/E_n$
1	3/2	0.285	10	3.281(8)	1.819
2	3/2	1.71	20	3.578(22)	1.290
3	27/8	5.13	30	3.708(38)	1.174
4	795/16	2.79	40	9.152(55)	1.124
5	3843/8	2.71	50	2.521(74)	1.101
6	11166/2	2.52	60	4.849(93)	1.080
7	$\frac{9543339}{128}$	2.31	80	1.886(134)	1.059

Note: The coefficients of the PT series (21) are given with opposite sign (here, we take the opportunity to correct a misprint made in [28] concerning E_5). Henceforth, we use the following notation: n is the PT order, and $a(b) \equiv a \times 10^b$.

In quantum mechanics, the last example is analogous to the harmonic oscillator

$$V(x) = V_0 + \delta V = \frac{1}{2}g(1 + \mu g^{2N})x^2, \qquad (16)$$

where $V_0(x) = gx^2/2$ and the remainder is interpreted as a perturbation δV. In this case, the energy spectrum possesses the same properties as (15); however, the PT series here is not divergent but has a finite radius of convergence: $|g| < g_0 = \mu^{-1/2N}$.

3. ON THE SUMMATION OF PT SERIES IN QUANTUM MECHANICS

Now, we pass from mathematical examples to the discussion of specific problems of nonrelativistic quantum mechanics that involve divergent PT series. The analysis of these problems, which is of interest in itself, is also useful because the character of the divergence of PT series in these problems is the same as in quantum field theory, while the calculation of higher orders of PT is much simpler.

(a) We begin with the anharmonic oscillator

$$H = p^2 + \frac{1}{4}(x^2 + gx^4), \quad D = 1, \qquad (17)$$

for which 75 orders of PT for the ground-state energy (see Table 1 in [21]), as well as their asymptotics $\tilde{E}_n$, were calculated with high accuracy in the pioneer work of Bender and Wu [21]:

$$E^{(0)}(g) = \sum_{k=0}^{\infty} E_n(-g)^n,$$

$$E_n \approx -\Gamma\left(n + \frac{1}{2}\right) \times 3^n \sqrt{\frac{6}{\pi^3}} \qquad (18)$$

$$\times\left(1 - \frac{95}{72n} - \frac{20\,099}{10\,368n^2} - \ldots\right).$$

Later on [22], using 150 orders of PT, Bender and Wu determined ten power corrections c_j in (18) by numerical fitting (for $n \gg 1$). The asymptotics of energy as $g \longrightarrow \infty$ has the form [23, 24]

$$E^{(0)}(g) = c_\infty g^\alpha[1 + O(g^{-2\alpha})], \quad \alpha = 1/3, \qquad (19)$$

where c_∞ is a constant whose numerical value is given in [23] (in a different normalization of the Hamiltonian; however, it can easily be recalculated for (17)).

In [25], the following value was obtained for the asymptotic coefficient:[7] $c_\infty = 1.048$ (the exact value [23] is equal to 1.060362…); but the method of [13] yields the value of $\alpha = 0.317 \pm 0.032$ for the exponent of power asymptotics (19) and the coefficient c_∞ with an error of 10%. Thus, in this case, the asymptotics of $E^{(0)}(g)$ is recovered by the coefficients of PT series with a satisfactory accuracy.

Sometimes, this problem is considered as a test example for testing the summation methods for diverging series. Without denying such a possibility, we would like to stress that, in this case, the calculation involves up to 75 orders of PT, which already "catch" the asymptotics: for example, the ratio[8] $\delta_n = \tilde{E}_n/E_n - 1 =$ 20, 2.8, and 1.8% for $n = 10$, 50, and 75, respectively. For $n \lesssim 10$, the exact coefficients E_n are still rather far from $\tilde{E}_n$ (especially if we consider the Hamiltonian (17) in a D-dimensional space with the dimension $D \geq 3$). Here, the coefficient c_1 of the first power correction rapidly increases with D (see Fig. 2 in [26]) and also increases while passing from the ground to excited levels [22]. The calculations performed do not show the presence of an intermediate asymptotics in this problem. Moreover, it was rigorously proved in [23] that, in the case of (17), the PT series is summed to give the

[7] In this case, a modified PT was used in which the exponent $\alpha = 1/3$ was chosen in the original approximation.

[8] We everywhere use the definition of asymptotic coefficients given in (4) for the function β.

exact energy $E^{(0)}(g)$ everywhere in the complex plane with a cut $-\infty < g < 0$.

(b) The anharmonic oscillator represents a one-dimensional model $\varphi^4_{(1)}$ of the scalar field theory $\varphi^4_{(4)}$, and the ordinary integral

$$J(g) = (2\pi)^{-1/2}\int_{-\infty}^{\infty} d\varphi \exp\left\{-\left(\frac{1}{2}\varphi^2 + \frac{g}{4!}\varphi^4\right)\right\} = \left(\frac{2\pi}{3}g\right)^{-1/2}\exp\left(\frac{3}{4g}\right)K_{1/4}\left(\frac{3}{4g}\right) \equiv \sum_{n=0}^{\infty} J_n(-g)^n \quad (20)$$

can be considered as a zero-dimensional analogue ($\varphi^4_{(0)}$) of a functional integral in the $\varphi^4_{(4)}$ theory. For the PT coefficients, we have [4, 8]

$$J_n = \frac{\Gamma(2n+1/2)}{\sqrt{\pi}6^n\Gamma(n+1)}, \qquad \tilde{J}_n = \frac{1}{\pi\sqrt{2}}\Gamma(n)\left(\frac{2}{3}\right)^n, \quad (21)$$

$$J_n/\tilde{J}_n = \frac{\sqrt{2\pi}\Gamma(2n+1/2)}{2^{2n}n!(n-1)!} = 1 - \frac{3}{16n} + \ldots; \quad (22)$$

here, even the first orders of PT are rather close to the asymptotics: $J_n/\tilde{J}_n = 0.833$, 0.911, 0.940, and 0.963 for $n = 1$, 2, 3, and 5, respectively. Hence, one can conclude that the power corrections here are numerically small. Moreover, there is no intermediate asymptotics for $J(g)$.

Example (20) makes clear that the function $J(g)$ and the Borel sum $\tilde{J}(g)$ may have different asymptotics as $g \longrightarrow \infty$ and may even differ in sign:

$$J(g) = c_\infty g^{-1/4}[1 - kg^{-1/2} + O(g^{-1})],$$
$$\tilde{J}(g) = \sum_n \tilde{J}_n(-g)^n = -(\pi\sqrt{2})^{-1}\ln g + \ldots, \quad (23)$$

where $c_\infty = 1.600714\ldots$ and $k = 0.827901\ldots$; see also formula (B.10) in Appendix B.

Integral (20) was considered for testing a summation procedure for divergent series. In [9], the function $J(g)$ was recovered in the interval $0 < g < 100$ by the Padé–Borel technique (see (28) and (29) below) with the use of 20 orders of PT. Introducing up to 50 PT coefficients, Suslov [13] obtained a value of $\alpha = -0.235 \pm 0.025$, which is close to the exact value $\alpha = -1/4$, for the exponent of the asymptotics at infinity $J(g) \propto g^\alpha$. In light of the aforesaid, this fact does not seem surprising.

The function

$$J(g; K) = (2\pi)^{-1/2}\int_{-\infty}^{\infty} d\varphi \exp\left\{-\left(\frac{\varphi^2}{2} + g\frac{\varphi^{2K}}{(2K)!}\right)\right\} \quad (24)$$

is a generalization of (20) (the model $\varphi^{2K}_{(0)}$, $K = 2, 3, \ldots$). Here, the PT coefficients J_n are calculated analytically (see Appendix B).

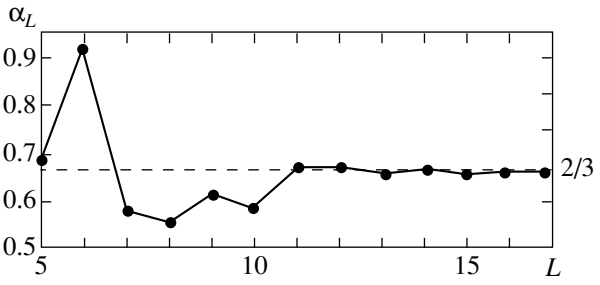


Fig. 1. Determination of the exponent α of the power-law asymptotics (32) for the ground-state energy in a funnel potential. The values of α_L are calculated by formula (31) for $x = 100$ (dots).

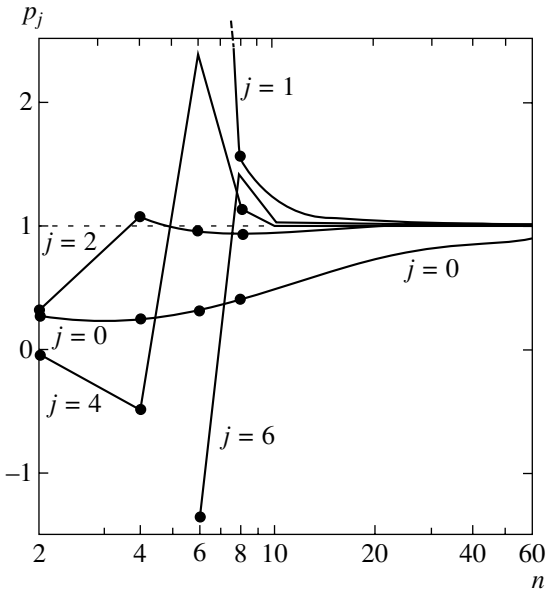


Fig. 2. The ratios $\rho_j(n) = \alpha_n/\tilde{\alpha}_n^{(j)}$ (see (35)) versus PT order n for the Stark effect in a hydrogen atom (ground state). The value of parameter j at the curves indicates the number of terms retained in expansion (35).

Let us show that series (3) in powers of $1/n$ is divergent for any K. Setting

$$\frac{J_n}{\tilde{J}_n} \equiv \eta(z) = \sum_{j=0}^{\infty} c_j z^j, \quad z = \frac{1}{n} \longrightarrow 0, \quad (25)$$

we can see from (B.3) that the function $\eta(z)$ has poles at the points $z = z_s$ that condense to zero:

$$z_s = -2K/(2s+1), \quad s = 0, 1, 2, \ldots. \quad (26)$$

Therefore, $z = 0$ is an essential singularity for $\eta(z)$, and the power series (25) diverges for any $z \neq 0$. The same

result should be expected in the case of more realistic theories, including $\varphi^4_{(4)}$ and QED (see also [16]). One can also determine the growth rate of the coefficients c_j as $j \longrightarrow \infty$ (see Appendix B).

Integral (24) can be considered as a zero-dimensional analogue of a functional integral in the field theory with self-interaction $\varphi^{2K}_{(D)}$, which can be renormalized [27] in the dimension $D = 2K/(K-1)$. The asymptotics of $J(g, K)$ as $g \longrightarrow \infty$ exhibits power-law behavior (B.10), and is the slower, the greater the nonlinearity exponent $2K$ in (24).

(c) The funnel potential

$$V(r) = -\frac{1}{r} + gr, \quad 0 < r < \infty, \tag{27}$$

was used in quantum chromodynamics for describing the states of heavy quarkonia (the so-called Cornell potential; see, for example, [28–30]). In this case, about 90 PT coefficients E_k were calculated for the ground state [31]; the parameters of the asymptotics (4) were $c_0 = -18/\pi e^3 = -0.2852\ldots$, $\beta = 1$, and $\alpha = 3/2$. Part of the calculated coefficients E_n are presented in Table 1, which shows that the difference between $\tilde{E}_n$ and E_n is less than 10% only when $n > 50$, whereas, for $n \le 10$, these coefficients differ severalfold.

The PT series was summed by the Padé–Borel technique:

$$\begin{aligned} E^{(0)}(g) &= \sum_{k=0}^{\infty} E_k(-g)^k \\ &= g^{-1}\int_0^{\infty} \exp\left(-\frac{x}{g}\right) B(x)\,dx \end{aligned} \tag{28}$$

(a generalization is given below in (49)); here, the Borel transformant $B(x)$ was replaced by the diagonal Padé approximant $[L/L]$:

$$B(x) = \sum_{k=0}^{\infty} \frac{E_k}{k!}(-x)^k \approx B_L(x) \equiv [L/L](x). \tag{29}$$

It follows from (28) that the asymptotics $E(g) \approx c_\infty g^\alpha$ as $g \longrightarrow \infty$ corresponds to the behavior of $B(x) \approx c_\infty x^\alpha/\Gamma(\alpha+1)$ as $x \longrightarrow \infty$ when $\alpha > -1$. Therefore, the exponent α of the asymptotics can be found from the equation $\alpha = \lim\limits_{L\to\infty} \alpha_L$, where

$$\alpha_L = \lim_{x\to\infty} [xB'_L(x)/B_L(x)] \tag{30}$$

or

$$\alpha_L = -1 + \lim_{x\to\infty}\left\{ xB_L/\int_0^x B_L(x')dx' \right\}, \tag{31}$$

after which one also determines the asymptotic coefficient c_∞.

Note that, in this case, the exact asymptotics of the energy levels $E_{nl}(g)$ is obtained by the scaling transformation $r \longrightarrow g^{-1/3}r$ in the Schrödinger equation:

$$E_{nl}(g) = c_\infty^{(nl)} g^{2/3}[1 + O(g^{-1/3})], \quad g \longrightarrow \infty \tag{32}$$

(a procedure suggested by Symanzik; see the remark on p. 85 in [23]). For the *ns* level ($l = 0$), the radial Schrödinger equation in the limit as $g \longrightarrow \infty$ is reduced after scaling to the Airy equation; therefore, $c_\infty^{(ns)} = -2^{-1/3}\xi_n$, where ξ_n stands for the nth zero of the Airy function Ai(ξ). For example, $\xi_1 = -2.3381\ldots$ and $c_\infty^{(1s)} = 1.855756\ldots$ for the ground state, $n = 1$.

The results of numerical calculations are illustrated in Fig. 1.[9] The application of the Padé–Borel approximants with $L \le 15$ guarantees an accuracy of 1% in the calculation of energy $E_{1s}(g)$ up to $g = 500$; for $g \sim 1$, the error does not exceed 0.01%. Naturally, the worst accuracy is achieved in the determination of the parameters of asymptotics (32). If we average α_L over the interval $11 \le L \le 17$, we obtain $\langle\alpha\rangle = 0.667 \pm 0.003$, which differs from (32) by 0.5%. Accordingly, for the asymptotic coefficient for $L = 17$, we obtain $c_\infty \approx 1.78$ (with an error of about 4%). On the other hand, the values of α_L for $L < 10$ are still far from the exact value $\alpha = 2/3$; this is natural since the coefficients $\tilde{E}_n$ for small orders of PT are several times greater than E_n (see Table 1). In the case of the potential $V(r) = -r^{-1} + gr^2$, the convergence of α_L is somewhat worse: the use of 30 orders of PT yielded $\langle\alpha\rangle = 0.47 \pm 0.08$ (the exact value is $\alpha = 1/2$).

Thus, the summation of a divergent PT series for the funnel potential yields quite satisfactory results, including the results for the strong-coupling region. However, the remarks made in the case of an anharmonic oscillator still remain in force: the calculation involves many coefficients E_n that have already reached the asymptotics $\tilde{E}_n$; in addition, the comparison with the numerical calculation shows that there is no region of intermediate asymptotics for $E(g)$ in this problem.

(d) The problem of the Stark effect in a strong field represents an instructive example from atomic physics.

In an electric field F (which is assumed to be constant and homogeneous), atomic energy levels are transformed into quasistationary states with complex energy $E(F) = E_r - i\Gamma/2$. We will restrict ourselves to

[9] In [32], diagonal Padé approximants were used up to $L = 17$, i.e., up to $2L + 1 = 35$ PT terms. The dots in Fig. 1 represent the results of calculation by formula (31) for $x = 100$ (formula (30) yields analogous results; however, the dispersion of α_L is somewhat greater).

Table 2. Higher orders of PT for the Stark effect

n	2	4	6	8
$D = 1$	5/4	55/4	10 625/16	1078125/16
$D = 3$	1/4	3/2	735/16	13265/4
H atom	21/8	$22947/2^9$	$48653931/2^{14}$	3.81903(5)
	9/2	3555/32	2512779/256	$13012777803/2^{13}$
n	10	20	50	100
$D = 1$	1.1467(7)	2.4187(20)	2.4126(71)	2.3883(173)
$D = 3$	4.3437(5)	4.2599(18)	1.6420(69)	8.0422(170)
H atom	7.7836(7)	2.6720(21)	4.5180(72)	6.4634(174)
	3.8906(8)	2.2421(22)	6.5583(73)	1.3647(176)
n	150	180	190	200
$D = 1$	6.2389(286)	3.5099(358)	9.2425(382)	4.1262(407)
$D = 3$	1.3958(284)	6.5364(355)	1.6301(380)	6.9118(404)
H atom	2.0826(288)	1.2865(360)	3.4826(384)	–
	5.4360(289)	3.6898(361)	1.0271(386)	4.8327(410)

Note: The table presents the hyperpolarizabilities α_n in the interval n = 2–200 for a δ potential in dimensions D = 1 and 3 and for the ground state of a hydrogen atom for D = 2 and 3 (the first and second rows for H, respectively).

one-dimensional and three-dimensional δ potentials[10] and the ground state of a hydrogen atom, where the PT series is given by

$$\varepsilon(F) = \frac{E_r(F)}{E_0} = \sum_{k=0}^{\infty} \alpha_{2k} F^{2k}, \tag{33}$$

where F is the expansion parameter; $\alpha_0 = 1$; α_2 is the polarizability; and α_4, α_6, … are the hyperpolarizabilities of atomic energy levels.[11]

The results of calculation of higher orders of PT (the details of these calculations are given in Appendix C) are presented in Table 2. This table displays an extremely rapid growth of the coefficients α_n, which is consistent with the asymptotics

$$\alpha_n \approx \frac{1 + (-1)^n}{2}\Gamma(n + b + 1)\left(\frac{3}{2}\right)^n c_0\left(1 + \frac{c_1}{n} + \frac{c_2}{n^2} + \ldots\right), \quad n \longrightarrow \infty, \tag{34}$$

and shows that the PT series diverges for any $F \neq 0$ (for excited states, as well as in the case of a PT expansion for the wave function, odd orders of PT are also different from zero). This formula follows from the dispersion relation (see Appendix C). The parameters b and c_j are given in Table 3.

Consider the asymptotic behavior of the PT coefficients with regard to power corrections. Denote by $\tilde{\alpha}_n^{(j)}$ a partial sum of series (34):

$$\tilde{\alpha}_n^{(j)} = \tilde{\alpha}_n\left(1 + \frac{c_1}{n} + \ldots + \frac{c_j}{n^j}\right), \quad \tilde{\alpha}_n^{(0)} \equiv \tilde{\alpha}_n = \frac{12}{\pi} n!\left(\frac{3}{2}\right)^n. \tag{35}$$

Figure 2 represents the ratios $\alpha_n/\tilde{\alpha}_n^{(j)}$, which characterize the accuracy of the jth approximation. One can see that the asymptotics $\tilde{\alpha}_n$ appreciably differs from the exact PT coefficients even for $n \sim 100$; however, the introduction of power corrections improves the situation. In particular, the inclusion of two power corrections (i.e., $\tilde{\alpha}_n^{(2)}$) yields a good accuracy even starting from $n = 4$ (however, a further increase in the number of terms retained in series (35) sharply degrades the accuracy for $n \lesssim 10$). It would seem that this situation is favorable for recovering the energy asymptotics $E(F)$ as $F \longrightarrow \infty$. However, this is not the case. The summation of PT series for various states of a hydrogen atom [35, 36], including the ground state, has shown that the

[10] Each such potential has one discrete level with energy $E_0 = -\kappa_0^2/2$ in the absence of a field (κ_0 is a constant entering the boundary condition at zero; see, for example, [33]). Below, without loss of generality, we set $\kappa_0 = 1$ and use atomic units $\hbar = m = e = 1$ in this section.

[11] In contrast to the previous examples and the $\varphi^4_{(4)}$ theory, all the coefficients α_n are of the same sign; therefore, the Borel transformant (29) has a pole on the integration path. This fact presents certain difficulties for numerical calculation, which can be overcome by different methods (see, for example, [34–36]).

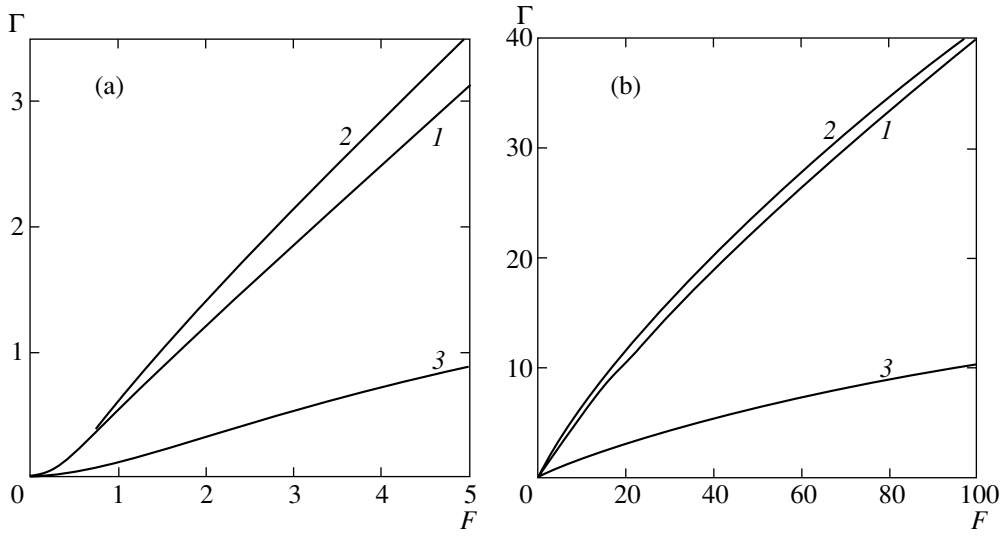


Fig. 3. The energy-level width $\Gamma(F)$ in the case of the Stark effect (a) above the barrier and (b) for $0 < F < 100$. Curves *1* and *3* correspond to one- and three-dimensional δ potentials, respectively, and curve *2* corresponds to the Rydberg states of a hydrogen atom [36, 37]. Quantities Γ and F are measured in atomic units.

dependence of the width Γ on the field strength above the barrier is very close to a linear function:

$$\Gamma(F) \approx k(F - F_0), \quad F_0 < F \lesssim 1.5\text{–}2.0 \qquad (36)$$

(parameters k and F_0 are determined by numerical fitting; for example, $k = 1.47$ and $F_0 = 0.122$ for the ground state, and $k = 0.81$ and $F_0 = 0.260$ for the Rydberg states of a hydrogen atom [37]). The linear dependence (36) takes place neither in the region of a weak field, where the width $\Gamma(F)$ is exponentially small [38], nor in the limit of very strong fields, when [39]

$$\Gamma(F) = \tilde{\Gamma}(F)\left\{1 + q\frac{\ln\ln F}{\ln F} + O\left(\frac{1}{\ln F}\right)\right\}, \quad F \longrightarrow \infty,$$

where

$$\tilde{\Gamma}(F) = \gamma_\infty (F\ln F)^{2/3}, \qquad \gamma_\infty = 2^{-5/2}\times 3^{1/2}, \quad q = 8/3. \qquad (37)$$

The asymptotics of $\Gamma(F)$ for a δ potential with $D = 1$ has the same form [37], only differing in the values of the coefficients $\gamma_\infty = 2^{-5/3}\times 3^{1/2}$ and $q = 2/3$; when $D = 3$, the form of this asymptotics is slightly changed (see (C.5)).

Note that the linear regime (36) also holds for the ground state of a helium atom [40] and for a δ potential (Fig. 3a). This is an intermediate asymptotics in the problem of the Stark effect in a strong field, which was qualitatively explained in [37] within the $1/n$ expansion. A deviation from (36) in the case of a δ potential manifests itself only for $F \gtrsim 10$ (cf. Figs. 3a and 3b). In this case, the function $\Gamma(F)$ reaches the asymptote as $F \longrightarrow \infty$ extremely slowly (Fig. 4): the relative deviation $(\Gamma - \tilde{\Gamma})/\Gamma$ becomes less than 10% only for $F > 5\times 10^7$ in the case of the one-dimensional δ potential and for $F \gtrsim 10^{26}$ in the case of a hydrogen atom (!). Therefore, the final asymptotics for the function $\Gamma(F)$ at infinity cannot be obtained by summing any (within reasonable limits) finite number of PT terms; it can only be

Table 3

Parameter	$D = 1$	$D = 2$	$D = 3$	H	j	h_j	c_j
a	3/2	3/2	3/2	3/2	1	107/12	–5.944
b	–1	–3/2	–2	0	2	7363/288	11.363
A	1	$\sqrt{\pi}/2$	1/2	4	3	158.7	–35.66
c_0	$2/\pi$	$\sqrt{2/3\pi}$	$2/3\pi$	$12/\pi$	4	469.0	–37.1
γ_∞	0.5456	–	0.6067	0.3062	5	1.025(4)	–1.12(3)
k	2/3	–	1.343	8.3	6	–	–1.0(4)

Note: The table presents the parameters entering formulas (26), (C6), and (C8) for a δ potential in a D-dimensional space, as well as (column H) for the ground state of a hydrogen atom for $D = 3$; the coefficients h_j and c_j refer to the latter case.

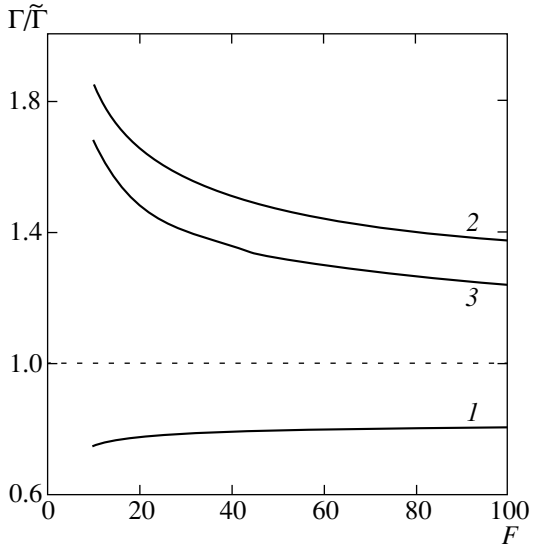


Fig. 4. Asymptotic behavior as $F \longrightarrow \infty$: the ratio $\Gamma/\tilde{\Gamma}$ versus field F. The numbering of curves is the same as in Fig. 3. The asymptotics $\tilde{\Gamma}(F)$ is defined in (37) and (C5).

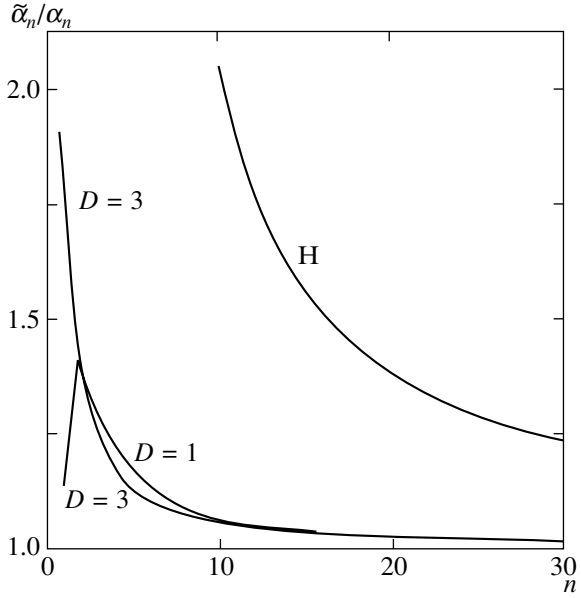


Fig. 5. The ratios $\tilde{\alpha}_n/\alpha_n$ in the case of a δ potential (of dimensions $D = 1$ and 3). Curve (H) corresponding to the ground state of a hydrogen atom is given for comparison.

obtained from the analysis of exact solutions. This also applies to the real part of energy, i.e., to the Stark shift of energy levels. In the case of short-range potentials, the hyperpolarizabilities reach their asymptotes somewhat faster as compared with the hydrogen atom; however, for $n \lesssim 10$, the difference between α_n and $\tilde{\alpha}_n$ is still rather large (Fig. 5).

(e) Just as in the case of the anharmonic oscillator, the dependence of the structure of a PT series on the dimension of a space is also of interest. For the ground state of a D-dimensional "hydrogen atom," in (33)

$$E_0 = -2/(D-1)^2, \quad D > 1,$$

we have

$$\alpha_0 = 1, \quad \alpha_2 = \frac{1}{8}(D+1)(2D+3),$$
$$\alpha_4 = \frac{1}{512}(D+1)(96D^3 + 645D^2 + 1522D + 1257), \ldots. \tag{38}$$

It is more convenient to apply the recurrence relations presented in Appendix C to the calculation of higher orders of PT. The asymptotics of the PT coefficients is determined by the discontinuity $\mathrm{disc}E(F) \equiv \Gamma(F)$ in the neighborhood of the singular point $F = 0$, which is calculated by the WKB method. Taking into consideration (C.5) and (C.10), we arrive at formulas (4) and (34) with

$$a = \frac{3}{2}, \quad b = \frac{D-3}{2},$$
$$c_0 = 2^{(D+1)/2} \times 3^{(D-1)/2}/\pi\Gamma\left(\frac{D+1}{2}\right),$$
$$c_1 = -\frac{1}{36}(12D^2 + 39D - 11), \tag{39}$$
$$c_1' = -\frac{1}{72}(33D^2 + 42D + 5)$$

(the subsequent power corrections for $D \neq 3$ are unknown). These power corrections are not small: $|c_1| > 3.2$ for $D \geq 2$. The rapid growth of the coefficient c_j in absolute value (see Table 3 for the case $D = 3$) provides evidence for the fact that the series of power corrections (3) is divergent.

As the dimension D increases, the PT coefficients reach their asymptotics more slowly, at greater n; taking into account the first correction c_1/n improves the accuracy of approximation as $n \longrightarrow \infty$; however, it sharply degrades the situation for small n (Fig. 6). It is interesting to note that, for $D = 10$, the asymptotic parameter $b = 7/2$, just as in the case of the $\varphi^4_{(4)}$ theory.

One could give more similar examples. The general conclusion is that the asymptotics of the PT coefficients in quantum theory is established, as a rule, only for $n \gg 1$, and lower orders of PT are not matched to it (except for extremely rare cases, such as the Heisenberg–Euler Lagrangian in QED, when there are no power corrections at all [8]:

$$\tilde{a}_n/a_n = 1 + (-1)^{2s} \times 2^{-2n} + \ldots,$$

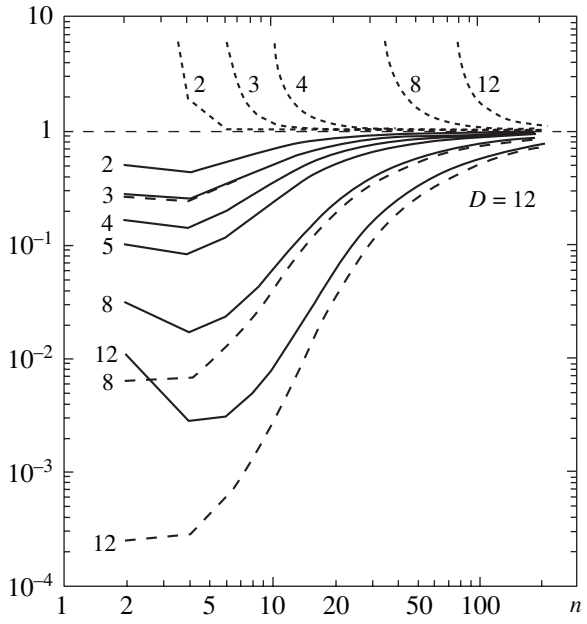


Fig. 6. The asymptotic behavior of PT coefficients for a D-dimensional hydrogen atom. Solid curves correspond to $\alpha_n/\bar{\alpha}_n$, dashed curves correspond, to $\alpha_n/\tilde{\tilde{\alpha}}_n$, and dotted curves are obtained with regard to the first power correction to $\bar{\alpha}_n$. The numbers at the curves indicate the dimension D of the space.

where $s = 0$ or 1/2 is the spin of a particle). The example with the Stark effect (where the field F plays the same role as the coupling constant g in (1)) shows how slowly the final asymptotics can be established in the strong-coupling regime when there is a region of intermediate asymptotics in the problem and the asymptotics itself contains $\ln F$.

4. ON THE GELL-MANN–LOW FUNCTION IN THE $\varphi^4_{(4)}$ THEORY

Finally, let us consider the problem of recovering the Gell-Mann–Low function in the $\varphi^4_{(4)}$ scalar field theory. The asymptotic series for the Gell-Mann–Low function is given by (2)–(4) with the parameters

$$a = 1, \quad b = \frac{7}{2}, \quad c_0 = \frac{1.096}{16\pi^2}, \quad c_1 = -4.7, \quad (40)$$

obtained by calculating the functional integral by the saddle-point method [5, 11].

Here, the situation is complicated due to the fact that, starting from the third order of PT, the coefficients of the Gell-Mann–Low function in quantum field theory depend on the subtraction scheme. The examples given above refer to the so-called symmetric MOM scheme, i.e., to the renormalization scheme with subtraction at a symmetric point ($s = t = u = -4\mu^2/3$). The original calculations in lower orders of perturbation theory were carried out precisely in this scheme (see Table 4, which also presents the number of Feynman diagrams that should be calculated in the nth order of PT). Further calculations were carried out in a modified scheme of minimal subtractions (the $\overline{\text{MS}}$ scheme). In this case, the asymptotic parameters a and b remain

Table 4. Lower orders of PT for the Gell-Mann–Low function in the $\varphi^4_{(4)}$ theory

n	2	3	4	5	6
Number of diagrams	1	2	7	23	135
β_n^{MOM}	3/2	17/6	19.27	146.1	–
$\tilde{\beta}_n/\beta_n$	0.104	0.687	1.106	1.593	–
$\bar{\beta}_n/\beta_n$	0.0978	0.659	1.072	1.554	–
$\tilde{\tilde{\beta}}_n/\beta_n$	1.323	4.582	5.053	5.666	–
$\beta_n^{\overline{\text{MS}}}$	3/2	17/6	16.27	135.8	1424.3
$\tilde{\beta}_n/\beta_n$	0.0080	0.0526	0.100	0.131	0.142
$\bar{\beta}_n/\beta_n$	0.0075	0.0505	0.097	0.128	0.139
$\tilde{\tilde{\beta}}_n/\beta_n$	0.101	0.351	0.459	0.467	0.423

Note: The table presents PT coefficients β_n in the renormalization schemes MOM and $\overline{\text{MS}}$, as well as the ratios of asymptotic coefficients to the exact ones.

unchanged, while the coefficients c_i are varied. Below, we calculate the constant c_0 in the $\overline{\text{MS}}$ scheme.

The constant c_0 in asymptotic formula (4) is calculated in the quasiclassical approximation and is determined from the classical action S, which appears in the argument of the exponential function, and the coefficient multiplying the exponential function. A transition from one scheme of subtractions to another in the leading order is associated with the variation of the classical action

$$\exp(-S_{cl}) = \exp\left(-\frac{1}{g_{\text{MOM}}}\right) \longrightarrow \exp\left(-\frac{1}{g_{\overline{\text{MS}}}}\right); \quad (41)$$

therefore, to determine c_0, one should know the relation between the charges g in these two subtraction schemes.

To this end, we consider a four-point vertex in the one-loop approximation (Fig. 7)

$$\Gamma_4 = g - \frac{1}{2}g^2[f(s) + f(t) + f(u)],$$

where s, t, and u are the Mandelstam variables and the function $f(s)$ is given by

$$f(s) = \frac{1}{\pi^2}\int \frac{d^{4-2\varepsilon}k}{k^2(k-p)^2} = \frac{1}{\varepsilon(1-2\varepsilon)}\left(\frac{-\mu^2}{s}\right)^{\varepsilon}$$
$$= \frac{1}{\varepsilon} + 2 + \ln\left(-\frac{\mu^2}{s}\right) + O(\varepsilon),$$

where $s = p^2$ and $\varepsilon \longrightarrow 0$. Thus, in a dimensional regularization, we have

$$\Gamma_4 = g - \frac{1}{2}g^2$$
$$\times\left[\frac{3}{\varepsilon} + 6 + \ln\left(-\frac{\mu^2}{s}\right) + \ln\left(-\frac{\mu^2}{t}\right) + \ln\left(-\frac{\mu^2}{u}\right)\right].$$

Let us renormalize the coupling constants and set the "bare" charge g_B in the $\overline{\text{MS}}$ and MOM schemes equal to

$$g_B = g_{\overline{\text{MS}}} + \frac{3}{2}g^2_{\overline{\text{MS}}}\left(\frac{1}{\varepsilon}\right),$$
$$g_B = g_{\text{MOM}} + \frac{3}{2}g^2_{\text{MOM}}\left(\frac{1}{\varepsilon} + \Delta_{\text{MOM}}\right), \quad (42)$$

respectively. Then, we obtain

$$\Gamma_4^{(\overline{\text{MS}})} = g_{\overline{\text{MS}}} + \frac{3}{2}g^2_{\overline{\text{MS}}}\frac{1}{\varepsilon} - \frac{1}{2}g^2_{\overline{\text{MS}}}$$
$$\times\left[\frac{3}{\varepsilon} + 6 + \ln\left(-\frac{\mu^2}{s}\right) + \ln\left(-\frac{\mu^2}{t}\right) + \ln\left(-\frac{\mu^2}{u}\right)\right]$$

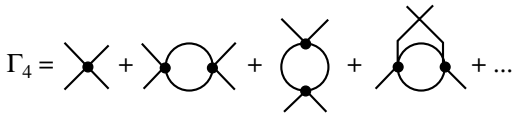


Fig. 7. Vertex function in the $\varphi^4_{(4)}$ theory in a one-loop approximation.

$$= g_{\overline{\text{MS}}} - \frac{1}{2}g^2_{\overline{\text{MS}}}$$
$$\times\left[6 + \ln\left(-\frac{\mu^2}{s}\right) + \ln\left(-\frac{\mu^2}{t}\right) + \ln\left(-\frac{\mu^2}{u}\right)\right]$$

and

$$\Gamma_4^{\text{MOM}} = g_{\text{MOM}} + \frac{3}{2}g^2_{\text{MOM}}\left(\frac{1}{\varepsilon} + \Delta_{\text{MOM}}\right) - \frac{1}{2}g^2_{\text{MOM}}$$
$$\times\left[\frac{3}{\varepsilon} + 6 + \ln\left(-\frac{\mu^2}{s}\right) + \ln\left(-\frac{\mu^2}{t}\right) + \ln\left(-\frac{\mu^2}{u}\right)\right]$$
$$= \left(g_{\text{MOM}} - \frac{1}{2}g^2_{\text{MOM}}\right)$$
$$\times\left[6 - 3\Delta_{\text{MOM}} + \ln\left(-\frac{\mu^2}{s}\right) + \ln\left(-\frac{\mu^2}{t}\right) + \ln\left(-\frac{\mu^2}{u}\right)\right].$$

By definition,

$$\Gamma_4^{\text{MOM}}\left(s = t = u = -\frac{4}{3}\mu^2\right) = g_{\text{MOM}}$$

in the MOM scheme; hence, $\Delta_{\text{MOM}} = 2 + \ln(3/4)$. Then, from (42), we have

$$\frac{1}{g_{\text{MOM}}} = \frac{1}{g_{\overline{\text{MS}}}} + \frac{3}{2}(\Delta_{\text{MOM}}). \quad (43)$$

Substituting this expression into (41), we obtain

$$\exp\left(-\frac{1}{g_{\text{MOM}}}\right) = \exp\left(-\frac{1}{g_{\overline{\text{MS}}}} - \frac{3}{2}\Delta_{\text{MOM}}\right)$$
$$= \exp\left(-\frac{1}{g_{\overline{\text{MS}}}}\right)\exp\left(-3 - \frac{3}{2}\ln\frac{3}{4}\right);$$

finally, we have

$$c_0^{\text{MOM}} = \frac{1.096}{16\pi^2},$$
$$c_0^{\overline{\text{MS}}} = c_0^{\text{MOM}}e^{-3}\left(\frac{4}{3}\right)^{3/2} \approx 0.07665c_0^{\text{MOM}}. \quad (44)$$

It would be interesting to compare the asymptotic estimates obtained with the results of calculations in lower orders of PT. Four terms of the PT series were calculated in the MOM scheme [41] and five in the $\overline{\text{MS}}$

scheme [42, 43]. The results of calculations, as well as a comparison with the Lipatov asymptotics, are presented in Table 4, which shows that the exact values of β_n are still far from the asymptotic ones; taking into consideration the first power correction, which is known in the MOM scheme [11], only degrades the situation in lower orders of PT. Indeed, for $n = 2$, 3, and 4, the coefficients $\tilde{\beta}_n^{(1)} = \tilde{\beta}_n(1 - 4.7/n) < 0$, while, for $n = 5$, we have $\tilde{\beta}_n^{(1)}/\beta_n = 0.096 \ll 1$. It is clear that β_n and $\tilde{\beta}_n$ strongly depend on the subtraction scheme. Moreover, the very expression of the asymptotics allows for certain freedom: one can rewrite (3) as

$$\beta_n \approx \Gamma(n + d + 1)a^n n^{b-d}\left(1 + \frac{k_1}{n} + \frac{k_2}{n^2} + \ldots\right) \qquad (45)$$

with an arbitrary parameter d, which is equivalent to (3) for $n \longrightarrow \infty$ but leads to the substantial dependence on d in lower orders of PT (cf. the values of $\tilde{\beta}_n/\beta_n$ and $\tilde{\tilde{\beta}}_n/\beta_n$ in Table 4). For example,

$$\frac{\tilde{\beta}_n}{\tilde{\tilde{\beta}}_n} = \frac{n!n^{7/2}}{\Gamma(n + 9/2)} \times \begin{cases} 0.0786, & n = 2, \\ 0.150, & n = 3, \\ 1 - \dfrac{63}{8n} + \dfrac{5313}{128n^2} - \dfrac{184.5}{n^3} + \ldots, & n \longrightarrow \infty, \end{cases} \qquad (46)$$

which accounts for the difference (by an order of magnitude!) between the values of $\tilde{\beta}_n/\beta_n$ and $\tilde{\tilde{\beta}}_n/\beta_n$ in Table 4 and shows that a concrete expression for the asymptotics of the Gell-Mann–Low function becomes insignificant only for higher orders of PT.

Note that $k_1 = c_1 - d(d + 1)/2$ in (45); therefore, one can reduce the first power correction to zero by an appropriate choice of the parameter d. Such a choice is likely to be optimal. If, as in the present case, $c_1 < -1/8$, then always $k_1 < 0$, and the minimal value of $|k_1|$ is attained for $d = -1/2$:

$$\beta_n \approx \bar{\beta}_n\left(1 + \frac{\bar{k}_1}{n} + \frac{\bar{k}_2}{n^2} + \ldots\right),$$

$$\bar{\beta}_n = \Gamma\left(n + \frac{1}{2}\right)a^n n^{b+1/2}c_0, \qquad (47)$$

$$\bar{k}_1 = c_1 + \frac{1}{8}, \quad \bar{k}_2 = c_2 + \frac{1}{8}c_1 - \frac{5}{128}, \ldots.$$

Table 4 shows that the coefficients $\bar{\beta}_n$ are close to $\tilde{\beta}_n$ for the $\varphi^4_{(4)}$ theory, and the power corrections are not numerically small: $k_1 = -4.6$, -4.7, and -12.6 for the parametrizations $\bar{\beta}_n$, $\tilde{\beta}_n$, and $\tilde{\tilde{\beta}}_n$, respectively. Hence, the asymptotics is reached at greater n, which, as we have noticed, is also characteristic of many quantum mechanical problems. The calculation of further orders of PT in quantum field theory presents a very difficult problem: the number of diagrams increases factorially (see, for example, Table 4), and there is no efficient method for their calculation in higher orders. Therefore, one can hardly expect any appreciable progress in this direction.

Thus, the situation in the $\varphi^4_{(4)}$ theory cannot be regarded as favorable. Here, none of the conditions necessary for determining the Gell-Mann–Low function as $g \longrightarrow \infty$ are satisfied: the calculated coefficients β_n are far from asymptotic ones, there is no information on the presence or absence of an intermediate asymptotics, and, in general, the analytic properties of the required function $\beta(g)$ are unknown.

In such a situation, it is appropriate to speak of the recovery of a function to a given accuracy in a certain finite interval of the expansion parameters, although the accuracy estimate depends on the summation method. It is this statement in which the problem in question was considered in [8–11].

In [8], the authors used the so-called improved perturbation theory, in which the sum of the divergent series $F(g) = \sum_n a_n(-g)^n$ is approximated by the functions

$$F(z)^{(N)} = \tilde{\tilde{F}}(z) + R_N(z),$$

$$R_N(z) = \sum_{n \le N}(a_n - \tilde{\tilde{a}}_n)(-z)^n, \qquad (48)$$

where $\tilde{\tilde{F}}(z) = \sum_n \tilde{\tilde{a}}_n(-z)^n$ with asymptotic coefficients of the type (4) is calculated exactly, using the Borel transformation (see (A.6)). This simple and, so to speak, naive summation method enables one to approximate the exact solution in a larger interval of g as compared with the polynomials of the ordinary PT. The summation of the remainder $R_N(z)$ by the Padé–Borel technique extends the recovery domain of the Gell-Mann–Low function to the values of $g \sim 1$ [8]. In this way, it was demonstrated that the Gell-Mann–Low function in the $\varphi^4_{(4)}$ theory cannot have a nontrivial zero for $g = g_0 \approx 0.63$ (as was pointed out in [5]); later on, this result was confirmed. However, we cannot speak of real progress toward the domain $g \gg 1$.

In [10], a conform Borel method was applied to summing series (2). The main difficulty associated with the factorial growth of the coefficients is overcome in

this case by using the Laplace integral transform. Then, the required function is represented as

$$f(g) = \frac{1}{g}\int_0^\infty dx e^{-x/g}\left(\frac{x}{g}\right)^{\mu-\nu}\left(x\frac{\partial}{\partial x}\right)^\nu B^{(\mu,\nu)}(x), \qquad (49)$$

where the Borel transform $B^{(\mu,\nu)}(x)$ is defined by a convergent power series, while the parameters μ and ν are chosen to guarantee the fastest decay of the coefficients of this series. To analytically continue the sum of the series to the whole infinite integration interval in (49), one applies a conformal mapping technique. Usually, this conformal mapping is chosen in the form

$$w(x) = \frac{\sqrt{1+ax}-1}{\sqrt{1+ax}+1},$$

which maps the complex plane with a cut $(-\infty, -1/a)$ to a unit disc $|w| < 1$ and takes into account the rooted branching of the function $B(x)$ at the point $x = 1/a$. This conformal mapping is applied to the function

$$B_\lambda[x(w)] = \frac{x^\lambda}{w^\lambda}\sum_k^N w^k B_k^{(\lambda)},$$

which contains an additional parameter λ. The introduction of this parameter allows one to take into account the asymptotics of the function $B(x)$ as $x \longrightarrow \infty$ or of the function $f(g)$ as $g \longrightarrow \infty$. Although the exact answer does not depend on λ, there is certain arbitrariness in the approximate expression that takes into account only a finite number of series terms; this arbitrariness can be canceled, for example, by minimizing the relative error [10].

In [11], a PT series was summed by using the Sommerfeld–Watson integral transformation (under the assumption that the coefficients β_n are analytic in n)

$$\beta(g) = -\frac{1}{2i}\int_{\sigma-i\infty}^{\sigma+i\infty} dz \frac{\beta(z)}{\sin\pi z} g^z, \quad \beta(n) = \beta_n; \qquad (50)$$

the integrand $\beta(z)$ was replaced by the Padé approximants $[L/M](z)$ constructed by the first terms of the PT series. The values of $\beta(g)$ thus calculated are in agreement with the results of [10] within the calculation accuracy (about 10% for $g = 40$): in the interval $0 < g \lesssim 50$, the Gell-Mann–Low function monotonically increases and behaves as g^α with the exponent $\alpha = 1.95 \pm 0.2$. The extrapolation of this function to the case of $g \longrightarrow \infty$ would imply that the $\varphi^4_{(4)}$ theory is self-contradictory. Naturally, in light of the aforesaid, one cannot definitely insist on this fact.

In [44], the authors proposed a method for constructing convergent PT series in quantum mechanics and field theory. The application of this method to the $\varphi^4_{(4)}$ theory has shown [12] that, for small values of g, the function β is recovered with a good accuracy and coincides with the results of earlier calculations; however, the effect of rejected terms of the PT series is large for large values of g, which prevents the application of this method to the case of strong coupling.

5. CONCLUSIONS

Let us formulate the main results of our work.

1. Strictly speaking, the asymptotics of the function β in the strong-coupling regime cannot be recovered by its first PT coefficients β_n and their asymptotic values $\tilde{\beta}_n$ without invoking additional information.

2. Nevertheless, sometimes, this problem can be solved at a physical level of rigor. The necessary (although not sufficient) conditions are the presence of a large number of calculated PT series terms that have already reached their asymptotics and the absence of an intermediate asymptotics of the required function.

3. In quantum field theory, the calculation of PT series coefficients is complicated due to the factorial growth of the number of Feynman diagrams, the ultraviolet divergences, and a lack of efficient methods for calculating higher order diagrams. The progress made over more than 50 years after the calculation of the Schwinger correction $\alpha/2\pi$ to the anomalous magnetic moment of electron in QED has only allowed one to reach the terms of order $(\alpha/\pi)^4$, and one can hardly expect significant changes in this field [45]. Similarly, the calculation of each subsequent term of a PT series for the function β in quantum field theory has taken ten years on the average.

4. Taking into consideration power corrections to $\tilde{\beta}_n$ improves the accuracy of approximating the coefficients of the Gell-Mann–Low function for $n \gg 1$; however, this sharply degrades the accuracy in lower orders of PT (see Figs. 2 and 6). In this case, the problem of matching the first coefficients β_n to the asymptotics $\tilde{\beta}_n$ is by no means facilitated. The situation is aggravated by the fact that series (3) in powers of $1/n$ are divergent (see, for example, [15] and Appendix B). Moreover, the determination of higher order power corrections in quantum field theory is associated with the calculation of higher order corrections to the quasiclassical approximation for a functional integral and is extremely laborious.

5. Currently available methods for summing asymptotic series (see, for example, [4, 9–12]) allow one to significantly extend the applicability of PT and to recover the required function to a given accuracy in a finite interval of its argument but do not allow one to determine its asymptotics at infinity.

Table 5. Determination of coefficients c_j by (51)

	c_1	c_2	c_3	c_4	c_5	c_6
k	Anharmonic oscillator ($\varphi^4_{(1)}$)					
1	–0.359	–	–	–	–	–
3	–2.551	4.00	–1.81	–	–	–
5	–1.626	–5.58	32.5	–49.8	24.1	–
6	–0.206	–26.9	1.53(2)	–3.69(2)	4.13(2)	–1.70(2)
7	0.334	–38.2	2.48(2)	–7.66(2)	1.29(3)	–1.12(3)
Exact [22]	–1.319	–1.939	–7.014	–40.12	–3.055(2)	–2.808(3)
k	Stark effect (ground state)					
1	–1.476	–	–	–	–	–
3	–7.108	21.3	–20.0	–	–	–
5	–5.067	–21.1	2.84(2)	–8.83(2)	8.56(2)	–
6	–3.763	–60.2	7.27(2)	–3.23(3)	6.57(3)	–5.01(3)
Exact [46, 57]	–5.944	11.36	–35.7	–37.1	–1.12(3)	–1.0(4)

Note: The table presents the coefficients of power corrections calculated by the procedure indicated in (51) with regard to k lower orders of PT; the exact values of these coefficients are presented for comparison.

6. Finally, we make several remarks concerning the studies [13, 14]. In these works, Suslov used the same information on the PT series coefficients for the Gell-Mann–Low function as those used in [10], and the summation method represented a certain modification of the conform Borel method. The fitting procedure used by Suslov is based on the statistical processing of the PT series coefficients β_n in the interval $20 \le n \le 40$, where the corresponding exact values are unknown; therefore, there are no grounds to assume that these coefficients have already reached the Lipatov asymptotics. As we have seen above, the first terms of the series are still far enough from the asymptotic terms, while the correction of order $1/n$ only degrades the situation for small n.

The requirement set forth in [13] that the interpolating function between the first calculated terms of the PT series and their asymptotic values should be smooth,[12] although it looks natural, is far from being sufficient for the unique determination of the coefficients β_n in the interval $10 \lesssim n \lesssim 50$, which is essential for calculating the asymptotics of the Gell-Mann–Low function. The test examples considered in [13] (a zero-dimensional analogue $\varphi^4_{(0)}$ and a one-dimensional anharmonic oscillator) are too simplified; moreover, the author introduces tens of exactly known terms of the PT series for which the Lipatov asymptotics has already been established to an accuracy of 1%.

The method adopted in [13] of using lower orders of PT that are known from diagrammatic calculations seems doubtful: the author used these terms to determine the coefficients c_j of power corrections from equations of the type

$$\frac{c_1}{n} + \frac{c_2}{n^2} + \ldots + \frac{c_l}{n^l} = \frac{\beta_n}{\tilde{\beta}_n} - 1, \quad 2 \le n \le l+1, \qquad (51)$$

which contain the first l PT terms. Obviously, such fitting makes sense in the region $n \ge n_0 \gg 1$ (and was earlier used in [22, 46]); however, it is not justified for $n \sim 1$ and leads to incorrect results. To illustrate this fact, using (51), we calculated the coefficients c_j for two examples for which the exact values of these coefficients are known: a one-dimensional anharmonic oscillator and the Stark effect. One can see from Table 5 that the convergence, as l increases, of the power corrections thus calculated is out of the question. As a matter of fact, this is clear enough without calculations: series (3) in powers of $1/n$ diverges, and its truncation at the lth term can only be justified if $1/n \ll 1$.

It should also be pointed out that, for the scalar $\varphi^4_{(4)}$ theory, the results of [13] for the Gell-Mann–Low function in the interval $5 \lesssim g < 20$ differ significantly from similar results of other authors that were obtained by independent methods [10, 11]. It would be interesting

[12] To put it more precisely, it is assumed that there exists an analytic function such that $\beta(n) = \beta_n$. Note that the analyticity in n of the coefficients of the Gell-Mann–Low function has not been rigorously proved. Moreover, these coefficients are defined at integer points $n = 2, 3, \ldots$, analytic continuation from which is ambiguous.

to find out the reasons for the discrepancy between these results.

The same remarks, even to a greater extent, apply to the study [14], which is devoted to the asymptotics of the Gell-Mann–Low function in QED. In this case, only the first four terms of the PT series in the MOM scheme are known [47], and the form of the asymptotic coefficients $\tilde{\beta}_n$ obtained within the framework of a certain ansatz [48] is not yet completely determined. Under these conditions, it seems untimely to speak of the possibility of recovering the Gell-Mann–Low function in the strong-coupling regime.

Thus, in our opinion, the statement set forth in [13, 14] about the absence of zero charge behavior in the $\varphi^4_{(4)}$ theory and in QED can by no means be considered justified. The results available on the summation of asymptotic PT series rather suggest that these theories are internally contradictory. A meaningful assertion concerning the asymptotics of the function $\beta(g)$ as $g \longrightarrow \infty$ can only be made after calculating a much greater number of PT series terms or by going beyond the weak-coupling approximation and applying methods that are not based on perturbation theory.

ACKNOWLEDGMENTS

We express our sincere thanks to M.I. Vysotskiĭ, Yu.A. Kubyshin, V.D. Mur, V.A. Novikov, and L.B. Okun' for discussing the paper and for many useful remarks; to S.G. Pozdnyakov for assistance in numerical computations; and B.M. Karnakov and M.N. Markina for assistance in preparing the manuscript.

This work was supported in part by the Russian Foundation for Basic Research (project nos. 02-02-16889, 00-15-96691, and 01-02-16850).

APPENDICES

APPENDIX A

Let us continue the discussion of the examples from Section 2. Functions (7) and (8) have poles on the negative half-axis $g < 0$. A singular point can be placed at zero; for example, for the function

$$f_N(g) = (-1)^{N+1}\sum_{n=N+1}^{\infty}(n-N-1)!(-g)^n = g^N\int_0^{\infty}\frac{\exp(-x/g)}{1+x}dx = -g^N\exp(g^{-1})\mathrm{Ei}(-g^{-1}), \quad \text{(A.1)}$$

the branching point $g = 0$ is the only singularity in the finite part of the g plane; a discontinuity across the cut (see (13)) is given by

$$\mathrm{disc}\, f_N(g) = \pi g^{N+1}\exp(g^{-1}), \quad -\infty < g < 0. \quad \text{(A.2)}$$

Since

$$f_N = g^N[\ln g - C + O(\ln g/g)], \quad C = -\psi(1) = 0.5772\ldots, \quad \text{(A.3)}$$

as $g \longrightarrow \infty$, the first N PT coefficients of the functions $\bar{\beta}(g) = \beta(g) + cf_N(g)$ and N power corrections are the same as those of $\beta(g)$, whereas the asymptotics as $g \longrightarrow \infty$ may be substantially different.

The PT coefficients in (A.1) vanish for $n \le N$. Although the analyticity of the coefficients of the Gell-Mann–Low function with respect to n has not been proved in quantum field theory and is actually an assumption, this condition can easily be satisfied, as is shown by examples (9)–(11). A generalization of these examples is given by the function

$$f(g, s) = \sum_{n=0}^{\infty}[\Gamma(n-N)]^{-s}g^n, \quad s = 1, 2, \ldots, \quad \text{(A.4)}$$

which, for large g, is reduced to the Mittag–Leffler function

$$E_s(z) = \sum_{n=0}^{\infty}\frac{z^n}{\Gamma(ns+1)} \approx s^{-1}\exp(z^{1/s}), \quad z \longrightarrow \infty \quad \text{(A.5)}$$

(for $|\arg z| < s\pi/2$; [1, Ch. VIII]); here, $z^{1/s} \approx sg^{1/s}$. For $s = 1$ and 2, this formula corresponds to (10) and (11).

The coefficients in (A.4) can be replaced by any integer function of n that has zeros of arbitrary multiplicity (possibly, depending on n) at the points $n = N$, $N-1$, $N-2$, …. Such functions are defined explicitly with the use of the Weierstrass representation known from the theory of analytic functions.

A generalization of formula (A.1) has the form

$$S_b(g) \equiv \sum_{n=0}^{\infty}\Gamma(n+b+1)(-g)^n = \Gamma(b+1)g^{-(b+1)}\exp(g^{-1})\Gamma(-b, g^{-1}), \quad \text{(A.6)}$$

where $\Gamma(-b, x)$ is the incomplete gamma function; for integer $b = 0, 1, 2, \ldots$, it is expressed in terms of an integral exponent:

$$\Gamma(-b, x) = \frac{(-1)^{b+1}}{b!}\left\{\mathrm{Ei}(-x) - e^{-x}\sum_{k=1}^{b}(k-1)!(-x)^{-k}\right\}$$

(for $b = 0$, the sum over k should be dropped). Consider the asymptotics

$$S_b(g) = \frac{1}{g}\begin{cases}\Gamma(b)(1-\delta+\ldots), & b>0,\\ \ln g - C + \ldots, & b = 0,\end{cases} \tag{A.7}$$

where $g \longrightarrow \infty$, $\delta = [b/(b-1)]g^{-1}$ for $b > 1$, $\delta = \ln g/g$ for $b = 1$, and $\delta = \Gamma(1-b)g^{-b}$ for $0 < b < 1$. Finally, if $b = -b' < 0$, then

$$S_b(g) = [\pi/\sin\pi b']g^{b'-1} + \ldots, \quad g \longrightarrow \infty. \tag{A.8}$$

For $b = -1, -2, \ldots$, the first few terms of series (A.6) become infinite; therefore, one should start the summation from $n = -b$. In this case, we return to (A.1), and the asymptotics $S_b(g)$ contains $\ln g$ as in (A.3). Relation (A.6) can easily be proved using the Borel transformation.

APPENDIX B

Here, we discuss the divergence of the series of power corrections (3) by an example of the zero-dimensional model $\varphi^{2K}_{(0)}$.

For the nth order of PT, from (24) we have

$$J_n = \frac{\Gamma(Kn+1/2)}{\sqrt{\pi}n!}\left[\frac{2^K}{(2K)!}\right]^n, \qquad \tilde{J}_n = [(K-1)n]!a^n n^{-1}c_0, \tag{B.1}$$

where

$$a = \frac{1}{(2K-1)!}\left(\frac{2K}{K-1}\right)^{K-1}, \qquad c_0 = 1/\pi\sqrt{2(K-1)}; \tag{B.2}$$

hence, we obtain

$$J_n/\tilde{J}_n = \sqrt{2\pi(K-1)}\frac{\Gamma(Kn+1/2)}{\Gamma((K-1)n)\Gamma(n+1)} \times\left[\frac{(K-1)^{K-1}}{K^K}\right]^n = 1 - \frac{2K^2+K-1}{24K(K-1)}\frac{1}{n} + O\left(\frac{1}{n^2}\right). \tag{B.3}$$

Taking into account the Barnes formula (see 1.18.12 in the reference book [49]), we obtain

$$\frac{\Gamma(n+b+1)}{\Gamma(n+1)n^b} = \exp\left(\sum_{k=1}^{\infty}\frac{q_k}{n^k}\right), \qquad q_k = \frac{(-1)^{k+1}}{k(k+1)}[B_{k+1}(b+1) - B_{k+1}], \tag{B.4}$$

where $B_{k+1}(x)$ and B_{k+1} are the Bernoulli polynomials and the Bernoulli numbers, respectively. Hence, we obtain

$$\frac{J_n}{\tilde{J}_n} = \exp\left(-\sum_{m-1}^{\infty}\frac{a_m}{n^{2m-1}}\right) = 1 - \frac{a_1}{n} + \frac{a_1^2}{2n^2} - \left(a_2 + \frac{1}{6}a_1^3\right)\frac{1}{n^3} + \ldots, \tag{B.5}$$

$$a_m = \frac{B_{2m}}{2m(2m-1)} \times\{1 + (K-1)^{1-2m} + K^{1-2m} - (2K)^{1-2m}\}.$$

Since [49]

$$B_{2m} \approx (-1)^{m+1}2(2m)!(2\pi)^{-2m}[1 + O(2^{-2m})], \quad m \gg 1, \tag{B.6}$$

the series in the exponent of (B.5) is alternating, and

$$\frac{J_n}{\tilde{J}_n} \equiv \sum_{j=0}^{\infty}\frac{c_j}{n^j} \approx \exp\left\{-\frac{s}{\pi}\sum_{k=1}^{\infty}\sin\left(\frac{k\pi}{2}\right)\frac{(k-1)!}{(2\pi n)^k}\right\} \tag{B.7}$$

as $n \longrightarrow \infty$, where $s = 2$ for $K = 2$ and $s = 1$ for $K \geq 3$. Thus, the asymptotics of the coefficients c_j is virtually independent of the nonlinearity exponent of the field $\varphi^{2K}_{(0)}$ in (24).

It remains to expand the exponential function in (B.7) in a power series. To this end, we apply the formula [50] relating the coefficients c_j of the power series to the coefficients a_m in (B.5):

$$c_j = \sum_{k_1+2k_2+\ldots+jk_j=j}\frac{1}{k_1!k_2!\ldots k_j!}b_1^{k_1}b_2^{k_2}\ldots b_j^{k_j}, \quad k_i \geq 0, \tag{B.8}$$

where $b_i = a_{(i+1)/2}$ for $i = 2m-1$ and $b_j = 0$ for $i = 2m$. Equation (B.8) allows one to determine the asymptotics of the coefficients c_j as $j \longrightarrow \infty$ from the asymptotics of b_k. Indeed, if the coefficients b_k grow factorially, $b_k \sim [1 + (-1)^k](k-1)!(2\pi)^{-k}$ as $k \longrightarrow \infty$, then the leading contribution to sum (B.8) is made by b_j. Other contributions of the type $b_{j-1}b_1$, $b_{j-2}b_2$, etc., are power suppressed, while contributions of the form $b_{j/2}b_{j/2}$ are suppressed exponentially ($b^2_{j/2}/b_j \sim 2^{-j}$). When the coefficients b_j are missing, as in the present case for even j, the leading contribution is determined by the next term, $b_{j-1}b_1$. Therefore, the coefficients c_j behave as

$$c_j \sim \left(\frac{j-1}{2}\right)!(2\pi)^{-(j+1)/2}, \quad j \longrightarrow \infty \tag{B.9}$$

for odd j and are relatively suppressed as $1/j$ as compared with (B.9) for even j, which is also confirmed by numerical calculations.

Finally, note that, for integral (24) as $g \longrightarrow \infty$, we have

$$J(g; K) = c_\infty g^{-1/2K}[1 - kg^{-1/K} + O(g^{-2/K})], \quad \text{(B.10)}$$

$$c_\infty = \sqrt{\frac{2}{\pi}}\Gamma\left(\frac{2K+1}{2K}\right)[(2K)!]^{1/2K}, \quad k = \frac{\Gamma(3/2K)}{2\Gamma(1/2K)}[(2K)!]^{1/K}. \quad \text{(B.11)}$$

For $K = 2$, i.e., for the $\varphi^4_{(0)}$ model, these formulas coincide with (23).

It is clear from (B.10) that the asymptotics $\tilde{J}(g; K) = c_\infty g^{-1/2K}$ is established only for $g \gg g_0 = K^K \sim (K!)^2(2/3)^K$, where $g_0 \gg 1$ if $K \geq 3$.

APPENDIX C

Here, we will discuss the calculation of higher orders of PT and asymptotics for the problem of the Stark effect.

In the model of a three-dimensional δ potential, the complex energy $E(F)$ of a level can be determined by solving the equation [51, 52]

$$\frac{1}{\sqrt{4\pi i}}\int_0^\infty \frac{d\tau}{\tau^{3/2}}\left\{1 - \exp\left[-i\left(\varepsilon\tau + \frac{1}{3}F^2\tau^3\right)\right]\right\} = 1, \quad \text{(C.1)}$$

where $\varepsilon = E/E_0 = -2E_r + i\Gamma$ is the reduced energy and $F = \mathscr{E}/\kappa_0^3$; henceforth, we assume that $\kappa_0 = 1$ and $E_0 = -1/2$. Since $\mathrm{Im}\varepsilon = \Gamma > 0$, this integral is divergent in the upper limit (this divergence is by no means accidental; it is associated with the exponential growth of the Gamov wave function of a quasistationary state as $r \longrightarrow \infty$ [33]). Therefore, Eq. (C.1) must be regularized; to this end, it suffices to shift the integration contour with the real axis τ to the lower half-plane within the angle $-\pi/3 < \arg\tau < 0$.[13] It is natural to choose the fastest descent line by setting $\tau = t\exp(-i\pi/6)$, $0 < t < \infty$, which yields

$$\int_0^\infty \frac{dt}{t^{3/2}}\left\{1 - \exp\left[-\left(e^{i\pi/3}\varepsilon t + \frac{F^2t^3}{3}\right)\right]\right\} = \sqrt{4\pi\exp\left(\frac{i\pi}{3}\right)}. \quad \text{(C.2)}$$

Now, the integral converges for any complex ε, and Eq. (C.2) can easily be solved by a computer (see curve *3* in Fig. 3).

Expanding $\exp(-F^2t^3/3)$ in a series and integrating term by term, we arrive at the equation

$$\sqrt{\varepsilon}\left\{1 - \sum_{k=1}^\infty \frac{\Gamma(3k-1/2)}{k!3^k\sqrt{4\pi}}\left(\frac{F^2}{\varepsilon^3}\right)^k\right\} = 1, \quad \text{(C.3)}$$

from which one can successively determine the higher orders of PT α_n. For $n \leq 14$, the hyperpolarizabilities α_n were calculated exactly (in the form of rational fractions), and, up to $n = 200$, they were calculated numerically (see Table 2).

It follows from (C.2) that

$$\varepsilon(F) = F^{2/3} \times\left\{\xi_0\exp\left(\frac{2\pi i}{3}\right) + \xi_1 F^{-1/3}\exp\left(-\frac{\pi i}{6}\right) + O(F^{-2/3})\right\}$$

as $F \longrightarrow \infty$, where $\xi_0 = 3^{-1/3}\xi$ and ξ is a root of the equation

$$\int_0^\infty [1 - \exp(\xi x - x^3)]x^{-3/2}dx = -0, \quad \xi = 1.010396\ldots; \quad \text{(C.4)}$$

hence, we finally obtain

$$\Gamma(F) = \mathrm{Im}\varepsilon(F) = \gamma_\infty F^{2/3}\{1 - qF^{-1/3} + O(F^{-2/3})\}, \quad \text{(C.5)}$$

where $\gamma_\infty = 2^{-1} \times 3^{1/6}\xi = 0.606711\ldots$ and $q = 1.342964\ldots$. This asymptotics does not contain $\ln F$, which distinguishes a three-dimensional δ potential from the one-dimensional one (cf. (37)). In the case of dimension $D = 1$, the equation for $E(F)$ is expressed in terms of the Airy functions (see (A.3) and (A.4) in [37]). Using asymptotic expansions for these functions, we obtain

$$\sqrt{\varepsilon} = \sum_{k=0}^\infty a_k\left(\frac{F^2}{\varepsilon^3}\right)^k, \quad \text{(C.6)}$$

where

$$a_0 = 1, \quad a_1 = 5/8,$$

$$a_{k+1} = [(2k+1)(6k+1)(6k+5)/8(k+1)]a_k.$$

Hence, one can easily determine the hyperpolarizabilities α_n for a one-dimensional δ potential.

[13] We are grateful to V.D. Mur for discussing this question. One could also apply a more universal regularization method of Zel'dovich [33, 53]; however, this method requires much larger computing time as compared with formula (C.2).

Finally, the coefficients α_n for a hydrogen atom have been calculated by the recurrence relations presented below (for $n \le 160$, our results agree with those of [46, 54], whereas, for $n \le 30$, they agree with the results of [55]). Part of the results obtained are presented in Table 2, which clearly demonstrates a factorial growth of higher orders of PT. Of the three systems considered above, the hydrogen atom is the "loosest" one: its polarizability α_2 is 18 times greater than that of an energy level bound by zero-radius forces [51]. This tendency is preserved as the PT order n increases; for $n \sim 200$, the difference between appropriate numerical values of α_n reaches six orders of magnitude. Note that all the coefficients α_n are real. The imaginary part of the energy, $\Gamma(F)$, is different from zero and only appears in the summation of PT series due to the divergence of these series for any $F \neq 0$.

The asymptotics of α_n follows from the dispersion relation [56, 57]

$$\alpha_n = [1 + (-1)^n]\frac{1}{\pi}\int^{\infty}\Gamma(F)F^{-(n+1)}dF \qquad \text{(C.7)}$$

with regard to the fact that, in weak fields, where a tunneling ionization occurs, the energy-level width is equal to

$$\Gamma(F) = A\exp(-1/aF)F^{-(b+1)}(1 - h_1F + h_2F^2 - \ldots), \quad F \longrightarrow 0 \qquad \text{(C.8)}$$

(the parameters A, a, etc., depend on a specific atomic energy level; see Table 3). Substituting this expansion into (C.7), we obtain

$$\alpha_n = \frac{2A}{\pi}a^{n+b+1}\sum_{j=0}^{\infty}(-1)^j\Gamma(n+b+1-j)a^{-j}h_j$$
$$= \tilde{\tilde{\alpha}}_n\left\{1 - \frac{h_1}{a(n+b)} + \frac{h_2}{a^2(n+b)(n+b-1)}\right.$$
$$\left. - \frac{h_3}{a^3(n+b)(n+b-1)(n+b-2)} + \ldots\right\}$$

and, reexpanding it in powers of $1/n$, arrive at the standard form (3). Here, n is even (for the ground state), and

$$\tilde{\tilde{\alpha}}_n = \frac{2A}{\pi}a^{b+1}\Gamma(n+b+1)a^n \equiv (n+b)!a^nc_0, \qquad \text{(C.9)}$$

while the coefficients c_j and h_j are connected by linear relations. For instance, for a hydrogen atom (the ground state), we have $a = 3/2$, $b = 0$, $\tilde{\alpha}_n \equiv \tilde{\tilde{\alpha}}_n$, and

$$c_1 = u_1, \quad c_2 = u_2, \quad c_3 = u_3 + u_2,$$
$$c_4 = u_4 + 3u_3 + u_2,$$
$$c_5 = u_5 + 6u_4 + 7u_3 + u_2,$$
$$c_k = \sum_{j=2}^{k}\sigma_{k-1}^{(j-1)}u_j = u_k + \frac{1}{2}(k-1)(k-2) \qquad \text{(C.10)}$$
$$\times u_{k-1} + \ldots + (2^{k-2} - 1)u_3 + u_2,$$

where $u_j = h_j(-2/3)^j$, while $\sigma_n^{(m)}$, $n \ge m$, are the Stirling numbers of the second kind known from the combinatorics (in particular, $\sigma_n^{(1)} = \sigma_n^{(n)} = 1$). The coefficients h_j are borrowed from [46, 57], and c_j are calculated by the above formulas. Note that the leading term in (C.8) is determined by the quasiclassical barrier factor and the calculation of h_j involves the determination of higher order corrections in the quasiclassical parameter.

A hydrogen atom in D dimensions was considered by many authors (see, for example, [58–60]). In the case of the ground state, $E_0 = -2/(D-1)^2$, while the PT series for the Stark effect is given by (25), where

$$F = (-2E_0)^{-3/2}\mathscr{E} = \frac{1}{8}(D-1)^3\mathscr{E}$$

is the reduced electric field. The ionization probability of this level in a weak field is determined by the WKB method:

$$\Gamma_D(F) = \frac{2^{D+1}}{(D-1)^2\Gamma((D+1)/2)}F^{-(D-1)/2} \qquad \text{(C.11)}$$
$$\times \exp\left(-\frac{2}{3F}\right)\left[1 - \frac{1}{48}(33D^2 + 42D + 5)F + O(F^2)\right],$$

which, with regard to (C.7), yields asymptotics (31) for the hyperpolarizabilities α_n.

The exact calculation of α_n can be reduced to the recurrence relations that define the descent procedure with respect to index j:

$$A_n^{(n)} = (2n-2)!/n!(n-1)!,$$
$$A_{j-1}^{(n)} = \left(j + \frac{D-1}{2}\right)A_j^{(n)} + \sum_{m=1}^{n-1}\sum_{p+q=j-1}A_p^{(m)}A_q^{(n-m)}, \qquad \text{(C.12)}$$
$$1 \le j \le n,$$

starting from $j = n$ ($A_n^{(n)} = 1, 1, 2, 5, 14, \ldots$ for $n = 1, 2, 3, 4, 5, \ldots$). Here, $p, q \ge 0$, and the nth PT order is given by[14]

[14]Here, we corrected the misprints made in [60]. If $D = 3$, then $F \equiv \mathscr{E}$, and (C.12) and (C.13) are reduced, after certain transformations, to equations from [46, 54] for nodeless states with parabolic quantum numbers $n_1 = n_2 = 0$ and $m = (D-3)/2$.

$$\alpha_n = \begin{cases} 2^{1-2n}\left[(D+1)A_1^{(n)} + 2\sum_{m=1}^{n-1} A_0^{(m)} A_0^{(n-m)}\right], & n = 2, 4, 6, \ldots, \\ 0, & n \text{ is odd.} \end{cases} \quad \text{(C.13)}$$

Here, for odd n, we have $A_0^{(n)} = (D+1)A_1^{(n)}/2$, whereas, for even n, one should set $A_0^{(n)} \equiv 0$ in Eqs. (C.12) and (C.13).

Thus, we have a purely algebraic calculation scheme, which is very convenient for computer calculations (all $A_j^{(n)}$ are positive rational fractions; therefore, the calculation by recurrence relations does not involve cancellations and does not lead to a loss of accuracy). This scheme yielded record-breaking values of $n \sim 200$ (see Table 2), which is hardly possible in quantum field theory in the foreseeable future (in fact, in quantum mechanics, such a simplification in the calculation of higher orders of PT is only possible for simple systems with a polynomial or Coulomb potential). In the case of an excited state, the equations become, as a rule, more complicated, and the odd orders of PT α_n become different from zero.

Note in addition that, for the three parametrizations

$$\begin{aligned} \bar{\alpha}_n &= \Gamma(n+1/2)a^n n^{(D-2)/2} c_0, \\ \tilde{\alpha}_n &= n! a^n n^{(D-3)/2} c_0, \\ \tilde{\tilde{\alpha}}_n &= \Gamma\left(n + \frac{D-1}{2}\right) a^n c_0 \end{aligned} \quad \text{(C.14)}$$

for $n \longrightarrow \infty$, we have

$$\begin{aligned} \frac{\tilde{\tilde{\alpha}}_n}{\tilde{\alpha}_n} &= 1 + \frac{(D-1)(D-3)}{8n} + \ldots, \\ \frac{\tilde{\tilde{\alpha}}_n}{\bar{\alpha}_n} &= 1 + \frac{(D-2)^2}{8n} + \ldots \end{aligned} \quad \text{(C.15)}$$

up to the terms of order $1/n^2$. Thus, it is clear why all the parametrizations for the dimensions $D = 2$ and 3 are close to each other, as one can see from Fig. 6.

REFERENCES

1. G. H. Hardy, *Divergent Series* (Clarendon, Oxford, 1949; Inostrannaya Literatura, Moscow, 1951).
2. M. A. Evgrafov, Itogi Nauki Tekh., Ser.: Sovrem. Probl. Mat. **13**, 5 (1986).
3. F. J. Dyson, Phys. Rev. **85**, 631 (1952).
4. D. I. Kazakov and D. V. Shirkov, Fortschr. Phys. **28**, 465 (1980).
5. L. N. Lipatov, Pis'ma Zh. Éksp. Teor. Fiz. **25**, 116 (1977) [JETP Lett. **25**, 104 (1977)]; Zh. Éksp. Teor. Fiz. **72**, 411 (1977) [Sov. Phys. JETP **45**, 216 (1977)].
6. E. Brezin, J. C. Le Guillou, and J. Zinn-Justin, Phys. Rev. D **15**, 1544 (1977); **15**, 1558 (1977).
7. J. Zinn-Justin, *Quantum Field Theory and Critical Phenomena* (Clarendon, Oxford, 1989).
8. V. S. Popov, V. L. Eletsky, and A. V. Turbiner, Phys. Lett. B **72**, 99 (1977); Zh. Éksp. Teor. Fiz. **74**, 445 (1978) [Sov. Phys. JETP **47**, 232 (1978)].
9. V. L. Eletsky and V. S. Popov, Phys. Lett. B **77**, 411 (1978).
10. D. I. Kazakov, O. V. Tarasov, and D. V. Shirkov, Teor. Mat. Fiz. **38**, 15 (1979).
11. Yu. A. Kubyshin, Teor. Mat. Fiz. **58**, 137 (1984).
12. D. I. Kazakov and A. I. Onishchenko, Teor. Mat. Fiz. **110**, 291 (1997).
13. I. M. Suslov, Pis'ma Zh. Éksp. Teor. Fiz. **71**, 315 (2000) [JETP Lett. **71**, 217 (2000)]; Zh. Éksp. Teor. Fiz. **120**, 5 (2001) [JETP **93**, 1 (2001)].
14. I. M. Suslov, Pis'ma Zh. Éksp. Teor. Fiz. **74**, 211 (2001) [JETP Lett. **74**, 191 (2001)].
15. I. M. Suslov, Zh. Éksp. Teor. Fiz. **117**, 659 (2000) [JETP **90**, 571 (2000)].
16. I. M. Suslov, Usp. Fiz. Nauk **168**, 503 (1998) [Phys. Usp. **41**, 441 (1998)].
17. L. D. Landau, A. A. Abrikosov, and I. M. Khalatnikov, Dokl. Akad. Nauk SSSR **95**, 773 (1954); **95**, 1177 (1954); **96**, 261 (1954).
18. E. S. Fradkin, Zh. Éksp. Teor. Fiz. **28**, 750 (1955) [Sov. Phys. JETP **1**, 604 (1955)].
19. A. A. Vladimirov and D. V. Shirkov, Usp. Fiz. Nauk **129**, 407 (1979) [Sov. Phys. Usp. **22**, 860 (1979)].
20. Ya. B. Zel'dovich and V. V. Sokolov, Usp. Fiz. Nauk **146**, 493 (1985) [Sov. Phys. Usp. **28**, 608 (1985)].
21. C. M. Bender and T. T. Wu, Phys. Rev. **184**, 1231 (1969).
22. C. M. Bender and T. T. Wu, Phys. Rev. Lett. **27**, 461 (1971); Phys. Rev. D **7**, 1620 (1973).
23. B. Simon, Ann. Phys. **58**, 76 (1970).
24. F. T. Hioe, D. McMillan, and E. W. Montroll, Phys. Rep. **43**, 305 (1978).
25. A. D. Dolgov and V. S. Popov, Phys. Lett. B **79**, 403 (1978).
26. V. L. Eletsky and V. S. Popov, Phys. Lett. B **94**, 65 (1980).
27. L. N. Lipatov, Zh. Éksp. Teor. Fiz. **71**, 2010 (1976) [Sov. Phys. JETP **44**, 1055 (1976)].
28. E. Eichten, K. Gottfried, T. Kinoshita, *et al.*, Phys. Rev. D **17**, 3090 (1978); **21**, 203 (1980); **21**, 312 (1980).
29. C. Quigg and J. Rosner, Phys. Rep. **56**, 169 (1979).
30. E. Eichten and C. Quigg, Phys. Rev. D **49**, 5845 (1994).
31. V. L. Eletsky, V. S. Popov, and V. M. Weinberg, Phys. Lett. A **84**, 235 (1981); Zh. Éksp. Teor. Fiz. **81**, 1567 (1981) [Sov. Phys. JETP **54**, 833 (1981)].

32. V. S. Popov and V. M. Weinberg, Phys. Lett. A **90**, 107 (1982).
33. A. I. Baz', Ya. B. Zel'dovich, and A. M. Perelomov, *Scattering, Reactions, and Decays in Nonrelativistic Quantum Mechanics* (Nauka, Moscow, 1971, 2nd ed.; Israel Program for Scientific Translations, Jerusalem, 1966).
34. V. Franceschini, V. Grecchi, and H. J. Silverstone, Phys. Rev. A **32**, 1338 (1985).
35. V. M. Vaĭnberg, V. D. Mur, and V. S. Popov, Zh. Éksp. Teor. Fiz. **93**, 450 (1987) [Sov. Phys. JETP **66**, 258 (1987)].
36. V. S. Popov, V. D. Mur, A. V. Sergeev, *et al.*, Phys. Lett. A **124**, 77 (1987); **149**, 418 (1990).
37. V. S. Popov, Phys. Lett. A **173**, 63 (1993).
38. L. D. Landau and E. M. Lifshitz, *Course of Theoretical Physics*, Vol. 3: *Quantum Mechanics: Non-Relativistic Theory* (Nauka, Moscow, 1974; Pergamon, New York, 1977).
39. L. Benassi, V. Grecchi, E. Harrell, and B. Simon, Phys. Rev. Lett. **42**, 704 (1979); **42**, 1430 (1979).
40. A. Scrinzi, M. Geissler, and T. Brabec, Phys. Rev. Lett. **83**, 706 (1999).
41. F. M. Dittes, Yu. A. Kubyshin, and O. V. Tarasov, Teor. Mat. Fiz. **37**, 66 (1978).
42. A. A. Vladimirov, D. I. Kazakov, and O. V. Tarasov, Zh. Éksp. Teor. Fiz. **77**, 1035 (1979) [Sov. Phys. JETP **50**, 521 (1979)].
43. K. G. Chetyrkin, S. G. Gorishny, S. A. Larin, and F. V. Tkachov, Phys. Lett. B **132**, 351 (1983); D. I. Kazakov, Phys. Lett. B **133**, 406 (1983); H. Kleinert, J. Neu, V. Schulte-Frohlinde, *et al.*, Phys. Lett. B **272**, 39 (1991); **319**, 545 (1993); hep-th/9503230.
44. V. V. Belokurov, Yu. P. Solov'ev, and E. T. Shavgulidze, Teor. Mat. Fiz. **109**, 51 (1996).
45. V. W. Hughes and T. Kinoshita, Rev. Mod. Phys. **71**, S133 (1999).
46. V. L. Eletskiĭ and V. S. Popov, Dokl. Akad. Nauk SSSR **250**, 74 (1980) [Sov. Phys. Dokl. **25**, 27 (1980)].
47. S. G. Gorishny, A. L. Kataev, S. A. Larin, and L. R. Surguladze, Phys. Lett. B **256**, 81 (1991).
48. E. B. Bogomolny and V. A. Fateyev, Phys. Lett. B **76**, 210 (1978).
49. *Heigher Transcendental Functions (Bateman Manuscript Project)*, Ed. by A. Erdelyi (McGraw-Hill, New York, 1953; Nauka, Moscow, 1974).
50. A. P. Prudnikov, Yu. A. Brychkov, and O. I. Marichev, *Integrals and Series* (Nauka, Moscow, 1986; Gordon and Breach, New York, 1989), Supplementary Chapters.
51. Yu. N. Demkov and G. F. Drukarev, Zh. Éksp. Teor. Fiz. **47**, 918 (1964) [Sov. Phys. JETP **20**, 614 (1964)].
52. S. P. Andreev, B. M. Karnakov, V. D. Mur, and V. A. Polunin, Zh. Éksp. Teor. Fiz. **86**, 866 (1984) [Sov. Phys. JETP **59**, 506 (1984)].
53. V. D. Mur, S. G. Pozdnyakov, V. S. Popov, and S. V. Popruzhenko, Pis'ma Zh. Éksp. Teor. Fiz. **75**, 294 (2002) [JETP Lett. **75**, 249 (2002)].
54. S. P. Alliluev, V. L. Eletsky, and V. S. Popov, Phys. Lett. A **73**, 103 (1979).
55. V. Privman, Phys. Rev. A **22**, 1833 (1980).
56. I. W. Herbst and B. Simon, Phys. Rev. Lett. **41**, 67 (1978).
57. H. J. Silverstone, B. G. Adams, J. Čizek, and P. Otto, Phys. Rev. Lett. **43**, 1498 (1979).
58. S. P. Alliluev, Zh. Éksp. Teor. Fiz. **33**, 200 (1957) [Sov. Phys. JETP **6**, 156 (1957)]; G. Gyorgyi and J. Revai, Zh. Éksp. Teor. Fiz. **48**, 1445 (1965) [Sov. Phys. JETP **21**, 967 (1965)]; M. Bander and C. Itzykson, Rev. Mod. Phys. **38**, 330 (1966); **38**, 346 (1966); D. R. Herschbach, J. Chem. Phys. **84**, 838 (1986); X. L. Yang, S. H. Guo, F. Y. Chan, *et al.*, Phys. Rev. A **43**, 1186 (1991).
59. V. S. Popov, in *Physics of High Energies and Theory of Elementary Particles* (Naukova Dumka, Kiev, 1967), p. 702.
60. S. P. Alliluev and V. S. Popov, Zh. Éksp. Teor. Fiz. **104**, 3569 (1993) [JETP **77**, 701 (1993)].

Translated by I. Nikitin

Journal of Experimental and Theoretical Physics, Vol. 95, No. 4, 2002, pp. 601–604.
Translated from Zhurnal Éksperimental'noĭ i Teoreticheskoĭ Fiziki, Vol. 122, No. 4, 2002, pp. 696–699.

NUCLEI, PARTICLES, AND THEIR INTERACTION

Comments on the Article by D. I. Kazakov and V. S. Popov

I. M. Suslov

Kapitza Institute of Physical Problems, Russian Academy of Sciences, Moscow, 117334 Russia

e-mail: suslov@kapitza.ras.ru

Received June 7, 2002

Abstract—The article by D.I. Kazakov and V.S. Popov [1] is devoted almost entirely to a criticism of my works [2–4]. Some of the questions raised by them are not without interest, but they have virtually all been discussed in detail in my publications. © *2002 MAIK "Nauka/Interperiodica".*

1. Numerous examples given in [1] essentially boil down to the following. If we know the first few terms of the diverging series

$$W(g) = \sum_{N=N_0}^{\infty} W_N(-g)^N \tag{1}$$

and the asymptotic form of W_N for $N \longrightarrow \infty$, the asymptotics of the sum $W(g)$ in the strong coupling limit can be changed drastically by varying the values of the unknown intermediate coefficients. It was concluded that it is not possible in principle to reconstruct the asymptotic form of $W(g)$ on the basis of this information.

In fact, there is no need to consider so many examples since a more stringent statement is made in [3]: "a function with a predetermined behavior at infinity can be constructed on the basis of a finite number of coefficients and their asymptotic form." The algorithm for solving this problem is also given: "A comprehensive formulation of the problem is possible when all values of W_N are defined approximately; in this case, the sum $W(g)$ can be reconstructed to a certain degree of accuracy. For this reason, a necessary stage in the solution of the problem is the interpolation of the coefficient function; naturally, this is possible only under the assumption of its analyticity."

These citations reveal the conceptual difference between the approach used in [2–4] and the position of the authors of [1]. If the information on the intermediate expansion coefficients is absent indeed, it is impossible to reconstruct the asymptotic form of $W(g)$. However, the smoothness of the coefficient function makes it possible to predict (using interpolation) the unknown W_N to a certain degree of accuracy and to present them in the form $W_N^0 + \delta W_N$, where W_N^0 are exact coefficients and δW_N is a small perturbation. By hypothesis, coefficients W_N^0 give a power behavior for large values of g ($W(g) = W_\infty g^\alpha$ with $W_\infty \sim 1$), while δW_N generate a generally more rapidly increasing function of g containing a small parameter as a coefficient. Consequently, there exists an interval of g in which the true asymptotic form $W_\infty g^\alpha$ can be reconstructed (naturally, with a certain error in W_∞ and α); as the information on coefficients W_N becomes more extensive, their uncertainty δW_N decreases, and the above-indicated interval of g increases indefinitely. Consequently, there are no basic limitations for determining the asymptotic behavior of $W(g)$: the unattainability of the asymptotics has the same sense as unattainability of infinity. The strategy of choosing the appropriate interval for processing was discussed in detail in [3], but in a somewhat different terminology (see below).

2. The analyticity of the coefficient function has not been proved rigorously, but serious arguments exist in favor of this property. According to Lipatov [5], the coefficients of expansion of the functional integral

$$W(g) = \int D\varphi \exp(-S_0\{\varphi\} - gS_{\text{int}}\{\varphi\}) \tag{2}$$

in g can be written in the form

$$W_N = \int_C \frac{dg}{2\pi i g} \int D\varphi \times \exp(-S_0\{\varphi\} - gS_{\text{int}}\{\varphi\} - N\ln g) \tag{3}$$

(C is a contour in the complex plane enclosing the point $g = 0$), and analyticity in N takes place under the condition of convergence of integrals. The integrals converge at least in the steepest descent approximation, which is

valid for large values of N. Unfortunately, the analyticity region cannot be determined in this way.

It follows from the representation of series (1) in the form of a Sommerfeld–Watson integral that, in the case of a power asymptotic form $W(g) \propto g^\alpha$, the point $N = \alpha$ is the extreme right singularity of the coefficient function in the complex plane of N [3, 5, 6]. Consequently, the analyticity region can be controlled by the final result: if α turns out to be smaller than N_0, the coefficient function is analytic in the domain $\mathrm{Re}\,N > \alpha$, and the assumption concerning its smoothness on the real axis for $N \geq N_0$, which is required for interpolation, is self-consistent.

The analyticity of the coefficient function is explicitly violated in a number of examples presented in [1] (see formulas (6) and (7)). Such perturbations cannot appear as a result of a smooth interpolation, and, hence, their discussion has no sense.

Examples of type (9), in which the correction δW_N to the coefficients is an integral function and the corresponding dependence on g is exponential (the impossibility of power dependence is proved by contradiction), are more interesting. Such perturbations actually appear as interpolation errors and are manifested in the form of an exponential component in the coefficients U_N (see Fig. 10 in [3]; the definition of U_N will be given below). For large values of such errors, the results cannot be interpreted at all in terms of a power law. This forms the basis of "filtration" of such errors, which is proposed in [3]: the interpolation method is chosen in such a way as to ensure the minimum values of χ^2 in the course of fitting by the power law. As a result, these errors can be reduced to such an extent that they hardly affect the accuracy of the results.[1]

Footnote 12 in [1] contains a reference to the problem of ambiguity of the analytic continuation from integral points to the complex plane, which emerges due to the fact that, generally speaking, the point $N = \infty$ is a singular point.[2] However, there are theorems guaranteeing the uniqueness of the analytic continuation under the condition that the singularity at $N = \infty$ is quite weak (the rate of increase for $|N| \longrightarrow \infty$ is limited by a certain exponential function); the standard interpolation schemes automatically converge to this unique function [7]. The singularity of the above coefficient function in actual field problems is very weak: a regular expansion in $1/N$ is valid, but has zero convergence radius [8]; this is apparently sufficient for proving the uniqueness.

3. I agree with the authors of [1] that, in the case of insufficient information, any method gives incorrect results for the asymptotic form if the transition to it is strongly delayed. Processing is always carried out on a finite interval of g, although it is not always evident.

The algorithm used in [2–4] is based on the fact that, in the case of a power asymptotic form ($W(g) \propto g^\alpha$), the coefficients U_N of the converging series obtained as a result of a certain resummation of series (1) behave as $N^{\alpha-1}$ in the region of large N. It can be proved that the knowledge of expansion coefficients with $N \leq N_{\max}$ determines the sum of the series for $g \lesssim N_{\max}$. The typical working interval $20 \leq N \leq 40$ used in [3] effectively corresponds to the region $20 \lesssim g \lesssim 40$. However, I did not suggest that the asymptotic form had already attained in this region since the processing was carried out not according to the purely power law Ag^α, but taking into account the first correction of the form $A'g^{\alpha'}$ to this law. This correction was poorly reproduced and depended strongly on the specific procedure, but the main asymptotic form turned out to be very stable. Consequently, the results obtained in [3] effectively correspond to the range of rather high values of g. The normalization of charge used in [3] was chosen from the condition that the nearest singularity in a Borel plane lies at a unit distance from the origin. In this case, the characteristic scale over which the variation of the β function occurs is found to be of the order of unity, and there are all grounds to state that the working interval lies in the asymptotic region.

In [9], the expansion coefficients with $N \leq 5$ were used, and interpolation of the coefficient function was not carried out. The information on the intermediate expansion coefficients was not used, and the arguments given in [1] are applicable to [9] in full extent. The bundle of curves presented in Fig. 9 [9] and demonstrating a 10% accuracy for $g < 50$ should not be taken seriously; these curves were obtained for a certain fixed summation procedure chosen in the course of "guessing" the asymptotic form. If another asymptotic form is used, the summation procedure will change, leading to considerable changes in the results in the range of large values of g. In my method, coefficients U_N display a clearly manifested intermediate asymptotics $U_N \sim N$ (see Subsection 8.3 in [3]), which corresponds exactly to the result obtained in [9]. If the values of W_N with $N \leq 10$ are used for reconstructing the asymptotic form, my method leads to results completely identical to those in [9].

[1] If the coefficient function varies slowly on a scale of the order of unity and the interpolation curve possesses the same property, the amplitude of such errors is found to be exponentially small. Indeed, an exponential increase in g occurs only if the correction to the coefficient function contains an oscillatory factor $(-1)^N = e^{i\pi N}$ (see Eqs. (9) and (10) in [1]), which is a "high-frequency" factor in this case; it is well known, however, higher-order Fourier harmonics are contained in a smooth function with an exponentially small weight.

[2] The conventional uniqueness theorem refers to the analytic continuation from a set of points containing a limiting point in the regularity region.

In [6], the interpolation of the coefficient function was carried out formally, but index α was determined from the position of the extreme right singularity obtained as a result of construction of Padé approximants. If only the known expansion coefficients are used for constructing such approximants, the order of approximation turns out to be quite low and corresponds effectively to an analysis in the range of comparatively small values of g. In my opinion, it is more reasonable to choose first a certain bundle of interpolation curves and construct high-order approximants for each curve, using the spread in the results for different curves as a measure of their uncertainty. Such an approach makes it possible to reproduce the results obtained in [3], although with a considerably larger error.

In my opinion, the above arguments clearly explain the discrepancy between the results obtained in [3] and [6, 9]; additional discussion can be found in [3], where the special Subsection 8.3 is devoted to this question. It should be noted that the methods of reconstructing the asymptotic form proposed by Kazakov [9, 10] and Kubyshin [6] are of certain interest and deserve more attention: the discrepancy in the results is not due to drawbacks of these methods, but is associated precisely with the above-mentioned conceptual difference in the approaches.

4. The authors of [1] question the validity of my method of interpolation, which was carried out on the basis of the formula

$$W_N = W_N^{as}\left\{1 + \frac{A_1}{N} + \frac{A_2}{N^2} + \ldots\right\} \qquad (4)$$

by truncating the series and by choosing parameters A_K from a correspondence with the first L coefficients W_N; the value of A_1 and the parameters of the asymptotic form W_N^{as} were assumed to be known. For small L, this method is quite effective: in zero dimension, the interpolation error is of the order of 10^{-4} for $L = 1$ and of the order of 10^{-9} for $L = 5$. For an anharmonic oscillator, the error is of the order of 10^{-2} for $L = 5$ and of the order of 10^{-3} for $L = 9$. The interpolation error for the φ^4 theory is estimated at a few percent; this can be done by varying the interpolation scheme or on the basis of formula (14) from [2].

I have never stated that the values of coefficients A_K obtained in the course of such a procedure are close to actual values.[3] I also admit that this algorithm may become unsatisfactory upon an increase in L. Ideally, the method of interpolation must be based on analytic properties of the coefficient function (see Section 6 in [8]) and possess a guaranteed convergence rate for $L \longrightarrow \infty$.

[3] This is true only under certain constraints.

5. Some remarks made by the authors of [1] lead to confusion. For example, it is said in Section 3 that, "introducing up to 50 PT coefficients in calculations," it was obtained in [3] for the zero-dimensional case "the value of $\alpha = -0.235 \pm 0.025$, which is close to the exact value of $\alpha = -1/4$, "which is not surprising" in view of the large number of coefficients used. This result was indeed obtained in [3] at the first stage of testing (Section 4); however, in Section 5, the use of only one (!) coefficient gives approximately the same result $\alpha = -(0.218–0.271)$. This result refutes the main statement made in [1] and stipulating a large number of expansion coefficients. The amount of information required for reconstructing the asymptotic form can be determined only empirically, but not on the basis of general principles.

It is stated in Section 3 that the value $c_\infty = 1.048$ was obtained in [11] instead of 1.0603… for an anharmonic oscillator, while in [3] this value was obtained with a 10% error. However, index α in [11] was assumed to be equal to the exact value 1/3, while in [3] it was determined in the course of data processing. When the exact value of α is used, the method developed in [3] gives the value of c_∞ with a relative error of 6×10^{-3}. Such details create an impression that the method developed in [3] is not superior to many other methods and does not lead to any progress. It should be emphasized in this connection that this method was claimed from the very outset not to be record exact, but to be a robust method; i.e., this method possesses an elevated stability under unfavorable conditions (see Subsection 2.3 in [3]).[4]

It should be noted in conclusion that the results obtained in [2–4] are quite natural: incomplete information on the asymptotic form $W(g)$ is extracted from incomplete information on coefficients W_N. My aim was to reflect adequately the uncertainty of the initial information in the uncertainty of the results. There are all grounds to believe that this aim was reached: within their uncertainty, the results are independent of the interpolation method. Index α in the φ^4 theory changes slightly even after a significant decrease in the amount of information [4]; this is an indication that the information is sufficient. In addition, the results match the available analytic estimates to form a coherent picture (see Subsection 8.2 in [32] and [4]).

REFERENCES

1. D. I. Kazakov and V. S. Popov, Zh. Éksp. Teor. Fiz. **122**, 675 (2002) [JETP **95**, 581 (2002)].

[4] In addition, there are no references in [1] that the optimal parametrization of the asymptotic form (47) was established in [3] or that corrections to the asymptotic form in the zero-dimensional φ^4 theory were obtained in [8]. The result (B.9) obtained in Appendix B on the basis of cumbersome calculations for $K = 2$ contradicts the result obtained in [8] (see formulas (8) and (41)) by using two different (and almost trivial) methods.

2. I. M. Suslov, Pis'ma Zh. Éksp. Teor. Fiz. **71**, 315 (2000) [JETP Lett. **71**, 217 (2000)].
3. I. M. Suslov, Zh. Éksp. Teor. Fiz. **120**, 5 (2001) [JETP **93**, 1 (2001)].
4. I. M. Suslov, Pis'ma Zh. Éksp. Teor. Fiz. **74**, 211 (2001) [JETP Lett. **74**, 191 (2001)].
5. L. N. Lipatov, Zh. Éksp. Teor. Fiz. **72**, 411 (1977) [Sov. Phys. JETP **45**, 216 (1977)].
6. Yu. A. Kubyshin, Teor. Mat. Fiz. **58**, 137 (1984).
7. A. O. Gel'fond, *Calculation of Finite Differences* (Nauka, Moscow, 1967).
8. I. M. Suslov, Zh. Éksp. Teor. Fiz. **117**, 659 (2000) [JETP **90**, 571 (2000)].
9. D. I. Kazakov, O. V. Tarasov, and D. V. Shirkov, Teor. Mat. Fiz. **38**, 15 (1979).
10. D. I. Kazakov, Teor. Mat. Fiz. **46**, 227 (1981).
11. A. D. Dolgov and V. S. Popov, Phys. Lett. B **79**, 403 (1978).

Translated by N. Wadhwa

Journal of Experimental and Theoretical Physics, Vol. 95, No. 4, 2002, pp. 605–610.
Translated from Zhurnal Éksperimental'noĭ i Teoreticheskoĭ Fiziki, Vol. 122, No. 4, 2002, pp. 700–706.

ATOMS, SPECTRA, RADIATION

Application of the United Atom Model for Estimating the Lifetime of Negative Molecular Ions Relative to Electron Autodetachment

E. P. Nafikova*, N. L. Asfandiarov, A. I. Fokin, and G. S. Lomakin

Institute of Molecular and Crystal Physics, Ufa Scientific Center,
Russian Academy of Sciences, pr. Oktyabrya 151, Ufa, 450075 Bashkortostan, Russia
**e-mail: nail@anrb.ru*

Received February 28, 2002

Abstract—A model proposed for describing the scattering of low-energy electrons (whose energy ranges between thermal energy and several electronvolts) from polyatomic molecules makes it possible to estimate the lifetime of shape resonances. The parameters of the model are determined by specific structural and experimental characteristics of molecules. The results of approximate computations of the lifetimes for negative ions of molecules with different symmetries (diatomic halogens, parabenzoquinone, fullerene C_{60}, benzothiadiazoles, anthraquinone derivatives, and substituted benzene forms) are presented. The obtained data show that the lifetimes are sufficient for the formation of fragment ions observed in the mass spectra of negative ions. © *2002 MAIK "Nauka/Interperiodica".*

1. INTRODUCTION

Resonant electron capture negative ion mass spectrometry [1] developed about forty years ago under the guidance of Khvostenko is a powerful tool for investigating processes of interaction between electrons and molecules. Obviously, the advances in application of any physical method are determined by the level of knowledge of the processes occurring in a given experiment. The variety of mechanisms of interaction of low-energy (up to 15 eV) electrons with polyatomic molecules renders the mass spectroscopy of negative ions formed as a result of resonant capture of electrons very informative. The interpretation of the spectra, which are three-dimensional by nature (mass number of an ion–energy of projectile electrons–intensity), is a nontrivial problem. The dissociative capture cross section $\sigma_D(E)$ can be presented as the product of the capture cross section $\sigma_C(E)$ and the ion survival probability $\rho(E)$ [1]:

$$\sigma_D = \sigma_C(E)\rho(E).$$

The first cofactor is studied in the theory of scattering, which made considerable advances in describing interactions between electrons and molecules [2–5, 13]. However, small molecules remain typical objects of theoretical investigations (except the recent publication by Fabrikant and Hotop [7]). This is not surprising, but does not help the experimenters dealing with complex polyatomic objects. In the interpretation of mass spectra, the resonant states formed as a result of scattering of slow electrons from molecules are of special interest. The most probable are resonant states emerging during the motion of an incoming electron in the field of central forces; these states were called shape resonances [8]. In the literature on scattering theory, the term resonance at a quasi-discrete level is often used. It is very important to obtain at least a very rough quantitative estimate of the lifetime τ of shape resonances. This is so important since the electron-excited states of negative molecular ions are quite stable to the emission of an extra electron as a rule [1]. It is essential to estimate the lifetime of a shape resonance to obtain a correct interpretation in the case when this lifetime directly determines the possibility of its dissociation into fragments. The dissociation is possible if the lifetime exceeds the characteristic period of nuclear vibrations in the ion. Otherwise, dissociation is suppressed by autodetachment, and the survival factor $\rho(E)$ for the ion is too small. It is not completely clear whether a negative molecular ion, which is formed in the electron state stable to dissociation, can be transformed, as a result of a nonradiative transition, into a decay term having a different symmetry during the lifetime of the shape resonance [9]. Theoretical estimates of the probability of nonradiative transitions are difficult to obtain and scarce. The only thing known about its value is that it ranges from 10^{-1}–10^{-2} for a CO_2 molecule to 10^{-6} for complex aromatic systems [10]. Only indirect experimental evidence exists for the nonradiative transition, i.e., observation of the vibrational structure in the threshold excitation spectrum of a negative molecular ion of chlorobenzene [11]. The existence of a vibrational structure only indicates that the lifetime of the ion is longer than several vibrational periods. The existence of a nonradiative transition is of fundamental

importance since it determines the processes of formation and decomposition of negative ions.

The main difficulty encountered in the theoretical description of collisions between slow electrons and polyatomic molecules is that, in the scattering problem, these molecules as targets do not possess spherical symmetry. Tossati and Manini [12] developed an approach to simulation of electron scattering by a C_{60} molecule possessing a nearly spherical symmetry. However, their results were criticized sharply because they disregarded the polarizability of the target molecule [13]. In addition, it should be noted that Tossati and Manini [12] analyzed elastic scattering, while long-lived ($\tau > 10^{-4}$ s [13]) negative molecular ions are observed in experiments; obviously, an electron transfers its energy to the nuclear subsystem during such a long time interval. It will be shown that the above approach deserves attention and leads to the results matching the experimental data if the model parameters typical of scattering process are taken into account correctly.

2. FORMULATION OF THE PROBLEM AND MODEL

A model is simple and universal if it contains a minimal number of measurable physical parameters typical of a scattering process. Insignificant details can be disregarded. Let us formulate the basic concepts of the proposed model.

1. The classical turning point for low-energy electrons is at a considerable distance at which the nonsphericity of the potential is unnoticeable.

2. The electron affinity determines the position of the energy level at which the electron is captured.

3. Boundary orbitals play a decisive role in the process of formation of negative ions.

4. In the process of formation of negative ions, the nodal properties of the molecular orbital to which the capture occurs (i.e., the behavior of the wave function at large distances) play a significant role.

As a basis, we take the model of a spherical potential well [12], which roughly approximates the potential of electron attraction to a molecule. This approach basically differs from that used earlier in that the spherical potential well is not treated as a formal mathematical model, but is interpreted in the united atom approximation [14] introduced by Herzberg in the 1930s. The energy levels appearing in a well of depth U and radius r are compared in this case with the molecular orbitals of the molecule being simulated. This molecule is presented in the form of a united atom, viz., pseudoatom with the nucleus containing a number of protons equal to the sum of the charges of the nuclei of the atoms constituting the molecule.

The first parameter to be determined is the radius of the potential well simulating the target molecule. Obviously, a captured electron occupies one of the vacant molecular orbitals and, hence, is spatially localized in a volume comparable with the volume of the molecule itself. Consequently, it is expedient to choose the radius of the well so that its volume is equal to the volume of the molecule being simulated.

There exist a large number of approaches for determining the intrinsic volume of a molecule. Most of them involve model computation [15]. One of possible approaches, which was described by us in [16, 17], makes use of available structural data. The volume of a molecule can be estimated proceeding from empirical data such as the molecular weight μ and density ρ_0 of a substance: $V = \mu/N_A\rho_0$, where N_A is Avogadro's number. Subsequent calculations of the lifetime proved that the latter method of determining volume is preferable.

The spatial size of a negative ion are obviously slightly larger than the size of a target molecule. In order to take into account the change in the volume in the framework of the model used, we introduced earlier [16] the concept of a united anion, which can be described as follows. The molecule being modeled is treated as a united atom whose volume is equal to the volume of the molecule. For estimating the volume of an ion, we introduce the empirical scale factor $k = k_0\exp(A_0/A)$ determined from the correlation dependence for the ratio of the ionic and atomic radii of halogen atoms, where $k_0 = 1.478$, $A_0 = 3.481$, and A is the atomic number of the united atom. The scale correction in the case of a C_{60} molecule is $k = 1.49$. It should be noted that this value is close to the corrections introduced empirically in [12, 13] to take into account the increase in the size of negative ions as compared to the corresponding molecule. For example, $k = 1.4343$ in [13].

The depth of the potential well is chosen using the following criterion. One of the fundamental parameters of a molecule is its electron affinity (EA), which is equal, by definition, to the difference in the total energies of the molecule and the anion. If EA > 0, the molecule is capable of forming a long-lived negative molecular ion. This quantity is undoubtedly one of the most important characteristics determining the electron scattering by the molecule [1] and, hence, must form the basis of our calculations.

The simplest way to take into account the value of EA of the target molecule is to consider a shallow potential well with only one energy level whose energy is equal in magnitude to the energy of electron affinity of the molecule being modeled (the symmetry of the wave function for this energy level is 1*s*).

We can also choose a potential well of such a depth that it contains $N + 1$ levels, the energy of the $(N + 1)$th level being equal to EA. Here, N is the number of occupied molecular orbitals of the molecule. However, such a choice does not guarantee the proper symmetry of the wave function of the $(N + 1)$th level corresponding to the lower vacant molecular orbital. In this case, propositions 3 and 4 forming the basis of the model are

ignored. The version corresponding to the proposed model can be described as follows.

The orbitals of a united atom correlate with the molecular orbitals of the molecule being simulated [14]. Each molecular orbital is put in one-to-one correspondence with one of atomic orbitals of the united atom or (in our case) with the energy levels in a spherical potential well; the orbital angular momentum of these levels is denoted by L. The order in which the levels appear in the potential well upon an increase in its depth is known from the course of quantum mechanics [18]: 1s, 1p, 1d, 2s, 1f, 2p, 1g, …. For example, the level with $L = 5$ corresponding to the lowest unoccupied orbital T_{1u} of the C_{60} molecule is the ninth level with an energy equal to the reciprocal value of EA for C_{60}: $\varepsilon = -2.666$ eV [19]. Eight deeper lying levels with various degeneracies correspond to occupied molecular orbitals of the fullerene molecule, while the next two levels correspond to vacant molecular orbitals of C_{60} lying in the discrete spectrum (the so-called second and third electron affinity) (Fig. 1). Similar correlations can be found for any molecule if we know the symmetry of its lowest unoccupied molecular orbital. The model takes into account only a part of the occupied and vacant molecular orbitals of the target molecule, but reflects correctly the symmetry of the lowest unoccupied molecular orbital (to be more precise, the nodal properties of the wave function of a given energy level in the united atom approximation).

Thus, we have a simple and unambiguous criterion for choosing the potential well depth, which takes into account the electron affinity energy of the molecule and the symmetry of the lowest unoccupied molecular orbital. It is clear intuitively that this approach is quite justified since the interaction of slow electrons with molecules, in analogy with chemical reactions, is mainly controlled by the boundary orbitals. In this case, there is no need to take into account all electrons in the molecular system. The most illustrative example in this respect is the Hückel method.

The method of expansion of the wave function of a projectile electron in spherical harmonics (partial waves) in the low-energy region makes it possible to estimate the lifetime of a shape resonance. The delay time τ of a scattered electron is determined by the phase shift (1), i.e., the phase difference δ_l between the asymptotic solution to Eq. (2) and the asymptotic solution (3) to the same equation in zero field; ultimately, the value of τ is determined by the potential $V(r)$ [20]:

$$\tau = 2\hbar\frac{d\delta_l}{dE}, \tag{1}$$

$$\frac{d^2u_l}{dr^2} + \left[k^2 - U(r) - \frac{l(l+1)}{r^2}\right]u_l = 0,$$
$$U(r) = \frac{2m}{\hbar^2}V(r), \tag{2}$$

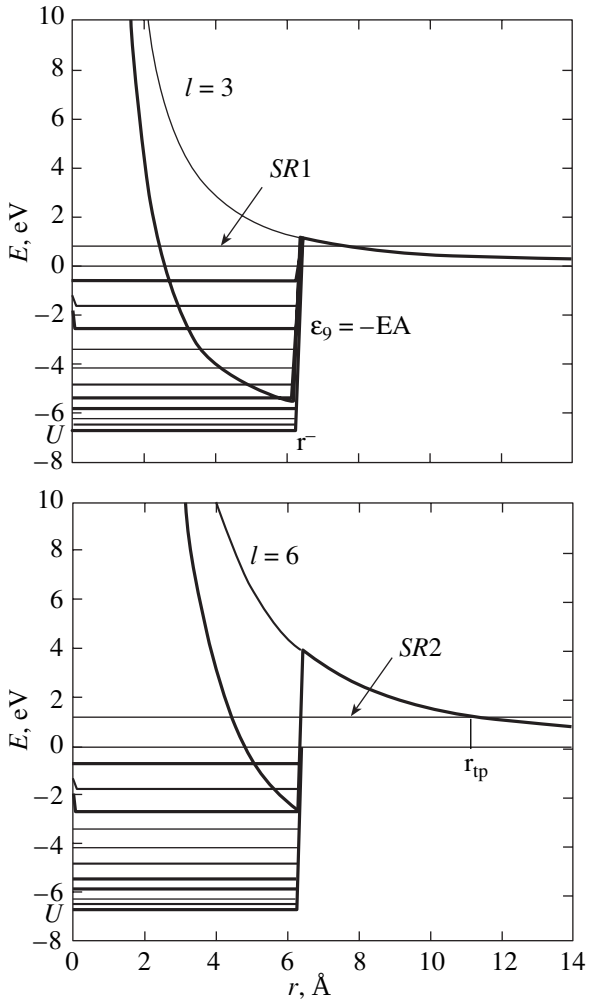


Fig. 1. Models of a spherical rectangular potential well for two shape resonances SR1 and SR2 for scattering of slow electrons by a C_{60} molecule; l is the orbital quantum number for an electron.

$$\frac{d^2v_l}{dr^2} + \left[k^2 - \frac{l(l+1)}{r^2}\right]v_l = 0. \tag{3}$$

As a result of summation of repulsive and attractive forces, an electron finds itself in a potential well separated from the continuum by a finite-width barrier (Fig. 2). The barrier configuration and the electron energy E determine the lifetime of the negative ion. Figure 2 shows the profiles of the electron capture cross sections according to the shape resonance mechanism for a molecule of fullerene C_{60}. It can be seen from the table that matching with the experiment is quite satisfactory. It should be noted that the calculation of cross sections is required for obtaining a quantitative estimate of the lifetime of the shape resonance. It is immaterial that we formally deal with the calculation of the cross section of the electron elastic scattering at a model potential, while in experiment we register nega-

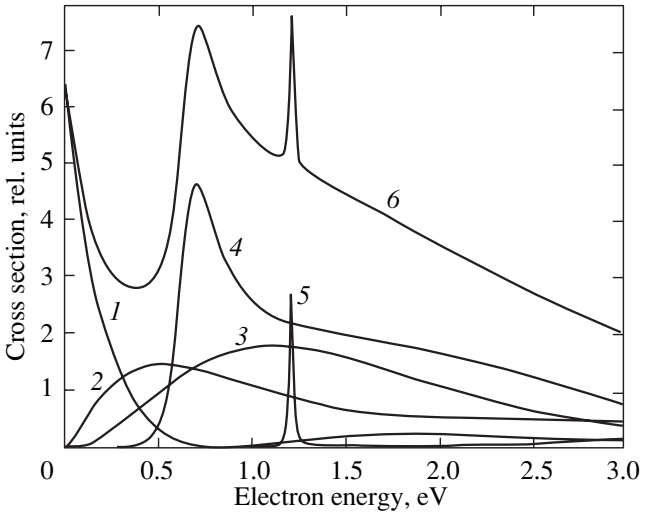


Fig. 2. Total and partial electron capture cross sections according to the shape resonance mechanism for a C_{60} molecule. The orbital quantum number for an electron $l = 0$ (*1*), 1 (*2*), 2 (*3*), 3 (*4*), and 6 (*5*); curve *6* describes the total capture cross section.

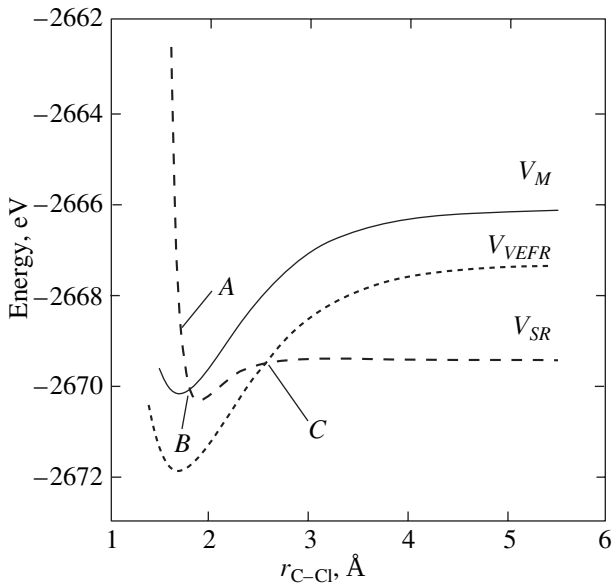


Fig. 3. Schematic representation of the potential energy curves for a molecule and the negative ion for the first two resonances in the case of dissociative capture of an electron as a function of the length of the C–Cl bond of a molecule $1Cl–C_{14}O_2H_8$.

tive fragments of negative ions or long-lived molecular ions, i.e., the result of inelastic scattering. Indeed, Compton [9, 13] demonstrated that, if the lifetime of a shape resonance is longer than the period of vibrations of nuclei in an ion, changes in the geometry of the ion, which render it stable to electron autodetachment, become possible, and the term of the ion lies lower than the term of the molecule (Fig. 3). The ion is formed at point *A* according to the mechanism of a shape resonance. For $\tau > 10^{-14}$ s, the ion starts moving in the term V_{SR} towards the crossing point *B* with the term V_M of the molecule. Having passed point *B*, the ion becomes stable with respect to autodetachment of an electron and can relax, through a nonradiative transition in the vicinity of point *C*, to the ground electron state (term V_{VEFR} of vibration-excited Feshbach resonance). Thus, long-lived negative molecular ions can be observed experimentally as, for example, in the case of anthraquinone derivatives [17, 22]. The estimation of the lifetime of a shape resonance carried out formally for elastic scattering is quite correct.

Thus, the given model has only three parameters: (i) the potential well width, (ii) the potential well depth describing the attraction of an electron to the molecule, and (iii) the symmetry of the wave function of an extra electron in a temporary negative ion. The first two parameters are determined by the volume of the molecule and the electron affinity energy. It should be noted that neither of these parameters is fitting and taken from some spectral data associated with the measurement of the cross section of electron capture by the molecule. Consequently, fitting is ruled out. Nevertheless, the results of estimation of the lifetime of a shape resonance appear to be quite reasonable (see table).

3. RESULTS OF CALCULATION

The developed technique was used for interpreting the mass spectra of negative ions resulting from resonant capture of electrons by the following molecules: diatomic halogen molecules [1–3], parabenzoquinone [4], anthraquinone derivatives [5–11], substituted benzene forms [12–14], fullerene C_{60} [15], and benzothiadiazoles [16, 17]. The results are compiled in the table. In all the above cases, the theoretical values of the resonance energies are in reasonable agreement with experimental data, and the calculated lifetimes of the shape resonance are sufficient for the formation fragment ions observed in the mass spectra of negative ions.

Figure 1 clearly shows that the partial wave with $l = 6$ encounters a barrier at a distance of approximately 12 Å (classical turning point r_{tp}), where the nonsphericity of the scattering potential is not very noticeable. Obviously, the lower the shape resonance energy, the longer the distance at which the interaction between an electron and the molecule begins, and the more reliable the result of calculation. Naturally, we do not claim an exact calculation of the electron capture cross section since the main task is to estimate the lifetime of the shape resonance. The above examples confirm that the model of a united anion provides quite adequate results for large molecules, which can hardly be investigated by exact methods, as in the *R*-matrix-theory.

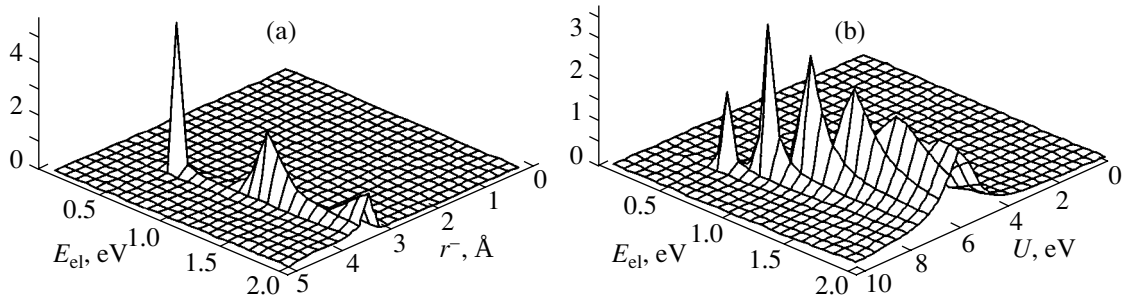


Fig. 4. Capture cross section of the f ($l = 3$) wave of a projectile electron (a) as a function of its energy E_{el} and the potential well radius r^- for a fixed depth $U = 7.9$ eV and (b) as a function of the electron energy and the well depth for a fixed radius $r^- = 3.78$ Å.

It should be emphasized that the methods of calculation of lifetimes of shape resonance are critical to the choice of model parameters. Figure 4 shows the dependences of the electron capture cross section on the radius of the spherical potential well (a) and its depth (b). It can easily be seen that resonance peaks appear only in a very narrow range of radii and only for quite definite depths of the well. This means that these parameters must be chosen with great care. Our experience shows that the best agreement with experiments is attained when the radius of the well is calculated from the data on the density of the substance and using the value the electron affinity measured experimentally by the method of photoelectron spectroscopy of negative ions.

Parameters of model potentials and results of calculation of scattering of slow electrons from polyatomic molecules according to the shape resonance mechanism

No.	Molecule	EA (eV)	U (eV)	r^- (Å)	r_{tp} (Å)	l	E_{exp} (eV)	E_{calc} (eV)	τ_{calc} ($s \times 10^{-15}$)	τ^0_{calc} ($s \times 10^{-15}$)
1	Cl_2	2.40	10.12	2.38	3.5	2	2.50	1.93	3	1.2
2	Br_2	2.42	8.93	2.60	4.2	2	1.40	1.31	5	1.6
3	I_2	2.33	7.52	2.93	5.5	2	0.90	0.75	12	2.2
4	p-$C_6H_4O_2$	1.99	14.02	2.75	5.2	3	1.35	1.69	14	1.4
5	AQ	1.59	8.27	3.72	9.2	3	0.44	0.54	110	3.4
6	2OH–AQ	1.64	8.07	3.79	10.0	3	0.34	0.46	170	3.8
7	$2NH_2$–AQ	1.49	7.80	3.82	9.1	3	0.40	0.54	100	3.6
8	$1NH_2$–AQ	1.46	7.90	3.78	8.7	3	0.56	0.60	80	3.2
9	1Cl–AQ	1.71	8.22	3.77	10.4	3	0.27	0.43	200	3.8
10	2Br–AQ	1.81*	8.04	3.86	13.0	3	0.20	0.27	600	4.6
11	1I–AQ	1.55*	6.82	4.2	14.5	3	0.25	0.21	1200	3.4
12	ClC_6H_5	–0.55*	3.23	2.94	9.5	3	0.74	0.5	6	2.8
13	p-$ClC_6H_4NO_2$	1.26	9.69	3.28	5.9	3	0.90	1.29	15	2.0
14	p-$IC_6H_4NO_2$	2.05*	11.15	3.18	7.4	3	0.7	0.85	60	2.4
15	C_{60}	2.66	6.79	6.32	8.0	3	0.70	0.70	12	5.2
					11.6	6	1.50	1.20	140	4.0
16	$4NH_2$, $5NH_2$–$C_6H_2N_2S$	2.06*	5.67	5.15	9.0	4	1.0	0.93	40	3.6
17	$5NH_2$–$C_6H_3N_2S$	3.22*	7.87	4.55	9.8	4	1.1	0.78	140	3.4

Note: EA is the electron affinity energy of a molecule, U is the potential well depth, r^- is the well radius, r_{tp} is the classical turning point, l is the electron orbital angular momentum, E_{exp} and E_{calc} are the experimental and calculated electron energies corresponding to the resonance state, τ_{calc} is the electron delay time due to resonance in the region of a molecule, τ^0_{calc} is the doubled mean free time for the region in the potential barrier, and AQ = $C_{14}O_2H_8$. Experimental data were obtained in our laboratory (Institute of Molecular and Crystal Physics, Ufa Scientific Center, Russian Academy of Sciences) except for compounds 1–3 [21], 4 [9], and 15 [13].

4. CONCLUSIONS

The model of a spherical potential well is applied for obtaining quantitative estimates of the shape resonance lifetime. The analysis is carried out in the united atom approximation, which makes it possible to single out three parameters determining the electron scattering by molecules: the electron affinity energy, the radius of the well, and the symmetry of the lowest unoccupied molecular orbital. These quantities are "observable" in contrast to the existing approaches using formal model parameters.

For all molecules except diatomic halogen molecules, shape resonances with lifetimes sufficient for the formation of long-lived negative ions are observed in the superthermal energy range. This indicates that the model of a united anion operates the better, the larger the size of the target molecule.

The estimates of the lifetime show that the shape resonance can survive without the emission of an electron untill the beginning of nuclear movements. In accordance with the Compton model [9], a negative molecular ion is transformed, after relaxation to the ground electron state, into a vibrationally excited negative molecular ion with a store of vibrational energy consisting of the electron affinity energy, the thermal energy of the target molecule, and the kinetic energy of a captured electron.

ACKNOWLEDGMENTS

This study was supported financially by the Russian Foundation for Basic Research (project no. 00-02-16578).

REFERENCES

1. V. I. Khvostenko, *Mass-Spectrometry of Negative Ions in Organic Chemistry* (Nauka, Moscow, 1981), p. 159.
2. I. I. Fabrikant, *Dissociative Recombination*, Ed. by B. R. Rove *et al.* (Plenum, New York, 1993), p. 195.
3. G. A. Gallup, Y. Xu, and I. I. Fabrikant, Phys. Rev. A **57**, 2596 (1998).
4. R. S. Wilde, G. A. Gallup, and I. I. Fabrikant, J. Phys. B **32**, 663 (1999).
5. I. I. Fabrikant, S. A. Kalin, and A. K. Kazansky, J. Chem. Phys. **95**, 4966 (1991).
6. J. Simons and K. D. Jordan, Chem. Rev. **87**, 535 (1987).
7. I. I. Fabrikant and H. Hotop, Phys. Rev. A **63**, 022706 (2001).
8. H. S. W. Massey, *Negative Ions* (Cambridge Univ. Press, Cambridge, 1976; Mir, Moscow, 1979).
9. C. D. Cooper, W. T. Naff, and R. N. Compton, J. Chem. Phys. **63**, 2752 (1975).
10. E. E. Nikitin, *Theory of Elementary Atomic-Molecular Processes in Gases* (Khimiya, Moscow, 1970), p. 228.
11. T. Skalicky, C. Chollet, N. Pasquier, and M. Allan, in *Proceedings of the International Symposium on Electron-Molecule Collisions and Swarms, Lincoln, 2001*, Ed. by I. Fabrikant *et al.*, p. 103.
12. E. Tossati and N. Manini, Chem. Phys. Lett. **223**, 61 (1994).
13. J. Huang, H. S. Carman, and R. N. Compton, J. Phys. Chem. **99**, 1719 (1995).
14. G. Herzberg, *Molecular Spectra and Molecular Structure*, Vol. 3 (Van Nostrand, New York, 1966; Mir, Moscow, 1969).
15. L. P. Filippov, *Computing and Forecasting Methods of Properties of Substances* (Mosk. Gos. Univ., Moscow, 1988), p. 252.
16. A. I. Fokin, E. P. Nafikova, G. S. Lomakin, and N. L. Asfandiarov, http://zhurnal.ape.relarn.ru/articles/2000/029, p. 412.
17. N. L. Asfandiarov, A. I. Fokin, V. G. Lukin, *et al.*, Rapid Commun. Mass Spectrom. **13**, 1116 (1999).
18. L. D. Landau and E. M. Lifshitz, *Course of Theoretical Physics*, Vol. 3: *Quantum Mechanics: Non-Relativistic Theory* (Nauka, Moscow, 1989, 4th ed.; Pergamon, New York, 1977, 3rd ed.).
19. C. Brink, L. H. Andersen, P. Hvelplud, *et al.*, Chem. Phys. Lett. **233**, 52 (1995).
20. T.-Y. Wu and T. Ohmura, *Quantum Theory of Scattering* (Prentice-Hall, London, 1962; Nauka, Moscow, 1969).
21. W. C. Tam and S. F. Wong, J. Chem. Phys. **68**, 5626 (1978).
22. N. L. Asfandiarov, V. S. Fal'ko, V. G. Lukin, *et al.*, Rapid Commun. Mass Spectrom. **15**, 1869 (2001).

Translated by N. Wadhwa

Journal of Experimental and Theoretical Physics, Vol. 95, No. 4, 2002, pp. 611–619.
Translated from Zhurnal Éksperimental'noĭ i Teoreticheskoĭ Fiziki, Vol. 122, No. 4, 2002, pp. 707–716.

ATOMS, SPECTRA, RADIATION

Micromaser Operation in Correlated Atoms

V. N. Gorbachev* and A. I. Trubilko**

Northwest State Institute of Printing, St. Petersburg State University of Technology and Design, St. Petersburg, 191180 Russia

**e-mail: vn@vg3025.spb.edu*

***e-mail: tai@at3024.spb.edu*

Received March 4, 2002

Abstract—Stimulated emission of two-level atoms in a high-Q cavity is considered under conditions when pumping produces correlated many-atomic states. Based on the many-body problem, a kinetic equation is obtained for the Glauber–Sudarshan probability, which describes the field in the Fokker–Planck approximation. The statistics of light in this approximation are determined by the atomic correlation functions of the order of no higher than the second. The noise of light is found in the regime of micromaser operation for two types of pumping producing the initial separable states with a classical correlation and entangled states. It is shown that the presence of the initial diatomic correlation enhances the intensity noise. The entangled state of atoms is found from which nonclassical light is generated with a steady-state phase and noise, which can be almost completely suppressed in the low-frequency spectral region. *© 2002 MAIK "Nauka/Interperiodica".*

1. INTRODUCTION

Micromaser operation of light [1, 2] has recently attracted a lot of attention. On the one hand, this is explained by the extensive study of the properties of micromasers and, on the other, by the possibility of the experimental verification of models in which the quantum description of the interaction of atoms with a field is required. The special feature of micromaser sources is the experimental realization of the conditions when the atomic relaxation is hardly manifested for the time of interaction of atoms with the field, so that the contribution from spontaneous processes is small, while stimulated emission dominates. The statistical properties of stimulated emission are determined by atomic correlations or atomic noise, for which an important role belongs to the pump mechanism. Upon regular pumping, a laser or a micromaser can generate light, which is characterized by a strict regularity of a photon flux in time or by sub-Poisson statistics of photons [3]. Such properties of light are of direct practical interest because in this case the noise is virtually absent.

However, the pump can produce correlated atoms. The theory of micromaser operation based on the standard Lamb–Scully approach [4] proves to be invalid for the description of a radiation source pumped in this way. This is explained by the fact that the Lamb–Scully method is based on a one-particle model, where the atoms are assumed to be independent, and their correlation, for example, of the "regular pump" type is introduced phenomenologically [3]. Our approach is based on the initial many-body problem describing the interaction of N two-level atoms with one mode of a high-Q cavity. In paper [5], where the many-body approach was also used, the nontrivial dynamics of operation of a micromaser field was found at the initial (but at the same time quasi-stationary) stage of the evolution for the Poisson statistics of injection of clusters consisting of excited independent atoms. Therefore, the question of how the initial state of atoms affects the properties of radiation is, in our opinion, of independent interest.

The process of obtaining the states with a multiparticle correlation is a separate problem, which we will not consider here. We restrict ourselves only to some types of the states that can be produced, for example, using protocols of the quantum theory of information. In this case, all N atoms are at once transferred to a correlated state, and therefore there is no need to discuss separately the question about the regularization inside such a cluster. Based on the many-body approach, we derive the kinetic Fokker–Planck equation for the field by using the formalism developed in paper [6]. However, because we assume that atoms do not interact with each other, although they are in the correlated state, no additional complications appear in Bardeen–Bogoliubov–Green–Kirkwood–Ivone chains. A special feature of the description of the field in the Fokker–Planck approximation is that the statistics of radiation are determined by atomic correlation functions of the order no higher than the second. It follows from this, in particular, that two different states with the same diatomic density matrix give the same statistics of radiation.

We consider the example of pumping that produces two types of initial states. The first type is a separable state with a classical correlation, which is described by the N-particle density matrix

$$f(1\ldots N) = \lambda_0(|0\rangle\langle 0|)^{\otimes N} + \lambda_1(|1\rangle\langle 1|)^{\otimes N}, \qquad (1)$$

where $\lambda_1 + \lambda_1 = 1$ and $A^{\otimes N}$ is the tensor product $A \otimes A \ldots \otimes A$. In this state, any pair of atoms has a two-particle density matrix of the form

$$f(1, 2) = \lambda_0|00\rangle\langle 00| + \lambda_1|11\rangle\langle 11|,$$

where $\lambda_{0,1}$ are the level populations. It seems that it is impossible to prepare the N-particle state (1) within the framework of a simple pump model, when atoms are irradiated by a classical monochromatic wave. Indeed, consider a set of N independent atoms interacting with a classical wave. The initial entropy in this case is $S = NE_1$, where E_1 is the entropy of an individual atom. Let

$$E_1 = -\lambda_1 \log \lambda_1 - (1 - \lambda_1)\log(1 - \lambda_1).$$

The pumping process can be considered as a unitary evolution retaining entropy, which is $S' = E_1$ for the state described by expression (1). Therefore, no correlations will appear between the atoms. In other words, a classical monochromatic field does not produce correlations.

The states of the second type are inseparable or entangled, the particles having a special quantum correlation. Such states were obtained experimentally [7]. A nontrivial example of the decay of two atoms interacting with a common thermostat, when entangled atomic states appear even in the absence of a direct interatomic interaction, was considered in [8].

One of the states considered by us belongs to the GHZ (Greenberger–Horne–Zeilinger) class:

$$|\mathrm{GHZ}_N\rangle = \alpha|0\rangle^{\otimes N} + \beta|1\rangle^{\otimes N}, \quad (2)$$

where $|\alpha|^2 + |\beta|^2 = 1$. For $\alpha = \beta = 1/\sqrt{2}$, the degree of entanglement is maximum, while, for $N = 2$, the Einstein–Podolsky–Rosen (EPR) pair appears. The state (2) is pure, with a two-particle matrix (1). Assuming $\alpha = 0$, we obtain the case of N independent atoms, each of them being in the state $|1\rangle$, which we consider the upper state for definiteness. If the pump produces such atoms, then micromaser emission will have the sub-Poisson photon statistics, as in the case of a model with regular pumping, the Fokker–Planck equations being completely coincident in these two cases. This means that the monoatomic model of a maser with regular pumping is equivalent to the case considered by us, when all the independent atoms produced upon pumping occupy the upper level.

The second entangled state, which we consider here, has the form

$$|\zeta\rangle = \alpha|bb\rangle + \beta|\Psi^+\rangle, \quad (3)$$

where

$$|\alpha|^2 + |\beta|^2 = 1, \quad |\Psi^+\rangle = \frac{|01\rangle + |10\rangle}{\sqrt{2}},$$

$b = 0, 1$. Each atom in this state has coherence or polarization. Therefore, the phase of light generated in this case will be constant. We found the conditions under which the field generated from the cavity had the sub-Poisson photon statistics.

The main goal of our paper is to derive the kinetic equation for the operation field when pumping produces correlated atoms. Based on the equation obtained, we considered the characteristics of the steady-state micromaser operation.

The paper is organized as follows. In Section 2, we derive the kinetic equation in the Fokker–Planck approximation by the method of adiabatic exclusion of atomic variables. The behavior of atomic averages, which determine the field statistics, is considered in Section 3. In Section 4, we obtain the diffusion coefficients of the kinetic equation for different initial atomic states, In Sections 5 and 6, we calculated noise of light for two types of pumping producing classical and quantum correlation.

2. BASIC EQUATIONS

Consider N identical two-level atoms in a high-Q cavity, which interact with one cavity mode of the radiation field during the time T. Let the frequency of a mode of the electromagnetic field be equal to the working transition frequency. For the micromaser operation, the approximations $1/T \gg \gamma \gg C$ are typical, where γ is the decay rate of atomic levels and C is the decay rate of the field mode. The density matrix F describing atoms and the field at times shorter than T, when the atomic relaxation can be neglected, satisfies the equation

$$\begin{aligned} \frac{\partial}{\partial t}F &= [\vartheta, F], \\ \vartheta &= -i\hbar^{-1}V, \\ V &= -i\hbar g(S_{10}b - S_{01}b^\dagger). \end{aligned} \quad (4)$$

Here,

$$g = d\hbar^{-1}\sqrt{\frac{\hbar\omega}{2L^3\epsilon_0}};$$

d is the transition dipole moment; $b^\dagger$ and b are the operators of photon creation and annihilation, respectively; and the atomic operators S_{xy} are defined by the relations

$$\begin{aligned} S_{xy} &= \sum_{a=1}^{N} s_{xy}(a), \\ s_{xy}(a) &= |x\rangle_a\langle y|, \\ x, y &= 0, 1, \end{aligned} \quad (5)$$

where $|0\rangle_a$ and $|1\rangle_a$ are the upper and lower states of the atom a, respectively. Such collective atomic operators obey the commutation relations

$$[S_{ij}, S_{kl}] = \delta_{jk}S_{il} - \delta_{il}S_{kj}.$$

We will obtain the kinetic equation for the electromagnetic field by introducing for the density matrix F the representation over the coherent states of the electromagnetic field [6]

$$F = \int d^2\alpha\Phi(\alpha)|\alpha\rangle\langle\alpha|. \quad (6)$$

The new density matrix Φ is an operator in atomic variables and a numerical function of the field. Its averaging over atoms leads to the known Glauber–Sudarshan quasi-probability or the P function. We can obtain the equation for Φ [4], taking into account that

$$[\vartheta, F] \longleftrightarrow [\vartheta_0, \Phi] + \partial_\alpha(D\Phi),$$

where the field operators are replaced by the complex amplitudes α and α^* and the derivatives ∂_α from them. The Hamiltonian ϑ_0 obtained in this way describes the interaction of atoms in a "classical" field with the complex amplitude α and has the form $\vartheta_0 = g(S_{01}\alpha^* - \text{c.c.})$. The term containing the derivatives with respect to complex amplitudes is defined by the relation

$$\partial_\alpha(D\Phi) = -\frac{g\partial(S_{01}\Phi)}{\partial\alpha} + \text{H.c.}$$

Let us represent Φ in the form

$$\Phi = P \otimes f + \Pi, \quad (7)$$

where f is the atomic density matrix for which $\mathrm{Sp}_A f = 1$, Π is the correlation density matrix, $\mathrm{Sp}_A\Pi = 0$, and the averaging is performed over atomic variables. For the quasi-probability $P = \mathrm{Sp}_A\Phi$, we can write the equation

$$\frac{\partial}{\partial t}P = -g\frac{\partial}{\partial\alpha}(\langle S_{01}\rangle P + \mathrm{Sp}_A(S_{01}\Pi)) + \text{H.c.} \quad (8)$$

Hereafter, the angle brackets denote the averaging of atomic operators over the state that is described by the density matrix f: $\langle S_{xy}\rangle = \mathrm{Sp}_A(S_{xy}f)$, $x, y = 0, 1$. The problem of the interaction of atoms with the field described by the Hamiltonian ϑ_0 in which the field is represented only by the complex number α has for f the form

$$\frac{\partial}{\partial t}f = [\vartheta_0, f]. \quad (9)$$

Based on (4) and taking into account (8) and (9), we can write the exact equation for Π. However, of great interest is an approximate equation, which gives for the correlation function Π a solution containing derivatives with respect to α of the order no higher than the first. Then, only second-order derivatives appear in equation (8) for the field, and therefore P satisfies the Fokker–Planck equation. This leads to the following problem for Π:

$$\frac{\partial}{\partial t}\Pi = [\vartheta_0, \Pi] - g\left[\frac{\partial}{\partial\alpha}(S_{01} - \langle S_{01}\rangle) f + \text{H.c.}\right]P. \quad (10)$$

Assuming that the characteristic rate of the field development is smaller than that of atoms (this approximation is valid under conditions of the problem), we can obtain a closed kinetic equation for P. Because T is the characteristic time of the interaction of atoms with the field, it is necessary to integrate the equations for P in the interval from t to $T + t$. Let us introduce the large-scale derivative

$$\frac{P(T+t) - P(t)}{T} = \frac{\partial P}{\partial t}.$$

Taking into account that, in the time interval T, the field does not develop, we obtain the kinetic equation

$$\frac{\partial}{\partial t}P(t) = -g\frac{1}{T}\int_t^{T+t} dt' \times\left\{\frac{\partial}{\partial\alpha}[\langle S_{01}(t')\rangle P(t) + \mathrm{Sp}_A(S_{01}\Pi(t')) + \text{H.c.}]\right\}. \quad (11)$$

For simple relaxation models, the decay of a field mode can be taken into account in this equation with the help of an additional term, whose contribution will be presented below. Coefficients in (11) can be determined from the solution to equations for density matrices f and Π for the initial values $f(0)$ and $\Pi(0)$, for example, at $t = 0$. Then, we can obtain the states of the field in the cavity at the next moment T, assuming that T is much smaller than the characteristic time of the atomic relaxation, which is neglected here. We can also specify a state at the moment t, by introducing in this way the pumping mechanism. Then, the atoms prepared by the pump interact with a cavity mode during the time T, which can be achieved in two ways. In the first case, the atoms fly through the cavity for the time T by interaction with the mode. The same time of interaction of different clusters with the mode can be achieved in modern experiments with a high precision. In the second case, Q-switching is performed. In this case, the atoms can be located, for example, in traps in the cavity, and the Q-switching of the cavity is performed at the time intervals T. In both cases, we can set $t = 0$ in the integral in (11), assuming that the atomic state $f(0)$ is specified and the correlation of atoms and the field is $\Pi(0) = 0$.

The methods for producing active atoms play an important role because they determine to a great extent the properties of generated light. In the theoretical description of the pumping mechanism, the method applied for the solution of the problem proves to be important. We derived the master equation within the framework of an *a priori* many-atomic problem, where the pump is introduced as the initial condition for the

atomic density matrix and is explicitly taken into account with the help of such parameters as the atomic-level populations and the number N of atoms. To write the kinetic equation for the field, common approximations should be used, assuming that the time T of the interaction of the atoms with the cavity mode is much shorter than the lifetime of the atoms at working levels and than the photon lifetime in the cavity. However, during times of the order of T, the number of atoms as a parameter specified by the pump can change, for example, arbitrarily [$N = N(t)$]. This circumstance was taken into account, in the case of the pump described by the Poisson statistics, in [3], however, within the framework of the initial monoatomic approach, in which $N = rT$, where $r = r(t)$ is the rate of injection of active atoms to the cavity. This method was used for the description of a regular excitation of atoms by modifying in a phenomenological way the procedure of deriving the equation presented in [3]. Physically, regular pumping means that each atom flying into the cavity is in the upper level, so that $N = rT$ of identical atoms are always found in the cavity after the time T of interaction with the mode. In the case under study, we assume that the number N of atoms after pumping does not depend on time. Physically, this is equivalent to a regular excitation, which produces the sub-Poisson photon statistics. However, we can consider a different optical scheme in which N atoms are in the cavity rather than injected to it. Then, using, for example, Q-switching, we can switch the interaction with the field during the time intervals T separated by the time interval, which is far longer that all other characteristic times, and the initial state of the atoms can be prepared by cloning. The cloning of telecloning protocols allows one to obtain at once N copies or transfer all the atoms to the same state. This property is attractive from the point of view of regular pumping, which solves the problem of preparing many atoms in the same state.

3. ATOMIC AVERAGES

According to (8), the behavior of the electromagnetic field is governed by the atomic averages of two types, $\langle S_p(t)\rangle = \mathrm{Sp}_A(S_p f(t))$ and $\mathrm{Sp}_A(S_p\Pi(t))$, where $p = 0, 1, 2, 3$ or $p = 00, 01, 10, 11$ in the binary representation. The atomic averages describe a semiclassical behavior of light and its noise and are calculated from one- and two-particle density matrices.

The averages of the type $\langle S_p(t)\rangle = N\langle s_p(t; 1)\rangle$ for identical atoms are expressed in terms of monoatomic averages $\langle s_p(t; 1)\rangle = \mathrm{Sp}(s_p f(t; 1))$, which are calculated from the one-particle matrix $f(t; 1) = \mathrm{Sp}_{2\ldots N} f(t; 1\ldots N)$. Here, the total density matrix $f = f(t; 1\ldots N)$ describing N atoms cannot be factorized due to the properties of pumping, which can produce correlated atomic states.

A different situation appears in calculating averages over the correlation matrix Π. By integrating (10), we find that they are determined by the variances of many-atomic operators

$$\mathrm{Sp}_A(S_p\Pi(t)) = -g\int_0^t dt' \times\left[\frac{\partial}{\partial\alpha}D_{p01}(t, t') + \frac{\partial}{\partial\alpha^*}D_{10p}(t', t)\right]P(t), \tag{12}$$

where the two-dimensional correlation functions $D_{pq}(t, t')$ ($p, q = 0, 1, 2, 3$) have the form

$$D_{pq}(t, t') = \langle S_p(t)S_q(t')\rangle - \langle S_p(t)\rangle\langle S_q(t')\rangle, \tag{13}$$

and the Heisenberg operators $S_p(t)$ are averaged over the initial state $f(t = 0; 1\ldots N)$ of atoms. It is assumed that atomic operators in (12) commute with operators of differentiation with respect to the complex amplitude, which is a typical condition for the Fokker–Planck approximation [6] used in the derivation of the kinetic equation. Variances can be expressed in terms of one- and diatomic correlators

$$D_{pq}(t, t') = N[\langle s_p(t; 1)s_q(t'; 1)\rangle - \langle s_p(t; 1)\rangle\langle s_q(t'; 1)\rangle] + N(N-1)[\langle s_p(t; 1)s_q(t'; 2)\rangle - \langle s_p(t; 1)\rangle\langle s_q(t'; 1)\rangle]. \tag{14}$$

Here, to calculate averages appearing in the diffusion coefficients in the kinetic equation, the two-particle matrix $f(t = 0; 12)$ is required along with the one-particle matrix. This means that, in the Fokker–Planck approximation, the noise of light interacting with atoms is determined only by diatomic correlation functions.

The evolution operator for the problem (9) can be readily written in the case under study. The Hamiltonian ϑ_0 is a sum of monoatomic interaction Hamiltonians $\sum_a \vartheta_0(a)$, $\vartheta_0(a) = g(s_{01}(a)\alpha^* - \mathrm{H.c.})$. Therefore, the evolution operator U(1…N) represents a product of one-particle operators, for which the closed expressions

$$U(1\ldots N) = U(1)^{\otimes N}, \qquad U(1) \equiv U = \exp(t\vartheta_0(1)) = \mu + \nu s_{10}(1) - \nu^* s_{01}(1) \tag{15}$$

can be written. Here,

$$\mu = \cos(g|\alpha|t), \quad \nu = -\frac{\alpha}{|\alpha|}\sin(|\alpha|t).$$

By using (15), we can calculate the atomic correlation functions in the Heisenberg representation, in which the one-particle operator has the form

$$s_p(t; 1) = \sum_q R_{pq}(t)s_q(1), \tag{16}$$

where $p, q = 0, 1, 2, 3$ and the unitary matrix R_{pq} is defined by the expression

$$R_{pq} = \begin{pmatrix} \mu^2 & -\mu\nu & -\mu\nu^* & |\nu|^2 \\ \mu\nu & \mu^2 & -\nu^2 & -\mu\nu \\ \mu\nu^* & -\nu^{*2} & \mu^2 & -\mu\nu^* \\ |\nu|^2 & \mu\nu & \mu\nu^* & \mu^2 \end{pmatrix}. \quad (17)$$

By using (16), we can find the variances of atomic operators $D_{pq}(t, t')$, which are expressed in terms of the initial monoatomic averages as

$$D_{pq}(t, t') = \sum_{PQ} R_{pP}(t) R_{qQ}(t') D_{PQ}(0, 0), \quad (18)$$

where $p, q, P, Q = 0, 1, 2, 3$. The initial correlation function $D_{PQ}(0, 0)$ is defined by expression (14), where $t = t' = 0$. Its form is determined by the mechanism of the pump producing the initial state $f(t = 0; 1\ldots N)$ of the atoms.

4. FOKKER–PLANCK EQUATION AND DIFFUSION COEFFICIENTS

The kinetic equation for the field in the Fokker–Planck approximation in polar coordinates $I = |\alpha|^2$, $\varphi = \arg\alpha$ takes the form

$$\frac{\partial}{\partial t} P = \left(\sum_{u, v = I, \varphi} \frac{\partial}{\partial u} A_u + \frac{1}{2} \frac{\partial^2}{\partial u \partial v} Q_{uv} \right) P. \quad (19)$$

Here, coefficients at the first derivatives depend only on one-particle averages

$$A_I = -\frac{g\sqrt{I}}{T} \int_0^T dt \exp(-i\varphi) \langle s_{01}(t; 1) \rangle + \text{H.c.},$$
$$A_\varphi = -\frac{g}{2\sqrt{I}T} \int_0^T dti \exp(-i\varphi) \langle s_{01}(t; 1) \rangle + \text{H.c.}, \quad (20)$$

unlike the diffusion coefficients, which are expressed in terms of diatomic variance

$$Q_{uv} = \theta_{uv} \frac{g^2}{T} \int_0^T dt \int_0^t dt' q_{uv}(t, t'), \quad (21)$$

where $\theta_{uv} = \theta_{vu}$, $\theta_{I\varphi} = 1$, $\theta_{\varphi\varphi} = 1/(4I)$, $\theta_{II} = I$,

$$q_{I\varphi} = q_{\varphi I} = -i \exp(-2i\varphi) D_{11}(t, t') + \text{c.c.},$$
$$q_{\varphi\varphi} = D_{21}(t, t') - \exp(-2i\varphi) D_{11}(t, t') + \text{c.c.}, \quad (22)$$
$$q_{II} = D_{21}(t, t') + \exp(-2i\varphi) D_{11}(t, t') + \text{c.c.}$$

To find the coefficients in the Fokker–Planck equation in the explicit form, it is necessary to specify the initial state $f(t = 0; 1\ldots N)$ of atoms prepared by the pump. By using the evolution law (18), we can obtain general expressions for the diffusion coefficients. Thus,

$$q_{II} = \sum_{Q = 0, \ldots, 3} \{2R_{20}(t) R_{1Q}(t')(D_{0Q} - D_{3Q}) + [2R_{11}(t) - 1] R_{1Q}(t')[D_{2Q} + \exp(-2i\varphi) D_{1Q}] + \text{c.c.} \quad (23)$$

The expression for $q_{\varphi\varphi}$ is obtained from q_{II} by replacing the sign $\exp(-2i\varphi) \longrightarrow -\exp(-2i\varphi)$,

$$q_{I\varphi} = -i \sum_{Q = 0, \ldots, 3} (2R_{11}(t) - 1) R_{1Q}(t') \times \exp(-2i\varphi) D_{1Q} + \text{c.c.} \quad (24)$$

If the emitting atoms are placed in the cavity, we should include in the Fokker–Planck equation [19] the relaxation term, which describes the radiation emergence from the cavity. For simple models, the decay of the electromagnetic field can be taken into account by adding to (4) the relaxation operator in the Lindblood form

$$\mathcal{L}(F) = -C \frac{b^\dagger b F - b F b^\dagger + \text{H.c.}}{2},$$

where the constant $C/2$ is determined by the transmission coefficients of cavity mirrors and is equal to the decay rate of the field amplitude. The consideration of the field decay leads to the substitution $A_I \longrightarrow A_I + CI$. As a result, steady-state operation regimes can appear, in which the diffusion coefficients directly determine the noise of light.

One of the main characteristics of the radiation statistics is the Mandel parameter ξ, which determines the deviation of the variance of the number of photons from the Poisson distribution:

$$\langle n^2 \rangle - \langle n \rangle^2 = \langle n \rangle (1 + \xi),$$

where $n = b^\dagger b$ is the operator of the number of photons. In the representation adopted, $\xi = \langle \epsilon^2 \rangle / \langle I \rangle$, where $\epsilon = I - \langle I \rangle$ describes the deviation of the intensity from its average value

$$\langle I \rangle = \int d^2\alpha P(\alpha) |\alpha|^2.$$

If the intensity fluctuations are small and are independent of phase fluctuations, then it follows from (19) that the diffusion coefficient Q_{II} determines the Mandel parameter

$$\xi = Q_{II} (\langle I \rangle \Gamma)^{-1}, \quad (25)$$

where

$$\Gamma = C + \left(\frac{\partial A_i}{\partial I}\right)_{I = \langle I\rangle}$$

is the rate of decay of intensity fluctuations. If a field state is classical, then its Mandel parameter is nonnegative. For nonclassical states, $\xi < 0$, and therefore the main criterion for the presence of the sub-Poisson statistics is the inequality $Q_{II} < 0$.

5. OPERATION IN ATOMS WITH A CLASSICAL CORRELATION

Let us assume that the pump produces correlated atoms in the states of type (1) or (2) for $N \geq 3$. Then,

$$f(t = 0, 1) = \lambda_0|0\rangle\langle 0| + \lambda_1|1\rangle\langle 1|, \quad (26)$$

$$f(t = 0, 12) = \lambda_0|00\rangle\langle 00| + \lambda_1|11\rangle\langle 11|, \quad (27)$$

where $\lambda_0 = |\alpha|^2$ and $\lambda_1 = |\beta|^2$ are the populations of the lower and upper levels, respectively. In this case, the diffusion coefficients in [19] take the form

$$Q_{II} = \frac{N}{T}\left[-\frac{1}{2}\sin^4 b + (2\lambda_1 - 1)b\sin b\cos b + (1-\lambda_1)\sin^2 b + 2\lambda_1(1-\lambda_1)\sin^4 b\right] + \mathcal{W},$$
$$Q_{\varphi\varphi} = \frac{N}{8I^2T}[b^2 + 2\lambda_1\sin^2 b - \sin^2 b], \quad (28)$$
$$Q_{I\varphi} = 0,$$

where $b = g\sqrt{I}\,T$ and the contribution from the diatomic correlation, which is produced by pumping, for both states is represented by the term

$$\mathcal{W} = \frac{N}{T}2(N-1)\lambda_1(1-\lambda_1)\sin^4 b.$$

If all the atoms are independent or only one atom is present, then $\mathcal{W} = 0$. Note that the atomic correlation in the example under study, which is produced by pumping, affects only the intensity noise, while the phase undergoes a usual diffusion and proves to be insensitive to these properties of the initial state of the atoms.

Consider the question about the appearance of nonclassical states of the operation field. The master equation (19), in which diffusion coefficients are defined according to (28), describes the steady-state regime of maser operation, which can be analyzed in the approximation of small fluctuations. Then, assuming that $\epsilon = I - \langle I\rangle \ll \langle I\rangle$, where the average intensity can be determined from the semiclassical operation equation

$$\frac{\partial}{\partial t}\langle I\rangle = -C\langle I\rangle + \frac{N}{T}(\lambda_1 - \lambda_0)\sin^2 B, \quad (29)$$

we obtain for the Mandel parameter the expression

$$\xi = \frac{Q_{II}}{(2\lambda_1 - 1)\sin^2 B(1 - B\cot B)}\frac{T}{N}, \quad (30)$$

where $B = g\sqrt{\langle I\rangle}\,T$. By using (30), we find the photocurrent spectrum or the noise of light, which is detected with a photodetector in the usual scheme of homodyne detection

$$i^{(2)} = 1 + 2\xi\frac{1}{1 - B\cot B}. \quad (31)$$

Here, the shot noise level is assumed equal to unity, and a low-frequency part of the spectrum near the zero frequency is chosen. For the operation regime to be stable, the conditions $\lambda_1 > \lambda_0$ and $1 - B\cot B > 0$ should be satisfied.

Due to the positive contribution $\mathcal{W}$ of diatomic correlations produced by pumping, the Mandel parameter is always nonnegative, so that no nonclassical states appear.

Let $\mathcal{W} = 0$. This means, for example, that the pump produces N independent atoms in a pure state, which occupy the upper level $f(t = 0, 1\ldots N) = (|1\rangle\langle 1|)^{\otimes N}$. Then, due to the negative term $-\sin^4 b/2$ in the expression for the diffusion coefficient Q_{II}, $\xi \longrightarrow -1/2$, and the shot noise can be almost completely suppressed. For the maser operation regime, this result was obtained assuming regular pumping, which was introduced phenomenologically [3]. It follows from the above discussion that nonclassical states of the field with suppressed noises appear when the pump produces independent atoms in a pure state at the upper level, i.e., when $\lambda_1 = 1$, or there is only one excited atom ($N = 1$, $\lambda_1 = 1$) in the interaction region. If the initial state of atoms has a classical correlation of the type (27), then this correlation enhances the operation-field noise by destroying a nonclassical state of light.

6. OPERATION IN ENTANGLED ATOMS

Recall that micromaser operation can occur even when the number of atoms in the active region is small [2]. It would be interesting, in our opinion, to study the characteristics of emission of two entangled atoms with a quantum correlation of the type of the EPR pair, whose state is described by the function $|\mathrm{ENT}\rangle = \alpha|00\rangle + \beta|11\rangle$, which is obtained from (2) for $N = 2$. The nonclassical nature of the pair will be manifested the most completely when the particles are entangled to the greatest extent ($|\alpha| = |\beta| = 1/\sqrt{2}$). The latter equality means equal populations of the upper and lower levels of an atom in a entangled state with the density matrix of type (26), where $\lambda_0 = \lambda_1 = 1/2$. Such an atom has no polarization or coherence, so that no steady-state oper-

ation appears in most strongly entangled EPR atoms in the above approximations.

Consider now the entangled atomic state ζ defined by expression (3), for which

$$f(t = 0; 1) = |\alpha|^2|b\rangle\langle b| + \frac{|\beta|^2}{2}(|0\rangle\langle 0| + |1\rangle\langle 1|) + \frac{\alpha^*\beta}{\sqrt{2}}(b|0\rangle\langle 1| + (1-b)|1\rangle\langle 0|) + \frac{\alpha\beta^*}{\sqrt{2}}((1-b)|0\rangle\langle 1| + b|1\rangle\langle 0|), \quad f(t = 1, 12) = |\zeta\rangle\langle\zeta|, \tag{32}$$

where $b = 0, 1$. In this state, in an atom, along with the coherence

$$\langle s_{01}(0)\rangle = \frac{1}{\sqrt{2}}[(1-b)\alpha^*b + b\alpha\beta^*] = \frac{\mathcal{U}}{\sqrt{2}}[(1-b)\exp(i\Psi_0) + b\exp(-i\Psi_0)],$$

where $\mathcal{U} = |\alpha\beta|$, there exists the population inversion at the working transition:

$$\langle s_{11}(1) - s_{00}(1)\rangle = (-1)^{1-b}|\alpha|^2.$$

For $b = 1$, the working transition will be amplifying. The specific feature of the atomic states under study is manifested already in calculation of monoatomic averages. The values of populations and polarization for a two-level atom cannot be arbitrary. Thus, the maximum value of coherence, which appears upon irradiation of the atom by a classical monochromatic wave, is $|\langle s_{10}(1)\rangle| \le 1/2$. In the case under study, the value of polarization is lower by a factor of $\sqrt{2}$, which is typical for a entangled state.

The presence of an atomic coherence leads to two specific features. First, a steady-state operation can appear if the population of the upper level is smaller than that of the lower level. Second, because of the phase locking, instead of diffusion the distribution of the phase φ of the generated field becomes steady-state. The phase distribution yields the equation

$$\frac{\partial}{\partial t}\varphi = -\sqrt{2}\frac{g}{\sqrt{I}}\mathcal{U}\sin(\varphi - \Psi_0), \tag{33}$$

which has a steady-state solution leading to a certain phase in the steady-state operation regime:

$$\varphi_0 = \Psi_0. \tag{34}$$

This equality follows from the condition of the steady-state-phase stability. In the approximation of small phase fluctuations, $\mu = \varphi - \varphi_0$, $\mu \ll 1$, we have

$$\frac{\partial}{\partial t}\mu = -\Gamma_\mu\mu, \tag{35}$$

where the coefficient

$$\Gamma_\mu = \sqrt{2}\frac{g}{\sqrt{I}}\mathcal{U}\cos(\varphi_0 - \Psi_0)$$

determines the decay rate of phase fluctuations and should always be positive, $\Gamma_\mu > 0$.

The initial atomic coherence results in the appearance of the additional term in the equation for the average intensity of radiation [cf. (29)]:

$$\frac{\partial}{\partial t}\langle I\rangle = -C\langle I\rangle + \frac{1}{T}2(-1)^{1-b}|\alpha|^2\sin^2 B + \frac{1}{T}2\sqrt{2}\mathcal{U}\sin B\cos B. \tag{36}$$

Here, the first term corresponds to the mode decay caused by emission from the cavity. The second term describes emission or absorption of a photon due to inversion at the working transition, while the influence of the atomic coherence is represented by the last term. Note that, in (36), we have already used the steady-state condition (34) for the phase. The gain, which is determined by the sum of the last two terms in (36), differs from that calculated for the monoatomic model [9]. First, in the case under study, the parameter α cannot be zero, and therefore operation of light based directly on the atomic coherence is impossible. Second, the term describing the influence of coherence differs by a factor of $\sqrt{2}$ from the value obtained in other known cases, which is caused by the specific properties of the state (3). For $b = 0$, the second term in (36) becomes negative, which does not mean, however, that the operation is impossible. As in a usual situation for a micromaser, a series of almost periodic steady-state states appear during the operation at any of the transitions under study. The stability of these states can be analyzed based on the equation for the intensity fluctuations

$$\frac{\partial}{\partial t}\epsilon = -\Gamma_\epsilon\epsilon, \tag{37}$$

where, taking into account the steady-state condition $\partial\langle I\rangle/\partial t = 0$, the decay rate of the intensity fluctuations

$$\Gamma_\epsilon = C\left\{1 - B(-1)^{1-b} \times \frac{2|\alpha|^2\cot B + (-1)^{1-b}\sqrt{2}\mathcal{U}(\cot^2 B - 1)}{2(-1)^{1-b}|\alpha|^2 + \sqrt{2}\mathcal{U}\cot B}\right\} \tag{38}$$

obeys the condition $\Gamma_\epsilon > 0$.

The diffusion coefficients of the Fokker–Planck equation (19), which determine the statistical charac-

teristics of the generated radiation, taking into account the steady-state solutions, have the form

$$Q_{II} = \frac{1}{T}\{2|\beta|^2\sin^2 B\cos^2 B + |\alpha|^2[\sin^2 B(\sin^2 B + 2(1-b)) + (-1)^{1-b}B\sin(2B)] + \sqrt{2}\mathcal{U}[\cos(2B)(B-(-1)^{1-b}\sin B\cos B) - (1-b)\sin 2B] - \mathcal{Q}\}, \quad (39)$$

$$Q_{\varphi\varphi} = \frac{1}{4I^2T}\{2|\beta|^2B^2 + |\alpha|^2[B^2 + (-1)^{1-b}\sin^2 B] + \sqrt{2}\mathcal{U}(\sin B\cos B - B)\},$$

$$Q_{I\varphi} = 0,$$

where

$$\mathcal{Q} = 2\{(\sqrt{2}\mathcal{U})^2\sin^2 B\cos^2 B - \sqrt{2}\mathcal{U}|\alpha|^2\sin B\cos B[1-\cos(2B)] + |\alpha|^4\sin^4 B\}.$$

The independence of the intensity and phase fluctuations follows from the steady-state condition (34). Our calculations showed that, for both initial states of the atoms considered here and for all allowed steady-state states of the light field, the diffusion coefficient $Q_{\varphi\varphi}$ is always positive. This means that no operation of phase-squeezed light occurs, which is caused by a linear phase locking of the atomic coherence and the field.

By using (25), where $\Gamma \longrightarrow \Gamma_\epsilon$, we will analyze the radiation statistics. Of direct interest for observation is not the Mandel parameter ξ but the value of the photocurrent from the emission detected at the cavity output in the low-frequency region of the noise spectrum

$$i^{(2)} = 1 + 2\xi\frac{C}{\Gamma_\epsilon}. \quad (40)$$

Here, unity still corresponds to the detected shot noise, and the quantum efficiency of a photodetector is assumed equal to unity. The negative Mandel parameter causes the suppression of the shot noise, which implies the appearance of a nonclassical state of light with the sub-Poisson photon statistics. The numerical analysis of the expressions obtained suggests the following. In the case of the initial state of atoms with the negative inversion at the transition, $b = 0$, of the most interest are the steady-state conditions with the maximum value of the gain profile, which are observed for the atomic coherence $\mathcal{U} \approx 0.49$. In this case, a series of almost periodic operation regimes are realized, for which $i^{(2)} = 0.27$, corresponding to the sub-Poisson light with a noise suppression of 63%. This value is optimal for such an initial atomic state. Recall that, in the monoatomic model [9], the optimum ultimate value of the noise suppression is $i^{(2)} = 0.5$. This improvement is associated namely with the state under study and the manifestation of its characteristics both in steady-state semiclassical equations (33), (36) and in the diffusion coefficients. For an amplifying transition, $b = 1$, and expression (39) for $\beta = 0$ transforms to (28) for $N = 2$ independent atoms. Recall that for independent atoms, the conditions can be created under which the noise is almost completely suppressed. In the case under study, the noise is also almost completely suppressed in different almost periodic steady-state regimes. The noise suppression is achieved by preparing the initial state in such a way that only the upper level is predominantly populated for both atoms ($|\alpha|^2 \gg |\beta|^2$). The presence of even weak coherence at the transition results in a strictly definite phase of the operation field, while the negative diatomic contribution produces the sub-Poisson photon statistics and almost complete suppression of the shot noise ($i^{(2)} \longrightarrow 0$).

7. CONCLUSIONS

We have obtained the kinetic equation for a single-mode electromagnetic field interacting with N two-level atoms in a high-Q cavity. Although it is assumed that the atoms do not interact with each other, their initial state prepared by pumping can be uncorrelated in the general case. This requires the consideration of a many-atomic problem. The kinetic equation was written for the Glauber P function in the Fokker–Planck approximation, where coefficients are determined only by one- and diatomic correlation functions. This means that, in this approximation, all the many-atomic states with identical one- and two-particle density matrices give the same radiation statistics. We considered the micromaser operation when the relaxation from the working levels can be neglected. The presence of a classical diatomic correlation enhances the noise of the generated light compared to the case of independent atoms whose emission has the sub-Poisson photon statistics. In a special case of a entangled state of type (3), a field can be generated with a certain steady-state phase and ultimately suppressed noise of the homodyne detection.

ACKNOWLEDGMENTS

This work was partially supported by the Delzell Foundation, the Russian Foundation for Basic Research (project no. 01-02-17059), and the INTAS program (grant no. 00-479).

REFERENCES

1. P. Filipowicz, J. Javanainen, and P. Meystre, Phys. Rev. A **34**, 3077 (1986); J. Krause, M. O. Scully, and H. Walther, Phys. Rev. A **34**, 2032 (1986); L. Lugiato, M. O. Scully, and H. Walther, Phys. Rev. A **36**, 740 (1987).
2. H. Walther, Usp. Fiz. Nauk **166**, 777 (1996) [Phys. Usp. **39**, 727 (1996)].

3. Yu. M. Golubev and I. V. Sokolov, Zh. Éksp. Teor. Fiz. **87**, 408 (1984) [Sov. Phys. JETP **60**, 234 (1984)]; F. Haake, S. M. Tan, and D. F. Walls, Phys. Rev. A **34**, 4025 (1986); Yu. M. Golubev, Zh. Éksp. Teor. Fiz. **106**, 1031 (1994) [JETP **79**, 561 (1994)].

4. M. O. Scully and W. E. Lamb, Phys. Rev. **159**, 208 (1967).

5. G. P. Miroshnichenko, I. P. Vadeĭko, A. V. Rybin, and J. Timonen, Pis'ma Zh. Éksp. Teor. Fiz. **72**, 647 (2000) [JETP Lett. **72**, 449 (2000)].

6. V. N. Gorbachev and A. I. Zhiliba, Quantum Opt. **5**, 193 (1993).

7. B. Julsgaard, A. Kozhekin, and E. S. Polzik, Nature **413**, 400 (2001); quant-ph/0106057.

8. A. M. Basharov, Pis'ma Zh. Éksp. Teor. Fiz. **75**, 151 (2002) [JETP Lett. **75**, 123 (2002)].

9. V. N. Gorbachev and A. I. Trubilko, Zh. Éksp. Teor. Fiz. **115**, 1605 (1999) [JETP **88**, 882 (1999)]; Opt. Spektrosk. **89**, 460 (2000) [Opt. Spectrosc. **89**, 384 (2000)].

Translated by M. Sapozhnikov

Journal of Experimental and Theoretical Physics, Vol. 95, No. 4, 2002, pp. 620–624.
Translated from Zhurnal Éksperimental'noĭ i Teoreticheskoĭ Fiziki, Vol. 122, No. 4, 2002, pp. 717–722.

ATOMS, SPECTRA, RADIATION

Theoretical Investigation of the $p + He \longrightarrow H + He^+$ and $p + He \longrightarrow H + He^{++} + e$ Reactions at Very Small Scattering Angles of Hydrogen

Yu. V. Popov[a, *], O. Chuluunbaatar[b], S. I. Vinitsky[b], L. U. Ancarani[c], C. Dal Cappello[c], and P. S. Vinitsky[d]

[a]*Institute of Nuclear Physics, Moscow State University, Vorob'evy gory, Moscow, 119899 Russia*
[b]*Joint Institute for Nuclear Research, Dubna, Moscow oblast, 141980 Russia*
[c]*Institute of Physics, University of Metz, Metz, France*
[d]*Moscow State University, Vorob'evy gory, Moscow, 119899 Russia*
**e-mail: popov@srdlan.npi.msu.su*
Received March 18, 2002

Abstract—A hypothesis is considered that the reactions $p + He \longrightarrow H + He^+$ and $p + He \longrightarrow H + He^{++} + e$ at very small scattering angles of hydrogen can be used for the angular spectroscopy of electron correlations in a target. It is shown that this hypothesis is inconsistent. *© 2002 MAIK "Nauka/Interperiodica".*

1. INTRODUCTION

A number of sufficiently fine experiments have been carried out over the last several years on the capture of an electron by a fast proton from a helium target, including those with a simultaneous transfer ionization process. With the use of a unique cold target recoil ion momentum spectroscopy (COLTRIMS), all independent kinematic characteristics of the final products of reactions were measured: the polar and azimuthal angles of hydrogen, as well as the momentum and energy of a He^{++} ion [1–3]. The azimuthal scattering angle θ_p of hydrogen was as small as 0.1–0.5 mrad, which is about 100 times less than that obtained in earlier experiments of this type. The proton energy E_p ranged from 0.15 to 1.4 MeV.

The singly differential cross section $d\sigma/d\theta_p$ of these reactions in the range of θ_p from 10 to 1000 mrad represents a sufficiently smoothly and rapidly decreasing function, which is satisfactorily described within the continuum distorted wave formalism [4, 5]. At angles of θ_p = 0.1–0.3 mrad, this curve attains its principal maximum (not counting relatively small Thomas peaks at appropriate scattering angles [6, 7] greater than 0.5 mrad).

The authors of the experimental works [1–3] suggested that the reaction $p + He \longrightarrow H + He^{++} + e$ in this range of extremely small scattering angles θ_p can be used for obtaining new and extraordinary information on the structure of the wave function of the target in the momentum representation. The present paper is devoted to the theoretical analysis of this concept. Throughout this paper, we use atomic units.

2. THEORY

For brevity, we will call the reaction

$$p + He \longrightarrow H + He^+ \tag{1}$$

a simple capture (SC) reaction and

$$p + He \longrightarrow H + He^{++} + e \tag{2}$$

a transfer ionization (TI) reaction. We will also use the following notation: $\mathbf{v}_p(\mathbf{p}_p)$ is the proton velocity (momentum), $\mathbf{v}_H(\mathbf{p}_H)$ is the hydrogen velocity (momentum), $\mathbf{k}$ is the momentum of the escaped electron, $\mathbf{K}$ is the momentum of the residual ion, and E is the total energy of the system. In atomic units, the proton mass is m = 1836.15 and the ion mass is $M \approx 4m$. In addition, we introduce the transferred momentum

$$\mathbf{q} = \mathbf{p}_H - \mathbf{p}_p = (m+1)\mathbf{v}_H - m\mathbf{v}_p.$$

First, consider the TI reaction. In the notation introduced, the energy and momentum conservation laws in the laboratory system of coordinates are expressed as

$$\mathbf{K} + \mathbf{k} + \mathbf{q} = 0 \tag{3}$$

and

$$E = \frac{p_p^2}{2m} + E_0^{He} = \frac{p_H^2}{2(m+1)} + \frac{k^2}{2} + \frac{K^2}{2M} + E_0^H, \tag{4}$$

respectively. Here, $E_0^{He} = -2.903$ and $E_0^H = -0.5$. For convenience, we introduce the quantity $Q = E_0^{He} - E_0^H = -2.403$.

The proton energy E_p ranges from 0.15 to 1.4 MeV, which corresponds to v_p = 2.45–7.49. At the same time,

it follows from the experiments that the measured values of the ion momentum and the transferred momentum at very small angles θ_p only amount to a few atomic units; this fact allows one to neglect the energies $K^2/2M$ and $q^2/2m$ as compared with other terms in Eq. (4). We stress that this can only be done at very small angles $\theta_p = 0.1$–0.5 mrad, when the helium ion remains at rest. At larger scattering angles, the proton–nucleon (pN) Coulomb interaction begins to play an increasing role, which significantly increases the transferred momentum and the momentum of the residual ion, which starts to move. Under the approximations made, it follows from (4) that

$$\mathbf{v}_p\mathbf{q} = \frac{1}{2}v_p^2 - \frac{1}{2}k^2 + Q. \tag{5}$$

If we choose the proton velocity vector as the z axis, we have $\mathbf{q} = (\mathbf{q}_\perp, q_z)$, where

$$q_z = \frac{v_p}{2} - \frac{k^2 - 2Q}{2v_p} \tag{6}$$

and $q_\perp = mv_p\sin\theta_p \approx mv_p\theta_p$. Note in passing that Eqs. (3)–(5) allow one to determine the total momentum of the electron provided that the momentum of the ion has been measured, to obtain appropriate constraints, etc.

For the SC reaction, Eqs. (3) and (4) are rewritten as

$$\mathbf{K} + \mathbf{q} = 0 \tag{7}$$

and

$$\frac{p_p^2}{2m} + E_0^{\mathrm{He}} = \frac{p_\mathrm{H}^2}{2(m+1)} + \frac{K^2}{2(M+1)} + E_0^{\mathrm{H}} + E_0^{\mathrm{He}^+}, \tag{8}$$

where $E_0^{\mathrm{He}^+} = -2$. Here, it is convenient to introduce a quantity $Q' = E_0^{\mathrm{He}} - E_0^{\mathrm{H}} - E_0^{\mathrm{He}^+} = -0.403$. Equation (6) is modified as follows:

$$q_z = \frac{v_p}{2} + \frac{Q'}{v_p}. \tag{9}$$

Now, we pass on to the dynamics of the processes. Let us write the Hamiltonian of the system p + He as

$$\mathcal{H} = \mathcal{H}_p + \mathcal{H}_{\mathrm{He}} + V_{p\mathrm{He}}, \tag{10}$$

where

$$\begin{aligned} \mathcal{H}_p &= p_p^2/2m, \\ \mathcal{H}_{\mathrm{He}} &= k_1^2/2 + k_2^2/2 + V_{Ne_1} + V_{Ne_2} + V_{ee}, \\ V_{p\mathrm{He}} &= V_{pe_1} + V_{pe_2} + V_{Np}, \\ V &= V_{p\mathrm{He}} + V_{Ne_1} + V_{Ne_2} + V_{ee}. \end{aligned} \tag{11}$$

Next, we use the following notation: $|\Phi_0\rangle$ is the wave function of the helium atom at rest in the ground state; $|\mathbf{p}_\mathrm{H}, \varphi_\mathrm{H}\rangle$ is the wave function of the hydrogen atom in the ground state that moves at velocity $\mathbf{v}_\mathrm{N}$; and $|\mathbf{K}, \varphi^-(\mathbf{k})\rangle$ is the wave function of the electron in the field of the ion He^{++} with momentum $\mathbf{K}$ (in the case of the SC reaction, one should take the function $|\varphi_0\rangle$ of the bound electron instead of $|\varphi^-(\mathbf{k})\rangle$).

We consider a helium atom in a singlet state; therefore, it is convenient to represent the amplitude of the TI reaction with regard to all the necessary symmetries as follows:

$$\mathcal{T} = \langle \mathbf{p}_p\Phi_0|V_{p\mathrm{He}}[1 + G(E)V_{\mathrm{out}}]|\Psi_{\mathrm{out}}\rangle, \tag{12}$$

where

$$G(E) = (E - \mathcal{H}_p - \mathcal{H}_{\mathrm{He}} - V + i\varepsilon)^{-1}$$

is the full Green's function of the problem and $|\Psi_{\mathrm{out}}\rangle$ is determined from the equation

$$[E - H_0 - (V - V_{\mathrm{out}})]|\Psi_{\mathrm{out}}\rangle = 0,$$

where $V_{\mathrm{out}} = V_{ee} + V_{pN}$.

The amplitude (12) is exact, and the final state represents the wave function of two noninteracting electrons in the field of two centers that move relative to each other. Here, we consider the approximation of this complicated function by its asymptotic value, i.e., by the normalized combination of functions

$$\langle \mathbf{r}_p, \mathbf{r}_1, \mathbf{r}_2, \mathbf{r}_N|\mathbf{p}_\mathrm{H}, \varphi_\mathrm{H}; \mathbf{K}, \varphi^-(\mathbf{k})\rangle,$$

which is symmetric with respect to the electron coordinates. Even this simplification leaves open the problem of correct normalization, which requires the orthogonalization of all components. However, considering the velocity v_p as a large parameter in the problem, one can neglect the cross-terms in the normalization integral and obtain an approximate normalization factor of $N = 1/\sqrt{2}$.

Formula (12) clearly shows that the term in the first Born approximation in $V_{p\mathrm{He}}$ in problems with rearrangement is largely determined by the choice of the output potential V_{out}, i.e., by the interaction in the final state; therefore, this term does not provide a suitable approximation for the amplitude $\mathcal{T}$ even for a large energy E. However, we will consider this term because it is this term that includes the simplest mechanism when one electron is captured immediately by a proton from an atom, while the other is emitted by the He^+ ion due to the shake-off process of the internal electric field in the atom. Indeed, leaving only $V_{p\mathrm{He}}$ in (12) and calculating the matrix element, we obtain

$$\begin{aligned} \mathcal{T}_0 = -4\pi\sqrt{2}\int\frac{d\mathbf{x}}{(2\pi)^2}\frac{\tilde{\varphi}_\mathrm{H}(\mathbf{x})}{|\mathbf{v}_p - \mathbf{q} - \mathbf{x}|^2}[F(\mathbf{q}; 0; \mathbf{k}) \\ + F(\mathbf{v}_p - \mathbf{x}; -\mathbf{v}_p + \mathbf{q} + \mathbf{x}; \mathbf{k}) - 2F(\mathbf{v}_p - \mathbf{x}; 0; \mathbf{k})], \end{aligned} \tag{13}$$

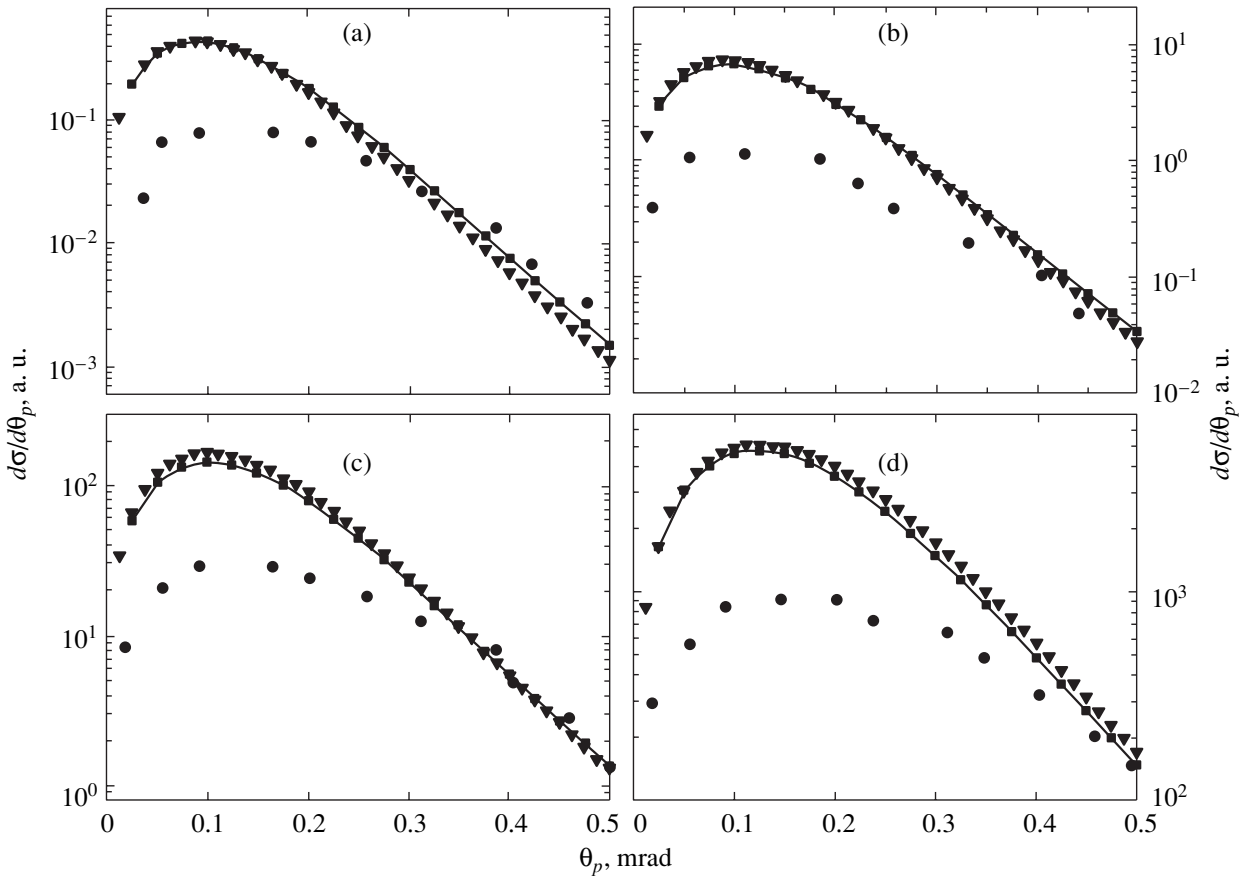


Fig. 1. Singly differential cross section $d\sigma/d\theta_p$ for the reaction $p + \mathrm{He} \longrightarrow \mathrm{H} + \mathrm{He}^+$ calculated with the use of the following functions: (solid curve) CVP, (squares) BK, and (triangles) Hy for four different values of the collision energy E_p: (a) 1.4, (b) 0.8, (c) 0.4, and (d) 0.15 MeV. Relative error of the experiment (circles) in case (a) is no greater than 10% in the range of angles considered.

where

$$F(\mathbf{y};\,\boldsymbol{\eta};\,\mathbf{k}) = \int \exp(-i\mathbf{y}\cdot\mathbf{r}_1 - i\boldsymbol{\eta}\cdot\mathbf{r}_2) \times \varphi^{-*}(\mathbf{k},\mathbf{r}_2)\Phi_0(\mathbf{r}_1,\mathbf{r}_2)d\mathbf{r}_1 d\mathbf{r}_2 \quad (14)$$
$$= \int \frac{d\boldsymbol{\xi}}{(2\pi)^3}\tilde{\varphi}^{-*}(\mathbf{k},\boldsymbol{\xi})\tilde{\Phi}_0(\mathbf{y},\boldsymbol{\eta}+\boldsymbol{\xi})$$

(the tilde over the functions denotes the momentum representation of these functions). Invoking the equation for the hydrogen wave function in the momentum representation,

$$\left(E_0^{\mathrm{H}} - \frac{x^2}{2}\right)\tilde{\varphi}_{\mathrm{H}}(\mathbf{x}) = \int \frac{d\mathbf{x}'}{(2\pi)^3}\frac{-4\pi}{|\mathbf{x}-\mathbf{x}'|^2}\tilde{\varphi}_{\mathrm{H}}(\mathbf{x}'), \quad (15)$$

whose normalized solution

$$\tilde{\varphi}_{\mathrm{H}}(\mathbf{x}) = \frac{8\sqrt{\pi}}{(1+x^2)^2} \quad (16)$$

is well known, we obtain

$$\mathcal{T}_0^{(1)} = -\frac{4\sqrt{2\pi}}{1+|\mathbf{v}_p-\mathbf{q}|^2}F(\mathbf{q};\,0;\,\mathbf{k}). \quad (17)$$

This is nothing but the first term in the sum (13), which corresponds to the shake-off process described above. Note that the electron captured by a proton has momentum $\mathbf{q}$, and the process described by formula (17) is a purely quantum one, which has nothing to do with the classical resonant capture. The second term in (13) represents the exchange term, and the third term corresponds to the Coulomb interaction of a proton with the nucleus in the first Born approximation.

Here, we restrict our attention to the amplitude (17) since the quantity

$$F(\mathbf{q};\,0;\,\mathbf{k}) = \int \exp(-i\mathbf{q}\cdot\mathbf{r}_1) \times \varphi^{-*}(\mathbf{k},\mathbf{r}_2)\Phi_0(\mathbf{r}_1,\mathbf{r}_2)d\mathbf{r}_1 d\mathbf{r}_2 \quad (18)$$

represents the Fourier transform of the coordinate wave function of the helium atom; according to the hypothesis of the authors of the experiments, it is this term that should dominate in the general amplitude (12) and pro-

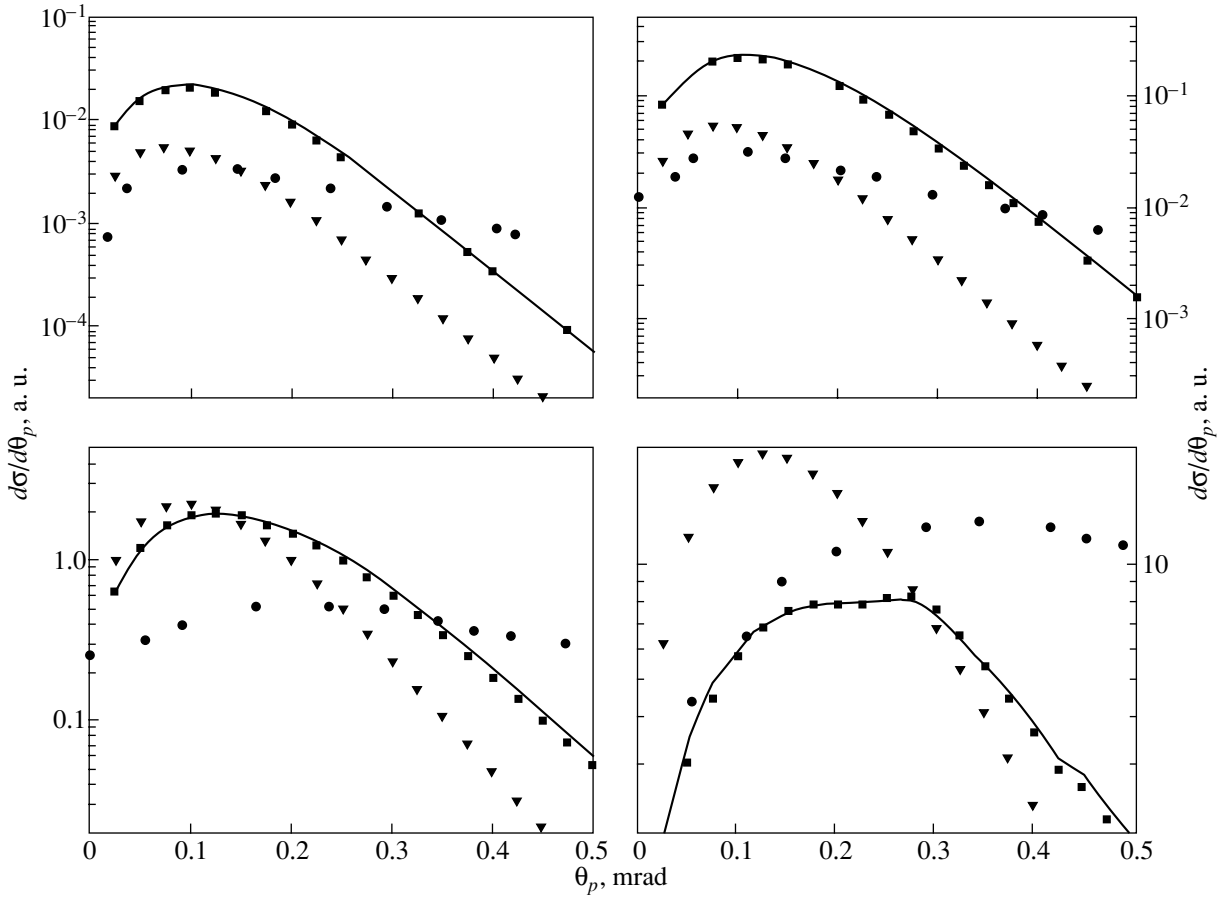


Fig. 2. Same as in Fig. 1 for the reaction p + He ⟶ H + e + He^{++}. Relative error of the experiment (circles) in case (a) is no greater than 20% in the range of angles considered.

vide information on the correlation structure of the wave function $\Phi_0(\mathbf{r}_1, \mathbf{r}_2)$ in the range of very small angles θ_p. Note that a function of type (18) appears in the amplitude of the process (e, 3e) [8, 9]; it was demonstrated that this process provides a powerful tool for the angular spectroscopy of e–e correlations in the target under investigation.

In the case of the SC reaction, the amplitude is given by (17); however, one should replace the amplitude $\varphi^{-*}(\mathbf{k}, \mathbf{r})$ in integral (18) by $\varphi_0(r) = \sqrt{(8/\pi)}\exp(-2r)$.

The differential cross section of the TI process is represented as

$$d^5\sigma = \frac{|\mathcal{T}|^2}{v_p^2}\frac{d\mathbf{q}_\perp}{(2\pi)^2}\frac{d\mathbf{k}}{(2\pi)^3} = \frac{m^2}{(2\pi)^5}|\mathcal{T}|^2 d\Omega_H d\mathbf{k}. \quad (19)$$

The singly differential cross section, which we calculated, follows from (19):

$$\frac{d\sigma}{d\theta_p} = \frac{m^2\theta_p}{(2\pi)^4}\int|\mathcal{T}|^2 d\mathbf{k}. \quad (20)$$

For the SC reaction, formula (20) reduces to

$$\frac{d\sigma}{d\theta_p} = \frac{m^2\theta_p}{2\pi}|\mathcal{T}|^2. \quad (21)$$

3. RESULTS OF CALCULATIONS AND DISCUSSION

Figures 1 and 2 represent the calculated cross sections (20) and (21) with the amplitude $\mathcal{T}_0^{(1)}$. In these calculations, we used the following three functions $\Phi_0(\mathbf{r}_1, \mathbf{r}_2)$:

(1) the simplest of the Hylleraas functions [10]

$$\Phi_0(\mathbf{r}_1, \mathbf{r}_2) = \frac{Z^3}{\pi}\exp[-Z(r_1 + r_2)], \quad Z = \frac{27}{16}$$

(indicated by the symbol Hy);

(2) one of the best functions of Bonham and Kohl [11] (number 17, denoted by BK);

(3) yhe factorized 12-component correlated variational function [12], denoted by CVP, which was designed by the authors specially for the present study. The experimental results are borrowed from [1].

We should immediately point out several features of the processes considered. First, the value $v_p \approx 7$ in the upper limit of the range of proton energies allows us to treat the proton velocity as a large parameter in the problem. The transferred momentum $q \geq v_p/2$ also is a large parameter; this fact distinguishes the capture reaction from the reactions (e, 2e) and (e, 3e), where this parameter is small ($q \sim 0$–2). On the other hand, if, by analogy with the processes (e, 2e) and (e, 3e), we calculate the velocity v_p for the energy of the initial electron, then this energy will correspond to 700 eV in the upper limit, which is obviously insufficient for the impulse approximation.

Second, one can easily show that $\mathcal{T}_0^{(1)} \propto v_p^{-6}$ in the case of a simple separable function Hy. Other, more correlated, functions also have the same order of smallness. However, if we analyze (at least qualitatively) the second Born approximation [13], or more accurately calculate the two-center function Ψ_{out} in (12), then we can see more terms of the same order. Thus, one should not expect that $\mathcal{T}_0^{(1)}$ is a suitable approximation even for asymptotically large v_p.

These features are manifested in the figures. In the case of the SC reactions, all the calculated results virtually coincide but appreciably differ from the experimental results: (1) the absolute values in the maximum for E_p = 1.4 MeV differ approximately by a factor of six; (2) the maximum itself is shifted toward smaller scattering angles; and (3) the shapes of the curves do not coincide (the experimental curve decays much more slowly after the maximum).

The same applies to the TI reactions, although the results of calculations for correlated and uncorrelated functions are different, as was to be expected.

The comparison of the theoretical and experimental results shows that, even for very small angles when the amplitude $\mathcal{T}_0^{(1)}$ seems to dominate and carry certain exclusive information on electron–electron correlations in the target, corrections due to other mechanisms in the first and second Born approximations are sufficiently large so that the reaction $p + \text{A} \longrightarrow \text{H} + e + \text{A}^{++}$ cannot be considered useful for the method of angular spectroscopy of correlations.

ACKNOWLEDGMENTS

We are grateful to H. Schmidt-Böcking for useful discussions and kindly providing us with the experimental data.

This work was supported in part by the Federal Scientific Research Program "Fundamental Spectroscopy" of the Ministry of Industry and Science of the Russian Federation (contract no. 108-39(00)-П).

REFERENCES

1. V. Mergel, PhD Thesis (Univ. Frankfurt/Main, Shaker Verlag, 1996).
2. V. Mergel, R. Dörner, M. Achler, *et al.*, Phys. Rev. Lett. **79**, 387 (1997).
3. V. Mergel, R. Dörner, Kh. Khayyat, *et al.*, Phys. Rev. Lett. **86**, 2257 (2001).
4. Dr. Belkiĉ, R. Gayet, and A. Salin, Phys. Rep. **56**, 281 (1979).
5. V. I. Lend'el, V. Yu. Lazur, M. I. Karbovanets, and R. K. Yanev, *Introduction in Theory of Atomic Collisions* (Vyshcha Shkola, L'vov, 1989), Chap. 5.
6. L. H. Thomas, Proc. R. Soc. London, Ser. B **114**, 561 (1927); J. S. Briggs and K. Taulbjerg, J. Phys. B **12**, 2565 (1979).
7. S. G. Tolmanov and J. H. McGuire, Phys. Rev. A **62**, 32711 (2000).
8. V. G. Neudachin, Yu. V. Popov, and Yu. F. Smirnov, Usp. Fiz. Nauk **169**, 1111 (1999).
9. Yu. V. Popov, C. Dal Cappello, and K. Kouzakov, J. Phys. B **29**, 5901 (1996).
10. H. A. Bethe and E. E. Salpeter, *Quantum Mechanics of One- and Two-Electron Atoms* (Academic, New York, 1957; Fizmatgiz, Moscow, 1960).
11. R. A. Bonham and D. A. Kohl, J. Chem. Phys. **45**, 2471 (1966).
12. O. Chuluunbaatar, I. V. Puzynin, and S. I. Vinitsky, J. Phys. B **34**, L425 (2001).
13. A. M. Brodskiĭ, V. S. Potapov, and V. V. Tolmachev, Zh. Éksp. Teor. Fiz. **58**, 264 (1970) [Sov. Phys. JETP **31**, 144 (1970)].

Translated by I. Nikitin

Journal of Experimental and Theoretical Physics, Vol. 95, No. 4, 2002, pp. 625–638.
From Zhurnal Éksperimental'noĭ i Teoreticheskoĭ Fiziki, Vol. 122, No. 4, 2002, pp. 723–737.

ATOMS, SPECTRA, RADIATION

Amplification of Short Laser Pulses by Raman Backscattering in Capillary Plasmas¶

I. Y. Dodin[a, *], G. M. Fraiman[a], V. M. Malkin[b], and N. J. Fisch[b]

[a]*Institute of Applied Physics, Russian Academy of Sciences, Nizhni Novgorod, 603950 Russia*
[b]*Princeton Plasma Physics Laboratory, Princeton, NJ 08543, USA*
**e-mail: idodin@pppl.gov*

Received April 1, 2002

Abstract—Short laser pulses can be significantly amplified in the process of Raman backscattering in plasma inside an oversized dielectric capillary. A dielectric capillary allows obtaining high intensities of the output radiation by sustaining efficient amplification at large distances compared to the diffraction length. The efficiency of the interaction between the pump wave and the amplified pulse is shown not to be critically sensitive to the transverse structure of the wave fields. For a quasi-single-mode initial seed pulse and a low pump intensity, the amplified pulse tends to preserve its transverse structure due to nonlinear competition of the capillary eigenmodes. At a high power of the pump wave, multimode amplification always takes place but the growth of the front peak of the pulse still follows the one-dimensional model. The Raman backscattering instability of the pump wave resulting in the noise amplification can be suppressed in detuned interaction by chirping the pump wave or arranging an inhomogeneous plasma density profile along the trace of amplification. The efficiency of the desired pulse amplification does not significantly depend on detuning in the case of a smooth detuning profile. Density inhomogeneities are shown to exert less influence on the amplification within a capillary than in the one-dimensional problem. Parameters of a future experiment on the Raman amplification of a short laser pulse inside a capillary are proposed. © *2002 MAIK "Nauka/Interperiodica".*

1. INTRODUCTION

Laser intensities inside conventional amplifiers are limited to gigawatts (GW = 10^9 W) per cm^2, above which a nonlinear modification of the material refraction index causes unacceptable distortions of the laser pulses [1]. The chirp pulse amplification technique allows increasing the output intensities by means of the longitudinal compression of laser pulses after their amplification [2]. The compression is usually performed by means of metallic diffraction gratings, which can survive intensities not larger than tens of TW/cm^2 (TW = 10^{12} W) [1]. One of the most promising ways for further increasing the output intensities consists in using the advantages of plasma technology [3]. Replacing all the major elements of the amplification–compression scheme by one element containing fully ionized plasma capable of acting as the stretcher, the nonlinear amplification medium, and the compressor simultaneously is cheaper and more adequate compared to the extensive development of traditional solid-state devices.

Currently, significant attention is attracted to the problem of generating ultraintense laser pulses in plasmas by means of the Raman backscattering process [3]. In this process, the seed pulse amplification follows the resonant excitation of a plasma wave provided by the beating of the seed pulse and the counterpropagating pump wave. The pump wave energy is primarily absorbed by the front part of the amplified pulse, which results in compression of the latter. By means of the resonant mechanism discussed in this paper, the amplified pulse duration can be decreased to the period of Langmuir oscillations. In what follows, we term such pulses as short, which corresponds to a femtosecond laser pulse duration for realistic experimental conditions. (As shown in [4], amplification of even shorter pulses is possible via Compton backscattering, which remains out of the scope of our study, although it represents a process complementary to the Raman interaction of laser waves.)

Compared to its solid-state analogues or plasma amplifiers utilizing the interaction of copropagating pulses, the scheme allows faster amplification, higher maximum output wave intensities, higher thresholds for developing plasma instabilities, and better limits for the nonlinear pulse compression. Because of a relative simplicity of the experimental implementation, the Raman backscattering pulse amplification in plasmas can successfully compete with more complicated techniques of generating femtosecond laser pulses [2].

Conventionally, the problem of short laser pulse amplification in the Raman backscattering process in plasmas is considered within the framework of a one-dimensional (1D) model, and the transverse structure of the pulse is neglected [1, 3, 5]. But the transverse effects

¶This article was submitted by the authors in English.

can become important in the experimental implementation of the amplification scheme and further practical applications. The study of the transverse effects was recently started for the pulse interaction in vacuum [6], where the amplification efficiency is significantly limited by the transverse diffraction of the amplified pulse. An efficient interaction in a boundless medium is only possible at distances small compared to the diffraction (Rayleigh) length $z_R \sim kR^2$, where $k = 2\pi/\lambda$ is a characteristic wave number of the seed pulse and R is its characteristic transverse scale. After the amplified pulse passes the distance $z \gg z_R$, diffraction increases the transverse scale of the pulse and, therefore, lowers its intensity, which results in a decrease in the interaction efficiency.

In order to maintain high interaction efficiency at large spatial scales compared to z_R, additional laser pulse focusing must be applied. Because of the high intensities of the amplified radiation, conventional dielectric lenses cannot adequately focus the amplified pulse. The problems of the refraction index distortion or even the dielectric medium breakdown, which might occur, can be eliminated using the channeling properties of a dielectric capillary that plays the role of an optical waveguide for both the pump wave and the amplified pulse. (A similar technique is often used in other Raman media for pulse amplification with significantly lower wave intensities [7, 8].) In oversized ($R \gg \lambda$) dielectric capillaries, the field amplitude decreases to the edges of the transverse waveguide cross section and almost equals zero on the inner wall of the tube [9]. Therefore, it is possible to have a field amplitude higher than critical (with respect to the breakdown of the dielectric material of the waveguide walls) in the center of the capillary without damaging its walls. These and other properties of channeling laser pulses in the process of the Raman backscattering amplification within a dielectric capillary are the main subject of this paper.

The paper is organized as follows. In Section 2, we give the basic equations describing Raman backscattering in plasmas. In Section 3, we revise some aspects of the 1D Raman amplification problem. We consider the capillary problem in Section 4, where we develop a mode approach allowing quantitative and simple qualitative understanding of some phenomena occuring during the laser pulse interaction inside a capillary. We also generalize the conventional 1D linear theory of pulse amplification by considering the interaction between the capillary modes of the amplified pulse and discuss some aspects of selective mode discrimination in capillaries. Single- and multimode amplification regimes are discussed in Section 5 in detail. In Section 6, we discuss the problem of detuned amplification. Some numerical estimates and the summary of the main ideas are given in Section 7. Specific features of the cylindric dielectric capillary are discussed in the Appendix.

2. BASIC EQUATIONS

Equations for vector electric fields describing paraxial propagation of laser pulses along the z axis can be written as (see, e.g., [10, 11])

$$\partial_t \mathbf{a} + c\partial_z \mathbf{a} - \frac{ic^2}{2\omega_a}\nabla_\perp^2 \mathbf{a} = \omega_p \mathbf{b} f, \tag{1}$$

$$\partial_t \mathbf{b} - c\partial_z \mathbf{b} - \frac{ic^2}{2\omega_b}\nabla_\perp^2 \mathbf{b} = -\omega_p \mathbf{a} f^*, \tag{2}$$

$$\partial_t f + i\delta\omega f = -\frac{\omega}{2}\mathbf{b}^\dagger \cdot \mathbf{a}, \tag{3}$$

where the vectors $\mathbf{a}$ and $\mathbf{b}$ represent the slowly changing amplitudes of the respective electric fields

$$\begin{aligned} \mathbf{E}_a &= \frac{m_e c\omega_a}{e}\{i\mathbf{a}\exp(ik_a z - i\omega_a t) + \text{c.c.}\}, \\ \mathbf{E}_b &= \frac{m_e c\omega_b}{e}\{i\mathbf{b}\exp(ik_b z - i\omega_b t) + \text{c.c.}\} \end{aligned} \tag{4}$$

of the pump and the seed pulse, and f is the normalized potential of the plasma wave electric field

$$\mathbf{E}_f = \mathbf{k}_f^{(0)}\frac{m_e c\omega_p}{e}\{f\exp(ik_f z - i\omega_f t) + \text{c.c.}\}, \tag{5}$$

where

$$\mathbf{k}_f^{(0)} = \mathbf{k}_f/k_f, \quad \mathbf{k}_f = \mathbf{z}^{(0)}(k_a - k_b),$$
$$\omega_f = \omega_a - \omega_b = \omega_p - \delta\omega.$$

Here,

$$\omega_p = \sqrt{\frac{4\pi n_e e^2}{m_e}}$$

is the plasma frequency, n_e is the electron density, and e and m_e are the electron charge and mass respectively. We assume the rare plasma conditions $\omega_p \ll \omega_a \approx \omega_b \equiv \omega$ and $k_f\lambda_D \ll 1$, and therefore, $k_{a,b} \approx \omega_{a,b}/c$ and the dispersion of plasma waves can be neglected ($\partial_k\omega_f \approx 0$). In terms of the dimensionless amplitude a, the pump intensity is

$$I_a = \pi c(m_e c^2/e)^2|a|^2/\lambda^2$$
$$= 2.736\times 10^{18}|a|^2/\lambda^2\ [\mu\text{m}]\ \text{W/cm}^2,$$

see [6].

It is useful to introduce the dimensionless equations

$$\partial_\tau \mathbf{a} + \partial_z \mathbf{a} - i(1+\sigma)\nabla_\perp^2 \mathbf{a} = \mathbf{b} f, \tag{6}$$

$$\partial_\tau \mathbf{b} - \partial_z \mathbf{b} - i\nabla_\perp^2 \mathbf{b} = -\mathbf{a} f^*, \tag{7}$$

$$\partial_\tau f + i\delta\omega f = -\mathbf{b}^\dagger \cdot \mathbf{a}, \tag{8}$$

where

$$\sigma = \frac{\omega_a - \omega_b}{\omega_b} \sim \frac{\omega_p}{\omega} \ll 1,$$

the time τ is measured in the units $t_0 = \sqrt{2/\omega\omega_p}$, the longitudinal coordinate z is measured in the units ct_0, f is measured in $\sqrt{\omega/2\omega_p}$, the transverse coordinate ρ is measured in the units $c(2\omega_p\omega^3)^{-1/4}$, and the detuning $\delta\omega$ is measured in the units t_0^{-1}.

For further analysis, it is convenient to introduce the coordinate $\zeta = \tau + z$ (in what follows, this change of variables is called the shift to the reference frame moving together with the amplified pulse at the speed of light). To describe the strongly nonlinear regime of the amplification of a compressed pulse, it suffices to keep only the ζ derivatives of $\mathbf{a}$ and f (the so-called quasistatic approximation [1, 3]); the basic equations then become

$$2\partial_\zeta \mathbf{a} - i(1+\sigma)\nabla_\perp^2 \mathbf{a} = \mathbf{b}f, \tag{9}$$

$$\partial_\tau \mathbf{b} - i\nabla_\perp^2 \mathbf{b} = -\mathbf{a}f^*, \tag{10}$$

$$\partial_\zeta f + i\delta\omega f = -\mathbf{b}^\dagger \cdot \mathbf{a}. \tag{11}$$

In the case of zero detuning, the basic equations are invariant under the transformation

$$\mathbf{a} \longrightarrow C\mathbf{a}, \quad \mathbf{b} \longrightarrow C\mathbf{b},$$

$$f \longrightarrow Cf, \quad \tau \longrightarrow \tau/C, \quad \zeta \longrightarrow \zeta/C, \quad \rho \longrightarrow \rho/\sqrt{C}.$$

Therefore, the specific value of the pump amplitude $a_0 = a(z \longrightarrow -\infty)$ is in fact not important in the sense that the field dynamics for another value of a_0 can be obtained by a simple rescaling.

3. THE ONE-DIMENSIONAL PROBLEM

For better understanding of the qualitative phenomena to be discussed in relation to the Raman backscattering inside a capillary, it is useful to revise the basic aspects of the conventional 1D problem first (see [1, 3] for a detailed discussion). During the linear stage of amplification, when the pump depletion is negligible, $a \approx a_0 = \text{const}$, the solution to Eqs. (6)–(8) can be obtained by the Laplace transformation and is given by [1]

$$\begin{aligned} b(\zeta, z) &= \frac{\partial}{\partial\zeta}\int G(\zeta - \zeta', z)b(\zeta', 0)d\zeta', \\ G(\zeta, z) &= I_0(2\sqrt{\eta}), \\ \eta &= -a_0^2\zeta z, \end{aligned} \tag{12}$$

where we assume zero detuning (a constant detuning can be removed from the evolution equations; the case of the linear detuning $\partial_z\delta\omega = \text{const}$ is considered in detail in [1, 5]). We note that the spatial coordinate $-z$ plays the role of time in Eqs. (12), measuring the interval between the initial and the current positions of the amplified pulse propagating along the z axis with a fixed velocity equal to the speed of light.

For $\eta \gg 1$, we have

$$G \approx \exp(2\sqrt{\eta})/2\sqrt{\pi\sqrt{\eta}}.$$

In the original variables,

$$\eta = a_0^2\omega\omega_p(t + z/c)(-z)/2c,$$

and the maximum of G is therefore reached at $z = -ct/2$; it increases with the peak growth rate $\gamma = a_0\sqrt{\omega\omega_p/2}$ as $\exp(\gamma t)$.

The linear approximation Eqs. (12) is valid until

$$\epsilon(\tau) = \int b(z, \tau)dz$$

remains small compared to unity; for larger ϵ, a nonlinear solution is formed. Because of the pump depletion, only the front part of the seed pulse is then amplified, which leads to the effective compression of the pulse. Eventually, as the pulse becomes sufficiently short, the quasistatic approximation (Eqs. (9)–(11)) becomes valid, and for the real constant pump, the solution is therefore given by

$$\begin{aligned} a &= a_0\cos(U/2), \\ f &= -\sqrt{2}a_0\sin(U/2), \\ b &= \partial_\zeta U/\sqrt{2}, \end{aligned} \tag{13}$$

where U satisfies the sine-Gordon equation

$$\partial_{\tau\zeta}^2 U = a_0^2\sin U. \tag{14}$$

Equation (14) has a family of self-similar solutions (Fig. 1) $U(\tau, \zeta) = U(\eta)$ that satisfy the equations

$$\eta U_{\eta\eta} + U_\eta = \sin U \tag{15}$$

or

$$U_{\xi\xi} + U_\xi/\xi = \sin U, \quad \xi = 2\sqrt{\eta}, \tag{16}$$

where we equate η with $a_0^2\tau\zeta$ because of the quasistatic approximation. It is convenient to consider the solution to Eq. (16) in the plane (U, U_ξ), which can approximately be treated as the phase plane of a nonlinear oscillator with the effective dissipation determined by the term U_ξ/ξ (Fig. 1). The absolute maximum of the self-similar solution grows in time as

$$b_{\max} \approx a_0^2\tau(1 + \ln(4\sqrt{2\pi}/\epsilon_0))^{-1}, \quad \epsilon_0 \equiv \epsilon(0) \ll 1,$$

and the locations of the pulse maxima change as

$$\zeta_{\max} \sim 1/b_{\max}.$$

The self-similar solution $U(\eta)$ in Eqs. (15), (16) corresponds to the initial conditions

$$b(z, \tau = 0) \equiv b_0(z) = \epsilon_0\delta(z), \tag{17}$$

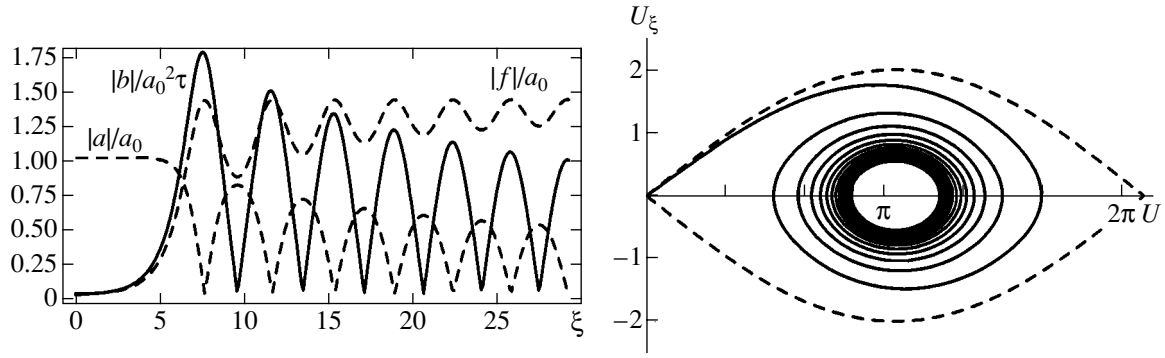


Fig. 1. Self-similar profiles of $|a(\xi)|/a_0$ (dotted decay), $|f(\xi)|/a_0$ (dotted growth), $|b(\xi)|/\tau a_0^2$ (solid line) for $\epsilon_0 = 0.01$ and the behavior of the self-similar solution on the (U, U_ξ) plane (dashed line represents the solution without the "friction" term U_ξ/ξ; $\xi = 2(a_0^2\omega\omega_p(t + z/c)(-z)/2c)^{1/2}$).

which imply that

$$U(\xi = 0+) = \epsilon_0, \quad U'(\xi = 0+) = 0,$$

and which are therefore applicable for all ζ to the left of the initial location of the seed pulse in the frame moving together with the amplified pulse. Consider now what happens when the spatial scale of the amplified pulse $\Delta(\tau)$ becomes comparable to its initial spatial scale $\Delta_0 \equiv \Delta(0)$, so that the delta approximation for initial conditions (17) therefore becomes invalid. In this case,

$$\epsilon_0 = \int_{-\infty}^{+\infty} b_0(z')dz'$$

does not determine the solution, and new initial conditions for a self-similar profile must then be applied. The front pulse faces the unperturbed profile of the pump wave. Qualitatively, pump depletion becomes significant ($\delta a/a_0 \sim 1$) starting only with $\zeta = \zeta_*$, where ζ_* is determined by the condition $\epsilon(\zeta_*(\tau) \sim 1$, with

$$\epsilon(\zeta) = \int_{-\infty}^{\zeta} b(\zeta', \tau)d\zeta'. \tag{18}$$

(To make a rough estimate, one can equate ζ_* to the location of the first maximum of $b(\zeta, \tau)$ at a current instant τ.) Therefore, linear solution (12) remains valid for $\zeta < \zeta_*$. In the case where the spatial scale of the Green's function G is large compared to the spatial scale of the initial pulse, Eq. (12) can be written as

$$U = \epsilon(\zeta)G(\eta). \tag{19}$$

On the other hand, because the self-similar solution represents an attractor, its formation still occurs starting from the end of the linear stage, where its profile can be obtained from linearized Eq. (15) and is given by

$$U = \epsilon_{\text{eff}}G(\eta), \tag{20}$$

where ϵ_{eff} is some constant. At the location where the linear stage ends and the self-similar solution starts, i.e., at $\zeta = \zeta_*(\tau)$, the two solutions in Eqs. (19) and (20) must match, which defines $\epsilon_{\text{eff}}(\tau)$,

$$\epsilon_{\text{eff}}(\tau) = \epsilon(\zeta_*(\tau)). \tag{21}$$

As long as $\Delta(\tau) \gg \Delta_0$, we have

$$\epsilon_{\text{eff}} \approx \epsilon_0 = \text{const}.$$

But when the nonlinear compression makes $\Delta(\tau)$ comparable to or less than Δ_0, ϵ_{eff} can become significantly smaller than ϵ_0. If $\epsilon_{\text{eff}}(\tau)$ is changing sufficiently slowly, so that the self-similar profile has enough time to set up on the entire length of the pulse, the entire solution remains close to the self-similar one with the only change that it is now parameterized by time-dependent

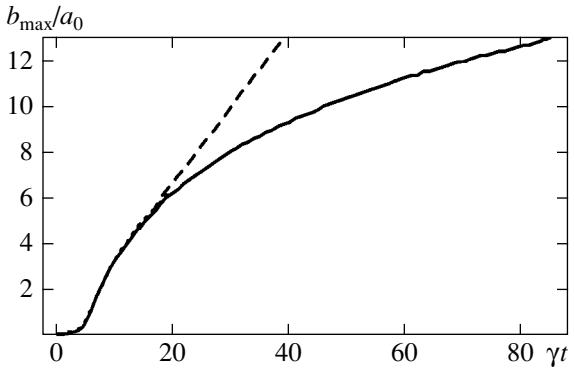


Fig. 2. The amplitude of the amplified 1D pulse maximum normalized to the amplitude of the pump wave ($b_{\max}/|a_0|$) as a function of $\gamma t \equiv |a_0|\tau$ for $b_0(z) = \epsilon_0\exp(-z^2/z_0^2)/\sqrt{\pi}z_0$: a delta-shaped initial pulse ($z_0 \longrightarrow 0$, self-similar profile with fixed $\epsilon_0 = 1.3 \times 10^{-2}$, dashed line) and a finite-width initial pulse ($z_0 = \sqrt{5}$, quasi-self-similar profile with $\epsilon_{\text{eff}}(\tau)$, solid line).

quantity (21) (Fig. 2). But if $\epsilon_{\text{eff}}(\tau)$ changes fast, the self-similar solution may not be able to form and, therefore, does not represent an attractor. Stochastic behavior of the amplified pulse structure is observed in this case.

In the approximation of geometric optics, when the pulse propagation is considered at small distances compared to the Rayleigh length $z_R \sim kR^2$, the diffraction-caused distortion of the transverse structure of laser pulses can be neglected. In this case, 1D quasi-self-similar solutions are formed on geometric rays constituting the field of the amplified pulse. The spatial profiles of the amplified pulse that are then formed have a shape similar to nested horseshoes. But for pulse traces $z > z_R$, the diffraction terms in Eqs. (6)–(11) become significant and must therefore be taken into account (see Section 5).

4. THE MODE APPROACH TO THE NON-ONE-DIMENSIONAL PROBLEM

We consider the pump wave **a** and the amplified pulse **b** given by a series in the normalized eigenmodes $\boldsymbol{\psi}_s$, $\langle \boldsymbol{\psi}_m | \boldsymbol{\psi}_n \rangle = \delta_{mn}$

$$\begin{aligned} \mathbf{a} &= R\sum_n a_n(z,\tau)\boldsymbol{\psi}_n(\mathbf{r}_\perp), \\ \mathbf{b} &= R\sum_m b_m(z,\tau)\boldsymbol{\psi}_n(\mathbf{r}_\perp), \end{aligned} \tag{22}$$

where R is the radius of the capillary, which we include as a normalization factor to make the amplitudes a_n and b_m dimensionless (see the Appendix for the explicit form of $\boldsymbol{\psi}_n$ for a dielectric capillary). By definition, the eigenfunctions $\boldsymbol{\psi}_s$ satisfy the equation

$$\nabla_\perp^2 \boldsymbol{\psi}_s + \chi_s^2 \boldsymbol{\psi}_n = 0, \tag{23}$$

where χ_s is the transverse wavenumber of the sth eigenmode. From Eqs. (6)–(8), we obtain the equations for the amplitudes a_n and b_m,

$$(\partial_\tau + \partial_z + i\delta\Omega_n^{(a)})a_n = \sum_m f_{nm} b_m, \tag{24}$$

$$(\partial_\tau - \partial_z + i\delta\Omega_n^{(b)})b_m = -\sum_n a_n f_{nm}^*, \tag{25}$$

where

$$\delta\Omega_n^{(a)} = \chi_n^2(1+\sigma), \quad \delta\Omega_m^{(b)} = \chi_m^2$$

(or $\delta\Omega_n^{(a,b)} = (c\chi_n)^2/2\omega_{a,b}$ in dimensional variables) and

$$f_{nm} = \langle \boldsymbol{\psi}_n | f | \boldsymbol{\psi}_m \rangle$$

are dimensionless transverse moments of the plasma wave profile satisfying

$$(\partial_\tau + i\delta\omega) f_{nm} = -\sum_{k,l} C_{nklm} a_l b_k^*, \tag{26}$$

with constant dimensionless coefficients given by

$$\begin{aligned} C_{nklm} &= R^2 \langle \boldsymbol{\psi}_n | \boldsymbol{\psi}_k^\dagger \cdot \boldsymbol{\psi}_l | \boldsymbol{\psi}_m \rangle \\ &= R^2 \int d^2 r_\perp (\boldsymbol{\psi}_n^\dagger \cdot \boldsymbol{\psi}_m)(\boldsymbol{\psi}_k^\dagger \cdot \boldsymbol{\psi}_l). \end{aligned} \tag{27}$$

The eigenmode approach can be useful only in the case where the modes are coupled weakly, which corresponds to the case of a strong waveguide dispersion. Otherwise, the number of modes to be taken into consideration becomes infinite. If the characteristic trace z_0 of the pulse evolution is large compared to the Rayleigh length z_R, Eq. (26) can be reduced to

$$\begin{aligned} (\partial_\tau + i\partial\omega) f_{nm} &= -\frac{C_{nm}}{1+\delta_{nm}}(a_n b_m^* + a_m b_n^*), \\ C_{nm} &\equiv C_{nmnm} \geq 0, \end{aligned} \tag{28}$$

where we assume the eigenfunctions to contain complexity in polarization factors at most, but not in the functional dependence of the transverse coordinates (for simplicity, below we ignore the fact that C_{nm} can be equal to zero for modes of the opposite polarization). Equations (24), (25), and (28) represent a completely defined Lagrangian set of equations that can be used for obtaining the amplitudes of resonantly interacting modes of the three waves **a**, **b**, and f in the case of weak coupling (see below).

The important conclusion following from Eqs. (24), (25), and (28) is that for every pair of modes of the pump and the seed, a_n and b_m, the resonant plasma wave harmonic f_{nm} can be generated to provide coupling of the two electromagnetic waves. This is a specific feature of light scattering on a cold plasma wave for which the spatial resonance condition

$$\mathbf{k}_a = \mathbf{k}_b + \mathbf{k}_f$$

is satisfied automatically because the wave vector $\mathbf{k}_f$ remains arbitrary for the given frequency $\omega_f \approx \omega_p$. For scattering on any other low-frequency wave f for which the wavevector depends on its frequency ω_f, the multiple-mode interaction on the quadratic nonlinearity is impossible.

The presence of the $a_m b_n^*$ term in Eq. (28) is responsible for a possible parasitic resonance, which can be explained as follows. We consider the interaction between the nth mode of the pump a_n and the mth mode of the seed pulse b_m generating the resonant plasma wave f_{nm} with the longitudinal wave number

$$h_{nm} = k_a - k_b - \delta\Omega_n^{(a)} - \delta\Omega_m^{(b)}.$$

For very small $\sigma \sim \omega_p/\omega_b$ (namely, for $\sigma \lesssim z_R/z_0$, where z_0 is the characteristic spatial scale of the pulse evolution), we have $h_{nm} \approx h_{mn}$, where h_{mn} is the wave number of the plasma wave f_{mn} resonant to the beating wave of the modes a_m and b_n, which provides an additional coupling of these two pairs of electromagnetic waves. For example, in the case where the pump contains the modes a_1 and a_2 and the seed pulse contains only b_1, the second seed harmonic $b_2 = O(a_2^* a_1 b_1)$ is generated. This effect can already become important at the linear stage of the interaction in a multimode pump, because it alters the increments of the linear Raman amplification.

We now use the developed mode approach to consider the linear stage of the pulse amplification inside a capillary in terms of the equation

$$\partial_\tau(\partial_\tau - \partial_z - i\nabla_\perp^2)\mathbf{b} = \mathbf{a}(\mathbf{a}^\dagger \cdot \mathbf{b}), \tag{29}$$

which directly follows from Eqs. (7) and (8) with zero detuning $\delta\omega$ and with a constant pump $\mathbf{a}$. The right-hand side of Eq. (29) can be considered as the result of applying the linear operator $\hat{\mathbf{A}} = \mathbf{a}\mathbf{a}^\dagger$ to the vector $\mathbf{b}$, and therefore, Eq. (29) can be rewritten as

$$[\partial_\tau(\partial_\tau - \partial_z + i\delta\Omega_m^{(b)}) - \gamma_m^2]b_m = \sum_{n \neq m} A_{mn} b_n,$$
$$A_{mn} = \left\langle \boldsymbol{\psi}_m \middle| \hat{\mathbf{A}} \middle| \boldsymbol{\psi}_n \right\rangle, \tag{30}$$

where $\gamma_m = \sqrt{A_{mm}}$ represents the increment of the linear amplification of the mth partial waveguide mode. In an arbitrary waveguide, for a single-mode pump, $a_n = \delta_{ns} a$, the matrix elements A_{mn} are of the order of a^2 for $n, m \sim 1$ and $A_{mn} = A(|m - n|)$ for $n, m \gg 1$, where the function $A(k) \sim a^2$ for $k \sim 1$ and decays as its argument grows.

The eigenmodes of the empty waveguide are coupled via the pump inhomogeneity. Only for the uniform pump is the matrix A_{mn} diagonal, and the right-hand side of Eq. (30) is therefore zero. For a nonuniform pump, which is only possible inside a capillary, the effect of mode coupling always occurs. In the case of a weak interaction ($\gamma_m \ll \delta\Omega_m^{(b)}$), the eigenwaves of system (30) are close to its partial waves, and the increments of the eigenwaves are approximately given by γ_m, $m = 1, 2, \ldots, \infty$ (here, we neglect the effect of the parasitic resonance discussed above). For the single-mode pump, $a_n = \delta_{ns} a$, all the increments are of the order of a and are independent of m for $m \gg s$. Specifically, for pulse amplification on the lowest mode of the pump in a dielectric capillary, $s = 1$, we have $(\gamma_m - \gamma_1)/\gamma_1 < 0.16$. Hence, the increments of amplication of all the waveguide modes are close to each other at the linear stage of interaction.

Variations of the pump transverse structure do not change the interaction efficiency significantly. For example, without the possible parasitic resonance taken into account, the amplification increment of the mth partial mode (equal to the increment of the mth eigenmode in the case of weak interaction) is given by

$$\gamma_m = \sqrt{\sum_n C_{nm} |a_n|^2}, \tag{31}$$

which implies that each mode of the pump amplifies each mode of the seed, because $C_{nm} > 0$ for all n and m. The higher modes of the pump amplify the seed with approximately the same efficiency as the lower ones, because

$$C_{nm}/C_{mm} \approx \text{const} \sim 1 \quad \text{for} \quad n \gg m.$$

This effect originates in the fact that the wave interaction inside a capillary is not a three-wave but a multiwave process, where the effective energy exchange between every pair of the pump and the seed modes is possible.

Because the increments of the linear amplification are approximately the same for all waveguide modes, the linear stage of pulse amplification cannot provide significant enhancement of the signal-to-noise ratio. This is true, however, only if the energy losses (which have not been taken into account yet) are negligible at the distance of pulse propagation, which might not be the case in real experiments. In an oversized cylindrical dielectric capillary, the energy losses are mostly radiative and can be incorporated into the model by introducing the spatial decrements of individual modes [9],

$$\alpha_{nm} \sim \left(\frac{\mu_{m,n}}{2\pi}\right)^2 \frac{\lambda^2}{R^3} \tag{32}$$

(see Appendix for the notation). The spatial scale of the exponential decay α_s^{-1} decreases with the mode number s roughly as s^{-2}, and for $\gamma_1 \gtrsim \alpha_1$, only the lowest mode can be amplified and the amplification of the higher modes is suppressed. This implies that the radiative energy losses essentially result in a selective mode discrimination, which can provide the single-mode operation regime.

Additional mode discrimination can occur in relatively narrow waveguides, where the group velocity substantially differs from mode to mode. After the amplified pulse passes the distance $z \gtrsim L_{\text{pulse}}(kR)^2$, where L_{pulse} is the length of the pulse, the wave envelope corresponding to the lowest mode leaves the envelopes of the higher modes behind. The front envelope then has a preferential opportunity of absorbing the energy from the pump wave. Because the pump is significantly

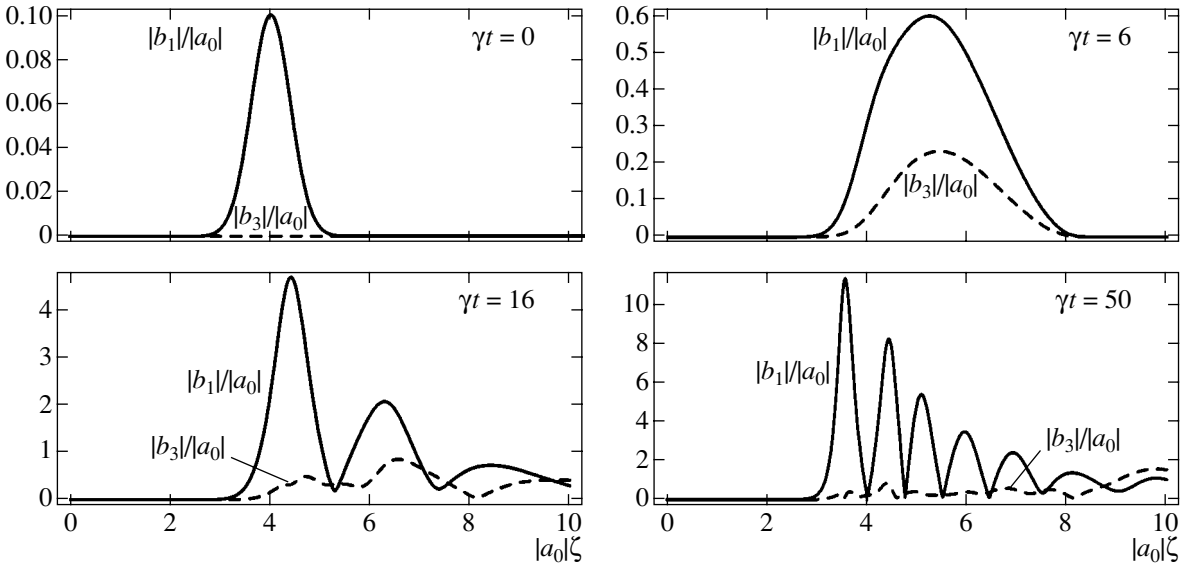


Fig. 3. Evolution of $|b_1(\zeta)|/|a_0|$ (solid lines) and $|b_3(\zeta)|/|a_0|$ (dashed lines); the planar-waveguide scalar problem; $b_m^{(0)} = 0.1a_0\delta_{m1}$, $a_n^{(0)} = a_0\delta_{n1}$, $R = \pi/\sqrt{a_0}$, $b_2 \equiv 0$ because of symmetry; the shots correspond to $\gamma t \equiv |a_0|\tau = 0, 5, 15, 50$. Strong waveguide dispersion provides nonlinear competition of the modes in the nonlinear regime of amplification. Although the small amplitude b_3 appears at the linear stage, it is left behind the wave envelope b_1 later. Amplification of b_3 is then slowed down by the pump depletion provided by b_1.

depleted by the lowest mode, the higher ones are left with less energy to absorb, which also maintains the single-mode amplification regime.

5. SINGLE- AND MULTIMODE AMPLIFICATION

The condition of a weak interaction (or the condition of a strong waveguide dispersion)

$$\gamma \ll \delta\Omega^{(b)},$$

$$a \ll a_{\text{crit}}, \quad a_{\text{crit}} = \sqrt{2/\omega\omega_p}(c\chi)^2/2\omega, \quad \chi = \pi/R,$$

can be treated as follows. The increment of the pulse amplification $\gamma \sim a\sqrt{\omega\omega_p}$ determines the spread of the amplified pulse spectrum $\delta h \sim \gamma/c$. As long as δh remains small compared to the spectral gap between the individual modes,

$$\Delta h \sim \delta\Omega^b/c \sim 1/kR^2,$$

the waveguide eigenmodes do not overlap, and hence, they represent a good basis for developing the mode approach in the linear theory. In this case, the eigenmodes of coupled system (30) remain close to the partial waves of the empty waveguide. This implies that an initially single-mode seed pulse remains single-mode over the entire duration of the linear stage.

The next question is what happens after the linear stage, when nonlinear compression comes into play, providing its own spectrum broadening. We consider the single-mode initial conditions for the seed pulse, e.g., $b_m^{(0)} = b_1^{(0)}\delta_{m1}$. Until the end of the linear stage, the higher mode amplitudes remain small compared to b_1. Then, it is the mode b_1 that passes from the linear to the nonlinear regime first, because its amplitude is the largest. (Here, by the nonlinear regime of an individual mode, we mean the ability of this particular mode to deplete the pump, which might have already been distorted by other modes at the moment.) In the laboratory frame, the maximum of the wave envelope moves approximately with the speed of light in the nonlinear regime, but in the linear one, the effective pulse velocity is substantially lower. For example, as follows from the linear theory of pulse propagation in a constant pump (Section 3), the maximum travels with the speed equal to half the speed of light. The higher mode envelopes (remaining in the linear regime) are therefore left behind the envelope of the first mode. The effective amplitude of the pump $a_{\text{eff}} < a_0$ determining the increments of the higher modes is decreased by the first mode. Because the first mode suppresses the growth of the higher modes, the waveguide dispersion effectively results in a nonlinear competition of the modes tending to sustain the single-mode operation. We call this effect the mode elasticity, because the strongest mode tends to dominate in the nonlinear stage of amplification, thereby preserving the transverse structure of the pulse.

The evolution of the two lowest modes having the highest amplitudes is shown in Fig. 3. (To show the robustness of the mode competition mechanism, numerical calculations demonstrating the single-mode

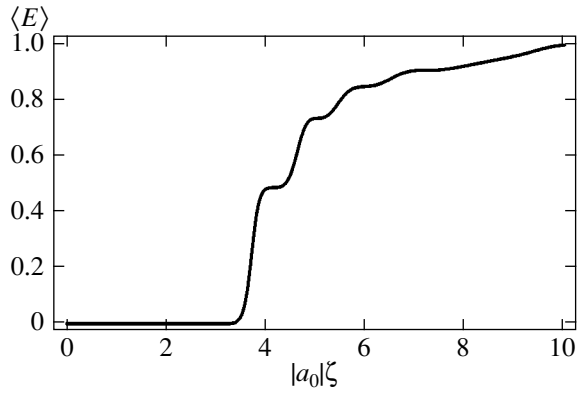


Fig. 4. The normalized energy integral distribution within a quasi-single-mode amplified pulse (averaged over the capillary cross section): up to 50% of the pulse total energy is contained within the first peak; the parameters are the same as in Fig. 3, $\gamma t = 40$.

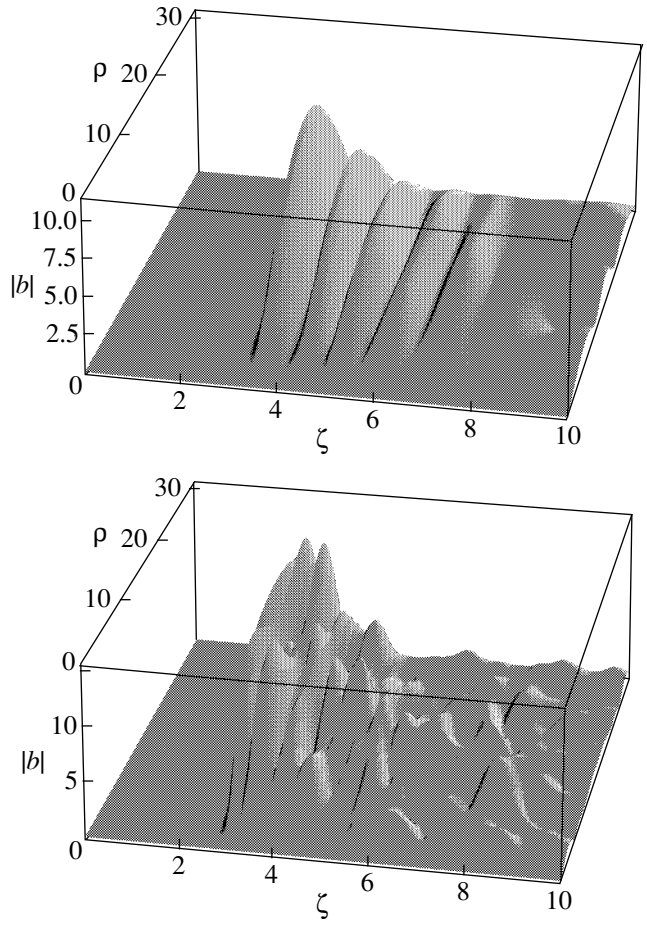


Fig. 5. Characteristic spatial profiles of the amplified pulse $|b(\zeta, \rho)|$ in the case of a strong waveguide dispersion (planar waveguide). At the first stage of the nonlinear amplification, the waveguide dispersion leads to the competition of modes, which supports the single-mode amplification. Later, the higher modes also enter the nonlinear regime, catch up with the wave envelope of the first mode, and ruin the structure of its tail. The front of the pulse always remains single-mode, however, because it always stays in the linear regime, where the growth of the higher modes is suppressed by a strong waveguide dispersion.

amplification were performed for $a \sim a_{crit}$.) In this case, the single-mode amplification also continues in the nonlinear regime, ensuring that the problem remains essentially one-dimensional. We can see the formation of the self-similar profile, which represents the attractor of the single-mode operation, similarly to the 1D problem. The energy distribution inside the amplified pulse (which determines the effective pulse length) averaged over the capillary cross section is given in Fig. 4.

The qualitative arguments given above lead to the conclusion that the formation of the single-mode operation regime in the case of a strong waveguide dispersion is stable with respect to fluctuations of the seed pulse. Neither can the fluctuations of the pump transverse structure influence the single-mode operation because all the modes of the pump wave provide approximately equal efficiencies of the energy transfer into the amplified pulse, as discussed in Section 4.

The nonlinear competition of the modes constituting the amplified pulse remains efficient only until the higher modes enter the nonlinear stage of amplification. After that, their envelopes catch up with the wave envelope of the first mode and ruin the tail of the single-mode structure (Fig. 5). But the front of the amplified pulse always remains in the linear regime (see also Section 3), which provides its single-mode structure.

In the other limiting case, where the interaction between the pump and the amplified pulse is strong ($\gamma \gg \delta\Omega^{(b)}$, or $a \gg a_{crit}$), the pulse is significantly amplified on a small distance compared to z_R, i.e., before the diffraction effects come into play. The waveguide walls cannot then influence the formation of the pulse structure at the first stage of amplification, and a solution close to those formed in a boundless vacuum is produced. Vacuum solutions [6] are shaped as nested horseshoe structures resulting from the transverse inhomogeneity of the pulse and the pump (Fig. 6). On every geometric ray, a self-similar profile is formed with its own $\epsilon_0(\rho)$ (or $\epsilon_{eff}(\rho)$), which determines the longitudinal spatial structure of the pulse at given p. At the edges of the amplified pulse, the amplitudes of both **a** and **b** are smaller than in the center of the system, and the longitudinal spatial scales are correspondingly larger.

In the frame moving together with the front of the amplified pulse (at the speed of light), the longitudinal locations of the pulse maxima $\zeta_{max}(\rho)$ are bounded by the position of the front of the seed pulse ζ_0. On the other hand, the nonlinear compression provided by the preferential amplification of the front of the pulse "pushes" the tail of the pulse from behind to $\zeta = \zeta_0$, which implies that ζ_0 represents the limit of $\zeta_{max}(\rho)$ for all ρ. The front of the horseshoe structure therefore tends to flatten as $\tau \longrightarrow \infty$.

Although stable on small distances compared to z_R and robust with respect to the structure of the seed (see also [6]), the horseshoe solution deteriorates inside the waveguide at $z \gtrsim z_R$, where the diffraction becomes sig-

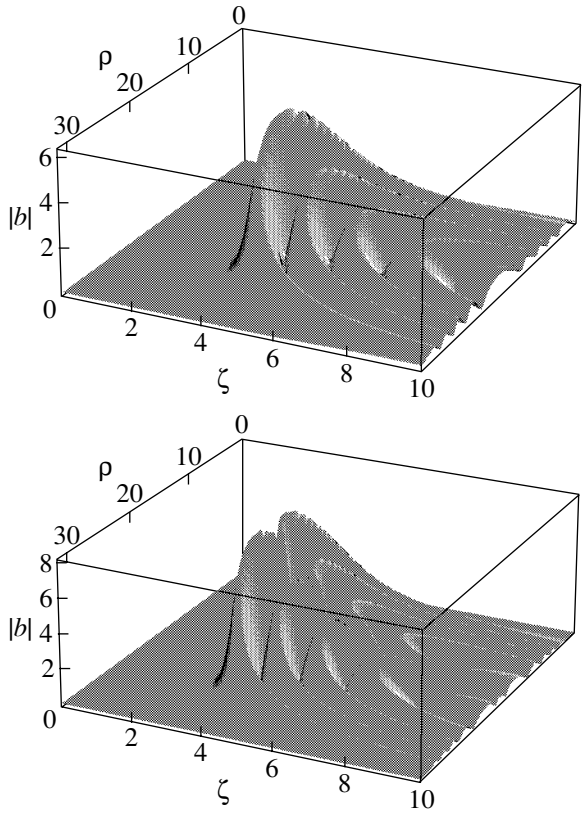


Fig. 6. Quasi-vacuum (horseshoe) nonlinear solutions for $|b(\zeta, \rho)|$ in the case of the strong pump ($a \gg a_{crit}$); the planar-waveguide scalar problem: (upper) $R = 10\pi$, $a_0(\rho) = \sin(\pi\rho/R)$, $b_0(\zeta, \rho) = 0.1\sin(\pi\rho/R)\exp(-(\zeta - 4)^2/0.5)$, $\tau = 20$; (lower) $R = 100\pi$, $a_0(\rho) = 2\sin(\pi\rho/R)$, $b_0(\zeta, \rho) = 0.1\sin(2\pi\rho/R)\exp(-(\zeta - 4)^2/0.5)$, $\tau = 10$.

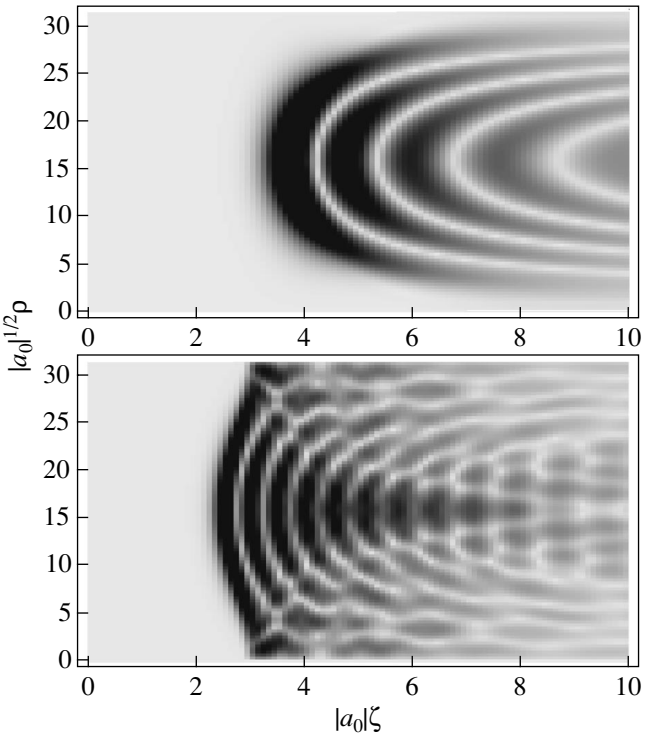


Fig. 7. Deterioration of the horseshoe solution $|b(\zeta, \rho)|$ as $\tau \longrightarrow \infty$ and flattening of the front of the amplified pulse (the planar-waveguide scalar problem, $R = 10\pi/\sqrt{a_0}$, $a_0(\rho) = a_0\sin(\pi\rho/R)$, $b_0(\zeta, \rho) = 0.1a_0\sin(2\pi\rho/R)\exp(-(\zeta - 4)^2/0.5)$: $\gamma t \equiv a_0\tau = 20, 90$, respectively; dark regions correspond to larger $|b|$.

nificant (Fig. 7). The very front of the horseshoe, however, always remains in the linear regime and therefore maintains its regular shape. In the center of the waveguide, the front peak of the amplified pulse grows similarly to the self-similar solution of 1D problem (16) (Fig. 8), which allows using the 1D model for estimating the maximum amplitude of the amplified pulse. The energy distribution inside the amplified pulse (which determines the effective pulse length) averaged over the capillary cross section is given in Fig. 9. At large t, the averaged energy longitudinal distribution becomes a smooth function (cf. Fig. 4), and it is therefore difficult to distinguish the individual peaks of the amplified pulse. On average, the energy becomes distributed over a length that is significantly larger than the length of the first peak.

6. SUPPRESSING NOISE AMPLIFICATION IN DETUNED INTERACTION

Because of the extreme efficiency of the Raman backscattering, which makes fast compression possible, delivering the pump wave energy to the seed pulse through the amplifying plasma layer represents a significant challenge. As the pump traverses the plasma layer towards the seed pulse, the fast Raman backscattering of the pump by thermal Langmuir waves or electromagnetic fluctuations existing inside the plasma layer or coming from outside can lead to a premature pump depletion. The problem is aggravated by the fact that the linear Raman backscattering instability of the pump (responsible for the unwanted noise amplification) has a larger growth rate than its nonlinear counterpart (responsible for the useful amplification of the seed laser pulse).

To see how significantly the thermal fluctuations can limit the maximum amplification gain of the seed pulse, consider the amplification at the identically zero detuning of the three-wave interaction. After a certain period of time t_m, the amplification gain

$$D_m \sim e^{G_m}, \quad G_m = \gamma t_m,$$

becomes sufficient for thermal fluctuations to deplete the pump wave substantially, and further amplification of the seed pulse is then suppressed. The dimensionless

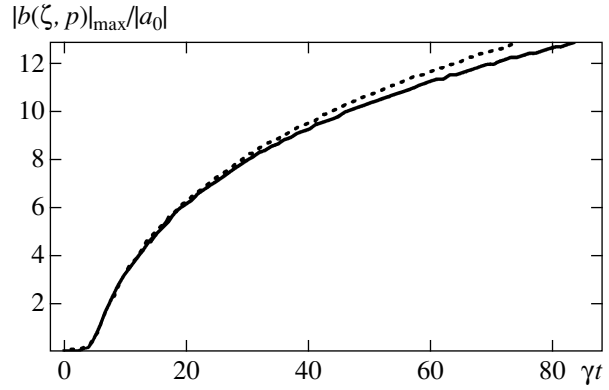


Fig. 8. The maximum amplitude of a horseshoe-type pulse normalized to the amplitude of the pump wave ($|b(\zeta, \rho)|_{max}/|a_0|$) as a function of time $\gamma t \equiv |a_0|\tau$ (solid line). The dotted line represents a 1D solution with $\epsilon_{eff}(\tau)$ for $b_0(\zeta, R/2)$ (the same initial conditions as in Fig. 7). The front peak of the amplified pulse grows similarly to the one of the 1D self-similar profile with decreasing ϵ_{eff}.

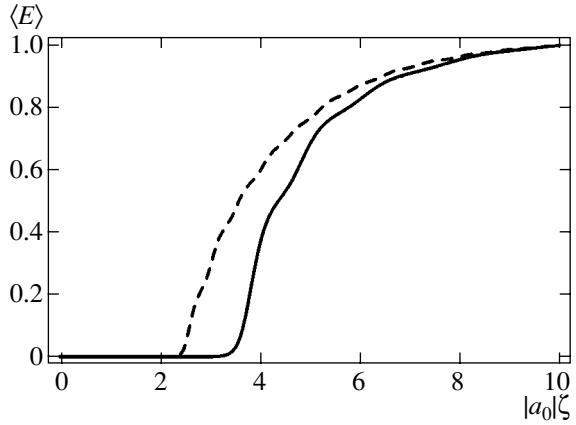


Fig. 9. The normalized energy integral distribution within a horseshoe-type amplified pulse (averaged over the capillary cross section): $\gamma t \equiv a_0\tau = 20$ (solid line) and $\gamma t = 90$ (dashed line); the same parameters as in Fig. 7. For larger γt, the averaged energy distribution becomes a smooth function (cf. Fig. 4), and it is therefore difficult to distinguish the individual peaks of the amplified pulse. On average, the energy is distributed over a length that is significantly larger than the length of the first peak.

quantity G_m depends on the plasma temperature and does not depend on the amplitude of the pump wave. The maximum amplification of the desired signal with respect to a_0 is then given by

$$\frac{b_{max}}{a_0} \approx \frac{2G_m\sqrt{2}}{1 + \ln\left(\frac{4}{\epsilon_0}\sqrt{2\pi}\right)} \tag{33}$$

and is independent of the amplitude of the pump. For $G_m \approx 20$, the electromagnetic wavelength $\lambda = 1$ μm, the initial pulse duration of 50 fs, and the initial pulse power density $P = 10^{13}$ W/cm^3, we find that the maximum amplification that can be achieved in a pump of an arbitrary intensity before the noise is amplified to the level of suppressing the pump is $b_{max}/a_0 \approx 6$.

Nevertheless, through a nonlinear filtering mechanism identified in [5], it is possible to suppress the unwanted instability of the pump wave without suppressing the desirable seed pulse amplification. The filtering effect occurs because the pumped pulse duration decreases inversely proportional to the pulse amplitude in the nonlinear regime. The pulse frequency bandwidth increases with the pulse amplitude, and the growing nonlinear instability can therefore tolerate larger and larger external detuning from the backscattering resonance. Because the linear instability, i.e., the exponential growth of thermal fluctuations, has a narrower bandwidth, filtering the desired signal can be achieved by arranging for an appropriate combination of the detuning and nonlinear effects. A slight frequency detuning can be equivalently provided either by pump chirping or by inhomogeneity of the plasma density along the trace of the pulse amplification resulting in variations of the plasma frequency involved in the three-wave resonance condition.

While the exact solution for a delta-pulse amplification problem obtained in [5] precisely deals with the linear profile of frequency detuning, we use an approximate analysis in this section to describe how the pulse amplification develops in the case of an arbitrary detuning profile. To do this, first consider the linear stage of amplification of a weak pulse b governed by the equation

$$(\partial_\tau - i\delta\omega)(\partial_\tau - \partial_z)b = |a_0|^2 b \tag{34}$$

(without the loss of generality, we temporarily neglect the transverse structure and the polarization of the amplified pulse for qualitative conclusions). Using the quasistatic approximation and assuming the detuning to change slowly along the trace of the pulse propagation, we can treat $\delta\omega$ as a slow function of time τ [5]. Perform the Fourier transformation of Eq. (34),

$$b = \int b_{\Delta k} \exp(i\Delta k z) d\Delta k,$$

and take

$$b_{\Delta k}(\tau) = \psi(\tau)\exp\left(i\int_0^\tau \frac{\delta\omega(\tau') + \Delta k}{2} d\tau'\right) \tag{35}$$

to transform the equation for the amplitude of the pulse spatial harmonic ψ to the form

$$\left[\frac{d^2}{d\tau^2} + w^2(\tau)\right]\psi = 0,$$
$$w^2 + i\Omega_\tau + \Omega^2 - |a_0|^2, \quad \Omega = (\delta\omega - \Delta k)/2. \tag{36}$$

In accordance with the assumption of a smooth detuning profile, take

$$q(\tau) = \frac{\delta\omega_\tau}{|a_0|^2} \ll 1.$$

Outside the regions where Ω^2 is close to $|a_0|^2$, the effective frequency w can be estimated as

$$w = \sqrt{\Omega^2 - |a_0|^2} + \frac{i\Omega_\tau}{2\sqrt{\Omega^2 - |a_0|^2}}, \quad (37)$$

and the amplification gain is given by

$$D \sim e^G, \quad G \approx \int \mathrm{Im} w d\tau$$

in the WKB approximation. At $\Omega^2 > |a_0|^2$, G depends on the length of the trace of the pulse propagation logarithmically, and the amplification gain is therefore negligible in the adopted approximation. Thus, the total amplification gain is given by

$$G \approx \int\limits_{\Omega^2 < |a_0|^2} \sqrt{|a_0|^2 - \Omega^2(\tau)} d\tau. \quad (38)$$

For the detuning monotonically changing along the trace of the pulse propagation, Eq. (38) can be written as

$$G \approx \frac{2}{|a_0|} \int\limits_{\Omega^2 < |a_0|^2} \sqrt{1 - \frac{\Omega^2}{|a_0|^2}} \frac{d\Omega}{|q(\Omega)|} \le \frac{\pi}{|q_{\min}|}, \quad (39)$$

where $|q_{\min}|$ stands for the minimum rate of the detuning evolution on the trace of amplification. As can be seen from Eq. (39), the upper limit of the total amplification gain on the entire trace of the pulse propagation is independent of Δk (included in the definition of Ω over which the integration is performed). For q = const, we have

$$D \sim \exp(\pi/|q|),$$

as obtained in [5], and therefore D itself is independent of Δk.

We can also generalize Eqs. (38), (39) to the case of oblique propagation of the pulses, describing the amplification of the electromagnetic noise coming from outside the system. The only difference is then that the group velocity of the amplified harmonic differs from the speed of light, which results only in a redefinition of Δk and does not affect the form of the final result in Eqs. (38) and (39) if $q(\tau)$ is calculated relative to the actual trajectory of the amplified pulse.

Equations (38) and (39) predict that each harmonic of a given frequency and a wave number is amplified only inside the region where the three-wave resonance conditions are satisfied in the sense that $\Omega^2 < |a_0|^2$ (or, in dimensional variables, $(\delta\omega - c\Delta k)^2/4 < \gamma^2$). The idea of the approach given here is similar to the one proposed by Rosenbluth and Pilia (see, e.g., [12]), who estimated the total linear amplification gain for stationary waves in an inhomogeneous medium with the wave-number detuning but with the temporal resonance condition satisfied exactly. The difference between the two cases is that instead of the wave-number detuning, the frequency detuning is important for the Raman pulse amplification in inhomogeneous plasmas. For the Raman backscattering in a cold plasma, the wave-number resonance condition is satisfied automatically, because a plasma wave is allowed to have an arbitrary wave number, although it oscillates at a certain frequency ω_p.

The conclusion that follows from the obtained result is that the detuning profile along the pulse amplification trace can be chosen such that the noise amplification is suppressed above a certain level determined by Eq. (38). Monotonically changing the detuning allows a stronger suppression, because there exists only one region for a given harmonic where the amplification occurs. In this case, the requirement for the characteristic $|q|$ to ensure that the noise is not amplified up to the transition to the nonlinear stage but the desired signal is ($\int b(z)dz \gtrsim 1$, see [1, 3]) can be formulated as

$$\frac{\pi}{G_m} \ll |q| \ll \frac{\pi}{\ln\frac{1}{\epsilon_0}}, \quad (40)$$

$$q = \frac{c}{\gamma^2}\frac{\partial\delta\omega}{\partial z} \approx \frac{10^{14}}{L_\delta \lambda P},$$
$$\frac{1}{L_\delta} = \frac{1}{\omega_p}\frac{\partial\omega_p}{\partial z} + \frac{1}{2\omega_p c}\frac{\partial\omega_a}{\partial t}, \quad (41)$$

where the characteristic spatial scale L_δ of the detuning evolution due to the plasma inhomogeneity (the first term in Eq. (41)) and the pump chirping (the second term) is measured in centimeters, the wavelength λ is measured in microns, and the pump power is measured in W/cm^2.

The next problem is how the frequency detuning influences the desired signal amplification in the nonlinear regime. We now show that it does not as long as these variations remain sufficiently smooth. To prove this, we consider the change of variables

$$a = \tilde{a},$$
$$b = \tilde{b}\exp(i\delta\omega(\tau + z)), \quad (42)$$
$$f = \tilde{f}\exp(-i\delta\omega(\tau + z)),$$

leading to the following 1D form of Eqs. (6)–(8):

$$\partial_\tau\tilde{a} + \partial_z\tilde{a} = \tilde{b}\tilde{f},$$
$$\partial_\tau\tilde{b} - \partial_z\tilde{b} - i(\tau + z)qa_0^2\tilde{b} = -\tilde{a}\tilde{f}^*, \quad (43)$$
$$\partial_\tau\tilde{f} = -\tilde{a}\tilde{b}^*.$$

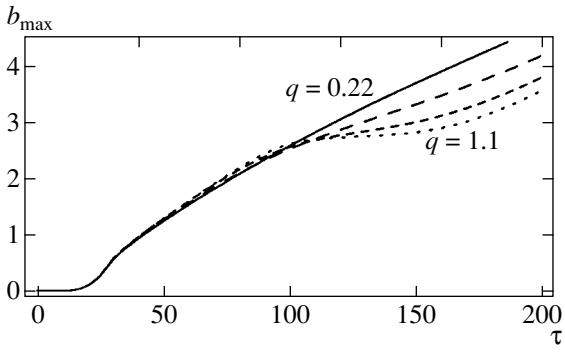


Fig. 10. Pulse-detuned nonlinear amplification $b_{max}(\tau)$ at different frequency detuning profiles $\delta\omega(\tau) = 2(1 + \tanh((\tau - 100)/\tau_0))$: $q = 2/(a_0^2\tau_0) = 0.22, 0.44, 0.74, 1.1$; $a_0 = 0.3$. (The larger q is, the lower the graph goes at $\tau > 100$.) For $q \sim 1$, the amplification efficiency decreases in the region where the detuning evolves relatively fast ($100 < \tau < 150$), while, at small q (e.g., for $q = 0.22$), the amplification proceeds exactly as in the case of zero detuning for all τ.

These equations are equivalent to Eqs. (6)–(8) with zero detuning if $q = 0$. The physical meaning of the formal change of variables (42) is as follows. The carrier frequencies of the seed pulse and the plasma wave are chosen such that the three-wave resonance condition is satisfied locally,

$$\tilde{\omega}_a - \tilde{\omega}_b(z) = \tilde{\omega}_f(z), \tag{44}$$

where

$$\tilde{\omega}_b(z) = \omega_b + \delta\omega(z), \quad \tilde{\omega}_f(z) \equiv \omega_p(z)$$

are functions of space, and the carrier frequency of the pump wave $\tilde{\omega}_a = \omega_a$ is left unchanged. It is only the gradient of the detuning that enters Eqs. (43), and the constant part of $\delta\omega$ enters the initial conditions for the seed pulse only.

In the frame moving together with the seed pulse ($\zeta = z + \tau$), in the quasistatic approximation [3], the basic equations can be written as

$$\begin{aligned} 2\partial_\zeta \tilde{a} &= \tilde{b}\tilde{f}, \\ \partial_\tau \tilde{b} - i\zeta q a_0^2 \tilde{b} &= -\tilde{a}\tilde{f}^*, \\ \partial_\zeta \tilde{f} &= -\tilde{a}\tilde{b}^*, \end{aligned} \tag{45}$$

which implies that the term corresponding to the detuning is negligible compared to the nonlinear drive when the overfall of the detuning $\Delta(\delta\omega)$ on the length of the pulse is small compared to $1/\tau$. Because

$$\Delta(\delta\omega) \sim q a_0^2 \zeta_{pulse},$$

where the characteristic length of the pulse is $\zeta_{pulse} \sim 1/a_0^2\tau$ at the nonlinear stage of interaction [3], the condition of negligible detuning becomes

$$|q| \ll 1. \tag{46}$$

The obtained condition for efficient amplification of short pulses was tested numerically. It can be seen from Fig. 10 that, for $q \sim 1$, the amplification efficiency decreases in the region where the detuning evolves relatively fast, while, at small q (e.g., for q = 0.22), the amplification proceeds exactly as in the case of zero detuning for all τ, exactly as predicted by the qualitative arguments given above.

The obtained results imply that, for short pulses, amplification can be efficient on the entire trace of the interaction with the pump wave. The integral variation of $\delta\omega$ (or the maximum frequency detuning amplitude experienced by the pulse on its trace of amplification) does not significantly influence the amplification efficiency if the detuning evolves smoothly along the trace of amplification. Condition (46) only requires the bandwidth of the wave envelope $\Delta\omega_b$ to grow due to the nonlinear compression sufficiently fast for the local-resonance frequency $\omega_a - \omega_p(z)$ to lie within the amplification line. For growing $|q|$ that approaches unity, the interaction becomes nonresonant, and the pulse amplification ceases. If q decreases, the pulse amplification develops similarly to the solution with a constant detuning. The degenerate case where q = const and the amplification efficiency depends on the amplitude of the initial pulse logarithmically is discussed in detail in [1, 5].

In a real experiment, transverse plasma inhomogeneities must be taken into account in addition to the detuning provided by pump chirping and longitudinal variations of the plasma density. It is important that the dependence of $\delta\omega$ on the transverse location lowers the sensitivity of the interaction efficiency to the average detuning (over the cross section). In the 1D problem, as shown above, the pulse amplification can be entirely suppressed by large gradients of the plasma density. But in the case where the plasma density also changes in the transverse direction, a radial position ρ_* such that $\delta\omega(\rho_*) = 0$ exists at every cross section of the pulse trajectory. The pulse can extract energy from the pump wave in the vicinity of $\rho = \rho_*$, although the interaction remains inefficient far from this point. This local pulse amplification cannot be entirely suppressed by large detuning that might exist at other radial positions. This fact determines a higher robustness of the pulse amplification in inhomogeneous plasmas in 2D or 3D systems than in the 1D case. In the case where the amplification occurs inside a capillary, the pulse energy is mixed in the transverse direction because of the reflection of electromagnetic waves from the walls of the waveguide, which eventually results in a nonlocal amplification of the waveguide eigenmodes, i.e., in the amplification of the entire pulse.

7. DISCUSSION

Characteristic parameters of the proposed Raman-backscattering pulse amplification experiment are given in table. For the wavelength $\lambda \approx 1$ μm and the radius of the capillary sufficiently large for the radiation energy losses to be negligible, the single-mode operation can only be provided by low pump intensities, which do not allow significant amplification on a reasonable (centimeter size) interaction length. At pump intensities higher than the critical one, multimode solutions are formed.

The parameters given in table correspond to the maximum possible amplification gain at the given wavelength and the electron density limited by such effect as the Langmuir wave breaking and the forward Raman scattering instability [1, 3], which remained out of the scope of our study and represent a field of further research in the context of the 3D Raman scattering problem. As regards the modulation instability, it is expected to be suppressed for the proposed parameters because the critical power of the amplified pulse self-focusing $P_{\text{crit}} = 17(\omega/\omega_p)^2$ GW [3] is equal to 1.7×10^{12} W, which is less than the power of the amplified pulse.

In summary, using a dielectric capillary for channeling laser radiation in a Raman amplifier provides a significant advantage as regards maintaining high interaction efficiency at distances larger than the diffraction length, which allows obtaining higher intensities of the output radiation. In addition, various mechanisms of selective mode discrimination and nonlinear competition of capillary modes are provided by the transverse waveguide dispersion, but cannot be achieved in a boundless vacuum. Although the presence of the capillary walls can influence the structure of the pulse, it does not alter the amplification of the front peak of the pulse, which carries a significant amount of the total energy of the pulse.

We find that, depending on the intensity of the pump, two possible regimes of operation can be realized within a capillary, namely, the single-mode and the multimode pulse amplification. For a low pump wave intensity, when the single-mode operation is possible, the problem admits the resonant mode approach that we develop in this paper. We also develop the linear theory of pulse amplification inside a capillary by generalizing the 1D linear problem. Contrary to the intuitive expectations, we show that the pulse amplification efficiency is not critically sensitive to the transverse structure of the pump wave, and therefore, both lower and higher modes of the pump provide approximately the same amplification rates of the seed pulse.

We generalize the mechanism of avoiding the pump wave instability (resulting in noise amplification) by chirping the pump wave or inhomogeneous plasma profile along the trace of the pulse propagation [5] in the case of an arbitrary smooth detuning profile. We show that, as the noise amplification can be suppressed by detuning, the latter does not alter the amplification of the desired pulse as long as the detuning profile remains sufficiently smooth. We conclude that guiding laser pulses through the capillary provides an additional robustness of the interaction efficiency with respect to transverse inhomogeneities of the plasma density.

Sample parameters for the Raman amplification inside an oversize dielectric capillary

Wavelength λ	1 μm
Electron density n_e	10^{19} cm^{-3}
ω/ω_p	10
Radius of capillary R	50λ
Diffraction length z_R	0.16 cm
Inverse decay rates α_{nm}^{-1}	60/40 cm
Trace of amplification	1.2 cm
Pulse duration	40 ps
a_0	0.006
Pump intensity	10^{14} W/cm^2
Pump power	4×10^9 W
Amplification length c/γ	0.12 mm
Seed pulse duration	100 fs
Seed pulse intensity	10^{14} W/cm^2
ϵ_0	0.25
Amplification factor $b_{\max}/a_0$	20
Amplified pulse intensity	3.5×10^{16} W/cm^2
Amplified pulse power	1.4×10^{12} W

The refraction index of capillary walls is taken to be $n = 1.5$; the pump wave intensity corresponding to $a = a_{\text{crit}}$ is 1.4×10^{11} W/cm^2, and the amplified pulse is therefore of the horseshoe type; the inverse spatial decay rates α_{nm}^{-1} are calculated for the two most slowly decaying modes.

ACKNOWLEDGMENTS

The authors thank A.A. Balakin for help in performing numerical simulations and A.N. Stepanov and V.A. Mironov for fruitful discussions. This work was supported by the Russian Foundation for Basic Research (project nos. 01-02-17388 and 02-02-06258).

APPENDIX

Waveguide Modes of a Dielectric Capillary

The waves channeled by a dielectric capillary can be separated into surface and waveguide-type waves [13]. A slow surface wave propagates without dissipation inside the dielectric walls of the tube with the wave number

$$h = \sqrt{\epsilon k^2 + \kappa_\epsilon^2},$$

where ϵ is the dielectric permittivity, $\kappa_\epsilon \sim 1/d$ is the wave transverse wave number, and d is the width of the capillary wall. Outside the dielectric, the field of the surface wave decays exponentially with the spatial decrement

$$\kappa_0 = \sqrt{h^2 - k^2} = \sqrt{(\epsilon - 1)k^2 - \kappa_\epsilon^2} \sim k$$

for $kd \gg 1$. Therefore, at the distance of several wavelengths from the wall, the surface wave field essentially equals zero, and as regards the interaction of pulses inside the capillary, the impact of the surface wave field can be neglected.

Waveguide-type waves propagate inside the capillary, with the channeling provided by reflection of waves from the inner surface of the capillary dielectric wall. For paraxial propagation ($k \gg 1/R$), the reflection coefficients of most of the waveguide-type waves are close to unity. The only exception is given by several waves with transverse wave numbers close to the resonant ones, for which the dielectric walls of the given width are transparent. Unless the capillary transverse sizes are maintained with high precision, which is not usually the case for the applications similar to the Raman amplifier, these resonances disappear because of the random corrugation of the wall surface. In this case, one can therefore treat all the waveguide-type waves as decaying slowly.

In the first-order approximation, the boundary conditions for the electric and magnetic fields on the inner wall of the dielectric capillary (under the assumption of a negligible decay rate) are given by

$$E_r(R) = H_r(R) = 0$$

(see [13]). The transverse structure of the electric field is then given by

$$\boldsymbol{\psi}_{m,n,\pm 1} = \frac{\mathbf{p}_{\pm 1}}{\sqrt{\pi}} \frac{J_{m\pm 1}(\mu_{m\pm 1,n} r/R)}{R J_m(\mu_{m\pm 1,n})} \exp(im\theta), \qquad (47)$$

where

$$\mathbf{p}_{\pm 1} = \frac{\boldsymbol{\theta}^{(0)} \pm i\mathbf{r}^{(0)}}{\sqrt{2}} = \frac{(\mathbf{y}^{(0)} \pm i\mathbf{x}^{(0)})\exp(\pm i\theta)}{\sqrt{2}}$$

are unit polarization vectors and $\mu_{m\pm 1,n}$ are the roots of the Bessel functions ($J_{m\pm 1}(\mu_{m\pm 1,n}) = 0$). Eigenmodes (47) are normalized such that

$$\langle \boldsymbol{\psi}_{m_1,n_1,j_1} | \boldsymbol{\psi}_{m_2,n_2,j_2} \rangle = \delta_{m_1,m_2}\delta_{n_1,n_2}\delta_{j_1,j_2}, \qquad (48)$$

where $m_{1,2}$ stand for the azimuthal indices, $n_{1,2}$ stand for the radial indices, and $j_{1,2}$ determine the polarization of the modes. The decay rate α_n for the nth mode can be obtained in the second order of perturbation theory under the assumption of the known transverse structure of the mode. Explicit expressions for α_n are given in [9] (see also Section 4).

REFERENCES

1. V. M. Malkin, G. Shvets, and N. J. Fisch, Phys. Plasmas **7**, 2232 (2000).
2. G. A. Mourou, C. P. J. Barty, and M. D. Perry, Phys. Today **51**, 22 (1998).
3. V. M. Malkin, G. Shvets, and N. J. Fisch, Phys. Rev. Lett. **82**, 4448 (1999).
4. G. Shvets, N. J. Fisch, A. Puknov, and J. Meyer-ter-Vehn, Phys. Pev. Lett. **81**, 4879 (1998).
5. V. M. Malkin, G. Shvets, and N. J. Fisch, Phys. Rev. Lett. **84**, 1208 (2000).
6. G. M. Fraiman, N. A. Yampolsky, V. M. Malkin, and N. J. Fisch, submitted to Phys. Plasmas (2002).
7. N. S. Vorob'ev, A. B. Grudinin, E. M. Dianov, *et al.*, Pis'ma Zh. Éksp. Teor. Fiz. **44**, 15 (1986) [JETP Lett. **44**, 17 (1986)].
8. A. B. Grudinin, E. M. Dianov, D. V. Korobkin, *et al.*, Pis'ma Zh. Éksp. Teor. Fiz. **45**, 211 (1987) [JETP Lett. **45**, 260 (1987)].
9. E. A. J. Marcatili and R. A. Schmeltzer, Bell. Syst. Tech. J. **43**, 1783 (1964).
10. A. G. Litvak, Zh. Éksp. Teor. Fiz. **57**, 629 (1970) [Sov. Phys. JETP **30**, 344 (1970)].
11. W. L. Kruer, *The Physics of Laser Plasma Interactions* (Addison-Wesley, Redwood City, 1988).
12. A. A. Galeev and R. Z. Sagdeev, in *Reviews of Plasma Physics*, Ed. by M. A. Leontovich (Atomizdat, Moscow, 1973; Consultants Bureau, New York, 1979), Vol. 7.
13. L. A. Vaĭnshteĭn, *Electromagnetic Waves* (Radio i Svyaz, Moscow, 1988).

Journal of Experimental and Theoretical Physics, Vol. 95, No. 4, 2002, pp. 639–644.
Translated from Zhurnal Éksperimental'noĭ i Teoreticheskoĭ Fiziki, Vol. 122, No. 4, 2002, pp. 738–745.

ATOMS, SPECTRA, RADIATION

Biphoton Light Generation in Polarization-Frequency Bell States

A. V. Burlakov*, S. P. Kulik, G. O. Rytikov, and M. V. Chekhova

Moscow State University, Vorob'evy gory, Moscow, 119899 Russia
**e-mail: postmast@qopt.phys.msu.su*

Received May 22, 2002

Abstract—Four polarization-frequency Bell states are obtained experimentally for photon pairs (biphotons) emitted during spontaneous parametric scattering from continuous pumping in the collinear frequency-nondegenerate regime. The polarization properties of such states are investigated. It is shown that biphoton light in the singlet Bell state is not polarized in the second or fourth order in the field. *© 2002 MAIK "Nauka/Interperiodica".*

1. BELL STATES

Entangled states of quantum systems occupy a significant place in quantum optics and especially in quantum informatics. The concept of entanglement of a quantum system was proposed for the first time by Schrödinger (in connection with the well-known Einstein–Podolsky–Rosen paradox) in [1]; however, this property was not defined exactly in that publication. After some time, entangled states ceased to be the object of just philosophical discussions and gedanken experiments since various methods of their experimental preparation had been developed. Accordingly, more rigorous definitions of such states were proposed. If we confine ourselves to the case of a pure state of a complete quantum-mechanical system consisting of several parts, the property of entanglement is defined as nonfactorizability of the overall wave function and can be reduced to the existence of quantum correlations between the parts of the system [2]. Among pure entangled states of two-quantum systems, the so-called Bell states [3],

$$\Phi^{\pm} \equiv \frac{1}{\sqrt{2}}(|\uparrow\uparrow\rangle + |\downarrow\downarrow\rangle),$$
$$\psi^{\pm} \equiv \frac{1}{\sqrt{2}}(|\uparrow\downarrow\rangle + |\uparrow\downarrow\rangle), \quad (1)$$

play a special role. Here, we assume that each quantum system has two eigenstates $|\uparrow\rangle$ and $|\downarrow\rangle$ (this may be a particle with a spin of 1/2, an atom in a resonant field, a polarized photon, etc.) It is these states that are used for verifying Bell inequalities, in experiments on quantum teleportation, in a number of protocols of quantum cryptography, and other trends in quantum optics. In particular, such states form a convenient basis for describing an arbitrary quantum state of two two-level systems. The state Ψ^- is often referred to as a singlet state since it is similar to the antisymmetric state of two particles with a spin of 1/2.

2. GENERATION OF BELL STATES OF PHOTONS DURING SPONTANEOUS PARAMETRIC SCATTERING OF LIGHT

Entangled states have been realized experimentally for various quantum systems such as two atoms, an atom and a photon, and two ions. The experiments with entangled (correlated) states of photons have become the most popular. The most effective method for generating correlated photon pairs is that involving spontaneous parametric scattering (SPS) [4].

In the case of spontaneous parametric scattering, the pumping radiation with frequency ω_p and wave vector $\mathbf{k}_p$ incident on a crystal with a quadratic nonlinearity of χ leads to the emergence of scattered radiation at the crystal exit; the state of the latter radiation can be presented in the form [5]

$$|\Psi\rangle = |\text{vac}\rangle + \frac{1}{2}\sum_{\mathbf{k},\,\mathbf{k}'} F_{\mathbf{k}\mathbf{k}'}|1_{\mathbf{k}}, 1_{\mathbf{k}'}\rangle, \quad (2)$$

where $|\text{vac}\rangle$ stands for the vacuum state and $|1_{\mathbf{k}}, 1_{\mathbf{k}'}\rangle$ is the state with one photon in mode $\mathbf{k}$ (signal photon) and one photon in mode $\mathbf{k}'$ (idler photon), which is often referred to as a biphoton. Indices $\mathbf{k}$ and $\mathbf{k}'$ label the frequency, spatial, and polarization modes. The quantity $F_{\mathbf{k}\mathbf{k}'}$ is often called the biphoton amplitude. In the stationary case, when radiation emitted by a CW single-mode laser is used for pumping and the medium parameters do not depend on time, the biphoton amplitude is proportional to $\delta(\omega + \omega' - \omega_p)$, where ω and ω' are the frequencies of the signal and idler photons. If, in addition, the scattering occurs in a plane layer unbounded in the directions transverse to the wave vector of pumping, $F_{\mathbf{k}\mathbf{k}'}$ is also proportional to $\delta(\mathbf{k}_\perp + \mathbf{k}'_\perp)$, where $\mathbf{k}_\perp$ and

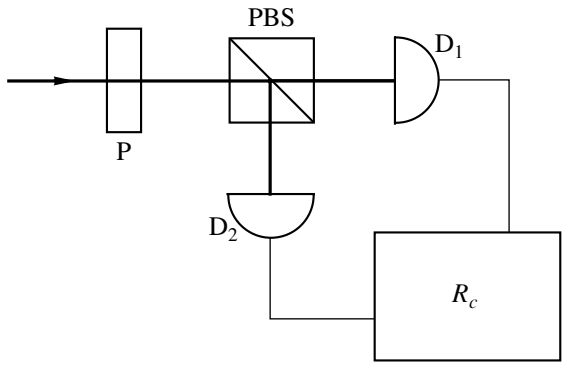


Fig. 1. Schematic diagram for observation of "latent polarization" [15]. After the polarization transformer P, the beam splits by the polarization beam splitter PBS and is directed to photodetectors D_1 and D_2. The coincidence count rate R_c may depend on the polarization transformation even if the intensities registered by the detectors are independent of it.

$\mathbf{k}'_\perp$ are the transverse components of the wave vectors. In this case, the summation over $\mathbf{k}'$ in Eq. (2) disappears, and the second term describes an entangled state of the two photons. This state becomes a Bell state if the sum over $\mathbf{k}$ contains only two terms.

Depending on the parameters in which "entanglement" occurs (frequency, direction of the wave vector (scattering angle), polarization), we can single out the following three classes of Bell states generated during spontaneous parametric scattering.

1. Polarization-angle Bell states. The signal and idler photons are emitted at different angles θ and θ' to the pumping wave vector; polarization for each photon is not specified, but there exists a correlation (entanglement) between the two polarizations. The two-photon part of the vector of state in this case has the form

$$|H_\theta V_{\theta'}\rangle \pm |V_\theta H_{\theta'}\rangle \quad \text{or} \quad |H_\theta H_{\theta'}\rangle \pm |V_\theta V_{\theta'}\rangle,$$

where symbols H and V denote the horizontal and vertical polarizations. Such states were realized for the first time by using type II synchronism[1] [6]. Subsequently, a more convenient scheme [7] was proposed, in which analogous states were obtained as a result of interference of biphotons generated in two successively arranged crystals with type I synchronism.

2. Frequency-angle Bell states. In the case of noncollinear nondegenerate spontaneous parametric scattering with type I synchronism for small frequency detuning of the signal and idler photons from the pumping frequency, we can single out such directions of scattering θ and θ' in which a signal photon of frequency ω and an idler photon of frequency ω' are emitted. In this case, the two-photon part of the vector of state has the form

$$|\omega_\theta \omega'_{\theta'}\rangle \pm |\omega'_\theta \omega_{\theta'}\rangle;$$

i.e., Bell states $\Psi^\pm$ are generated. The experimental realizations of such states are described in [8].

3. Finally, it is possible to prepare polarization-frequency Bell states of the form

$$|H_\omega V_{\omega'}\rangle \pm |V_\omega H_{\omega'}\rangle \quad \text{or} \quad |H_\omega H_{\omega'}\rangle \pm |V_\omega V_{\omega'}\rangle. \qquad (3)$$

An experiment in which such states were obtained is described in [9], where the interferometric method of preparing Bell states was used. The modes ω and ω' differed not only in frequency, but also in the direction of propagation (two-beam regime). However, the preparation of states (3) in the one-beam regime realized in a collinear spontaneous parametric scattering is of special interest. The one-beam regime of two-photon light generation is used comparatively rarely in quantum optics, although it is this regime which is of interest for data transmission. In addition, it will be shown below that such Bell states may possess interesting features as regards their polarization properties. Namely, light in one of the states (3) turns out to be nonpolarized in all orders in the field.

3. LATENT POLARIZATION OF LIGHT

Many authors (see, for example, [10–12]) considered the polarization of light in higher (than second) orders in the field. It was shown in [11] that a situation is possible when light is not polarized in the second order in the field, but exhibits polarization dependences in the fourth-order (in field) correlation functions. Such a property (latent polarization) is observed, for example, in parametric scattering radiation with a collinear frequency-degenerate type II synchronism, which was demonstrated experimentally in [13]. A classical analogue of this effect was also proposed in [11], and its experimental realization is described in [14].

The schematic diagram of experimental observation of latent polarization is shown in Fig. 1 [15]. The radiation under investigation is directed to the polarization beam splitter PBS with a pair of photodetectors D1 and D2 mounted at the exits. In front of the beam splitter, a system of phase plates (two such plates are sufficient) P is mounted, which makes it possible to carry out any polarization transformation. If light is not polarized in the second order in the field, the intensity registered by each of the detectors remains unchanged under any polarization transformation.

Let us suppose that a correlation function of the form

$$G^{(2)}_{HV}(\tau) = \langle E^{(-)}_H(t) E^{(-)}_V(t+\tau) E^{(+)}_H(t) E^{(+)}_V(t+\tau)\rangle \quad (4)$$

is measured in an experiment, where $E^{(-)}$ and $E^{(+)}$ are, respectively, the negative-frequency and positive-frequency fields, the subscripts H and V denoting the lin-

[1] In the case of type I synchronism, the signal and idler photons are polarized identically; in the case of type II synchronism, their polarizations are orthogonal.

ear horizontal and vertical polarization modes. This quantity, which characterizes intensity correlation in the polarization modes, may depend on polarization transformation in front of the beam splitter even for light nonpolarized in the second order. In order to measure the correlation function (4), signals from detectors are directed to the input of the photocounting coincidence circuit; the coincidence count rate R_c is determined by the value of $G^{(2)}_{HV}$. For example, the correlation function (4) in [13] was measured for polarization transformations carried out by rotating the half-wave plate. If two-photon light was present at the input, modulation of the coincidence count rate was observed with a high visibility.[2] In the presence of a classical source (e.g., radiation from two orthogonally polarized lasers with independent phase fluctuations [14]) at the input, a modulation of $G^{(2)}_{HV}$ with a 50% visibility is observed in the experiment.

Let us now consider four polarization-frequency Bell states (3). It is well known that the singlet state Ψ^- is invariant to any polarization transformations [16]. It should not display any polarization dependence during measurements of moments of any order in the field, including the correlation function (4). It can also be noted that the state Φ^+ is invariant to rotations of the polarization plane. Consequently, no modulation in the number of coincidences must be observed for such a state (or for the Ψ^- state) in the experiment described in [13].

Thus, light in the singlet Bell state Ψ^- is nonpolarized in the second as well as fourth order in the field. Since all moments for two-photon light can be expressed in terms of second- and fourth-order moments, light in state Ψ^- is nonloparized in all orders in the field. Such light can be referred to as completely nonpolarized.

4. "SCALAR LIGHT"

The state of light Ψ^- is close to the state of polarization-scalar light proposed in [10] (see also [17]). A transition from the state of polarization-scalar light to state Ψ^- occurs in the limit of a low pumping power or a small parametric transformation coefficient. For polarization-scalar light, fluctuations of all Stokes parameters must be suppressed.

The fluctuations of Stokes parameters can be easily calculated for all four polarization-frequency Bell states. It turns out as a result that fluctuations of the third Stokes parameter are suppressed in state Φ^+ ($\Delta S_3^2 = 0$); $\Delta S_2^2 = 0$ in state Φ^-; $\Delta S_1^2 = 0$ in state Ψ^+; and in the singlet state Ψ^-, the fluctuations of all three Stokes parameters are suppressed: $\Delta S_1^2 = \Delta S_2^2 = \Delta S_3^2 = 0$. However, the measurements of fluctuations of Stokes parameters for biphoton light would require very rapid detectors (with a time resolution of the order of reciprocal width of the parametric scattering spectrum, i.e., of the order of hundreds of femtoseconds). Such detectors are not available at present. The suppression of fluctuations of the Stokes parameters can be obtained with the help of detectors with a nanosecond time resolution, but only in the case of parametric oscillation of light, when the emission spectrum is considerably narrower than for spontaneous parametric scattering. In this study, we used the spontaneous scattering regime; accordingly, we measured not the fluctuations of the Stokes parameters, but correlation functions of intensity.

[2] In such an experiment, the visibility must be 100% according to the theory.

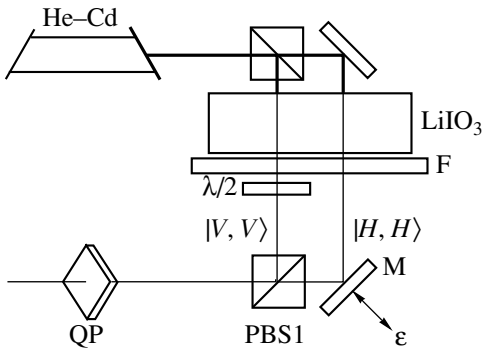


Fig. 2. Schematic diagram of experiment. Parametric scattering in the collinear frequency-nondegenerate regime with type I synchronism occurs in two spatially separated regions in a lithium iodate crystal. Biphoton radiation at the crystal exit is in the state $|H_\omega H_{\omega'}\rangle$ in both beams. In the right beam, polarization is rotated with the help of a λ/2 plate. The phase ε between the beams is controlled by mirror M. The phase incursion between the extraordinarily polarized wave and the ordinarily polarized wave at frequency ω in the quartz plate QP exceeds the corresponding phase incursion at frequency ω' by π.

5. PREPARATION AND ANALYSIS OF POLARIZATION-FREQUENCY BELL STATES

The schematic diagram of the experiment on obtaining four polarization-frequency Bell states is presented in Fig. 2. Continuous pumping (radiation of helium–cadmium laser at a wavelength of 325 nm) through a nonpolarization beam splitter is directed to the interferometer whose both arms contain a lithium iodate crystal. In the crystal, lights experiences spontaneous parametric scattering with a nondegenerate collinear type I synchronism; as a result, biphoton radiation in state $|H_\omega H_{\omega'}\rangle$ is present in both arms behind the crystal. The wavelength of the signal and idler photons are 635 and 665 nm, respectively. Pumping radiation behind the crystal is cut off by filter F. A half-wave plate in one of the arms rotates polarization through π/2, transforming

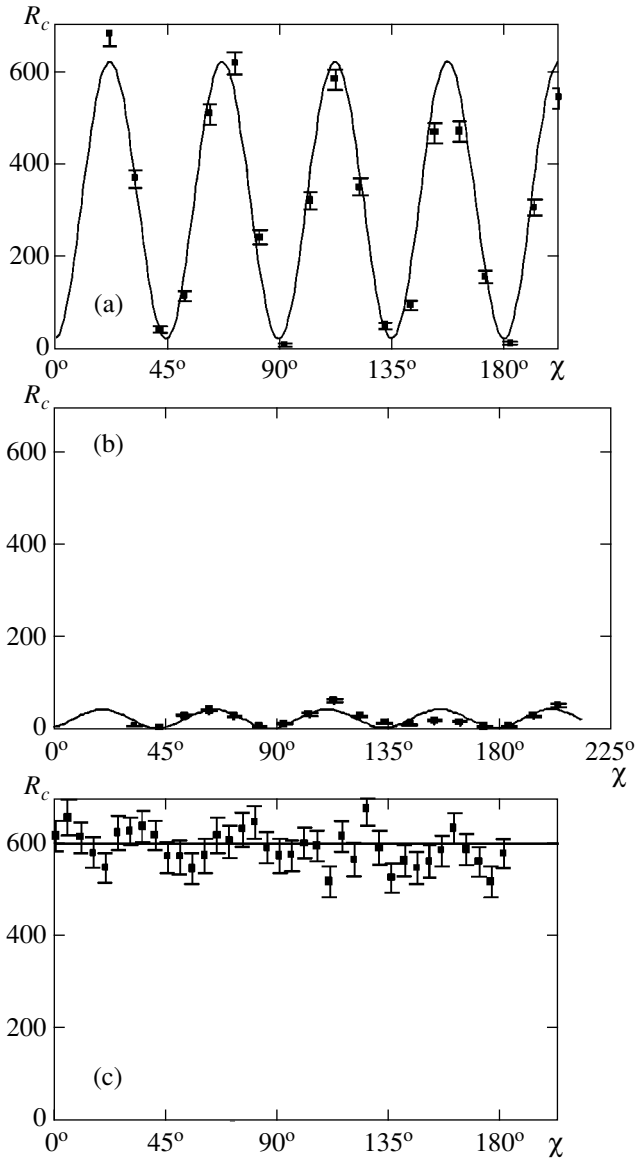


Fig. 3. Dependence of the coincidence count rate (number of coincidences in 200 s) on the angle of rotation χ of the $\lambda/2$ plate for (a) Φ^-, (b) Φ^+, and (c) Ψ^-. The solid curve in (a) corresponds to relation (10) with an added constant background, which corresponds to a visibility of 94%. The solid curve in (b) is plotted under the assumption that the intensities of biphoton beams differ by 20%; in this case, calculations give the same dependence as in (a), but with an amplitude smaller by a factor of 20. The theoretical dependence in (c) is depicted by the straight line.

the beam state into $|V_\omega V_{\omega'}\rangle$, and both beams meet without loss in the polarization beam splitter PBS1. Mirror M in the interferometer can be displaced with the help of a piezoelectric feed. The state of the biphoton field at the exit of the interferometer has the form

$$|\Psi\rangle = \frac{1}{\sqrt{2}}(|H_\omega H_{\omega'}\rangle + e^{-i\varepsilon}|V_\omega, V_{\omega'}\rangle), \qquad (5)$$

where phase ε can be varied by displacing mirror M. For $\varepsilon = 0$, Bell state Φ^+ is formed, while, for $\varepsilon = \pi$, we have Bell state Φ^-. State Ψ^+ can be obtained from Φ^- with the help of a half-wave plate oriented at an angle of $\pi/8$.

Bell states Φ^+, Φ^-, and Ψ^+ are analogues of states obtained in [18] by using the same experimental setup, but for the degenerate regime of parametric scattering. As we pass to the degenerate scattering regime, state Φ^+ is transformed into a pair of correlated photons with right-circular and left-circular polarizations, state Φ^- is converted into a pair of linearly polarized photons at angles $\pm\pi/4$, and state Ψ^+ is transformed into a pair of photons polarized along the vertical and the horizontal.

The singlet state Ψ^- has no analogue in the degenerate regime since it is antisymmetric relative to the transposition of photons in a pair. In order to prepare this state, a special phase plate made of a quartz crystal (QP) was used in experiments. The thickness of this plate satisfied the following condition: the phase incursion between the ordinary and extraordinary waves at frequency ω differs from the corresponding phase incursion at frequency ω' by π. If the state $\Psi^+ = |H_\omega V_{\omega'}\rangle + |V_\omega H_{\omega'}\rangle$ exists at the entrance of such a plate and its optical axis is oriented along the vertical or horizontal, the state behind the plate has the form $\Phi^- = |H_\omega V_{\omega'}\rangle - |V_\omega H_{\omega'}\rangle$ to within an insignificant common phase. In order to obtain state Ψ^-, phase ε in the interferometer was set equal to π so that the state $\Phi^- = |H_\omega H_{\omega'}\rangle - |V_\omega V_{\omega'}\rangle$ was formed at the exit of the interferometer. In the basis *XY* turned through $\pi/4$ relative to the basis *HV*, state Φ^- is transformed into Ψ^+:

$$|H_\omega H_{\omega'}\rangle - |V_\omega V_{\omega'}\rangle = |X_\omega Y_{\omega'}\rangle + |Y_\omega X_{\omega'}\rangle.$$

Plate QP is mounted at the exit of the interferometer so that its optical axis is oriented along direction *X*. Behind the plate, the state in the basis *XY* was transformed into Ψ^-; consequently, in view of its invariance to polarization transformations, this state remained unchanged in any polarization basis.

For the four polarization-frequency Bell states obtained by us, measurements were made according to the scheme proposed in [15]: the value of $G_{HV}^{(2)}$ was measured depending on the polarization transformation in front of the beam splitter (see Fig. 1). In order to single out small scattering angles, an aperture was used. Since noise radiation was also present at the entrance of the detecting elements of the setup in addition to radiation from spontaneous parametric scattering, an interference filter of width 40 nm with transmittance peak at a wavelength of 650 nm was used. The filter transmitted both signal and idler radiation. Avalanche photodiodes operating in the photon count mode were used as detectors, and the resolution of the coincidence scheme was 1.5 ns.

The role of polarization transformers was played by $\lambda/2$ and $\lambda/4$ plates. Figure 3 shows the dependences

obtained for $G_{HV}^{(2)}$ during the rotation of the λ/2 plate for Bell states Φ^-, Φ^+, and Ψ^-. The dependence for the Ψ^+ state is not shown since this state is transformed to Φ^- by rotating the basis through π/4; consequently, the angular dependence for this state is the same as for Φ^- to within a shift by π/8 along the abscissa axis.

It should be noted at the very outset that all the experimentally obtained entangled states were nonpolarized in the second order in the field: under polarization transformations by the λ/4 and λ/2 plates, the intensity of the beam detected by each detector (see Fig. 1) remained practically unchanged.[3]

It can be seen that the state Φ^- possesses "latent polarization" (Fig. 3a): upon the rotation of the λ/2 plate, the number of coincidences R_c oscillates with a high visibility (94%). The dependence of R_c on the angle of rotation χ of the plate can be easily derived using expression (4) for the correlation function and writing state Φ^- in the form

$$\begin{aligned}\Phi^- &= |H_\omega H_{\omega'}\rangle - |V_\omega V_{\omega'}\rangle \\ &= [a_H^\dagger(\omega)a_H^\dagger(\omega') - a_V^\dagger(\omega)a_V^\dagger(\omega')]|\text{vac}\rangle,\end{aligned} \tag{6}$$

where $a_{H,V}^\dagger(\omega, \omega')$ are the photon creation operators in the polarization modes H and V and in frequency modes ω and ω'. Now, we express the fields in relation (4) in terms of the creation operator and consider that the left-hand side of Eq. (4) is Hermitian conjugate to the right-hand side. Averaging in Eq. (4) should be carried out over the state $\Phi(\chi)$ obtained from Φ^- as a result of action of the plate. This gives

$$\begin{aligned}&G_{HV}^{(2)}(\tau) \\ &= \left|\iint d\omega_1 d\omega_2 a_H^\dagger(\omega_1) e^{-i\omega_1 t} a_V^\dagger(\omega_2) e^{-i\omega_2(t+\tau)}|\Phi(\chi)\rangle\right|^2.\end{aligned} \tag{7}$$

Using the Jones matrix for the λ/2 plate oriented at angle χ [15],

$$D = \begin{pmatrix} i\cos(2\chi) & i\sin(2\chi) \\ i\sin(2\chi) & -i\cos(2\chi) \end{pmatrix},$$

and expressing the creation operators in front of the plate in terms of the creation operators behind the plate, we obtain the following expression for state $\Phi(\chi)$ accurate to an insignificant phase factor:

$$\begin{aligned}\Phi(\chi) &= \cos(4\chi)(|H_\omega H_{\omega'}\rangle - |V_\omega V_{\omega'}\rangle) \\ &\quad - \sin(4\chi)(|H_\omega V_{\omega'}\rangle + |V_\omega H_{\omega'}\rangle).\end{aligned} \tag{8}$$

After integration, this equation will contain only the terms corresponding to frequencies ω and ω'. Substituting Eq. (8) into relation (7), we obtain

$$G_{HV}^{(2)}(\tau) = 4\cos^2\frac{\omega-\omega'}{2}\tau\sin^2(4\chi) \tag{9}$$

[3] Intensity modulation did not exceed 15%.

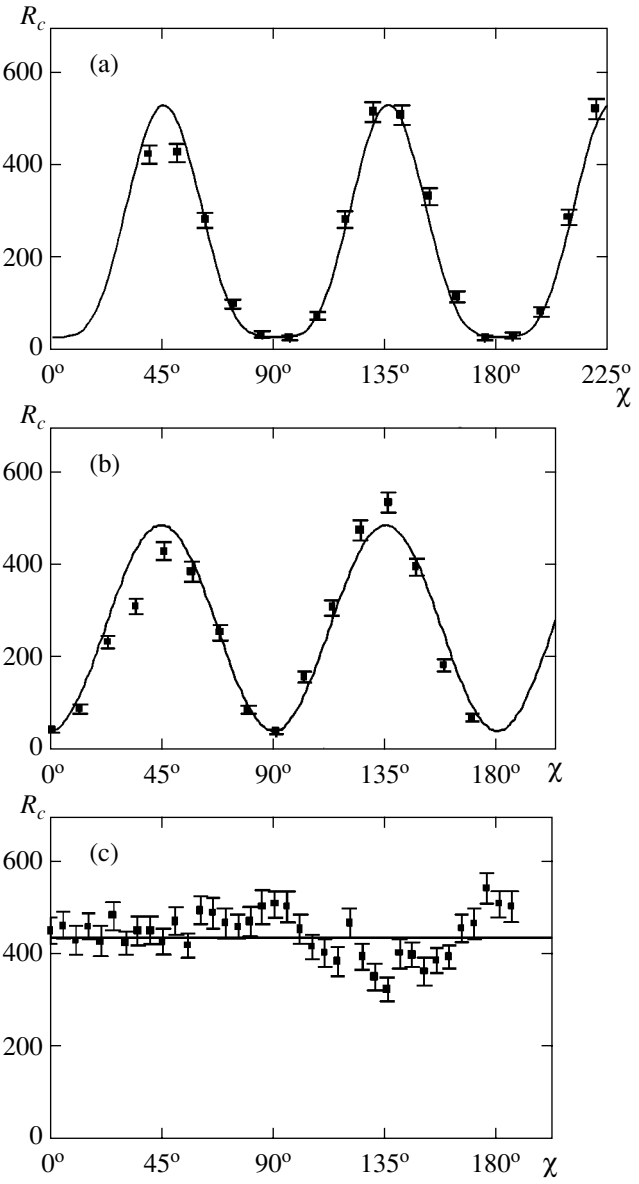


Fig. 4. Dependence of the number of coincidences on the angle of rotation χ of a λ/4 plate for (a) Φ^-, (b) Φ^+, and (c) Ψ^-. Theoretical curves are plotted from relations (11) (a) and (12) (b) with an added background, taking into account the departure of visibility from 100%. The visibility is 93%. The theoretical dependence in (c) is depicted by the straight line.

(the contribution resulting from averaging gives only the second term in Eq. (8)).

The rate of coincidence count in the scheme presented in Fig. 1 is determined by the integral of $G_{HV}^{(2)}(\tau)$ with respect to τ in the limits determined by the time resolution T of the coincidence circuit, which is considerably longer than the period of oscillations of the first factor in Eq. (9). This gives

$$R_c\left(\Phi^-, \frac{\lambda}{2}\right) \propto \sin^2(4\chi), \tag{10}$$

which is in good agreement with the experimental results (Fig. 3a).

If state Φ^+ is formed at the entrance of the polarization transformer, the coincidence count rate is independent of the angle of rotation of the half-wave plate; it can easily be verified that state Φ^+ does not change in this case. The coincidence count rate remains close to zero since the polarization anticorrelation effect takes place [19]. Accordingly, the coincidence count rate in Fig. 3b is smaller than the count rate for state Φ^- by more than an order of magnitude (Fig. 3a). The observed modulation is apparently due to the fact that biphoton beams in different arms of the interferometer had slightly different intensities. For example, if the contributions to the coincidence count rate from the two arms differ by 20%, we can obtain for R_c a dependence analogous to (10), but with an amplitude smaller by a factor of 20.

For the case when state Ψ^- was formed in front of the plate, a high coincidence count rate independent of χ was observed (Fig. 3c). Calculations similar to those described above give in this case a constant value for R_c, equal to the maximum of function (10).

The results presented in Fig. 4 were obtained by using a $\lambda/4$ plate as a polarization transformer. In this case, calculations give the following dependence of state Φ^- at the entrance to the plate:

$$R_c\left(\Phi^-, \frac{\lambda}{4}\right) \propto \sin^4(2\chi). \quad (11)$$

This dependence is in good agreement with the experimental dependence (Fig. 4a).

The change in the value of χ for state Φ^+ at the plate entrance leads to complete modulation of the coincidence count rate (Fig. 4b). Calculations for this case give the following dependence:

$$R_c\left(\Phi^+, \frac{\lambda}{4}\right) \propto \sin^2(2\chi). \quad (12)$$

Finally, for state Ψ^- at the entrance of the $\lambda/4$ plate, the coincidence count rate modulation during the rotation of the plate is virtually absent (Fig. 4c).

Our measurements revealed that biphoton light in state Ψ^- does not possess "latent polarization"; i.e., it is nonpolarized in the second as well as fourth order in the field. All the remaining Bell states exhibit latent polarization.

6. CONCLUSIONS

Thus, biphoton light in a singlet polarization-frequency Bell state turns out to be completely nonpolarized and is polarization-scalar light. Such a state of light was obtained experimentally, as well as the remaining three polarization-frequency Bell states. It is convenient to use "one-beam" geometry in this case since it makes it possible to use biphoton light for data transmission by sending the signal and idler photons through the same optical fiber. It can be expected that such states will be applied for coding and transmission of quantum information.

ACKNOWLEDGMENTS

This study was supported by INTAS (project no. 01-2122) and the Russian Foundation for Basic Research (project no. 02-02-16664).

REFERENCES

1. E. Schrödinger, Proc. Am. Philos. Soc. **124**, 323 (1980).
2. I. V. Bargatin, B. A. Grishanin, and V. N. Zadkov, Usp. Fiz. Nauk **171**, 625 (2001).
3. S. L. Braunstein, A. Mann, and M. Revzen, Phys. Rev. Lett. **68**, 3259 (1992).
4. D. N. Klyshko, Pis'ma Zh. Éksp. Teor. Fiz. **6**, 490 (1967) [JETP Lett. **6**, 23 (1967)].
5. A. V. Belinsky and D. N. Klyshko, Laser Phys. **4**, 663 (1994).
6. P. G. Kwiat, K. Mattle, H. Weinfurter, *et al.*, Phys. Rev. Lett. **75**, 4337 (1995).
7. P. G. Kwiat, E. Waks, A. White, *et al.*, Phys. Rev. A **60**, R773 (1999).
8. J. G. Rarity and P. R. Tapster, Phys. Rev. Lett. **64**, 2495 (1990).
9. Y. Kim, S. P. Kulik, and Y. Shih, Phys. Rev. A **63**, 060301 (2001).
10. V. P. Karasev, J. Sov. Laser Res. **12**, 147 (1991).
11. D. N. Klyshko, Phys. Lett. A **163**, 349 (1992).
12. A. S. Chirkin, A. A. Orlov, and D. Yu. Parashchuk, Kvantovaya Élektron. (Moscow) **20**, 999 (1993).
13. P. Usachev, J. Söderholm, G. Björk, and A. Trifonov, Opt. Commun. **193**, 161 (2001).
14. D. I. Guzun and A. N. Penin, Proc. SPIE **2799**, 249 (1996).
15. D. N. Klyshko, Zh. Éksp. Teor. Fiz. **111**, 1955 (1997) [JETP **84**, 1065 (1997)].
16. *The Physics of Quantum Information: Quantum Cryptography, Quantum Teleportation, Quantum Computation*, Ed. by D. Bouwmeester, A. K. Ekert, and A. Zeilinger (Springer-Verlag, Berlin, 2001; Postmarket, Moscow, 2002).
17. V. P. Karasev and A. V. Masalov, Opt. Spektrosk. **74**, 928 (1993) [Opt. Spectrosc. **74**, 551 (1993)].
18. A. V. Burlakov, M. V. Chekhova, O. A. Karabutova, *et al.*, Phys. Rev. A **60**, R4209 (1999).
19. A. V. Burlakov, M. V. Chekhova, O. A. Karabutova, and S. P. Kulik, Phys. Rev. A **64**, 041803 (2001).

Translated by N. Wadhwa

Journal of Experimental and Theoretical Physics, Vol. 95, No. 4, 2002, pp. 645–661.
From Zhurnal Éksperimental'noĭ i Teoreticheskoĭ Fiziki, Vol. 122, No. 4, 2002, pp. 746–764.

ATOMS, SPECTRA, RADIATION

Semiclassical Approach to States Near the Potential Barrier Top¶

V. A. Benderskiĭ[a, b, *], E. V. Vetoshkin[a], and E. I. Kats[c, d]

[a]*Institute for Problems of Chemical Physics, Russian Academy of Sciences, Moscow oblast, Chernogolovka, 142432 Russia*
[b]*Lab. Spectrometrie Physique, UJF BP 87, St. Martin d'Heres, Cedex, France*
[c]*Institute Laue-Langevin F-38042, Grenoble, France*
[d]*Landau Institute for Theoretical Physics, Russian Academy of Sciences, Moscow oblast, Chernogolovka, 142432 Russia*
**e-mail: bender@icp.ac.ru*

Received May 30, 2002

Abstract—Within the framework of the instanton approach, we present analytical results for the following model problems: (i) particle penetration through a parabolic potential barrier, where the instanton solution practically coincides with the exact (quantum) one; (ii) descriptions of highly excited states in anharmonic potentials of two types: double-well X^4 and decay X^3. For the former potential, the instanton method accurately reproduces not only single-well and double-well quantization, but also a crossover region (in contrast to the standard WKB approach that fails to describe the crossover behavior); for the latter potential, the instanton method allows studying the resonance broadening and collapse phenomena. We also investigate resonance tunneling that plays a relevant role in many semiconducting devices. We show that the instanton approach gives exact (quantum) results in a broad range of energies. Applications of the method and of the results are applicable to various systems in physics, chemistry, and biology exhibiting double-level behavior and resonance tunneling.

1. INTRODUCTION

Semiclassical mechanics has a long history. Surprisingly, however, some long-standing problems still exist in the theory. One of them—the description of states near a potential barrier top with a sufficient accuracy—is the subject of this paper. It is known that the commonly used WKB method (phase integral approach) [1, 2] amounts to matching the wave functions for the classically allowed and forbidden regions. Technically, the procedure works for linear (or first-order) turning points and can be relatively simply performed only in one-dimensional problems. But one-dimensional problems are not of great physical importance, not only because the reduced dimensionality does not allow modelling many relevant experimental situations, but also (at least partially) because one-dimensional quantum mechanical problems can be rather easily solved numerically. Unfortunately, efficiency and accuracy of direct numerical methods in quantum mechanics rapidly degrade for multidimensional systems possessing many degrees of freedom because of an extraordinary amount of computational work required for calculations. Furthermore, an extension of the WKB procedure to multidimensional systems encounters fundamental difficulties because of the still unsolved matching problem for multidimensional WKB solutions, which become singular on caustic lines separating manifolds in phase space with real and imaginary momenta for each among N coordinates. Because the number of these domains increases as $N!$, it is a tremendous task for $N > 2$. After several decades of efforts, a complete and unifying description of multidimensional WKB solutions is still unavailable.

The problem was first addressed long ago, and some attempts to overcome the difficulties of the WKB approach and to improve the accuracy of the method have been performed quite successfully. We note, e.g., [3], where the authors additionally included trajectories of a special type on the complex phase plane in the standard WKB method; the semiclassical motion along these trajectories is described by Weber functions (also see [4]). But the choice of these additional special trajectories (which must be included in order to improve the accuracy of the WKB method near the barrier top) depends on the detailed form of the potential far from the top, and in each particular case, a nonuniversal procedure must therefore be performed from the very beginning (also see more recent publications [5], where the authors use some distortion of Stokes diagrams, or [6], where time-dependent quantum mechanical calculations for anharmonic and double-well oscillators have been performed).

Evidently, therefore, there is some need for a semiclassical approach different from the WKB method. One of the alternatives to the WKB semiclassical formalism, the so-called extreme tunneling trajectory, or instanton [7–9], could be very effective in calculating a

¶This article was submitted by the authors in English.

globally uniform wave function of the ground state (i.e., a wave function without singularities). It allows finding semiclassical wave functions for a very broad class of potentials with arbitrary combinations of the first- and the second-order turning points. The method was recently adapted for the description of low-energy excited states [10, 11]. One of the main advantages of the instanton approach is that it can be readily extended to multidimensional systems using perturbative techniques (see [12] and references therein).

But before investigating multidimensional problems, we must study one-dimensional potentials and one-dimensional problems that cannot be accurately solved by the standard WKB method. These problems are the subject of this paper. The generalization of the instanton procedure to highly excited states is not straightforward at all and requires additional analysis. We consider only few relatively simple examples, but this analysis is useful for gaining insight into more complex systems for which even approximate theoretical results are not available.

In many interesting physical problems, high-accuracy calculations are out of reach of the standard WKB method, but as we see in what follows, the instanton approach offers a solution to the difficulties inherent to the WKB procedure. Since this fact was largely unnoticed in the previous studies, we found it worthwhile to present the investigation of several simple examples in a short and explicit form and to point out practical usability of the instanton approach. Apart from the aim to illustrate the efficiency of the instanton approach, our study is a prerequisite for an explanation and successful description of many relevant physical phenomena (for example, low-temperature quantum kinetics of phase transitions—see, e.g., [13]) where an active (reaction) path is effectively confined to one dimension.

All examples considered in this paper are related to the fundamental problems of chemical dynamics and molecular spectroscopy (see, e.g., [9] and references therein). Symmetric or slightly asymmetric double-well potentials are characteristic of molecules and van der Waals complexes with more than one stable configuration [14–17]. The states of such systems close to the barrier top (theoretically described by the instanton approach in this paper) are most relevant for radiationless evolution of highly excited states. These states have a double (localized–delocalized) nature, which manifests itself in the wave functions that simultaneously contain both components, the localized component in one of the wells and the delocalized component between the two wells. The states close to the barrier top of decay potentials govern thermally activated overbarrier transition amplitudes. For low-energy states, the main reduction factor is the tunneling exponent, while the contribution of highly excited states is limited by the Boltzmann factor. Our instanton calculations demonstrate that there is no sharp boundary between quasistationary and delocalized states. Two of us (V. B. and E. K.) recently investigated [18] the eigenstates of a highly asymmetric double-well potential. We have shown that quantum irreversibility phenomena occur when the spacing between neighboring levels of the deeper well becomes smaller than the typical transition matrix element. Obviously, this criterion can also be applied to the states near the barrier top. We note that, for low-energy states, the asymmetry leading to irreversible behavior must be very large, whereas, for states near the barrier top, the condition of the ergodic behavior is not very severe, and it is sufficient for the asymmetry of the potential to be comparable to the barrier height.

This paper is organized as follows. Section 2 contains the basic equations of the instanton method that are necessary for our investigation. As an illustration of the method, we consider a touchstone quantum mechanical problem—penetration of a particle through a parabolic potential barrier. The instanton solutions that are the asymptotic forms of the Weber equation are then exact. Section 3 is devoted to the investigation of highly excited states in a double-well potential. For concreteness and simplicity, we study the quartic anharmonic potential X^4. The instanton approach allows us to accurately reproduce not only the asymptotic behavior but also the crossover region from the single-well to the double-well quantization. In Section 4, a similar problem for the X^3 anharmonic potential is studied. Section 5 is devoted to the so-called resonance tunneling phenomena, which are not only interesting in their own right but also play a relevant role in many semiconducting double-barrier structures. In Section 6, we discuss the results. In the Appendix, we compute the so-called connection matrices that provide a very efficient method of finding semiclassical solutions to the Schrödinger equation in potentials having several turning points. Knowing the connection matrices is also important and significant for developing a good analytical approximation. The readers not interested in the mathematical derivation can skip the Appendix and find all the results in the main body of the paper.

2. PENETRATION THROUGH THE PARABOLIC POTENTIAL BARRIER

2.1. Instanton Approach

For convenience, we recall the main ideas of the instanton approach. The first step of the approach in [7, 8] is the so-called Wick rotation of phase space corresponding to the transformation to imaginary time $t \longrightarrow -it$. The potential and the kinetic energy change their signs after the transformation, and the Lagrangian is replaced by the Hamiltonian in the classical equation of motion. By this Wick rotation, the standard oscillating WKB wave functions are transformed into exponentially decaying functions that vanish as $X \longrightarrow \pm\infty$. Following [10, 11], we use a slightly different formulation of the instanton method, assuming exponentially

decaying real-valued wave functions from the very beginning. Taking into account that the wave functions of bound states can be chosen as real quantities, we can therefore seek a solution to the Schrödinger equation in the form

$$\Psi = \exp(-\gamma\sigma(X)), \tag{2.1}$$

where γ is the semiclassical parameter that is assumed to be sufficiently large ($\gamma \equiv m\Omega_0 a_0^2/\hbar$, where m is the mass of a particle; a_0 is a characteristic length in the problem, e.g., the tunneling distance; and Ω_0 is a characteristic frequency, e.g., the oscillation frequency around the potential minimum; in what follows, we set $\hbar = 1$, measuring energies in the units of frequency) and σ can be called the action; this function must satisfy the first-order differential equation of the Ricatti type,

$$\gamma^2\left[-\frac{1}{2}\left(\frac{d\sigma}{dX}\right)^2 + V(X)\right] + \gamma\left[\frac{1}{2}\frac{d^2\sigma}{dX^2} - \epsilon\right] = 0, \tag{2.2}$$

where $V(X)$ is the potential and ϵ gives particle eigenstates (energies). Here and in what follows, we use dimensionless variables ($\epsilon = E/\Omega_0$ for the energy, $V = U/\gamma\Omega_0$ for the potential, and $X = x/a_0$ for the coordinate, where E and U are the corresponding dimensional values of the energy and potential). We believe that $\gamma \gg 1$, and $\sigma(X)$ can therefore be expanded in the asymptotic series

$$\sigma(X) = W(X) + \gamma^{-1}W_1(X) + \gamma^{-2}W_2(X) + \ldots \tag{2.3}$$

The first- and the second-order terms in γ^{-1} become identically zero if the time-independent Hamilton–Jacobi equation (HJE) and the so-called transport equation (TE) are satisfied,

$$\frac{1}{2}\left(\frac{dW}{dX}\right)^2 = V(X) \tag{2.4}$$

and

$$\frac{dW}{dX}\frac{dA}{dX} + \frac{1}{2}\frac{d^2W}{dX^2}A = \epsilon A, \tag{2.5}$$

where

$$A(X) \equiv \exp(-W_1(X)). \tag{2.6}$$

An essential advantage of the instanton method in comparison to the standard WKB is that, in the former approach, the HJE is solved at $E = 0$, and the classically allowed regions therefore disappear. The price to be paid for this is the appearance of second-order turning points (in contrast to the WKB method, where all turning points are linear).

It is well known that the WKB wave functions are singular at the turning points, and therefore, different approximations represent the same wave function in different domains. The famous Stokes phenomenon [2] is related to the distribution of the turning points; Stokes and anti-Stokes lines emanate from each turning point. By definition, Stokes lines are the lines where the dominance of the dominant exponential semiclassical solution to the Schrödinger equation becomes strongest, and anti-Stokes lines are the lines on which the dominance and subdominance of the solutions interchange. Evidently, the WKB approximation does not work and must be refined near Stokes and anti-Stokes lines [2]. On the contrary, because classically accessible regions do not exist in the instanton formalism, the Stokes lines continuously pass through second-order turning points, and globally uniform real solutions to the Schrödinger equation can be constructed using the asymptotically smooth transformation of the instanton wave functions into the Weber functions. This global uniformity is the principal advantage of the instanton method.

A clearer idea of the instanton approach is obtained by the derivation of the well-known [1] quantization rules for the harmonic oscillator ($V(X) = X^2/2$). For a given energy ϵ, any solution to the Schrödinger equation can be represented as a linear combination of the solutions to the Weber equation [19]

$$\frac{d^2\Psi}{dz^2} + \left(\nu + \frac{1}{2} - \frac{z^2}{4}\right)\Psi(z) = 0, \tag{2.7}$$

where $z \equiv X\sqrt{\gamma}$ and $\nu = \epsilon - 1/2$. The basic solutions of (2.7) are the parabolic cylinder functions [19], and only the function $D_\nu(-z)$ vanishes as $z \longrightarrow \infty$ for $\arg z = 0$. For $\arg z = \pi$, the asymptotic behavior of this function as $z = \infty$ is given by [19]

$$D_\nu(-z) = \exp(i\pi\nu)z^\nu\exp\left(-\frac{z^2}{4}\right) - \frac{\sqrt{2\pi}}{\Gamma(-\nu)}z^{-\nu-1}\exp\left(\frac{z^2}{4}\right). \tag{2.8}$$

It can vanish as $z \longrightarrow \infty$ only at the poles of $\Gamma(-\nu)$, and this vanishing condition gives the exact eigenvalues of the harmonic oscillator

$$\epsilon = n + \frac{1}{2}.$$

Moreover, because $D_\nu(-z)$ coincide with the known harmonic oscillator eigenfunctions for positive integer ν [1], the instanton approach to the harmonic oscillator is exact.

2.2. Tunneling Through the Harmonic Barrier

As a less trivial illustration of the instanton approach efficiency, we apply the method to the problem of quantum mechanical tunneling through the parabolic potential

$$U(x) = U_0 - \frac{m\Omega_0^2}{2}x^2, \tag{2.9}$$

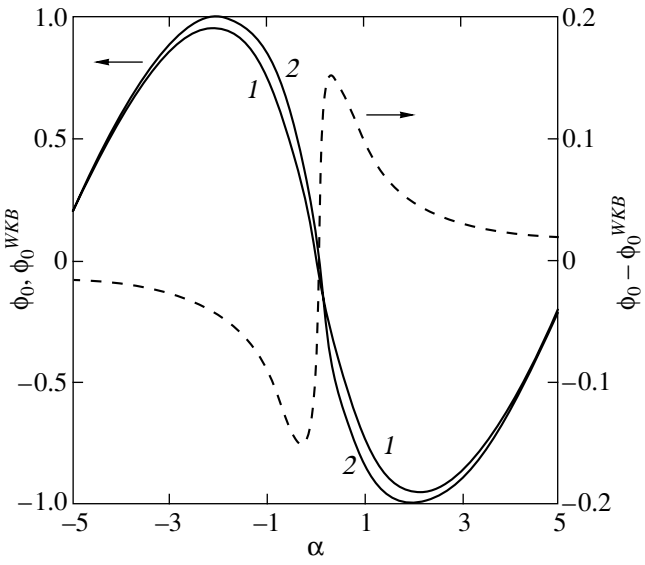


Fig. 1. The phase of the wave reflected from the parabolic barrier: (*1*) the exact quantum and instanton solutions ϕ_0; (*2*) the WKB solution ϕ_0^{WKB}; the dashed line is the difference $\phi_0 - \phi_0^{WKB}$.

where m is the mass of the tunneling particle and Ω_0 is a characteristic frequency (the curvature of the potential). The potential involves an additional characteristic space scale a_0. Using Ω_0 and a_0 to set the corresponding scales, we can rewrite parabolic potential (2.9) in the dimensionless form

$$V(X) = V_0 - \frac{1}{2}X^2. \tag{2.10}$$

In these variables, the Schrödinger equation is given by

$$\frac{d^2\Psi}{dX^2} + (\gamma^2X^2 - \alpha\gamma)\Psi(X) = 0, \tag{2.11}$$

where

$$\alpha = 2\frac{U_0 - E}{\Omega_0}, \tag{2.12}$$

and $\gamma \gg 1$ semiclassical parameter introduced above.

Schrödinger equation (2.11) can be transformed into the Weber equation [19] by a $\pi/4$ rotation in the complex plane,

$$X = \frac{1}{\sqrt{2}\gamma}z\exp\left(\frac{i\pi}{4}\right),$$

and the solution to (2.11) can therefore be represented as a linear combination of the parabolic cylinder functions D_ν [19],

$$\Psi_\nu(z) = c_1D_\nu(z) + c_2D_\nu(-z), \tag{2.13}$$

where $\nu = -1/2 - i\alpha/2$.

As $X \longrightarrow \infty$, only the transmitted wave exists with the amplitude (the transmission coefficient) T,

$$\Psi \approx T\exp\frac{i\gamma X^2}{2}. \tag{2.14}$$

As $X \longrightarrow -\infty$, both the incident wave ($\propto \exp(-i\gamma X^2/2)$) and the reflected wave proportional to $R\exp(i\gamma X^2/2)$ exist. By a standard quantum mechanical procedure [1], the transmission coefficient T and the reflection coefficient R can be found using the known asymptotic form of the parabolic cylinder functions [19] at fixed energy (i.e., at fixed α). This leads to the well known expression [1]

$$|T|^2 = \frac{1}{1 + \exp(\pi\alpha)}. \tag{2.15}$$

We note that solutions (2.13) are the exact solutions to the Schrödinger equation in parabolic potential (2.10). We now apply the instanton approach described above to the same problem. The solutions to HJE (2.4) and TE (2.5), which are milestones of the method, can easily be found as

$$\begin{aligned} W &= \pm i\frac{X^2}{2}, \\ A &= A_0X^{-1/2}\exp\left(\pm i\alpha\ln\frac{X}{2}\right), \end{aligned} \tag{2.16}$$

where the integration constant A_0 determines energy-dependent phases of the wave functions. Comparing (2.16) and (2.13), we can see that the instanton wave functions are the asymptotic forms of the parabolic cylinder functions, and therefore, because the transmission (T) and reflection (R) coefficients are determined only by the asymptotic behavior, the values of T and R found in the framework of the instanton approach coincide with the exact quantum mechanical ones at any value of the energy (of the parameter α). We recall that the instanton and the exact quantum mechanical solutions for the harmonic oscillator also coincide for any energy.

To finish this subsection, we mention for the skeptical reader that the WKB wave functions coincide with the exact solutions only at $\alpha \ll -1$. In the region where $|\alpha| \leq 1$, i.e., where the characteristic size of the forbidden region becomes comparable to the particle wavelength, specific interference phenomena between the transmitted and reflected waves occur, and phenomena of this kind cannot be reproduced in the standard WKB approach assuming that all turning points are independent.

As an illustration, in Fig. 1, we show the energy (α) dependence of the phase for the wave function reflected by the parabolic potential. The exact quantum mechanical and the instanton solutions (ϕ_0 in Fig. 1) are indistinguishable over a broad region of energies, while the

WKB solution (ϕ_0^{WKB} in Fig. 1) deviates from both of them.

2.3. Connection Matrices

Our analysis can be recast into a more elegant form by introducing the so-called connection matrices. In the instanton approach (as in any semiclassical treatment of the scattering or transition processes), we deal with only the asymptotic solutions and their connections on the complex coordinate plane. It is therefore important to know the connection matrices. These connection matrices provide a very efficient method of finding semiclassical solutions to the Schrödinger equation in potentials with several turning points. This is also a relevant starting point for developing a good analytical approximation.

It is convenient to formulate the general procedure for calculating the connection matrices for an arbitrary combination of the first- and second-order turning points. The procedure can then be applied to any particular problem under investigation. Technically, this requires extending the procedure known for linear turning points [2]. All the necessary details of the generalization are given in the Appendix, and we present here only the main definitions and results. In the semiclassical limit $\gamma \gg 1$, the Stokes and anti-Stokes lines for the equation

$$\frac{d^2\Psi}{dz^2} + \gamma^2 q(z)\Psi(z) = 0 \qquad (2.17)$$

are determined by the respective conditions

$$\mathrm{Re}\, W(z) = 0 \qquad (2.18)$$

and

$$\mathrm{Im}\, W(z) = 0, \qquad (2.19)$$

where

$$W(z) = \int_{z_0}^{z} \sqrt{q(z)}\, dz \qquad (2.20)$$

and z_0 is the turning point under consideration.

For the harmonic potential, there are only linear turning points for real ($\alpha > 0$) and imaginary ($\alpha < 0$) energies. In the Appendix, we calculate all the connection matrices that we need. To fully analyze the problem for the entire range of parameters, we must therefore know only the distributions of turning points and the Stokes and anti-Stokes lines on the complex plane. At real turning points ($\alpha > 0$) where $X_{1,2} = \pm(\alpha/\gamma)^{1/2}$, there are four Stokes and four anti-Stokes lines and two cuts in the complex plane (see Fig. 2).

For $\alpha \gg 1$, the connection matrix can be easily calculated as the direct product of the connection matrices

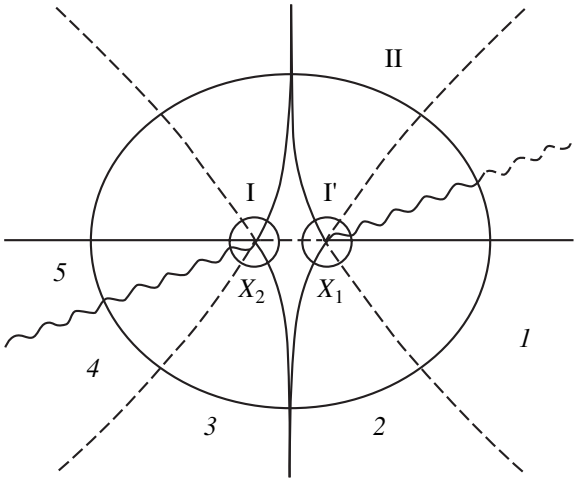


Fig. 2. The Stokes (solid) and anti-Stokes (dashed) lines for the two real-valued turning points $X_{1,2}$ with the surrounding contours *I* and *I'*. On contour *II*, the Stokes lines for the Airy equation asymptotically matches the lines for the Weber equation. The cut is depicted by the wavy line.

found in the Appendix ($\hat{M}^-$ in Eq. (A.4) and the Hermitian conjugate matrix $\hat{M}^+$) and the diagonal shift matrix

$$\begin{pmatrix} \exp(\pi\alpha/2) & 0 \\ 0 & \exp(-\pi\alpha/2) \end{pmatrix}. \qquad (2.21)$$

This leads to the transmission coefficient $T \approx \exp(-\pi\alpha/2)$, which coincides with (2.15) with exponential accuracy in the limit $\alpha \gg 1$. To improve the accuracy at smaller values of α, we must calculate the connection matrices taking into account not only the contributions of the contours encompassing the turning points, but also the additional contribution to the action of the closed path (with a radius $\gg |X_{1,2}|$) encompassing both points X_1 and X_2 (see Fig. 2). The procedure changes the Stokes constant T_3 (on the dashed line separating regions *3* and *4* in Fig. 2), which becomes

$$|T_3| = [1 + \exp(-\pi\alpha)]^{1/2}.$$

This finally leads to the correct transmission coefficient

$$T = iT_3^{-1}\exp\left(-\frac{\pi\alpha}{2}\right),$$

which is identical to (2.15).

In the case where $\alpha < 0$, the entire picture (see Fig. 3) of the Stokes and of the anti-Stokes lines and turning points is rotated by the angle $\pi/2$ with respect to the picture in Fig. 2. If we bluntly take the point $X = 0$ as the low integration limit for the action W^* in Eq. (A.11), we obtain the transmission coefficient

$$T = 1 - \frac{1}{2}\exp(-\pi|\alpha|),$$

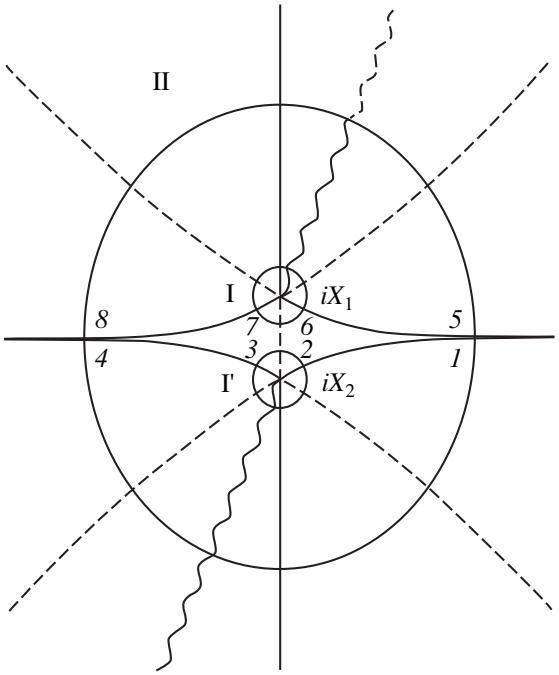


Fig. 3. The same as Fig. 2 for the case of two purely imaginary turning points $iX_{1,2}$.

which can be reliable (with the accuracy $\exp(-2\pi|\alpha|)$) only for $|\alpha| \gg 1$. As in the case where $\alpha > 0$, the accuracy can be improved by taking into account the contribution of the path surrounding both imaginary turning points (this fact was noticed by Pokrovskii and Khalatnikov long ago [3]).

At the isolated linear imaginary turning point iX_1, the connection matrix is found from (A.4),

$$\tilde{M}_1^+ = \begin{pmatrix} 1 & i\exp(-\pi|\alpha|/2) \\ 0 & 1 \end{pmatrix}. \quad (2.22)$$

Similarly, the Hermitian conjugate matrix $\tilde{M}_1^-$ comes from the contribution of the closed path surrounding $-iX_1$. These contours provide only the amplitude of the dominant (exponentially increasing) wave. But the accuracy is insufficient for finding the amplitude of the corresponding subdominant solution (the exponentially decaying wave function), and we obtain the incorrect transmission coefficient $T = 1$. To improve the accuracy and to find T correctly, we must include the connection matrix for the isolated second-order turning point in the procedure (in this particular example, this turning point is the maximum of the potential). Using (A.9), we can explicitly find this matrix as

$$\tilde{M}_2 = \begin{pmatrix} [1+\exp(-\pi|\alpha|)]^{1/2} & i\exp(-\pi|\alpha|/2) \\ -i\exp(-\pi|\alpha|/2) & [1+\exp(-\pi|\alpha|)]^{1/2} \end{pmatrix}. \quad (2.23)$$

In principle, similar calculations can be performed in the adiabatic perturbation theory (which in fact employs the Planck constant smallness equivalent to $\gamma \gg 1$). We note that, in [20], the contributions of the contours surrounding turning points (analogous to those presented above) were taken into account. It seems very plausible that, following this way, it will be possible to combine the instanton approach and the adiabatic perturbation theory, but this issue is beyond the scope of this paper and will be discussed elsewhere.

3. HIGHLY EXCITED STATES IN THE DOUBLE-WELL POTENTIAL

Literally, the instanton approach described in the previous section is valid for states with characteristic energies that are small compared to the barrier height. But as we show in this section, the instanton method works sufficiently well for the energy states near the barrier top V_0. As an illustration, we consider the symmetric double-well potential (quartic anharmonic X^4 potential)

$$V_0 - V(X) = \frac{1}{2}X^2(1-X^2). \quad (3.1)$$

The Schrödinger equation with potential (3.1) can be rewritten in dimensionless variables in the form

$$\frac{d^2\Psi}{dX^2} + [2\gamma^2(V_0 - V(X)) - \alpha\gamma]\Psi(X) = 0, \quad (3.2)$$

which is most convenient in applications of the instanton approach. The HJE and TE then become

$$\frac{1}{2}\left(\frac{dW}{dX}\right)^2 = V_0 - V(X), \quad (3.3)$$

and

$$\frac{dW}{dX}\frac{dA}{dX} + \frac{1}{2}\left(\frac{d^2W}{dX^2} + i\alpha\right)A = 0. \quad (3.4)$$

Formal solutions to the set of equations (3.3), (3.4) are the even and odd instanton wave functions

$$\Psi_I^\pm = A_\pm(X)\exp(i\gamma W_\pm(X)), \quad (3.5)$$

where the action $W_\pm$ (a solution of the HJE) is to be determined from

$$\frac{dW_\pm}{dX} = \pm\sqrt{2(V_0 - V(X))}, \quad (3.6)$$

and the amplitude (prefactor) is given by

$$A_\pm = \left|\frac{dW_\pm}{dX}\right|^{-1/2}\exp\left[-i\alpha\int\left(\frac{dW_\pm}{dX}\right)^{-1}dX\right]. \quad (3.7)$$

The quantization rules [1] are related to continuous matching of the solutions at the turning points (the second-order turning point $X = 0$ and the linear turning

points $X = \pm 1$ for $\alpha > 0$ and $X = \pm i$ for $\alpha < 0$). A crucial advantage of instanton solutions (3.5) is that these functions have no singularities inside the barrier, because the corresponding exponents are pure imaginary in the classically accessible regions (unlike the WKB solutions). In addition, the general form of the instanton wave functions does not noticeably depend on whether $E < V_0$ or $E > V_0$. This advantage allows us to include the instanton wave functions in the basis of globally uniform functions diagonalizing the Hamiltonian even for highly excited states.

The above general procedure for searching instanton solutions to the Schrödinger equation with model potential (3.1) has a subtle point, which motivates giving the explicit searching procedure in some detail; new results follow from our investigation. The procedure includes several steps.

1. Near the second order turning point, exact solution (2.14) to the Schrödinger equation can be used with $c_1 = \pm c_2$ for the even and odd solutions respectively. For $|X| \gg 1$, it follows from (2.14) and from the known asymptotic forms of the parabolic cylinder functions [19] that

$$\Psi(X) = \frac{c_1}{\sqrt{X}}\left[\frac{\exp(if(X))}{\Gamma((1-i\alpha)/4)} + \frac{\exp(-if(X))}{\Gamma((1+i\alpha)/4)}\right], \quad (3.8)$$

where

$$c_1 = -\frac{2\pi}{\Gamma((3+i\alpha)/4)}\exp\left(-\frac{\pi\alpha}{8}\right)2^{-i\alpha/4}(2\gamma)^{-1/4} \quad (3.9)$$

and

$$f(X) = \frac{\gamma}{2}X^2 - \frac{\alpha}{2}\ln X - \frac{\alpha}{4}\ln\gamma - \frac{\pi}{8}. \quad (3.10)$$

To obtain the correct even and odd linear combinations conforming to (3.5), we set

$$c_\pm = c_1\frac{\exp(\pm if_1)}{\Gamma((1 \pm i\alpha/\sqrt{2})/4)}, \quad (3.11)$$

where

$$f_1 = \frac{\alpha\ln\gamma}{4} + \frac{\pi}{8}.$$

2. Near the linear turning point $X = \pm 1$, the Schrödinger equation reduces to the Airy equation [19]

$$\frac{d^2\Psi}{dy^2} - y\Psi(y) = 0, \quad (3.12)$$

where

$$y = \gamma^{2/3}\left|X + 1 + \frac{\alpha}{\gamma}\right| \quad (3.13)$$

for $X < 0$. The solution that vanishes as $y \longrightarrow \infty$ is given by [19]

$$\Psi(y) = |y|^{-1/4}\sin\left(\frac{2}{3}|y|^{3/2} + \frac{\pi}{4}\right). \quad (3.14)$$

Continuing this solution to the regions $(X \pm 1)\sqrt{2\gamma} \gg 1$ and sewing there with (3.8), we obtain

$$\frac{c_+}{c_-} = \exp\left(-i2\gamma W^* + i\frac{3\pi}{2}\right), \quad (3.15)$$

where W^* is the energy-dependent action on the interval $[X = 0, X = 1]$.

3. Comparing (3.15) and (3.11), we find the quantization rules

$$\frac{\Gamma((1+i\alpha)/4)}{\Gamma((1-i\alpha)/4)} = \exp\left(-2i\gamma W^* - i\frac{3\pi}{2}\right) \quad (3.16)$$

for the even states and

$$\frac{\Gamma((3+i\alpha)/4)}{\Gamma((3-i\alpha)/4)} = \exp(-2i\gamma W^* - i\pi) \quad (3.17)$$

for the odd states.

4. From (3.16) and (3.17), we finally obtain the quantization rule that can be written in the single form for both the even and the odd states as

$$2\gamma W^* + 2\phi(\alpha) \equiv \begin{pmatrix} \frac{5\pi}{4} + 2\pi n - \arctan\left(\tanh\left(\frac{\pi\alpha}{4}\right)\right) \\ \frac{3\pi}{4} + 2\pi n - 2\arctan\left(\tanh\left(\frac{\pi\alpha}{4}\right)\right) \end{pmatrix}. \quad (3.18)$$

Relation (3.18) is the sought quantization rule that now allows us to use the advantages of the instanton method. For highly excited states (i.e., for $\alpha \ll -1$), it follows from (3.18) that

$$2\gamma W^* + 2\phi(\alpha) = \pi\left(n + \frac{1}{2}\right),$$

where n is an integer. For low-energy states ($\alpha \gg 1$), Eq. (3.18) reproduces the known quantization rule

$$\gamma W_L^* = \pi\left(n + \frac{1}{2}\right) \pm \frac{1}{2}\exp\left(-\frac{\pi\alpha}{2}\right),$$

where W_L^* is the action in the classically admissible region between the linear turning point in the left well,

$$\gamma W_L^* = \gamma W^* + \phi(\alpha). \quad (3.19)$$

We note an essential advantage of instanton quantization rule (3.18) compared to the traditional WKB formalism, where the quantization rules are totally different [1] in the tunneling and overbarrier regions. The

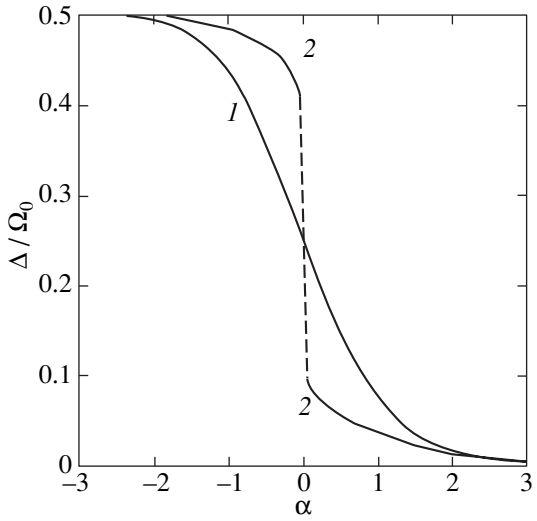


Fig. 4. The dimensionless tunneling splitting Δ/Ω_0 for the anharmonic X^4 potential near the barrier top: (*1*) exact quantum and instanton calculations; (*2*) the WKB result.

instanton approach gives a single quantization rule, Eq. (3.18), which is valid in both regions and, in addition, quite accurately describes the crossover behavior near the barrier top, where periodic orbits localized at separate wells transform into a common figure-eight orbit enclosing both wells.

We illustrate the results of this section in Fig. 4, where we plot the universal dependence of the eigenvalues in symmetric double-well potential (3.1) on α. For comparison, we show in the same figure the eigenvalues found by the conventional WKB procedure and by the exact quantum mechanical computation. It is clear from the figure that the WKB method errors are maximal in the region of small $|\alpha|$, because the oscillation period logarithmically diverges in this region (the particle spends infinitely long time near the second-order turning point). On the contrary, the errors of the instanton approach are minimal near the barrier top (small $|\alpha|$).

As mentioned at the beginning of this section, the instanton approach is also very accurate near the potential minimum. Generally speaking, the instanton solutions are always correct when the deviation from the corresponding extremum is of the order of characteristic zero-point amplitudes. Mathematically, the accuracy of the instanton approach is based on the transformation of the semiclassical solutions into the harmonic oscillator eigenfunctions (which also ensures the correct normalization of the instanton wave functions). It is therefore natural to expect that the instanton method is very accurate near the barrier top and near the potential minimum. On the contrary, in the intermediate region, where the anharmonic shape of the potential is relevant, we should expect poor accuracy of the instanton method. Fortunately, it turns out that the mathematical nature of the problem is on our side, and the instanton approach has a reasonable accuracy (of the order of the accuracy of the WKB method) even in this region. The fact is that the instanton wave functions are exact not only in zero but also in the first order with respect to anharmonic corrections to the potential approximation. This can be shown using the anharmonic perturbative procedure that was proposed by Avilov and Iordanskii for the WKB functions [21] and was generalized for the instanton wave functions in [22].

For practical computations, it is also relevant that the instanton wave functions (unlike the WKB ones) are continuous near their "own" minimum. Numerical estimate shows that, in the intermediate energy region, the instanton wave functions reproduce exact quantum results with the accuracy about 5–10%.

To finish the section, we present the connection matrices needed to find semiclassical solutions to the Schrödinger equation in the double-well potential. Similarly to the results in Section 2, the connection matrix for the instanton solutions is the product of connection matrices (A.4) for the linear turning points and the connection matrix for the second-order turning point, which is the maximum of the double-well potential in the case under consideration. Using (A.9), we can find this latter connection matrix as

$$\begin{pmatrix} 2\left[\exp\left(\frac{\pi\alpha}{2}\right) + (1+\exp(\pi\alpha))^{1/2}\cos(2\gamma W^*)\right] & -(1+\exp(\pi\alpha))^{1/2}\sin(2\gamma W^*) \\ (1+\exp(\pi\alpha))^{1/2}\sin(2\gamma W^*) & \frac{1}{2}\left[-\exp\left(\frac{\pi\alpha}{2}\right) + (1+\exp(\pi\alpha))^{1/2}\cos(2\gamma W^*)\right] \end{pmatrix}. \quad (3.20)$$

It is worth noting that the reflected wave acquires a nontrivial phase factor near the barrier top. This phenomenon is related to the interference of the incident, reflected, and transmitted waves, and the phase therefore has some geometrical meaning, similarly to the famous Berry phase [23]. The geometrical origin of the phase manifests itself more clearly if we recall that the semiclassical phase factor is determined by the probability density flow through the barrier,

$$J = i\Psi^*\frac{d\Psi}{dX}.$$

We can consider this phase factor from a slightly different standpoint, because tunneling results in a phase shift related to the change of the eigenvalues. Quantization rules (3.18) and (3.19) can be rewritten as

$$\epsilon_n = n + \frac{1}{2} + \chi_n,$$

which is the definition of the eigenvalues ϵ_n and where n is an integer labeling the eigenvalues and χ_n is determined by the exponentially small phase shift due to the existence of the barrier between the two wells. The phase shift χ_n has the same functional form (and physical meaning) as the geometrical phase factor (appearing because of the interference phenomena) acquired by a quantum mechanical wave function upon a cyclic evolution [23–25].

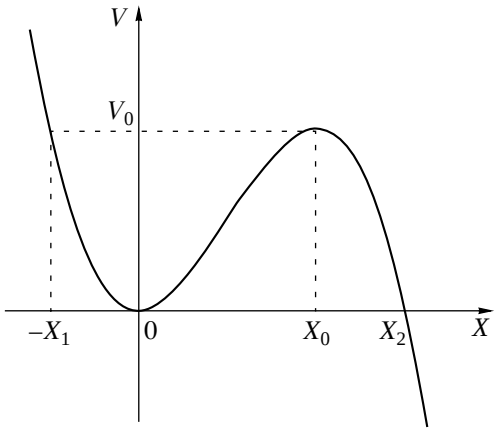


Fig. 5. The X^3 anharmonic decay potential.

4. THE DECAY POTENTIAL

In this section, we study highly excited states in a decay potential, which we choose as the anharmonic X^3 potential for definiteness,

$$V(X) = \frac{1}{2}X^2(1 - X). \tag{4.1}$$

As a first (but compulsory) step, we investigate the low-lying tunneling states.

4.1. Tunneling Decay of Metastable States

We start from this simple case to pick low-hanging fruit first, i.e., to describe the states under the conditions

$$V_0 \gg \epsilon_n \gg V(X \longrightarrow \infty), \tag{4.2}$$

which imply that a local minimum is separated from the continuum spectrum by a high energetical barrier, and the quasistationary states ϵ_n are therefore characterized by good quantum numbers n. We note that a generic decay potential shown in Fig. 5 is determined by the positions of the barrier top X_0 and the three turning points $-X_1$, $X = 0$, and $+X_2$; near these points, we have

$$V(X) = \begin{cases} V_0 - \left(\dfrac{dV}{dX}\right)_{X=-X_1}(X + X_1), & |X + X_1| \longrightarrow 0, \\ \dfrac{1}{2}X^2, & |X| \longrightarrow 0, \\ V_0 - \dfrac{1}{2}(X - X_0)^2, & |X - X_0| \longrightarrow 0, \\ -\left(\dfrac{dV}{dX}\right)_{X=X_2}(X - X_2), & |X - X_2| \longrightarrow 0. \end{cases} \tag{4.3}$$

Potential (4.1) is a particular example of the generic decay potential in Eq. (4.3) (with $X_1 = 1/3$, $X_0 = 2/3$, $X_2 = 1$, and $V_0 = 2/27$); we use it only as an explicit illustration, and, in order to be specific, while all the results given below are equally valid for the generic potential. As a note of caution, we also remark that, in the instanton approach to this problem, we must always deal with only two turning points. For low-energy states, the points are X_2 and the potential minimum $X = 0$, and for high-energy states, the points are $-X_1$ and the potential maximum X_0.

In accordance with (4.1), there are no turning points at $X > X_2$; at $X \gg X_2$, the potential can be considered as a constant, and therefore the wave functions must asymptotically coincide with plane waves for $X \gg X_2$. Furthermore, near the linear turning point $X = 1$, the Schrödinger equation with the X^3 anharmonic potential (4.1) is reduced to Airy equation (3.12), whose solutions are linear combinations of the Bessel functions with the indices $\pm 1/3$ at real (for $X < 1$) and imaginary (for $X > 1$) values of the arguments,

$$\Psi(u) = \sqrt{u}\left[B_+ I_{1/3}\left(\frac{2u^{3/2}}{3}\right) + B_- I_{-1/3}\left(\frac{2u^{3/2}}{3}\right)\right] \tag{4.4}$$

and

$$\Psi(\zeta) = \sqrt{\zeta}\left[-B_+ J_{1/3}\left(\frac{2\zeta^{3/2}}{3}\right) + B_- J_{-1/3}\left(\frac{2\zeta^{3/2}}{3}\right)\right], \tag{4.5}$$

where

$$u = (2\gamma)^{2/3}\left[1 - X - \frac{\nu + 1/2}{\gamma}\right] \tag{4.6}$$

for $X < 1$ and

$$\zeta = (2\gamma)^{2/3}\left[X - 1 + \frac{\nu + 1/2}{\gamma}\right] \tag{4.7}$$

for $X > 1$ (we recall that $\nu = \epsilon_n - 1/2$ here).

The coefficients $B_\pm$ must be chosen such that Eq. (4.5) gives plane waves for $\zeta \gg 1$; using the known

asymptotic forms of the Bessel functions [19], we thus obtain

$$B_+ = B_-\exp\left(-i\frac{\pi}{3}\right).$$

In the classically forbidden region $u \gg 1$, the instanton solutions to HJE (2.4) and TE (2.5) that continuously match the quantum mechanical solutions of the Schrödinger equation near the turning points are given by

$$\Psi_\pm = A_\pm\exp(\pm\gamma W), \tag{4.8}$$

where

$$A_\pm = X^{-1/2}(1-X)^{-1/4}\left(\frac{1-\sqrt{1-X}}{1+\sqrt{1-X}}\right)^{n\epsilon_n} \tag{4.9}$$

and

$$W = \frac{8}{15} - \frac{4}{3}(1-X)^{3/2} + \frac{4}{5}(1-X)^{5/2}. \tag{4.10}$$

The sought wave functions of the quasistationary states are linear combinations of instanton solutions (4.8), with the coefficients in the linear combinations determined from the condition of asymptotically matching the parabolic cylinder functions in Eq. (2.13) and the Airy functions. This leads to the following equation for complex eigenvalues ν:

$$-\frac{\sqrt{2\pi}}{\Gamma(-\nu)}\exp\left(\frac{16}{15}\gamma\right) = i\gamma^{\nu+(1/2)}2^{6\nu+3}. \tag{4.11}$$

Because the function $\Gamma(z)$ has a simple pole at $z = -n$, we can easily find the leading contribution to the decay rate Γ_n of the quasistationary state ϵ_n,

$$\frac{\Gamma_n}{\Omega_0} = \sqrt{\frac{2}{\pi}}\frac{\gamma^{\nu+1/2}2^{6\nu+3}}{n!}\exp\left(-\frac{16}{15}\gamma\right). \tag{4.12}$$

We note that, for the ground state, with $n = 0$, Eq. (4.12) coincides with the result found by Caldeira and Legget [26]. On the other hand, the decay rate is related to the current flow [1] as $X \longrightarrow +\infty$, providing the constant amplitude of the outgoing wave,

$$\frac{\Gamma_n}{\Omega_0} = (2i\sqrt{\gamma}\int|\Psi|^2dX)^{-1}\left(-\Psi^*\frac{d\Psi}{dX} + \Psi\frac{d\Psi^*}{dX}\right). \tag{4.13}$$

Inserting the explicit forms of the wave functions in Eqs. (4.5)–(4.7) in (4.13), we obtain

$$\frac{\Gamma_n}{\Omega_0} = \frac{9}{8}\frac{\gamma^{1/6}}{2^{1/3}}|B_+|^2. \tag{4.14}$$

In accordance with (4.14), the decay rate depends only on the normalization of the instanton wave function and on the amplitude of the outgoing wave. Both characteristics are determined essentially by the behavior of the instanton wave function in the vicinity of the turning points only. We note, however, that, in this approximation, the instanton computation of decay rate (4.13) or (4.14) is satisfactory only for the ground state, because corrections of the order γ^{-1} rapidly increase with the quantum number n. The method can be improved by taking the X^3 anharmonic contribution to the potential into account as a perturbation,

$$\frac{\Gamma_n}{\Omega_0} = \sqrt{\frac{2}{\pi}}\frac{\gamma^{\nu+1/2}2^{6\nu+3}}{n!}\exp\left(-\frac{16}{15}\gamma\right) \times\left[1-\frac{1}{576\gamma}(164n^3+246n^2+1216n+567)\right]. \tag{4.15}$$

The decay rate calculated in accordance with (4.15) is of the same accuracy level as the WKB and the exact quantum mechanical computations for $\gamma \geq 5$. Outside the regime of interest, the instanton theory loses all pretense of predictability.

4.2. Highly Excited States for the Anharmonic X^3 Potential

In Subsection 4.1, we calculated the decay rate of low-energy metastable states. In this case (where the states ϵ_n can be characterized by the good quantum number n), the period of oscillations in the well is smaller than the inverse decay rate ($\epsilon_n \gg \Gamma_n/\Omega_0$) and Γ_n is determined by the probability current density flowing from the well into the classically admissible region ($X > X_0$ for a given energy n—see Fig. 5) under the condition of the vanishing backflow from this region to the barrier. Evidently, the method does not work for highly excited states with $\Gamma_n \geq \epsilon_n\Omega_0$. In this section, we go one step further with respect to Section 4.1 in extending the instanton approach to the decay of highly excited states.

First, it is worth noting that the wave functions must vanish as $X \longrightarrow -\infty$ and, moreover, can always be chosen as real-valued quantities as $X \longrightarrow +\infty$. From these two conditions, we can find the relations between the instanton wave functions in the regions $X < -X_1$ and $X > X_0$ (see the notation in Fig. 5) and, consequently, calculate the phase $\delta(\alpha)$ (counted from the barrier top) of the standing wave in the region $X > X_0$. It is given by

$$\exp(i2\delta) = -i\exp(-i2\gamma W^*) \times\left\{1+\frac{\sqrt{2\pi}\exp(-\pi\alpha/4)\exp(i2\gamma W^*)}{\Gamma(1-i\alpha/2)}\right\} \times\left\{1+\frac{\sqrt{2\pi}\exp(-\pi\alpha/4)\exp(-i2\gamma W^*)}{\Gamma(1+i\alpha/2)}\right\}^{-1}. \tag{4.16}$$

According to the standard quantum mechanics [1], phase (4.16) determines the scattering amplitude. We can therefore find the scattering amplitude deep in the classically forbidden region from (4.16) and, hence, compute the eigenvalues in this region. For the calculation, we must know the terms of the order $\exp(-\pi|\alpha|)$ in

the expansion of the Γ functions (these terms are beyond the standard Stirling formula) [19],

$$\Gamma\left(\frac{1 \pm i\alpha}{2}\right) \approx \sqrt{2\pi}\exp\left(-\frac{\pi\alpha}{4} \pm i\phi\right) \times \left[1 - \frac{1}{2}\exp(-\pi|\alpha|)\right], \tag{4.17}$$

where

$$\phi(\alpha) \equiv \frac{\alpha}{2}\left[\ln\frac{|\alpha|}{2} - 1\right]. \tag{4.18}$$

Finally, taking Eqs. (4.17) and (4.18) into account, we find the poles of the scattering amplitude from (4.16) (with the required exponential accuracy) as

$$2\gamma W_L^* = 2\gamma W^* + \phi(\alpha) = \pi(2n+1) - i\left[\frac{\pi}{4}(|\alpha| - \alpha) + \frac{1}{2}\exp(-\pi|\alpha|)\right]. \tag{4.19}$$

Explicitly solving (4.19), we find the complex eigenvalues and, in particular, the decay rate for highly excited states in the anharmonic decay potential.

In the same way as for low-energy tunneling states, the real part of the eigenvalues ϵ_n for highly excited states (i.e., for $|\alpha| \gg 1$) is determined by the action along closed trajectories in the well, whereas the imaginary part (i.e., the decay rate Γ_n) is related to the probability current density flow from the well to the barrier.

Using the instanton approach procedure described in Sections 2 and 3 (see [10, 11] for the details), we can find not only the eigenvalues but also the eigenstates. The real-valued instanton wave functions are determined by the action $W(X_1, X)$, which is counted from the linear turning point X_1,

$$\Psi(X) = A(\alpha)|X - X_1|^{-1/4}\sin\left(\gamma W(X_1, X) + \frac{\pi}{4}\right), \tag{4.20}$$

where the amplitude $A(\alpha)$ of the wave function acquires maximum values at the poles of (4.16) with the widths proportional to Γ_n. We plot the functions $|A(\alpha)|^2$ in Fig. 6.

5. RESONANCE TUNNELING

The phenomenon of the electron resonance tunneling is familiar [27] and was observed (see, e.g., [28] and also [29] for more recent references) in semiconducting heterostructures possessing the so-called double-barrier potentials (see Fig. 7). This phenomenon manifests itself as peaks in the tunneling current at voltages near the quasistationary states of the potential well. The physical mechanism of the resonance tunneling can be understood as a constructive interference between the wave reflected from the left barrier and the wave outgoing to the left of the well.

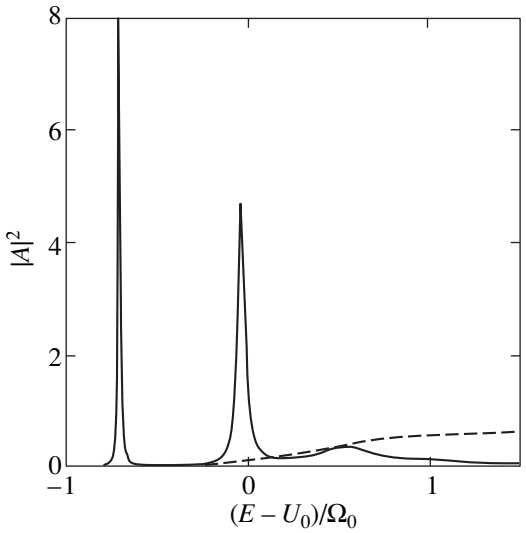


Fig. 6. The amplitude of the wave function localized in the potential shown in Fig. 5 (the dashed line is a nonresonant part of the amplitude and $\gamma = 101$).

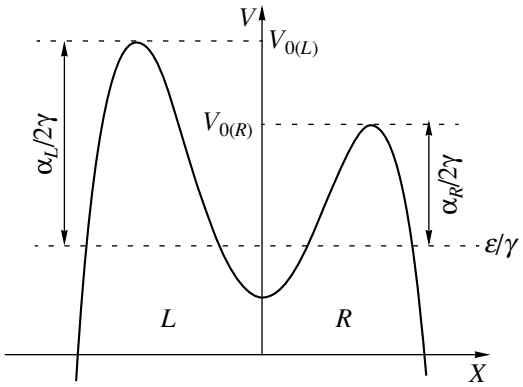


Fig. 7. The model two-barrier potential structure for the resonance tunneling.

In the instanton method, the total transmission coefficient T is determined by the second-order turning points of the double-barrier potential (i.e., the maxima of the potential); in accordance with the procedure described in the Section 2, T is given by

$$|T|^2 = \pi^2\Gamma_L\Gamma_R\{[1 - \sqrt{(1 + \pi\Gamma_L)(1 + \pi\Gamma_R)}]^2 + 4\sqrt{(1 + \pi\Gamma_L)(1 + \pi\Gamma_R)}\cos^2(\gamma W_R^*)\}^{-1}, \tag{5.1}$$

where we use the notation

$$\Gamma_{L,R} = \frac{1}{\pi}\exp(-\pi\alpha_{L,R}) \tag{5.2}$$

and, similarly to (2.12),

$$\alpha_{L,R} = 2\frac{U_{0(L,R)} - E}{\Omega_{0(L,R)}}. \tag{5.3}$$

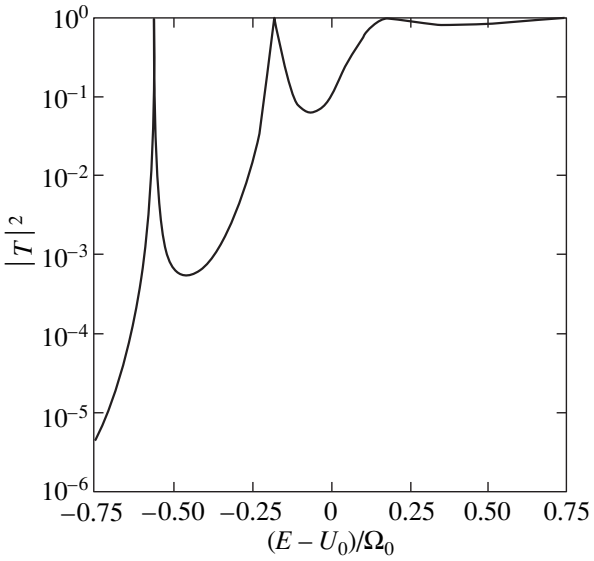


Fig. 8. The transmission coefficient for the potential shown in Fig. 7 ($\gamma = 54$).

Similarly to (3.19), the action in the classically admissible region is given by

$$\gamma W_R^* = \gamma W^* - \phi(\alpha_L) - \phi(\alpha_R). \tag{5.4}$$

In the resonance region, where

$$\gamma W_R^* = \pi\left(n + \frac{1}{2}\right)$$

in accordance with the stationary quantization rule, the transmission coefficient in (5.1) is given by

$$|T|^2 = \frac{4\Gamma_L\Gamma_R}{(\Gamma_L + \Gamma_R)^2}. \tag{5.5}$$

Far from the resonance, it is given by

$$|T|^2 = \frac{\pi^2\Gamma_L\Gamma_R}{4\cos^2(\gamma W_R^*)}. \tag{5.6}$$

We thus found the resonance amplification of the transmission. For the symmetric case at the resonance, $T = 1$, and the interference completely suppresses the reflection. In the opposite case of strongly asymmetric barriers, T in (5.1) is almost coincident with the transmission coefficient for the highest barrier, and the influence of the lower barrier is suppressed by the interference. In Fig. 8, we show the energy dependence of T for the symmetric structure of the barriers. The resonances become broader as the energy approaches the top of the potential barriers and disappear at higher energies (above the top). It is worthwhile to stress that the instanton solution of the resonance tunneling problem allows us to study the phenomenon in a very broad energy region, including the states near the tops of barriers.

We finally present the connection matrices for the instanton solutions found above. The corresponding matrix can be found as the product of two connection matrices connecting instanton solutions near the second-order turning points (see (A.9) and (3.20)) and diagonal shift matrix (A.6),

$$\begin{pmatrix} 2\pi\exp\left(\pi\frac{\alpha_L+\alpha_R}{4}\right)\exp(i\gamma W^*)\left[\Gamma\left(\frac{1+i\alpha_L}{2}\right)\Gamma\left(\frac{1+i\alpha_R}{2}\right)\right]^{-1} + \exp\left(\pi\frac{\alpha_L+\alpha_R}{2}\right)\exp(-i\gamma W^*) \\ -i\sqrt{2\pi}\exp\left(3\pi\frac{\alpha_R+\alpha_L}{8}\right)\left[\exp(i\gamma W^*)\Gamma^{-1}\left(\frac{1+i\alpha_R}{2}\right) + \exp(-i\gamma W^*)\Gamma^{-1}\left(\frac{1-i\alpha_L}{2}\right)\right] \\ i\sqrt{2\pi}\exp\left(3\pi\frac{\alpha_R+\alpha_L}{8}\right)\left[\exp(i\gamma W^*)\Gamma^{-1}\left(\frac{1+i\alpha_L}{2}\right) + \exp(-i\gamma W^*)\Gamma^{-1}\left(\frac{1-i\alpha_R}{2}\right)\right] \\ 2\pi\exp\left(\pi\frac{\alpha_L+\alpha_R}{4}\right)\exp(-i\gamma W^*)\left[\Gamma\left(\frac{1+i\alpha_L}{2}\right)\Gamma\left(\frac{1+i\alpha_R}{2}\right)\right]^{-1} + \exp\left(\pi\frac{\alpha_L+\alpha_R}{2}\right)\exp(i\gamma W^*) \end{pmatrix}. \tag{5.7}$$

Here, W^* denotes the action between the turning points (in this case, between the second-order turning points).

6. CONCLUSIONS

This paper could be considered as a formal one, in the sense that we asked theoretical questions that most solid-state or chemical physics experimentalists would not think to ask. But answering these very basic questions can be illuminating.

We first summarize the results of the paper. Within the framework of the instanton approach, we derived accurate analytical solutions for a number of one-dimensional semiclassical problems and checked the results numerically. As an illustration of the method, we considered a simple quantum mechanical problem, penetration of a particle through the parabolic potential

barrier. In this case, the instanton solutions (which are asymptotic solutions of the Weber equation) are exact. We also considered the description of highly excited states in a double-well potential. For definiteness and simplicity, we studied the quartic anharmonic X^4 potential. The instanton approach enables us to accurately reproduce not only the asymptotic behavior but also the crossover region from the single-well to the double-well quantization (in contrast to the standard WKB approach, which fails to describe the crossover behavior). A similar problem for the X^3 anharmonic potential was also studied, and the instanton method allowed us to study the resonance broadening and collapse phenomena. In addition, we investigated the so-called resonance tunneling phenomena, not only interesting in their own right but also playing a relevant role in many semiconducting double-barrier structures. We also computed the connection matrices that provide a very efficient method of finding semiclassical solutions to the Schrödinger equation in potentials with several turning points (they are also useful in developing a good analytical approximation).

All the examples selected to illustrate the efficiency of the instanton approach belong to the fundamental problems in chemical dynamics and molecular spectroscopy (see, e.g., [9] and references therein). Symmetric or slightly asymmetric double-well potentials are characteristic of molecules and van der Waals complexes with more than one stable configurations [14–17]. The states of such systems that are close to the barrier top (theoretically described by the instanton approach in our paper) are not easy to investigate experimentally, because optical transitions between these states and the localized ones are typically inactive. But precisely these states are most relevant for radiationless evolution of highly excited states. In a certain sense, these states have a double (localized–delocalized) nature that manifests itself in the form of wave functions that simultaneously contain both components, the one localized in one of the wells and the other delocalized between the two wells. Consequently, any initially prepared localized state evolves via formation and decay of these states. Our calculations are intended to pave the way to the investigation of this class of problems using the wave functions computed within the instanton approach.

The states that are close to the barrier top of decay potentials govern thermally activated overbarrier transition amplitudes. For low-energy states, the main reduction factor is the tunneling exponent, while the contribution of highly excited states is limited by the Boltzmann factor. The energy width of the region dominating in the total transition rate is traditionally postulated in the transition rate theory [30] to be of the order of the temperature T. But our results in Section 3 predict another estimate. Instanton calculations demonstrate that the intermediate region between the quasistationary ($\Gamma \ll \Omega$) and the delocalized states could be much larger than T, namely, of the order Ω. This implies that no sharp boundary exists between the quasistationary and the delocalized states and that all of the states within the interval $V^* - \Omega$, $V^* + T$ equally contribute to the total rate constant for the penetration through the barrier.

One more point should be emphasized. It was recently shown in [18] that quantum irreversibility phenomena occur when the spacing between neighboring levels of the deeper well becomes smaller than the typical transition matrix element. Obviously, this criterion can also be applied to the states near the barrier top. We note that, for low-energy states, the asymmetry providing the irreversible behavior must be very large, whereas, for states near the barrier top, the condition of the ergodic behavior is not very severe; it suffices that the asymmetry of the potential is comparable to the barrier height.

The method and the results can also be applied to various systems in physics, chemistry, and biology exhibiting double-level behavior and resonance tunneling. Literally speaking, we dealt with the microscopic Hamiltonians in this paper. But because of rapid development of electronics and cryogenic technologies, it has become possible to apply the same Hamiltonians to study cases where the corresponding variables are macroscopic (e.g., the magnetic flux through a SQUID ring, or charge or spin density wave phase in certain one-dimensional solids). In this paper, we studied the example of a tunneling processes in which the system penetrates into a classically forbidden region (a potential barrier). It is an intrinsically quantum effect with no classical counterpart, but it can nevertheless occur for macroscopic systems, and the tunneling of a macroscopic variable of the macroscopic system (e.g., spin or charge tunneling in atomic condensates trapped in a double-well potential [31]) can also be investigated by our method.

With this background in mind, our results are also intended to clarify different subtle aspects of tunneling phenomena. An example was given at the end of Section 3, where we found the geometrical phase acquired by a particle tunneling through a potential barrier. This phase can be tuned by the particle energy and by the barrier shape, and specific interference phenomena might occur. The observation of oscillations related to this geometrical phase in real systems has proved challenging. Evidently, because the forms of the model potentials that we used are rather special (and, in addition, only one-dimensional), we cannot discuss the behavior for general cases with full confidence. Nevertheless, we believe that the instanton approach employed in this work should also be useful in deriving valuable results for the general and multidimensional potentials.

It is essential that, in the instanton method discussed in this paper, a mere observation of several classical trajectories suffices to develop a qualitative insight for the

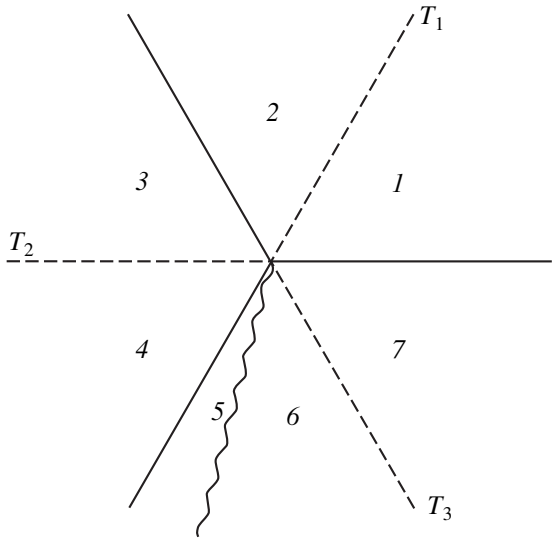


Fig. 9. The Stokes (solid) and anti-Stokes (dashed) lines in the vicinity of the linear turning point $V(X) = -X$. The cut is depicted by the wavy line, and the Stokes constant are T_1, T_2, and T_3.

quantum behavior. Although the semiclassical instanton approach is reliable in this context (as we illustrated in a number of particular examples), it is much more than a qualitative picture. As an approximation, the instanton method can be surprisingly precise. We also note that the results presented here are not only interesting in their own right (at least in our opinion) but may also be directly tested experimentally, because there are many systems where the model investigated in the paper is a reasonable model of reality.

The theory presented in our paper could be extended in several directions. One very interesting question is how our quantum mechanical instanton formulas (e.g., Eqs. (4.12)–(4.14) for the tunneling rate in the anharmonic X^3 decay potential) are modified by interactions with the surrounding media (see, e.g., [32], where the WKB approach was used to study the time evolution of quantum tunneling in a thermally fluctuating medium). Theoretical modelling of this case is hampered by the absence of detailed knowledge of the medium and of the interaction with it. A more specific study might become appropriate once suitable experimental results become available. A simple criterion for the strength of the interaction with the environment (in other words, for the effective temperature) for the crossover from the thermally activated classical to quantum mechanical decay can easily be found by equating the corresponding Arrhenius factor and the characteristic frequency oscillations inside the barrier Ω_* (see (A.10)).

All of the potentials investigated in this paper can be considered in a number of realistic cases as effectively resulting from avoiding the adiabatic level crossing in the situation where the adiabatic splitting is so large that any influence of the upper adiabatic states on the lower states can be neglected. Certainly, in the general case of an arbitrary coupling strength, this interaction of higher and lower adiabatic states must be taken into account, and the tunneling matrix elements must be accompanied by the corresponding Franck–Condon factors arising because of the violation of the Born–Oppenheimer approximation. We defer these problems to the future, although there is no doubt that the instanton approach is also useful in problems of this kind.

ACKNOWLEDGMENTS

The research described in this publication was made possible in part by the Russian Foundation for Basic Research (project nos. 00-03-32938 and 00-02-11785).

APPENDIX

Following [2], we introduce short notation for the standard basic WKB solutions,

$$(\circ, z) \equiv (q(z))^{-1/4}\exp(i\gamma W(z)) \tag{A.1}$$

and

$$(z, \circ) \equiv (q(z))^{-1/4}\exp(-i\gamma W(z)). \tag{A.2}$$

The position of the turning point is denoted by $\circ$ and is inessential if we seek solutions in the region $|z| \gg 1$. In accordance with definitions (2.18) and (2.19), we must add the dominant solution times a certain constant (the Stokes constant) to the subdominant (decaying) solution on the Stokes lines; the dominant and the subdominant solutions are exchanged on the anti-Stokes lines. To find the Stokes constant, we must match both solutions by encompassing the turning point and taking the cut on the complex z plane into account (see Fig. 9).

We first consider the linear turning point

$$q(z) = -z, \tag{A.3}$$

with the classically admissible region corresponding to $X > 0$. In this case, we have three Stokes lines, three anti-Stokes lines, one cut, and therefore seven different regions on the complex z plane where functions (A.1) and (A.2) must be matched; as a result, three Stokes constants must be determined. After not very sophisticated but rather tedious algebraic calculations, we find all the three Stokes constants

$$T_1 = T_2 = T_3 = i$$

and the connection matrix

$$\hat{M}^- = \exp\left(-i\frac{\pi}{4}\right) \times \begin{pmatrix} \exp(i\pi/4) & (1/2)\exp(-i\pi/4) \\ \exp(-i\pi/4) & (1/2)\exp(i\pi/4) \end{pmatrix} \tag{A.4}$$

relating the coefficients of the linear combinations of basic solutions (A.1) and (A.2) in the classically forbidden region (A_1, A_2) and in the classically admissible region (A_2, B_2) as

$$\begin{pmatrix} A_2 \\ B_2 \end{pmatrix} = \hat{M}^{-}\begin{pmatrix} A_1 \\ B_1 \end{pmatrix}. \qquad \text{(A.5)}$$

For the other linear turning point $q(z) = +z$, the connection matrix $\hat{M}^{+}$ is Hermitian conjugate to $\hat{M}^{-}$. The variation of the coefficients in the region between the two independent linear turning points z_1 and z_2 is determined by the diagonal matrix

$$\hat{L} = \begin{pmatrix} \exp(-i\gamma W^*) & 0 \\ 0 & \exp(i\gamma W^*) \end{pmatrix}, \qquad \text{(A.6)}$$

where

$$W^* = \int_{z_1}^{z_2} \sqrt{q(z)}\,dz.$$

Finally, for the solutions in the classically forbidden regions $X < X_1$ and $X > X_2$, the connection matrix is the direct matrix product of the above matrices,

$$\hat{M} = \hat{M}^{+}\hat{L}\hat{M}^{-}.$$

To generalize the procedure to second-order turning points, we must find the connection matrices relating the basic solutions to the Weber equation, namely,

$$(\circ, z) \equiv (z)^{\nu}\exp\left(-\frac{z^2}{4}\right) \qquad \text{(A.7)}$$

and

$$(z, \circ) \equiv (z)^{-\nu-1}\exp\left(\frac{z^2}{4}\right). \qquad \text{(A.8)}$$

In this case, we have four Stokes lines, four anti-Stokes lines, and one cut, and therefore nine different regions where the solutions must be matched (see Fig. 10a as an illustration). Four Stokes constants are given by

$$\tilde{T}_1 = \tilde{T}_2^{-1}[\exp(i2\pi\nu) - 1], \quad \tilde{T}_3 = \tilde{T}_1^*,$$

$$\tilde{T}_4 = -\tilde{T}_2\exp(-i2\pi\nu).$$

From the known asymptotic form of the parabolic cylinder functions, we can obtain the remaining Stokes constant

$$\tilde{T}_2 = \frac{\sqrt{2\pi}}{\Gamma(-\nu)}.$$

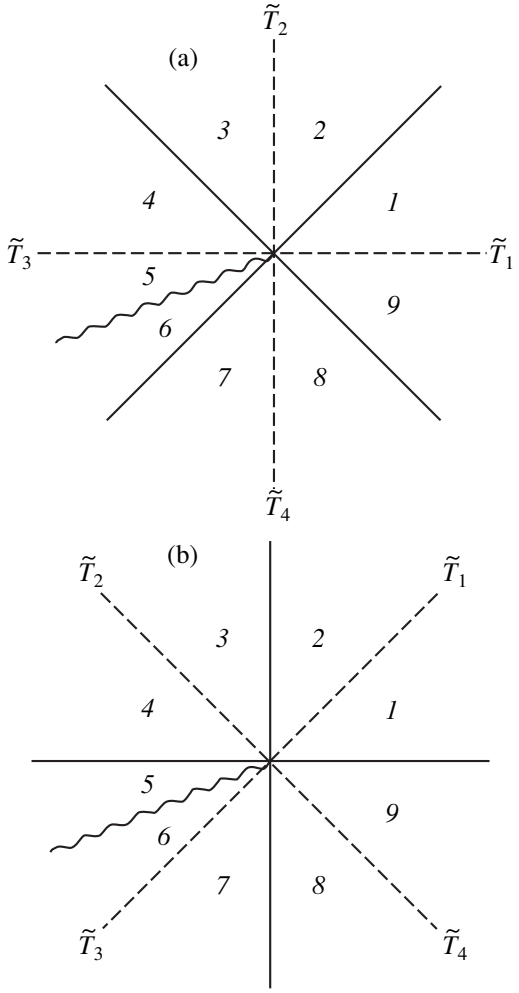


Fig. 10. The Stokes and anti-Stokes lines in the vicinity of the second-order turning points (with the same notation as in Fig. 9); (a) $V(X) = (1/2)X^2$; (b) $V(X) = -(1/2)X^2$.

The connection matrix for an isolated second-order turning point can therefore be represented as

$$\begin{pmatrix} -\tilde{T}_2 & \cos(\pi\nu) \\ \cos(\pi\nu) & -\dfrac{\sin^2(\pi\nu)}{\tilde{T}_2} \end{pmatrix}. \qquad \text{(A.9)}$$

This matrix depending on the energy ϵ determines, e.g., the instanton semiclassical solutions for the harmonic oscillator, $\epsilon = \nu + 1/2$. It can be verified by explicit calculations that, for the harmonic oscillator, the connection matrix also has the same form (A.9) in the WKB approach. The difference could appear only from anharmonic terms in the potential. But for low-energy states with $\epsilon/\gamma \ll 1$, the anharmonic corrections are small and the instanton and WKB connection matrices coincide up to the second order in these correction terms.

For the symmetric double-well potential, the connection matrix describing the variation of the coefficients at the basic solutions in Eqs. (A.7) and (A.8) between the two second-order turning points X_2^0 and X_3^0 is given by

$$\begin{pmatrix} \frac{n!}{\sqrt{2\pi}}\left(\frac{\Omega_0\gamma}{\Omega_*}\right)^{-\nu+1/2}\exp(\gamma W_E^*) & 0 \\ 0 & \frac{\sqrt{2\pi}}{n!}\left(\frac{\Omega_0\gamma}{\Omega_*}\right)^{\nu+1/2}\exp(-\gamma W_E^*) \end{pmatrix}, \tag{A.10}$$

where the instanton action is

$$W_E^* = \int_{X_2^0}^{X_3^0}\sqrt{2\left(V(X)-\frac{\epsilon}{\gamma}\right)}dX \tag{A.11}$$

and Ω_* is the characteristic *oscillation* frequency in the barrier (i.e., in the classically forbidden region). The connection matrix in Eq. (A.10) must be compared with a similar matrix in Eq. (A.6) for two linear turning points. For the asymmetric double-well potential in the region between the second-order and the linear turning points, the matrix analogous to (A.10) is

$$\begin{pmatrix} \left(\frac{n!}{\sqrt{2\pi}}\right)^{1/2}\left(\frac{\Omega_0\gamma}{\Omega_*}\right)^{-(1/2)(\nu+1/2)}\exp(\gamma W_E^*) & 0 \\ 0 & \left(\frac{\sqrt{2\pi}}{n!}\right)^{1/2}\left(\frac{\Omega_0\gamma}{\Omega_*}\right)^{(1/2)(\nu+1/2)}\exp(-\gamma W_E^*) \end{pmatrix}. \tag{A.12}$$

All the above matrices allow us to find any other connection matrix that we need in the particular examples considered in the main text of the paper. Any of them can be constructed as a corresponding product of the matrices in (A.4)–(A.6), (A.9), (A.10), and (A.12). It is worth noting a general property of the connection matrices that the connection matrix is real-valued for all bound states, and off-diagonal elements of the connection matrix are complex for continuum spectrum states.

Similarly to the problem of tunneling through the potential barrier $V(X) = -(1/2)X^2$, all the Stokes and anti-Stokes lines are rotated by the angle $\pi/4$ (see Fig. 10b) with respect to the corresponding lines for the parabolic well ($V(X) = (1/2)X^2$ considered above—see Fig. 10a). The connection matrix for the tunneling through the barrier is given by

$$\begin{pmatrix} S_1 & -i\exp\left(\frac{\pi\alpha}{2}\right) \\ i\exp\left(\frac{\pi\alpha}{2}\right) & S_1^{-1}(\exp(\pi\alpha)+1) \end{pmatrix}, \tag{A.13}$$

where $\alpha = i(2\nu + 1)$ and S_1 is the Stokes constant on the first quadrant bisectrix (see Fig. 10b). To find the Stokes constant S_1, we must match the sum of the incident and reflected waves to the solutions of the Weber equation at $X \longrightarrow -\infty$ and to the transmitted wave as $X \longrightarrow \infty$. This gives

$$S_1 = \frac{\sqrt{2\pi}}{\Gamma[(1+i\alpha)/2]}\exp\left(\frac{\pi\alpha}{4}\right).$$

REFERENCES

1. L. D. Landau and E. M. Lifshitz, *Course of Theoretical Physics*, Vol. 3: *Quantum Mechanics: Non-Relativistic Theory* (Nauka, Moscow, 1974; Pergamon, New York, 1977).
2. J. Heading, *An Introduction to Phase-Integral Methods* (Wiley-Interscience, London, 1962).
3. V. L. Pokrovskii and I. M. Khalatnikov, Zh. Éksp. Teor. Fiz. **40**, 1713 (1961) [Sov. Phys. JETP **13**, 1207 (1961)].
4. N. T. Maintra and E. J. Heller, Phys. Rev. A **54**, 4763 (1996).
5. C. S. Park and M. C. Jeong, Phys. Rev. A **58**, 3443 (1998).
6. A. K. Roy, N. Gupta, and D. M. Deb, Phys. Rev. A **65**, 012109 (2002).
7. A. M. Polyakov, Nucl. Phys. B **129**, 429 (1977).
8. S. Coleman, *Aspects of Symmetry* (Cambridge Univ. Press, Cambridge, 1985).
9. V. A. Benderskii, D. E. Makarov, and C. A. Wight, *Chemical Dynamics at Low Temperatures* (Wiley-Interscience, New York, 1994).

10. V. A. Benderskii, E. V. Vetoshkin, and H. P. Trommsdorf, Chem. Phys. **244**, 273 (1999).
11. V. A. Benderskii and E. V. Vetoshkin, Chem. Phys. **257**, 203 (2000).
12. V. A. Benderskii, E. V. Vetoshkin, L. S. Irgebaeva, and H. P. Trommsdorff, Chem. Phys. **262**, 369 (2000).
13. I. M. Lifshitz and Yu. Kagan, Zh. Éksp. Teor. Fiz. **62**, 385 (1972) [Sov. Phys. JETP **35**, 206 (1972)].
14. M. Grifoni and P. Hanggi, Phys. Rep. **304**, 229 (1998).
15. J. E. Avron and E. Gordon, Phys. Rev. A **62**, 062504 (2000).
16. J. Ankerhold and H. Grabert, Europhys. Lett. **47**, 285 (1999).
17. Y. Kayanuma and Y. Mizumoto, Phys. Rev. A **62**, 061401 (2000).
18. V. A. Benderskii and E. I. Kats, Phys. Rev. E **65**, 036217 (2002).
19. V. V. Avilov and S. V. Iordanskiĭ, Zh. Éksp. Teor. Fiz. **69**, 1338 (1975) [Sov. Phys. JETP **42**, 683 (1975)].
20. V. A. Benderskii, E. V. Vetoshkin, L. von Laue, and H. P. Trommsdorff, Chem. Phys. **219**, 143 (1997).
21. A. M. Dykhne, Zh. Éksp. Teor. Fiz. **41**, 1324 (1961) [Sov. Phys. JETP **14**, 941 (1961)].
22. *Heigher Transcendental Functions (Bateman Manuscript Project)*, Ed. by A. Erdelyi (McGraw-Hill, New York, 1953, 1953, 1955; Nauka, Moscow, 1965, 1966, 1967), Vols. 1–3.
23. M. Berry, Proc. R. Soc. London, Ser. A **392**, 45 (1984).
24. M. Wilkinson, J. Phys. A **17**, 3459 (1984).
25. H. Karatsuji, Prog. Theor. Phys. **74**, 439 (1985).
26. A. O. Caldeira and A. J. Legget, Ann. Phys. **149**, 374 (1983).
27. D. Bohm, *Quantum Theory* (Prentice-Hall, Englewood Clifs, 1951; Nauka, Moscow, 1965).
28. L. L. Chang, L. Esaki, and R. Tsu, Appl. Phys. Lett. **24**, 593 (1974).
29. J. P. Eisenstein, L. N. Pfeiffer, and K. W. West, Phys. Rev. Lett. **74**, 1419 (1995).
30. Sh. Matsumoto and M. Yoshimura, Phys. Rev. A **63**, 012104 (2000).
31. M. Baer, Phys. Rep. **358**, 75 (2002).
32. H. Pu, W. Zhang, and P. Meystre, Phys. Rev. Lett. **87**, 140405 (2001).

Journal of Experimental and Theoretical Physics, Vol. 95, No. 4, 2002, pp. 662–672.
Translated from Zhurnal Éksperimental'noĭ i Teoreticheskoĭ Fiziki, Vol. 122, No. 4, 2002, pp. 765–777.

PLASMA, GASES

Statistical and Dynamic Intensities of Atomic Spectral Lines in Plasma

L. A. Bureyeva[a, *], V. S. Lisitsa[b], and D. A. Shuvaev[b]

[a]*Institute of Spectroscopy, Russian Academy of Sciences, Troitsk, Moscow oblast, 142190 Russia*
[b]*Russian Research Centre Kurchatov Institute, pl. Kurchatova 1, Moscow, 123181 Russia*
**e-mail: bureyeva@sci.lebedev.ru*

Received February 8, 2002

Abstract—Based on the quasi-classical approach, we solve the problem of a radiative cascade between Rydberg states in the space of parabolic quantum numbers that correspond to atom (or ion) quantization in a plasma microfield. We consider the level population source associated with the dielectronic recombination of electrons on ions with the cores in an external electric field. We determine the populations of states and the intensities of transitions between levels with a small change in principal quantum number. These "dynamic" intensities are compared with the "statistical" intensities that correspond to an equidistribution in Stark sublevels. An appreciable discrepancy was found between the two types of intensities in plasma of low density that corresponds to the dominance of radiative transitions over collisional transitions. © *2002 MAIK "Nauka/Interperiodica".*

1. INTRODUCTION

The problem of the statistical and dynamic intensities of atomic spectral lines in an electric field observed in a dense or rarefied medium arose at the first stages of Stark effect measurements. It was discussed in detail by Bethe and Salpeter [1] in their well-known monograph. The essence of this problem is that statistical intensities emerge in a dense medium where each excitation is either proportional to the statistical weight of the states or (if this is not the case) is accompanied by collisional "mixing" in statistical weights after excitation in a time shorter than the time of radiative decay from these states. In this case, the observed line intensities can generally be obtained by summing the intensities of the transitions from individual sublevels proportionally to their statistical weights. In contrast, dynamic intensities emerge at a low density when, after the population of a particular sublevel, the radiative transition occurs directly from this sublevel in a fairly short time when collisions have no time to establish a statistical equilibrium between the sublevels. Under these conditions, the intensity of radiation from the entire line is the sum of the products of the intensities of individual components by the component distribution function determined by the radiative cascade. The explicit form of this distribution function and, hence, the difference between the statistical and dynamic intensities can be calculated for hydrogen-like states in plasma. This is the subject of our study.

Note that most calculations of spectral line profiles were performed for a statistical distribution of atomic populations (see [2]). A few calculations of the profiles for low-excitation lines of multiply charged ions, where the populations and line profiles were calculated simultaneously based on the density-matrix formalism, constitute an exception. However, these calculations are cumbersome even for these low-excitation lines, and it is unlikely that they can be applied to highly excited atomic states.

It is pertinent to point out that the population of atomic states in the space of quantum numbers is multidimensional in nature. For a statistically equilibrium population, it will suffice to restrict our analysis to a one-dimensional cascade that includes transitions only between various principal quantum numbers n (the n cascade), which is widely used in atomic kinetics [3]. At a low plasma density and in the absence of an electric field, a two-dimensional spherically symmetric radiative cascade in the space of principal (n) and orbital (l) quantum numbers is a natural approximation [4]. An important factor in plasma is the electric field produced by the surrounding electric charges. In this field, spherical symmetry is lost, so the radiative cascade in the space of principal (n), electric (k), and magnetic (m) quantum numbers is three-dimensional.

Note that the direct quantum calculation of a radiative cascade in an electric field is cumbersome when transitions between highly excited (Rydberg) states are included, because it necessitates taking into account an extremely large number of matrix elements for which there are no strict selection rules in the parabolic basis that diagonalizes the interaction with the field. Indeed, for the principal quantum number $n = 10^2$, the number of matrix elements between adjacent levels is of the order of $10^4 \times 10^4 = 10^8$. Since the number of such sublevels is of the order of 10^2 and since of the order of 10^3 matrix elements must also be known for their population to be taken into account, we conclude that the total

number of such matrix elements is of the order of 10^{13}. This value is unrealistic for its direct allowance in the chain of cascades. Therefore, below, we use the classical approximation, which is valid for large quantum numbers, to calculate the cascade.

The plasma electric microfield produced by plasma ions is assumed to be constant (static), while the electrons are responsible for the collisional mixing of sublevels. The static nature of the field implies that the Stark splitting of atomic sublevels in it is large compared to the characteristic frequency of its variation; i.e., the atom adiabatically "follows" the field while conserving its parabolic quantum numbers k and m. The following standard condition serves as the criterion for the field being static:

$$g = \Delta\omega_S \tau_F \gg 1, \quad (1)$$

where

$$\Delta\omega_S = \frac{3}{2}\frac{\hbar}{m_e}\frac{n(n-1)}{Z}F \sim \frac{n^2}{Z}F,$$

$$\tau_F^{-1} \sim \frac{v_i}{r} \sim N_i^{1/3} v_i,$$

whence we obtain

$$g = \frac{Zn^2}{Z} N_i^{2/3} \frac{1}{N_i^{1/3} v_i} = \left(\frac{N_e}{Z}\left(\frac{n^2}{v_i}\right)^3\right)^{1/3} \gg 1.$$

Here, Z is the ion charge, $F = ZN_i^{2/3}$ is the plasma ion microfield strength, τ_F is the characteristic time of the ion field variation (N_i is the ion density, and v_i is their thermal velocity), and m_e is the electron mass. Clearly, this condition corresponds to a fairly high plasma electron density N_e.

The second condition is the population "nonmixing" of atomic states in the time of their radiative decay. This implies that we consider ions with a sufficiently large charge Z, for which the radiative width A proportional to Z^4 is large enough compared to the frequency of the electron collisions Γ responsible for the mixing,

$$\Gamma \ll A.$$

In contrast, this condition corresponds to a low density and it is given, for example, in the monograph [3]. The set of these conditions limits the electron density both from above (collisions are disregarded) and from below (the field is static). Substituting an electron temperature of the order of the ionization potential, $T_e = Z^2\mathrm{Ry}$, into these conditions yields the following inequalities for the electron density:

$$\left(\frac{3Z}{n^2}\right)^3 \ll \frac{N_e}{10^{18}\ \mathrm{cm}^{-3}} \ll \left(\frac{Z}{n}\right)^7. \quad (2)$$

For example, for ions with a charge of the order of 30 and for principal quantum numbers of the order of 10^2, the density range is

$$10^{12}\ \mathrm{cm}^{-3} \ll N_e \ll 10^{15}\ \mathrm{cm}^{-3}. \quad (3)$$

Comparing the left- and right-hand sides of inequalities (2), we find the ranges of atomic parameters in our approximation

$$Z^4 \gg 10n. \quad (4)$$

Clearly, these ranges of plasma and atomic-ion parameters refer to thermally equilibrium plasma. However, there is a considerable number of objects in which multiply charged ions and excited atoms exist at low plasma temperatures and densities, at which the radiative cascade plays a dominant role. Here, we primarily note the plasma of storage rings and the rarefied plasma of gaseous nebulae.

Thus, we consider below the population of highly excited states for multiply charged ions via the radiative cascade produced by a population source. We consider the dielectronic recombination of electrons on ions with complex cores having transitions without any change in principal quantum number n_c (the $\Delta n_c = 0$ transitions) as this source. Dielectronic recombination on such transitions populates the highly excited ion states with the principal quantum number $n = 10^2$. The hydrogen-like nature of such Rydberg states allows us to determine universal distributions for the intensities of the cascade-populated lines and to derive an explicit relation between the statistical and dynamic intensities of these lines.

2. THE KRAMERS FORMULAS IN PARABOLIC COORDINATES: THE INTENSITIES OF RADIATIVE TRANSITIONS

The presence of an external electric field acting on an atom in plasma or produced by external sources necessitates considering radiative transitions in parabolic quantum numbers, in contrast to the spherical basis for a free atom. The passage from the spherical basis to the parabolic one is known [5] to be determined by the transformation of the wave functions with the Clebsch–Gordan coefficients:

$$\Psi(n, n_1, n_2, m) = \sum_l C(n, n_1, n_2, l, m)\Psi(n, l), \quad (5)$$

where n_1 and n_2 are the parabolic quantum numbers; m is the magnetic quantum number; and $C(n, n_1, n_2, l, m)$ are the Clebsch–Gordan coefficients.

Below, we use the "electric" quantum number k that defines the level energy shift in an electric field:

$$k = n_1 - n_2. \quad (6)$$

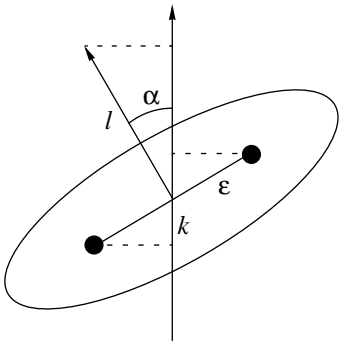


Fig. 1. The trajectory of an atomic electron in a Coulomb field and the angles defining the relationship between the spherical and parabolic variables that describe the motion.

The passage to relations (5) and (6), which are similar to the relations for a hydrogen-like atom, occurs for a sufficiently strong electric field, when the level quantum defects may be ignored and the states may be assumed to be completely mixed by this field. In practice, complete mixing occurs for the Rydberg states under consideration at low field strengths (of the order of 10 V cm^{-1}; see [6]), whereas the plasma microfield strengths are several orders of magnitude higher. In addition, the quantum defect rapidly decreases with increasing orbital angular momentum, and the statistical weights of states with low angular momenta are also low. Therefore, we use below the approximation of mixed hydrogen-like atomic states.

As was noted above, the direct calculations with the wave functions (5) are cumbersome. Therefore, we use the quasi-classical and purely classical approximations. In these approximations, the transition probabilities in the spherical and parabolic bases are related by the squares of the Clebsch–Gordan coefficients. The latter are [5]

$$[C(n,k|l,m)]^2 = \frac{2l}{\pi[(l^2-l_{\min}^2)(l_{\max}^2-l^2)]^{1/2}}, \quad (7)$$

Table 1. Comparison of the Kramers probabilities for the radiative transition from the $n = 10$ level described by the quantum numbers n, k, and $m = 0$ to all the low-lying levels with numerical results [7]

k	$B(k, m = 0)$	Exact
1	0.73	0.48
3	0.77	0.57
5	0.88	0.77
7	1.15	1.17
9	2.63	2.15

where

$$l_{\min}^2 = \frac{1}{2}\{[(n-1)^2+m^2-k^2]$$

$$-[[(n-1)^2+m^2-k^2]-4(n-1)^2m^2]^{1/2}\},$$

$$l_{\max}^2 = \frac{1}{2}\{[(n-1)^2+m^2-k^2]$$

$$+[[(n-1)^2+m^2-k^2]^2-4(n-1)^2m^2]^{1/2}\},$$

and we use the electric quantum numbers. The quantities $l_{\max}$ and $l_{\min}$ define the relationship between the parabolic and spherical variables. The geometric meaning of this relationship directly follows from an examination of the trajectory of an atomic electron in a Coulomb field shown in Fig. 1.

The relationship between the spherical (orbital angular momentum) and parabolic (the projection of orbital eccentricity ε) variables directly follows from Fig. 1:

$$\frac{k}{n} = \varepsilon\sin\alpha = \varepsilon\sqrt{1-\frac{m^2}{l^2}}. \quad (8)$$

It is of interest to establish the explicit form of the classical radiative transition probabilities in parabolic coordinates. To this end, we transform the standard Kramers formulas for the radiative decay probability in the spherical basis by using the quasi-classical representation of the Clebsch–Gordan coefficients (7) according to the relation

$$B(k,m) = \int_{l_{\min}}^{l_{\max}} A(n,l)[C(n,k|l,m)]^2 dl. \quad (9)$$

Directly calculating the integrals in (9) yields

$$B(k,m) = \frac{4Z^4}{\pi\sqrt{3}c^3n^3(n-1)|m|}, \quad (10)$$

$$B(k,m=0) = \frac{4Z^4}{\pi^2\sqrt{3}c^3n^3}\frac{2}{[(n-1)^2-k^2]^{1/2}} \quad (11)$$

(c is the speed of light). Formulas (10) and (11) are the "parabolic analogs" of the Kramers formulas in the spherical basis. Note that, for $m \neq 1$, these probabilities do not depend on the electric quantum number k.

It is of interest to compare the calculations in the Kramers approximation with accurate numerical calculations of the transition probabilities in the parabolic basis [7]. The results of this comparison are presented in Table 1 for $m = 0$ and Table 2 for $m \neq 0$. As follows from the tables, the Kramers probabilities have a high accuracy, particularly for $m \neq 0$. Note that these data refer to moderate principal quantum numbers, $n = 10$.

Next, let us consider the probabilities of radiative transitions in parabolic coordinates between two levels with high n and various parabolic quantum numbers. These probabilities are defined by the matrix elements of the radius-vector components for an atomic electron between states with certain parabolic quantum numbers. The corresponding results are known to be described by the Gordon formulas, which are expressed in terms of cumbersome hypergeometric functions. However, these formulas can be significantly simplified in the quasi-classical range or, to be more precise, in the range of large quantum numbers with their small change [8]:

$$n \gg 1, \quad \Delta n \ll n. \tag{12}$$

In this range, the array of the matrix elements that correspond to the upper parabolic quantum numbers n_1, n_2, m and the lower parabolic quantum numbers n_1', n_2', m' can be broken down into large-scale and small-scale blocks. The large scale corresponds to the combination of quantum numbers

$$K = (n_1 - n_2) - (n_1' - n_2'). \tag{13}$$

At the same time, the intensities of the Stark components within each block in K are described by the small-scale variable

$$k = n_1' - n_2', \tag{14}$$

which is assumed below to change continuously between $-n$ and n.

The intensity of the large-scale blocks decreases sharply as their number increases, so only the first two or three blocks must be taken into account. The small-scale dependence admits a simple analytic representation in the form of power functions. Thus, for the $\Delta n = 1$ ($\Delta n = n - n'$) transitions, which will be widely used below for illustration, this dependence is [9]

$$X_m^{m-1} \approx b\sqrt{(n_1 + m)(n_2 + m)} \times\left(1 - \frac{(n_1 + 1)(n_2 + 1)}{b^2}\right), \tag{15}$$

$$X_m^{m+1} \approx b\sqrt{n_1 n_2}\left(1 - \frac{(n_1 + m)(n_2 + m)}{b^2}\right), \tag{16}$$

$$Z_m^m \approx b[\sqrt{n_1(n_1 + m)}\delta_{K,+1} + \sqrt{n_2(n_2 + m)}\delta_{K,-1}], \tag{17}$$

$$b = \frac{4nn'}{(n - n')^2} = 4n(n-1).$$

It should be noted that the above results can also be obtained from a purely classical analysis of the electron trajectory in a Coulomb field in parabolic coordinates (see [10]). Indeed, the relative radiation amplitudes calculated by Born [10] are the double Fourier expansion in parabolic variables and are given in the Appendix.

Table 2. Same as Table 1 for n, k, and $m > 0$

m	$\bar{B}^{\mathrm{Kr}}(m)$	$\bar{B}^{0}(m)$	$\bar{B}$
1	1.129	1.045	1.0871
2	0.556	0.518	0.552
3	0.363	0.341	0.366
4	0.265	0.254	0.272

Note: $\bar{B}^{0}(m)$ is the quasi-classical formula from [8], $\bar{B}$ is the quantum calculation, and $\bar{B}^{\mathrm{Kr}}(m)$ was calculated using formula (10) with averaging over k.

Integrating these formulas over the magnetic quantum number m yields the total relative intensities of the Stark π and σ components for $\Delta n = 1$:

$$I_k^{\sigma} \sim \sum_m [(X_m^{m-1})^2 + (X_m^{m+1})^2] \approx \frac{1}{3n^2}(2n^3 - 3n^2k + k^3),$$

$$I_k^{\pi} \sim \sum_m (Z_m^m)^2 \approx \frac{1}{6n^2}(n^3 + 3n^2k - 3nk|k| - k^2|k|).$$

These expressions are used below to calculate the intensities of spectral lines.

3. THE KINETIC EQUATION

The kinetic equation for the populations of atomic states that describes the radiative cascade in the space of quantum numbers is the relation for balance between the probabilities of arrival in a given quantum state and departure from this state through photon emission. As was noted above, the quantum kinetic equation is cumbersome because of the large number of transitions that must be taken into account in solving it. For highly excited states, the quantum equation can be transformed into the classical continuity equation, which expresses the flow of an electron fluid in the space of the corresponding quantum numbers. The passage to this equation was closely followed in spherical coordinates (see [4]). The corresponding results match the classical kinetic equation for a radiative cascade that was first derived by Belyaev and Budker [11]:

$$\dot{n}\frac{\partial}{\partial n}f + \dot{l}\frac{\partial}{\partial l}f = q(n, l). \tag{18}$$

Here, f is the distribution function in energy (principal quantum number) and angular-momentum (orbital quantum number) space that determines the atomic level populations. The quantity $q(n, l)$ is the population

source that depends on the specific population mechanism for excited atomic states during collisions with plasma particles. The changes in quantum numbers [the derivatives in Eq. (18)] are defined by purely classical relations for the loss of energy and angular momentum through radiation in a Coulomb field [12]:

$$\dot{n} = \frac{Z^4}{c^3 l^5}\left(3 - \frac{l^2}{n^2}\right), \tag{19}$$

$$\dot{l} = \frac{2Z^4}{c^3 n^3 l^2}. \tag{20}$$

The classical kinetic equation describes the population of atomic states with a high accuracy even for small principal quantum numbers, as was shown in [4] by a detailed comparison with the results of quantum calculations (see also [8]).

The plasma microfield is taken into account by passing from the spherically symmetric system "atom in the absence of field" to the system "atom in field," which is quantized in parabolic coordinates. The classical picture of the motion of an atomic electron in an electric field consists in the orbital precession around the field direction with the frequency determined by the angular momentum component along the field direction (i.e., Stark splitting) and a slow (compared to this frequency) decrease in the parameter of a Keplerian ellipse due to the radiative friction force. The structure of this force (and, hence, the intensities of the corresponding radiative transitions) is determined by the electron motion in a Keplerian ellipse unperturbed by the electric field, whose strength is assumed to be small compared to the Coulomb field strength. For a quantum treatment, this is expressed in that, to a first approximation, the intensities of the split Stark components do not depend on the electric-field strength and are defined by the standard Gordon formulas. The field dependence of the radiation intensity arises in a higher order of the perturbation theory and leads to small corrections to the intensities of the Stark components calculated in [13]. Thus, in our approximation, the classical orbital precession results in Stark state splitting, while the intensities of radiative transitions are determined by unperturbed Keplerian motion as before. Nevertheless, for the intensities of radiative transitions to be related to certain Stark splitting, this motion must be considered in the parabolic variables that correspond to a certain component of the dipole moment along the field direction.

Let us transform the kinetic equation (18) to the parabolic variables. Clearly, this requires knowing the Jacobian of the corresponding transformation. It can be easily found from the above relations (8) between the spherical and parabolic variables. The relation between these variables is (see also Fig. 1)

$$k^2 = (n^2 - l^2)\left(1 - \frac{m^2}{l^2}\right). \tag{21}$$

Given this relation, the classical equation for a radiative cascade in parabolic variables takes the form

$$\dot{n}\frac{\partial}{\partial n} f + \dot{k}\frac{\partial}{\partial k} f + \dot{m}\frac{\partial}{\partial m} f = q(n, k, m). \tag{22}$$

Here, all the derivatives (19) and (20) are expressed in terms of the parabolic variables by substituting in the dependence $l = l(n, k, m)$ that follows from (21):

$$\dot{k} = \dot{n}\frac{\partial k}{\partial n} + \dot{l}\frac{\partial k}{\partial l} + \dot{m}\frac{\partial k}{\partial m} = Z^4 \frac{l^2(n, k, m) - m^2}{c^3} \times \frac{n^2(3n^2 - l^2(n, k, m)) - 2l^4(n, k, m)}{n^3 l^7(n, k, m) k}. \tag{23}$$

Equation (22) is three-dimensional, in contrast to the two-dimensional spherical case, because the spherical symmetry of atomic states breaks down in an external electric field.

The right-hand side of Eq. (22) is the population source in parabolic variables. This source cannot always be obtained from the corresponding spherical source by its transformation with the Clebsch–Gordan coefficients. As we show below, its calculation is a problem of its own. The characteristics are introduced by reducing the left-hand side of the equation to the total derivative:

$$\left\{\frac{\partial}{\partial n} + \frac{\dot{k}}{\dot{n}}\frac{\partial}{\partial k} + \frac{\dot{m}}{\dot{n}}\frac{\partial}{\partial m}\right\} f = \frac{d}{dn} f = \frac{q(n, k, m)}{\dot{n}}. \tag{24}$$

The equation is integrable if the variables k and m satisfy the characteristic equations

$$\frac{\partial k}{\partial n} = \frac{\dot{k}}{\dot{n}}, \quad \frac{\partial m}{\partial n} = \frac{\dot{m}}{\dot{n}}. \tag{25}$$

Their solutions are the functions

$$m^2 = \frac{k^2}{\ln\dfrac{n^2}{n^2 - k^2} - k^2 C_2}, \tag{26}$$

$$k^2 = n^2(1 - C_1^{2/3} m^2) + m^2, \tag{27}$$

where the constants C_1 and C_2 are the characteristics of the equations. Given (24)–(27), the distribution function takes the form

$$f(n, k, m) = \varphi(n, k, m) + \int_{n+1}^{\infty} \frac{q(n', k[n', m, C_1], m[n', k, C_2])}{|\dot{n}(n', k[n', m, C_1], m[n', k, C_2])|} dn', \tag{28}$$

where $\varphi(n, k, m)$ is the boundary conditions.

In solving Eq. (22), we chose the boundary conditions in such a way that, when letting n tend to infinity, the solution matched the direct population determined by the ratio of the source to the total radiative decay

probability of a given state. Clearly, the latter is defined by the Kramers formulas in parabolic variables.

Thus, the problem of calculating the populations of highly excited states in the presence of a known population source is solved.

4. THE POPULATION SOURCES

The main population sources for highly excited states are associated with the recombination to highly excited levels. As for the radiative recombination, it can be obtained from the corresponding results for the spherical basis by their transformation with the Clebsch–Gordan coefficients, much as we did in Section 2 when deriving the Kramers formulas. In this case, since the Clebsch–Gordan coefficients are normalized, the total recombination flux into the atom, i.e., the total rate of radiative recombination, is conserved. The rate of recombination to a given level summed over all the spherical (or parabolic) quantum numbers that belong to this level is also conserved. Differences are possible for the populations of separate Stark components, because the radiative transitions even from a statistically equilibrium continuum do not result in the statistical-weight population of individual sublevels (see [4, 8]).

As we pointed out above, the main population source considered here is dielectronic recombination. Dielectronic recombination for transitions in the ion core without any change in principal quantum number n_c populates the highly excited (Rydberg) states of the ion up to principal quantum numbers of the order of 10^2 (see [3]).

Dielectronic recombination is peculiar in that the corresponding results for the population source in the parabolic basis cannot be obtained from the results in the spherical basis by the direct transformation with the Clebsch–Gordan coefficients. Indeed, the rate of dielectronic recombination is the probability of an impinging electron being captured on a doubly excited level multiplied by the probability of radiative stabilization of the ion core excited during the capture. The latter is determined by the so-called branching factor, which depends on the ratio of the radiative stabilization rate for the core to the autoionization rate (inverse decay) of the captured electron [3]. Of these two quantities, only the autoionization rate can be transformed to the parabolic basis with the Clebsch–Gordan coefficients. At the same time, the ion core is not affected noticeably by the electric field, and its radiative stabilization probability is constant in both bases. In fact, the ion is the part associated with the highly excited electron that is strongly affected by the electric field and associated with the core that is virtually unaffected by the field.

Thus, the dielectronic recombination rate for an ion with charge Z at temperature T in the parabolic basis can be written as

$$Q_{DR}(n,k,m) = \left(\frac{4\pi\mathrm{Ry}}{T}\right)^{3/2}\frac{g_f}{g_i} \times a_0^3 W_R e^{-\hbar\omega_c/T + Z^2\mathrm{Ry}/2n^2T}\frac{W_A(n,k,m)}{W_R + W_A(n,k,m)}, \quad (29)$$

where $W_R = 2\omega^2 f_{ij}/c^3$ is the radiative stabilization rate of the core; f_{ij} is the transition oscillator strength in the core; ω_c is the transition frequency in the core; a_0 is the Bohr radius; g_i and g_f are the statistical weights of the initial and final states, respectively; T is the electron temperature; and W_A is the autoionization rate in the parabolic basis, which is related by the Clebsch–Gordan coefficients to the autoionization rate in the spherical basis:

$$W_A(n,k,m) = \int_{l_{\min}}^{l_{\max}} W_A(n,l)[C(n,k|l,m)]^2 dl. \quad (30)$$

The autoionization width in the spherical basis can be calculated exactly with the Coulomb wave functions [14]. Passing to large quantum numbers, we can obtain simple analytic results for the autoionization widths. The same results can be obtained in a purely classical way if it is considered that the matrix elements in the quasi-classical case must transform into the Fourier coefficients for the electron trajectory in a Coulomb field. The latter are known to be expressed in terms of Bessel functions [12]. We are concerned with the limiting forms of these functions that correspond to highly curved trajectories of the electron, where it undergoes the highest acceleration and, thus, the most intense radiative transitions. In this region, the transition probability is determined by the orbital angular momentum of the electron alone and does not depend on its initial energy. This region serves as the basis for the so-called Kramers electrodynamics, which allows even highly inelastic transitions to be described classically [15]. Formally, the following simple considerations form the basis for using classical methods to calculate the rates of dielectronic recombination. Indeed, the recombination of an electron requires that its initial energy E be lower than the core excitation energy, which is equal to Z Ry in order of magnitude. In that case, the following conditions are clearly satisfied:

$$E \ll Z\mathrm{Ry} \ll Z^2\mathrm{Ry}, \quad \frac{Ze^2}{\hbar v} \gg 1, \quad (31)$$

i.e., exactly the conditions for the electron motion in a Coulomb field being quasi-classical.

Note that the autoionization width is uniquely related to the threshold inelastic excitation cross section [3]. For the $\Delta n_c = 0$ transitions, these cross sections are also well described in the quasi-classical approximation (see [16]).

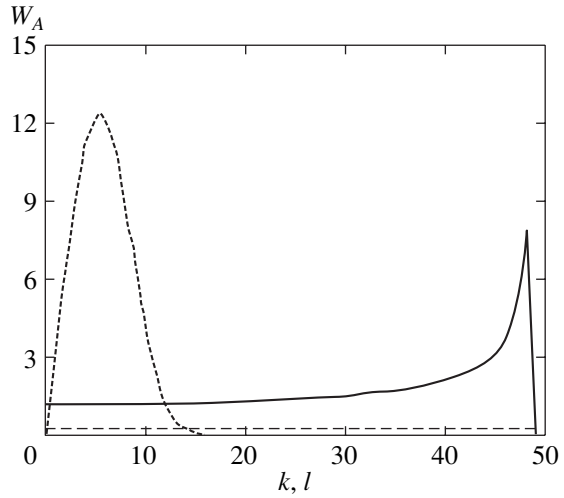


Fig. 2. Autoionization widths of the level with the principal quantum number $n = 50$ versus orbital quantum number l (dashed line) and versus electric quantum number k (solid line) for a lithium-like ZnXXVIII ion (in arb. units). The lower straight line represents the radiative width.

Using any of the above methods of calculation, we derive the following expression for the autoionization width in spherical coordinates [8]:

$$W_A = \frac{f_{if}}{\pi n^3} l G\left(\frac{\omega_c (l+1/2)^3}{3Z^2}\right), \tag{32}$$

where

$$G(x) = x[K_{1/3}^2(x) + K_{2/3}^2(x)].$$

We see that the effective values of the orbital angular momentum that contribute to the autoionization width are of the order of

$$l_{\text{eff}} = \left(\frac{3Z^2}{\omega_c}\right)^{1/3}. \tag{33}$$

Note that expression (32) is valid if the orbital angular momentum is not too large; otherwise, the pattern of decrease in autoionization width will be different (see [8]), but this region does not contribute appreciably to the cascade.

For the transformation to the parabolic basis, let us integrate (32) with the square of the Clebsch–Gordan coefficients similar to (9). The result can be expressed in terms of the universal function

$$W_A(n, k, m) = \frac{f_{if}}{\pi n^3} I(n, k, m), \tag{34}$$

where

$$I(n, k, m) = \frac{2 l_{\text{eff}}}{\pi} \int_{t_{\min}}^{t_{\max}} \frac{t^2 G(t^3)}{\sqrt{(t^2 - t_{\min}^2)(t_{\max}^2 - t^2)}} dt,$$

$$t_{\min} \approx \frac{(n-1)m}{\sqrt{(n-1)^2 + m^2 - k^2}}, \quad t_{\max} \approx \frac{(n-1)m}{t_{\min}},$$

$$t_{\max} = \frac{l_{\max}}{l_{\text{eff}}}.$$

Formula (34) is a universal representation of the autoionization width in parabolic variables for all the ions that have transitions without any change in principal quantum number within the core. Its dependence on the electric quantum number is much less sharp than the dependence on the orbital quantum number in the spherical basis. In Fig. 2, the autoionization widths are plotted against the corresponding quantum numbers for an ion with $Z = 30$ and for the $n = 100$ level. It is seen that the region of phase space occupied by the electric quantum numbers is much wider than the region occupied by the orbital quantum numbers. This circumstance underlies the fact that the dielectronic recombination in an electric field significantly exceeds this recombination in the absence of an electric field.

The result (34) should be substituted into the general formula (29) for the dielectronic recombination rate. It is convenient to write the rate of recombination to a specified parabolic state by introducing the effective principal quantum number of the sublevel being populated given by the relation

$$q_{DR} = \frac{B(Z, T_e)}{c^3\left(1 + \dfrac{n^3}{n^{*3}}\right)}, \tag{35}$$

$$B(Z, T_e) = 2\frac{g_f}{g_i}\left(\frac{2\pi}{T_e}\right)^{3/2} \omega^2 f_{ij} e^{-\omega/T_e},$$

$$n^{*3} = \frac{c^3 l_{\text{eff}}^2}{\omega_c^2 \pi l_{\max}}\left(\frac{l_{\min}}{l_{\text{eff}}}\right)^2 \int_1^\infty \frac{t^2 e^{-2(l_{\min}/l_{\text{eff}})^3 t^3}}{\sqrt{t^2 - 1}} dt. \tag{36}$$

The result (35) defines the population source of a specified parabolic state. It should be substituted into the general solution to the kinetic equation to obtain the level populations (the distribution functions in parabolic quantum numbers).

5. LEVEL POPULATIONS

For the level populations to be calculated, we must perform the integration in Eq. (28). The characteristic that corresponds to the correct initial conditions is identical to the function l_{eff}. It is easy to derive the limiting expressions for the populations in the range of quantum

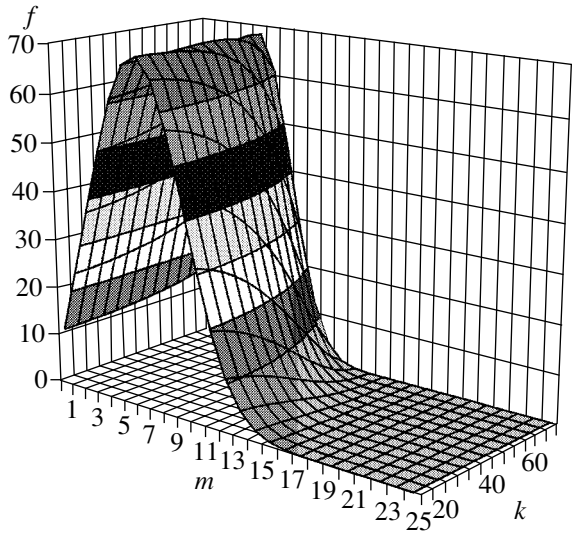


Fig. 3. The population distribution function for the $n = 100$ level versus quantum numbers k and m (in arb. units).

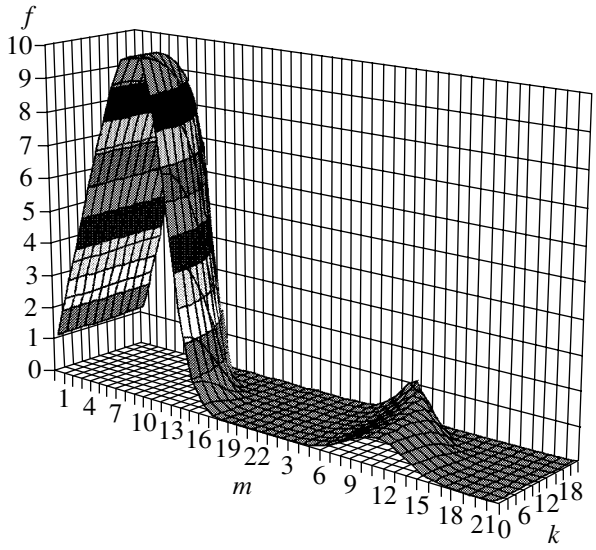


Fig. 4. Same as Fig. 3 for $n = 30$. The left and right parts on the m scale correspond to the direct and cascade populations, respectively.

numbers that are small and large compared to the effective quantum number (36):

$$f(n,k,m,Z) = \frac{B(Z,T)}{Z^4}\left(\frac{\pi\sqrt{3}n^4|m|}{4} + \frac{\sqrt{2}}{4}\pi\frac{l_{\min}^5}{3}\right.$$
$$\left.\times\left(\frac{c^3 l_{\min}^3}{\omega^2\pi m}\int_1^\infty \frac{t^2 e^{-2(l_{\min}/l_{\text{eff}})^3 t^3}}{\sqrt{t^2-1}}dt\right)^{1/4}\right), \quad n \ll n^*, \tag{37}$$

$$f(n,k,m,Z) = \frac{B(Z,T)}{Z^4}\frac{c^3}{\pi\omega^2}$$
$$\times\left(\frac{\pi\sqrt{3}|m|^3}{4} + \frac{1}{9}\frac{|m|^7}{(n+1)^3}\right) \tag{38}$$
$$\times\int_1^\infty \frac{t^2 e^{-2(l_{\min}/l_{\text{eff}})^3 t^3}}{\sqrt{t^2-1}}dt, \quad n \gg n^*.$$

The calculations in the entire range of parabolic quantum numbers are performed numerically using formulas (37) and (38).

Let us consider the population structure of the parabolic states for a lithium-like zinc ion with the nuclear charge $Z = 30$ populated by dielectronic recombination followed by a radiative cascade in high-temperature plasma. Figures 3–5 show the level population distributions in the space of electric (k) and magnetic (m) quantum numbers for various principal quantum numbers. For comparison, Fig. 6 shows a plot of the population distribution in the spherical basis against orbital angular momentum. We see from comparison of these data that the distribution in magnetic quantum number in the parabolic basis is similar in pattern to the distribution in orbital angular momentum in the spherical basis. At the same time, the distribution in electric quantum number is much wider, of the order of the principal quantum number n. Therefore, it is clear that the total (integrated over all quantum numbers) population in the parabolic basis is considerably larger than that in the spherical basis. This increase is determined by the factor $n/l_{\text{eff}} \gg 1$, which is several tens. However, this increase is virtually unobservable, because no highly excited states take place

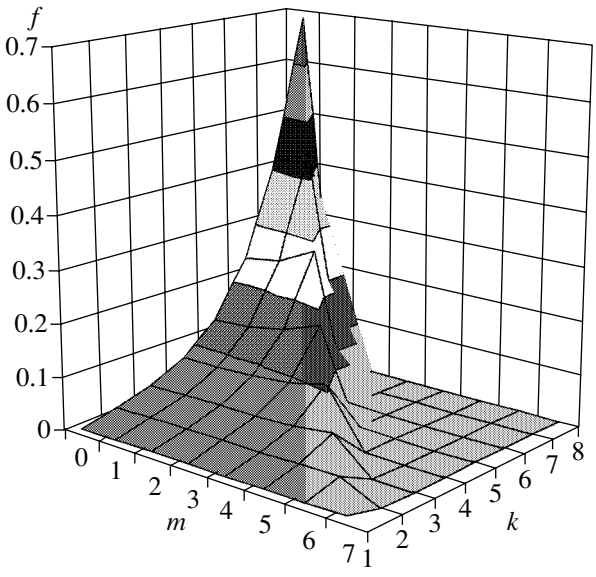


Fig. 5. Same as Fig. 3 for $n = 10$.

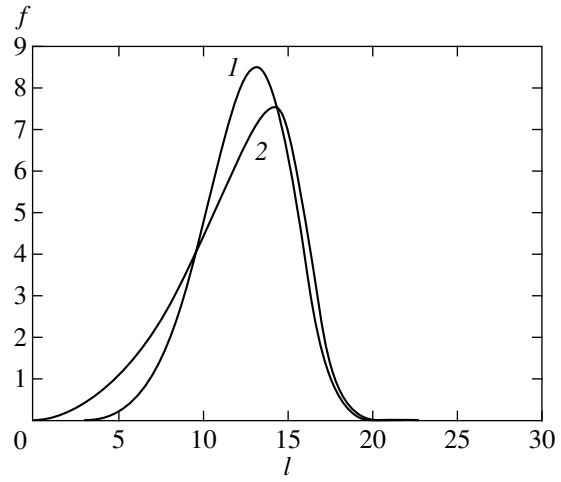


Fig. 6. Direct (1) and cascade (2) terms of the population distribution function for the $n = 30$ level in the spherical basis versus orbital quantum number l for a lithium-like ZnXXVIII ion (in arb. units).

under actual plasma conditions because of their ionization both in the electric field and by secondary collisions with electrons. Nevertheless, the increase in recombination in the field can be significant (see 6, 17]).

The relationship between the direct and cascade populations in the parabolic basis also changes compared to the spherical basis. Thus, the cascade term is of the order of the direct population term even for $n = 30$ in the spherical basis and only for $n = 20$ in the parabolic basis.

Interestingly, the parabolic distribution cannot be obtained from the spherical distribution by directly substituting the characteristic $l_{\min}(n, k, m)$ for the orbital angular momentum. This is because it is nontrivial to introduce the population source in parabolic coordinates (as we noted above).

6. THE STATISTICAL AND DYNAMIC INTENSITIES OF SPECTRAL LINES

Let us determine the statistical and dynamic intensity distributions in parabolic quantum numbers by using the results obtained above. As was pointed out in Section 1, the statistical intensities I^{stat} can be calculated by assuming a statistical population distribution in all parabolic variables. Consider the transitions with a small change in principal quantum number by using the transition intensity distributions in large-scale and small-scale blocks from Section 2. The total line intensity can then be calculated by summing the individual intensities over the large-scale blocks (denoted by the quantum number K), over the small-scale blocks (denoted by the electric quantum number k) with allowance made for their shift in an electric field, and over all magnetic quantum numbers:

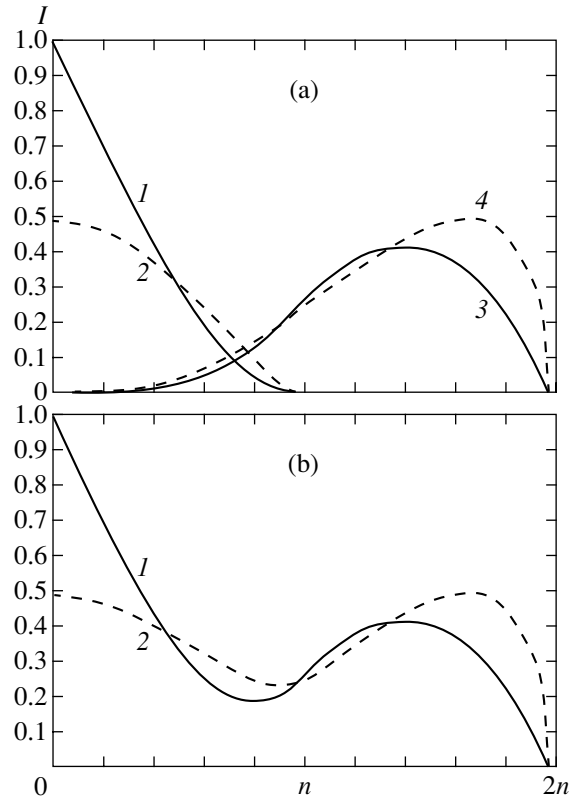


Fig. 7. The intensity distribution $I(\omega)$ (in arb. units) in an electric field F as a function of the frequency shift $\Delta\omega$ [in $(3ea_0/2\hbar)F$] for $H_{n\alpha}$ lines with $n \gg 1$: (a) the separate distributions for the Stark components (*1*, the statistical σ component; *2*, the dynamic σ component; *3*, the statistical π component; *4*, the dynamic π component); (b) the total distribution (*1*, the statistical total intensity; *2*, the dynamic total intensity).

$$I^{\text{stat}}(n, \Delta n) = \sum_{K, k, m} I(K, n, \Delta n, k, m) \times \delta\left(\frac{\Delta\omega}{\omega_F} - Kn - \Delta nk\right). \quad (39)$$

Here, the dimensionless frequency shift $\Delta\omega/\omega_F$ in an electric field F is expressed in units of the typical splitting in the field,

$$\omega_F = 3ea_0F/2\hbar Z.$$

The summation in (39) is performed over all quantum numbers with the same weight, which reflects the equidistribution between the Stark sublevels.

The dynamic intensities I^{dyn} can be calculated by the summation of the intensities of individual transitions multiplied by the normalized distribution function (or population) $f(n, k, m)$ of the initial state, which reflects

the pattern of Stark sublevel population by the radiative cascade:

$$I^{\mathrm{dyn}}(n, \Delta n) = \sum_{K, k, m} I(K, n, \Delta n, k, m) \times \frac{f(n, k, m)}{\sum_{k, m} f(n, k, m)} \delta\left(\frac{\Delta\omega}{\omega_F} - Kn - \Delta nk\right). \quad (40)$$

The results of our calculations of the statistical and dynamic intensities for the $\Delta n = 1$ transitions performed with formulas (39) and (40) by taking into account the results (37) and (38) for the distribution function are shown in Fig. 7a separately for the σ and π line components and in Fig. 7b for the total intensity distribution. We see an appreciable difference between the two types of intensities.

Note that the observed intensities can be obtained by the summation over all magnetic quantum numbers. Clearly, this corresponds to a spherically symmetric distribution of the electric field in equilibrium plasma. At the same time, the intensity of the radiation from an individual atom, as well as its distribution function (sublevel population), significantly depends on the magnetic quantum number, which reflects the conservation of its component along the field direction when the field varies slowly (adiabatically).

The statistical intensities have been commonly used to calculate the line profiles in plasma. As is clear from our analysis, this is by no means always proper.

7. DISCUSSION

We have determined the distribution function (level populations) in parabolic quantum numbers in an electric microfield that corresponds to the population by dielectronic recombination and by a radiative cascade. We also calculated the dynamic intensities of the spectral lines that emerge during radiative transitions from these levels.

One of our main conclusions is that the populations of excited states in an electric field increase compared to the calculations without any field. This may affect the interpretation of the spectroscopic data on absolute spectral line intensity measurements for multiply charged ions in rarefied plasma. Indeed, an increase in state populations, together with line intensities, implies that the ion abundance in the interpretation of observational data in the spherical basis may prove to be significantly overestimated. Actually, the observed increase in intensities is attributable to the plasma microfield.

The difference between the statistical and dynamic intensity distributions is also of great importance in calculating the spectral line shape for which this distribution serves as the starting point (see [2]). As follows from our analysis, the differences between the line profiles can change in moderately dense plasma, where radiative transitions dominate over collisional transitions. However, the calculation of the profiles for specific lines is beyond the scope of this study. Note only that, in our case of highly excited (Rydberg) states, for which the condition (2) is satisfied, the Stark splitting significantly exceeds the Doppler splitting.

The difference between the statistical and dynamic intensities also implies that the population kinetics of atomic states is essentially multidimensional when radiative transitions dominate over collisional transitions in low-density plasma.

ACKNOWLEDGMENTS

We thank G.V. Sholin for valuable remarks and discussions.

This study was supported by the Russian Foundation for Basic Research (project no. 01-02-16305) and the Science Ministry of Russia on the subject "The Development of Methods for Optical Spectroscopy of Ultrahigh Spectral Resolution."

APPENDIX

Below, we give an expression for the relative amplitudes:

$$Z^{\Delta m = 0}_{\Delta n_1 \Delta n_2} \approx \frac{1}{\Delta n}\{\sigma_2 J_{\Delta n_1}(\Delta n\sigma_1) J'_{\Delta n_2}(\Delta n\sigma_2) - \sigma_1 J'_{\Delta n_1}(\Delta n\sigma_1) J_{\Delta n_2}(\Delta n\sigma_2)\}$$

$$= \frac{1}{\Delta n}\left\{-\frac{K}{\Delta n} J_{\Delta n_1}(\Delta n\sigma_1) J_{\Delta n_2}(\Delta n\sigma_2) - \sigma_2 J_{\Delta n_1}(\Delta n\sigma_1)\Delta J_{\Delta n_2 + 1}(\Delta n\sigma_2) + \sigma_1 J_{\Delta n_1 + 1}(\Delta n\sigma_1) J_{\Delta n_2}(\Delta n\sigma_2)\right\}, \quad (A.1)$$

$$X^{\Delta m = \pm 1}_{\Delta n_1 \Delta n_2} \approx \frac{1}{\Delta n} \times \left\{\frac{n\sigma_1\sigma_2}{\sqrt{n_1 n_2}} J_{\Delta n_1}(\Delta n\sigma_1) J_{\Delta n_2}(\Delta n\sigma_2) - \frac{\sqrt{n_1 n_2}}{n} J_{\Delta n_1 \pm 1}(\Delta n\sigma_1) J_{\Delta n_2 \pm 1}(\Delta n\sigma_2)\right\},$$

where $J_p(g)$ is the Bessel function with integer indices p,

$$\sigma_1 = \frac{\sqrt{n_1(n_1+m)}}{n} = \frac{1}{2\sqrt{nn'}}\sqrt{(n+k-1)^2-m^2},$$
$$\sigma_2 = \frac{\sqrt{n_2(n_2+m)}}{n} = \frac{1}{2\sqrt{nn'}}\sqrt{(n-k-1)^2-m^2},$$
$$\Delta n_1 = \frac{1}{2}(\Delta n + K \mp \Delta m), \quad \Delta n_2 = \frac{1}{2}(\Delta n - K \mp \Delta m), \tag{A.2}$$
$$K = (n_1 - n_1') - (n_2 - n_2') = \Delta n_1 - \Delta n_2,$$
$$s = n_1 - n_2, \quad s' = n_1' - n_2' = k.$$

It is easy to notice that the π and σ components correspond to even and odd $\Delta n + K$, respectively. It is also easy to show that the terms with small K mainly contribute to the intensities. $\mp$

Expanding the Bessel functions in the expression for the squares of the relative amplitudes (A.1) in the small parameter $\Delta n^2/4nn'$, we can easily obtain simple universal expressions for the relative intensities of the π and σ components for transitions with any $\Delta n \ll n$:

$$I^{\pi}_{K=0}(m) \approx Z^2_{K=0} \approx \left(\frac{1}{2}\right)^{2(\Delta n+1)}\left(\frac{\Delta n^2}{4nn'}\right)^{\Delta n} \times \frac{[(n^2-k^2)^2 - 2m^2(n^2+k^2) + m^4]^{\Delta n/2}}{(\Delta n/2)!^2((\Delta n+1)/2)!^2}\left(\frac{k}{n}\right)^2, \tag{A.3}$$

$$I^{\pi}_{K\neq 0}(m) \approx Z^2_{K\neq 0} \approx \left(\frac{1}{2}\right)^{2\Delta n}\left(\frac{\Delta n^2}{4nn'}\right)^{\Delta n}\left(\frac{K}{\Delta n^2}\right)^2 \times \frac{[(n^2+k^2)^2 - m^2]^{|\Delta n+K|/2}[(n^2-k^2)^2 - m^2]^{|\Delta n-K|/2}}{(|\Delta n-K|/2)!^2(|\Delta n-K|/2)!^2}, \tag{A.4}$$

$$I^{\sigma}_K(m) \approx (X^{\Delta m=+1})^2 + (X^{\Delta m=-1})^2 \approx \left(\frac{1}{2}\right)^{2(\Delta n-3)}\left(\frac{\Delta n^2}{4nn'}\right)^{\Delta n}\frac{1}{\Delta n^4} \times \frac{[(n+k)^2 - m^2]^{|\Delta n-1+K|/2}[(n-k)^2 - m^2]^{|\Delta n-1-K|/2}}{(|\Delta n-1+K|/2)!^2(|\Delta n-1-K|/2)!^2}. \tag{A.5}$$

REFERENCES

1. H. A. Bethe and E. E. Salpeter, *Quantum Mechanics of One- and Two-Electron Atoms* (Academic, New York, 1957; Fizmatgiz, Moscow, 1960).
2. H. R. Griem, *Spectral Line Broadening by Plasmas* (Academic, New York, 1974; Mir, Moscow, 1978).
3. I. I. Sobelman, L. A. Vainshtein, and E. A. Yukov, *Excitation of Atoms and Broadening of Spectral Lines* (Nauka, Moscow, 1979; Springer-Verlag, Berlin, 1981).
4. A. B. Kukushkin and V. S. Lisitsa, Zh. Éksp. Teor. Fiz. **88**, 1570 (1985) [Sov. Phys. JETP **61**, 937 (1985)].
5. D. A. Varshalovich, A. N. Moskalev, and V. K. Khersonskii, *Quantum Theory of Angular Momentum* (Nauka, Leningrad, 1975; World Sci., Singapore, 1980).
6. Y. Hanh, Rep. Prog. Phys. **60**, 691 (1997).
7. D. R. Herrick, Phys. Rev. A **12**, 1949 (1975).
8. L. A. Bureeva and V. S. Lisitsa, *Perturbed Atom* (IzdAt, Moscow, 1997).
9. S. A. Gulyaev, Astron. Zh. **53**, 1010 (1976) [Sov. Astron. **20**, 573 (1976)].
10. M. Born, *The Mechanics of the Atom* (G. Bell and Sons, London, 1927; Fizmatgiz, Moscow, 1936).
11. S. T. Belyaev and G. I. Budker, in *Plasma Physics and Problem of Controlled Thermonuclear Reactions*, Ed. by M. A. Leontovich (Akad. Nauk SSSR, Moscow, 1958), Vol. 3, p. 41.
12. L. D. Landau and E. M. Lifshitz, *Course of Theoretical Physics*, Vol. 2: *The Classical Theory of Fields* (Nauka, Moscow, 1988; Pergamon, Oxford, 1975).
13. A. A. Kamenski and V. D. Ovsiannikov, J. Phys. B **33**, 491 (2000).
14. I. L. Beĭgman, L. A. Vaĭnshteĭn, and B. N. Chichkov, Zh. Éksp. Teor. Fiz. **80**, 964 (1981) [Sov. Phys. JETP **53**, 490 (1981)].
15. V. I. Kogan, A. B. Kukushkin, and V. S. Lisitsa, Phys. Rep. **213**, 1 (1992).
16. V. A. Bazylev and M. I. Chibisov, Usp. Fiz. Nauk **133**, 617 (1981) [Sov. Phys. Usp. **24**, 276 (1981)].
17. L. A. Bureyeva, T. Kato, V. S. Lisitsa, and C. Namba, Phys. Rev. A **65**, 032702 (2002).

Translated by V. Astakhov

Journal of Experimental and Theoretical Physics, Vol. 95, No. 4, 2002, pp. 673–681.
Translated from Zhurnal Éksperimental'noĭ i Teoreticheskoĭ Fiziki, Vol. 122, No. 4, 2002, pp. 778–788.

PLASMA, GASES

The Dynamics of Macroparticles in a Direct Current Glow Discharge Plasma under Microgravitation Conditions

A. P. Nefedov[a], O. S. Vaulina[a], O. F. Petrov[a, *], V. I. Molotkov[a], V. M. Torchinskiĭ[a], V. E. Fortov[a], A. V. Chernyshev[a], A. M. Lipaev[a], A. I. Ivanov[b], A. Yu. Kaleri[b], Yu. P. Semenov[b], and S. V. Zaletin[c]

[a]*Institute of Thermal Physics of Extremal States, Russian Academy of Sciences, Moscow, 127412 Russia*
[b]*Korolev Space Rocket Energiya Corporation, Korolev, Moscow oblast, 141070 Russia*
[c]*Gagarin Center for Training Cosmonauts, Zvezdnyi gorodok, Moscow oblast, 141160 Russia*
**e-mail: ipdustpl@redline.ru*
Received March 29, 2002

Abstract—The dynamics of large-sized (70–180 μm) spherical bronze particles in a direct current glow discharge plasma was studied experimentally under microgravitation conditions. The temperatures, velocities, pair correlation functions, and self-diffusion coefficients of macroparticles were measured at various discharge currents. The charges of dust particles (on the order of 10^6 *e*) corresponded to high surface potentials of about 30–40 V. The experimental data were in close agreement with the simulation data on Yukawa systems with weak screening of dust charges. The influence of macroparticles on equilibrium ionization in a dense dust cloud was considered. © *2002 MAIK "Nauka/Interperiodica".*

1. INTRODUCTION

A dust plasma is a partially ionized gas containing charged particles of matter (dust) of micron dimensions. The size of dust particles allows them to be recorded on a videotape, which significantly simplifies the use of direct contactless methods for their diagnostics. A dust plasma is therefore a convenient experimental model for studying various transport phenomena in systems of interacting particles, such as charging and mass transfer and the formation of dust structures and complex oscillatory modes [1–10]. In recent years, considerable attention has been given to experimental studies of dust plasmas under microgravitation conditions [6–10]. Such experiments allow a wide range of phenomena that cannot be observed in laboratories on the Earth to be studied [6–8]. These are photoemission charging of aerosols in the upper layers of the atmosphere, ambipolar diffusion, the dynamics of large-sized (larger than 100 μm) dust particles in a plasma, etc. Recent experiments performed by the Russian–German team on the International Space Station revealed the occurrence of several new phenomena in an RF discharge plasma, such as dust beats, the formation of complex crystalline lattices and opposite charging of macroparticles. No analogues of such processes are observed under usual laboratory conditions [10].

In this work, we describe some of the earliest experiments on the dynamics of macroparticles under microgravitation conditions performed aboard the Mir space station during the 28th orbital expedition. The experiments were performed for large-sized (hundred micron) bronze particles in a direct current glow discharge plasma.

2. EXPERIMENTAL

The experimental unit is schematically shown in Fig. 1. The main element of the working chamber was a gas discharge tube of radius $R_t \approx 1.6$ cm filled with neon (Ne) to a pressure of $P = 1$ Torr. The distance between a plane anode and a cathode was 28 cm. An insulated electrode was mounted at a distance of 4.5 cm from the anode. The electrode was made as two steel grids (wire 60 μm in diameter) with 150 × 150 μm meshes, and the distance between the grids was 1 cm. During experiments, the electrode was under a floating potential and prevented negatively charged macroparticles from escaping to the anode. Bronze spherical particles (fraction of particles 70–180 μm in diameter, mean radius $\langle a_p \rangle = 62.5$ μm, density of the material $\rho \approx 8.2$ g/cm^3) were placed between the grid electrode and the cathode.

The diagnostics of macroparticles was performed with the use of a plane laser beam ("laser knife" about 300 μm wide, wavelength 0.67 μm) and additional illumination of the dust cloud by an incandescent lamp. In the latter case, the number of detected particles was determined by the depth of field of the video system, which allowed us to keep track of particle positions for a long time sufficient for analyzing their transport characteristics. The image was recorded on a videotape at a 50-s^{-1} frame frequency. The video data were processed using a special program for identifying the displace-

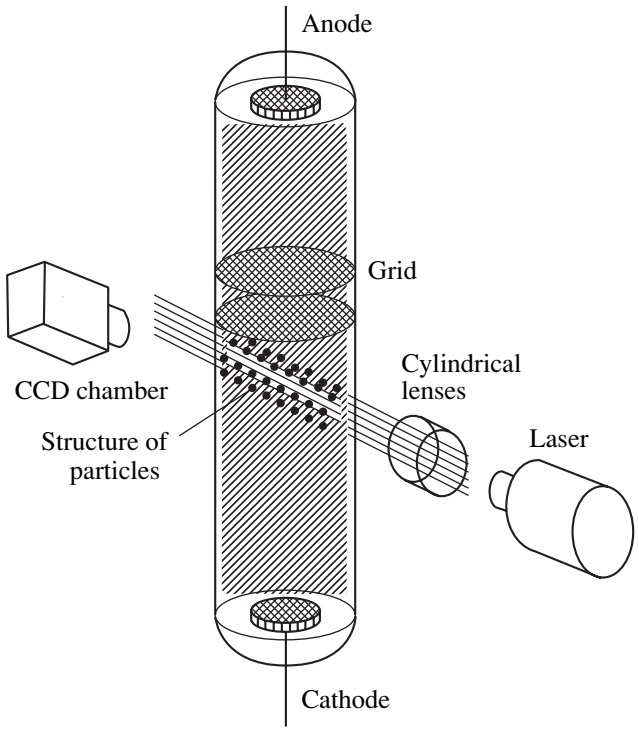


Fig. 1. Schematic drawing of the experimental unit.

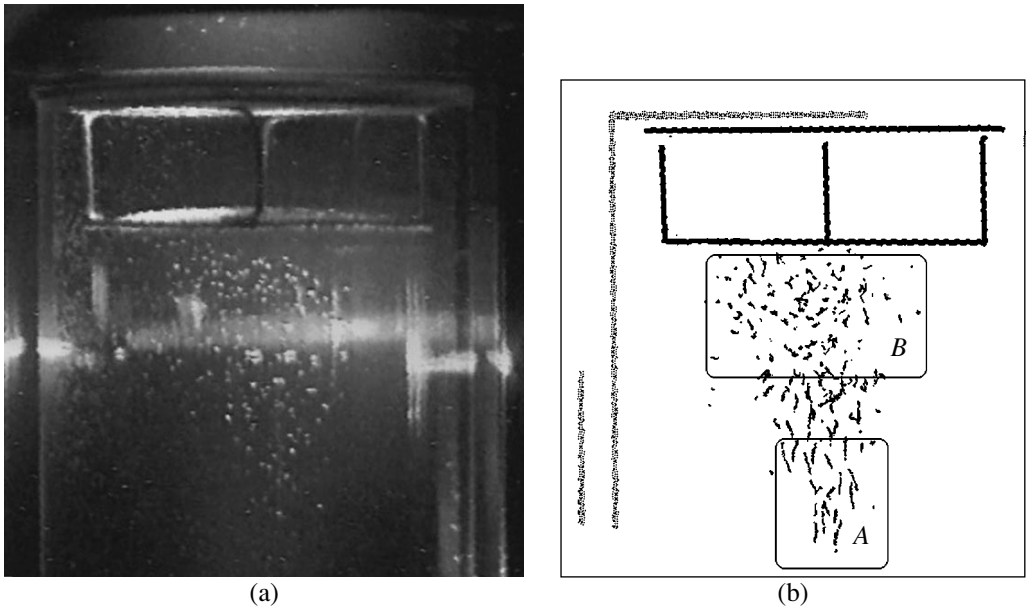
ments of separate particles in the visual field of video recording.

Experiments were performed at discharge currents (I = 0.1–0.8 mA) varied by the current source. Bronze particles were initially situated on tube walls. For this reason, the system was subjected to a dynamic action (pushed) after switching on a discharge with a given current I to shake off particles from tube walls. After the dynamic action, bronze particles moved toward insulated grid electrodes, in the vicinity of which ordered structures were formed (Figs. 2a, 2b). The discharge was then "quenched," the particles relaxed to the initial state (went away to tube walls), and the experiment was repeated at a new gas discharge current value.

Plasma concentration $n^0 \approx n_e^0 \approx n_i^0$ [here, $n_{e(i)}^0$ is the concentration of electrons (ions) in the absence of dust particles] in the region of the positive discharge column can be estimated from discharge current density measurements, $j = I/\pi R_t^2 \approx$ 12–96 μA/cm^2; this value should remain constant over the whole tube length:

$$j = e\mu_e E n_e^0, \quad (1)$$

where μ_e is the mobility of electrons [for neon, $\mu_e P \approx 1.5 \times 10^6$ Torr cm^2/(V s)]. In tubes with radii of 1–2 cm and $P \approx 1$ Torr, field, of the positive column in a normal glow discharge in neon is $E \approx 1$ V/cm, the normal current density is $j_n \approx$ 5–6 μA/cm^2, and the mean electron energy characterizing electron temperature T_e is about 3–5 eV [11]. Under our experimental conditions, discharges occurred in an anomalous mode (which followed from both the recorded current densities and the current–voltage characteristics), and the electric field strength obeyed the equation $E \propto j^{1/2}$ [11]. At E = 1.5 V/cm and j = 12 μA/cm^2, the mean plasma concentration calculated by (1) was $n^0 \approx (5–14) \times 10^7$ cm^{-3} ($j \approx$ 12–96 μA/cm^2). Taking into account the Bessel radial profile, the plasma density along the tube axis, where



(a) (b)

Fig. 2. Video images of (a) gas-discharge tube and (b) particle trajectories from the discharge positive column toward insulated grid electrodes.

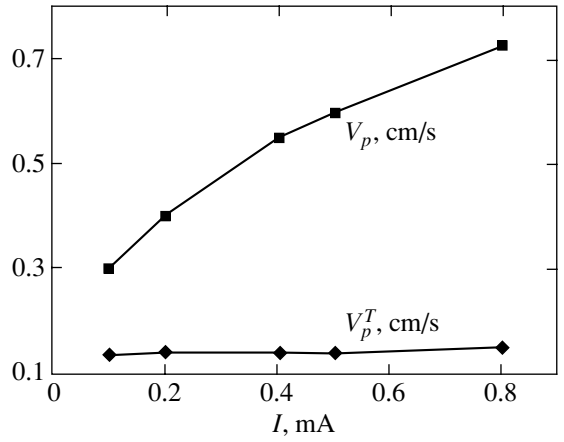


Fig. 3. Mean thermal V_p^T and drift V_p velocities of particles from the positive column region (Fig. 2b, field *A*) to grid electrodes at various discharge currents *I*.

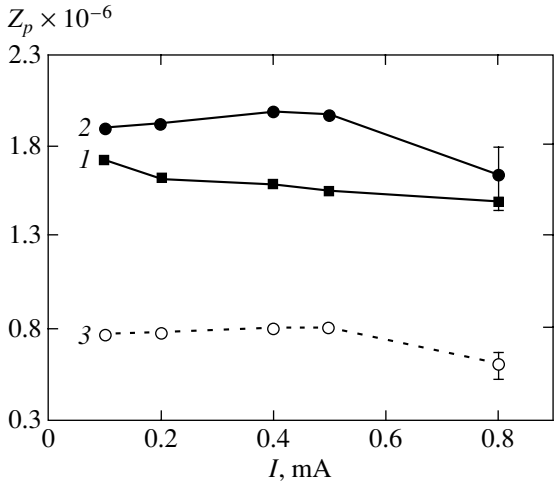


Fig. 4. Particle charge Z_p as a function of discharge current I obtained from (*1*) equation of motion (2) for $\langle a_p \rangle$ = 62.5 μm, (*2*) diffusion measurements by (5) and (6) for a_p = 62.5 μm, and (*3*) diffusion measurements for a_p = 35 μm.

the parameters of dust particles were measured, reached values on the order of $2.4n^0$ [11]. The further estimates will therefore be based on the assumption that, in the absence of particles, the $n^0_{e(i)}$ concentration of electrons (ions) varied in the range from 10^8 to 3×10^8 cm^{-3} under current density variations from 12 to 96 μA/cm^2.

3. MEASUREMENT RESULTS AND THEIR ANALYSIS

3.1. The Determination of Dust Charges from the Drift Velocity of Macroparticles

The mean drift velocity V_p of particles from the positive column (Fig. 2b, field *A*) to the grid electrode is shown in Fig. 3 as a function of discharge current *I*. As the velocity of particle motions in the measurement volume remained virtually constant, their charges eZ_p could be found from the equation of motion [8]

$$m_p dV_p/dt = -\nu_{fr} m_p V_p + EeZ_p \equiv 0, \qquad (2)$$

where m_p is the mass of the particle and ν_{fr} is the friction coefficient of dust particles, which determines the frequency of their collisions with surrounding gas neutrals [12, 13],

$$m_p \nu_{fr} = 6\pi a_p \eta (1 - l_g(1 - \exp(-2a_p/l_g))/2a_p). \quad (3)$$

Here, l_g [μm/Torr] ≈ 125/*P* is the free path of neon neutrals, and $\eta \approx 3.17 \times 10^{-4}$ g/(cm s) is the viscosity of neon at $l_g \ll a_p$. Setting *E* [V/cm] = $1.5\sqrt{j\ [\mu\text{A/cm}^2]/12}$ and $\langle a_p \rangle$ = 62.5 μm in (2) and (3), we find that particle charge $Z_p \approx 10^6$ is virtually independent of discharge current (Fig. 4, curve *1*) and corresponds to fairly high surface potentials $\varphi_s = eZ_p/a_p \approx$ 35–37 V. The errors in the φ_s values obtained by this method are determined by the availability of *a priori* information on the field and friction coefficient values in (2).

3.2. The Formation of Liquid Dust Structures

The characteristic dimensions of dust clouds formed in the vicinity of the grid electrode (Fig. 2b, field *B*) were about 2 cm in the radial direction and 0.7–1.3 cm from the grid edge along the tube axis. Pair correlation functions $g(r)$ for various discharge currents *I* are shown in Fig. 5. It is easy to see that, at low *I* values, the degree of ordering of particles in the dust cloud

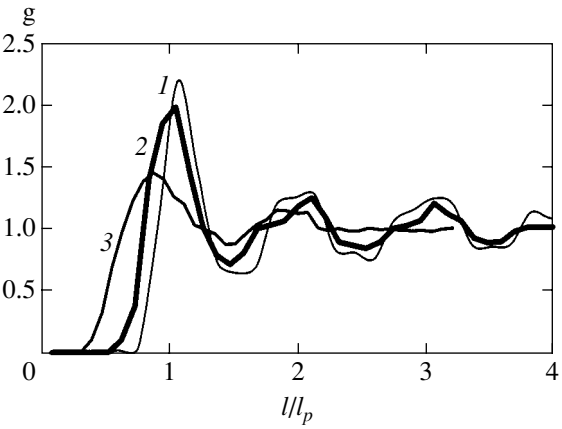


Fig. 5. Pair correlation functions g versus l/l_p for dust structures formed close to grid electrodes (Fig. 2b, field *B*) at discharge currents *I* of (*1*) 0.1, (*2*) 0.4, and (*3*) 0.8 mA.

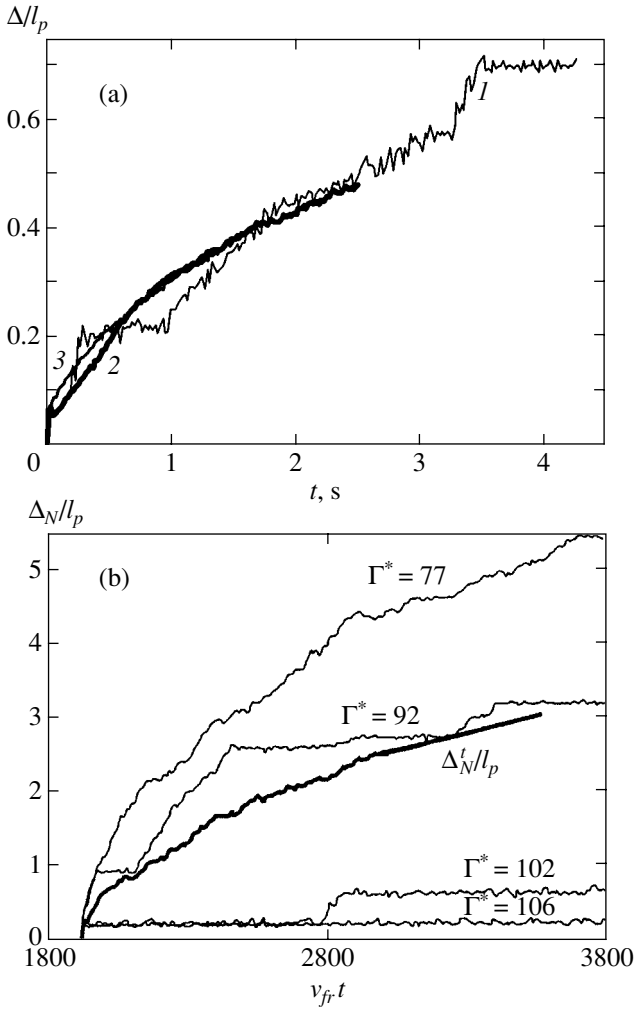


Fig. 6. (a) Measurement time dependences of Lindemann parameter $\delta_c = \Delta(t)/l_p$ (I = 0.4 mA) obtained using various averaging techniques: (*1*) $\Delta(t) \equiv \Delta_N(t)$, N_p = 16; (*2*) $\Delta(t) \equiv \Delta^t_N(t)$, N_p = 16; and (*3*) $\Delta(t) \equiv \Delta^t_N(t)$, N_p = 96 and (b) "jumps" in simulated Yukawa systems (thick line corresponds to $\Delta^t_N(t)$ for Γ^* = 92).

increases as their concentration n_p grows. An increase in n_p is accompanied by a decrease in $l_{\max}$,

$$l_{\max} \approx l_p = n_p^{-1/3},$$

where $l_{\max}$ is the position of the main $g(r)$ maximum and l_p is the mean interparticle distance. Because of a fairly small number of particles in the laser knife plane, a quantitative analysis of the experimental $g(r)$ functions is virtually impossible. Nevertheless, the presence of fairly well-defined secondary $g(r)$ maxima is evidence of strong interparticle interactions and weak screening of particles (screening parameter $\kappa = l_p/\lambda < 1$, where λ is the screening length).

According to the Lindemann criterion, a solid phase melts if the ratio between the root of the mean-square displacement Δ_0 of particles from their equilibrium positions and the mean interparticle distance l_p reaches about 0.15 [14]. In experiments, particle displacements are usually measured from the center of mass of the system, that is, $\Delta = \sqrt{2}\Delta_0$, and the Lindemann parameter at the melting curve is therefore $\delta_c = \Delta/l_p \approx 0.21$. The $\delta_c(t) = \Delta_N(t)/l_p$ time dependence [here, $\Delta_N(t) = \sqrt{\langle l(t) - l(0)\rangle^2_N}$, $l(t)$ is the displacement of a separate particle, and $\langle\rangle_N$ denotes averaging over the ensemble of N particles] for 16 particles (N_p = 16), which remain in the visual field of the video system for about 4.5 s, is shown in Fig. 6a. Note that, during some time intervals, the $\Delta_N(t)/l_p$ ratio remains constant and corresponds to the Lindemann criterion. When the observation time was decreased to about 1 s, the number of identified trajectories increased to N_p = 70–120 at the expense of particles that experienced transitions between "settled" states from one cloud region to another. Averaging of particle displacements $\Delta_N(t)$ over the ensemble was then close to time averaging,

$$\Delta^t_N = \sqrt{\langle\langle l(t) - l(0)\rangle_N\rangle^2_t}$$

(Fig. 6a). Here, $\langle\rangle_{N,t}$ denotes averaging over the ensemble and time, respectively. A similar picture was observed in numerical simulations of the dynamics of macroparticles with the Yukawa interaction potential $\varphi = eZ_p\exp(-l/\lambda)/l$ in a strongly nonideal dust liquid [15]. "Jumps" observed in simulated systems are illustrated by Fig. 6b, where the difference between the ensemble $\Delta_N(t)/l_p$ and time Δ^t_N/l_p averages close to the crystallization curve of the system, where the Γ^* normalized nonideality parameter tends to 102, is shown. This difference vanishes as Γ^* decreases, and the system under study becomes ergodic.

The Γ^* normalized parameter value determines the degree of ordering and the dynamics of particles in nonideal Yukawa systems [15, 16],

$$\Gamma^* = (1 + \kappa + \kappa^2/2)\exp(-\kappa)\Gamma, \qquad (4)$$

where

$$\Gamma = \frac{(eZ)^2 n_p^{1/3}}{T_p}.$$

Here, T_p is the temperature of macroparticles in energy units.

3.3. Diffusion of Macroparticles

The self-diffusion coefficient of macroparticles D_p can be found from measured mean-square displacements Δ^t_N as

$$D_p = \lim_{t\to\infty} D(t) = \langle\langle l(t) - l(0)\rangle^2_N\rangle_t/6t. \qquad (5)$$

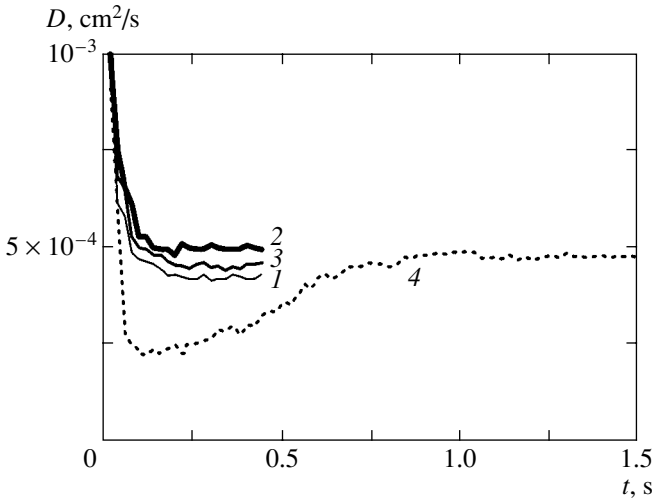


Fig. 7. Dependences $D_p(t)$ in a discharge with $I = 0.4$ mA for particles in the visual field of the video system during the (*1*) first, (*2*) second, and (*3*) third seconds of observation and (*4*) for 16 particles in the visual field for 4.5 s.

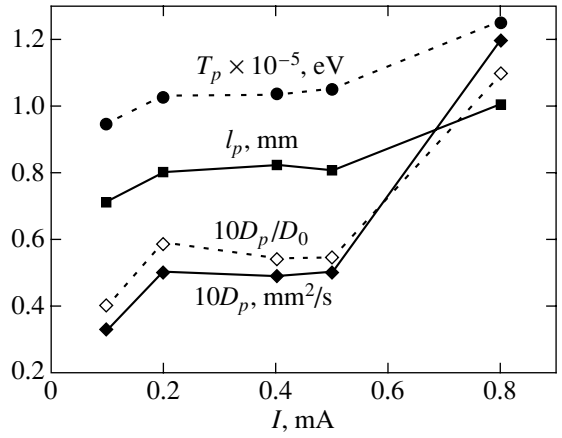


Fig. 8. Mean interparticle distance l_p, temperature T_p, and diffusion coefficients D_p of macroparticles and the ratio between the D_p coefficient and the D_0 diffusion coefficient of noninteracting particles as functions of discharge current I.

The $D_p(t)$ values measured in an $I = 0.4$ mA discharge are shown in Fig. 7 for various numbers of particles $N_p = 70$–120 in the visual field of the video system during the first, second, and third seconds of observations. The same figure contains the $D_p(t)$ dependence for 16 macroparticles in the visual field during 4.5 s of observations. It is easy to see that the $D_p = \lim_{t\to\infty} D(t)$ coefficient is virtually independent of the size of the sample of particles (N_p) in spite of different behaviors of the systems at initial time moments.

The relation between the D_p self-diffusion coefficient given by (5) and the Γ^* value (4) in strongly correlated ($\Gamma^* > 30$) liquid Yukawa systems can be written as

$$D_p \approx \frac{T_p\Gamma^*}{12\pi(\omega_l + \nu_{fr})m_p}\exp\left(-3\frac{\Gamma^*}{\Gamma_c}\right), \qquad (6)$$

where $\Gamma^* \approx 102$ is the Γ^* parameter value at the crystallization point [16] and

$$\omega_l = eZ\sqrt{\frac{n_p}{\pi m_p}}\sqrt{1+\kappa+\kappa^2/2}\exp\left(-\frac{\kappa}{2}\right) \equiv \sqrt{\frac{\Gamma^* T_p}{\pi l_p^2 m_p}}$$

is the normalized dust frequency. The error involved in approximation (6) sharply decreases as Γ^* increases; it amounts to about 30% for $\Gamma^* \approx 30$ and less than 3% at $\Gamma^* > 50$ [15]. Equation (6) allows Γ^* to be determined fairly easily from the results obtained in measurements of the mean interparticle distance l_p, temperature T_p, and the D_p diffusion coefficient of macroparticles. The discharge current dependences of these parameters (l_p, T_p, and D_p) and of the ratio between D_p and the diffusion coefficient of noninteracting particles $D_0 = T_p/m_p\nu_{fr}$ are shown in Fig. 8; the D_0 value was found to be more than an order of magnitude larger than the D_p value calculated by (5). The temperature of macroparticles was established by finding the Maxwell distribution that best described the recorded spectrum of instantaneous velocities of particles obtained by analyzing video records. A similar procedure for determining temperature was described in [2, 7]. The mean velocities of chaotic movement of particles were about 1 mm/s (see Fig. 3). Note that this procedure could underestimate the mean chaotic energy of particles (kinetic temperature) if the frame frequency w (50 s^{-1} in our experiments) did not satisfy the condition $w \gg \nu_{fr}$ [2].

3.4. The Nonideality Parameter and Macroparticle Charges

The Γ^* normalized parameter values reproduced by (6) from diffusion coefficient measurements are shown in Fig. 9 for the mean ($\langle a_p\rangle = 62.5$ μm) and minimal [$(a_p)_{\min} = 35$ μm] particle sizes. Note that Γ^* determinations are only based on measurement results. The ν_{fr} friction coefficient value depends on the radius of particles, but it does not significantly influence the Γ^* value because $\nu_{fr} < \omega_l$ under the experimental conditions.

If the effects of screening macroparticles in a plasma–dust system are significant, an analysis of diffusion of macroparticles on the assumption of negligibly weak screening ($\Gamma^* = \Gamma$, $\kappa < 1$) gives a minimal estimate of the charge Z_p of dust macroparticles irrespective of the form of the interparticle interaction potential. Parameter $\Gamma^* \approx \Gamma$ measurements for $\kappa = 0$ allow surface potentials at $I \le 0.5$ mA to be determined. This gives $\varphi_s \approx 42$ V for medium-sized particles $\langle a_p\rangle = 62.5$ μm

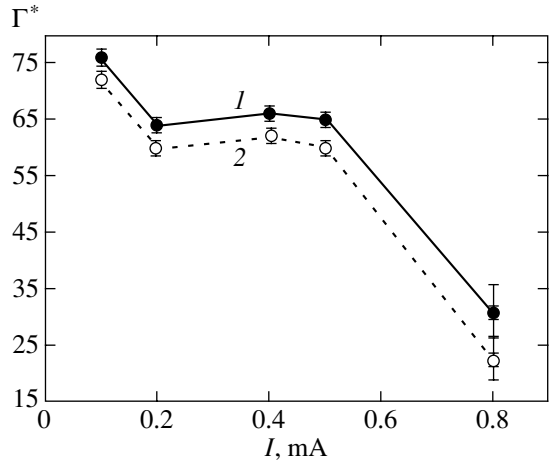


Fig. 9. Normalized nonideality parameters Γ* reproduced from diffusion coefficient measurements for particles of different sizes: (*1*) $\langle a_p \rangle$ = 62.5 μm and (*2*) $(a_p)_{\min}$ = 35 μm; error intervals for reproducing Γ* by (4) are shown.

and $\varphi_s \approx 30$ V for particles of radius $(a_p)_{\min} = 35$ μm (see Fig. 4). For $I = 0.8$ mA, the recorded surface potential value is lower by approximately 30%. The error in determining the charge of macroparticles is then substantially larger, because, at low Γ* < 35 values, (6) is fairly inaccurate (see Fig. 9). The obtained macroparticle charge estimates are in agreement with the results of determining Z_p from equation of motion (2). We can therefore hope that the surface potentials $\varphi_s \approx 30$–42 V ($I \le 0.5$ mA) and $\varphi_s \approx 22$–30 V ($I = 0.8$ mA) correctly reproduce real Z_p charge values, and, accordingly, the assumption of weak screening ($\kappa < 1$) of macroparticles corresponds well with the conditions existing in the dust system under study.

3.5. Screening of Macroparticles

The suggestion of weak screening ($\kappa < 1$) of dust charges in our experiments is indirectly substantiated by the behavior of correlation functions (Fig. 5) and is in agreement with numerical simulation results. For instance, it was shown in [17] that screening of isolated particles of size $a_p \approx 2\lambda_i - \lambda_e$, where $\lambda_{i(e)}$ was the Debye ionic (electronic) radius, was determined by the λ_e value at the distances

$$l < l_D \approx (5\text{–}7)\lambda_i$$

from the surface of the particle. As the size of particles increases to $a_p > \lambda_e$, the effective screening length λ becomes substantially larger than the λ_e value.

According to the condition of electroneutrality of the plasma–dust system,

$$Z_p n_p + n_e = n_i, \tag{7}$$

where $n_{e(i)}$ is the concentration of electrons (ions) in the dust cloud. Note that condition (7) is necessary for the confinement of negatively charged particles in the absence of other nonelectrostatic forces that can compensate mutual repulsion of dust particles. The $n_{e(i)}$ concentration of electrons (ions) in a dust cloud under the conditions of our experiments can therefore be estimated as

$$n_e < Z_p n_p \approx (2\text{–}4) \times 10^9 \text{ cm}^{-3} < n_i.$$

This gives $\lambda_i < 20$–30 μm and $\lambda_e > 200$–300 μm for ions with temperature $T_i \approx 0.03$ eV and electrons with $T_e > 3$ eV. The obtained λ_e value may be strongly underestimated because of a possible increase in the temperature of electrons in the dust cloud and because of the law of current density conservation in a gas-discharge tube, $j = \text{const}$ [see (1)]. It follows that, if electric field variations are ignored, the n_e mean concentration of electrons in the dust cloud should be comparable in order of magnitude with the n^0 concentration of plasma in the absence of particles, that is, $n_e \approx 10^8$ cm^{-3}. The screening parameter is then $\kappa = l_p/\lambda \le 1$, because $l_p \approx 7200$–1000 μm. Note also that the surface potential ceases to obey the exponential law as the distance from a macroparticle increases to $l > l_D \approx (5\text{–}7)\lambda_i \approx 100$–200 μm because of effective electron–ion recombination on the surface of the macroparticle. At such distances, the potential exhibits a power dependence on l [17, 18]. This effect weakens screening of large-sized particles in the dust cloud when the l_p mean interparticle distance exceeds l_D [19].

3.6. The Kinetics of Charging Dust Particles

Equilibrium charge Z_p of dust particles in a plasma is determined by the equation

$$\frac{dZ_p}{dt} = \sum_j I_j \equiv 0, \tag{8}$$

where the summation is over all flows I_j of charged particles absorbed or emitted by the dust particle. Most of the theoretical models of charging particles of size $a < l_{e(i)}$ [$l_{e(i)}$ is the free path of electrons (ions)] are based on the restricted orbit approximation (OML theory). In a gas-discharge plasma, where emission processes are of no significance, the charge of a macroparticle is negative because of a higher mobility of electrons. OML calculations of flows $I_{e(i)}$ of electrons (ions) per negatively charged spherical particle yield [20]

$$I_e = \pi a_p^2 n_e v_{T_e} \exp(e\varphi_s/T_e), \tag{9}$$

$$I_i = \pi a_p^2 n_i v_i \left(1 - \frac{2e\varphi_s}{m_i v_i^2}\right), \tag{10}$$

where $T_{e(i)}$, $m_{ie(i)}$, and $n_{e(i)}$ are the temperature, mass, and concentration of electrons (ions) in the dust cloud, respectively; $v_{T_e} = \sqrt{8T_e/\pi m_e}$ is the thermal velocity

of electrons; and v_i is the velocity of ions, which, depending on the conditions, can be set equal to either the thermal velocity ($v_i \equiv \sqrt{8T_i/\pi m_i}$) or the directed velocity ($v_i \equiv v_s$). Simultaneously solving (7)–(10) yields the equilibrium charge of particles in the form

$$\exp(-z) = \frac{v_i}{v_{T_e}}\left(1 + z\frac{2T_e}{v_i^2 m_i}\right)(1 + z\chi), \qquad (11)$$

where

$$\chi = \frac{Z_p n_p}{z n_e} \qquad (12)$$

and the z parameter is proportional to the ratio between the surface potential φ_s of the dust particle and the temperature of electrons,

$$z = \frac{e\varphi_s}{T_e} \equiv \frac{|Z_p|e^2}{a_p T_e}. \qquad (13)$$

The z value is determined by buffer gas ions. For rare gases such as neon and argon, $z \approx 2$–4 for most of the experimental conditions in a direct current glow discharge [21, 22]. Note that, in the presence of large-sized particles with $a_p > 10$–20 μm, the l_i free path of ions (Ar, Ne) becomes comparable with or smaller than a_p at room temperature [11] and typical buffer gas pressures $P \approx 1$ Torr. This decreases flow I_i of ions onto the surface of a dust particle [Eq. (10)] and, therefore, increases φ_s and z [Eq. (13)]. The corresponding effect was observed experimentally. For instance, it was shown in [8] that the surface potential of a particle grew approximately from 9 to 30 V as the particle radius increased from 2 to 7 μm in neon at $P \approx 0.5$–1.5 Torr. This corresponded to changes in z approximately from 3 to 10 at $T_e = 3$ eV. In our experiments, the mean size of particles ($\langle a_p \rangle = 62.5$ μm) was much larger than in [8]; we could therefore expect that z should be close to 10. It follows that the obtained $\varphi_s \approx 30$–42 V values are in close agreement with the results of measurements performed in [8]. It should, however, be borne in mind that the z and φ_s values decrease in a dense dust cloud with $\chi > 1$ [see (12)] [22]. This decrease can in part be compensated by an increase in electronic temperature under the conditions of changes in equilibrium ionization processes considered below.

Note also that high macroparticle charges $Z_p \approx 10^6$ at their concentrations $n_p \approx 10^3$ cm^{-3} do not correspond with the suggestion of the electroneutrality of the system, Eq. (7), if we assume that the presence of dust particles has no significant influence on discharge conditions and the concentration of ions in the dust cloud is comparable with their concentration in the absence of macroparticles, $n_i \approx n_i^0 \approx 10^8$ cm^{-3}. Such a discrepancy has already been observed in laboratory conditions on the Earth. For instance, heavy glass particles hovered in weak fields (about 1–12 V/cm) of glow discharge strata and positive columns in neon [19], which led the authors of [19] to suggest either the existence of higher charges Z_p or a sharp increase in electric field E caused by a change in the discharge conditions in the dust cloud.

3.7. The Influence of Macroparticles on Equilibrium Ionization

The influence of macroparticles on kinetic processes in a glow discharge can be substantial, because a considerable decrease in the concentration n_e of electrons as a consequence of their effective loss on dust cloud particles results in an increase in electric field strength ($j = \text{const} \longrightarrow En_e = \text{const}$). The mean electron energy then increases, which causes an increase in ionization frequency ν_i, which in turn increases the concentration of electrons to its equilibrium value.

As glow discharges in rare gases are usually controlled by ambipolar diffusion (plasma recombination on gas-discharge tube walls), this sequence is only observed if the electron loss frequency ν_{ep} on dust cloud particles is comparable with or much larger than the ν_{ab} frequency of diffusion electron loss [11],

$$\nu_{ab} \approx (2.4)^2 D_a/R_{\text{tr}}^2, \qquad (14)$$

where $D_a \approx \mu_i T_e/e$ is the ambipolar diffusion coefficient at $T_e \gg T_i$ and $\mu_i \approx 3200$ cm^2/(V s) is the mobility of singly charged ions in neon at $P = 1$ Torr.

The ν_{ep} electron loss frequency on dust cloud particles can be estimated as

$$\nu_{ep} \approx \sqrt{8\pi T_e/m_e}\, n_p a_p^2 \exp(-z). \qquad (15)$$

Substituting the conditions of our experiments [$a_p \equiv \langle a_p \rangle = 62.5$ μm and $n_p \approx (1$–$3) \times 10^3$ cm^{-3}] into (15), we find that, at $z = 3$–4, the ν_{ep} electron loss frequency on dust cloud particles far exceeds the ν_{ab} diffusion loss frequency ($\nu_{ep}/\nu_{ab} > 10$ at $T_e < 15$ eV). We can therefore assume that discharge conditions in our experiments are determined by plasma recombination on the surface of particles.

Under steady state conditions, the rate of electron loss should be compensated by the rate of particle creation, and, accordingly, the ionization frequency ν_i should coincide with the electron loss frequency. If the shape of the spectrum of electron velocities is close to a Maxwell distribution, the temperature of electrons is related to the ionization frequency as [11]

$$\nu_i \approx \gamma C_i n_g \sqrt{8T_e/\pi m_e}(W + 2T_e)\exp(-W/T_e), \qquad (16)$$

where $W \approx 21.6$ eV is the ionization potential of neon, $C_i \approx 1.6 \times 10^{-18}$ cm^2/eV is the experimental ionization constant of neon, and γ is the coefficient that takes into account the depletion of the spectrum of fast electrons with energies $\varepsilon > I$ in the Maxwell distribution tail

caused by their loss in excitation of atoms (for neon, $\gamma < 0.1$). The assumption of a Maxwell spectrum of electron velocities is valid when the $\nu_u \approx \delta\nu_{T_e}/l_e$ frequency of electron energy loss in collisions with neutrals is lower than the frequency of their collisions with each other ($\nu_u < \nu_{ee}$). Here, δ is the coefficient of the effectiveness of collisions (for neon, $\delta \approx 10^{-4}$) [11]. Equation (16) is also used to estimate the T_e temperature of electrons in a weakly ionized glow discharge plasma when the positive column length L is larger than the characteristic energy relaxation length $\Lambda = 0.8 l_e \delta^{1/2}$. For neon at $P = 1$ Torr, $\Lambda \approx 9.7$ cm [11], which is smaller than the length $L \approx 20$ cm observed experimentally. The temperature of electrons in the absence of dust particles can therefore be obtained from the equation $\nu_i = \nu_{ep}$, which gives $T_e \approx 3.1$ eV. This value is in close agreement with the results of T_e measurements for tubes with 1–2 cm radii [11].

The determination of the spectrum of electron velocities in a dense dust cloud in the presence of additional ionization sources is a complex independent problem. For this reason, the mean electron energies in a dust cloud will also be qualitatively estimated by (16) taking into account that the ν_u electron energy loss frequency in collisions with neutrals is much lower than the $n_{ed} \approx \pi a_p^2 v_{T_e} n_p$ frequency of casual collisions between electrons and dust cloud macroparticles under our experimental conditions ($\nu_u \ll \nu_{ed}$). The electron velocity spectrum in the electric field of macroparticles whose surface potentials are comparable with electron energy ε_e can then be fairly close to the Gaussian spectrum with the variance proportional to $\varepsilon_e \approx T_e$ [23]. The temperature of electrons can therefore be estimated from the equality $\nu_i = \nu_{ep}$ by simultaneously solving (15) and (16). For $\gamma = 0.1$ and $z = 3$–4, this yields $T_e \approx$ 8–10 eV and $\varphi_s \approx$ 30–32 V if $n_p \approx (2\text{–}3) \times 10^3$ cm^{-3}. A decrease in the concentration to $n_p \approx 1.2 \times 10^3$ cm^{-3} ($I =$ 0.8 mA) results in $T_e \approx$ 6–8 eV and $\varphi_s \approx 24$ V. Lowering γ increases the T_e and φ_s values.

Electric field E necessary for electrons to acquire mean energy T_e can be estimated on the assumption that $T_e \propto \tau^{1/2} E$, where τ is the characteristic time required for the electron to acquire its mean energy [11]. In diffusion-controlled discharges, the ν_{ab} diffusion escape frequency given by (14) is much smaller than the $\nu_u \approx 10^5$ s^{-1} electron energy loss frequency in collisions with neutrals ($P = 1$ Torr), and we have $\tau \approx \nu_u^{-1}$. The characteristic τ value in a dense dust cloud is determined by the ν_{ep} electron loss frequency on macroparticles [Eq. (15)], $\tau \approx \nu_{ep}^{-1}$ ($\nu_{ep} > \nu_u$). The change in field E^p in a dust cloud with respect to positive column field E^0 is then, to within an order of magnitude, given by

$$\frac{E^p}{E^0} \approx \sqrt{\frac{\nu_{ep}}{\nu_u}}\frac{T_e^p}{T_e^0} = (\pi a_p^2 n_p \exp(-z) l_e/\delta)^{-0.5}\left(\frac{T_e^p}{T_e^0}\right)^{1.25},$$

where T_e^p and T_e^0 are the temperatures of electrons in the dust cloud and positive column, respectively. According to this equation, the field increases 15–20 times in the dust cloud if $a_p = 62.5$ μm, $n_p \approx (1\text{–}2) \times 10^3$ cm^{-3}, $z = 3$, and $T_e^p/T_e^0 = 3$. That is, at $E^0 = 1.5$ V/cm, $E^p \approx$ 22–30 V/cm, which is comparable with the $eZ_p n_p^{2/3} \approx$ 14–45 V/cm value for two-particle interactions in the system under consideration ($Z_p \approx (1\text{–}2) \times 10^6$).

Note that, in a positive column free of particles, ionization equilibrium is attained by equalizing the rates of plasma ionization ($n_e^0 \nu_i$) and recombination on tube walls ($\beta_{ab} n_e^0 n_i^0$),

$$n_e^0 \nu_i \equiv n_e^0 \nu_{ab} = \beta_{ab} n_e^0 n_i^0,$$

and the condition of ionization equilibrium in a dust cloud is

$$n_e \nu_i \equiv n_e \nu_{ep} = \beta_{ep} n_e n_i,$$

where β_{ab} and β_{ep} are the corresponding recombination coefficients. Suppose that the plasma recombination coefficients on tube walls and on the surface of dust particles are approximately equal, $\beta_{ab} \approx \beta_{ep}$. The ratio between the n_i^p concentration of ions in a dust cloud and the n_i^0 concentration of ions in a plasma without macroparticles is then given by

$$n_i/n_i^0 \equiv \nu_{ep}/\nu_{ab}, \tag{17}$$

where the $n_i^0 \approx (1\text{–}3) \times 10^8$ cm^{-3} value corresponds to the experimental current density. Using (17) and taking into account (14) and (15), we find that

$$n_i \approx (5\text{–}7.5) \times 10^9 \text{ cm}^{-3}$$

at various discharge currents at $z = 4$. At $z = 3$, n_i is approximately three times smaller. The condition $Z_p n_p < n_i$ [see (7)] necessary for the confinement of negatively charged particles is therefore quite attainable in the plasma–dust system under consideration. Note in conclusion that plasma parameter estimates considered in this section are used to analyze a qualitative picture of the influence of particles on ionization processes under the conditions of a dense dust cloud.

4. CONCLUSION

This work describes the results of experimental studies of the dynamics of bronze macroparticles in a direct current glow discharge plasma under microgravitation conditions. The temperatures, concentrations, pair correlation functions, self-diffusion coefficients, and charges of macroparticles were measured in various discharge currents. An analysis of the measurement results showed that the observed structures of macroparticles corresponded to a strongly nonideal dust liquid ($\Gamma^* \approx 60$–75), and their dynamics was in close agreement with the results of simulating Yukawa systems characterized by weak screening ($\kappa < 1$). The recorded time dependence of mean-square displacements of separate macroparticles (Fig. 6) was characteristic of the behavior of the system close to the liquid–crystal phase transition line. The normalized nonideality parameter was noticeably smaller than its $\Gamma_c \approx 102$ value on the system crystallization line. This circumstance could be explained by either polydispersity of macroparticles or a possible decrease in Γ_c caused by finite dust cloud dimensions.

Experimental estimates of macroparticle charges corresponded to surface potentials of the order of 30–40 V, which far exceeded values predicted by the OML theory. This discrepancy could be related to both a decrease in the effective flow of ions onto particles whose radii were larger than the free path of ions and an increase in the temperature of electrons in the dust cloud. It was shown that changes in equilibrium ionization processes in dense dust clouds could cause a substantial increase in the mean energy of electrons and in the concentration of the plasma component.

Note once more in conclusion that experiments on studying interaction potentials and charging processes for large-sized hundred-micron particles in a dust plasma cannot be performed under Earth gravitation conditions.

ACKNOWLEDGMENTS

The authors are deeply indebted to cosmonauts S.V. Avdeev, Yu.M. Baturin, T.A. Musabaev, and G.I. Padalka for the considerable work done to prepare the Plazmennyi Kristall-2 cosmic experiment aboard the Mir orbital station.

This work was financially supported by the Russian Foundation for Basic Research (project nos. 01-02-16658, 00-02-81036, and 00-02-17520) and INTAS (project no. 2000-0522).

REFERENCES

1. H. Thomas, G. E. Morfill, V. Demmel, *et al.*, Phys. Rev. Lett. **73**, 652 (1994).
2. A. Melzer, A. Homann, and A. Piel, Phys. Rev. E **53**, 2757 (1996).
3. V. V. Zhakhovskiĭ, V. I. Molotkov, A. P. Nefedov, *et al.*, Pis'ma Zh. Éksp. Teor. Fiz. **66**, 392 (1997) [JETP Lett. **66**, 419 (1997)].
4. S. Nunomura, T. Misawa, N. Ohno, and S. Takamura, Phys. Rev. Lett. **83**, 1970 (1999).
5. O. Vaulina, S. Khrapak, A. Nefedov, *et al.*, Phys. Rev. E **60**, 5959 (1999).
6. V. E. Fortov, A. P. Nefedov, O. S. Vaulina, *et al.*, Zh. Éksp. Teor. Fiz. **114**, 2004 (1998) [JETP **87**, 1087 (1998)].
7. O. S. Vaulina, A. P. Nefedov, O. F. Petrov, *et al.*, Zh. Éksp. Teor. Fiz. **119**, 1129 (2001) [JETP **92**, 979 (2001)].
8. V. E. Fortov, A. P. Nefedov, V. I. Molotkov, *et al.*, Phys. Rev. Lett. **87**, 205002 (2001).
9. G. Morfill, H. Thomas, U. Konopka, *et al.*, Phys. Rev. Lett. **83**, 1598 (1999).
10. T. Stuffler, G. Schmitt, H. Pfeuffer, *et al.*, in *Proceedings of the 52nd International Astronautical Congress, Toulouse, France, 2001*, IAF-01-J.6.02.
11. Yu. P. Raizer, *Gas Discharge Physics* (Nauka, Moscow, 1987; Springer-Verlag, Berlin, 1991).
12. E. M. Lifshitz and L. P. Pitaevskiĭ, *Course of Theoretical Physics*, Vol. 10: *Physical Kinetics* (Nauka, Moscow, 1979; Pergamon, Oxford, 1981).
13. N. A. Fuchs, *The Mechanics of Aerosols* (Dover, New York, 1964).
14. F. A. Lindemann, Z. Phys. **11**, 609 (1910).
15. O. S. Vaulina and S. V. Vladimirov, Plasma Phys. **9**, 835 (2002).
16. O. S. Vaulina and S. A. Khrapak, Zh. Éksp. Teor. Fiz. **119**, 264 (2001) [JETP **92**, 228 (2001)].
17. J. E. Daugherty, R. K. Porteos, M. D. Kilgore, *et al.*, J. Appl. Phys. **72**, 3934 (1992).
18. J. E. Allen, Phys. Scr. **45**, 497 (1992).
19. A. M. Lipaev, V. I. Molotkov, A. P. Nefedov, *et al.*, Zh. Éksp. Teor. Fiz. **112**, 2030 (1997) [JETP **85**, 1110 (1997)].
20. T. Nitter, Plasma Sources Sci. Technol. **5**, 93 (1996).
21. O. S. Vaulina, S. A. Khrapak, A. A. Samarian, and O. F. Petrov, Phys. Scr. **84** (2000).
22. J. Goree, Plasma Sources Sci. Technol. **3**, 400 (1994).
23. A. A. Ovchinnikov, S. F. Timashev, and A. A. Belyĭ, *Kinetics of Diffusely Controlled Processes* (Khimiya, Moscow, 1986).

Translated by V. Sipachev

Journal of Experimental and Theoretical Physics, Vol. 95, No. 4, 2002, pp. 682–696.
Translated from Zhurnal Éksperimental'noĭ i Teoreticheskoĭ Fiziki, Vol. 122, No. 4, 2002, pp. 789–804.

PLASMA, GASES

On a Kinetic Justification of the Generalized Nonequilibrium Thermodynamics of Multicomponent Systems

V. M. Zhdanov[a] and V. I. Roldugin[b, *]

[a]*Moscow Engineering Physics Institute, Moscow, 115409 Russia*
[b]*Institute of Physical Chemistry, Russian Academy of Sciences, Moscow, 119991 Russia*
**e-mail: roldugin@phyche.ac.ru*

Receiver April 19, 2002

Abstract—The problem of the kinetic justification of the generalized thermodynamics of nonequilibrium processes using the method of moments for solving the kinetic equation for a multicomponent gas mixture is examined. Generalized expressions are obtained for the entropy density, entropy flux density, and entropy production as functions of an arbitrary number of state variables (moments of the distribution function). Different variants of writing the relations between fluxes and thermodynamic forces are considered, which correspond to the Onsager version for spatially homogeneous systems and, in a more general case, lead to the generalized thermodynamic forces of a complicated form, including derivatives of the fluxes with respect to time and spatial coordinates. Some consequences and new physical effects, following from the obtained equations, are analyzed. It is shown that a transition from results of the method of moments to expressions for the entropy production and the corresponding phenomenological relations of the generalized nonequilibrium thermodynamics is possible on the level of a linearized Barnett approximation of the Chapman–Enskog method. © *2002 MAIK "Nauka/Interperiodica".*

1. INTRODUCTION

In recent years, there were successful attempts at going beyond the conventional limits of applicability of the classical linear nonequilibrium thermodynamics (LNT) [1–10]. The main idea underlying generalization of the classical approach consists in increasing the number of variables describing nonequilibrium states of the system: the set of usual classical variables employed in the LNT is supplemented with new independent variables representing dissipative fluxes (e.g., heat flux $\mathbf{q}$, tensor of viscous stresses π_{ik}, or diffusion flux $\mathbf{J}_\alpha$). In this formulation, a local entropy of the system also becomes a function of the new variables; expressions for the entropy flux density and entropy production (entering into the general entropy balance equation) acquire a different (generalized) form.

Another important circumstance is that, in contrast to the classical variables obeying the laws of conservation and varying relatively slowly, the fluxes are "fast" variables satisfying equations of the evolution type. Note that the relaxation character of these equations eliminates the paradox of infinite velocity of the propagation of thermal and shear perturbations, which follows from the usual linear Fourier and Newton laws for the heat flux and the viscous stress tensor. Realization of the above concepts led to the construction of the so-called extended irreversible thermodynamics [1, 2].

It should be noted that the idea of generalizing the LNT was already present in the early works of Grad [11] devoted to development of the method of moments in the kinetic theory of gases. Grad pointed out the possibility of applying nonequilibrium thermodynamics to more general cases, when a nonequilibrium state of a homogeneous gas (and the nonequilibrium entropy of the system) depends not only on the local density and internal energy (temperature), but on an arbitrary number of additional state variables (moments of the distribution function) as well. The latter variables satisfy quasi-linear differential equations of the relaxation type obtained on the basis of the Boltzmann equation using expansion of the distribution function of the gas molecules into series in Hermite polynomials [11].

The kinetic aspect of the problem still plays an important role in development of generalized nonequilibrium thermodynamics (GNT). Previously [5–8], it was demonstrated that the GNT principles are compatible with the highest order approximations of the Chapman–Enskog method on a level of the linearized Barnett approximation. Upon this transition, additional terms enter into the expression for the entropy production, which contain higher (second-order and above) derivatives of the principal macroscopic variables with respect to coordinates. Accordingly, the system of phenomenological equations for the fluxes and thermodynamic forces is expanded.

An important role of the spatial derivatives of fluxes (high-order moments) in the linearized equations of moments was demonstrated within the framework of the method of moments [9, 10, 12]. It should be noted that the idea of including new quantities proportional to the derivatives of fluxes and possessing the correspond-

1063-7761/02/9504-0682$22.00

ing tensor dimension into the number of independent thermodynamic forces involved in the transfer equations was put forward in 1962 by Kagan in the course of the study [12] and then elaborated in [5, 6, 8–10]. The introduction of such terms brings the GNT relations outside the framework of the usual linear equations of the classical LNT, but, despite certain complication of the generalized thermodynamic forces, an expression for the entropy production still appears as a bilinear combination of thermodynamic forces and fluxes.

Until now, the GNT relations were formulated and analyzed mostly in application to the case of a simple (single-component) gas. Only Heckl and Muller [13] considered the corresponding equations on the kinetic level for a gas mixture, in a particular case of so-called Maxwellian molecules with the interaction potential inversely proportional to the fourth power of the interparticle distance. Moreover, the investigation was restricted for the most part to an approximation of 13 moments with the additional variables representing the vector of the heat flux and the viscous stress tensor. However, an allowance for the high-order moments becomes important in a number of cases. For example, realistic transfer coefficients as functions of the frequency and wavenumber for the so-called generalized hydrodynamics can be obtained within the framework of an approximation using at least 26 moments [14, 15].

This paper considers problems pertaining to the kinetic justification of the GNT in the most general form based on a linearized variant of the method of moments applied to a multicomponent gas mixture with an arbitrary potential of the intermolecular interaction. A linearized Boltzmann equation is used to obtain an infinite chain of linked equations for the coefficients of expansion of the distribution function in a system of tensor polynomials (equations of moments). Using these equations, we obtain generalized expressions for the entropy density, entropy flux density, and entropy production, which allow different variants of the relations between fluxes and thermodynamic forces to be considered. For spatially homogeneous systems, these relations correspond to the Onsager version of the LNT. In a more general case, thermodynamic forces involved in the transfer equations are significantly redefined to include, besides the usual gradients of the initial macroscopic parameters, derivatives of the dissipative fluxes with respect to time and coordinates. Some consequences and new physical effects, following from the obtained equations, are analyzed. It is shown that a transition from results of the generalized method of moments to the well-known results obtained by the Chapman–Enskog method [8, 15, 16] is possible on the level of both the first and the next (Barnett) approximation.

It should be noted that the possibility of using the methods of nonequilibrium thermodynamics on the level of transfer equations in the Barnett approximation was repeatedly discussed [17, 18]. Below, we reveal the source of such misunderstanding, obtain a generalized expression for the entropy production, and construct a system of phenomenological equations of the linearized Barnett approximation completely conforming to the principles of nonequilibrium thermodynamics. A series of "nonphysical" fluxes involved, together with the usual ones, in this scheme provides for completeness of the system and ensures validity of the necessary Onsager relations between the cross coefficients. The proposed system of equations admits obvious generalization to the cases of super-Barnett and higher approximations.

2. LINEARIZED KINETIC EQUATION FOR A GAS MIXTURE

Consider an N-component gas mixture consisting of monoatomic molecules of various types. Let the state of the mixture deviate slightly from equilibrium, in which case the velocity distribution function for particles of the α type can be represented as

$$f_\alpha = f_\alpha^{(0)}(1 + \phi_\alpha),$$
$$f_\alpha^{(0)} = n_\alpha\left(\frac{\beta_\alpha}{\pi}\right)^{3/2}\exp(-\beta_\alpha c_\alpha^2), \tag{1}$$

where $f_\alpha^{(0)}$ is the local Maxwell distribution, ϕ_α is a small correction ($|\phi_\alpha| \ll 1$), $\beta_\alpha = m_\alpha/2kT$, n_α is the number density of particles of the α type, k is the Boltzmann constant, T is the absolute temperature, m_α is the mass of particles of the α type, $\mathbf{c}_\alpha = \mathbf{v}_\alpha - \mathbf{u}$ is the relative velocity of a particle, and $\mathbf{u}$ is the mass-average velocity of the mixture. The values of n_α, $\mathbf{u}$, and T are determined by the following relations:

$$n_\alpha = \int f_\alpha d\mathbf{c}_\alpha, \quad \mathbf{u} = \rho^{-1}\sum_\alpha\int m_\alpha \mathbf{v}_\alpha f_\alpha d\mathbf{c}_\alpha,$$
$$\frac{3}{2}nkT = \sum_\alpha\int\frac{m_\alpha}{2}c_\alpha^2 f_\alpha d\mathbf{c}_\alpha. \tag{2}$$

Here, $\rho = \sum_\alpha \rho_\alpha$, where $\rho_\alpha = m_\alpha n_\alpha$ is the density of molecules of the α type, and $n = \sum_\alpha n_\alpha$ is the number density of particles in the mixture.

Introducing the internal scalar product of functions in the Hilbert space,

$$n_\alpha(g_\alpha \cdot h_\alpha) = \int f_\alpha^{(0)} g_\alpha h_\alpha d\mathbf{c}_\alpha, \tag{3}$$

we can determine the values of n_α, $\mathbf{u}$, and T in a similar form for both f_α and $f_\alpha^{(0)}$, by virtue of which

$$(1, \phi_\alpha) = 0, \quad \sum_\alpha n_\alpha (m_\alpha \mathbf{c}_\alpha, \phi_\alpha) = 0,$$
$$\sum_\alpha n_\alpha (m_\alpha \mathbf{c}_\alpha^2, \phi_\alpha) = 0. \tag{4}$$

Let us define nonequilibrium macroscopic parameters of the mixture components, including the diffusion flux $\mathbf{J}_\alpha$, the partial viscous stress tensor $\hat{\boldsymbol{\pi}}_\alpha$, and the partial heat flux $\mathbf{q}_\alpha$ as

$$\mathbf{J}_\alpha = \rho_\alpha \mathbf{w}_\alpha = n_\alpha m_\alpha (\mathbf{c}_\alpha, \phi_\alpha),$$
$$\hat{\boldsymbol{\pi}}_\alpha = n_\alpha m_\alpha (\overline{\mathbf{c}_\alpha \mathbf{c}_\alpha}, \phi_\alpha), \tag{5}$$
$$\mathbf{q}_\alpha = \frac{1}{2} n_\alpha m_\alpha (c_\alpha^2 \mathbf{c}_\alpha, \phi_\alpha),$$

where $\mathbf{w}_\alpha = \mathbf{u}_\alpha - \mathbf{u}$ is the diffusion velocity of particles of the α type, $\pi_{\alpha ik} = P_{\alpha ik} - p_\alpha \delta_{ik}$, $P_{\alpha ik}$ is the partial stress tensor, and p_α is the scalar pressure of particles of the α type. Here and below, expression $\overline{\mathbf{a}\mathbf{a}\ldots}$ refers to irreducible symmetric tensors, for example,

$$(\overline{\mathbf{c}_\alpha \mathbf{c}_\alpha})_{ik} = c_{\alpha i} c_{\alpha k} - \frac{1}{3} \delta_{ik} c_\alpha^2.$$

The quantities ρ_α, $\mathbf{J}_\alpha$, $P_{\alpha ik}$, and $2\mathbf{q}_\alpha$ represent the first several moments of the distribution function. Note also that p_α can be represented in the form of the product

$$p_\alpha = n_\alpha k T_\alpha,$$

where T_α is the temperature of the α component of the mixture defined as

$$k T_\alpha = kT + \frac{1}{3} m_\alpha (c_\alpha^2, \phi_\alpha). \tag{6}$$

The correction $\phi_\alpha(\mathbf{c}_\alpha, \mathbf{r}, t)$ in expression (1) satisfies the linearized kinetic Boltzmann equation [19, 20]

$$D_\alpha \ln f_\alpha^{(0)} + D_\alpha \phi_\alpha = \sum_\beta L_{\alpha\beta} \phi_\beta, \tag{7}$$

where the operators are defined as

$$D_\alpha = \frac{d}{dt} + (\mathbf{c}_\alpha \cdot \nabla) + \left(\frac{\mathbf{F}_\alpha}{m_\alpha} \cdot \nabla_v\right),$$
$$\frac{d}{dt} = \frac{\partial}{\partial t} + (\mathbf{u} \cdot \nabla). \tag{8}$$

Here, $\mathbf{F}_\alpha$ is the external force acting upon a particle of the α type, ∇_v is the gradient operator in the space of velocities, and $L_{\alpha\beta}$ is the linearized operator of collisions [19]:

$$L_{\alpha\beta}\phi_\beta = \int f_\beta^{(0)} (\phi'_\alpha + \phi'_{1\beta} - \phi_\alpha - \phi_{1\beta}) g \sigma_{\alpha\beta} d\boldsymbol{\Omega} d\mathbf{v}_{1\beta}, \tag{9}$$

where g is the relative velocity of colliding particles, $\sigma_{\alpha\beta}(g, \boldsymbol{\Omega})$ is the differential scattering cross section, and $\boldsymbol{\Omega}$ is the solid angle (primes indicate the values upon collision and subscript 1 is introduced to distinguish the colliding particles when $\alpha = \beta$).

The mass, momentum, and kinetic energy of the particles are the invariants of collisions, which can be expressed as

$$(m_\alpha, L\phi_\alpha) = 0, \quad \sum_\alpha n_\alpha (m_\alpha \mathbf{c}_\alpha, L\phi_\alpha) = 0,$$
$$\sum_\alpha n_\alpha (m_\alpha c_\alpha^2, L\phi_\alpha) = 0, \tag{10}$$

where we introduce the notation

$$L\phi_\alpha = \sum_\beta L_{\alpha\beta} \phi_\beta.$$

Below, we will use two well-known properties of the linearized operator of collisions [19], including the symmetry relation

$$\sum_\alpha n_\alpha (\psi_\alpha, L\phi_\alpha) = \sum_\alpha n_\alpha (\phi_\alpha, L\psi_\alpha) \tag{11}$$

and the condition

$$\sum_\alpha n_\alpha (\phi_\alpha, L\phi_\alpha) \le 0 \tag{12}$$

(where equality corresponds to the case of satisfied conditions (10)).

Using the expression for $f_\alpha^{(0)}$, the first term in the left-hand part of the kinetic equation (7) can be represented as

$$D_\alpha \ln f_\alpha^{(0)} = \left\{ \left(\frac{1}{n_\alpha}\frac{dn_\alpha}{dt} + \nabla \mathbf{u}\right) \right.$$
$$+ \left(\beta_\alpha c_\alpha^2 - \frac{3}{2}\right)\left(\frac{1}{T}\frac{dT}{dt} + \frac{2}{3}\nabla \mathbf{u}\right)$$
$$+ \mathbf{c}_\alpha \frac{1}{p_\alpha}\left[\nabla p_\alpha + \rho_\alpha \left(\frac{d\mathbf{u}}{dt} - \frac{\mathbf{F}_\alpha}{m_\alpha}\right)\right] \tag{13}$$
$$\left. + \mathbf{c}_\alpha \left(\beta_\alpha c_\alpha^2 - \frac{5}{2}\right)\frac{1}{T}\nabla T + 2\beta_\alpha \overline{\mathbf{c}_\alpha \mathbf{c}_\alpha} : \overline{\nabla \mathbf{u}} \right\},$$

where

$$(\overline{\nabla \mathbf{u}})_{ik} = \frac{1}{2}\left(\frac{\partial u_i}{\partial x_k} + \frac{\partial u_k}{\partial x_i}\right) - \frac{1}{3}\delta_{ik} \nabla \mathbf{u}. \tag{14}$$

3. EXPANSION OF THE DISTRIBUTION FUNCTION: LINEARIZED EQUATIONS OF MOMENTS

Let us expand the small nonequilibrium correction ϕ_α in a series with respect to an orthonormalized system of irreducible tensor polynomials $\mathbf{P}_\alpha^{mn}(\mathbf{W}_\alpha)$:

$$\phi_\alpha = \sum_{m=0}^{\infty}\sum_{n=0}^{\infty} \mathbf{a}_\alpha^{mn}(\mathbf{r}, t)\mathbf{P}_\alpha^{mn}(\mathbf{W}_\alpha), \tag{15}$$

where $\mathbf{W}_\alpha = \beta_\alpha^{1/2}\mathbf{c}_\alpha$ is the dimensionless velocity of particles of the α type. To within the normalization, the $\mathbf{P}_\alpha^{mn}(\mathbf{W}_\alpha)$ polynomials represent the products of the Sonin polynomials $S_{m+1/2}^n(W_\alpha^2)$ by the tensor spherical harmonics $\mathbf{R}^{(m)}(\mathbf{W}_\alpha)$ [20, 21],

$$\mathbf{P}_\alpha^{mn}(\mathbf{W}_\alpha) = \gamma_{mn}S_{m+1/2}^n(W_\alpha^2)\mathbf{R}^{(m)}(\mathbf{W}_\alpha), \tag{16}$$

where γ_{mn} are the normalization factors:

$$\gamma_{mn} = (-1)^n\sqrt{\frac{2^{m+n}n!(2m+1)!!}{m!(2m+2n+1)!!}}.$$

The first several $\mathbf{P}^{mn}(\mathbf{W})$ polynomials are as follows:

$$\begin{gathered} P^{00} = 1, \quad P^{01} = \sqrt{\frac{2}{3}}\left(W^2 - \frac{3}{2}\right), \\ \mathbf{P}^{10} = \sqrt{2}\mathbf{W}, \\ \mathbf{P}^{11} = \sqrt{\frac{4}{5}}\mathbf{W}\left(W^2 - \frac{5}{2}\right), \quad \mathbf{P}^{20} = \sqrt{2}\overline{\mathbf{WW}}, \\ \mathbf{P}^{30} = \sqrt{\frac{4}{3}}\overline{\mathbf{WWW}}. \end{gathered} \tag{17}$$

A normalization condition for the $\mathbf{P}_\alpha^{mn}(\mathbf{W}_\alpha)$ polynomials is

$$(\mathbf{P}_\alpha^{mn}, \mathbf{P}_\alpha^{m'n'}) = \delta_{mm'}\delta_{nn'}\Delta^{(m)}, \tag{18}$$

where $\Delta^{(m)}$ is the unit projection tensor [20].

According to the condition of orthogonality of the polynomials, the coefficients of expansion (15) are given by the relations

$$n_\alpha\mathbf{a}_\alpha^{mn} = \int \mathbf{P}_\alpha^{mn} f_\alpha^{(0)}\phi_\alpha d\mathbf{c}_\alpha = n_\alpha(\mathbf{P}_\alpha^{mn}, \phi_\alpha). \tag{19}$$

Using these formulas, the expansion coefficients can be expressed through the corresponding moments of the distribution function. The first five coefficients are as follows:

$$\begin{gathered} a_\alpha^{00} = 0, \quad a_\alpha^{01} = \sqrt{\frac{3}{2}}\frac{T_\alpha - T}{T}, \\ \mathbf{a}_\alpha^{10} = \sqrt{2}\sqrt{\beta_\alpha}\mathbf{w}_\alpha, \\ \mathbf{a}_\alpha^{11} = \sqrt{\frac{4}{5}}\sqrt{\beta_\alpha}\frac{\mathbf{h}_\alpha}{p_\alpha}, \quad \hat{\mathbf{a}}_\alpha^{20} = \frac{\sqrt{2}}{2}\frac{\hat{\boldsymbol{\pi}}_\alpha}{p_\alpha}, \end{gathered} \tag{20}$$

where

$$\mathbf{h}_\alpha = \mathbf{q}_\alpha - (5/2)p_\alpha\mathbf{w}_\alpha \tag{21}$$

is the reduced partial heat flux.

The system of equations for the coefficients $n_\alpha\mathbf{a}_\alpha^{mn}$ (the system of equations of moments) is constructed by multiplying the linearized kinetic equation (7) by $f_\alpha^{(0)}\mathbf{P}_\alpha^{mn}$, followed by integrating with respect to velocities, which yields

$$n_\alpha(\mathbf{P}_\alpha^{mn}, D_\alpha\ln f_\alpha^{(0)}) + n_\alpha(\mathbf{P}_\alpha^{mn}, D_\alpha\phi_\alpha) = \mathbf{R}_\alpha^{mn}, \tag{22}$$

where

$$\mathbf{R}_\alpha^{mn} = \sum_\beta \int \mathbf{P}_\alpha^{mn} f_\alpha^{(0)} L_{\alpha\beta}\phi_\beta d\mathbf{c}_\alpha = n_\alpha(\mathbf{P}_\alpha^{mn}, L\phi_\alpha) \tag{23}$$

is the moment relative to the collision integral.

Note that each term in the expression (13) for $D_\alpha\ln f_\alpha^{(0)}$ contains some $\mathbf{P}_\alpha^{mn}$ polynomial (17) as a factor. Then, integrating the first term in (7) and using the orthogonality condition (18), we obtain

$$\begin{aligned} n_\alpha(\mathbf{P}_\alpha^{mn}, D_\alpha\ln f_\alpha^{(0)}) = n_\alpha\Bigg\{&\left(\frac{1}{n_\alpha}\frac{dn_\alpha}{dt} + \nabla\mathbf{u}\right)\delta_{m0}\delta_{n0} \\ &+ \sqrt{\frac{3}{2}}\left(\frac{1}{T}\frac{dT}{dt} + \frac{2}{3}\nabla\mathbf{u}\right)\delta_{m0}\delta_{n1} \\ &+ \frac{\sqrt{2}}{2}\beta_\alpha^{-1/2}\frac{1}{p_\alpha}\left[\nabla p_\alpha + \rho_\alpha\left(\frac{d\mathbf{u}}{dt} - \frac{\mathbf{F}_\alpha}{m_a}\right)\right]\delta_{m1}\delta_{n0} \\ &+ \sqrt{\frac{5}{4}}\beta_\alpha^{-1/2}\nabla\ln T\delta_{m1}\delta_{n1} + \sqrt{2}\overline{\nabla\mathbf{u}}\delta_{m2}\delta_{m0}\Bigg\}. \end{aligned} \tag{24}$$

In the expression obtained upon integration of the second term in (7), we omit all nonlinear terms including the products of moments (or the $n_\alpha\mathbf{a}_\alpha^{mn}$ coefficients) by small gradients of the corresponding thermodynamic variables, as well as the small potential gradients (weak external fields). As a result, this yields

$$\begin{aligned} n_\alpha(\mathbf{P}_\alpha^{mn}, D_\alpha\phi_\alpha) &= \frac{dn_\alpha\mathbf{a}_\alpha^{mn}}{dt} \\ &+ \nabla n_\alpha(\mathbf{c}_\alpha\mathbf{P}_\alpha^{mn}, \phi_\alpha) + \text{nonlinear terms}. \end{aligned} \tag{25}$$

By substituting the expansion (15) for ϕ_α, the second term (flux term) in the right-hand part of Eq. (25) can be represented as a linear combination of derivatives with respect to coordinates of the coefficients at the $(m+1)$th and $(m-1)$th tensor dimensions [20, 21]:

$$\nabla n_\alpha(\mathbf{c}_\alpha \mathbf{P}_\alpha^{mn}, \phi_\alpha) = \sum_k \sum_l (\mathbf{c}_\alpha(\mathbf{P}_\alpha^{mn}, \mathbf{P}_\alpha^{kl}))\nabla n_\alpha \mathbf{a}_\alpha^{kl} \tag{26}$$

$$= \sum_l (\mathbf{A}_{\alpha mn}^{m+1,l} \nabla n_\alpha \mathbf{a}_\alpha^{m+1,l} + \mathbf{B}_{\alpha mn}^{m-1,l} \overline{\nabla n_\alpha \mathbf{a}_\alpha}^{m-1,l}),$$

where $\mathbf{A}_\alpha$ and $\mathbf{B}_\alpha$ are operators expressed as

$$\mathbf{A}_{\alpha mn}^{m+1,l} = \frac{1}{2m+3}(\mathbf{c}_\alpha \mathbf{P}_\alpha^{mn}, \mathbf{P}_\alpha^{m+1,l}),$$
$$\mathbf{B}_{\alpha mn}^{m-1,l} = \frac{1}{2m+1}(\mathbf{c}_\alpha \mathbf{P}_\alpha^{mn}, \mathbf{P}_\alpha^{m-1,l}), \tag{27}$$

and the notation $\overline{\nabla n_\alpha \mathbf{a}_\alpha}^{m-1,l}$ corresponds to a symmetric irreducible tensor. For example, if $\mathbf{a}_\alpha^{m-1,l}$ is a vector (for $m = 2$), $\overline{\nabla n_\alpha \mathbf{a}_\alpha}^{1,l}$ is an irreducible second-order tensor with the components

$$\frac{1}{2}\left(\frac{\partial n_\alpha \mathbf{a}_{\alpha i}^{1l}}{\partial x_j} + \frac{\partial n_\alpha \mathbf{a}_{\alpha j}^{1l}}{\partial x_i}\right) - \frac{1}{3}\delta_{ij}\frac{\partial n_\alpha \mathbf{a}_{\alpha l}^{1l}}{\partial x_l}. \tag{28}$$

Now, let us transform the quantity $\mathbf{R}_\alpha^{mn}$ by substituting the corrections ϕ_α and ϕ_β, expanded according to (15), into the linearized collision integral. This yields

$$\mathbf{R}_\alpha^{mn} = -\sum_{\beta=1}^{N} \sum_k \sum_l n_\alpha n_\beta \times ([\mathbf{P}^{kl}, \mathbf{P}^{mn}]'_{\alpha\beta} \mathbf{a}_\alpha^{kl} + [\mathbf{P}^{kl}, \mathbf{P}^{mn}]''_{\alpha\beta} \mathbf{a}_\beta^{kl}), \tag{29}$$

where we have introduced the so-called integral brackets of the $\mathbf{P}_\alpha^{mn}$ polynomials. A general definition of the integral brackets is as follows [16, 19]:

$$[F, G]'_{\alpha\beta} = \frac{1}{n_\alpha n_\beta} \times \int f_\alpha^{(0)} f_\beta^{(0)} (F_\alpha - F'_\alpha) G_\alpha g \sigma_{\alpha\beta} d\mathbf{\Omega} d\mathbf{c}_\alpha d\mathbf{c}_\beta,$$
$$[F, G]''_{\alpha\beta} = \frac{1}{n_\alpha n_\beta} \times \int f_\alpha^{(0)} f_\beta^{(0)} (F_\beta - F'_\beta) G_\alpha g \sigma_{\alpha\beta} d\mathbf{\Omega} d\mathbf{c}_\alpha d\mathbf{c}_\beta. \tag{30}$$

Using definition (16) of the $\mathbf{P}_\alpha^{mn}(\mathbf{W}_\alpha)$ polynomials and the relation valid for the integral brackets of the $\mathbf{R}^{(m)}(\mathbf{W})$ polynomials,

$$[\mathbf{R}^{(k)}, \mathbf{R}^{(m)}] = \frac{1}{2m+1}[\mathbf{R}^{(m)}, \mathbf{R}^{(m)}]\delta_{km}, \tag{31}$$

the final expression for $\mathbf{R}_\alpha^{mn}$ can be written as

$$\mathbf{R}_\alpha^{mn} = -\sum_\beta \sum_l C_{\alpha\beta}^{mnl} \mathbf{a}_\beta^{ml}, \tag{32}$$

where

$$C_{\alpha\beta}^{mnl} = \delta_{\alpha\beta} \sum_\gamma A_{\alpha\gamma}^{mnl} + B_{\alpha\beta}^{mnl}, \tag{33}$$

$$A_{\alpha\beta}^{mnl} = n_\alpha n_\beta Q_{mnl} \times [S_{m+1/2}^l(W^2)\mathbf{R}^m(\mathbf{W}), S_{m+1/2}^n(W^2)\mathbf{R}^{(m)}(\mathbf{W})]'_{\alpha\beta},$$
$$B_{\alpha\beta}^{mnl} = n_\alpha n_\beta Q_{mnl} \times [S_{m+1/2}^l(W^2)\mathbf{R}^m(\mathbf{W}), S_{m+1/2}^n(W^2)\mathbf{R}^{(m)}(\mathbf{W})]''_{\alpha\beta}. \tag{34}$$

$$Q_{mnl} = \frac{1}{2m+1}\gamma_{mn}\gamma_{ml}. \tag{35}$$

The integral quantities $[\ldots]'_{\alpha\beta}$ and $[\ldots]''_{\alpha\beta}$ correspond to the well-known integral brackets of the Sonin polynomials introduced in the Chapman–Enskog theory [16, 19].

Taking into account the above expressions, the system of linearized equations of moments can be written in the following final form:

$$\frac{dn_\alpha \mathbf{a}_\alpha^{mn}}{dt} + \sum_l (\mathbf{A}_{\alpha mn}^{m+1,l} \nabla n_\alpha \mathbf{a}_\alpha^{m+1,l} + \mathbf{B}_{\alpha mn}^{m-1,l} \overline{\nabla n_\alpha \mathbf{a}_\alpha}^{m-1,l})$$
$$+ n_\alpha \left\{ \left(\frac{1}{n_\alpha}\frac{dn_\alpha}{dt} + \nabla \mathbf{u}\right)\delta_{m0}\delta_{n0} + \sqrt{\frac{3}{2}}\left(\frac{1}{T}\frac{dT}{dt} + \frac{2}{3}\nabla \mathbf{u}\right)\delta_{m0}\delta_{n1} \right.$$
$$+ \frac{\sqrt{2}}{2}\beta_\alpha^{-1/2}\frac{1}{p_\alpha}\left[\nabla p_\alpha + \rho_\alpha\left(\frac{d\mathbf{u}}{dt} - \frac{\mathbf{F}_\alpha}{m_\alpha}\right)\right]\delta_{m1}\delta_{n0}$$
$$\left. + \sqrt{\frac{5}{4}}\beta_\alpha^{-1/2}\nabla \ln T \delta_{m1}\delta_{n1} + \sqrt{2}\overline{\nabla \mathbf{u}}\delta_{m2}\delta_{n0} \right\}$$
$$= -\sum_\beta \sum_l C_{\alpha\beta}^{mnl} \mathbf{a}_\beta^{ml}. \tag{36}$$

These equations should be supplemented with the relations following from definitions (4) and (6):

$$\sum_\alpha m_\alpha n_\alpha \beta_\alpha^{-1/2} \mathbf{a}_\alpha^{10} = 0, \quad \sum_\alpha n_\alpha a_\alpha^{01} = 0. \tag{37}$$

Equations (36) form an infinite system of linked equations for scalar ($m = 0$), vector ($m = 1$), and tensor ($m = 2, 3, \ldots$) quantities. The search for particular solutions is possible upon restricting expansion (15) to a finite number of terms. Depending on the m and n values adopted, the general system (36) splits into independent sets of equations for the scalar, vector, and tensor coefficients $n_\alpha \mathbf{a}_\alpha^{mn}$. For example, the case of $m = 0$ and $n = 0$ corresponds to the continuity equation. Summing the expressions for $m = 1$, $n = 0$ or $m = 0$, $n = 1$ with respect to α leads to an equation of motion and a linearized equation of energy of the gas mixture. With an allowance for conditions (10) and (37), the corresponding conservation equations take the form

$$\frac{dn_\alpha}{dt} + n_\alpha \nabla \mathbf{u} + \nabla n_\alpha \mathbf{w}_\alpha = 0, \tag{38}$$

$$\rho \frac{du}{dt} + \nabla p + \nabla \hat{\boldsymbol{\pi}} + \sum_\alpha n_\alpha \mathbf{F}_\alpha = 0, \tag{39}$$

$$nk\frac{dT}{dt} + \frac{2}{3}p\nabla \mathbf{u} + \frac{2}{3}\nabla \mathbf{q} - kT\sum_\alpha \nabla n_\alpha \mathbf{w}_\alpha = 0, \tag{40}$$

where $p = nkT$ is the total pressure, $\hat{\boldsymbol{\pi}} = \sum_\alpha \hat{\boldsymbol{\pi}}_\alpha$ is the viscous stress tensor, and $\mathbf{q} = \sum_\alpha \mathbf{q}_\alpha$ is the heat flux. Note that, using Eqs. (39) and (40), it is possible to exclude the time derivatives $d\mathbf{u}/dt$ and dT/dt from the left-hand part of Eq. (36), while the continuity equation (38) entering into Eqs. (36) is identically valid.

4. ENTROPY BALANCE EQUATION

Multiplying the linearized kinetic equation (7) by $f_\alpha^{(0)}(\ln f_\alpha - 1)$, integrating with respect to energies, and summing over α, we obtain the entropy balance equation [22]

$$\rho \frac{ds}{dt} + \nabla \mathbf{J}_s = \sigma, \tag{41}$$

where

$$\rho s = (\rho s)_0 - \frac{k}{2}\sum_\alpha n_\alpha (1, \phi_\alpha^2) \tag{42}$$

is the local entropy density, $(\rho s)_0$ representing the density of entropy in the state of local equilibrium:

$$\begin{aligned}(\rho s)_0 &= -\sum_\alpha \int f_\alpha^{(0)}(\ln f_\alpha^{(0)} - 1) d\mathbf{c}_\alpha \\ &= -\frac{1}{T}\left(\sum_\alpha \rho_\alpha \mu_\alpha - \rho U - p\right).\end{aligned} \tag{43}$$

For the ideal gas of monoatomic molecules [23], the quantity

$$\mu_\alpha = \frac{kT}{m_\alpha}\left(\ln n_\alpha - \frac{3}{2}\ln\frac{2\pi kT}{m_\alpha}\right) + \text{const} \tag{44}$$

represents the chemical potential (per unit mass) of particles of the α type and $U = 3kT/2m$ is the internal energy density.

The entropy flux $\mathbf{J}_s$ and the local entropy production σ are given by the expressions

$$\mathbf{J}_s = \frac{1}{T}\left(\mathbf{q} - \sum_\alpha \mu_\alpha \mathbf{J}_\alpha\right) - \frac{k}{2}\sum_\alpha n_\alpha (\mathbf{c}_\alpha, \phi_\alpha^2), \tag{45}$$

$$\sigma = -k\sum_\alpha (L\phi_\alpha, \phi_\alpha). \tag{46}$$

Using expansion (15), expressions (42), (45), and (46) can be written as

$$\rho s = (\rho s)_0 - \frac{k}{2}\sum_\alpha \sum_{m,n} n_\alpha \mathbf{a}_\alpha^{mn} \mathbf{a}_\alpha^{mn}, \tag{47}$$

$$\begin{aligned}\mathbf{J}_s &= \frac{1}{T}\left(\mathbf{q} - \sum_\alpha \mu_\alpha \mathbf{J}_\alpha\right) \\ &- \frac{k}{2}\sum_\alpha \sum_{m,n}\sum_{k,l} n_\alpha \mathbf{a}_\alpha^{mn}(\mathbf{P}_\alpha^{mn}, \mathbf{c}_\alpha \mathbf{P}_\alpha^{kl})\mathbf{a}_\alpha^{kl},\end{aligned} \tag{48}$$

$$\begin{aligned}\sigma &= -k\sum_\alpha \sum_m \sum_n n_\alpha \mathbf{a}_\alpha^{mn}(\mathbf{P}_\alpha^{mn}, L\phi_\alpha) \\ &= -k\sum_\alpha \sum_m \sum_n n_\alpha \mathbf{a}_\alpha^{mn} \mathbf{R}_\alpha^{mn}.\end{aligned} \tag{49}$$

In the last expression, R_α^{mn} can be replaced by the left-hand part of Eq. (22). Then, using formulas (25) and (26), we arrive at

$$\sigma = \sigma^{(1)} + \sigma^{(2)}, \tag{50}$$

$$\sigma^{(1)} = -k\sum_\alpha \sum_m \sum_n n_\alpha \mathbf{a}_\alpha^{mn}[(\mathbf{P}_\alpha^{mn}, D_\alpha \ln f_\alpha^{(0)})], \tag{51}$$

$$\begin{aligned}\sigma^{(2)} = -k\sum_\alpha \sum_m \sum_n n_\alpha \mathbf{a}_\alpha^{mn}\Bigg[&\frac{dn_\alpha \mathbf{a}_\alpha^{mn}}{dt} \\ &+ \sum_l (A_{\alpha mn}^{m+1,l}\nabla n_\alpha \mathbf{a}_\alpha^{m+1,l} + B_{\alpha mn}^{m-1,l}\overline{\nabla n_\alpha \mathbf{a}_\alpha}^{m-1,l})\Bigg].\end{aligned} \tag{52}$$

5. GENERALIZED NONEQUILIBRIUM THERMODYNAMICS

The entropy balance equation (41), together with the relations (42), (45), (46) or (47)–(49) constitute the basis of the generalized nonequilibrium thermodynamics of multicomponent systems, in which the state of a mixture is determined not only by local values of the density, mass-average velocity, and temperature, but also by an arbitrary number of additional state variables. The latter variables represent the coefficients of expansion of the nonequilibrium corrections to the distribution function ($\mathbf{a}_\alpha^{mn}$) or the corresponding moments of the distribution function. For a simple single-component gas ($N = 1$), the above derived expressions correspond to our results obtained previously [9, 10]. Analogous equations for a single-component gas were earlier considered in [1, 2] and, for the particular case of a gas mixture of Maxwellian molecules [13], within the framework of the extended irreversible thermodynamics developed in recent years.

Let us turn to relations (50)–(52) for the entropy production. In order to calculate $\sigma^{(1)}$, we can use an expression for $D_\alpha \ln f_\alpha^{(0)}$. This yields

$$\sigma^{(1)} = [\sigma^{(1)}]_0 + [\sigma^{(1)}]_1, \tag{53}$$

where

$$[\sigma^{(1)}]_0 = -\frac{1}{T^2}(\mathbf{h}, \nabla T) - \frac{1}{T}\sum_{\alpha=1}^{N}\left(\frac{\mathbf{J}_\alpha}{\rho_\alpha}, p\mathbf{d}_\alpha\right) - \frac{1}{T}\hat{\boldsymbol{\pi}} : \overline{\nabla \mathbf{u}}, \tag{54}$$

$$[\sigma^{(1)}]_1 = \frac{1}{T}\sum_{\alpha=1}^{N}\left(\frac{\mathbf{J}_\alpha}{\rho_\alpha}, \frac{\rho_\alpha}{\rho}\nabla\hat{\boldsymbol{\pi}}\right) + \frac{1}{T}\sum_{\alpha=1}^{N}\frac{n_\alpha}{n}\frac{T_\alpha - T}{T}\left(\nabla\mathbf{q} - \frac{3}{2}kT\sum_{\beta=1}^{N}\nabla\frac{\mathbf{J}_\beta}{\rho_\beta}\right). \tag{55}$$

The quantity

$$\mathbf{h} = \sum_\alpha \mathbf{h}_\alpha = \mathbf{q} - \frac{5}{2}\sum_\alpha p_\alpha \mathbf{w}_\alpha \tag{56}$$

represents the reduced total heat flux in the mixture, and $\mathbf{d}_\alpha$ is called the diffusive thermodynamic force [19, 22] and is defined as

$$\mathbf{d}_\alpha = \nabla y_\alpha + \left(y_\alpha - \frac{\rho_\alpha}{\rho}\right)\nabla \ln p - p^{-1}\left(n_\alpha \mathbf{F}_\alpha - \frac{\rho_\alpha}{\rho}\sum_\alpha n_\alpha \mathbf{F}_\alpha\right), \tag{57}$$

where $y_\alpha = n_\alpha/n$ is the relative number density of component α.

The expression for $[\sigma^{(1)}]_0$ exactly coincides with the well-known representation of the local entropy production in the classical LNT [22] in the form of a bilinear combination of the fluxes $\hat{\boldsymbol{\pi}}$, $\mathbf{h}$, and $\mathbf{w}_\alpha = \mathbf{J}_\alpha/\rho_\alpha$) and the conjugated thermodynamic forces $\overline{\nabla \mathbf{u}}$, $(1/T)\nabla T$, and $p\mathbf{d}_\alpha$. This expression essentially follows from (51), where $D_\alpha \ln f_\alpha^{(0)}$ is replaced by expression (13) with $d\mathbf{u}/dt$ and dT/dt taken from the equations of conservation (39) and (40) in the Euler form, whereby all dissipative terms proportional to the spatial derivatives of the fluxes are omitted. As can be readily seen, this is equivalent to using the linearized kinetic equation (7) with only the $D_\alpha \ln f_\alpha^{(0)}$ term retained in the left-hand part, which corresponds to the well-known first approximation of the Chapman–Enskog method [16, 19]. This approximation is known to yield results coinciding with conclusions of the classical LNT [22]. Taking into account the condition

$$\sum_{\alpha=1}^{N}\mathbf{J}_\alpha = 0,$$

the expression for $[\sigma^{(1)}]_0$ can be represented in an alternative form

$$[\sigma^{(1)}]_0 = -\frac{1}{T^2}(\mathbf{h}, \nabla T) - \frac{1}{T}\sum_{\alpha=1}^{N-1}\left(\left(\frac{\mathbf{J}_\alpha}{\rho_\alpha} - \frac{\mathbf{J}_N}{\rho_N}\right), p\mathbf{d}_\alpha\right) - \frac{1}{T}\hat{\boldsymbol{\pi}} : \overline{\nabla \mathbf{u}}. \tag{58}$$

The linear phenomenological relations corresponding to this representation appear as

$$\begin{aligned}\mathbf{h} &= -L_{00}\frac{\nabla T}{T} - \sum_{\beta=1}^{N-1}L_{0\beta}p\mathbf{d}_\beta,\\ \frac{\mathbf{J}_\alpha}{\rho_\alpha} - \frac{\mathbf{J}_N}{\rho_N} &= -L_{\alpha 0}\frac{\nabla T}{T} - \sum_{\beta=1}^{N-1}L_{\alpha\beta}p\mathbf{d}_\beta,\\ \hat{\boldsymbol{\pi}} &= -2K_{00}\overline{\nabla \mathbf{u}},\end{aligned} \tag{59}$$

where $L_{\alpha\beta}$ obey the Onsager relations [22]

$$L_{\alpha\beta} = L_{\beta\alpha}, \quad L_{\alpha 0} = L_{0\alpha}. \tag{60}$$

In the entropy production part $[\sigma^{(1)}]_1$, the thermodynamic forces are represented by spatial derivatives of the fluxes, appearing upon substitution of the equations of conservation (38)–(40) into expression (13). Analogous terms, proportional to the spatial derivatives of partial fluxes, enter into the expression for $\sigma^{(2)}$. Combining the corresponding terms from all such expressions leads to overdetermination and a certain complication of the thermodynamic forces. Nevertheless, an expression for the entropy production still retains the

form of a bilinear combination of the generalized thermodynamic forces and fluxes.

Another important feature of the total expression (50) for the entropy production is the presence of time derivatives $dn_\alpha \mathbf{a}_\alpha^{mn}/dt$. Note that, in the case of a spatially homogeneous gas mixture, the expression for σ takes the form

$$\sigma = -k\sum_\alpha \sum_m \sum_n n_\alpha \mathbf{a}_\alpha^{mn} \frac{dn_\alpha \mathbf{a}_\alpha^{mn}}{dt}. \quad (61)$$

This form corresponds quite well to the LNT version as formulated by Onsager [22]. In this case, the entropy production is described by an expression bilinear in fluxes and conjugated thermodynamic forces, whereby the fluxes are represented by the derivatives $dn_\alpha \mathbf{a}_\alpha^{mn}/dt$, and the forces, by the coefficients $-kn_\alpha \mathbf{a}_\alpha^{mn}$. In the Onsager theory, the thermodynamic forces are defined as the partial derivatives of the entropy with respect to the corresponding state variables $\mathbf{a}_\alpha^{mn}$. The quadratic form of expression (47) for the nonequilibrium entropy with respect to the state variables is just what leads to such a definition. According to the Onsager theory generalized to the case of multicomponent systems, the expression for the entropy production (61) must correspond to linear phenomenological relations of the type

$$\frac{dn_\alpha \mathbf{a}_\alpha^{mn}}{dt} = \sum_\beta \sum_l \Lambda_{\alpha\beta}^{mnl} \mathbf{a}_\beta^{ml}. \quad (62)$$

As can be readily seen, relations (62) in the case of a spatially homogeneous system are in perfect agreement with the equations of moments (36), provided that $\Lambda_{\alpha\beta}^{mnl} = -C_{\alpha\beta}^{mnl}$. Validity of the Onsager relations for the kinetic cross coefficients can be shown to follow from the symmetry relations (11) for the linearized operator of collisions. Thus, the method of moments provides for an independent corroboration of the Onsager theory in the case of spatially homogeneous states of the gas mixtures. In application to simple gases, this problem was previously discussed in [5, 6, 24].

More complicated is the situation with analysis of the GNT for spatially inhomogeneous systems. In this case, according to a conventionally accepted scheme, the fluxes are represented by the variables $n_\alpha \mathbf{a}_\alpha^{mn}$, and the thermodynamic forces, by gradients of the initial macroscopic parameters of the mixture (see the expression for $[\sigma^{(1)}]_0$). Here, according to relations (20), the first several coefficients $n_\alpha \mathbf{a}_\alpha^{mn}$ correspond to the fluxes possessing a clear physical sense (e.g., to the diffusion flux $\mathbf{J}_\alpha$, the reduced heat flux $\mathbf{h}_\alpha$, and the partial viscous stress tensor $\hat{\boldsymbol{\pi}}_\alpha$, as well as the relative temperature difference $(T_\alpha - T)/T$. Within the framework of the GNT, thermodynamic forces conjugated with these fluxes are complicated by an additional contribution related to derivatives of the fluxes with respect to time and spatial coordinates. In addition, there is an increase in number of the state variables for which their own phenomenological transfer equations can be written. Note that bilinearity of the expression for the entropy production implies the possibility of various definitions of the fluxes and conjugated thermodynamic forces, whereby these quantities are mutually interchanged. Below, we will analyze this situation in more detail for the well-known approximation of 13*N* moments.

6. THE APPROXIMATION OF 13*N* MOMENTS

In this approximation, expansion of the distribution function (15) retains the terms with the coefficients $\mathbf{a}_\alpha^{01}$, $\mathbf{a}_\alpha^{10}$, $\mathbf{a}_\alpha^{11}$, and $\mathbf{a}_\alpha^{20}$. The equations of moments are written for the variables n_α (or ρ_α), $\mathbf{u}$, and T, entering into the weighting function (local Maxwell distribution) corresponding to the equations of conservation (38)–(40), and for the expansion coefficients expressed through moments of the distribution function by formulas (20). Here, or we will digress from the case of scalar quantities, which will be specially considered below. This implies that $a_\alpha^{01} = 0$ and $T_\alpha = T$ for all components of the gas mixture.

Turning to expressions for the local density of entropy and entropy flux and the entropy production in the approximation adopted, we obtain from (47) and (48)

$$\rho s = \rho s_0 - \frac{1}{2T}\sum_\alpha \frac{1}{\rho_\alpha}\mathbf{J}_\alpha \mathbf{J}_\alpha - \frac{1}{4T}\sum_\alpha \frac{1}{p_\alpha}\hat{\boldsymbol{\pi}}_\alpha \hat{\boldsymbol{\pi}}_\alpha - \frac{1}{5T}\sum_\alpha \frac{\rho_\alpha}{p_\alpha^2}\mathbf{h}_\alpha \mathbf{h}_\alpha, \quad (63)$$

$$\mathbf{J}_s = \frac{1}{T}\left(\mathbf{q} - \sum_\alpha \mu_\alpha \mathbf{J}_\alpha\right) - \frac{2}{5}\frac{1}{T}\sum_\alpha \frac{1}{p_\alpha}\hat{\boldsymbol{\pi}}_\alpha \mathbf{q}_\alpha. \quad (64)$$

According to (50)–(52), the entropy production is expressed as

$$\sigma = -\frac{1}{T}\sum_{\alpha=1}^{N-1}\left(\left(\frac{\mathbf{J}_\alpha}{\rho_\alpha} - \frac{\mathbf{J}_N}{\rho_N}\right), \left(\frac{d\rho_\alpha \mathbf{w}_\alpha}{dt} + p\mathbf{d}_\alpha^*\right)\right) - \frac{1}{T}\sum_{\alpha=1}^{N}\left(\mathbf{h}_\alpha, \left(\frac{2}{5}\frac{m_\alpha}{p_\alpha kT}\frac{d\mathbf{h}_\alpha}{dt} + \frac{1}{T}\nabla T + \frac{2}{5p_\alpha}\nabla\hat{\boldsymbol{\pi}}_\alpha\right)\right) - \frac{1}{T}\sum_{\alpha=1}^{N}\pi_{\alpha ik}\left(\frac{1}{p_\alpha}\frac{d\pi_{\alpha ik}}{dt} + 2\varepsilon_{ik} + \frac{4}{5p_\alpha}\left\{\frac{\partial q_{\alpha i}}{\partial x_k}\right\}\right), \quad (65)$$

where

$$p\mathbf{d}_\alpha^* = p\mathbf{d}_\alpha + \nabla\hat{\boldsymbol{\pi}}_\alpha - (\rho_\alpha/\rho)\nabla\hat{\boldsymbol{\pi}} \quad (66)$$

and

$$\{A_iB_k\} = \frac{1}{2}(A_iB_k + B_iA_k) - \frac{1}{3}A_lB_l\delta_{ik},$$

$$\varepsilon_{ik} = \left\{\frac{\partial u_i}{\partial x_k}\right\}$$

(repeated italic subscripts imply summation).

Bilinearity of the expression for the entropy production indicates that the fluxes and conjugated thermodynamic forces can be selected in two alternative ways. The first possibility is to identify generalized thermodynamic forces with the expressions in parentheses in each term of the general expression for σ (65). In this case, the corresponding phenomenological relations acquire the following form:

$$\mathbf{h}_\alpha = -\sum_{\beta=1}^{N}\hat{L}_{\alpha\beta}\left(\frac{2}{5}\frac{m_\beta}{p_\beta kT}\frac{d\mathbf{h}_\beta}{dt} + \frac{1}{T}\nabla T + \frac{2}{5p_\beta}\nabla\hat{\boldsymbol{\pi}}_\beta\right) - \sum_{\beta=1}^{N-1}\check{L}_{\alpha\beta}\left(\frac{d\rho_\beta\mathbf{w}_\beta}{dt} + p\mathbf{d}_\beta^*\right), \quad (67)$$

$$\frac{\mathbf{J}_\alpha}{\rho_\alpha} - \frac{\mathbf{J}_N}{\rho_N} = -\sum_{\beta=1}^{N}\tilde{L}_{\alpha\beta}\left(\frac{2}{5}\frac{m_\beta}{p_\beta kT}\frac{d\mathbf{h}_\beta}{dt} + \frac{1}{T}\nabla T + \frac{2}{5p_\beta}\nabla\hat{\boldsymbol{\pi}}_\beta\right) - \sum_{\beta=1}^{N-1}L_{\alpha\beta}\left(\frac{d\rho_\beta\mathbf{w}_\beta}{dt} + p\mathbf{d}_\beta^*\right), \quad (68)$$

$$\pi_{\alpha ik} = -\sum_{\beta=1}^{N}K_{\alpha\beta}\left(\frac{1}{p_\beta}\frac{d\pi_{\beta ik}}{dt} + 2\varepsilon_{ik} + \frac{4}{5p_\beta}\left\{\frac{\partial q_{\beta i}}{\partial x_k}\right\}\right). \quad (69)$$

Note that the quantities $L_{\alpha\beta}$ coincide with the corresponding coefficients in Eqs. (59) and

$$\sum_{\alpha,\beta=1}^{N}\hat{L}_{\alpha\beta} = L_{00}, \quad \sum_{\alpha,\beta=1}^{N}K_{\alpha\beta} = K_{00}, \quad \sum_{\alpha=1}^{N}\check{L}_{\alpha\beta} = L_{0\beta}.$$

The alternative form of writing equations employs the possibility of interchanging the fluxes and forces in expression (65) for σ. In this case, expressions in the parentheses in each term of the general expression for the entropy production correspond to fluxes, rather than to the thermodynamic forces. The corresponding linear phenomenological relations are as follows:

$$\frac{d\rho_\alpha\mathbf{w}_\alpha}{dt} + p\mathbf{d}_\alpha^* = -\sum_{\beta=1}^{N}\Lambda_{\alpha\beta}\frac{\mathbf{J}_\beta}{\rho_\beta} - \sum_{\beta=1}^{N}\hat{\Lambda}_{\alpha\beta}\frac{\mathbf{h}_\beta}{p_\beta}, \quad (70)$$

$$\frac{d\mathbf{h}_\alpha}{dt} + \frac{5}{2}\frac{k}{m_\alpha}p_\alpha\nabla T + \frac{kT}{m_\alpha}\nabla\hat{\boldsymbol{\pi}}_\alpha = -\frac{5}{2}\frac{kT}{m_\alpha}p^2\left(\sum_{\beta=1}^{n}\check{\Lambda}_{\alpha\beta}\frac{\mathbf{J}_\beta}{\rho_\beta} + \sum_{\beta=1}^{N}\tilde{\Lambda}_{\alpha\beta}\frac{\mathbf{h}_\beta}{p_\beta}\right), \quad (71)$$

$$\frac{d\pi_{\alpha ik}}{dt} + 2p_\alpha\varepsilon_{ik} + \frac{4}{5}\left\{\frac{\partial q_{\alpha i}}{\partial x_k}\right\} = -p^2\sum_{\beta=1}^{N}\Omega_{\alpha\beta}\frac{\pi_{\beta ik}}{p_\beta}. \quad (72)$$

It should be noted that equations of an analogous structure follow directly from the system of equations of moments (36), whereby coefficients in the right-hand parts of equations no longer represent phenomenological quantities and are expressed in terms of the integral brackets of the Sonin polynomials (34). The corresponding explicit expressions for the right-hand parts of Eqs. (70)–(72) in the approximation of 13*N* moments were obtained in [12, 25].

There is one important circumstance that is worth mentioning: the method of moments encounters a problem of selecting polynomials to be used for constructing a system of the equations of moments for a given approximation of the distribution function. As can be seen from the above considerations, nonequilibrium thermodynamics allows a unique system of equations of moments to be constructed using only the form of the approximating function: it follows from relations (70)–(72) that the required system should be constructed using the moments in which the distribution function is expanded.

This approach offers two advantages. First, the entropy balance equation is exactly satisfied when the entropy, entropy flux, and entropy production are calculated using the approximate distribution function (the condition of strict validity of the entropy balance equation is analogous to the requirement of observation of the laws of mass, momentum, and energy conservation usually posed on approximate equations). Second, for this construction of a system of equations of moments, the Onsager symmetry relations are strictly obeyed both for the local kinetic coefficients of the equations of type (70)–(72) and for the integral kinetic coefficients in the macroscopic phenomenological equations, irrespective of the accuracy of calculations [26].

Let us consider a simple gas. In this case, Eqs. (67) and (69) can be reduced to the following form:

$$\tau_\lambda\frac{d\mathbf{q}}{dt} + \mathbf{q} = -\lambda\left(\nabla T + \frac{2}{5}\frac{T}{p}\nabla\boldsymbol{\pi}\right), \quad (73)$$

$$\tau_\eta \frac{d\hat{\boldsymbol{\pi}}}{dt} + \hat{\boldsymbol{\pi}} = -2\eta\left(\overline{\nabla \mathbf{u}} + \frac{2}{5}\frac{1}{p}\overline{\nabla \mathbf{q}}\right), \tag{74}$$

where $L_{00} = \lambda T$ and $K_{00} = \eta$. The quantities η and λ represent the coefficients of viscosity and thermal conductivity, respectively, related to the corresponding relaxation times by the formulas

$$\eta = p\tau_\eta, \quad \lambda = \frac{5k}{2m}p\tau_\lambda, \quad \tau_\lambda = \frac{15}{4}\tau_\eta. \tag{75}$$

Note that equations analogous to (73) and (74) also follow from relations (71) and (72), where the case of a simple gas corresponds to $\Omega_{11} = \eta^{-1}$ and $\tilde{\Lambda}_{11} = (\lambda T)^{-1}$.

Equation (73) with $\nabla\hat{\boldsymbol{\pi}} = 0$ is known as the Maxwell–Cattaneo equation [27]. Relations analogous to (73) and (74) were also considered within the framework of the extended irreversible thermodynamics [1, 2]. The presence of relaxation terms in these equations allows the propagation of thermal and shear perturbations in the gas to be analyzed in a more justified manner. The corresponding equations for the temperature and velocity acquire the form of hyperbolic differential equations. This eliminates the paradox of infinite velocity of the propagation of thermal and shear perturbations, which is inherent in the usual linear relations of classical nonequilibrium thermodynamics (the Fourier and Newton laws). Indeed, using Eq. (73) (with neglect of $\nabla\hat{\boldsymbol{\pi}}$) together with the energy conservation equation (40) leads to an equation of the type

$$\tau_\lambda \frac{\partial^2 T}{\partial t^2} + \frac{\partial T}{\partial t} = \frac{\lambda}{\rho c_v}\Delta T. \tag{76}$$

This telegraph equation predicts a temperature wave propagating at a velocity

$$V_T^0 = (\lambda/\rho c_v \tau_\lambda)^{1/2} = (5p/3\rho)^{1/2}.$$

An analogous equation for the pressure p, which follows from (39) and (74), describes a transverse shear wave propagating at a velocity

$$V_p^0 = (\eta/\rho\tau_\eta)^{1/2} = (p/\rho)^{1/2}.$$

An allowance for $\nabla\boldsymbol{\pi}$ and $\nabla\mathbf{q}$ in Eqs. (73) and (74) yields finite corrections to these velocities and their dispersions. In particular, the velocity of propagation of the transverse shear wave in a gas with a homogeneous temperature field in a high-frequency approximation is [1]

$$V_p^1 = (7p/5\rho)^{1/2}.$$

In the case when $\tau_\eta \ll \tau_L$ (τ_L is a characteristic time of the problem), the terms containing derivatives of the fluxes with respect to time can be omitted and we arrive at the well-known linear relations for $\mathbf{q}$ and $\hat{\boldsymbol{\pi}}$ [9, 11].

In the case of established viscous flows, the general equation of motion (39) yields $\nabla\hat{\boldsymbol{\pi}} = -\nabla p$ and the expression for $\mathbf{q}$ acquires the following form [9]:

$$\mathbf{q} = -\lambda\left(\nabla T - \frac{2}{5}\frac{T}{p}\nabla p\right).$$

Thus, an allowance for the spatial derivatives of fluxes in the transfer equations may lead to the appearance of additional terms linear in gradients of the thermodynamic parameters.

In the case of a multicomponent mixture, an allowance for the corresponding terms leads to the generalized linear relations for π_{ik} and $\pi_{\alpha ik}$, which can be presented in the following form (for simplicity, the nonstationary terms are omitted):

$$\begin{aligned}
\pi_{ik} &= -2K_{00}\varepsilon_{ik} + \sum_{\beta=1}^{N} K_{0\beta}\frac{4}{5p_\beta}\left\{\frac{\partial q_{\beta i}}{\partial x_k}\right\}, \\
\pi_{\alpha ik} &= -2K_{\alpha 0}\varepsilon_{ik} + \sum_{\beta=1}^{N} K_{\alpha\beta}\frac{4}{5p_\beta}\left\{\frac{\partial q_{\beta i}}{\partial x_k}\right\},
\end{aligned} \tag{77}$$

where $K_{\alpha 0} = K_{0\alpha}$. Analogous expressions can be written for $\mathbf{h} = \sum_\alpha \mathbf{h}_\alpha$ and $\mathbf{h}_\alpha$, in which case the Onsager symmetry relations are valid for the coefficients $L_{\alpha\beta}$ ($\hat{L}_{\alpha 0} = \hat{L}_{0\alpha}$).

Let us consider an equation following from relation (70) for the diffusion velocities of components. In the simplest case, when it is possible to ignore a contribution of the partial thermal fluxes in the right-hand part of (70) (this assumption is strictly valid for a mixture of Maxwellian molecules), the diffusion equation is as follows:

$$\frac{d\rho_\alpha \mathbf{w}_\alpha}{dt} + p\mathbf{d}_\alpha^* = -\sum_{\beta=1}^{N-1}\Lambda_{\alpha\beta}(\mathbf{w}_\alpha - \mathbf{w}_N). \tag{78}$$

The right-hand part of this equation is conveniently presented in the form whereby relations for determining the diffusion rates take the form known as the Stefan–Maxwell equation. From the conditions

$$\sum_{\alpha=1}^{N}\rho_\alpha \mathbf{w}_\alpha = 0, \quad \sum_{\alpha=1}^{N}\mathbf{d}_\alpha^* = 0,$$

it follows that

$$\sum_{\alpha=1}^{N}\sum_{\beta=1}^{N}\Lambda_{\alpha\beta}(\mathbf{w}_\beta - \mathbf{w}_N) = 0.$$

This is possible only provided that

$$\sum_{\alpha=1}^{N}\Lambda_{\alpha\beta} = \sum_{\beta=1}^{N}\Lambda_{\beta\alpha} = 0$$

or

$$\Lambda_{\alpha\alpha} = -\sum_{\beta\neq\alpha}^{N}\Lambda_{\beta\alpha}.$$

As a result, Eq. (78) acquires the following form:

$$\frac{d\rho_\alpha \mathbf{w}_\alpha}{dt} + p\mathbf{d}_\alpha^* = -\sum_{\beta\neq\alpha}^{N}\Lambda_{\alpha\beta}(\mathbf{w}_\alpha - \mathbf{w}_\beta). \qquad (79)$$

A comparison to the results of the kinetic theory [12, 25] yields

$$\Lambda_{\alpha\beta} = n_\alpha\mu_{\alpha\beta}\tau_{\alpha\beta}^{-1} = \frac{n_\alpha n_\beta kT}{n[D_{\alpha\beta}]_1}, \qquad (80)$$

where $\tau_{\alpha\beta}^{-1}$ is the effective frequency of collisions,

$$\mu_{\alpha\beta} = \frac{m_\alpha m_\beta}{m_\alpha + m_\beta}$$

is the reduced mass of particles of the α and β types, and $[D_{\alpha\beta}]_1$ is the coefficient of binary diffusion (the first Chapman–Cowling approximation) [16].

With neglect of the derivative $d\rho_\alpha\mathbf{w}_\alpha/dt$, that is, when $\tau_{\alpha\beta} \ll \tau_L$, and for $\mathbf{d}_\alpha^* = \mathbf{d}_\alpha$, Eq. (79) corresponds to a usual equation for the diffusion rates written in the Stefan–Maxwell form. Taking into account the time derivative of $\rho_\alpha\mathbf{w}_\alpha$ in Eq. (79) and considering this equation together with the continuity equation (38), we arrive at a hyperbolic equation for the particle concentration and obtain the following rate of propagation of the concentration perturbation:

$$V_c^0 = \left(\frac{\rho\mu_{12}[D_{12}]_1}{\rho_2 m_1\tau_{12}}\right)^{1/2} = \left(\frac{\rho kT}{nm_1m_2}\right)^{1/2}.$$

An interesting effect is observed with an allowance for the spatial derivatives of the viscous stress tensors in expression (66) for $p\mathbf{d}_\alpha^*$. In this case, a solution to Eqs. (72) with neglect of the contribution due to derivatives of the heat flux yields for the established viscous flows

$$\hat{\boldsymbol{\pi}}_\alpha = -2\eta_\alpha\overline{\nabla\mathbf{u}}, \quad \hat{\boldsymbol{\pi}} = -2\eta\overline{\nabla\mathbf{u}}, \qquad (81)$$

where η_α and η are the partial and total viscosity coefficients of the mixture [12]. Then, using relation $2\eta\nabla^2\mathbf{u} = \nabla p$ and omitting the time derivative of $\rho_\alpha\mathbf{w}_\alpha$, Eq. (79) can be reduced to

$$\nabla p_\alpha - \frac{\eta_\alpha}{\eta}\nabla p = \sum_{\beta\neq\alpha}^{N}\frac{n_\alpha n_\beta kT}{n[D_{\alpha\beta}]_1}(\mathbf{w}_\alpha - \mathbf{w}_\beta). \qquad (82)$$

Thus, an allowance for the spatial derivatives of $\hat{\boldsymbol{\pi}}_\alpha$ and $\hat{\boldsymbol{\pi}}$ in Eq. (79) leads to redefining the pressure diffusion constant, which becomes essentially a kinetic quantity [12].

In a more general case, the equations of transfer for a multicomponent gas mixture can be written using an iterative procedure with respect to a small parameter—the Knudsen number Kn = l/L, where l is the effective mean free path length of gas molecules and L is a characteristic linear parameter of the problem. In this procedure, the first approximation employs relations (59) based on the concept of entropy production (58). In the next approximation, the stationary transfer relations in the absence of external forces take the following form:

$$\sum_{\beta\neq\alpha}\frac{n_\alpha n_\beta kT}{n[D_{\alpha\beta}]_2}(\mathbf{w}_\alpha - \mathbf{w}_\beta) = -\left(\nabla p_\alpha - \frac{\rho_\alpha}{\rho}\nabla p\right) + \left(\eta_\alpha - \frac{\rho_\alpha}{\rho}\eta\right)\nabla^2\mathbf{u} + \sum_\beta\xi_{\alpha\beta}\left(\frac{\lambda_\alpha}{\rho_\alpha} - \frac{\lambda_\beta}{\rho_\beta}\right)\nabla T, \qquad (83)$$

$$\boldsymbol{\pi} = -2\eta\overline{\nabla\mathbf{u}} + \frac{4}{5}\sum_\alpha\frac{\eta_\alpha\lambda_\alpha}{p_\alpha}\overline{\nabla\nabla}T, \qquad (84)$$

$$\mathbf{q} = \frac{5}{2}\sum_\alpha p_\alpha\mathbf{w}_\alpha - \lambda\nabla T + T\sum_\alpha\sum_\beta\frac{\lambda_\alpha}{\rho_\alpha}\xi_{\alpha\beta}(\mathbf{w}_\alpha - \mathbf{w}_\beta) + \frac{2}{5}T\sum_\alpha\frac{n_\alpha\lambda_\alpha}{\rho_\alpha}\nabla^2\mathbf{u}. \qquad (85)$$

Here, λ_α is the partial coefficient of thermal conductivity [12, 25], $[D_{\alpha\beta}]_2$ is the binary diffusion coefficient (the second Chapman–Cowling approximation) [16, 19], and

$$\xi_{\alpha\beta} = \frac{n_\alpha n_\beta}{n[D_{\alpha\beta}]_1}\mu_{\alpha\beta}\left(\frac{6}{5}C_{\alpha\beta}^* - 1\right),$$

where the parameter $C_{\alpha\beta}^*$ is determined by the character of interparticle interaction [19]; in particular, for Maxwellian molecules,

$$(6/5)C_{\alpha\beta}^* - 1 = 0,$$

which implies vanishing of the terms related to thermal diffusion in the gas mixture.

Note the following circumstance: considering relations (83)–(85) as the phenomenological equations of nonequilibrium thermodynamics, generalizing Eqs. (59) to the case of thermodynamic forces of higher approximation and corresponding to Eqs. (67)–(69), we encounter the violation of the Onsager symmetry relations for the terms with second-order derivatives. Breakage of the Onsager symmetry inspired some researchers [17, 18] to suggest that nonequilibrium thermodynamics becomes inapplicable on

the passage to the Barnett approximation of the kinetic theory of gases. Misconception of this conclusion was demonstrated [5–9] by calculation of the entropy production based on the distribution function of the Barnett approximation and by construction of the corresponding phenomenological equations, which exhibit no discrepancies with nonequilibrium thermodynamics, provided that some nonphysical fluxes (see below) are naturally introduced into consideration.

In concluding this section, let us discuss the problem of a difference between the component temperatures T_α and the overall mixture temperature T. According to (61), the relative temperature difference can be expressed by a phenomenological relation of the type

$$\frac{T_\alpha - T}{T} = H_{\alpha 0}\frac{n_\alpha}{n}\left(\nabla \mathbf{q}_\alpha - \frac{3}{2}kT\sum_{\alpha=1}^{N}\nabla\frac{\mathbf{J}_\alpha}{\rho_\alpha}\right). \quad (86)$$

According to this, deviations of T_α from T are determined by the divergences of vector fluxes or by the second derivatives of the temperature and concentration, which usually corresponds to values of the second order of smallness with respect to the parameter Kn = l/L. However, this situation takes place only for a gas mixture of monoatomic molecules. In a mixture of diatomic and polyatomic gases, the molecules of which possess internal degrees of freedom, the quantity T_α in the transfer equations is replaced by T_α^{tr}, the temperature of the translational degrees of freedom. The latter value exhibits relaxation to the overall translational temperature of the mixture, T^{tr}, during a time period on the order of the effective elastic collision time $\tau_{\alpha\beta}$ of molecules. The value of T^{tr} differs from the temperatures of the internal degrees of freedom, T_α^{in}, which leads to the appearance of an additional scalar term in the diagonal part of the pressure tensor, this term being proportional to the divergence of the mass-average velocity (volume viscosity) (see [28]).

7. EQUATIONS OF GENERALIZED NONEQUILIBRIUM THERMODYNAMICS IN THE BARNETT APPROXIMATION OF THE CHAPMAN–ENSKOG METHOD

Previously [9], we demonstrated the possibility of reducing the GNT equations, obtained by the method of moments for a simple gas, to the corresponding equations of the linearized Barnett approximation of the Chapman–Enskog method. This procedure employed perturbation theory with respect to a small parameter (the Knudsen number) of the system of equations of moments (applied in the Chapman–Enskog method to the distribution function). Below, we show how to pass from equations of the method of moments directly to the Chapman–Enskog equations for a multicomponent gas mixture and how to derive an expression for the entropy production corresponding to a generalized system of phenomenological equations in the Barnett approximation. The approach employed here is somewhat different from that used in [9].

We proceed from the system of equations of moments (36) and the equations of conservation (38)–(40). Let us represent the coefficients a_α^{mn} in the form of a series

$$\mathbf{a}_\alpha^{mn} = \mathbf{a}_\alpha^{mn}(1) + \mathbf{a}_\alpha^{mn}(2) + \ldots. \quad (87)$$

Since the right-hand parts of the equations of moments are linear combinations of terms on the order of $\tau_{\alpha\beta}^{-1} n_\alpha \mathbf{a}_\alpha^{mn}$, where $\tau_{\alpha\beta}$ is the effective mean time between the collisions of molecules, expansion (87) is conveniently considered as a formal expansion with respect to a small parameter $\varepsilon_{\alpha\beta} = \langle v_{\alpha\beta}\rangle \tau_{\alpha\beta} L^{-1}$, where $\langle v_{\alpha\beta}\rangle$ is the average relative velocity of particles of the types α and β (for a simple gas, a small parameter in this case is the Knudsen number Kn = l/L, where $l = \langle v\rangle\tau$). Let us assume that the order of each succeeding term in expansion (87) with respect to ε decreases by unity. In addition, we assume (as in the Chapman–Enskog method) that the equations of conservation (38)–(40) are valid in each step with an accuracy corresponding to the approximation order number minus unity.

Let us consider first the equations of moments (36) for vector quantities ($m = 1$). To the first approximation, replacing $d\mathbf{u}/dt$ by expressions according to the equations of motion (39) in the Euler form, we obtain

$$\frac{\sqrt{2}}{2}\beta_\alpha^{-1/2}\frac{1}{p_\alpha}p\mathbf{d}_\alpha\delta_{n0} + \sqrt{\frac{5}{4}}\beta_\alpha^{-1/2}\nabla\ln T\delta_{n1} = -\sum_{\beta=1}^{N}\sum_{l=0}^{\infty} C_{\alpha\beta}^{1nl}\mathbf{a}_\beta^{1l}(1), \quad (88)$$

where $\mathbf{d}_\alpha$ is given by expression (57). Multiplying both sides of Eq. (88) by $f_\alpha^{(0)}\mathbf{P}_\alpha^{ln}(\mathbf{W}_\alpha)$, summing the products over n, and taking into account the property of orthogonal polynomials forming a complete system,

$$\sum_n \mathbf{P}_\alpha^{mn}(\mathbf{W}_\alpha) f_\alpha^{(0)}(\mathbf{W}_\alpha)\mathbf{P}_\alpha^{mn}(\mathbf{W}'_\alpha) = \delta(\mathbf{W}_\alpha - \mathbf{W}'_\alpha), \quad (89)$$

we arrive at the equation

$$\frac{1}{p_\alpha}\mathbf{c}_\alpha p\mathbf{d}_\alpha + \mathbf{c}_\alpha\left(\beta_\alpha c_\alpha^2 - \frac{5}{2}\right)\nabla\ln T = \sum_{\beta=1}^{N} L_{\alpha\beta}\left(\sum_l \mathbf{a}_\beta^{1l}(1)\mathbf{P}_\beta^{1l}\right). \quad (90)$$

As expected, this equation corresponds to the first approximation for vector quantities in the Chapman–Enskog method [19, 20]. Naturally, this equation can also be obtained directly, proceeding from the initial

linearized kinetic equation (7) and using the standard procedure of the Chapman–Enskog method. The above approach is more convenient, since it will be also employed for obtaining the following approximations. Applying this procedure to the equations of moments for the tensor component ($m = 2$), we arrive at

$$2\beta_\alpha \overline{\mathbf{c}_\alpha \mathbf{c}_\alpha} : \overline{\nabla \mathbf{u}} = \sum_{\phi=1}^{N} L_{\alpha\beta}\left(\sum_l \mathbf{a}_\beta^{2l}(1)\mathbf{P}_\beta^{2l}\right). \quad (91)$$

A solution to Eqs. (90) and (91) can be represented in the standard form [20]

$$\sum_l \mathbf{a}_\beta^{2l}(1)\mathbf{P}_\beta^{1l} = \mathbf{c}_\beta \Phi_{t\beta} \nabla \ln T + \mathbf{c}_\beta \sum_\gamma \Phi_{d\beta}^\gamma \mathbf{d}_\gamma,$$
$$\sum_l \mathbf{a}_\beta^{2l}(1)\mathbf{P}_\beta^{2l} = \overline{\mathbf{c}_\beta \mathbf{c}_\beta} \Phi_{p\beta}. \quad (92)$$

$\Phi_{t\beta}$, $\Phi_{p\beta}$, and $\Phi_{d\beta}^\gamma$ are the known solutions obtained within the Chapman–Enskog theory [19]. Thus, the distribution function in the first approximation takes the form

$$\Phi_\alpha^{(1)} = \Phi_{t\alpha}(\mathbf{c}_\alpha \cdot \nabla)\ln T + \Phi_{p\alpha}\overline{\mathbf{c}_\alpha \mathbf{c}_\alpha} : \overline{\nabla u} + \left(\mathbf{c}_\alpha, \sum_\beta \Phi_{d\alpha}^\beta \mathbf{d}_\beta\right). \quad (93)$$

Now we pass to the second approximation and retain the spatial derivatives of $\mathbf{a}_\alpha^{mn}$ (1) in the left-hand part. Putting aside equations for the scalar quantities, we write down a system for the vector and tensor moments:

$$\sum_l \mathbf{A}_{\alpha 1n}^{2l} \nabla n_\alpha \mathbf{a}_\alpha^{2l}(1) - \frac{\sqrt{2}}{2}\beta_\alpha^{-1/2}\frac{1}{p_\alpha}\frac{\rho_\alpha}{\rho}\nabla \boldsymbol{\pi}(1)\delta_{n0} = -\sum_{\beta=1}^{N}\sum_l C_{\alpha\beta}^{1nl}\mathbf{a}_\beta^{1l}(2),$$
$$\sum_l \mathbf{A}_{\alpha 2n}^{3l}\overline{\nabla n_\alpha \mathbf{a}_\alpha}^{1l}(2) + \sum_l \mathbf{B}_{\alpha 2n}^{1l}\overline{\nabla n_\alpha \mathbf{a}_\alpha}^{1l}(1) = -\sum_{\beta=1}^{N}\sum_l C_{\alpha\beta}^{2nl}\mathbf{a}_\beta^{2l}(2), \quad (94)$$
$$\sum_l \mathbf{B}_{\alpha 3n}^{2l}\nabla n_\alpha \mathbf{a}_\alpha^{2l}(1) = -\sum_{\beta=1}^{N}\sum_l C_{\alpha\beta}^{3nl}\mathbf{a}_\beta^{3l}(2).$$

For a correct transition to the linearized Barnett approximation of the Chapman–Enskog method, we added an equation for moments of the third-order tensor to the system of equations for the vector and second-order tensor moments. In writing the first of these equations, we took into account that the equation of conservation should be taken in the first approximation (in order to exclude $d\mathbf{u}/dt$), which accounts for the appearance of $\nabla\boldsymbol{\pi}(1)$. Equations of the Barnett approximation are obtained upon multiplying Eqs. (92) by $f_\alpha^{(0)}P_\alpha^{mn}$, summing over n, and using the property of orthogonal polynomials (89):

$$\left(\frac{1}{5}\mathbf{c}_\alpha \mathbf{c}_\alpha^2 \Phi_{p\alpha} + \mathbf{c}_\alpha \frac{1}{p_\alpha}\frac{\rho_\alpha}{\rho}\eta\right)\Delta\mathbf{u} = \sum_\beta L_{\alpha\beta}(\Phi_B^1),$$
$$\overline{\mathbf{c}_\alpha \mathbf{c}_\alpha}\Phi_{t\alpha}\overline{\nabla\nabla}\ln T + \overline{\mathbf{c}_\alpha \mathbf{c}_\alpha}\sum_\gamma \Phi_{d\alpha}^\gamma \overline{\nabla \mathbf{d}_\gamma} = \sum_\beta L_{\alpha\beta}(\Phi_B^2),$$
$$\overline{\mathbf{c}_\alpha \mathbf{c}_\alpha \mathbf{c}_\alpha} : \overline{\nabla\nabla \mathbf{u}}\Phi_{p\alpha} = \sum_\beta L_{\alpha\beta}(\Phi_B^3). \quad (95)$$

In writing this, we used solutions (92), the relation $\boldsymbol{\pi} = -2\eta\nabla\mathbf{u}$, and the notation Φ_B^i for corrections in the Barnett approximation of the ith-order tensor:

$$\Phi_\alpha^{(2)} = \Phi_{B\alpha}^1 + \Phi_{B\alpha}^2 + \Phi_{B\alpha}^3.$$

The equations analogous to (95) obtained by direct application of the Chapman–Enskog method were considered in [8]. Thus, we have demonstrated that, proceeding from a system of equations of the method of moments, it is possible to obtain equations of the Chapman–Enskog theory accurate to within the Barnett approximation.

Now, let us determine the entropy production in the Barnett approximation. Using the general expression (46) for σ and replacing the collision integral by the left-hand parts of equations taken from the corresponding approximation, we obtain

$$\sigma = \sigma(1) + \sigma(2), \quad (96)$$

where $\sigma(1) = \lfloor\sigma^{(1)}\rfloor_0$ corresponds to the classical entropy production (54) (the first Chapman–Enskog approximation) and $\sigma(2)$ corresponds to the Barnett entropy production [8]

$$\sigma(2) = -\frac{1}{T} \times \left(\mathbf{J}^v \Delta\mathbf{u} + \mathbf{J}^u \overline{\nabla\nabla\mathbf{u}} + \mathbf{J}^T \frac{1}{T}\overline{\nabla\nabla}T + \sum_\beta J_\beta^D \overline{\nabla \mathbf{d}_\beta}\right). \quad (97)$$

Additional thermodynamic fluxes appearing in this approximation are determined by the following expressions:

$$\mathbf{J}^{\nu} = kT\sum_{\alpha}\left(\left(\Phi_{p_\alpha}\frac{c_\alpha^2}{5}+\frac{1}{p_\alpha}\frac{\rho_\alpha}{\rho}\eta\right)\mathbf{c}_\alpha, \Phi_\alpha\right),$$

$$\mathbf{J}^{u} = kT\sum_{\alpha}(\Phi_{p\alpha}\overline{\mathbf{c}_\alpha\mathbf{c}_\alpha\mathbf{c}_\alpha}, \Phi_\alpha), \quad (98)$$

$$\mathbf{J}^{T} = kT\sum_{\alpha}(\Phi_{t\alpha}\overline{\mathbf{c}_\alpha\mathbf{c}_\alpha}, \Phi_\alpha),$$

$$\mathbf{J}^{D}_{\beta} = kT(\Phi_d^{\beta}\overline{\mathbf{c}_\alpha\mathbf{c}_\beta}, \Phi_\beta),$$

where

$$\Phi_\alpha = \Phi_\alpha^{(1)} + \Phi_\alpha^{(2)}.$$

The forms of the corresponding thermodynamic forces are clear from the structure of expression (97) for $\sigma(2)$.

Representation of the entropy production in the form of Eqs. (96) and (97) corresponds to a system of linear phenomenological equations. This system separates into three subsystems related to the vector fluxes ($\mathbf{w}_\alpha - \mathbf{w}_N$, $\mathbf{h}$, and $\mathbf{J}^{\nu}$), second-order tensor fluxes ($\hat{\boldsymbol{\pi}}$, $\hat{\mathbf{J}}^{T}$, and $\hat{\mathbf{J}}^{D}_{\alpha} - \hat{\mathbf{J}}^{D}_{N}$), and the third-order tensor flux $\hat{\mathbf{J}}^{u}$. Here are expressions for the vectors,

$$\mathbf{q} - \frac{5}{2}\sum_{\alpha} p_\alpha \mathbf{w}_\alpha = \lambda_{00}\nabla\ln T + \sum_{\beta}\lambda_{0\beta} p\mathbf{d}_\beta + \lambda_{0N+1}\Delta\mathbf{u},$$

$$\mathbf{w}_\alpha - \mathbf{w}_N = \lambda_{\alpha 0}\nabla\ln T + \sum_{\beta}\lambda_{\alpha\beta} p\mathbf{d}_\beta + \lambda_{\alpha N+1}\Delta\mathbf{u}, \quad (99)$$

$$\mathbf{J}^{\nu} = \lambda_{N+1,0}\nabla\ln T + \sum_{\beta}\lambda_{\alpha\beta} p\mathbf{d}_\beta + \lambda_{N+1,N+1}\Delta\mathbf{u};$$

for the second-order tensors,

$$\hat{\boldsymbol{\pi}} = \Lambda_{00}\overline{\nabla\mathbf{u}} + \sum_{\beta}\Lambda_{0\beta}\overline{\nabla\mathbf{d}_\beta} + \Lambda_{0N+1}\overline{\nabla\nabla}\ln T,$$

$$\mathbf{J}^{D}_{\alpha} - \mathbf{J}^{D}_{N} = \Lambda_{\alpha 0}\overline{\nabla\mathbf{u}} + \sum_{\beta}\Lambda_{\alpha\beta}\overline{\nabla\mathbf{d}_\beta} + \Lambda_{\alpha N+1}\overline{\nabla\nabla}\ln T, \quad (100)$$

$$\mathbf{J}^{T} = \Lambda_{N+1,0}\overline{\nabla\mathbf{u}} + \sum_{\beta}\Lambda_{\alpha\beta}\overline{\nabla\mathbf{d}_\beta} + \Lambda_{N+1,N+1}\overline{\nabla\nabla}\ln T;$$

and for the third-order tensors,

$$\mathbf{J}^{u} = l_{00}\overline{\nabla\nabla\mathbf{u}}. \quad (101)$$

As can be seen, the equations of thermal conductivity and diffusion contain additional terms related (with an allowance for the condition $\eta\Delta\mathbf{u} = \nabla p$) to the pressure-gradient-induced heat flux and diffusion (these additional terms were considered above). In order to provide for the Onsager symmetry, the system of phenomenological equations has to be supplemented with a new equation for the "nonphysical" flux $\mathbf{J}^{\nu}$. Direct calculation shows that the kinetic coefficients in system (99) obey the symmetry relations.

The system of phenomenological equations (100) for the tensor quantities contains two nonphysical fluxes corresponding to the appearance of two contributions to the stress tensor, related to the temperature and concentration stresses. In this system of equations, the kinetic coefficients can also be readily shown to obey the symmetry relations.

Finally, note that the nonequilibrium thermodynamics predicts the appearance of phenomenological equations for the third-order tensors in the Barnett approximation. However, presently it is difficult to reveal the physical sense of these quantities.

Evidently, increasing the order of approximation in the Chapman–Enskog theory is accompanied both by expansion of the systems of phenomenological equations (99)–(101) due to the appearance of new terms and by an increase in the number of systems due to growing dimension of the tensor corrections entering into the Chapman–Enskog distribution function. The appearance of additional (nonphysical) fluxes significantly expands the system of phenomenological equations and establishes the necessary relations between the cross coefficients, thus ensuring validity of the Onsager relations. By directly using the equations of moments, the analogous relations can be obtained only upon going outside the approximation of $13N$ moments. Equations of the linearized Barnett approximation obtained above formally correspond to an allowance for the infinite number of moments of the corresponding tensor dimension in the expansion of the distribution function.

8. CONCLUSION

Nonequilibrium thermodynamics and kinetic theory are two closely related approaches to description of nonequilibrium processes in various physical media. The well-developed methods available for an approximate solution to the Boltzmann kinetic equation for simple gases and gas mixtures (the Chapman–Enskog method, the Grad method of moments) provide a reliable basis for generalization of nonequilibrium thermodynamics and determination of the limits of applicability of the generalized theory. One of such generalizations is the extended irreversible thermodynamics developed in recent years [1, 2], which is now rather widely applied to the description of nonequilibrium phenomena in various media. However, in constructing this theory, many important problems encountered in the description of multicomponent systems were ignored, the consequences following from using an infinite system of equations of moments were not prop-

erly analyzed, and the possibilities of passing to description on the level of the canonical Chapman–Enskog method were underestimated.

An important, still unresolved question of the classical LNT is the problem of a relation between nonequilibrium thermodynamics in the Onsager formulation (whereby the fluxes are determined as derivatives of state variables with respect to time and the generalized thermodynamic forces, as partial derivatives of the entropy density with respect to these variables) and the variant in which the fluxes are real physical (vector and tensor) fluxes and the thermodynamic forces are represented by gradients of the initial macroscopic parameters. Application of the method of moments in the kinetic theory provides for progress in understanding this relation, since use of a system of equations of the relaxation type for the state variables allows the two variants of nonequilibrium thermodynamics to be considered as partial cases realized in spatially homogeneous and inhomogeneous systems. We hope that approaches proposed in this paper to the construction of GNT of multicomponent systems will additionally stimulate application of the GNT methods to investigation of a wide class of nonequilibrium phenomena.

ACKNOWLEDGMENTS

The authors are grateful to Yu. Kagan and L.A. Maksimov for their interest in this study and helpful remarks.

This study was supported by the Russian Foundation for Basic Research (project no. 00-01-00175) and by the Ministry of Education of the Russian Federation (Natural Sciences Program, project no. E00-4.0-118).

REFERENCES

1. I. Muller and T. Ruggeri, *Extended Thermodynamics* (Springer-Verlag, New York, 1992).
2. D. Jou, J. Casas-Vazquez, and G. Lebon, *Extended Irreversible Thermodynamics* (Springer-Verlag, New York, 1993).
3. B. C. Eu, *Kinetic Theory and Irreversible Thermodynamics* (Wiley, New York, 1992).
4. L. S. Garcia-Colin and F. J. Uribe, J. Non-Equilib. Thermodyn. **16**, 89 (1991).
5. V. I. Roldugin, Dokl. Akad. Nauk SSSR **263**, 606 (1982) [Sov. Phys. Dokl. **27**, 244 (1982)].
6. V. I. Roldughin, J. Non-Equilib. Thermodyn. **9**, 71 (1984).
7. I. Kuscer, Physica A (Amsterdam) **133**, 397 (1985).
8. V. M. Zhdanov and V. I. Roldugin, Zh. Éksp. Teor. Fiz. **109**, 1267 (1996) [JETP **82**, 683 (1996)].
9. V. M. Zhdanov and V. I. Roldugin, Zh. Éksp. Teor. Fiz. **113**, 2081 (1998) [JETP **86**, 1141 (1998)].
10. V. M. Zhdanov and V. I. Roldugin, Usp. Fiz. Nauk **168**, 407 (1998) [Phys. Usp. **41**, 349 (1998)].
11. H. Grad, Commun. Pure Appl. Math. **2**, 331 (1949); in *Handbuch der Physik* (Springer-Verlag, Berlin, 1958), Vol. 12, p. 205.
12. V. Zhdanov, Yu. Kagan, and A. Sazykin, Zh. Éksp. Teor. Fiz. **42**, 857 (1962) [Sov. Phys. JETP **15**, 596 (1962)].
13. M. Heckl and I. Muller, Acta Mech. **50**, 71 (1983).
14. R. M. Velasco and L. S. Garcia-Colin, Phys. Rev. A **44**, 4961 (1991).
15. R. M. Velasco and L. S. Garcia-Colin, J. Stat. Phys. **69**, 217 (1992).
16. S. Chapman and T. G. Cowling, *Mathematical Theory of Non-Uniform Gases* (Cambridge Univ. Press, Cambridge, 1952; Inostrannaya Literatura, Moscow, 1960).
17. J. R. Brock, J. Colloid Interface Sci. **20**, 570 (1965).
18. J. A. McLennan, Phys. Rev. A **10**, 1272 (1974).
19. J. H. Ferziger and H. G. Kaper, *Mathematical Theory of Transport in Gases* (North-Holland, Amsterdam, 1972; Mir, Moscow, 1976).
20. F. R. W. McCourt, J. J. M. Beenakker, W. E. Kohler, and I. Kuscer, *Non-Equilibrium Phenomena in Polyatomic Gases* (Clarendon, Oxford, 1990), Vols. 1, 2.
21. L. Waldmann, in *Handbuch der Physik* (Springer-Verlag, Berlin, 1958), Vol. 12, p. 295.
22. S. R. de Groot and P. Mazur, *Nonequilibium Thermodynamics* (North-Holland, Amsterdam, 1962; Mir, Moscow, 1964).
23. L. D. Landau and E. M. Lifshitz, *Course of Theoretical Physics*, Vol. 5: *Statistical Physics* (Nauka, Moscow, 1976; Pergamon, Oxford, 1980), Part 1.
24. R. M. Velasco and L. S. Garcia-Colin, J. Non-Equilib. Thermodyn. **18**, 157 (1993).
25. B. M. Zhdanov, *Transport Phenomena in Multicomponent Plasma* (Énergoizdat, Moscow, 1982).
26. V. I. Roldugin, Izv. Akad. Nauk SSSR, Mekh. Zhidk. Gaza, No. 6, 157 (1983).
27. C. Cattaneo, Atti Semin. Mat. Fis. Univ. Modena **3**, 3 (1948).
28. V. M. Zhdanov and M. Ya. Alievskiĭ, *Transport and Relaxation Phenomena in Molecular Gases* (Nauka, Moscow, 1989).

Translated by P. Pozdeev

Journal of Experimental and Theoretical Physics, Vol. 95, No. 4, 2002, pp. 697–704.
Translated from Zhurnal Éksperimental'noĭ i Teoreticheskoĭ Fiziki, Vol. 122, No. 4, 2002, pp. 805–813.

PLASMA, GASES

Electromagnetically Induced Transparency in High-Temperature Magnetoactive Plasma

A. Yu. Kryachko*, A. G. Litvak, and M. D. Tokman

Institute of Applied Physics, Russian Academy of Sciences, ul. Ul'yanova 46, Nizhni Novgorod, 603950 Russia
**e-mail: kryachko@appl.sci-nnov.ru*

Received April 23, 2002

Abstract—We investigate a classical analog of electromagnetically induced transparency (EIT) currently popular in quantum electronics. We consider EIT for electron cyclotron waves in finite-temperature plasma. We derive an expression for the effective refractive index of an electromagnetic wave and study the dispersion and absorption of this wave under EIT conditions. Allowance for thermal motion is shown to radically change the behavior of the dispersion curves for the signal wave in the EIT region compared to the case of cold plasma.

1. INTRODUCTION

A popular parametric effect in modern physics of nonlinear waves is the electromagnetically induced transparency[1] in ensembles of three-level atoms attributable to the interference of quantum electron states. This effect manifests itself as the formation of a transparency window within a resonance absorption line of a three-level system in the presence of a pump wave, which is accompanied by a record high deceleration of the signal light wave (see, e.g., [1] for a review). At present, using EIT in problems of spectroscopy and the formation and correction of the spatiotemporal structure of laser radiation is generally recognized to hold considerable promise [1]; in the EIT regime, the power thresholds of most known nonlinear optical effects were significantly lowered [1, 2]. In particular, EIT underlies the recent widely known "light stopping" experiments [3, 4].

From the viewpoint of the general theory for nonlinear wave processes, investigating classical analogs of EIT is obviously of interest. This interest also stems from the hopes to extend new ideas of quantum electronics to classical microwave electronics and plasma physics. Thus, for example, the authors of [5–7] analyzed the formation of a transparency region for a wave with a frequency lower than the plasma frequency in isotropic cold plasma through its parametric interaction with an intense pump wave. However, they considered a completely conservative system in which the processes significantly differ from the formation of a transparency window within a resonance absorption line typical of EIT [1]. Litvak and Tokman [8] investigated a closer (direct) classical analog of this effect in a wave system: EIT for an electromagnetic wave at the electron cyclotron resonance frequency in cold collisional plasma. They considered a situation where the beats between the signal and pump waves effectively excite a plasma electrostatic mode.

In addition, Litvak and Tokman [8] pointed out a classical analog of EIT for a lumped-parameter system: dynamic damping (see, e.g., [9]). Two coupled oscillators with a harmonically varying coupling coefficient form such a system. As was shown, an external action (an analog of the signal wave) on one of the oscillators for such a system produces no resonant excitation of its oscillations due to the resonant excitation of oscillations in the second oscillator (because of parametric coupling), i.e., through dynamic damping of this resonant excitation. In its standard case [9], this effect is produced by ordinary linear coupling between the two oscillators, which causes the damping system to be excited at the frequency of the driving generalized force. In the case under consideration, however, the parametric coupling between the oscillators leads to the excitation of the damping system at the combination frequency; this is the principal feature of wave interactions in the EIT regime.

Here, we investigate the effect of thermal plasma particle motion on EIT for electron cyclotron waves. The spatial dispersion attributable to thermal motion turns out to radically change the pattern of EIT, giving rise to a number of important features that are absent both in the original quantum case and in cold plasma.

The paper is structured as follows. The hydrodynamic theory of EIT is outlined in Section 2. In Sections 3 and 4, this effect is studied in terms of the kinetic theory: an expression for the effective refractive index of the signal wave is derived in Section 3, and properties of the dispersion and absorption of the signal wave for warm plasma are discussed in Section 4.

[1] Bearing in mind the formation of a transparency window precisely in the regime of a parametric wave interaction, we use the universally accepted abbreviation EIT (electromagnetically induced transparency) for this effect.

2. THE HYDRODYNAMIC THEORY

Although the hydrodynamic model disregards one of the most important plasma effects, collisionless damping, the qualitative behavior of the dispersion curves for the signal wave in warm plasma can be understood in terms of this model. Our statement of the problem is the same as that in [8]. Let two circularly polarized waves with the electric field rotating in the sense of the electron cyclotron rotation (extraordinary waves) propagate in magnetoactive plasma along a constant external magnetic field $\mathbf{H} = H\mathbf{z}_0$:

$$\begin{aligned}\mathbf{E}_\perp(z,t) &= \mathrm{Re}\{\mathbf{e}_+[E_1\exp(-i\omega_1 t + ik_1 z)\\ &\quad + E_2\exp(-i\omega_2 t + ik_2 z)]\},\end{aligned} \tag{1}$$

where $\mathbf{e}_+ = 2^{-1/2}(\mathbf{x}_0 + i\mathbf{y}_0)$ is the wave polarization vector and $\mathbf{x}_0$, $\mathbf{y}_0$, $\mathbf{z}_0$ are the unit vectors of the Cartesian axes.

The oscillations of the transverse and longitudinal (relative to the constant magnetic field) electron velocities are described by the Euler equations including the Lorentz force from the wave fields:

$$\begin{aligned}&\frac{\partial \mathbf{V}_\perp}{\partial t} + \omega_H[\mathbf{V}_\perp \times \mathbf{z}_0] + \gamma\mathbf{V}_\perp + V_\parallel\frac{\partial \mathbf{V}_\perp}{\partial z}\\ &= -\frac{e}{m}\left(\mathbf{E}_\perp(z,t) + V_\parallel\frac{\partial}{\partial z}\int_{-\infty}^{t}\mathbf{E}_\perp(z,t)dt\right),\\ &\frac{\partial V_\parallel}{\partial t} + \gamma V_\parallel + V_\parallel\frac{\partial V_\parallel}{\partial z} = -\frac{1}{mN_e}\frac{\partial p}{\partial z}\\ &\quad -\frac{e}{m}E_p(z,t) + \frac{e}{m}\mathbf{V}_\perp\frac{\partial}{\partial z}\int_{-\infty}^{t}\mathbf{E}_\perp(z,t)dt.\end{aligned} \tag{2}$$

Here, $\omega_H = eH/mc$ is the electron gyrofrequency, e and m are the electron charge and mass ($e > 0$), γ is the effective collision frequency, E_p is the plasma-wave electric field, p is the gas-kinetic pressure, and N_e is the electron density. System (2) must be supplemented with the continuity equation

$$\frac{\partial N_e}{\partial t} + \frac{\partial}{\partial z}(N_e V_\parallel) = 0 \tag{3}$$

and with the equation that describes the excitation of an electric field in the plasma wave:

$$\frac{\partial E_p}{\partial t} = -4\pi j_z = 4\pi e N_e V_\parallel. \tag{4}$$

Here, j_z is the longitudinal electron current component. Assuming the constancy of the ion density N_i and quasi-neutrality,

$$\begin{aligned}n &= N_e - N_i \ll N_e, N_i,\\ N_i &= N_0 = \text{const},\end{aligned} \tag{5}$$

we can derive a system of equations that describes the excitation of longitudinal collective oscillations by the ponderomotive force from the high-frequency fields (for definiteness, consider an adiabatic process) from Eqs. (2)–(4):

$$\begin{aligned}&\frac{\partial^2 n}{\partial t^2} + \gamma\frac{\partial n}{\partial t} + \omega_p^2 n - 3V_T^2\frac{\partial^2 n}{\partial z^2}\\ &= -\frac{eN_0}{m}\frac{\partial}{\partial z}\mathbf{V}_\perp\frac{\partial}{\partial z}\int_{-\infty}^{t}\mathbf{E}_\perp(z,t)dt,\\ &V_\parallel = -N_0^{-1}\int_{-\infty}^{z}\frac{\partial n}{\partial t}dz.\end{aligned} \tag{6}$$

Here, $\omega_p = (4\pi e N_0/m)^{1/2}$ is the electron plasma frequency, and $V_T = (T/m)^{1/2}$ and T are the thermal electron velocity and temperature, respectively.

We assume that the following synchronism conditions are satisfied:

$$|\omega_1 - \omega_H|, \gamma \ll \omega_1, \omega_H, \tag{7a}$$

$$|\omega_L - \omega_p|, \gamma \ll \omega_L, \omega_p, \tag{7b}$$

$$k_{1,2,L}V_T \ll \omega_{1,2,L}. \tag{7c}$$

Here, $\omega_L = \omega_1 - \omega_2$ and $k_L = k_1 - k_2$ are the frequency and wavenumber of the beats between the signal and pump waves, respectively. Inequality (7a) implies the satisfaction of the electron cyclotron resonance conditions for the signal wave and, hence, the cyclotron absorption of this wave in the absence of a pump wave. Inequality (7b) is the condition for the excitation of plasma waves at the combination frequency ω_L. Inequality (7c) is the condition for weak spatial dispersion.

Given the synchronism conditions (7a) and (7b), only the terms with the "resonance" frequencies can be retained in Eqs. (2)–(4); i.e., the method of "shortened" equations [9, 10] can be used. As a result, we have the system of equations

$$\begin{aligned}-iZ_1\hat{V}_1 &= -\frac{e}{m}E_1 + \frac{e}{2m}\frac{k_2}{\omega_2}E_2\hat{V}_\parallel - \frac{i}{2}k_2\hat{V}_2\hat{V}_\parallel,\\ -iZ_2\hat{V}_2 &= -\frac{e}{m}E_2,\\ -iZ_N\hat{V}_N &= \frac{e}{2m}\frac{k_2}{\omega_2}E_2^*\hat{V}_\parallel + \frac{i}{2}k_2\hat{V}_2^*\hat{V}_\parallel,\\ -i\frac{Z_p}{\omega_L}\hat{V}_\parallel &= -\frac{e}{2m}\frac{k_1}{\omega_1}E_1\hat{V}_2^* - \frac{e}{2m}\frac{k_2}{\omega_2}E_2^*\hat{V}_1\\ &\quad -\frac{e}{2m}\frac{k_2}{\omega_2}E_2\hat{V}_N, \quad \hat{n} = \frac{k_L}{\omega_L}N_0\hat{V}_\parallel.\end{aligned} \tag{8}$$

Here,

$$Z_{1,2} = \omega_{1,2} - \omega_H + i\gamma,\ Z_N = \omega_1 - 2\omega_2 + \omega_H + i\gamma,$$
$$\text{and } Z_p = \omega_L^2 - \omega_p^2 - 3k_L^2 V_T^2 + i\gamma\omega_L,$$

$Z_p = 0$ being the standard "hydrodynamic" dispersion relation for plasma waves. In (8), we introduced the complex amplitudes of the corresponding quantities:

$$\mathbf{V}_\perp = \mathrm{Re}\Bigg[\mathbf{e}_+ \sum_{j=1,2} \hat{V}_j \exp(-i\omega_j t + ik_j z) + \mathbf{e}_- \hat{V}_N \exp[-i(\omega_1 - 2\omega_2)t + i(k_1 - 2k_2)z]\Bigg], \quad (9)$$

$$V_\parallel = \mathrm{Re}[\hat{V}_\parallel \exp(-i\omega_L t + ik_L z)],$$
$$n = \mathrm{Re}[\hat{n}\exp(-i\omega_L t + ik_L z)],$$

where $\mathbf{e}_- = 2^{-1/2}(\mathbf{x}_0 - i\mathbf{y}_0)$ is the polarization vector for the circularly polarized wave with the electric field rotating oppositely to the sense of the electron cyclotron rotation (ordinary wave).

We can derive an expression for the complex oscillation amplitudes of the velocity $\hat{V}_{1,2}$ and density $\hat{n}$ from system (8) and an expression for the effective refractive index of the signal wave from the latter (see, e.g., [11]):

$$N_{\mathrm{hyd}}^2 = 1 - \frac{4\pi\hat{j}_1}{i\omega_1 E_1}. \quad (10)$$

Here, $\hat{j}_1$ is the complex amplitude of the electric current $\mathbf{j}_1$ at frequency ω_1:

$$\mathbf{j}_1 = \mathrm{Re}[\mathbf{e}_+ \hat{j}_1 \exp(-i\omega_1 t + ik_1 z)],$$
$$\hat{j}_1 = -e\left(N_0 \hat{V}_1 + \frac{\hat{n}\hat{V}_2}{2}\right). \quad (11)$$

The final expression for N_{hyd}^2 is

$$N_{\mathrm{hyd}}^2 = 1 - \frac{\omega_p^2}{\omega_1} \times \frac{Z_p + \xi_{EC}\omega_L\omega_2(2k_1/k_2 - 1 - \omega_1/2\omega_2)}{(\omega_1 - \omega_H + i\gamma)Z_p - \omega_L^2\omega_1\xi_{EC}}. \quad (12)$$

Here

$$\xi_{EC} = \left|\frac{eE_2}{2m(\omega_2 - \omega_H + i\gamma)}\right|^2 \left(\frac{k_2}{\omega_2}\right)^2 = \frac{|V_2|^2}{(2V_{\mathrm{ph}})^2}$$

is a dimensionless nonlinearity parameter representing the ratio of the squares of the oscillatory and phase velocities for the pump field $E_2\exp(-i\omega_2 t + ik_2 z)$. Expression (12) differs from that derived in [8] by the corrections $3k_L^2 V_T^2$, because we took into account thermal motion.

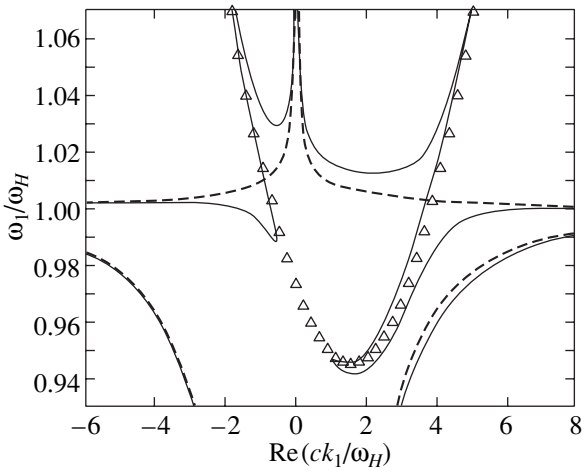


Fig. 1. The dispersion law for an electron cyclotron wave under EIT conditions, the hydrodynamic theory. $\xi_{EC} = 1 \times 10^{-3}$ (solid line) and 0 (dashed line); $Z_p = 0(\Delta)$; $\omega_p/\omega_H = 0.75$, $\omega_2/\omega_H = 0.2$, $\gamma/\omega_H = 2.5 \times 10^{-3}$, $V_T/c = 7.5 \times 10^{-3}$.

Let us now discuss the dispersion law of the signal wave under EIT conditions for warm plasma. To this end, we represent Eq. (12) as

$$Z_p(N_{\mathrm{hyd}}^2 - N_{0,\mathrm{hyd}}^2) = -\xi_{EC}\frac{\omega_2\omega_L^2}{Z_1} \times \left[1 - N_{\mathrm{hyd}}^2 + \frac{\omega_p^2}{\omega_1\omega_L}\left(\frac{2k_1}{k_2} - 1 - \frac{\omega_1}{2\omega_2}\right)\right]. \quad (13)$$

Here,

$$N_{0,\mathrm{hyd}}^2 = 1 - \frac{\omega_p^2}{\omega_1(\omega_1 - \omega_H + i\gamma)}$$

is the "linear" refractive index of the signal wave in cold magnetoactive plasma (see, e.g., [11]). The quantity ξ_{EC} is a small parameter of the problem: for reasonable pump intensities (of the order of 100 kW cm^{-2}), ξ_{EC} is 10^{-4}–10^{-6}. Therefore, the signal-wave dispersion curves defined by Eq. (13) pass either near the linear signal-wave dispersion curve defined by the relation $N_{\mathrm{hyd}}^2 = N_{0,\mathrm{hyd}}^2$ or near the plasma-wave dispersion curve $Z_p = 0$ (shifted by the pump frequency ω_2 and wavenumber k_2). This behavior of the dispersion curves is violated only near the points of intersection of the $N_{\mathrm{hyd}}^2 = N_{0,\mathrm{hyd}}^2$ and $Z_p = 0$ curves, where the signal-wave dispersion curves pass from one curve to the other. All of these results are confirmed by numerical calculations for relation (13) (see Fig. 1). Figure 2 shows the behavior of the signal-wave absorption line in the EIT frequency range. On the one hand, it follows from these

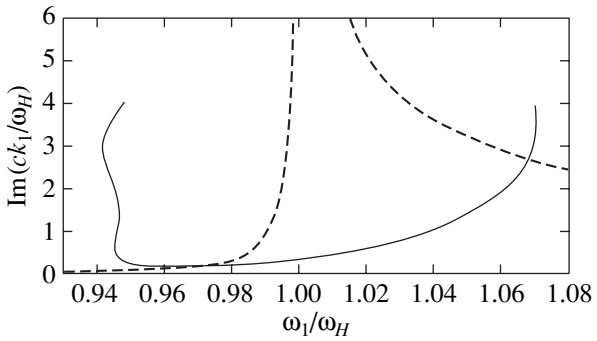


Fig. 2. The formation of a transparency window within the cyclotron absorption line [Re(k_1) > 0], the hydrodynamic theory. $\xi_{EC} = 1 \times 10^{-3}$ (solid line) and 0 (dashed line). For $\xi_{EC} \neq 0$, only the segment of the curve that corresponds to low absorption is shown. All parameters are the same as those in Fig. 1.

figures that EIT also exists in warm plasma (there is a high group deceleration of the signal wave and the appearance of a transparency region within the cyclotron absorption line). On the other hand, since the plasma wave frequency depends on wave number, the behavior of the dispersion curves in the EIT window significantly differs from that for cold plasma (see [8]). The signal-wave propagation is discussed in detail in Section 4 in terms of the kinetic theory.

3. THE KINETIC THEORY

As was noted above, collisionless wave damping is disregarded in the hydrodynamic theory. Therefore, EIT in high-temperature plasma can be properly studied only in terms of the kinetic theory. We consider the propagation of a bichromatic field (1) in plasma. Let us write the kinetic equation for the electron distribution function f,

$$\frac{\partial f}{\partial t} + \mathbf{V}\frac{\partial f}{\partial \mathbf{r}} - \omega_H[\mathbf{V}\times\mathbf{z}_0]\frac{\partial f}{\partial \mathbf{V}} - \frac{e}{m}\left(\mathbf{E} + \frac{1}{c}[\mathbf{V}\times\mathbf{B}]\right)\frac{\partial f}{\partial \mathbf{V}} = -\gamma(f - f_0) \tag{14}$$

(f_0 is the equilibrium distribution function), and the equation that describes the excitation of plasma waves,

$$\frac{\partial E_p}{\partial t} = -4\pi j_z = 4\pi e\int V_\parallel f d^3V. \tag{15}$$

As above, we assume that inequalities (7a)–(7c) hold. Let us consider the simplified collision integral in Eq. (14). This form of the collision integral allows us to properly pass to the Landau formula with a determining contribution of collisionless damping. In this case, in view of (7a), the collisionless damping decrement for the signal wave is large, while, in view of (7c), the collisionless damping decrement for the pump and plasma waves is exponentially small; for the parameters of interest, it is much lower than the transport frequency for Coulomb collisions of thermal particles. Thus, it would be reasonable to take γ in (14) to be equal to the transport frequency of Coulomb collisions at a given temperature (see, e.g., [12]):

$$\gamma = \frac{4\pi e^4 N_0 L_e}{m^2 V_T^3}. \tag{16}$$

Here, L_e is the Coulomb logarithm. Let us introduce the complex amplitudes of the harmonics of the distribution function f and electric field E_p:

$$f = f_0(V_\perp, V_\parallel) + \mathrm{Re}\left\{\sum_{j=1,2} f_j(V_\perp, V_\parallel, \varphi)\exp(-i\omega_j t + ik_j z) + f_\parallel(V_\perp, V_\parallel, \varphi)\exp(-i\omega_L t + ik_L z) + f_N(V_\perp, V_\parallel, \varphi) \times \exp[-i(\omega_1 - 2\omega_2)t + i(k_1 - 2k_2)z]\right\}, \tag{17}$$

$$E_p = \mathrm{Re}\{E_\parallel \exp(-i\omega_L t + ik_L z)\}.$$

Here, $V_\parallel$ and $V_\perp$ are the longitudinal and transverse electron velocity components relative to the external magnetic field, and φ is the azimuthal angle in the xy plane. Below, we give the φ dependence of the $f_{1,2,\parallel,N}$ components; since these components are periodic in φ, they can be represented as the Fourier series

$$f_{1,2,\parallel,N} = \sum_{n=-\infty}^{\infty} f^{(n)}_{1,2,\parallel,N}\exp(in\varphi). \tag{18}$$

Clearly, only the harmonic $f_1^{(1)} = \hat{f}_1$ for f_1 gives a nonzero contribution to the nonlinear current at frequency ω_1. Next, it follows from the specific form of the shortened equations following from (14) that only the harmonics $f_2^{(1)} = \hat{f}_2$, $f_\parallel^{(0)} = \hat{f}_\parallel$, and $f_N^{(-1)} = \hat{f}_N$ give a nonzero contribution to $f_1^{(1)}$. Thus, the expression for f in (17) can be represented as

$$f = f_0(V_\perp, V_\parallel) + \mathrm{Re}\left\{\sum_{j=1,2} \hat{f}_j(V_\perp, V_\parallel)\exp(-i\omega_j t + ik_j z + i\varphi) + \hat{f}_\parallel(V_\perp, V_\parallel)\exp(-i\omega_L t + ik_L z) + \hat{f}_N(V_\perp, V_\parallel) \times \exp[-i(\omega_1 - 2\omega_2)t + i(k_1 - 2k_2)z - i\varphi]\right\}. \tag{19}$$

The system of shortened equations follows from (14)–(19). To write it in a more compact form, let us introduce new variables:

$$\hat{p}_{1,2,N} = \int \hat{f}_{1,2,N} V_\perp d^2 V_\perp, \quad \hat{n} = \int \hat{f}_\parallel d^2 V_\perp. \qquad (20)$$

Our system then takes the form

$$\begin{aligned}
-iZ_1\hat{p}_1 &= -\frac{e}{m}E_1\left(1-\frac{k_1V_\parallel}{\omega_1}\right)N_0 \\
&+ \frac{e}{m}\frac{k_1}{\omega_1}E_1\frac{d}{dV_\parallel}\int V_\perp^2 f_0 d^2V_\perp \\
&- \frac{e}{2m}E_2\left(1-\frac{k_2V_\parallel}{\omega_2}\right)\hat{n} + \frac{e}{2m}\frac{k_2}{\omega_2}E_2 \\
&\times \frac{d}{dV_\parallel}\int V_\perp^2 f_\parallel d^2V_\perp + \frac{e}{2m}E_\parallel\frac{d\hat{p}_2}{dV_\parallel}, \\
-iZ_2\hat{p}_2 &= -\frac{e}{m}E_2\left(1-\frac{k_2V_\parallel}{\omega_2}\right)N_0 \\
&+ \frac{e}{m}\frac{k_2}{\omega_2}E_2\frac{d}{dV_\parallel}\int V_\perp^2 f_0 d^2V_\perp, \qquad (21) \\
-iZ_N\hat{p}_N &= -\frac{e}{2m}E_2^*\left(1-\frac{k_2V_\parallel}{\omega_2}\right)\hat{n} \\
&+ \frac{e}{2m}\frac{k_2}{\omega_2}E_2^*\frac{d}{dV_\parallel}\int V_\perp^2 f_\parallel d^2V_\perp + \frac{e}{2m}E_\parallel\frac{d\hat{p}_2^*}{dV_\parallel}, \\
-iZ_\parallel\hat{n} &= \frac{e}{m}E_\parallel\frac{dN_0}{dV_\parallel} + \frac{e}{2m}\frac{k_1}{\omega_1}E_1\frac{d\hat{p}_2^*}{dV_\parallel} \\
&+ \frac{e}{2m}\frac{k_2}{\omega_2}E_2^*\frac{d\hat{p}_1}{dV_\parallel} + \frac{e}{2m}\frac{k_2}{\omega_2}E_2\frac{d\hat{p}_N}{dV_\parallel}, \\
-i\omega_L E_\parallel &= 4\pi e\int_{-\infty}^{\infty}\hat{n}V_\parallel dV_\parallel.
\end{aligned}$$

Here, we denoted

$$\begin{aligned}
Z_{1,2} &= \omega_{1,2} - \omega_H - k_{1,2}V_\parallel + i\gamma, \\
Z_N &= \omega_1 - 2\omega_2 + \omega_H - (k_1 - 2k_2)V_\parallel + i\gamma, \qquad (22) \\
Z_\parallel &= \omega_L - k_L V_\parallel + i\gamma.
\end{aligned}$$

To solve system (21), we use the following considerations. The amplitude of the electric field $E_\parallel$ in a plasma wave is proportional to $F_\parallel/D(\omega_L, k_L)$, where $F_\parallel$ is the Lorentz force from the wave fields and $D(\omega_L, k_L)$ is such that $D(\omega_L, k_L) = 0$ is the dispersion relation for plasma waves [e.g., $D(\omega_L, k_L)$ corresponds to the previously introduced quantity Z_p in the hydrodynamic theory]. Since we assume the plasma waves to be effectively excited, $D(\omega_L, k_L)$ in our case is a small parameter. Therefore, apart from the linear terms, it would be appropriate to retain only the terms proportional to $E_\parallel$ in the equation for $\hat{p}_1$ in (21). This assumption allows us to find a solution to system (21) and to derive an expression for the effective refractive index of the signal wave from it in terms of the kinetic theory [cf. (10) and (11)]:

$$\begin{aligned}
N_{\text{kin}}^2 &= 1 - \frac{4\pi\hat{j}_1}{i\omega_1 E_1}, \\
\hat{j}_1 &= -e\int V_\perp e^{-i\varphi} f_1 d^3V = -e\int_{-\infty}^{\infty}\hat{p}_1 dV_\parallel.
\end{aligned} \qquad (23)$$

The final expression for N_{kin}^2 is

$$N_{\text{kin}}^2 = N_{0,\text{kin}}^2 - \xi_{EC}\frac{A(\omega_1, k_1)}{D(\omega_L, k_L) - \xi_{EC}B(\omega_1, k_1)}. \qquad (24)$$

Here, $N_{0,\text{kin}}^2$ is the expression for the "linear" refractive index of the extraordinary wave that follows from the kinetic theory (see, e.g., [11]):

$$\begin{aligned}
N_{0,\text{kin}}^2 &= 1 - \frac{\omega_p^2}{\omega_1^2} \\
&\times \int_{-\infty}^{\infty}\frac{(\omega_1 - k_1V_\parallel)F_0 - k_1\partial(\langle V_\perp^2\rangle F_0)/\partial V_\parallel}{\omega_1 - \omega_H - k_1V_\parallel + i\gamma}dV_\parallel;
\end{aligned} \qquad (25)$$

$D(\omega_L, k_L)$ is the "longitudinal" permittivity of magnetoactive plasma [11]:

$$D(\omega_L, k_L) = 1 + \frac{\omega_p^2}{\omega_L}\int_{-\infty}^{\infty}\frac{(\partial F_0/\partial V_\parallel)V_\parallel dV_\parallel}{\omega_L - k_L V_\parallel + i\gamma}. \qquad (26)$$

The equation $D(\omega_L, k_L) = 0$ describes the dispersion law of plasma waves, and A and B are functions that do not depend on ξ_{EC}:

$$\begin{aligned}
A &= \frac{\omega_p^2}{\omega_1^2}\left[\int_{-\infty}^{\infty}\frac{\omega_p^2}{\omega_L Z_\parallel}\left(\frac{k_1}{k_2}\frac{\omega_2}{Z_2^*} - \frac{\omega_1}{Z_1}\right)F_0 dV_\parallel\right] \\
&\times \left[\int_{-\infty}^{\infty}\frac{|Z_2^{(0)}|^2}{Z_1 Z_\parallel}\left(\frac{\omega_2}{Z_2} - 1 + \frac{k_L}{k_2}\frac{\omega_2}{Z_2}\frac{Z_1}{Z_\parallel}\right)F_0 dV_\parallel\right], \\
B &= \int_{-\infty}^{\infty}\frac{\omega_p^2}{\omega_L Z_\parallel}\frac{|Z_2^{(0)}|^2}{Z_1 Z_2} \\
&\times \left[\frac{\omega_2}{Z_2} - 1 - \left(\frac{3k_L}{k_2} - 1 - \frac{Z_\parallel}{\omega_2}\right)\frac{Z_1\omega_2}{Z_\parallel^2}\right]F_0 dV_\parallel,
\end{aligned} \qquad (27)$$

where

$$F_0(V_\parallel) = N_0^{-1}\int f_0 d^2V_\perp$$

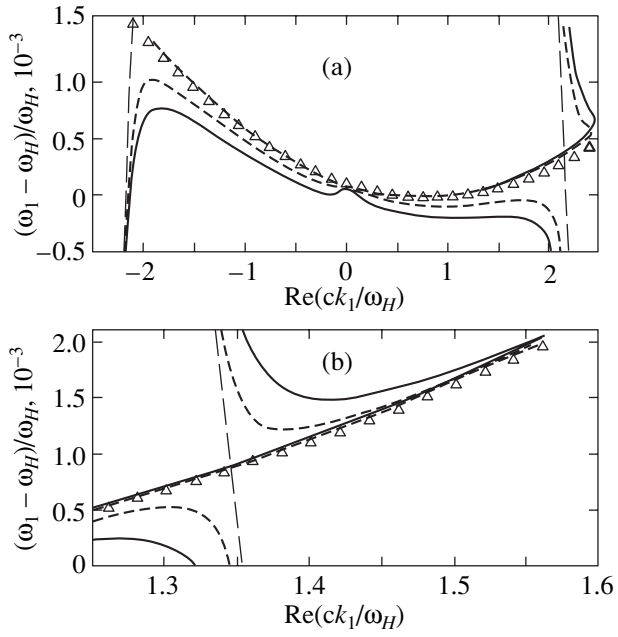


Fig. 3. The dispersion law for an electron cyclotron wave under EIT conditions, the kinetic theory. I_2 = 400 (a) and 300 (b) kW cm^{-2} (solid line), 200 (a), 100 (b) kW cm^{-2} (thick dashed line), and 0 (thin dashed line); $\mathrm{Re}D(\omega_L, k_L) = 0$ (Δ). Panel (b) shows in more detail a segment of the dispersion curve near the point of its intersection with the linear dispersion curve. T = 30 (a), 610 (b) eV; $N_0 = 10^{13}$ cm^{-3}; H = 35 kG ($\omega_H/2\pi$ = 94 GHz); $\gamma/\omega_H = 4.0 \times 10^{-6}$ (a), 4.0×10^{-8} (b); ω_2/ω_H = 0.7 (a), 0.72 (b).

is the unperturbed longitudinal velocity distribution function normalized to unity,

$$\langle V_\perp^2 \rangle = \frac{\int f_0 V_\perp^2 d^2 V_\perp}{\int f_0 d^2 V_\perp}, \quad Z_2^{(0)} = Z_2(V_T = 0).$$

In deriving (24), we disregarded small terms of the order of kV_T/ω and $(kV_T/\omega)^2$ because of condition (7c). For $\xi_{EC} = 0$, the linear expression $N^2_{0,\,\mathrm{kin}}$ for the refractive index clearly follows from (24), while, in the limiting case of cold plasma ($F_0(V_\parallel) = \delta(V_\parallel)$), expression (24) is identical to that for cold plasma from the hydrodynamic theory [expression (12), in which we must set $V_T = 0$].

4. SIGNAL-WAVE PROPAGATION

As was noted above, ξ_{EC} is a small parameter. Consequently, the contribution of the second term on the right-hand side of Eq. (24), which corresponds to the manifestation of EIT, is significant only if $D(\omega_L, k_L)$ is small enough. Therefore, the criterion for EIT to take place can be written as

$$\xi_{EC} \gg \max\{\mathrm{Re}D(\omega_L, k_L), \mathrm{Im}D(\omega_L, k_L)\}. \quad (28)$$

This implies that, just as in the hydrodynamic approach, the behavior of the dispersion curves for the signal wave in the EIT region is mainly determined by the dispersion law of plasma waves $\mathrm{Re}D(\omega_L, k_L) = 0$. This is in contrast to the case of cold plasma [8], where their behavior is unambiguously specified by the parameter ξ_{EC}. In general, this result could be predicted based on the theoretical oscillatory treatment of EIT formulated in [8]. In order for EIT to take place, of fundamental importance is the parametric excitation of "damping" oscillations, which in our case satisfy the dispersion relation $\mathrm{Re}D(\omega_L, k_L) = 0$.

Below, we present the results of our numerical analysis of expression (24); the unperturbed distribution function is assumed to be Maxwellian and isotropic:

$$f_0 = (2\pi)^{-3/2} V_T^{-3} \exp\left(-\frac{V_\perp^2 + V_\parallel^2}{2V_T^2}\right). \quad (29)$$

Figure 3 shows the signal-wave dispersion curves[2] in the EIT region for various pump intensities I_2. We see from Fig. 3 that the behavior of the dispersion curves in the EIT region is actually determined by the dispersion law of plasma waves. In particular, the dispersion curves contain segments that correspond to a high group deceleration of the signal wave down to a zero group velocity (which is not possible in a three-level quantum system). In this case, the frequency ω_1 and the wavenumber $\mathrm{Re}k_1$ that correspond to a zero group velocity are roughly determined by their values at the minimum of the dispersion curves for plasma waves:

$$\begin{aligned} \omega_1 &= \omega_2 + \omega_p + O(\xi_{EC}), \\ \mathrm{Re}k_1 &= k_2 + O(\xi_{EC}). \end{aligned} \quad (30)$$

Figure 4 shows the absorption-line profile[3] for the signal wave under EIT conditions. We see that, as in cold plasma, there is a frequency range where the resonant wave absorption is suppressed, with the absorption minimum roughly corresponding to the point on the dispersion curve with a zero group velocity. Note that the "linear" curves in Figs. 3 and 4 are defined by Eq. (25); i.e., their behavior significantly depends on the electron temperature (for $N^2_{0,\,\mathrm{kin}}$, there is a standard expression [11, 13] in terms of the Kramp function).

Let us estimate the characteristic width of the signal-wave transparency band $\Delta\omega$ under EIT conditions.

[2] The regions on the dispersion curves where the group velocity formally becomes infinite should not cause any misunderstanding, because the signal-wave absorption is large for these regions ($\mathrm{Re}k_1 \sim \mathrm{Im}k_1$) and there is no proper determination of the group velocity [13].

[3] For the curve in Fig. 4, the term "absorption line" is not completely correct, because the behavior of this curve depends not only on ω_1 but also on $\mathrm{Re}k_1$, which, in turn, depends on frequency. In particular, this is the reason why regions of ambiguity in ω_1 exist for this curve.

For cold plasma, as follows from (12), the wave absorption significantly decreases in the frequency band

$$\left|(\omega_1 - \omega_H)((\omega_1 - \omega_2)^2 - \omega_p^2)\right| \approx \omega_p^2 \omega_H \xi_{EC};$$

i.e., $\Delta\omega$ can be estimated as

$$\Delta\omega \sim (\omega_p \omega_H \xi_{EC})^{1/2}. \quad (31)$$

In the hydrodynamic theory for warm plasma, it follows from (12) that the transparency band width is given by the relation

$$\left|(\omega_1 - \omega_H)\mathrm{Re}Z_p\right| \approx \omega_p^2 \omega_H \xi_{EC},$$

which leads to the expression

$$\Delta\omega \sim \max\left\{(\omega_p \omega_H \xi_{EC})^{1/2}, \frac{k_L^2 V_T^2}{\omega_p}\right\}. \quad (32)$$

Since expressions (24)–(27) for $N_{0,\,\mathrm{kin}}^2$ are complex, a simple analytic expression for $\Delta\omega$ cannot be derived in terms of the kinetic theory. However, numerical calculations indicate that relation (32) is also valid in the kinetic theory. In particular, it follows from this relation that the transparency band width in warm plasma can be significantly larger than that in cold plasma if the plasma temperature is high enough:

$$\frac{T}{mc^2} > \omega_p^2 (ck_L)^{-2}\left(\frac{\omega_H \xi_{EC}}{\omega_p}\right)^{1/2}. \quad (33)$$

In this case, the transparency band width ceases to depend on ξ_{EC} altogether (see Fig. 4). Thus, for example, for a pump intensity of 100 kW cm^{-2} and for the remaining parameters corresponding to Fig. 4, the plasma temperature must exceed 150 eV.

The table presents the dependence of basic parameters for the dispersion law and the signal-wave absorption line in the EIT region on pump intensity (the remaining parameters correspond to Fig. 4): the characteristic group velocity[4] V_{gr}, the transparency band width $\Delta\omega$, and the absorption length $L = (2\mathrm{Im}\,k_1)^{-1}$ at the minimum of the absorption line. In this case, relations (28) and (33) are assumed to hold. As we see from the table, for a sufficiently wide intensity range, only L changes significantly, while V_{gr} and $\Delta\omega$, which are determined by the structure of the plasma-wave dispersion curve, are constant.

The above estimates of basic parameters for the EIT window allows us to formulate the most pessimistic requirements for the conditions for this effect to take place in plasma. The constraints on the admissible non-

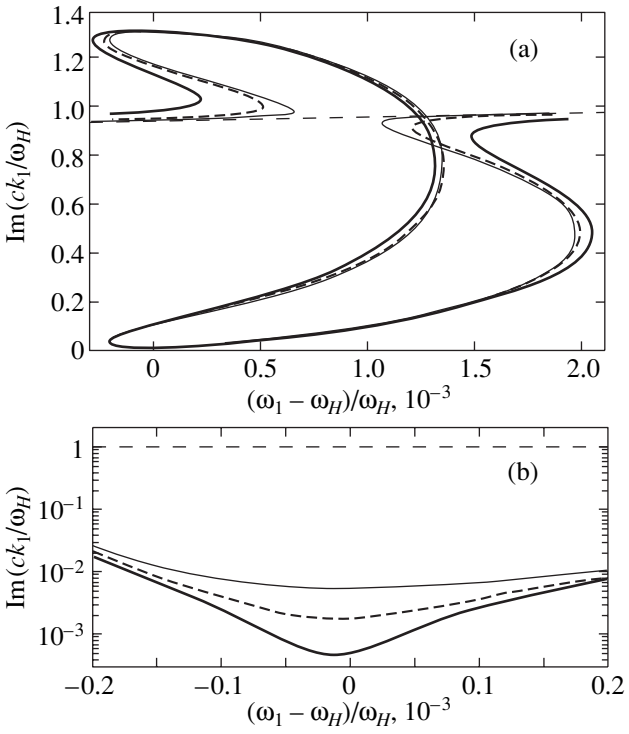


Fig. 4. The formation of a transparency window within the cyclotron absorption line [Re(k_1) > 0], the kinetic theory. I_2 = 300 kW cm^{-2} (thick solid line), 100 kW cm^{-2} (thick dashed line), 30 kW cm^{-2} (thin solid line), and 0 (thin dashed line). Panel (b) shows in more detail a portion of the absorption line that corresponds to minimum signal-wave absorption. We see that the wave absorption decreases with increasing pump intensity. All parameters are the same as those in Fig. 3b.

uniformity of the magnetic field H and density N_0 are given by relation (32):

$$\delta H, \delta N_0 \sim \max\left\{\left(\frac{\omega_H \xi_{EC}}{\omega_p}\right)^{1/2}, \frac{k_L^2 V_T^2}{\omega_p^2}\right\}. \quad (34)$$

Here, δH and δN_0 are the relative detunings of H and N_0, respectively. For a pump intensity of the order of 100 kW cm^{-2} and for temperatures of the order of 1 and 10 keV, δH and δN_0 must be no larger than 1 and 5%, respectively. In this case, the characteristic size of the working region must be of the order of one meter, as we

Table

I_2, kW cm^{-2}	300	100	10	0
L, cm	60	15	2	0.025
$\Delta\omega/\omega_1$ (at Im k_1/Re k_1 = 0.1)	$\approx 1 \times 10^{-3}$			–
V_{gr}/c	$\approx 3 \times 10^{-3}$			–

[4] Within the EIT window, the group velocity changes from zero to the value given in the table.

see from the table. However, it may well be that a more detailed EIT theory for an inhomogeneous layer will yield more relaxed constraints.

CONCLUSION

We have shown that EIT in plasma also exists when thermal electron motion is taken into account: there is a high group deceleration and the formation of a transparency band within the cyclotron absorption line of the signal wave. At the same time, thermal motion was found to significantly change the behavior of the dispersion curves and the signal-wave absorption line. These relationships are mainly determined by the dispersion law of the damping system (plasma waves) and are virtually independent of the pump wave intensity. EIT takes place in an invariable range of frequencies and wavenumbers of the signal wave for arbitrary pump intensities (in our examples, of the order of 10–100 kW cm^{-2}). At the same time, the corresponding frequency range in cold plasma narrows with decreasing pump intensity.

The possible applications of EIT in ensembles of classical electrons can be associated with its realization in plasma microwave electronic devices (e.g., to compress pulses) or with problems of thermonuclear plasma spectroscopy. However, a generalization of the theory developed here to wave propagation at an arbitrary angle to the magnetic field in inhomogeneous plasma sheets is required for more detailed applications.

ACKNOWLEDGMENTS

This study was supported by the Russian Foundation for Basic Research (project nos. 01-02-17388 and 01-02-17779).

REFERENCES

1. S. E. Harris, Phys. Today **50** (7), 36 (1997).
2. S. E. Harris, Phys. Rev. Lett. **82**, 4611 (1999).
3. C. Liu, Z. Dutton, C. H. Behroozl, *et al.*, Nature **409**, 490 (2001).
4. D. F. Phillips, A. Fleischhauer, A. Mair, *et al.*, Phys. Rev. Lett. **86**, 783 (2001).
5. S. E. Harris, Phys. Rev. Lett. **77**, 5357 (1996).
6. A. B. Matsko and Y. V. Rostovtsev, Phys. Rev. E **58**, 7846 (1998).
7. D. F. Gordon, W. B. Mori, and C. Joshi, Phys. Plasmas **7**, 3145 (2000); **7**, 3156 (2000).
8. A. G. Litvak and M. D. Tokman, Phys. Rev. Lett. **88**, 095003 (2002).
9. M. I. Rabinovich and D. I. Trubetskov, *Introduction to the Theory of Oscillations and Waves* (Nauka, Moscow, 1984).
10. A. A. Andronov, A. A. Vitt, and S. É. Khaĭkin, *Theory of Oscillations* (Nauka, Moscow, 1981).
11. A. I. Akhiezer, I. A. Akhiezer, R. V. Polovin, *et al.*, *Plasma Electrodynamics* (Nauka, Moscow, 1974; Pergamon, Oxford, 1975).
12. V. E. Golant, A. P. Zhilinskii, and I. E. Sakharov, *Fundamentals of Plasma Physics* (Atomizdat, Moscow, 1977; Wiley, New York, 1980).
13. V. L. Ginzburg, *The Propagation of Electromagnetic Waves in Plasmas* (Nauka, Moscow, 1967; Pergamon, Oxford, 1970).

Translated by V. Astakhov

Journal of Experimental and Theoretical Physics, Vol. 95, No. 4, 2002, pp. 705–709.
Translated from Zhurnal Éksperimental'noĭ i Teoreticheskoĭ Fiziki, Vol. 122, No. 4, 2002, pp. 814–819.

PLASMA, GASES

The Influence of the Surface Electron Layer on the Energy Spectrum of Photoelectrons

M. A. Yakovlev[a, *] and M. V. Fedorov[b]

[a]*Bauman State Technical University, Vtoraya Baumanskaya ul. 5, Moscow, 107005 Russia*
[b]*Institute of General Physics, Russian Academy of Sciences, ul. Vavilova 38, Moscow, 119991 Russia*
**e-mail: yakovlev@mx.bmstu.ru*

Received May 6, 2002

Abstract—The influence of the self-consistent electric field of a surface electron layer on the energy spectrum of photoelectrons emitted under the action of a picosecond laser pulse on a metallic target was determined.

1. INTRODUCTION

Experimental studies of the energy spectrum of photoelectrons formed under the action of picosecond laser pulses on a metallic target revealed the presence of photoelectrons with unexpectedly high energies, up to 600 eV [1]. Similar earlier studies performed using nanosecond laser pulses of the same energy and wavelength gave no evidence of the existence of photoelectrons with energies exceeding 10 eV [2].

The initial theoretical explanations related the appearance of "fast" photoelectrons to the influence of a surface electron cloud formed when laser radiation acted on the surface of a metal [1, 3]. However, the mechanism of the formation of the surface electron cloud was not suggested in these works, and its physical properties were not determined. In particular, the conditions imposed on laser radiation parameters (laser pulse intensity and width) under which the produced electron layer could noticeably influence surface processes were not determined. Accordingly, the obtained theoretical results were purely qualitative in character and were incapable of providing a convincing explanation of the appearance of fast electrons.

In the more recent theoretical work [4], the formation of fast electrons was related to surface polarization of metallic conductor electrons under the action of glancing laser radiation rather than to electronic cloud influence. The main claim in [4] was that phase coherence between the laser electric field and surface electron oscillations inside a metal was established when a glancing laser beam with the required polarization properties passed along the surface of the metal. The necessary condition for the establishment of phase coherence was the requirement of smallness of electron temperature under irradiation conditions (the temperature of electrons had to be of the order of the temperature of the lattice, that is 300 K).

Note, however, that the phase coherence condition cannot be satisfied in real experiments with picosecond laser pulses. Indeed, even for a Nd laser beam glancing along a metallic surface (at an angle of about 5° in real experiments [1]), we cannot ignore heating of electrons inside the metal in a layer of thickness of the order of laser radiation wavelength ($\lambda = 1.06 \times 10^{-4}$ cm, that is, heating depth of about 10^4 lattice layers). The gap between the temperature of electrons and the temperature of the lattice under Nd laser pulses of an $I \sim 10^{10}$ W/cm^2 intensity reaches values of about 10^5 K [5], which violates the phase coherence condition.

In this work, the appearance of fast photoelectrons under picosecond laser actions on a metallic target is related to the influence of the electric field formed under certain conditions in a surface electron layer on separate photoelectrons [6].

2. SURFACE ELECTRON LAYER

The main reason for the formation of a surface electron layer under the action of ultranarrow laser pulses on a metallic target is a substantial (of the order of 1 eV) gap between the temperature of the electronic component and the temperature of the lattice [5]. This causes a sharp increase in thermoemission current and the formation of a fairly extended volume negative charge region near the surface.

The double electric layer is known to exist near the surface of a metal also under the conditions of thermodynamic equilibrium between the electronic and lattice subsystems. The electron gas is then degenerate, and the concentration of electrons decreases at a very high rate as the distance from the surface increases,

$$n_e \propto z^{-2}\exp(-\beta z),$$

where β^{-1} is a value of the order of the mean interelectronic distance in the metal [7]. The surface layer of

degenerate electrons can therefore be considered arbitrarily thin, and its influence on laser beam penetration inside the metal can be ignored. The degree of degeneracy, however, decreases at a high rate as the concentration of electrons lowers, and electrons experience the transition to the classical state when the Fermi energy $E_F(n_e)$ becomes of the order of kT. A further decrease in the concentration of electrons follows a much smoother law [6],

$$n_e(z) = n_0\left(1 + \frac{z}{\sqrt{2}L_d}\right)^{-2}, \quad (1)$$

where

$$L_d = (\varepsilon_0 kT/e^2 n_0)^{1/2}$$

is the Debye screening length and n_0 is the concentration boundary determined from the degeneracy condition

$$E_F(n_0) \approx kT,$$

that is,

$$n_0 \propto T^{3/2}.$$

It follows that the dimensions of the surface electron layer region, where the concentration decreases according to (1), increase as the temperature of electrons rises. Simultaneously, concentration boundary n_0 also increases, and, at temperatures of the order of E_F, all electrons in the surface electron layer experience the transition to the classical state. Their distribution then obeys (1), and n_0 becomes of the order of the concentration of electrons in the metal. It follows that intense nonequilibrium heating of the electronic component of a metal can cause the formation of a fairly extended layer of electrons with a high concentration near the surface.

The formation of a surface electron layer whose electrophysical characteristics enable it to substantially influence near-surface processes under the action of ultranarrow laser pulses is only possible within limited laser pulse intensity I_{em} and width τ_p intervals,

$$I_{em}^{\min} \le I_{em} \le I_{em}^{\max}, \quad \tau_p^{\min} \le \tau_p \le \tau_p^{\max}.$$

The $I_{em}^{\min}$ lower intensity boundary is determined from the condition that a fairly effective gap between the T_e electronic component temperature and the T_l lattice temperature should be attained during a pulse,

$$\Delta T = T_e - T_l \approx T_F \sim 10^4\text{–}10^5 \text{ K},$$

where T_F is the temperature of the degenerate electronic subsystem; that is (see [8]),

$$I_{em}^{\min} \approx \alpha_l l \Delta T,$$

where $\alpha_l \sim 10^{10}$ W/cm^3 K) is the energy exchange rate between the electronic and lattice subsystems and l is the depth of the metal layer heated during a pulse, $l \sim \max[\delta, \sqrt{\chi\tau_p}]$ (δ is the skin depth, and χ is the electron thermal diffusivity). Picosecond pulses are characterized by $l \sim 10^{-5}$ cm, that is, $I_{em}^{\min} \sim 10^9$ W/cm^2.

According to [5], when picosecond laser pulses of intensity $I_{em} \sim 3 \times 10^9$ W/cm^2 act on the surface of typical metals, two competing processes of electron escape from the surface, namely, thermoemission and photoemission, have the same order of magnitude. We can therefore expect that the collective thermoemission process should most noticeably influence the one-particle photoemission process under these conditions. Thermoemission begins to predominate over photoemission as the intensity of radiation increases; that is, because of thermoemission, the formation of the surface electron layer at the specified radiation intensities occurs in time

$$\tau_s \approx L_d/v_T \sim 10^{-14}\text{–}10^{-13} \text{ s}.$$

Here, v_T is the thermal velocity of electrons [6]. Note in addition that the character of thermoemission is then substantially different from that of thermoemission from an electrode in a closed circuit, because thermoemission that we are considering occurs from an insulated metallic surface, on which an uncompensated positive charge remains, and a volume negative charge is formed near the surface, which in turn influences the thermoemission current. This considerably complicates the description of the formation of the surface electron layer. Nevertheless, the time τ_s of formation of the spatial distribution is much shorter than the picosecond laser pulse width τ_p. The spatial distribution of the surface electron layer can therefore be considered stationary with the corresponding layer temperature virtually over the whole pulse width.

The upper intensity boundary $I_{em}^{\max}$ corresponds to the prethreshold region of the beginning of melting and ablation of the target material. Such processes are observed when the density of the energy accumulated in the electronic subsystem under the action of an ultranarrow laser pulse exceeds a certain threshold,

$$I_{em}^{\max}\tau_p \le F_{abl},$$

where the threshold energy density of laser ablation $F_{abl} \approx 0.2$–0.5 J/cm^2 [9, 10]; that is, at a $\tau_p \sim 1$ ps pulse width, we have $I_{em}^{\max} \sim 10^{12}$ W/cm^2.

The limitations imposed on the width of laser pulses τ_p are determined from the conditions

$$\tau_p^{\min} \ge \tau_s,$$

where $\tau_s \sim 10^{-13}$ s is the time of formation of the surface electron layer with a nondegenerate electronic component, and

$$\tau_p^{\max} \le \tau_{el},$$

where τ_{el} is the characteristic time of energy transfer from electrons to the lattice. The second condition guarantees that the lattice is not heated during a laser pulse and the pulse width is insufficient for creating conditions of developed vaporization of the target material, when the role played by the surface electron layer loses significance.

To summarize, it follows from the results described above that the surface electron layer can exert substantial influence on surface processes within limited laser pulse intensity and width ranges, which are, however, fairly important for technical applications, namely, 10^{10} W/cm$^2 \le I_{em} \le 10^{12}$ W/cm^2 and $10^{-13} \le \tau_p \le 10^{-11}$ s, respectively.

3. THE ENERGY SPECTRUM OF FAST PHOTOELECTRONS IN THE ELECTRIC FIELD OF THE SURFACE ELECTRON LAYER

As in [11], the self-consistent electric field of the surface electron layer E_z created by the positive charge of the conductor and electrons under surface irradiation by picosecond laser pulses was found by solving the set of equations that described electronic component temperature variations with time and electron layer formation. The ionization of the neutral gas taken into account in [11] could, however, be ignored at pressures and picosecond laser pulse intensities under consideration (of about 1 atm and ~10^{10} W/cm^2, respectively).

Let us turn to the system of equations that describe surface electron layer formation near the surface of a condensed substance. First consider the heat equation. At $z < 0$ (a conducting condensed substance), we have

$$C_m \frac{\partial T_e}{\partial t} = \frac{\partial}{\partial z}\left[\chi_m \frac{\partial T_e}{\partial z}\right] - \alpha(T_e - T_l) + q(z, t). \qquad (2)$$

Here, C_m and χ_m are the heat capacity and the heat conductivity of the conducting condensed substance, respectively, and α is the coefficient of heat exchange between electrons and the lattice, which is virtually independent of the temperature of electrons [12],

$$\alpha = \frac{\pi^2}{6}\frac{m c_s^2 \nu_{\text{eff}} n_m}{T_l},$$

where $\nu_{\text{eff}} = \nu_{\text{eff}}(T_l)$ and c_s is the velocity of sound in the conducting condensed substance. For typical metals with $n_m \sim 10^{22}$ cm^{-3}, we have $\alpha \sim 10^{10}$ W/(cm^3 K). The heat capacity and heat conductivity of the condensed substance were approximated by the asymptotic equations [12]

$$C_m = \begin{cases} \dfrac{\pi^2}{2} n_m k\left(\dfrac{kT_e}{E_F}\right), & kT_e \ll E_F, \\ \dfrac{3}{2} n_m k, & kT_e \ge E_F, \end{cases}$$

$$\chi_m = \begin{cases} \chi_{m0}\dfrac{T_e}{T_l}, & kT_e \ll E_F, \\ \dfrac{1}{2} n_m k \dfrac{\bar{v}_e^2}{\nu_{\text{eff}}}, & kT_e \ge E_F, \end{cases}$$

where $\bar{v}_e$ is the mean thermal velocity of electrons and χ_{m0} is the equilibrium thermal conductivity coefficient of the condensed conducting substance at initial temperature T_0. The $q(z, t)$ volume energy release function has the form

$$q(z, t) = \kappa_r \kappa_i k_0 \varepsilon_0 c |E_0|^2 \exp(2k_0 \kappa_i z),$$

where κ_r and κ_i are the refractive indexes of the condensed conducting substance and E_0 is the amplitude of the wave field at $z = 0$.

The heat equation for a gas ($z > 0$) has the form

$$\frac{3}{2} k \frac{\partial T_e}{\partial t} = \frac{1}{n_e}\frac{\partial}{\partial z}\left[\chi_e \frac{\partial T_e}{\partial z}\right] - \frac{3m}{M} k (T_e - T_a)\nu_e + \frac{e^2 |E|^2 \nu_e}{2m(\omega^2 + \nu_e^2)}, \qquad (3)$$

where M is the atomic weight of the gas (in m units), $\nu_e = \sigma_a n_a \bar{v}_e$ is the frequency of electronic collisions in the gas, and $\sigma_a = \sigma_a(\bar{v}_e)$ is the transport cross section of electron scattering by atoms. The $\sigma_a(\bar{v}_e)$ dependence is well known for rare gases. We used the data from [13] in our calculations.

The equations describing variations in lattice temperature T_l and in the T_a temperature of the heavy gas component (atoms) have the form

$$C_l \frac{\partial T_l}{\partial t} = \alpha(T_e - T_l),$$

$$\frac{\partial T_a}{\partial t} = \frac{2m}{M}\frac{n_e}{n_a}\nu_{ea}(T_e - T_a),$$

where C_l is the heat capacity of the lattice:

$$C_l \approx 3 n_l k.$$

Problem conditions correspond to very fast processes ($t \ll \tau_l \sim 10^{-10}$ s), and, according to the equations given

Table 1

$I_{em}^{(1)} = 0.3 \times 10^{10}$ W/cm^2	$I_{em}^{(2)} = 10^{10}$ W/cm^2	$I_{em}^{(3)} = 3 \times 10^{10}$ W/cm^2
$T_1 = 4.6 \times 10^4$ K	$T_1 = 7.5 \times 10^4$ K	$T_3 = 11.3 \times 10^4$ K
$n_0^{(1)} = 3.8 \times 10^{17}$ cm^{-3}	$n_0^{(2)} = 2.2 \times 10^{19}$ cm^{-3}	$n_0^{(3)} = 6 \times 10^{19}$ cm^{-3}
$E_1 = 1.3 \times 10^7$ V/cm	$E_2 = 1.6 \times 10^7$ V/cm	$E_3 = 1.8 \times 10^7$ V/cm

above, the temperatures of the lattice and the heavy gas component change very insignificantly during such time intervals. For instance, even at the mean electron temperature $T_e \sim 10$ eV, the temperature of the lattice changes by a value of the order of 10^2 K in time $t \sim 10^{-13}$ s. For this reason, we assumed in our calculations that $T_l = T_a = T_0$.

The boundary conditions for (2) and (3) have the form

$$\frac{\partial T_e}{\partial z} = 0, \quad z = -l_m, l_a,$$
$$\left[\chi \frac{\partial T_e}{\partial z}\right] = 0, \quad z = 0, \tag{4}$$

where l_m and l_a are the boundaries of the region under consideration in the condensed conducting substance and gas, respectively (formally, $l_m, l_a \longrightarrow \infty$).

Next, consider the equations that describe the kinetics of surface electron layer formation. The continuity equation for the concentration of electrons should be solved simultaneously with the Maxwell equations in the $z < 0$ and $z > 0$ regions. It is, however, sufficient to consider the continuity equation in the $z > 0$ region rather than in the whole space. The boundary condition at $z = 0$ should then correspond to distributions (1) of surface electron layer electrons as functions of temperature (see above),

$$n_e|_{z=0} = n_0,$$

where n_0 is the boundary concentration of the surface electron layer at a given time moment (note that surface electron layer electrons form an ideal Coulomb system at $T \geq 4$–5 eV). There is sufficient time for the concentration profile of the surface electron layer near the surface to keep track of temperature variations, because the characteristic time of the formation of this layer is about ω_p^{-1}.

Table 2

$L_s^1 = 2.2 \times 10^{-5}$ cm	$L_s^2 = 10^{-6}$ cm	$L_s^3 = 8.5 \times 10^{-7}$ cm
$W_e^1 = 191$ eV	$W_e^2 = 46$ eV	$W_e^3 = 98$ eV

Consider the continuity equation for the electronic component,

$$\frac{\partial n_e}{\partial t} = \frac{\partial}{\partial z}\left[D_e \frac{\partial n_e}{\partial z} + \mu_e E_z n_e\right], \tag{5}$$

where D_e and μ_e are the diffusion and mobility coefficients defined as

$$\mu_e = \frac{e}{m_e \nu_e}, \quad D_e = \frac{v_e^2}{3\nu_e}.$$

The boundary conditions for (5) have the form

$$z = 0: \; n_e = n_0,$$
$$z = l_a: \; D_e \frac{\partial n_e}{\partial z} + \mu_e E_z n_e = 0. \tag{6}$$

The initial condition for n_e corresponds to the distribution of surface layer electrons at the initial temperature T_0. The distribution of the longitudinal electric field E_z is found from the equation

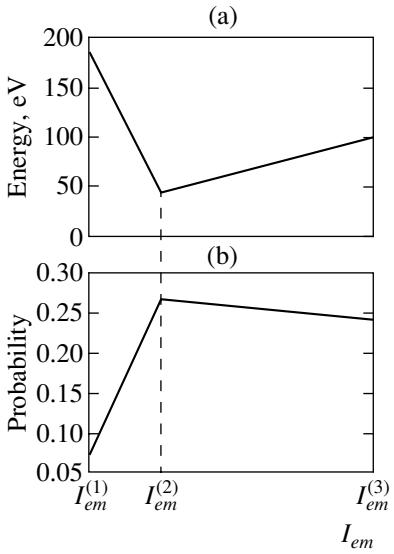
$$\frac{\partial E_z}{\partial z} = -\frac{e n_e}{\varepsilon_0}. \tag{7}$$

The boundary condition that corresponds to this equation at each time step has the form

$$E_z|_{z=0} = \frac{e}{\varepsilon_0}\int_0^{l_a} n_e dz.$$

The results obtained in calculating the electrophysical parameters of the surface electron layer near the surface of a metallic target are listed in Table 1 as a function of the intensity of the picosecond laser pulse polarized along the surface of the target ($T_{1,2,3}$ and n_0 are the temperature and the concentration of electrons, respectively, and $E_{1,2,3}$ is the self-consistent electric field in the surface electron layer directed normally to the surface of the target). Taking into account the Coulomb character of collisions between surface electron layer electrons, we can use the data given in Table 1 to determine the free path of electrons L_s and the energy W_e accumulated along the free path. These data are listed in Table 2.

The energy accumulated by a separate photoelectron in the electric field of the surface electron layer should be determined taking into account the special



(a) Maximum energy of fast photoelectrons and (b) probability of the appearance of electrons with a given energy as functions of the intensity of picosecond laser pulses ($I_{em}^{(1)} = 0.3 \times 10^{10}$ W/cm^2, $I_{em}^{(2)} = 1.1 \times 10^{10}$ W/cm^2, $I_{em}^{(3)} = 3 \times 10^{10}$ W/cm^2).

features of the spatial profile of the surface layer in the direction normal to the surface of the target (axis z). The concentration distribution of surface thermal electrons can conventionally be divided into two main regions. The first one (region 1) is the near-wall region whose size is close to the L_D Debye length. The major part of surface layer electrons are concentrated in this region. Note, however, that our problem has planar symmetry, and the electric field decreases fairly slowly within this layer. According to (1) and (7), electric field E_z is inversely proportional to z with scale L_D. It follows that the electric field of the positively charged surface of a conductor passes a plane Debye layer, within which it decreases insignificantly (approximately by a factor of 2), and passes into a relatively rarefied region outside the Debye layer, whose characteristic size is of the order of several times the free path L_s. This is region 2. It follows that a separate photoelectron can accumulate substantial energy in region 2 over a free path distance along the line of the electric field. After elastic reflection from region 1 with a change in the velocity direction, this electron can escape the near-wall region at a small angle with respect to the surface of the conductor with a complete store of accumulated energy. The probability of such an elementary event can be estimated at L_D/L_s.

4. CONCLUSION

The following conclusions can be drawn from a comparison of the numerical calculation results with the data of real experiments [1, 2].

(1) Fast photoelectrons were not emitted under the action of nanosecond (τ_p = 27 ns) Nd laser pulses of energy W_p = 0.05 J [2] because the surface electron layer was not formed under these conditions. Accordingly, there was no electric field capable of "heating" photoelectrons to substantial energies.

(2) The experimental data on the action of picosecond (τ_p = 8 ps) Nd laser pulses with intensities of 1.3×10^{10} to 2.5×10^{10} W/cm^2 showed that photoelectrons formed under these conditions had maximum energies of 100 to 600 eV [1]. These data were fairly well explained by numerical calculations of the influence of the electric field of a surface electron layer on separate photoelectrons; such a layer is formed at the specified laser radiation parameters, as is well seen from the figure.

ACKNOWLEDGMENTS

The authors thank A.A. Rukhadze for interest in this work and support.

REFERENCES

1. G. Farkas and C. Toth, Phys. Rev. A **41**, 4123 (1990).
2. E. M. Logothetis and P. L. Hartman, Phys. Rev. **187**, 460 (1969).
3. G. Petite, P. Agostini, R. Trainham, *et al.*, Phys. Rev. B **45**, 12210 (1992).
4. S. Varro and F. Ehlotzky, Phys. Rev. A **57**, 663 (1998).
5. S. I. Anisimov, B. L. Kapeliovich, and T. L. Perel'man, Zh. Éksp. Teor. Fiz. **66**, 776 (1974) [Sov. Phys. JETP **39**, 375 (1974)].
6. A. V. Ivlev, K. B. Pavlov, and M. A. Yakovlev, Zh. Tekh. Fiz. **64** (9), 50 (1994) [Tech. Phys. **39**, 888 (1994)].
7. A. K. Gupta and K. S. Singwi, Phys. Rev. B **15**, 1801 (1977).
8. S. I. Anisimov and B. Retfel'd, Izv. Akad. Nauk, Ser. Fiz. **61**, 1642 (1997).
9. Yu. V. Afanas'ev, N. N. Demchenko, I. N. Zavestkovskaya, *et al.*, Izv. Akad. Nauk, Ser. Fiz. **63**, 667 (1999).
10. D. Riley, A. J. Langley, P. F. Taday, *et al.*, J. Phys. D **31**, 515 (1998).
11. A. V. Ivlev, M. A. Yakovlev, and A. N. Bordenyuk, Zh. Tekh. Fiz. **68** (8), 42 (1998) [Tech. Phys. **43**, 921 (1998)].
12. S. I. Anisimov, B. I. Makshantsev, and A. V. Barsukov, Opt. Acoust. Rev. **1**, 251 (1990).
13. L. J. Kieffer, At. Data **2** (4), 293 (1971).

Translated by V. Sipachev

Journal of Experimental and Theoretical Physics, Vol. 95, No. 4, 2002, pp. 710–713.
From Zhurnal Éksperimental'noĭ i Teoreticheskoĭ Fiziki, Vol. 122, No. 4, 2002, pp. 820–823.

FLUIDS

A Liquid–Liquid Phase Transition in the "Collapsing" Hard Sphere System¶

V. N. Ryzhov[a, *] and S. M. Stishov[a, b]

[a]*Institute for High Pressure Physics, Russian Academy of Sciences, Troitsk, Moscow oblast, 142190 Russia*
[b]*Los Alamos National Laboratory, Los Alamos, NM 87545, USA*
**e-mail: ryzhov@hppi.troitsk.ru*

Received May 6, 2002

Abstract—A liquid–liquid phase transition is discovered in a system of collapsing hard spheres using the thermodynamic perturbation theory. This is the first evidence in favor of the existence of that kind of phase transition in systems with purely repulsive and isotropic interactions. *© 2002 MAIK "Nauka/Interperiodica".*

It has been known for many years that the system of hard spheres only experiences a phase transition at a high density, when $\sigma \approx l$, where l is the hard sphere diameter and $l = (V/N)^{1/3}$ is the average interparticle distance (V is the system volume and N is the number of particles). This transition corresponds to the ordering of the centers of gravity of the particles and can be called an order–disorder transition, or crystallization. In case of hard particles of different shapes such as hard rods, ellipses, and discs, a number of orientational phase transitions can occur in accordance with a hierarchy of characteristic lengths defined by particle shapes. A new situation arises when an extra interaction of a finite amplitude ε is added to the system of hard particles. As known from the van der Waals theory, a negative value of ε inevitably causes an instability of the system in a certain range of densities and generally leads to a first-order phase transition with no symmetry change (the order parameter characterizing this transition is simply the density difference of the coexisting phases, $\Delta\rho = \rho_1 - \rho_2$). This situation is almost universal and independent of the interaction length.

Much less is known about the case where the interaction parameter ε has a positive value. The simplest example of an interaction of that kind is the so-called repulsive step potential (Fig. 1):

$$\Phi(r) = \begin{cases} \infty, & r \le \sigma, \\ \varepsilon, & \sigma < r \le \sigma_1, \\ 0, & r > \sigma_1. \end{cases} \qquad (1)$$

In what follows, the system of particles interacting via potential (1) is called the system of "collapsing" hard spheres [1]. Systems of this type are studied in relation to anomalous melting curves, isostructural phase transitions, transformations in colloid systems, etc. (see, e.g., [2–5]). A general conclusion derived from numerous studies of the system is that the repulsive interaction of finite amplitude and length results in the melting curve anomaly and the isostructural solid–solid phase transition. The latter is a first-order phase transition and can end in a critical point, because there is no symmetry change across the phase transition line. The existence of a phase transition of that type is a direct consequence of the form of the interparticle interaction, and we see no particular reason why it cannot occur in a fluid phase.

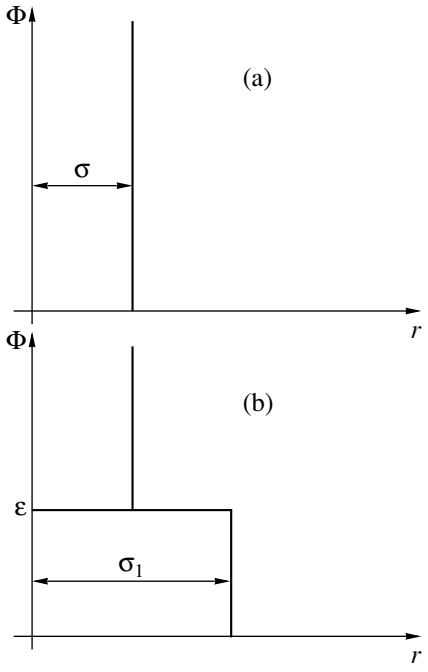


Fig. 1. (a) The hard-sphere potential with the hard-sphere diameter σ_1. (b) The repulsive step potential; σ is the hard-core diameter, σ_1 is the soft-core diameter, and ε is the height of the repulsive step.

¶This article was submitted by the authors in English.

Despite the growing interest in the possible polymorphic phase transitions in liquids and glasses (see, e.g., [6] for recent reviews), the nature of different phases that can be found in dense (and possibly metastable) liquids is still puzzling. In recent years, experimental evidence of such features of the phase diagram as liquid–liquid transitions and polymorphism appeared for a wide range of systems including water, Si, I, Se, S, C, P, etc. [6]. The complexity of the phase diagrams in these substances may result from complex interactions depending on the intermolecular orientations. At the same time, exploring the possibility that simple fluids interacting through isotropic potentials may exhibit a similar behavior represents a serious challenge for theorists.

The possibility of the existence of a liquid–liquid phase transition drastically depends on the shape of the interparticle potential. After the pioneering work by Hemmer and Stell [2], much attention has been paid to investigating the properties of the systems with the so-called core-softened potentials—the potentials that have a negative curvature region in their repulsive core. It has been shown that, depending on the parameters of the potentials, waterlike thermodynamic anomalies and the second critical point can be observed in this system [6–10]. It is widely believed, however (see, e.g., [7, 8]), that the existence of a fluid–fluid transition must be related to the attractive part of the potential. In this paper, we show that the purely repulsive step potential in Eq. (1) is sufficient to explain a liquid–liquid phase transition and the anomalous behavior of the thermal expansion coefficient.

We apply the second-order thermodynamic perturbation theory for fluids to this problem. The soft core of potential (1) (Fig. 1b) is treated as a perturbation with respect to the hard sphere potential (Fig. 1a). In this case, the free energy of the system can be written as [11, 12]

$$\frac{F-F_{HS}}{Nk_BT} = \frac{1}{2}\rho\beta\int u_1(r)g_{HS}(r)d\mathbf{r} - \frac{1}{4}\rho\beta^2\left[k_BT\left(\frac{\partial\rho}{\partial P}\right)_0\right]\int [u_1(r)]^2 g_{HS}(r)d\mathbf{r}, \quad (2)$$

where $\rho = V/N$ is the mean number density, $\beta = 1/k_BT$, $u_1(r)$ is the perturbation part of the potential $u_1(r) = \Phi(r) - \Phi_{HS}(r)$, $\Phi_{HS}(r)$ is the hard sphere singular potential, and $g_{HS}(r)$ is the hard sphere radial distribution function, which is taken in the Percus–Yevick approximation [13]. In the same approximation, the compressibility can be written as [12]

$$k_BT\left(\frac{\partial\rho}{\partial P}\right)_0 = \frac{(1-\eta)^4}{(1+2\eta)^2}. \quad (3)$$

We note that the actual small parameter in expansion (2) is the ratio $\varepsilon/(k_BT)$, and therefore, the perturbation scheme used in this paper works very well at high temperatures and gives quantitative agreement with computer simulations for intermediate temperatures and sufficiently high densities [11, 12]. We believe that, in the range of temperatures and densities considered in this paper, Eq. (2) gives correct results at least qualitatively.

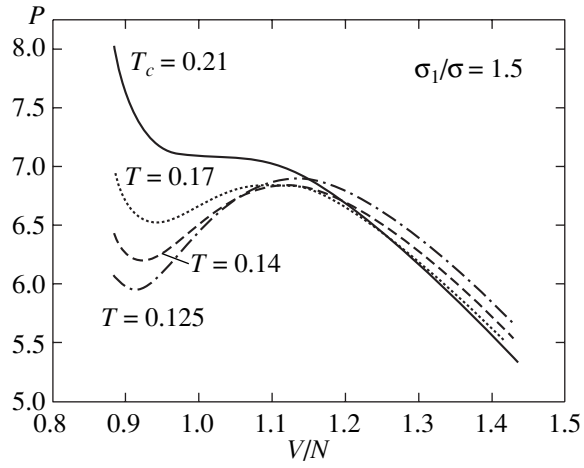


Fig. 2. Compression isotherms of the collapsing sphere system at various temperatures.

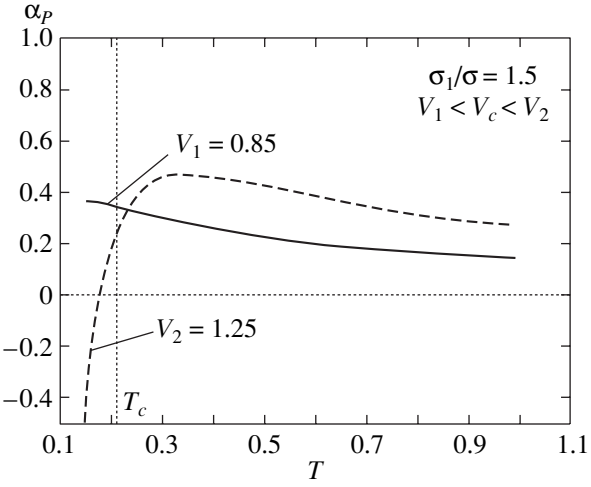


Fig. 3. The thermal expansion coefficient α_P as a function of temperature for two values of specific volume $V_1 = 0.85$ and $V_2 = 1.25$, $V_1 < V_c < V_2$.

To calculate F_{HS}, we can use, e.g., the approximate equation [12]

$$\frac{F_{HS}}{k_BTN} = 3\ln\lambda - 1 + \ln\rho + \frac{4\eta - 3\eta^2}{(1-\eta)^2}, \quad (4)$$

where $\lambda = h/(2\pi m k_BT)^{1/2}$ and $\eta = \pi\rho\sigma^3/6$. In what follows, we use the reduced quantities $\tilde{P} = P\sigma^3/\varepsilon$, $\tilde{V} = V/N\sigma^3 = 1/\tilde{\rho}$, and $\tilde{T} = k_BT/\varepsilon$, omitting the tildas.

The results of the calculations are demonstrated in Figs. 2 and 3. In Fig. 2, a family of pressure isotherms

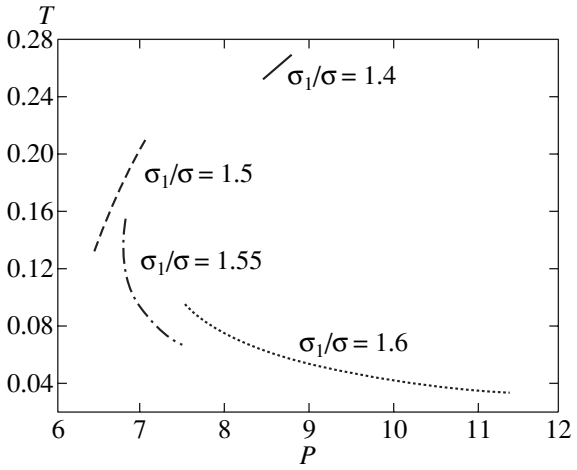


Fig. 4. Phase diagram of the liquid–liquid phase transition for different values of σ_1/σ.

is shown for the system with $\sigma_1/\sigma = 1.5$. The van der Waals loops in the isotherms are clearly seen at low temperatures, indicating the existence of a first-order liquid–liquid phase transition. A critical point is found at $T_c \approx 0.21$ and $V_c \approx 1.015$.

An interesting and unusual feature of the isotherms in Fig. 2 is their intersection in the low-density region. This behavior implies a negative value of the thermal expansion coefficient in a certain region of density and temperature. Most liquids contract upon cooling. But this is not the case for the most important liquid on Earth, water. It is well known that the specific volume of water increases when cooled below $T = 4C$. The existence of this anomaly can be related to the hypothetical liquid–liquid phase transition in supercooled water [14, 15]. This anomaly is not restricted to water but is also present in other liquids [16].

In Fig. 3, the thermal expansion coefficient $\alpha_P = V^{-1}(\partial V/\partial T)_P$ is shown as a function of the temperature for two values of specific volume $V_1 = 0.85$ and $V_2 = 1.25$, corresponding to high-density and low-density liquids, respectively. We can see that for the low-density liquid, there is a range of negative values of α_P below the critical temperature. The appearance of the negative thermal expansion coefficient can easily be understood for repulsive step potential (1). At low temperatures and densities, the particles do not penetrate the soft core of the potential. As the temperature increases, the particles can penetrate the soft core, and the average distance between particles can therefore decrease, resulting in the anomalous behavior of α_P. At high densities, this process is less pronounced because most of the particles are inside the soft core due to the external contraction. It must be noted that Stillinger and collaborators also found a negative thermal expansion coefficient $\alpha_P < 0$ for a purely repulsive Gaussian potential [17, 18], but without a liquid–liquid transition.

Using the Maxwell construction, we were able to calculate the equilibrium lines of the liquid–liquid phase transitions at different values of σ_1/σ (Fig. 4). We cannot extend the transition lines to zero temperature because of the limitations of the perturbative approach. We can see from Fig. 4 that the critical temperature decreases as the ratio σ_1/σ increases and the slope of the transition curves drastically changes with changes in the ratio σ_1/σ and the temperature. The decrease in the critical temperature may suggest that the transition ceases to exist at high values of the ratio σ_1/σ. This is not surprising because, for a long-range interaction, perturbation energy can be treated in the mean-field approximation. The perturbation energy is then a positive monotonic function of the volume ($\delta F \propto 1/V$) and cannot provide any grounds for the existence of a phase transition. In accordance with the Clausius–Clapeyron equation $dT/dP = \Delta V/\Delta S$ (where ΔV and ΔS are the variations of volume and entropy at the transition), the change in the slope of the transition line implies that the entropy jump at the transition changes its sign for different values of the ratio σ_1/σ and the temperature. This behavior of the entropy change can possibly be understood in terms of the entropy of mixing, implying that two states of particles in the system can be considered as two different species.

The liquid–liquid transition line found most probably lies below the melting curve and can be observed only in the metastable liquid state, as was discovered in supercooled water [14, 15]. On the other hand, the liquid–liquid transition can be observed in stable liquids in some cases [6, 19, 20]. It must be noted that computer simulations also show [8] that, for some choices of the parameters of the potential (and in the presence of the attractive part of the potential), a liquid–liquid phase transition can occur in the stable range of the phase diagram.

We note that a second phase transition corresponding to the liquid–gas transformation can be expected when an attractive tail is appended to the repulsive step potential, as was observed in the molecular dynamic simulations [9, 10]. We performed the corresponding calculations using the second-order perturbation scheme with the parameters of the core-softened potential proposed by Stanley and coauthors [9, 10] and found a second phase transition and a second critical point. That may be viewed as some sort of justification of our approach to phase transformations in liquids.

Finally, for the first time, we found essential evidence for a first-order phase transition in the liquid state of the system of collapsing hard spheres.

We thank V.V. Brazhkin, A.G. Lyapin, and E.E. Tareyeva for stimulating discussions. Thise work was supported by the Russian Foundation for Basic Research (project no. 02-02-16622).

REFERENCES

1. S. M. Stishov, Zh. Éksp. Teor. Fiz. **122**, 76 (2002) [JETP **95**, 64 (2002)].
2. P. C. Hemmer and G. Stell, Phys. Rev. Lett. **24**, 1284 (1970); G. Stell and P. C. Hemmer, J. Chem. Phys. **56**, 4274 (1972).
3. J. M. Kincaid, G. Stell, and E. Goldmark, J. Chem. Phys. **65**, 2172 (1976).
4. D. A. Young and B. J. Alder, Phys. Rev. Lett. **38**, 1213 (1977); J. Chem. Phys. **70**, 473 (1979).
5. P. Bolhuis and D. Frenkel, J. Phys.: Condens. Matter **9**, 381 (1997).
6. *Proceedings of the NATO Advanced Research Workshop "New Kinds of Phase Transitions: Transformations in Disordered Substances", Uglich, Russia, 2001*, Ed. by V. V. Brazhkin, S. V. Buldyrev, V. N. Ryzhov, and H. E. Stanley (Kluwer, Dordrecht, 2002).
7. E. A. Jagla, J. Chem. Phys. **111**, 8980 (1999).
8. E. A. Jagla, Phys. Rev. E **63**, 061501 (2001).
9. G. Franzese, G. Malescio, A. Skibinsky, *et al.*, Nature **409**, 692 (2001).
10. S. V. Buldyrev, G. Franseze, N. Gioivambattista, *et al.*, in *Proceedings of the NATO Advanced Research Workshop "New Kinds of Phase Transitions: Transformations in Disordered Substances", Uglich, Russia, 2001*, Ed. by V. V. Brazhkin, S. V. Buldyrev, V. N. Ryzhov, and H. E. Stanley (Kluwer, Dordrecht, 2002), p. 97.
11. J. A. Barker and D. Henderson, J. Chem. Phys. **47**, 2856 (1967).
12. J. A. Barker and D. Henderson, Rev. Mod. Phys. **48**, 587 (1976).
13. W. R. Smith and W. Henderson, Mol. Phys. **19**, 411 (1970).
14. O. Mishima and H. E. Stanley, Nature **396**, 329 (1998).
15. H. E. Stanley, S. V. Buldyrev, N. Gioivambattista, *et al.*, in *Proceedings of the NATO Advanced Research Workshop "New Kinds of Phase Transitions: Transformations in Disordered Substances", Uglich, Russia, 2001*, Ed. by V. V. Brazhkin, S. V. Buldyrev, V. N. Ryzhov, and H. E. Stanley (Kluwer, Dordrecht, 2002), p. 311.
16. P. G. Debenedetti, *Metastable Liquids* (Princeton Univ. Press, Princeton, 1996).
17. F. H. Stillinger and T. A. Weber, J. Chem. Phys. **68**, 3837 (1978).
18. F. H. Stillinger and T. A. Weber, J. Chem. Phys. **74**, 4015 (1981).
19. V. V. Brazhkin, S. V. Popova, and R. N. Voloshin, High Press. Res. **15**, 267 (1997).
20. Y. Katayama, T. Mizutari, W. Utsumi, *et al.*, Nature **403**, 170 (2000).

Journal of Experimental and Theoretical Physics, Vol. 95, No. 4, 2002, pp. 714–727.
From Zhurnal Éksperimental'noĭ i Teoreticheskoĭ Fiziki, Vol. 122, No. 4, 2002, pp. 824–839.

FLUIDS

Disclination Motion in Liquid Crystalline Films¶

E. I. Kats[a, b, *], V. V. Lebedev[a, c, **], and S. V. Malinin[a, d, *]**

[a]*Landau Institute for Theoretical Physics, Russian Academy of Sciences, Moscow, 117940 Russia*
[b]*Laue–Langevin Institute, F-38042, Grenoble, France*
[c]*Theoretical Division, Los Alamos National Laboratory, Los Alamos, NM 87545, USA*
[d]*Forschungszentrum Jülich, D-52425, Jülich, Germany*
**e-mail: kats@ill.fr*
***e-mail: lebede@landau.ac.ru*
****e-mail: malinin@itp.ac.ru*

Received May 24, 2002

Abstract—We theoretically study a single disclination motion in a thin free-standing liquid crystalline film. Backflow effects and the own dynamics of the orientational degree of freedom (bond or director angle) are taken into account. We find the orientation field and the hydrodynamic velocity distribution around the moving disclination, which allows us to relate the disclination velocity to the angle gradient far from the disclination. Different cases are examined depending on the ratio of the rotational and shear viscosity coefficients. © *2002 MAIK "Nauka/Interperiodica".*

1. INTRODUCTION

The physics of thin liquid-crystalline films has been a recurrent hot topic during the past decade because of their intriguing physical properties and a wide range of applications in display devices, sensors, and for many other purposes. Hexatic, nematic, and smectic-*C* liquid-crystalline films belong to two-dimensional systems with a spontaneously broken continuous rotational symmetry. An essential role in the behavior of the films is therefore played by vortexlike excitations (disclinations). Defects are almost necessarily present in liquid crystals, and their dynamics plays a crucial role in the overall pattern organization. Early studies of defects focused on classifying the static properties of the defects and their interactions [1, 2]. More recently, the focus has shifted to examining the dynamics of defects (see, e.g., [3] and references therein). We note that, although defects are undesirable in most practical applications of liquid crystals, such as traditional display devices, because they destroy an optical adjustment, there are novel display designs (bistable, multidomain liquid-crystalline structures) exploiting defect properties.

Although experimental dynamic studies are likely to be more fruitful than static ones, theoretical research of the film dynamics is in a rather primitive stage. This is largely accounted for by a complexity of dynamic phenomena in films, and a complete and unifying description of the problem is still unavailable. Moreover, some papers devoted to this problem (dynamics of defects) claim contradicting results. These contradictions come mainly from the fact that different authors take different microscopic dissipation mechanisms into account, but partially the source of controversy is related to semantics, because different definitions of the forces acting on defects are used (see, e.g., the discussion in [4]). We believe that such problems are irrelevant if the macroscopic (phenomenological) approach to the film dynamics is used.

In this paper, we theoretically examine the disclination dynamics in free-standing liquid-crystalline films at scales that are much larger than the film thickness, where the films can be treated as 2D objects. Our investigation is devoted to the first (but compulsory) step of defect dynamics studies: a single point disclination in a liquid-crystalline film. A number of theoretical efforts [5–9] deal with similar problems. Our justification for adding one more paper to the subject is the fact that, in the literature, we did not see a full investigation of the problem with the hydrodynamic backflow effects taken into account. Evidently, these effects can drastically modify the dynamics of defects. The goal of this work is to study the disclination motion in free-standing liquid crystalline films on the basis of hydrodynamic equations containing some phenomenological parameters (the elasticity modulus and shear and rotational viscosity coefficients).

In our approach, the disclination is assumed to be driven by a large-scale inhomogeneity in the bond or director angle, which leads to a motion of the disclination with a nonzero velocity relative to the film. As a physical realization of such a nonuniform angle field, a system of disclinations distributed with a finite density can be imagined. The inhomogeneity in the vicinity of a given disclination is then produced by fields of other disclinations. We can also think about a pair of discli-

¶This article was submitted by the authors in English.

nations of the opposite topological charges, in which case the inhomogeneity is related to the mutual orientational distortion fields created by each disclination at the point of its counterpart. In fact, the majority of experimental and numerical studies of disclination motions in liquid crystals [10–18] is devoted to the investigation of the dynamics of two oppositely charged defects. We solve the hydrodynamic equations and find the bond (director) angle and the flow velocity distributions around the moving disclination. The results enable us to relate the disclination velocity and the gradient of the angle far from the disclination.

An obvious context where our results can be applied is the film dynamics near the Berezinskii–Kosterlitz–Thouless phase transition. The static properties of the films near the transition have been investigated in a great number of papers starting from the famous papers by Berezinskii [19] and Kosterlitz and Thouless [20]. There are several works discussing the theory of dynamic phenomena associated with vortexlike excitations in condensed matter physics: vortices in type-II superconductors (see, e.g., [21]), vortices in superfluid ^{4}He and ^{3}He (see, e.g., [22, 23]), dislocations in 2D crystals, and disclinations (and other topological defects) in liquid crystals (see [10–14, 24–27]). But most of the theoretical works on the subject start from phenomenological equations of motion of the defects, and our aim is to derive the equations and to verify their validity.

The structure of our paper is as follows. Section 2 contains basic hydrodynamic equations for liquid-crystalline films necessary for our investigation. In Section 3, we find the bond (director) angle and the flow velocity around the uniformly moving disclination, which allows us to relate the disclination velocity to the angle gradient far from the disclination. Different cases, depending on the ratio of the rotational and shear viscosity coefficients, are examined in Section 4. Section 5 contains a summary and discussion. The appendices are devoted to the details of calculations of the velocity and bond angle fields around the moving disclination. Those readers who are not very interested in mathematical derivations can skip these appendices, finding all the essential physical results in the main text of the paper.

2. BASIC RELATIONS FOR LIQUID-CRYSTALLINE FILMS

We formulate the basic relations needed to describe a disclination motion in thin liquid-crystalline films. Here, we investigate freely suspended hexatic, nematic, and smectic-*C* films that can be pulled from 3D (bulk) smectics [3]. We examine scales larger than the film thickness, where the films can be treated as two-dimensional objects and can be described in terms of a macroscopic approach containing some phenomenological parameters.

Liquid crystalline films with the in-plane orientational ordering of different types (hexatic, nematic, and smectic-*C*) are observed experimentally. In these films, as in 3D nematic liquid crystals, the rotational symmetry is spontaneously broken. The general analysis of their symmetry can be found in [28]. The smectic-*C* films are characterized by the director that is tilted with respect to the normal to the film, which defines a preferred direction in the plane of the film. The ordering of this type can be described by a vector Q_α (the subscripts denoted by Greek letters take two values, because we treat the films as 2D objects). The nematic films have higher symmetry D_2, which corresponds to the 2D nematic phase. The order parameter of the nematic phase is the irreducible (traceless) symmetric tensor of the second rank $Q_{\alpha\beta}$. In the hexatic films (pulled from smectics-*B*), molecules are locally arranged in a triangular lattice, but the lattice is not an ideal one. The positional order does not extend over distances larger than several molecular sizes. Nevertheless, the bond order extends over macroscopic distances. The phase is therefore characterized by the D_{6h} point group symmetry, and hence, the order parameter for the case is the sixth-rank symmetric irreducible tensor $Q_{\alpha\beta\gamma\delta\mu\nu}$. In liquid crystalline films of all the types enumerated above, the order parameter Q has two independent components (e.g., Q_{xx} and Q_{xy} for the 2D nematics). We note that the order can be readily observed in the smectic-*C* or nematic films by looking for in-plane anisotropies in quantities such as the dielectric permeability tensor. Because of its intrinsic sixfold rotational symmetry, the hexatic orientational order is hardly observable. But it can be detected, e.g., as a sixfold pattern of spots in the in-plane monodomain *X*-ray structure factor, proportional to $Q_{\alpha\beta\gamma\delta\mu\nu}$(see, e.g., [3] and references therein).

In accordance with the Goldstone theorem, in films of all types with a broken rotational symmetry, the only degree of freedom of the order parameter that is relevant at large scales is an angle φ (like the phase of the order parameter for the superfluid ^{4}He). In hexatics, it is the bond angle, whereas, in 2D nematics and in smectic-*C* films, it is an angle related to the director. It is convenient to express a variation of the order parameter in terms of a variation of the angle φ. For the smectic-*C* films, the relation is

$$\delta Q_\alpha = -\delta\varphi\epsilon_{\alpha\mu}Q_\mu, \tag{2.1}$$

where $\epsilon_{\alpha\mu}$ is the two-dimensional antisymmetric tensor. For an orientational order with a higher symmetry, the relation has a similar form. For example, for hexatic films,

$$\delta Q_{\alpha\beta\gamma\delta\mu\nu} = -\delta\varphi\epsilon_{\alpha\rho}Q_{\rho\beta\gamma\delta\mu\nu} + \ldots, \tag{2.2}$$

where the dots represent the sum of all other possible combinations of the same structure. Therefore, for films of all types, the order parameter can be characterized by its absolute value $|Q|$ and the phase φ, which are traditionally represented as a complex quantity Ψ (see, e.g.,

[24]). The quantity is written as $\Psi = |Q|\exp(6i\varphi)$ for hexatic films, as $\Psi = |Q|\exp(2i\varphi)$ for 2D nematic films, and as $\Psi = |Q|\exp(i\varphi)$ for smectic-*C* films.

The angle φ should be included in the set of macroscopic variables of the films. A convenient starting point of the consideration is the energy density (per unit area) $\rho v^2/2 + \varepsilon$, where ρ is the 2D mass density, $\mathbf{v}$ is the film velocity, and ε is the internal energy density. The latter is a function of the mass density ρ, the specific entropy σ, and the angle φ. In fact, ε depends on $\nabla\varphi$, because any homogeneous shift of the angle φ does not affect the energy. For hexatic films, the leading terms of the energy expansion over gradients of φ are

$$\varepsilon = \varepsilon_0(\rho, \sigma) + \frac{K}{2}(\nabla\varphi)^2, \tag{2.3}$$

where *K* is the only (because of the hexagonal symmetry) orientational elastic module of the film. For low-symmetry films (2D nematic or smectic-*C* films), two orientational elastic modules are introduced, the longitudinal and transversal ones with respect to the specific in-plane direction (characterized by the so-called $\mathbf{c}$ director). But fluctuations of the director lead to a renormalization of the modules, and isotropization of the smectic-*C* or 2D nematic films [29] occurs at large scales. The same isotropic expression (2.3) for the elastic energy can therefore be used at large scales.

The complete dynamic equations for the freely suspended liquid-crystalline films, valid at scales larger than the film thickness, can be found in [30]. We consider a quasistationary motion of the disclination. Then, hard degrees of freedom are not excited. In other words, we can accept incompressibility and neglect bending deformations (which are suppressed by the presence of the surface tension in freely suspended films). Similarly, the thermodiffusive mode is not excited for the quasistationary disclination motion, which implies the isothermal condition. For freely suspended films, such effects as the substrate friction (relevant, e.g., for Langmuir films) are absent. In describing the disclination motion, we can therefore consider the system of equations for only the velocity $\mathbf{v}$ and the angle φ. The equations have to be formulated under the conditions $\rho = \text{const}$, $T = \text{const}$ (where *T* is the temperature), and $\nabla\mathbf{v} = 0$.

The equation for the velocity follows from the momentum density $\mathbf{j} = \rho\mathbf{v}$ conservation law,

$$\partial_t j_\alpha = -\nabla_\beta[T_{\alpha\beta} - \eta(\nabla_\alpha v_\beta + \nabla_\beta v_\alpha)], \tag{2.4}$$

where $T_{\alpha\beta}$ is the reactive (nondissipative) stress tensor and η is the 2D shear viscosity coefficient of the film. For two-dimensional hexatics, the reactive stress tensor is (see [30], Chapter 6)

$$\begin{aligned} T_{\alpha\beta} &= \rho v_\beta v_\alpha - \varsigma\delta_{\alpha\beta} + K\nabla_\alpha\varphi\nabla_\beta\varphi \\ &\quad - \frac{K}{2}\epsilon_{\alpha\gamma}\nabla_\gamma\nabla_\beta\varphi - \frac{K}{2}\epsilon_{\beta\gamma}\nabla_\gamma\nabla_\alpha\varphi, \end{aligned} \tag{2.5}$$

where $\varsigma = \varepsilon - \rho\partial\varepsilon/\partial\rho$ is the surface tension. We note that the ratio $K\rho/\eta^2$ is a dimensionless parameter that can be estimated by substituting 3D quantities instead of 2D ones (because all the 2D quantities can be estimated as the corresponding 3D quantities times the film thickness, and the latter drops from the ratio). For all known liquid crystals, the ratio is 10^{-3}–10^{-4}(see, e.g., [1–3, 31]) and can therefore be treated as a small parameter of the theory.

The second dynamic equation, the equation for the bond angle, is

$$\partial_t\varphi = \frac{1}{2}\epsilon_{\alpha\beta}\nabla_\alpha v_\beta - v_\alpha\nabla_\alpha\varphi + \frac{K}{\gamma}\nabla^2\varphi, \tag{2.6}$$

where γ is the so-called 2D rotational viscosity coefficient. We did not find the values of the coefficient γ for thin liquid-crystalline films in the literature. For bulk liquid crystals (see, e.g., [1–3, 31]), the 3D rotational viscosity coefficient is usually several times larger than the 3D shear viscosity coefficient. We can therefore expect that $\gamma > \eta$. But in order to span a wide range of possibilities, we treat the dimensionless ratio $\Gamma = \gamma/\eta$ as an arbitrary parameter in what follows.

If disclinations are present in the film, it is no longer possible to define a single-valued continuous bond-angle variable φ. But the order parameter is a well-defined function of coordinates that goes to zero at the disclination position. The gradient of $\varphi(t, \mathbf{r})$ is a single-valued function of $\mathbf{r}$ and is analytic everywhere except at an isolated point, the position of the disclination. The phase acquires a certain finite increment at each rotation around the disclination,

$$\oint dr_\alpha\nabla_\alpha\varphi = 2\pi s, \tag{2.7}$$

where the integration contour is a closed counterclockwise loop around the disclination position and *s* is the topological charge of the disclination: $s = (1/6)n$ for the hexatic ordering, $s = (1/2)n$ for the 2D nematic symmetry, and $s = n$ for the smectic-*C* films, where *n* is an integer. We can restrict ourselves to disclinations with the unitary charge $n = \pm 1$ only, because disclinations with larger $|s|$ possess a higher energy than the set of unitary disclinations with the same net topological charge, and defects with larger charges are therefore unstable with respect to the dissociation to the unitary ones. Therefore, disclinations with the charges $|n| > 1$ do not play an essential role in the physics of films [1–3, 31]. To write the expressions given below in a compact form, we keep the notation *s* for the topological charge, with the respective values $|s| = 1, 1/2, 1/6$ for the smectic-*C*, nematic, and hexatic films.

The static bond angle is determined by the stationary condition $\delta E/\delta\varphi = 0$, where

$$E = \int d^2r\left(\frac{\rho}{2}v^2 + \varepsilon\right)$$

is the energy of the film. For the energy density in Eq. (2.3), the condition is reduced to the Laplace equation $\nabla^2\varphi = 0$. For an isolated static disclination, there is a symmetric solution to this equation φ_0 that satisfies Eq. (2.7) and whose gradient is given by

$$\nabla_\alpha\varphi_0 = -s\epsilon_{\alpha\beta}\frac{r_\beta - R_\beta}{(\mathbf{r}-\mathbf{R})^2}, \tag{2.8}$$

where **R** is the position of the disclination. If the origin of the reference system is placed at this point, we can write $\varphi_0 = s\arctan(y/x)$, where x and y are coordinates of the observation point **r**. In dynamics, distribution (2.8) is disturbed as φ varies in time. It is also perturbed because of the presence of an angular distortion related to boundaries or other disclinations.

In what follows, we have in mind a case where a system of a large number of disclinations (with an uncompensated topological charge) is created. For 3D nematics, this can be done rather easily [1–3], because the energies of positive and negative defects are different due to the intrinsic elastic anisotropy. We are unaware of experimental or theoretical studies of defect nucleation mechanisms in free-standing films. Hopefully, the situation with a finite 2D density of defects can also be realized for films (for instance, the defects could even appear spontaneously as a mechanism to relieve frustrations in chiral smectic or hexatic films, similarly to the formation of the Abrikosov vortex lattice in superconductors [32]). Examining the motion of a disclination in this case, we investigate a vicinity of the disclination of the order of the interdisclination distance. Far from the disclination, the bond angle φ can then be written as const + **ur**, where **u** is much larger than the inverse interdisclination distance (because the number of disclinations is large). Near the disclination position, the bond angle φ can be approximated by expression (2.8). Our main problem is to establish a general coordinate dependence of φ and **v**, which, in particular, allows relating the bond (director) angle gradient **u** and the velocity of the disclination.

3. FLOW AND ANGULAR FIELDS AROUND A UNIFORMLY MOVING DISCLINATION

Here, we proceed to the main subject of our study, a single disclination driven by a large-scale inhomogeneity in the bond (director) angle φ. The disclination velocity is determined by an interplay of the hydrodynamic back-flow and the intrinsic dynamics of the angle φ. To find the disclination velocity, one has to solve the system of equations (2.4), (2.5), and (2.6) with constraint (2.7) ensuring a suitable asymptotic behavior. As we explained in the previous section, the angle φ is supposed to behave as const + **ur** at large distances from the disclination. We work in the reference system where the film as a whole is at rest. This means that the flow velocity excited by the disclination must tend to zero far from the disclination.

We consider the situation where the disclination moves with a constant velocity **V**. The angle φ and the flow velocity are then functions of $\mathbf{r} - \mathbf{V}t$ (where $\mathbf{R} = \mathbf{V}t$ is the disclination position). Equation (2.4) for the velocity can then be written as

$$\rho(V_\beta - v_\beta)\nabla_\beta v_\alpha + \eta\nabla^2 v_\alpha + \frac{K}{2}\epsilon_{\alpha\beta}\nabla_\beta\nabla^2\varphi - K\nabla_\alpha\varphi\nabla^2\varphi + \nabla_\alpha\left[\varsigma - \frac{K}{2}(\nabla\varphi)^2\right] = 0. \tag{3.1}$$

We can omit the first (inertial) term on the left-hand side of (3.1), which is small because of the smallness of the parameter $K\rho/\eta^2$. It then follows from Eqs. (2.4) and (2.5) that

$$\nabla^2 v_\alpha + \frac{K}{2\eta}\epsilon_{\alpha\beta}\nabla_\beta\nabla^2\varphi - \frac{K}{\eta}\nabla_\alpha\varphi\nabla^2\varphi + \nabla_\alpha\varpi = 0, \tag{3.2}$$

where $\varpi = \eta^{-1}[\varsigma - (K/2)(\nabla\varphi)^2]$. Under the same conditions, the equation for the angle φ following from Eq. (2.6) is

$$\nabla^2\varphi + \frac{\gamma}{K}V_\alpha\nabla_\alpha\varphi = \frac{\gamma}{K}v_\alpha\nabla_\alpha\varphi - \frac{\gamma}{2K}\epsilon_{\alpha\beta}\nabla_\alpha v_\beta. \tag{3.3}$$

We seek a solution characterized by the asymptotic behavior that the velocity **v** vanishes and $\nabla\varphi$ tends to a constant vector **u** as $r \longrightarrow \infty$. It is clear from the symmetry of the problem that the gradient **u** of the bond angle is directed along the Y axis if the velocity is directed along the X axis. Therefore, $\varphi \longrightarrow uy$ as $r \longrightarrow \infty$. Our problem is to find a relation between V and u, that is, between the disclination velocity and the bond angle gradient far from the disclination. There are two different regions: the region of large distances $r \gg u^{-1}$ and the region near the disclination $r \ll u^{-1}$. At large distances, corrections to the leading behavior $\varphi \approx uy$ are small and the problem can be treated in the linear approximation with respect to these corrections. In the region near the disclination, φ is close to static value (2.8) and the flow velocity **v** is close to the disclination velocity **V** (the special case where the ratio γ/η is extremely small is discussed in Subsection 4C). In what follows, these two regions are examined separately. The relation between u and V can be found by matching the asymptotics at $r \sim u^{-1}$. As a result, we obtain

$$V = \frac{K}{\eta}Cu, \tag{3.4}$$

where C is a dimensionless factor depending on the dimensionless ratio $\Gamma = \gamma/\eta$. This factor C is on the order of unity if $\Gamma \sim 1$. We are interested in the asymptotic behavior of C at small and large Γ.

A. The region near the disclination

We consider the region $r \ll u^{-1}$. Here, we can write

$$\varphi = \varphi_0(\mathbf{r}-\mathbf{R}) + \varphi_1(\mathbf{r}-\mathbf{R}), \tag{3.5}$$

where $\mathbf{R} = \mathbf{V}t$ is the disclination position, φ_0 is the static bond (director) angle with gradient (2.8), and φ_1 is a small correction to φ_0. The gradients of φ_0 are determined by Eq. (2.8).

Linearizing Eqs. (3.2) and (3.3) with respect to φ_1, we obtain

$$\eta\nabla^2 v_\alpha + \frac{K}{2}\epsilon_{\alpha\beta}\nabla_\beta\nabla^2\varphi_1 - K\nabla_\alpha\varphi_0\nabla^2\varphi_1 + \nabla_\alpha\left[\varsigma - \frac{K}{2}(\nabla\varphi)^2\right] = 0, \tag{3.6}$$

$$\nabla^2\varphi_1 - \frac{\gamma}{K}v_\alpha\nabla_\alpha\varphi_0 + \frac{\gamma}{2K}\epsilon_{\alpha\beta}\nabla_\alpha v_\beta = -\frac{\gamma}{K}V_\alpha\nabla_\alpha\varphi_0. \tag{3.7}$$

Introducing a new variable $\chi = (K/\eta)\nabla^2\varphi_1$, we rewrite Eqs. (3.6) and (3.7) as

$$\nabla^2 v_\alpha + \frac{1}{2}\epsilon_{\alpha\beta}\nabla_\beta\chi - \nabla_\alpha\varphi_0\chi + \nabla_\alpha\varpi = 0, \tag{3.8}$$

$$\chi - \Gamma v_\alpha\nabla_\alpha\varphi_0 + \frac{\Gamma}{2}\epsilon_{\alpha\beta}\nabla_\alpha v_\beta = -\Gamma V_\alpha\nabla_\alpha\varphi_0, \tag{3.9}$$

where $\Gamma = \gamma/\eta$, as above, and $\varpi = \eta^{-1}[\varsigma - K/2(\nabla\varphi)^2]$. It follows from Eq. (3.8) and $\nabla_\alpha v_\alpha = 0$ that $\nabla^2\varpi = \nabla_\alpha\varphi_0\nabla_\alpha\chi$. A solution of the system in Eqs. (3.8) and (3.9) can be written as

$$v_\alpha = V_\alpha + \epsilon_{\alpha\beta}\nabla_\beta\Omega, \tag{3.10}$$

where V_α is the obvious (because of the Galilean invariance) forced solution and the stream function Ω describes a zero mode of system (3.8) and (3.9). The system is homogeneous in r, and Ω is therefore a sum of contributions that are powerlike functions of r.

Taking the curl of Eq. (3.8), we obtain

$$-\nabla^4\Omega - \frac{1}{2}\nabla^2\chi - \epsilon_{\gamma\alpha}\nabla_\alpha\varphi_0\nabla_\gamma\chi = 0. \tag{3.11}$$

Substituting χ expressed in terms of $\mathbf{v}$ from Eq. (3.9) into Eq. (3.11) and using explicit expressions (2.8) for the derivatives of φ_0, we obtain

$$\left(1+\frac{\Gamma}{4}\right)\nabla^4\Omega + s\Gamma\left(\frac{2}{r^2}\partial_r^2\Omega - \frac{1}{r^2}\nabla^2\Omega - s\frac{1}{r^2}\partial_r^2\Omega + s\frac{1}{r^3}\partial_r\Omega\right) = 0 \tag{3.12}$$

in the polar coordinates (r, ϕ). Solutions to Eq. (3.12) are superpositions of the terms $\propto r^{\alpha+1}\exp(im\phi)$. Substituting this r, ϕ dependence into Eq. (3.12), we obtain an equation for α that has the roots

$$\alpha = \pm\frac{1}{\sqrt{2}}[2 + 2m^2 - s(1-s)\tilde{\Gamma} \pm \{(2+2m^2 - s(1-s)\tilde{\Gamma})^2 - 4s\tilde{\Gamma}(m^2 - 1 + s) - 4(m^2-1)^2\}^{1/2}]^{1/2}, \tag{3.13}$$

where $\tilde{\Gamma} = \Gamma(1+\Gamma/4)^{-1}$. Hence, $0 < \tilde{\Gamma} < 4$ for any γ and η. Evidently, all the roots in Eq. (3.13) are real. We emphasize that there is no solution $\alpha = 0$ (corresponding to a logarithmic behavior of the velocity in r) among the set (3.13). The first angular harmonic with $|m| = 1$ is of particular interest because $\varphi_1 = ur\sin\phi$ and $\Omega = -Vr\sin\phi$ far from the disclination. If Γ is small, there is a pair of small solutions among (3.13),

$$\alpha = \pm\alpha_1, \quad \alpha_1 = s\sqrt{\Gamma}/2, \tag{3.14}$$

for $m = \pm 1$. Otherwise, for any other relevant m, solutions (3.13) have no special smallness (terms with $m = 0$ are forbidden because of the symmetry).

We established that Ω is a superposition of the terms $\propto r^{\alpha+1}\exp(im\phi)$ with the exponents α determined by Eq. (3.13). The velocity can then be found from Eq. (3.10). To avoid a singularity in the velocity at small r, one should keep contributions with positive α only. In other words, the velocity field contains contributions with all powers α given by (3.13), but the factors at the terms with negative α are formed at $r \sim a$ (where a is the disclination core radius), and the corresponding contributions to the velocity are therefore negligible at $r \gg a$ (this statement must be clarified and refined for small negative exponents $-\alpha_1$ in the limit of small Γ; see Subsection 4C). We conclude that the correction to $\mathbf{V}$ in the flow velocity $\mathbf{v}$ related to Ω in Eq. (3.10) is negligible at $r \sim a$. We thus arrive at the nonslipping condition for the disclination motion: the disclination velocity $\mathbf{V}$ coincides with the flow velocity $\mathbf{v}$ at the disclination position.

Next, to find φ, one should solve the equation $(K/\eta)\nabla^2\varphi = \chi$, where χ is determined from Eq. (3.9). In addition to the part determined by the velocity, φ_1 can then involve zero modes of the Laplacian. The most dangerous zero mode is Uy, because it produces a nonzero momentum flux to the disclination core (and the Magnus force associated to it),

$$\oint dr_\alpha\epsilon_{\alpha\beta}T_{\beta\gamma} \sim KU. \tag{3.15}$$

But because of the condition $\alpha \neq 0$, all the contributions to the velocity correspond to zero viscous momentum flux to the origin. Consequently, it is impossible to compensate the Magnus force by other terms. The above reasoning leads us to the conclusion that the factor U (and therefore, the Magnus force) must be zero. Thus, φ_1 contains only terms proportional to $r^{\alpha+1}$ with

$\alpha > 0$. This conclusion is related to the fact that, for free-standing liquid-crystalline films, any distortion of the bond angle unavoidably produces hydrodynamic backflow motions (i.e., $\mathbf{v} \neq 0$). For liquid-crystalline films on substrates (Langmuir films), in contrast to free-standing films, hydrodynamic motions (backflows) are strongly suppressed by the substrate, and the situation where the backflow is irrelevant for the disclination motion can be realized.

B. The remote region

Let us consider the region $r \gg u^{-1}$, where we can write $\varphi = uy + \tilde{\varphi}$ and linearize the system of equations (3.2) and (3.3) with respect to $\tilde{\varphi}$. We then obtain the system of linear equations for $\mathbf{v}$ and $\tilde{\varphi}$,

$$\begin{aligned} \nabla^2 v_\alpha + \frac{K}{2\eta}(\epsilon_{\alpha\beta}\nabla_\beta\nabla^2\tilde{\varphi} - 2u_\alpha\nabla^2\tilde{\varphi}) + \nabla_\alpha\varpi &= 0, \\ (\nabla^2 + 2p\partial_x)\tilde{\varphi} + \frac{\gamma}{2K}(\epsilon_{\alpha\beta}\nabla_\alpha v_\beta - 2uv_y) &= 0, \end{aligned} \tag{3.16}$$

where $p = V\gamma/2K$. Taking the curl of the first equation and eliminating the Laplacian, we obtain

$$\epsilon_{\beta\alpha}\nabla_\beta v_\alpha = \frac{K}{2\eta}[(\nabla^2 + 2u\partial_x)\tilde{\varphi} + \Phi], \tag{3.17}$$

where Φ is a harmonic function. In terms of Φ, system (3.16) is reduced to

$$\left[\left(1 + \frac{\Gamma}{4}\right)\nabla^4 + 2p\nabla^2\partial_x - \Gamma u^2\partial_x^2\right]\tilde{\varphi} = \frac{\Gamma}{2}u\partial_x\Phi. \tag{3.18}$$

Equation (3.18) can be written as

$$(\nabla^2 + 2k_1\partial_x)(\nabla^2 - 2k_2\partial_x)\tilde{\varphi} = \frac{\tilde{\Gamma}}{2}u\partial_x\Phi, \tag{3.19}$$

$$k_{1,2} = \frac{1}{2(1 + \Gamma/4)}\left(\sqrt{p^2 + \Gamma\left(1 + \frac{\Gamma}{4}\right)u^2} \pm p\right). \tag{3.20}$$

The quantities k_1 and k_2 have the meaning of characteristic wave vectors. We conclude from Eq. (3.19) that zero modes of the operator on the left-hand side of the equation are proportional to

$$\exp(-k_1 r - k_1 x), \quad \exp(-k_2 r + k_2 x),$$

that is, they are exponentially small everywhere outside narrow angular regions near the X axis. The behavior of the zero modes inside the regions is powerlike in r. In addition, there is a contribution to $\tilde{\varphi}$ related to the harmonic function Φ. It contains a part that decays as a power of r (the leading term is $\propto r^{-1}$) at $r \gg u^{-1}$. This solution is examined in more detail in Appendix A.

4. DIFFERENT REGIMES GOVERNED BY Γ

The behavior of the velocity and the bond (director) angle fields around the moving disclination is sensitive to the ratio of the rotational and the shear viscosity coefficients $\Gamma = \gamma/\eta$. In this section, we examine different cases depending on the Γ value.

A. The case where $\Gamma \gtrsim 1$

We start analyzing different mobility regimes with the most probable case where $\Gamma \gtrsim 1$. If $\Gamma \sim 1$, then the factor C in Eq. (3.4) is on the order of 1 and $u \sim p$. It then follows from Eqs. (3.20) that $k_1, k_2 \sim u$. This is a manifestation of the fact that there is a unique characteristic scale in this case, given by u^{-1}. We can then estimate $\tilde{\varphi}$ by matching the solutions in the regions near the disclination and far from it at $r \sim u^{-1}$. We conclude that it is a function of the dimensionless parameter ur; the function is on the order of unity when its argument ur is on the order of unity.

For large Γ, there remains a unique characteristic scale u^{-1}, and consequently, $C \sim 1$ in this case. To prove this statement, we first treat small distances $r \ll u^{-1}$. As shown in Section 3A, the respective corrections φ_1 and $\delta\mathbf{v}$ to φ_0 and $\mathbf{V}$ are expanded in a series over the zero modes characterized by exponents (3.13). In particular, for $m = 1$, we can write $\varphi_1 \sim uy(ur)^\alpha$. In the large-$\Gamma$ limit, the exponents α given by (3.13) are regular because $\tilde{\Gamma} \longrightarrow 4$. From (3.13), we have $\alpha_1 \sim 1$, and in this case,

$$\chi \sim \frac{K}{\eta r^2}uy(ur)^{\alpha_1}.$$

Comparing Eqs. (3.8) and (3.9), we conclude that, for large Γ, the term involving χ can be omitted in Eq. (3.9), and the equation therefore becomes a constraint imposed on the velocity. Equation (3.8) then gives

$$|\delta v| \sim \frac{K}{\eta r}uy(ur)^{\alpha_1}.$$

The disclination velocity can now be found from the relation $V \sim |\delta v|$ at the scale u^{-1}, that is, $p \sim \Gamma u$, or $C \sim 1$. The complete analysis also covers the remote region. With the condition $p \sim \Gamma u$, it follows that $k_{1,2} \sim u^{-1}$. Using the procedure given in Appendix A, we can then prove that the solutions in the two regions can be matched at $r \sim u^{-1}$, and therefore, there are no new characteristic scales, indeed. We also note that the rotational viscosity γ drops from the hydrodynamic equations at large Γ. Although this is not true inside the disclination core (see Appendix D), the boundary conditions for $\mathbf{v}$ and φ on the core boundary reveal no dramatic changes in behavior. Consequently, it is the shear viscosity alone that determines the disclination mobility, which implies that $C \sim 1$.

We can therefore say that, in the limit as $\Gamma \longrightarrow \infty$, no additional features appear compared to $\Gamma \sim 1$. But this is not the case for small Γ, because $u \gg p$ for $\Gamma \ll 1$. We study this case in the next subsection.

B. Small Γ

Here, we consider the case where $\Gamma \ll 1$. This limit is physically attained at anomalously large η, with $K\rho/\eta^2$ still treated as the smallest dimensionless parameter. This justifies the use of the same equations (3.2) and (3.3) as in the previous subsections.

For $r \ll u^{-1}$, the analysis in Section A is correct. As we noted, the contributions to $\mathbf{v}$ and φ_1 related to the modes with negative α should not be taken into account there. For $\Gamma \ll 1$, the leading role is played by the mode with the smallest exponent ($\alpha_1 = s\sqrt{\Gamma}/2$), because the presence of modes with positive exponents $\alpha \sim 1$ would contradict the condition of smooth matching at $r \sim u^{-1}$. Strictly speaking, neglecting a small negative exponent $-\alpha_1$ is correct under the condition $\alpha_1|\ln(ua)| \gg 1$, where a is the core radius of the disclination. This is what is considered in this subsection. The opposite case, which we call the extremely small-Γ limit, is analyzed in Section 4C. At $r \ll u^{-1}$, we can therefore write

$$\varphi_1 \sim uy(ur)^{\alpha_1}, \quad V - v_x \sim \alpha_1 u\frac{K}{\gamma}(ur)^{\alpha_1}, \tag{4.1}$$

with the coefficient at $y(ur)^{\alpha_1}$ determined from matching at $r \sim u^{-1}$, where $\nabla\varphi \sim 1/r$. Similarly, matching $V - v_x \sim V$ at $r \sim u^{-1}$ gives $V \sim \alpha_1 uK/\gamma$. The relation can be rewritten as $p \sim \alpha_1 u \ll u$, and we therefore conclude that $C \sim 1/\sqrt{\Gamma}$.

In accordance with Eq. (3.20), the relation $p \sim \sqrt{\Gamma}\,u$ leads to $k_{1,2} \sim p \ll u$. In other words, a new scale p^{-1} (different from u^{-1}) appears in the problem. A detailed investigation of the remote region $r \gg u^{-1}$ is therefore needed to establish the $\mathbf{r}$ dependences of the both angle φ and the velocity field $\mathbf{v}$ there. This investigation can be based on the equations formulated in Section 3B, which are correct irrespective of the value of pr.

Explicit expressions describing the velocity and the angle are presented in Appendix A. They contain three dimensionless functions $\zeta_1(\nabla/u)$, $c_1(\nabla/u)$, and $c_2(\nabla/u)$. At $ur \gg 1$, only zero terms of the expansions of these functions in the Taylor series can be kept. Only one of these three coefficients is independent (see Eq. (A.10)). The general solution can therefore be expressed in terms of a single parameter, which we choose as $\zeta \equiv \zeta_1(0)$. The procedure corresponds to the following construction of the solutions to equations of motion (3.16) in the region $ur \gg 1$. We have to match the solutions in the outer and the inner regions (far from and close to the disclination respectively) at $ur \sim 1$. Technically, the matching is equivalent to the appropriate boundary conditions for the outer problem at $ur \sim 1$, and these boundary conditions can be formally replaced by the local source terms in the equations, acting at $ur \sim 1$. We can expand these sources in the standard multipolar series. We thus arrive at the expansion with respect to the gradients of the δ function. The gradients scale as u, and therefore, ζ, c_1, and c_2 are dimensionless functions of the dimensionless ratio ∇/u.

To find the asymptotic behavior of the angle φ and of the velocity $\mathbf{v}$, we first consider the region $u^{-1} \ll r \ll p^{-1}$. From Eqs. (A.4), (A.5), and (A.10), we then derive

$$v_x = \frac{K(2s-\zeta)}{\gamma u}k_1k_2\ln(pr), \tag{4.2}$$

where we keep only the leading logarithmic contribution of the zero harmonic in v_x. Matching the velocity derivatives determined by Eqs. (4.1) and (4.2) at $r \sim u^{-1}$, we find that $\zeta \sim 1$ (we imply that $s \sim 1$). Using Exps. (A.2), (A.5), and (A. 10), we obtain

$$\varphi = \varphi_0 + uy + spy\ln(pr) \tag{4.3}$$

in the region $u^{-1} \ll r \ll p^{-1}$. We see that there is only a small correction to the simple expression $\varphi_0 + uy$ in that region, because $p \ll u$.

In the region $pr \gg 1$, the expressions for the angle φ and the velocity $\mathbf{v}$ are more complicated. Using Eqs. (A.2)–(A.5), we obtain

$$\partial_x\varphi = -s\sqrt{\frac{\pi}{2}}[c_1\sqrt{k_1}\exp(-k_1r - k_1x) + c_2\sqrt{k_2}\exp(-k_2r + k_2x)]\frac{y}{r^{3/2}} - \frac{\zeta}{2}\frac{y}{r^2}, \tag{4.4}$$

$$\partial_y\varphi = u - 2s\sqrt{\frac{\pi}{2}}[c_1\sqrt{k_1}\exp(-k_1r - k_1x) - c_2\sqrt{k_2}\exp(-k_2r + k_2x)]\frac{y}{r^{1/2}} - \frac{\zeta}{2}\frac{x}{r^2}, \tag{4.5}$$

$$v_y = \frac{K}{\gamma u}\left\{2s\sqrt{\frac{\pi}{2}}[c_1k_2\sqrt{k_1}\exp(-k_1r - k_1x) - c_2k_1\sqrt{k_2}\exp(-k_2r + k_2x)]\frac{y}{r^{3/2}} - p\zeta\frac{y}{r^2}\right\}, \tag{4.6}$$

$$v_x = -\frac{K}{\gamma u}\left\{\sqrt{\frac{\pi}{2}}s\Gamma u^2\left[\frac{c_1}{\sqrt{k_1r}}\exp(-k_1r - k_1x) + \frac{c_2}{\sqrt{k_2r}}\exp(-k_2r + k_2x)\right] + p\zeta\frac{x}{r^2}\right\}, \tag{4.7}$$

where $c_1 \sim 1$ and $c_2 \sim 1$ are determined by Eq. (A.10) (we omitted the argument 0 to simplify the notation). Expressions (4.4)–(4.7) contain terms of two types, iso-

tropic and anisotropic ones. The anisotropic contributions are essential only in the narrow angular regions near the X axis, where they dominate. It is worth noting a very nontrivial structure of the flow, in which the isotropic flux to the origin is compensated by the anisotropic terms.

The expressions found in this subsection generalize the famous Lamb solution for the hydrodynamic flow around a hard cylinder (see, e.g., [33–35]), where the velocity field is exponentially small everywhere far from the cylinder except for the wake of the corps, i.e., in a very narrow angular sector ("tail"). Disclination motion in liquid-crystalline films can be regarded as the motion of a cylinder framed by a "soft" (i.e., deformable) orientational field φ. Because of the additional degree of freedom (compared to the classical Lamb problem), our solution has two tails around the moving disclination: a wake beyond the disclination and a precursor in front of it. In fact, both degrees of freedom (the flow velocity and the bond angle) are relevant.

C. Extremely small Γ

In the above analysis, we implied the condition $\alpha_1|\ln(ua)| \gg 1$ (we recall that $\alpha_1 = s\sqrt{\Gamma}/2$ at small Γ), imposing a restriction from below on Γ at a given u. If $\alpha_1|\ln(ua)| \ll 1$, the terms with both $\alpha = \pm\alpha_1$ determined by Eq. (3.14) must be taken into account near the disclination, which leads to a logarithmic behavior of the correction φ_1 to φ_0 in that region,

$$\varphi_1 \sim uy\ln\left(\frac{r}{a}\right)|\ln(au)|^{-1}, \tag{4.8}$$

instead of Eq. (4.1). Matching the derivatives of expressions (4.3) and (4.8) at $r \sim u^{-1}$ gives $p \sim u|\ln(au)|^{-1}$. In other words, $C \sim [\Gamma\ln(au)]^{-1}$. This case formally corresponds to the limit $\eta \longrightarrow \infty$ in our equations, where we can drop the backflow hydrodynamic velocity in the equation for the bond angle. The situation was examined in [6–9]. We present the simple analysis of the case in Appendix B. We also note that there is no crossover at $r \sim u^{-1}$ in the bond angle behavior in this situation.

We now clarify the question regarding the Magnus force in this case. In accordance with Eq. (4.8), the reactive momentum flux to the disclination core is

$$\oint dr_\alpha \epsilon_{\alpha\beta} T_{\beta\gamma} \sim Ku\ln\left(\frac{r}{a}\right)|\ln(au)|^{-1}.$$

The flux is therefore r-dependent, tending to zero as $r \longrightarrow a$. This reactive momentum flux is compensated by the viscous momentum flux (related to derivatives of the flow velocity $\mathbf{v}$), which is nonzero in this case because of the logarithmic behavior of the flow velocity in r near the disclination. The flow velocity can be found from Eqs. (3.6) and (4.8) as

$$v_\alpha \sim \frac{Ku}{\eta|\ln(au)|}\epsilon_{\alpha\beta}\nabla_\beta\left[y\ln^2\left(\frac{r}{a}\right)\right],$$

which is a generalization of the Stokes–Lamb solution [33, 34]. But unlike in the Lamb problem (a hard cylinder moving in a viscous liquid), $|\mathbf{V} - \mathbf{v}(\mathbf{r} = \mathbf{a})| \sim V$ in our case; i.e., we have a slipping of the core of the moving disclination. This slipping seems natural in the limit of extremely small values of Γ, corresponding to the limit $\eta \longrightarrow \infty$, that is, to a strongly suppressed hydrodynamic flow. Physically, this property implies that the disclination cannot be understood as a hard impenetrable object. It is also worth noting that the logarithmic behavior found above is similar to the general feature of two-dimensional hydrodynamic motion that comes from the well-known fact (see, e.g., [33–35]) that nonlinear terms cannot be neglected in a two-dimensional laminar flow even for a small Reynolds number; these terms become relevant for sufficiently large distances. But in our case, these nonlinear terms do not come from the convective hydrodynamic nonlinearity; they come from the terms in stress tensor (2.4) that are nonlinear in φ.

An explicit expression for φ and its asymptotic forms corresponding to the considered case are given in Appendix B. An expression for the flow velocity field induced by the disclination motion at extremely small Γ is derived in Appendix C.

5. CONCLUSIONS

We now summarize the results of our paper. To understand the physics underlying the freely suspended film dynamics, we studied the ground case—a single disclination motion in a thin hexatic, smectic-C, or nematic liquid-crystalline film, driven by an inhomogeneity in the bond (or director) angle. We investigated the uniform motion (the one with a constant velocity). In this case, we derived and solved the equations of motion and found the bond angle and hydrodynamic velocity distributions around the disclination. This allows us to relate the velocity of the disclination V to the bond angle gradient $u = |\nabla\varphi|$ in the region far from the disclination. So much effort is needed because the full set of equations must be solved everywhere, not only locally. We established the proportionality coefficient C (see Eq. (3.4)) in this nonlocal relationship; it has the meaning of an effective mobility coefficient. The coefficient C depends on the dimensionless ratio Γ of rotational (γ) and shear viscosity (η) coefficients.

There is little experimental knowledge of the values of the coefficients γ and η in liquid-crystalline films. It is generally believed that the corresponding values in a film (normalized by its thickness) and in a bulk material are not very different [31, 3], in which case we are in the regime of $\Gamma \sim 1$, where the coefficient C is on the

order of 1. But the case where $\Gamma \ll 1$ is not excluded from both theoretical and materials science standpoints. We found the coefficient $C \sim 1/\sqrt{\Gamma}$ in the small-Γ limit. We established a highly nontrivial behavior of the flow velocity and of the bond angle, which is powerlike in r near the disclination and extremely anisotropic far from it. Only for extremely small Γ, $\Gamma \ll 1/\ln^2(ua)$ (where a is the disclination core radius), did we find a logarithmic behavior $C \sim [\Gamma \ln(ua)]^{-1}$. The main message of our study is that the hydrodynamic motion (that is, the backflow), unavoidably accompanying any defect motion in liquid crystals, plays a significant role in the disclination mobility. Experimental evidence (see, e.g., the recent publication [36]) shows that this is indeed the case.

Our analysis can be applied to the motion of a disclination pair with the opposite topological charges. In this case, the role of the scale u^{-1} is played by the distance R between the disclinations. In accordance with Eq. (3.4), we then find that $\partial_t R \propto R^{-1}$ without a logarithm (provided the rotational viscosity coefficient γ is not anomalously small; see Section 4C for the quantitative criterion). This conclusion is confirmed by the results of numerical simulations for 2D nematics [15–18]. The authors of [15–18] consider the equations of motion in terms of the tensor order parameter, consistently taking the coupling between the disclination motion and the hydrodynamic flow into account. They simulated dynamics of the disclination pair annihilation and found that the distance R between the disclinations scales depends on time t as $t^{1/2}$, without logarithmic corrections (as follows from our theoretical analysis) for all values of the parameter Γ except extremely small ones. Unfortunately, we did not find in [16–18] the magnitudes of the shear viscosity that were used in the simulations. Lacking sufficient data on the values of γ and η, we can presently discuss only the general features of the disclination dynamics. For instance, the authors of [18] numerically found an asymmetry of the disclination dynamics with respect to the sign of the topological charge ($s = \pm 1/2$) in the one-constant approximation. In our approach, the asymmetry naturally appears from nonlinear terms in stress tensor (2.5) and from the first term on the right-hand side of Eq. (2.6) responsible for the different couplings of orientational and hydrodynamic flow patterns for positive and negative disclinations. This results in the fact that, for each m, the smaller positive exponents in Eq. (3.13) (corresponding to the minus in the brackets) are larger for $s = 1/2$ than for $s = -1/2$. The disclination with $s = 1/2$ therefore exerts a stronger influence on the flow velocity; this conclusion was qualitatively obtained in [18].

Although the theory presented in this paper is valid for free-standing liquid-crystalline films, the general scheme can be applied to liquid-crystalline films on solid or liquid substrates. Because such a film is arranged on the substrate surface, any of its hydrodynamic motions are accompanied by the substrate motion. For solid substrates, the situation where the hydrodynamic backflow is irrelevant for the disclination dynamics can therefore be realistic. In Section 4C (also see Appendix B), we examine this limit and reproduce the results in [6–9], where the hydrodynamic backflow was neglected from the very beginning. The case of the films on a liquid substrate requires a special investigation, but the approach and the main ideas of our paper could be useful there as well.

Our results can be directly tested by comparing with the experimental data for smectic-C or nematic films. The hexatic order parameter, which has a sixfold local symmetry, is not coupled to the light in any simple way (and, therefore, ideal hexatic disclinations are hardly observed in optics). But it is possible to observe the core splitting of the disclinations in tilted hexatic smectic films [26]. Indeed, because of discontinuity of the tilt direction (which is locked to the bond direction), the hexatic order and hexatic disclinations can be observed indirectly. The second possibility of detecting the defects of hexatic ordering and verifying our theoretical results is classical light scattering (where the wave vectors are $q = 10^2$–10^4 cm^{-1} and the frequency is $\omega \lesssim 10^8$ s^{-1} in typical experiments). For a reasonably thick film, the power spectrum of light scattering can have some additional structure revealing the disclination properties (e.g., defects are thought to be relevant to the very low frequency noise observed in thin films). Experimental studies of this type are highly desirable.

ACKNOWLEDGMENTS

The research described in this publication was made possible in part by the Russian Foundation for Basic Research (project no. 00-02-17785) and INTAS (grant no. 30-234). S.V.M. is grateful for support of this work by the Deutsche Forschungsgemeinschaft, grant no. KO 1391/4. Fruitful discussions with V.E. Zakharov, E.A. Kuznetzov, G.E. Volovik, and N.B. Kopnin are gratefully acknowledged.

APPENDIX A

Distances Far from the Disclination

Here, we derive some results for the region far from the disclination. These results are used in the case of small Γ considered in Section 4B.

We examine the harmonic function Φ in Eq. (3.17). Because the function is analytic in the region $r > u^{-1}$, it can be expanded in the derivatives of $\ln r$ there. Next, because of the symmetry of the problem, Φ is an antisymmetric function of y. At least one derivative ∂_y must therefore be present in each term of the expansion, that is,

$$\Phi = u\hat{\zeta}_1 \partial_y \ln r, \tag{A.1}$$

where $\hat{\zeta}_1 = \zeta_1(\nabla/u)$ and $\zeta_1(z)$ is a series in z converging in a circle with the radius on the order of 1. The expansion coefficients in the series $\zeta_1(\nabla/u)$ are determined by matching with the inner problem at $r \sim u^{-1}$.

Because of the symmetry, the angle $\tilde{\varphi}$ can be represented as

$$\partial_x\tilde{\varphi} = \partial_y B, \quad \partial_y\tilde{\varphi} = -(H + \partial_x B), \qquad \nabla^2 B + \partial_x H = 0. \tag{A.2}$$

The latter equation is the condition $\epsilon_{\alpha\beta}\nabla_\alpha\nabla_\beta\tilde{\varphi} = 0$. We note that $\nabla^2\tilde{\varphi} = -\partial_y H$. In the region far from the disclination, we can use Eqs. (3.16) and (3.17). The incompressibility condition $\nabla_\alpha v_\alpha = 0$ must also be taken into account. We thus obtain expressions for the velocity in terms of B and H,

$$\mathrm{curl}\,\mathbf{v} = \frac{K}{2\eta}\partial_y[-H + 2uB + u\hat{\zeta}_1\ln(pr)],$$

$$v_y = \frac{K}{\gamma u}\partial_y \times\left\{-H + 2pB + \frac{\Gamma}{4}[-H + 2uB + u\hat{\zeta}_1\ln(pr)]\right\}, \tag{A.3}$$

$$v_x = \frac{K}{\gamma u}\partial_x \times\left\{-H + 2pB + \frac{\Gamma}{4}[-H + 2uB + u\hat{\zeta}_1\ln(pr)]\right\} - \frac{K}{2\eta}[-H + 2uB + u\hat{\zeta}_1\ln(pr)]. \tag{A.4}$$

Solutions to Eq. (3.18) imply that

$$B = s[\hat{c}_1 K_0(k_1 r)e^{-k_1 x} + \hat{c}_2 K_0(k_2 r)e^{k_2 x}] - \frac{1}{2}\hat{\zeta}_1\ln(pr), \qquad H = 2s[k_1\hat{c}_1 K_0(k_1 r)e^{-k_1 x} - k_2\hat{c}_2 K_0(k_2 r)e^{k_2 x}]. \tag{A.5}$$

Here, the particular representation in Eq. (A.1) is used and an arbitrary function of y that can contribute to H is chosen to be zero because $\nabla\tilde{\varphi} \longrightarrow 0$ (and, hence, $H \longrightarrow 0$) as $r \longrightarrow \infty$. In (A.5), $\hat{c}_1$ and $\hat{c}_2$ are dimensionless differential operators that can be represented as Taylor series in ∇/u, i.e., $c_1(\nabla/u)$ and $c_2(\nabla/u)$. These functions must scale with u because the functions must be found from matching at $r \sim u^{-1}$.

Additionally, there are two conditions for the variables in the region $ur \gg 1$. First, the correct circulation around the origin leads to the effective δ-functional term in Eq. (A.2),

$$\nabla^2 B + \partial_x H = -2\pi s\delta(\mathbf{r}). \tag{A.6}$$

The second condition is the absence of the flux to the origin,

$$\int d\phi\, v_r(r, \phi) = 0. \tag{A.7}$$

Relations (A.6) and (A.7) lead to the conditions

$$c_1(0) + c_2(0) + \frac{\zeta_1(0)}{2s} = 1, \tag{A.8}$$

$$\left(1 + \frac{\Gamma}{4}\right)[k_1 c_1(0) - k_2 c_2(0)] - \left(p + \frac{\Gamma u}{4}\right) \times\left[c_1(0) + c_2(0) + \frac{\zeta_1(0)}{2s}\right] + \frac{\Gamma u}{8s}\zeta_1(0) = 0. \tag{A.9}$$

At small Γ, the solution to Eqs. (A.8) and (A.9) is

$$\zeta_1(0) = \zeta, \quad c_1(0) = \frac{k_1 - \zeta k_2/2s}{k_1 + k_2}, \qquad c_2(0) = \frac{k_2 - \zeta k_1/2s}{k_1 + k_2}. \tag{A.10}$$

We also assumed that $\zeta \lesssim 1$, which is justified in Subsection 4B.

APPENDIX B

Suppressed Flow

Here, we demonstrate how the disclination velocity V can be found if the hydrodynamic velocity $\mathbf{v}$ is negligible (e.g., because of substrate friction). We reproduce the results in [6–9].

In the absence of the hydrodynamic flow, the equation for the angle φ is purely diffusive,

$$\gamma\partial_t\varphi = K\nabla^2\varphi, \tag{B.1}$$

as follows from Eq. (1.6) with $\mathbf{v} = 0$. We assume that $\varphi \longrightarrow uy$ as $r \longrightarrow \infty$. The disclination motion is forced by the "external field" u. We seek a solution $\varphi(t, x, y) = \varphi(x - Vt, y)$. From Eq. (B.1), we then obtain

$$2p\partial_x\varphi + \nabla^2\varphi = 0, \quad \text{where} \quad 2p = \gamma V/K. \tag{B.2}$$

In what follows, we consider the solution corresponding to a single disclination with the circulation

$$\oint d\mathbf{r}\nabla\varphi = 2\pi s, \tag{B.3}$$

where the integral is taken along a contour encompassing the disclination counterclockwise. The quantity s in Eq. (B.3) is an arbitrary parameter (which is equal to $\pm 1/6$ for hexatic, $\pm 1/2$ for nematic, and ± 1 for smectic-C ordering). For a suitable solution to Eq. (B.2) corresponding to Eq. (B.3), we have

$$\partial_x\varphi = s\partial_y\int\frac{d^2q}{2\pi}\frac{1}{q^2 - 2ipq_x}\exp(i\mathbf{q}\cdot\mathbf{r}) = s\exp(-px)\partial_y K_0(pr). \tag{B.4}$$

This derivative tends to zero as $r \longrightarrow \infty$, as it should be.

Expression (B.4) does not determine φ unambiguously because $\partial_x(uy) = 0$, and we can therefore obtain a new solution by adding a term uy to a given solution. We note that uy is the zero mode of the Eq. (B.2). The solution can therefore be written as

$$\varphi = \tilde{\varphi}_L + uy,$$
$$\tilde{\varphi}_L(x', y) = -s\int_{x'}^{\infty} dx \exp(-px)\partial_y K_0(pr), \tag{B.5}$$

where $\tilde{\varphi}_L$ tends to zero as $r \longrightarrow \infty$. To relate p and u in Eq. (B.5), we must know the boundary conditions at $r \longrightarrow 0$, or, in fact, at $r \sim a$, where a is the core radius. At small r, the angle φ can be written as a series $\varphi = \varphi_0 + \varphi_1 + \ldots$, where φ_0 corresponds to the static disclination and φ_1 is the first correction to φ_0 related to the motion. Matching with the inner problem gives

$$\nabla\varphi_1(a) \sim p, \tag{B.6}$$

because the solution for the order parameter inside the core is an analytic function of r/a and the expansion in p is a regular expansion in pa (see [7] and Appendix D).

Expanding Eq. (B.4) in p, we obtain

$$\frac{1}{s}\partial_x\varphi \approx -\frac{y}{r^2} + \frac{pxy}{r^2}$$

at $pr \ll 1$. In accordance with Eq. (B.5), we then obtain with logarithmic accuracy (i.e., in the main approximation in $|\ln(pa)| \gg 1$) that

$$\varphi_1 = spy\ln(pr) + uy. \tag{B.7}$$

Using boundary condition (B.6), we now obtain

$$u = sp\ln\left(\frac{1}{pa}\right) \tag{B.8}$$

with the same logarithmic accuracy. This can be rewritten as

$$V = \frac{2Ku}{s\gamma\ln(1/pa)}. \tag{B.9}$$

The same answer (B.9) can be found from the energy dissipation balance. First of all, we can find the energy E corresponding to solution (B.5),

$$E = \int d^2r\frac{K}{2}(\nabla\varphi)^2$$
$$= K\int d^2r\left[\frac{1}{2}u^2 + \frac{1}{2}(\nabla\tilde{\varphi}_L)^2 + u\partial_y\tilde{\varphi}_L\right], \tag{B.10}$$

where the first term is the energy of the external field, the second term represents the energy of the disclination itself, and the third term is the coupling energy. Obviously, only the last cross-term depends on time. For $|x - Vt| \gg p^{-1}$,

$$\int_{-\infty}^{\infty} dy\partial_y\tilde{\varphi}_L = \begin{cases} 0 & \text{if } x > Vt, \\ -2\pi s & \text{if } x < Vt. \end{cases}$$

It then follows from Eq. (B.10) that

$$\partial_t E = -2\pi sKuV. \tag{B.11}$$

On the other hand, we can use Eq. (B.1) to obtain

$$\partial_t E = -\frac{K^2}{\gamma}\int d^2r(\nabla^2\varphi)^2. \tag{B.12}$$

Replacing $\nabla^2\varphi$ with $2p\partial_x\varphi$ here in accordance with Eq. (B.2), we obtain

$$\partial_t E = -\gamma V^2\int d^2r(\partial_x\varphi)^2.$$

The main logarithmic contribution to the integral comes from the region $a < r < p^{-1}$, where $\partial_x\varphi \approx -sy/r^2$. We thus obtain

$$\partial_t E = -\pi s^2\gamma V^2\ln\left(\frac{1}{pa}\right). \tag{B.13}$$

Comparing the expression with Eq. (B.11), we find the same answer (B.9).

APPENDIX C

Extremely Small Γ

Here, we consider the flow velocity induced by the moving disclination for extremely small Γ. The velocity is zero in the zero approximation in Γ (this case is considered in Appendix B), and we therefore examine the next, first-order, approximation in Γ. We use the same formalism and the same notation as in Appendix A.

In accordance with Appendix A, solutions to the complete set of nonlinear stationary equations can be represented as

$$\partial_x\tilde{\varphi} = \partial_y B, \quad \partial_y\tilde{\varphi} = -(H + \partial_x B), \tag{C.1}$$

$$\mathrm{curl}\mathbf{v} = \frac{K}{2\eta}[-\partial_y H + 2u\partial_y B + 2us\partial_y\ln r + \Phi'], \tag{C.2}$$

$$v_x = -\frac{K}{2\eta}\partial_y\nabla^{-2} \times [-\partial_y H + 2u\partial_y B + 2us\partial_y\ln r + \Phi'], \tag{C.3}$$

$$v_y = \frac{K}{2\eta}\partial_x\nabla^{-2} \times [-\partial_y H + 2u\partial_y B + 2us\partial_y\ln r + \Phi'], \tag{C.4}$$

where B, H, and Φ' are to be found from the equations

$$-\partial_y H + 2p\partial_y B + \frac{\Gamma}{4}\nabla^{-2}(\nabla^2 - 2u\partial_x) \times [-\partial_y H + 2u\partial_y B + 2us\partial_y\ln r + \Phi'] \tag{C.5}$$
$$= -\frac{\Gamma}{2}(\partial_y\nabla^{-2}[-\partial_y H + 2u\partial_y B + 2us\partial_y\ln r + \Phi']\partial_y B + \partial_x\nabla^{-2}[-\partial_y H + 2u\partial_y B + 2us\partial_y\ln r + \Phi'](\partial_x B + H)),$$

$$\Phi' = 2\nabla^{-2}[(\partial_x B + H)\partial_x\partial_y H + \partial_y B\partial_y^2 H], \quad \text{(C.6)}$$

$$\nabla^2 B + \partial_x H = -2\pi s\delta(\mathbf{r}). \quad \text{(C.7)}$$

If Γ is extremely small, $s^2\Gamma\ln^2(ua) \ll 1$, the solution to Eqs. (C.5)–(C.7) can be continued to the vicinity of the core. In the leading approximation, the solution for $\tilde{\varphi}$ coincides with the solution for the angle $\tilde{\varphi}_L$ in the absence of the backflow. This case, examined in [6–9], is described in Appendix B. The functions B_L and H_L corresponding to $\tilde{\varphi}_L$ are given by

$$2pB_L = H_L = 2spK_0(pr)\exp(-px). \quad \text{(C.8)}$$

This solution gives

$$\Phi' = 2s^2 p\frac{y}{r^2}\ln\left(\frac{\min\{r, p^{-1}\}}{a}\right). \quad \text{(C.9)}$$

Neglecting the nonlinear right-hand side of Eq. (C.5), we can then find

$$H(\mathbf{r}) = \frac{4\pi s}{1+\Gamma/4}\int\frac{d^2q}{(2\pi)^2}\exp(i\mathbf{q}\cdot\mathbf{r}) \quad \text{(C.10)}$$

$$\times\frac{pq^2 - (s\Gamma p/4)(q^2 + 2iuq_x)\ln(\min\{(qa)^{-1}, (pa)^{-1}\})}{(q^2 - 2ik_1q_x)(q^2 + 2ik_2q_x)}.$$

$B(\mathbf{r})$ can be found similarly. Using B and H in Eqs. (C.3) and (C.4), we calculate the flow velocity $\mathbf{v}(\mathbf{r})$ that vanishes at infinity.

For $r \gg p^{-1}$, this solution coincides with expressions (A.5), (A.8), and (A.9) with

$$\zeta_1(0) = 2s + \frac{2s^2 p}{u}\ln\left(\frac{1}{pa}\right).$$

For $pr \ll 1$, expression (C.10) is reduced to (C.8) and this region produces the main contribution to Φ' in (C.9). The following expressions are obtained in the inner region ($pr \ll 1$) from the solution in Eqs. (C.1)–(C.10):

$$\varphi_1 = \left(u - sp\ln\frac{1}{pa}\right)y + spy\ln\frac{r}{a}, \quad \text{(C.11)}$$

$$\mathrm{curl}\,\mathbf{v} = \frac{Ks^2 p}{\eta}\frac{y}{r^2}\ln\frac{r}{a}. \quad \text{(C.12)}$$

A relation between p and u is fixed by condition (B.6), leading to $u = sp\ln[1/(pa)]$, which is equivalent to Eq. (B.9). The flow velocity at $pr \ll 1$ and $\ln(r/a) \gg 1$ is

$$v_\alpha = -\frac{s^2\Gamma}{8}V\epsilon_{\alpha\beta}\nabla_\beta\left[y\ln^2\left(\frac{r}{a}\right)\right], \quad \text{(C.13)}$$

which corresponds to the stream function

$$\Omega = -Vy - \frac{Ks^2 p}{4\eta}y\ln^2\left(\frac{r}{a}\right). \quad \text{(C.14)}$$

The expansion with respect to Γ near the disclination is regular and can be derived from Eqs. (3.8) and (3.9) with the condition $\nabla\varphi_1(a) \sim p$: $\tilde{\varphi}_L + uy$ is the zero term of the series for φ, and expression (C.14) represents the zero and the first terms for Ω.

We note that, in accordance with Eq. (C.13) in the limit as $\Gamma \longrightarrow 0$, the flow velocity tends to zero near the disclination core, $v(a)/V = O(\Gamma)$, despite the fact that the disclination itself moves with the finite velocity V; thus, there is a slipping on the disclination core in this limit.

APPENDIX D

Solution with the Complete Order Parameter

Here, we consider the dynamic equations for the coupled velocity field $\mathbf{v}$ and the complete order parameter $\Psi = Q\exp(i\varphi/|s|)$ describing the 2D orientational order in liquid-crystalline films. These equations are needed to examine the velocity field close to the disclination position. We assume that the core size a is larger than characteristic molecular scales and work in the framework of the mean field theory.

Formally, the equations can be derived using the Poisson bracket method [30, 37]. In the mean field approximation, the energy associated with the order parameter is

$$\mathcal{H}_\Psi = \frac{Ks^2}{2}\int d^2r\left(|\nabla\Psi|^2 + \frac{1}{2a^2}(1-|\Psi|^2)^2\right),$$

its density becomes the K contribution in Eq. (2.3) at large scales $r \gg a$. The only nontrivial Poisson bracket that must be added to the standard expressions is [28]

$$\{j_\alpha(r_1), \Psi(r_2)\} = -\nabla_\alpha\Psi\delta(r_1 - r_2) + \frac{i}{2|s|}\Psi(r_2)\epsilon_{\alpha\beta}\nabla_\beta\delta(r_1 - r_2).$$

To be specific, we use the expressions for the energy and the Poisson bracket for hexatic films. The dynamic equations are given by

$$\rho\partial_t v_\alpha + \rho v_\beta\nabla_\beta v_\alpha = \eta\nabla^2 v_\alpha - \frac{s^2 K}{2} \times\left\{\nabla_\alpha\Psi^*\left(\nabla^2\Psi + \frac{1}{a^2}\Psi(1-|\Psi|^2)\right) + \nabla_\alpha\Psi\left(\nabla^2\Psi^* + \frac{1}{a^2}\Psi^*(1-|\Psi|^2)\right)\right\} - \frac{i|s|K}{4}\epsilon_{\alpha\beta}\nabla_\beta\{\Psi^*\nabla^2\Psi - \Psi\nabla^2\Psi^*\} + \nabla_\alpha\tilde{\zeta}, \quad \text{(D.1)}$$

$$\partial_t\Psi + v_\alpha\nabla_\alpha\Psi = \frac{i}{2|s|}\Psi\epsilon_{\alpha\beta}\nabla_\alpha v_\beta + \frac{Ks^2}{2\gamma_s}\left(\nabla^2\Psi + \frac{1}{a^2}\Psi(1-|\Psi|^2)\right),$$

the relation $\gamma_s = s^2\gamma/2$ ensures the reduction to Eq. (2.6) in the limit $|\Psi| = 1$, and the kinetic coefficients are believed to be independent of Q (otherwise, we can assume, for example, the dependence $\gamma_s = s^2\gamma|\Psi|^2/2$). The slow dynamics of a 2D liquid-crystalline system with disclinations can be described by Eqs. (D.1) with the additional incompressibility condition $\nabla\mathbf{v} = 0$ that allows excluding the passive variable $\tilde{\xi}$.

If the distance from the disclination to a boundary or other disclinations is much larger than a (i.e., the perturbation of the static solution $\Psi_0 = Q_0\exp(i\varphi_0/|s|)$ for a single defect is small), we can linearize Eqs. (D.1) with respect to the perturbation expressed in terms of the respective corrections Q_1 and φ_1 to Q_0 and φ_0,

$$\eta\nabla^2 v_\alpha - 2\gamma_s\Big\{\nabla_\alpha Q_0(v_\beta - V_\beta)\nabla_\beta Q_0 + \frac{1}{s^2}Q_0^2\nabla_\alpha\varphi_0\Big((v_\beta - V_\beta)\nabla_\beta\varphi_0 - \frac{1}{2}\epsilon_{\beta\gamma}\nabla_\beta v_\gamma\Big)\Big\} \quad \text{(D.2)}$$

$$+ 2\gamma_s\frac{1}{2s^2}\epsilon_{\alpha\beta}\nabla_\beta\Big[Q_0^2\Big((v_\mu - V_\mu)\nabla_\mu\varphi_0 - \frac{1}{2}\epsilon_{\mu\nu}\nabla_\mu v_\nu\Big)\Big] + \nabla_\alpha\tilde{\xi} = 0,$$

$$\frac{Ks^2}{2\gamma_s}\Big(\nabla^2 Q_1 - \frac{(\nabla\varphi_0)^2}{s^2}Q_1 - \frac{2\nabla_\alpha\varphi_1\nabla_\alpha\varphi_0}{s^2}Q_0 + \frac{1}{a_s^2}(1 - 3Q_0^2)Q_1\Big) = (v_\beta - V_\beta)\nabla_\beta Q_0, \quad \text{(D.3)}$$

$$\frac{Ks^2}{2\gamma_s}(\nabla^2\varphi_1 + 2Q_0^{-1}(\nabla_\alpha Q_1\nabla_\alpha\varphi_0 + \nabla_\alpha Q_0\nabla_\alpha\varphi_1)) = -\frac{1}{2}\epsilon_{\alpha\beta}\nabla_\alpha v_\beta + (v_\beta - V_\beta)\nabla_\beta\varphi_0. \quad \text{(D.4)}$$

In terms of the dimensionless quantities $L = \eta\Omega/K$, $R = r/a$, and $\Gamma = 2\gamma_s/(s^2\eta)$, Eq. (D.2) becomes (as previously, we consider a disclination with the unitary topological charge $|s|$ or $-|s|$)

$$\nabla_R^4 L + \frac{\Gamma}{4}\Big\{-4s^2\frac{(\partial_R Q_0)^2}{R^2}\partial_\phi^2 L + \Big(\nabla_R^2 + \frac{2s}{R}\partial_R\Big)\Big(Q_0^2\Big(\nabla_R^2 - \frac{2s}{R}\partial_R\Big)L\Big)\Big\} = 0, \quad \text{(D.5)}$$

where $\nabla_R^2 \equiv \partial_R^2 + \frac{1}{R}\partial_R + \frac{1}{R^2}\partial_\phi^2$ and Q_0 is found from

$$\Big(\partial_R^2 + \frac{1}{R}\partial_R - \frac{1}{R^2}\Big)Q_0 + Q_0(1 - Q_0^2) = 0,$$

$$Q_0(0) = 0, \quad Q_0(\infty) = 1.$$

If $\Gamma \gg 1$, as follows from Eq. (D.5), a new scale $R \sim 1/\sqrt{\Gamma} \ll 1$ appears inside the core, the first term in Eq. (D.5) can be neglected at larger scales, and there is no crossover at $R \sim 1$.

If $Q_0 \equiv 1$, Eq. (D.5) is reduced to Eq. (3.12). If $R \ll 1$, $Q_0 = AR$ ($A \approx 0.58$) and Eq. (D.5) can be rewritten as

$$\nabla_R^2\Big\{\nabla_R^2 L + \frac{A\Gamma}{4}(R^2\nabla_R^2 - 4s^2)L\Big\} = 0.$$

The solution to the equation is a superposition of the terms $\lambda(R)\sin(m\phi)$ with different m. After imposing the condition $\lambda(R) = 0$, two constants remain in the general solution of the ordinary differential equation for $\lambda(R)$; two partial solutions that are regular near $R = 0$ are given by

$$R^{|m|} \text{ and } R^{|m|}{}_2F_1\Big(\frac{|m| - \sqrt{m^2 + 4s^2}}{2}, \frac{|m| + \sqrt{m^2 + 4s^2}}{2}, 1 + |m|, -\frac{A^2\Gamma R^2}{4}\Big),$$

where ${}_2F_1$ is the hypergeometric function (${}_2F_1(a, b, c, z) = 1 + abz/c + \ldots$). Two constants (e.g., the derivatives $\lambda^{(|m|)}(0)$ and $\lambda^{(|m|+2)}(0)$) are chosen to ensure the slowest possible growth at $R \gg 1$ in order to eliminate the largest exponent among α in Eq. (3.13).

If $\Gamma \gg 1$, it is possible to derive a better approximation in the core region. We can expand $Q_0(R)$ in a series, seek a series solution $\lambda(R)$, and extract the terms of the highest order in Γ. For example, for $m = 1$, the series for $\lambda(R)$ begins with $l_1R + l_3R^3$, which fixes two constants in the partial solution,

$$\lambda(R) = l_1R\Big[1 + \frac{1}{A^2\Gamma s(2 - s^2)}\Big(-1 - \frac{s^2A^2\Gamma R^2}{8} + {}_2F_1\Big(\frac{1 - \sqrt{1 + 4s^2}}{2}, \frac{1 + \sqrt{1 + 4s^2}}{2}, 2, -\frac{A^2\Gamma R^2}{4}\Big)\Big)\Big] + l_3\frac{8}{A^2\Gamma s^2}R\Big[-1 + {}_2F_1\Big(\frac{1 - \sqrt{1 + 4s^2}}{2}, \frac{1 + \sqrt{1 + 4s^2}}{2}, 2, -\frac{A^2\Gamma R^2}{4}\Big)\Big].$$

The solutions to Eqs. (D.3) and (D.4) are given by

$$Q_1 = \vartheta(R)\partial_\phi\sin(m\phi), \quad \varphi_1 = \sigma(R)\sin(m\phi),$$

where ϑ and σ must be found from the equations

$$\vartheta'' + \frac{1}{R}\vartheta' - \frac{1+m^2}{R^2}\vartheta - \frac{2Q_0}{sR^2}\sigma + (1-3Q_0^2)\vartheta = \Gamma\frac{1}{R}\partial_R Q_0 \lambda,$$

$$\sigma'' + \frac{1}{R}\sigma' - \frac{m^2}{R^2}\sigma + \frac{2}{Q_0}\left(-\frac{sm^2}{R^2}\vartheta + \partial_R Q_0 \sigma'\right) = \frac{\Gamma}{2}\left(\lambda'' + \frac{1-2s}{R}\lambda' - \frac{m^2}{R^2}\lambda\right)$$

that generalize the expressions given in [7].

The dynamic equations with the complex order parameter demonstrate that, for all Γ, the boundary conditions for Eqs. (2.4)–(2.6) experience no significant changes on the core. The peculiarity of extremely small Γ leading to the nonslipping condition consists in a slow growth of $\nabla\Omega$ far from the disclination.

REFERENCES

1. P. G. de Gennes and J. Prost, *The Physics of Liquid Crystals* (Clarendon, Oxford, 1995).
2. S. Chandrasekhar, *Liquid Crystals* (Cambridge Univ. Press, Cambridge, 1992).
3. P. Oswald and P. Pieranski, *Les cristaux liquides* (Gordon and Breach, Paris, 2000), Vol. 1.
4. E. B. Sonin, Phys. Rev. B **55**, 485 (1997).
5. H. Pleiner, Phys. Rev. A **37**, 3986 (1988).
6. E. Bodenschatz, W. Pesch, and L. Kramer, Physica D (Amsterdam) **32**, 135 (1988).
7. L. M. Pismen and J. D. Rodriguez, Phys. Rev. A **42**, 2471 (1990).
8. G. Ryskin and M. Kremenetsky, Phys. Rev. Lett. **67**, 1574 (1991).
9. C. Denniston, Phys. Rev. B **54**, 6272 (1996).
10. R. Pindak, C. Y. Young, R. B. Meyer, and N. A. Clark, Phys. Rev. Lett. **45**, 1193 (1980).
11. R. Loft and T. A. DeGrand, Phys. Rev. B **35**, 8528 (1987).
12. P. E. Cladis, W. van Saarloos, P. L. Finn, and A. R. Kortan, Phys. Rev. Lett. **58**, 222 (1987).
13. A. Pargellis, N. Turok, and B. Yurke, Phys. Rev. Lett. **67**, 1570 (1991).
14. B. Yurke, A. N. Pargellis, T. Kovacs, and D. A. Huse, Phys. Rev. E **47**, 1525 (1993).
15. J.-I. Fukuda, Eur. Phys. J. B **1**, 173 (1998).
16. C. Denniston, E. Orlandini, and J. M. Yeomans, Europhys. Lett. **52**, 481 (2000).
17. C. Denniston, E. Orlandini, and J. M. Yeomans, Phys. Rev. E **64**, 021701 (2001).
18. G. Tóth, C. Denniston, and J. M. Yeomans, Phys. Rev. Lett. **88**, 105504 (2002).
19. V. L. Berezinskiĭ, Zh. Éksp. Teor. Fiz. **59**, 907 (1970) [Sov. Phys. JETP **32**, 493 (1970)].
20. J. M. Kosterlitz and D. J. Thouless, J. Phys. C **6**, 1181 (1973).
21. G. Blatter, M. V. Feigelman, V. B. Geshkenbein, *et al.*, Rev. Mod. Phys. **66**, 1125 (1994).
22. H. E. Hall and W. F. Vinen, Proc. R. Soc. London, Ser. A **238**, 204 (1956).
23. E. B. Sonin, Rev. Mod. Phys. **59**, 87 (1987).
24. A. Zippelius, B. I. Halperin, and D. R. Nelson, Phys. Rev. B **22**, 2514 (1980).
25. D. H. van Winkle and N. A. Clark, Bull. Am. Phys. Soc. **30**, 379 (1985).
26. S. B. Dierker, R. Pindak, and R. B. Meyer, Phys. Rev. Lett. **56**, 1819 (1986).
27. C. D. Muzny and N. A. Clark, Phys. Rev. Lett. **68**, 804 (1992).
28. E. V. Gurovich, E. I. Kats, and V. V. Lebedev, Zh. Éksp. Teor. Fiz. **100**, 855 (1991) [Sov. Phys. JETP **73**, 473 (1991)].
29. D. R. Nelson and R. A. Pelcovits, Phys. Rev. B **16**, 2191 (1977).
30. E. I. Kats and V. V. Lebedev, *Fluctuational Effects in the Dynamics of Liquid Crystals* (Springer-Verlag, New York, 1993).
31. L. M. Blinov and V. G. Chigrinov, *Electrooptic Effects in Liquid Crystal Materials* (Springer-Verlag, New York, 1996).
32. R. D. Kamien and J. V. Selinger, J. Phys.: Condens. Matter **13**, R1 (2001).
33. G. K. Batchelor, *An Introduction to Fluid Dynamics* (Cambridge Univ. Press, Cambridge, 1970; Mir, Moscow, 1973).
34. H. Lamb, *Hydrodynamics* (Cambridge Univ. Press, Cambridge, 1975).
35. L. D. Landau and E. M. Lifshitz, *Course of Theoretical Physics*, Vol. 6: *Fluid Mechanics* (Nauka, Moscow, 1986; Pergamon, New York, 1987).
36. E. J. Acosta, M. J. Towler, and H. G. Walton, Liq. Cryst. **27**, 977 (2000).
37. I. E. Dzyaloshinskii and G. E. Volovik, Ann. Phys. **125**, 67 (1980).

Journal of Experimental and Theoretical Physics, Vol. 95, No. 4, 2002, pp. 728–735.
From Zhurnal Éksperimental'noĭ i Teoreticheskoĭ Fiziki, Vol. 122, No. 4, 2002, pp. 840–848.

Temperature- and Field-Induced Transitions in Free-Standing Films of an Antiferroelectric Liquid Crystal¶

P. V. Dolganov[a, *], E. I. Demikhov[a], Y. Suzuki[b], and A. Fukuda[c]

[a]*Institute of Solid State Physics, Russian Academy of Sciences, Chernogolovka, Moscow oblast, 142432 Russia*
[b]*Central Research and Development Laboratory, Showa Shell Sekuyi, K. K., Atsugi Kanagawa 243-0206, Japan*
[c]*Department of Kansei Engineering, Shinshu University, Ueda 386-8567, Japan*
**e-mail: pauldol@issp.ac.ru*

Received May 17, 2002

Abstract—Thin free-standing films of a compound with the smectic-C_A^* and smectic-C_α^* phases were investigated by means of depolarized microscopy and optical reflectivity. In thin films, the smectic phase sequence $C_A^*-C_\alpha^*-A$ is replaced by a series of temperature- and field-induced transitions into states with the coplanar orientation of molecular tilt planes. Transitions are accompanied by a change in the direction of the electric polarization with respect to the tilt plane of molecules. The coplanar structure of these states is consistent with the Ising model. © *2002 MAIK "Nauka/Interperiodica"*.

1. INTRODUCTION

The synclinic smectic-*C* phase (Sm*C*) and the anticlinic smectic-C_A phase (SmC_A) are the fundamental mesophases with fluid layers that are observed in rodlike molecular systems. When the constituent molecules render the system chiral, the respective phases are denoted as Sm*C** and SmC_A^* and become ferroelectric, with the spontaneous polarization perpendicular to the tilt plane, and antiferroelectric [1–5]. The locally averaged molecular tilt direction, called the director, is specified by the polar angle θ and the azimuthal angle ϕ in the liquid crystal frame of reference with the *z* axis along the smectic layer normal. In the bulk sample, the tilt angle θ is constant at a given temperature. Chirality leads to the formation of a helicoidal structure, $\phi = 2\pi z/p$, where *p* is the helicoidal pitch. Because *p* is sufficiently larger than the smectic layer spacing *d* in general, the azimuthal angle difference between adjacent layers is $\Delta\phi \simeq 0°$ in Sm*C** and $\Delta\phi \simeq 180°$ in SmC_A^*; the respective phases therefore remain practically synclinic and anticlinic. In other words, the ordinary chiral intermolecular interaction is weak. The molecular origin of the synclinic and anticlinic ordering in Sm*C* and SmC_A has been studied in detail [6, 7]. It is concluded that the conventional dispersion and steric interactions stabilize Sm*C* and Sm*C**, while the orientational correlations of transverse dipoles in adjacent layers actually promote SmC_A and SmC_A^*. We note that these short-range interactions and correlations are not sensitive to the molecular chirality and favor the coplanar synclinic or anticlinic structure.

The transition between synclinic ferroelectric Sm*C** and anticlinic antiferroelectric SmC_A^* is of the first order. In many materials, a sequence of polar subphases with periods consisting of more than two layers are formed between Sm*C** and SmC_A^*. In this narrow temperature interval, the system is frustrated and both phase structures have nearly the same energy; the subphases must therefore be stabilized by some relatively weak additional factors [7]. Isozaki *et al.* [8, 9] systematically investigated the electric-field-temperature (*E*–*T*) phase diagram in several compounds and mixtures [5–8, 10]. They confirmed that at least five subphases can exist between SmC_A^* and Sm*C** and an additional one, denoted as SmC_α^*, just below Sm*A* on the high-temperature side of Sm*C**, suggesting that the subphases between SmC_A^* and Sm*C** constitute a part of a devil's staircase formed by frustration between ferro- and antiferroelectricity. Moreover, it was noticed by Takanishi *et al.* and Hiraoka *et al.* [11–15] that SmC_α^* itself is not a simple single phase but can constitute another devil's staircase [16] where the ferroelectric Sm*C** and/or antiferroelectric SmC_A^* soft mode fluctuations play an important role because of the extremely small tilt angle. Sophisticated experimental techniques, such as the polarized resonant X-ray scattering [17–19], precision ellipsometry and reflectometry [18, 20–22], and advanced polarizing microscopy [23], have recently been used to determine the detailed subphase

¶This article was submitted by the authors in English.

structures. It is unambiguously established that the subphases between SmC_A^* and SmC^* with three- and four-layer periodicities are not coplanar and that their azimuthal angle difference between adjacent layers considerably deviates from $\Delta\phi \simeq 0°$ (synclinic) or 180° (anticlinic) [18, 24]. With respect to SmC_α^*, a short pitch helical structure was recently emphasized in contrast with the previous suggestion of its devil's staircase character.

Two conflicting approaches have been proposed with the important factor being either the continuous short-pitch evolution of the SmC_α^* helical structure [25–29] or the devil's staircase character not only of the subphase emerging between SmC_A^* and SmC but also of SmC_α^* itself [5, 10, 30]. The first one, called the discrete, clock, or X–Y model, takes competing orientational interactions between nearest- and next-nearest-neighbor smectic layers into account. The minimum of the free energy then corresponds to a uniform rotation of the tilt plane about the layer normal. The formation of subphases can be qualitatively explained by introducing several much more complicated interactions with three- and four-layer periodicities and the continuous short-pitch evolution of SmC_α^*. The second approach is based on the microscopic Ising model with competing repulsive and attractive interactions between nearest and next-nearest neighbors. It is known that a sequence of subphases resembling the devil's staircase is indeed obtained in such a model [31, 32]. But the Ising-like Hamiltonian can hardly be applied to smectic liquid crystals [7]. Neither model can therefore appropriately explain the facts experimentally observed thus far.

Prost and Bruinsma [33, 34] proposed a more consistent model by taking long-range polarization fluctuations into account; they tried to explain the formation of SmC_α^* and its devil's staircase character. This mechanism seems to be very promising, and should be taken into consideration in a more general theory [7]. In order to understand the sequence of the subpahses and their nonplanar structures better, further careful experimental investigations are necessary in a variety of compounds and mixtures. The structures of the SmC_α^* phase in bulk samples and in thin films were recently reported to be substantially different [35, 36]. Structures without a short-pitched helix were found above the SmC^* and SmC_A^* phases. Another question is the origin and behavior of the electric polarization in thin films of structures without the net polarization in the bulk sample, in particular, in the SmC_A^* and SmC_α^* phases. The clock and Ising models predict a different behavior of spontaneous polarization as the temperature changes: a smooth variation of the polarization value in the clock model and its steplike changes at certain temperatures in the Ising model. Unusual behavior of antiferroelectric structures in the electric field was found in thin films [37, 38]. These structures possess a ferroelectric polarization: perpendicular (P_Y) to the average direction of the tilt planes in odd-N films (where N is the number of smectic layers) and parallel (P_X) to the average direction of the tilt planes in even-N films.

Because the free energy difference between subphases is intrinsically small, several interface effects in both homogeneous and homeotropic cells disturb the supbphase sequence and structures. Free-standing films [39, 40] are most suitable for making observations that are almost free of these effects. Moreover, applying an electric field can also seriously deform the structure. Consequently, it is essential to study free-standing films by applying a sufficiently weak electric field for controlling the director alignment. In this paper, we present such an example of the direct microscopic observation and optical reflectivity. We believe that the observed results may reflect the bulk property to some extent, although the interface effects must also be considerably large because the number of smectic layers is somewhat small.

2. EXPERIMENT

The material studied was 4-(1-trifluoromethyl-hepthyloxycarbonyl)phenyl 4'-octyl-biphenyl-4-carboxylate (TFMHPBC) [5]. Bulk TFMHPBC exhibits the following transition temperatures between smectic phases:

$$\mathrm{Sm}C_A^* \xrightarrow{74.3°\mathrm{C}} \mathrm{Sm}C_\alpha^* \xrightarrow{75.0°\mathrm{C}} \mathrm{Sm}A.$$

Two types of cells were used in optical measurements. For the first type, films were prepared in a rectangular frame with two mobile metallic blades. For the second type, free-standing films were drawn over a 4-mm-diameter circular hole in a glass plate. Electrodes on four sides of the hole were used to apply an in-plane electric field. The number N of smectic layers was determined by optical reflectivity [39, 41].

Two techniques were used for studies of the temperature- and field-induced phase transitions. The first one was the depolarized reflected light microscopy (DRLM) [42, 43]. For relatively thin films, the reflection intensity is proportional to $(n^2 - 1)^2$ [44],

$$I \propto (n^2 - 1)^2 \pi^2 N^2 d^2 / \lambda^2,$$

where n is the refraction index and λ is the light wavelength. The refraction index along the tilt plane, $n_\parallel$, is different from the refraction index in the perpendicular direction, n_0. The DRLM allows visualizing domains with different orientations of the tilt plane with respect to the direction of the electric polarization **P**. The tech-

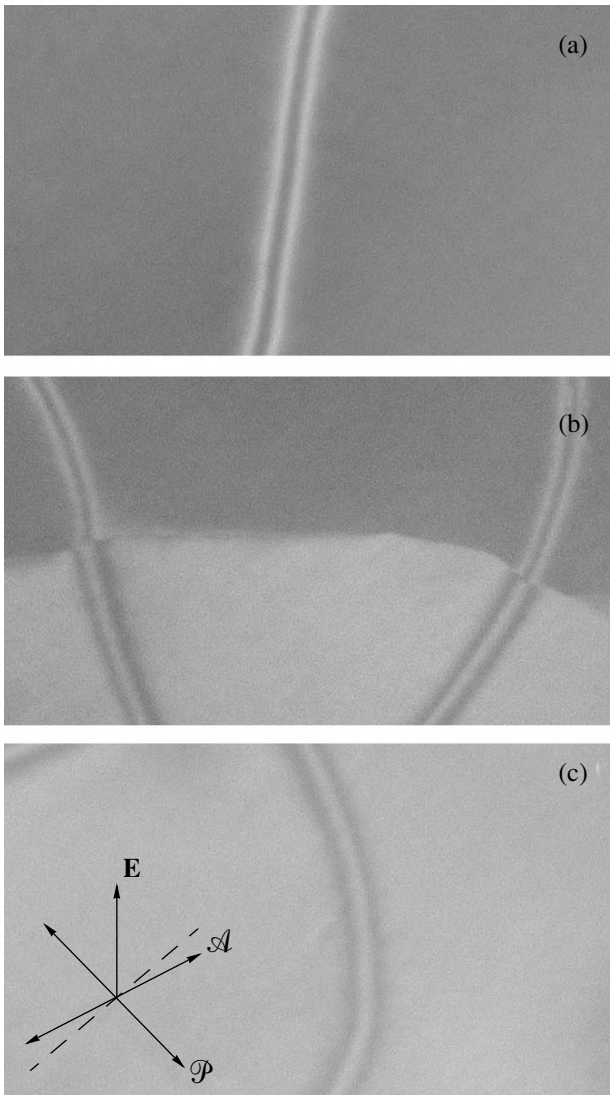


Fig. 1. Temperature-induced transition in even-*N* films. (a) The microscope image in reflection of an eight-layer TFMHPBC film in the antiferroelectric phase at 70.7°C. The *c*-director is aligned parallel to the electric field. The film appears dark with a bright 2π wall. (b) Transition on heating (T = 75 °C): the region of a high-temperature phase with different orientations of the *c*-director emerges. Upon heating, the front moves from bottom to top of the picture. (c) The state at higher temperature with polarization perpendicular to the tilt plane (T = 75.7°C). The electric field value is 8.8 V/cm. The orientation of the polarizer ($\mathcal{P}$), the analyzer ($\mathcal{A}$), and the electric field is shown in Fig. 1c. The horizontal size of each photograph is 370 μm.

nique uses a polarizer turned by 45° with respect to the electric field direction. The difference between the refraction indices for the two main directions results in an effective rotation of the polarization plane of the reflected light in the direction of the tilt plane. In slightly decrossed polarizers, the domains therefore look dark or bright depending on the orientation of the tilt plane with respect to the direction of the electric polarization. In this method, the low intensity of the image is compensated by a highly contrast pattern.

Another technique was used in measurements of the linearly polarized reflectivity from the films in "backward" geometry. In electrooptical measurements, we could change the direction of light polarization and the direction of the electric field in the plane of the film. In tilted smectic phases, the films are optically anisotropic in the plane of the layers. The reflection intensities with the light polarization parallel ($I_\parallel$) and perpendicular ($I_\perp$) to the direction of the electric field were measured. Using this technique, we could determine the orientation of the tilt plane in the films and their optical anisotropy.

3. RESULTS AND DISCUSSION

In the geometry used for imaging, SmC_A^* films with an even number of layers look like a dark field with two bright lines in the regions where 2π walls exist (Fig. 1a). Films have the polarization P_X parallel to the tilt plane, and the *c*-director (the projection of **n** on the film plane [3]) is parallel to the vertical axis. Across the 2π wall, the *c*-director rotates by the angle 2π; as a result, regions with a nearly horizontal orientation of the *c*-director in 2π walls look like two bright narrow lines. In microscopic investigations, we chose regions of the films with 2π walls. In these regions, it is easiest to judge the orientation of the *c*-director with respect to the electric field direction. Heating to 75 °C leads to the appearance of a new phase (Fig. 1b). In spite of a small tilt angle, the phase boundary and the phases themselves are easily visualized in films using DRLM above the bulk transition temperature. A transformation of the DRLM picture (dark 2π walls in the bright field, Figs. 1b, 1c) shows that the direction of the polarization with respect to the *c*-director is drastically different from that in the low-temperature state. The transition may be caused either by a substantial change of the relative orientation of tilt planes in neighboring layers or by the temperature variation of relative values of the polarizations P_X and P_Y and the rotation of the *c*-director of the film in the electric field. The existence of a sharp boundary and the behavior of 2π walls near it show that the transition is not caused by a smooth temperature decrease in the P_X component of the polarization with respect to the P_Y one. In that case, the transition front would be blurred and the regions of 2π walls with the horizontal orientation of the *c*-director could smoothly turn into a new oriented state.

In films with an odd number of layers, we also observed the transition to a new structure (Fig. 2). A transformation of the DRLM picture indicates the transition from the structure with the polarization P_Y perpendicular to the tilt plane to a structure with the dominant polarization P_X. Thus, films having polarizations

of different types in the SmC_A^* phase (P_Y in odd-N films and P_X in even-N films) after the transition also possess different polarization directions in the plane of the layers (P_Y in even-N films and P_X in odd-N films). We observed such a behavior in all investigated films with the thickness from three to more than twenty layers. But the transition temperatures strongly depend on the film thickness (in a three-layer film, the transition occurs at about 85°C). Figures 1b and 2b show two realizing possibilities for the coexistence of 2π walls and the boundary of two phases. In the antiferroelectric phase, the 2π walls are either continued into the new phase (Fig. 1b) or split into walls that are localized near the boundary as in Fig. 2b.

In films with $N > 10$, we found an increase in the number of structures above the SmC_A^* phase. These structures can be simultaneously observed in a sample with a temperature gradient. An image of such a film is shown in Fig. 3. The temperature increases from left to right. In the left part of the figure, the film is in the SmC_A^* phase *1*. In regions *2–4*, the sample is in different states with different directions of the polarization. This is clearly seen from measurements of the optical reflectivity from a film for two directions of the light polarization, perpendicular ($I_\perp$) and parallel ($I_\parallel$) to the electric field direction. Figure 4 shows the results of such measurements for an 18-layer film. In the SmC_A^* phase (Fig. 4a), $I_\perp < I_\parallel$, which implies that the polarization is parallel to the tilt plane. The same polarization direction is observed in state *3* (Fig. 4c). In states *2* and *4* (Figs. 4b, 4d), $I_\perp > I_\parallel$, which indicates that the tilt plane is perpendicular to the electric field direction and to the film polarization. The sequence of transitions from state *1* to state *4* therefore occurs with the change in the polarization direction. These measurements also provide information about optical anisotropy of the films. This enables us to answer one of the main questions: Do short-pitch helical structures emerge above the SmC_A^* phase? Formation of the short-pitched azimuthal helix with the pitch on the order of or less than the film thickness should lead to a substantial decrease in the reflection anisotropy $|I_\perp/I_\parallel - 1|$. In such a nearly optically uniaxial structure, the anisotropy can occur because the film thickness is not a multiple of the half-pitch. But the magnitude of the anisotropy in short-pitched structures must be several times smaller than that before the transition from the SmC_A^* phase. A smooth decrease in $|I_\perp/I_\parallel - 1|$ in Figs. 4b–4d is due to a well-known decrease in the molecular tilt angle on heating. We found that $I_\perp/I_\parallel - 1$ changes its sign as the result of transitions, but the absolute value changes only insignificantly. Therefore, even in relatively thick films (Fig. 4), transitions occur into structures without a short pitch. At high temperatures, state *4* (Figs. 3 and 4) can be switched to the state with the longitudinal polarization by decreasing the electric field value.

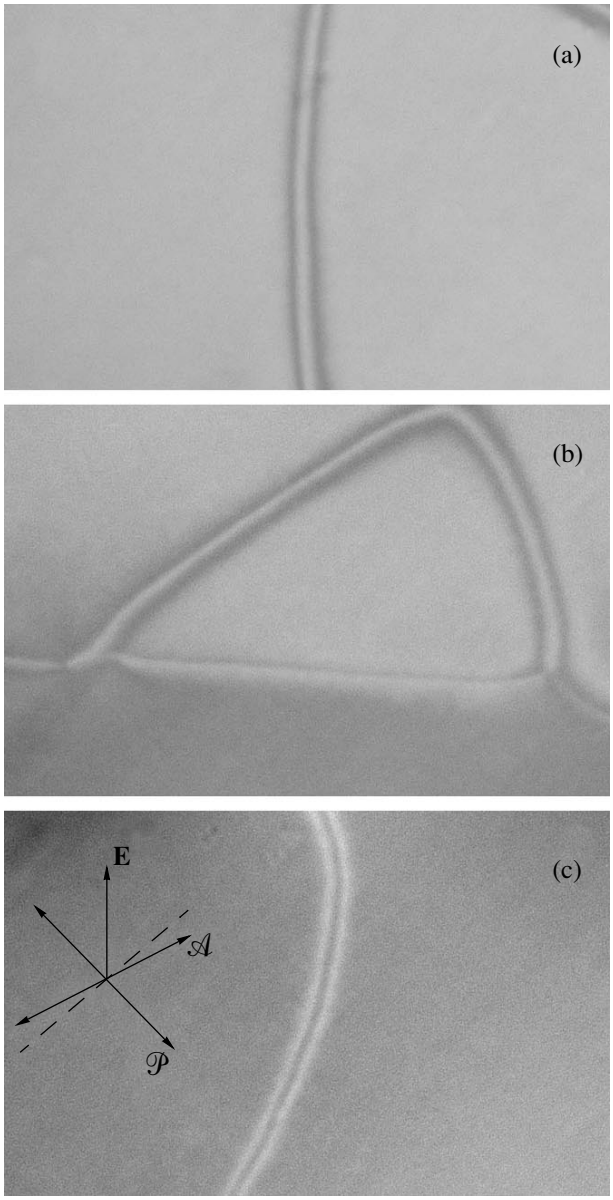


Fig. 2. Temperature-induced transition in odd-N films. (a) The optical microscope picture of a nine-layer TFMHPBC film in reflection in the low-temperature (antiferroelectric) phase (T = 70.4°C). In the odd-N film, the c-director is oriented along the horizontal axis. The film is bright with dark 2π walls. (b) On heating to 75°C, the film undergoes a transition to the phase with the polarization parallel to the tilt plane. This phase appears in the lower part of the photograph. (c) The film after the transition. The entire film appears dark with a bright 2π wall. The electric field value is 7.5 V/cm. The horizontal size of each photograph is 420 μm.

Four pictures in Fig. 5 demonstrate the behavior of an eight-layer film at high temperatures. In the low field, the film is in the state with the polarization parallel to the tilt plane (the bright 2π wall in a dark field, Fig. 5a). The increase in the field induces the appear-

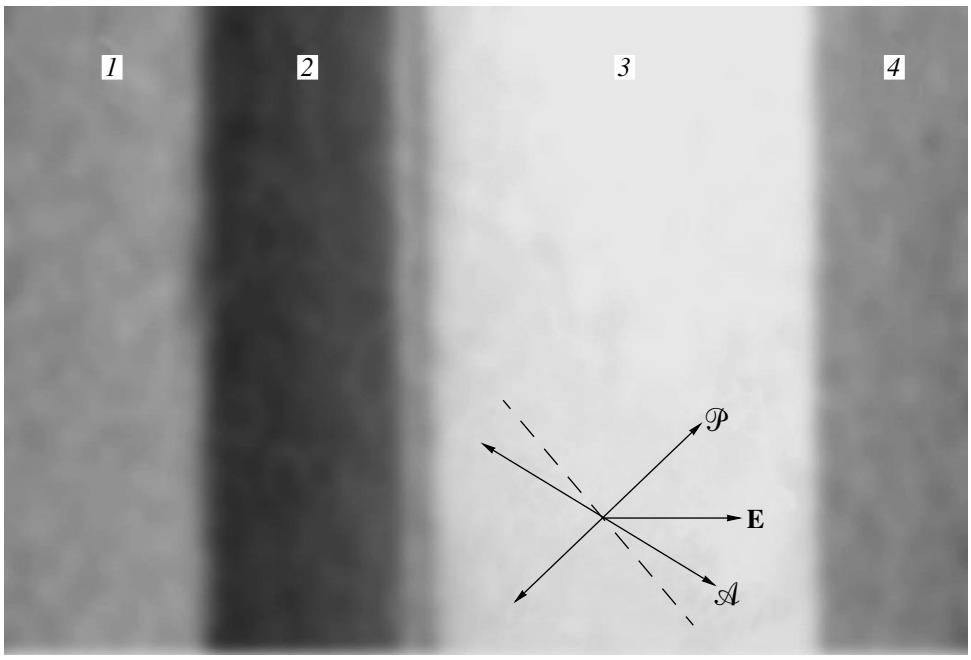


Fig. 3. States with the polarizations parallel and perpendicular to the tilt plane in a 12-layer film are observed as regions with different brightness levels. A sample with a temperature gradient was used for this observation. The left part of the film corresponds to a temperature of about 74°C and the right part to a temperature of about 76 °C. In region *1*, the film is in the Sm C_A^* phase. In regions *2–4*, states with different orientations of the tilt plane exist. These states can be aligned by the electric field. The narrow stripe between regions *2* and *3* is a line defect localized at the boundary of the states. The horizontal size of the image is 610 μm, E = 10 V/cm.

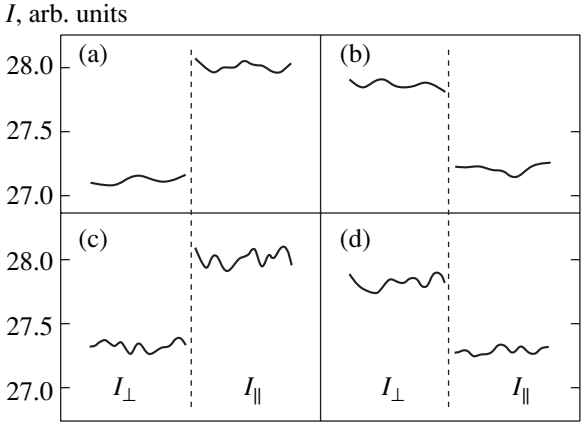


Fig. 4. Optical reflectivity from an 18-layer film for two directions of the light polarization: perpendicular ($I_\perp$) and parallel ($I_\parallel$) to the electric field. A sequence of four states was observed: the Sm C_A^* phase in state *1* (a) and states *2–4* (b–d) above the Sm C_A^* phase. Transitions between these states occur with the change in the direction of the electric polarization. The tilt plane of molecules is oriented parallel to the electric field and polarization ($I_\perp < I_\parallel$) in states *1* and *3* (a, c). In states *2* and *4* (b, d), the tilt plane is oriented perpendicular to the electric field and polarization ($I_\perp > I_\parallel$). In the Sm C_A^* phase, the measurements were made near the temperature T_0 of the transition to state *2*, such that $T_0 - T$ = 0.3°C; to states *2–4*, respectively, at T_0 + 0.15°C, T_0 + 0.2°C, and T_0 + 1.1°C. The data were obtained at the electric field value of 35 V/cm.

ance of the structure with the polarization perpendicular to the tilt plane (Figs. 5b, 5c). The switching electric field E_s is more than three orders of magnitude smaller than the field inducing the anticlinic–synclinic transition in the Sm C_A^* phase at low temperatures. Such a small value of E_s is due to the small tilt angle of the molecules. Seemingly, the change of interlayer organization (anticlinic–synclinic or vice versa) occurs in only several layers. This field-induced transition is reversible and the decreasing field returns the film to its initial state (Fig. 5d). It is worth mentioning that such a behavior at high temperatures is observed in films of all thicknesses investigated, is independent of the film oddity, and is similar to the behavior of films without the Sm C_α^* phase with a large value of the layer polarization [45, 46].

The number of transitions caused by heating depends on the film thickness, its oddity, and the value of the electric field. In thin odd-N films, one transition is observed in the low field and two in the high field. Thin even-N films undergo two transitions in the low field and only one in the high field. The crossover from low to high field is determined by the value of E_s, depends on the film thickness, and typically occurs at about 10 V/cm. In thicker films (N > 10), the number of transitions occurring with heating is increased. For odd-N antiferroelectric films, it is odd (even) in the low

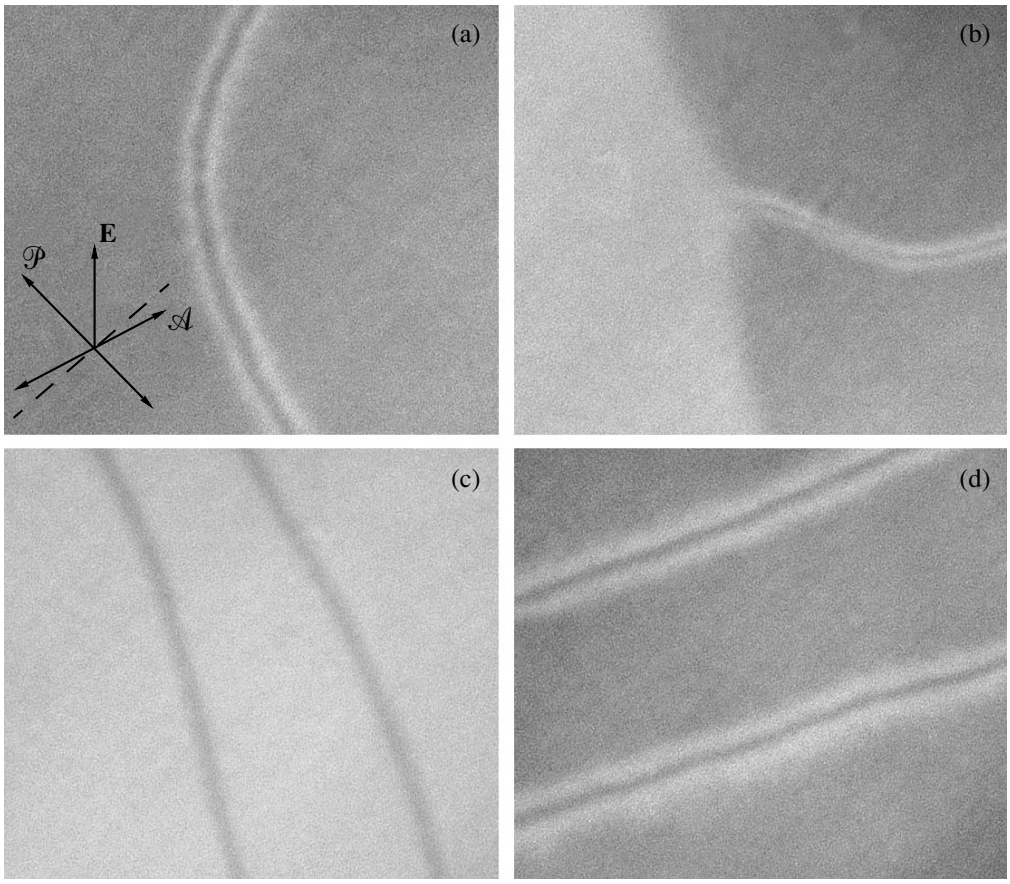


Fig. 5. An example of the field-induced transition at a high temperature in an eight-layer TFMHPBC film. (a) At the low electric field 4 V/cm, the film is in the state with the polarization parallel to the tilt plane. The *c*-director is aligned in the vertical direction (dark background with a bright 2π wall). (b) When the field is increased to 19 V/cm, the region of the high-field state with the polarization perpendicular to the tilt plane emerges as a bright one. This region gradually fills the entire film. The front moves from the left to the right side of the image. (c) The film in the high-field state. (d) When the electric field is decreased to 4 V/cm, the film returns to the low-field state. $T = 78.8°C$. The horizontal size of each image is about 300 μm.

(high) electric field. For even-*N* films, the number of transitions is even (odd) in the low (high) electric field. A simpler situation is expected in films with the low-temperature ferroelectric phase: an odd (even) number of transitions in the low (high) electric field in both odd-*N* and even-*N* films. As already mentioned, this is the case for compounds with a high layer polarization. For films with a small layer polarization in which the polarization of the high-temperature ground state is perpendicular to the tilt plane [45], the number of transitions is opposite to that described above.

Film surfaces can influence the structure of the film and phase transitions in different ways. The presence of a surface breaks the symmetry existing in bulk samples. For layers near the surface, the nearest and next-nearest interlayer interactions are realized only on one side (only on the side of the film). The missing interlayer interaction near the surface can influence both the modulus θ and the phase ϕ of the order parameter. Symmetry breaking in antiferroelectric films must lead to a smaller magnitude of the polar angle θ and to the low-temperature shift of phase transitions with respect to the bulk sample [47]. The second effect is related to the existence of the surface as a physical boundary between two media. In liquid crystals, this effect mainly manifests itself through the surface tension and the suppression of the layer displacement fluctuations at free surfaces that increase the modulus of the order parameter (surface freezing) and the phase transitions temperatures. To the best of our knowledge, the direct influence of the surface tension on the phase ϕ of the order parameter and suppression of the smectic-layer fluctuations have not been considered theoretically or observed experimentally. Presently, it is still debated which of the above effects prevails for antiferroelectric liquid crystals [47, 48]. The existence of a tilted structure above the bulk transition temperature to the Sm*A* phase and a high-temperature shift of the phase transitions in thin films show that, in antiferroelectric films, as in ferroelectric ones, the surface freezing effect dominates. In strong surface-freezing conditions, the most distinctive feature of the film is that the profile of θ is essentially inhomogeneous across the film. Above but near the temperature of the bulk transition to the SmC_α^* phase,

the modulus of the order parameter on the surface exceeds that in the bulk sample. At the same time, the tilt is smaller in the interior of the film; at some temperature, it corresponds to the value at which transitions with a change of ϕ occur in the bulk sample. In thin films, only a few interior layers can undergo the transition. In thick films, the conditions for a change of ϕ arise in a greater number of layers and the number of possible structures with the synclinic and anticlinic tilt is increased. This is related to the observed increase in the number of transitions in thicker films.

For the SmC_A^* structure, the appearance of the polarization P_Y perpendicular to the tilt plane in odd-*N* films and the polarization P_X parallel to the tilt plane in even-*N* films is related to different symmetries with respect to the center of the film [37]. Similar symmetry considerations apply to films consisting of synclinic and anticlinic pairs. The change in the number of pairs with synclinic and anticlinic tilts by one can lead to the transition from the P_Y net polarization to the P_X one and vice versa. In even-*N* films, this reorientation of the tilt can occur with the formation of a symmetric or antisymmetric structure with respect to the center of the film and, accordingly, to the P_X or P_Y net polarization. In odd-*N* films, formation of a synclinic pair destroys the symmetry with respect to the center of the film. In such films, the P_X or P_Y polarization can dominate. Transition to the P_X polarization with the appearance of a symmetric structure is possible when the molecular tilt becomes zero in the central layer of the odd-*N* film. We mention that, in this case, the decrease in the molecular tilt angle in the center of the film occurs gradually, because we did not observe a substantial decrease in anisotropy in thin films at transitions with the change in the *c*-director orientation.

Two reasons may be responsible for the formation of a coplanar structure in films above the SmC_A^* phase. At low temperatures, the condition for the reorientation of the tilt planes arises only in the central part of the film because of surface freezing, which prevents the formation of a short-pitch structure. The second reason is related to peculiarities of fluctuations of θ in Sm*A* films [47]. In the bulk sample, sinusoidal fluctuations in perpendicular planes can assemble with an arbitrary phase shift, which leads to coplanar, circular, or ellipsoidal fluctuations and to possibilities of the formation of tilted pitch structures at a second-order transition. As pointed out in [47], unlike in the bulk sample, two fluctuation modes have the same or the opposite phases in films because of the symmetry with respect to the center of the film, which leads to only coplanar fluctuations and to the transition to coplanar tilted structures. Coplanar structures in thin films may also reflect the interlayer molecular organization in thick films. In optical experiments [20–22] in the range of the SmC_α^* phase, the films possess optical anisotropy and polarization. These may not only be related to surface effects but also be an indication of a distorted staircase character of the SmC_α^* phase in thick free-standing films.

In summary, using depolarized light microscopy and polarized reflectivity, we observed coplanar structures in free-standing films above the SmC_A^* phase. The number of transitions depends on the film thickness, its oddity, and the value of the electric field. Temperature- and field-induced transitions occur as the direction of the electric polarization changes with respect to the tilt plane.

ACKNOWLEDGMENTS

This work was supported in part by the Russian Foundation for Basic Research (project no. 01-02-16507) and a Grant-in-Aid for Scientific Research (no. 12650010) through the Japan Society for the Promotion of Science. We acknowledge very stimulating discussions with D.R. Link and H. Takezoe. One of us (P.V.D.) thanks Shinshu University for his stay as a foreign visiting researcher.

REFERENCES

1. R. B. Meyer, L. Liebert, L. Strzelecki, and P. Keller, J. Phys. Lett. **36**, L69 (1975).
2. R. B. Meyer, Mol. Cryst. Liq. Cryst. **40**, 33 (1977).
3. P. G. de Gennes, *The Physics of Liquid Crystals* (Clarendon, Oxford, 1974; Mir, Moscow, 1977).
4. A. D. L. Chandani, E. Gorecka, Y. Ouchi, *et al.*, Jpn. J. Appl. Phys. **28**, L1265 (1989).
5. A. Fukuda, Y. Takanishi, T. Isosaki, *et al.*, J. Mater. Chem. **4**, 997 (1994).
6. M. A. Osipov and A. Fukuda, Phys. Rev. E **62**, 3724 (2000).
7. M. A. Osipov, A. Fukuda, and H. Hakoi, submitted to Mol. Cryst. Liq. Cryst. (2002).
8. T. Isozaki, T. Fujikawa, H. Takezoe, *et al.*, Jpn. J. Appl. Phys. **31**, L1435 (1992).
9. T. Isozaki, T. Fujikawa, H. Takezoe, *et al.*, Phys. Rev. B **48**, 13439 (1993).
10. T. Matsumoto, A. Fukuda, M. Johno, *et al.*, J. Mater. Chem. **9**, 2051 (1999).
11. Y. Takanishi, K. Hiraoka, V. Agrawal, *et al.*, Jpn. J. Appl. Phys. **30**, 2023 (1991).
12. K. Hiraoka, Y. Takanishi, K. Skarp, *et al.*, Jpn. J. Appl. Phys. **30**, L1819 (1991).
13. K. Hiraoka, Y. Takanishi, H. Takezoe, *et al.*, Jpn. J. Appl. Phys. **31**, 3394 (1992).
14. T. Isozaki, K. Hiraoka, Y. Takanishi, *et al.*, Liq. Cryst. **12**, 59 (1992).
15. K. Yamada, Y. Takanishi, K. Ishikawa, *et al.*, Phys. Rev. E **56**, R43 (1997).
16. Ch. Bahr, D. Fliegner, C. J. Booth, and J. W. Goodby, Phys. Rev. E **51**, R3823 (1995).
17. P. Mach, R. Pindak, A.-M. Levelut, *et al.*, Phys. Rev. Lett. **81**, 1015 (1998).

18. P. M. Johnson, D. A. Olson, S. Pankratz, *et al.*, Phys. Rev. Lett. **84**, 4870 (2000).
19. A.-M. Levelut and B. Pansu, Phys. Rev. E **60**, 6803 (1999).
20. D. Shlauf, Ch. Bahr, and H. T. Nguyen, Phys. Rev. E **60**, 6816 (1999).
21. D. A. Olson, S. Pankratz, P. M. Johnson, *et al.*, Phys. Rev. E **63**, 061711 (2001).
22. P. M. Johnson, S. Pankrats, P. Mach, *et al.*, Phys. Rev. Lett. **83**, 4073 (1999).
23. V. Laux, N. Isaert, H. T. Nguyen, *et al.*, Ferroelectrics **179**, 25 (1996).
24. T. Akizuki, K. Miyachi, Y. Takanishi, *et al.*, Jpn. J. Appl. Phys. **38**, 4832 (1999).
25. H. Sun, H. Orihara, and Y. Ishibashi, J. Phys. Soc. Jpn. **62**, 2706 (1993).
26. M. Čepič and B. Žekš, Phys. Rev. Lett. **87**, 085501 (2001).
27. A. Roy and N. V. Madhusudana, Eur. Phys. J. E **1**, 319 (2000).
28. M. Čepič and B. Žekš, Mol. Cryst. Liq. Cryst. **263**, 61 (1995).
29. M. Skarabot, M. Čepič, B. Žekš, *et al.*, Phys. Rev. E **58**, 575 (1998).
30. M. Yamashita and S. Miyazima, Ferroelectrics **148**, 1 (1993).
31. M. E. Fisher and W. Selke, Phys. Rev. Lett. **44**, 1502 (1980).
32. P. Bak and J. von Boehm, Phys. Rev. B **21**, 5297 (1980).
33. J. Prost and R. Bruinsma, Ferroelectrics **148**, 25 (1993).
34. R. Bruinsma and J. Prost, J. Phys. II **4**, 1209 (1994).
35. A. Fera, R. Opitz, W. H. de Jeu, *et al.*, Phys. Rev. E **64**, 021702 (2001).
36. P. V. Dolganov, Y. Suzuki, and A. Fukuda, Phys. Rev. E **65**, 031702 (2002).
37. D. R. Link, J. E. Maclennan, and N. A. Clark, Phys. Rev. Lett. **77**, 2237 (1996).
38. D. R. Link, G. Natale, N. A. Clark, *et al.*, Phys. Rev. Lett. **82**, 2508 (1999).
39. P. Pieranski, L. Beliard, J.-Ph. Tournellec, *et al.*, Physica A (Amsterdam) **194**, 364 (1993).
40. For reviews see, e.g., Ch. Bahr, Int. J. Mod. Phys. B **8**, 3051 (1994); T. S. Stoebe and C. C. Huang, Int. J. Mod. Phys. B **9**, 2285 (1995).
41. I. Kraus, P. Pieranski, E. I. Demikhov, and H. Stegemeyer, Phys. Rev. E **48**, 1916 (1993).
42. R. Pindak, C. Y. Young, R. B. Meyer, and N. A. Clark, Phys. Rev. Lett. **45**, 1193 (1980).
43. D. R. Link, G. Natale, R. Shao, *et al.*, Science **278**, 1924 (1997).
44. M. Born and E. Wolf, *Principles of Optics* (Pergamon, Oxford, 1964; Nauka, Moscow, 1973).
45. P. M. Johnson, D. A. Olson, S. Pankratz, *et al.*, Phys. Rev. E **62**, 8106 (2000).
46. C. Y. Chao, C. R. Lo, P. J. Wu, *et al.*, Phys. Rev. Lett. **86**, 4048 (2001).
47. B. Rovsek, M. Čepič, and B. Žekš, Phys. Rev. E **62**, 3758 (2000).
48. B. Rovsek, M. Čepič, and B. Žekš, Mol. Cryst. Liq. Cryst. **328**, 997 (1997).

Journal of Experimental and Theoretical Physics, Vol. 95, No. 4, 2002, pp. 736–747.
From Zhurnal Éksperimental'noĭ i Teoreticheskoĭ Fiziki, Vol. 122, No. 4, 2002, pp. 849–862.

SOLIDS
Structure

Pressure-Induced Transformations and Optical Properties of the Two-Dimensional Tetragonal Polymer of C_{60} at Pressures up to 30 GPa¶

K. P. Meletov[a, *], J. Arvanitidis[b], S. Assimopoulos[b], G. A. Kourouklis[b], and B. Sundqvist[c]

[a]*Institute of Solid State Physics, Russian Academy of Sciences, Chernogolovka, Moscow oblast, 142432 Russia*
[b]*Physics Division, School of Technology, Aristotle University of Thessaloniki GR-540 06, Thessaloniki, Greece*
[c]*Department of Physics, Umeå University S-901 87, Umea, Sweden*
**e-mail: mele@issp.ac.ru*

Received April 24, 2002

Abstract—The Raman spectra of the two-dimensional tetragonal (2D(T)) polymeric phase of C_{60} have been studied *in situ* at pressures up to 30 GPa and room temperature. The pressure dependence of the phonon modes shows an irreversible transformation of the material near 20 GPa into a new phase, most probably associated with the covalent bonding between the 2D polymeric sheets. The Raman spectrum of the high-pressure phase is intense and well resolved, and the majority of modes are related to the fullerene molecular cage. The sample recovered at ambient conditions is in a metastable phase and transforms violently under laser irradiation: the transformed material contains mainly dimers and monomers of C_{60} and small inclusions of the diamond-like carbon phase. The photoluminescence spectra of the 2D(T) polymer of C_{60} were measured at room temperature and pressure up to 4 GPa. The intensity distribution and the pressure-induced shift of the photoluminescence spectrum drastically differ from those of the C_{60} monomer. The deformation potential and the Grüneisen parameters of the 2D(T) polymeric phase of C_{60} have been determined and compared with those of the pristine material. © *2002 MAIK "Nauka/Interperiodica".*

1. INTRODUCTION

The polymeric forms of C_{60} have attracted considerable attention because of their interesting structure and properties [1]. Pristine C_{60} has a great potential for polymerization because of the existence of 30 double C=C bonds in the fullerene molecular cage. C_{60} has been found to polymerize under illumination with visible and ultraviolet light [2] and upon alkali metal doping [3, 4]. The treatment of C_{60} under various high-pressure and high-temperature conditions also leads to polymerization of the material (HPHT polymers) [5]. The covalent polymeric bonds are usually formed by the so-called [2 + 2] cycloaddition reaction via the formation of four-member rings between adjacent fullerene molecules, resulting in an appreciable decrease in the intermolecular distance [2].

The structure and the dimensionality of HPHT polymers strongly depend on the pressure (P) and temperature (T) treatment conditions. The C_{60} molecules form linear polymeric chains (one-dimensional polymer) having an orthorhombic crystal structure (1D(O)) and/or dimers and higher oligomers at lower P and T, two-dimensional polymeric layers that have either a rhombohedral (2D(R)) or a tetragonal (2D(T)) crystal structure at intermediate P and T, and face-centered cubic structures based on three-dimensional (3D) cross-linked polymerization of the material at higher P and T [1, 5–7]. In addition, the treatment of the pristine material under high nonuniform pressure and high temperature leads to the creation of several disordered polymeric phases, the so-called ultrahard fullerite phases [8, 9]. The detailed X-ray studies of these phases have revealed their 3D polymeric character [10, 11].

The polymerization of C_{60} is characterized by the destruction of a number of double C=C intramolecular bonds and the creation of intermolecular covalent bonds associated with sp^3-like fourfold coordinated carbon atoms in the fullerene molecular cage. Their number increases from 4 to 8 and to 12 per each cage for 1D(O), 2D(T), and 2D(R) polymeric phases, respectively, and is expected to further increase in the 3D polymeric phases. Theoretical studies by Okada *et al.* [12] have predicted that the 3D-polymerized C_{60} might be formed by the application of uniaxial pressure perpendicular to the polymeric sheets of the 2D(T) phase of C_{60}. According to their density-functional calculations, polymerization occurs at the lattice constant c = 10.7 Å, which is attainable at the pressure of approximately 20.2 GPa. This polymerization results in the formation of a stable metallic phase having 24 sp^3-like and 36 sp^2-like hybridized carbon atoms in each C_{60} molecule. Another theoretical study, by Burgos *et al.* [13],

¶This article was submitted by the authors in English.

predicted that uniaxial compression perpendicular to the chains in the 1D or to the polymeric planes in the 2D polymeric phases of C_{60} leads to 3D polymerization with 52, 56, and even 60 sp^3-like coordinated carbon atoms per C_{60} molecular cage. These transformations are expected to occur at pressures lower than 14 GPa, and the new phases are semiconducting with large bulk and shear moduli.

The Raman scattering and infrared absorption spectra of various polymeric phases prepared under carefully controlled conditions of HPHT treatment have a very rich and prominent structure. Their intensity distribution and peak positions differ significantly for the 1D(O), 2D(R), and 2D(T) polymeric phases, as has been shown by the detailed study of their optical spectra combined with their structural analysis [14]. The phonon spectra of these materials are very sensitive to any perturbation of the fullerene molecular cage caused by external disturbances such as pressure or chemical bond formation [15, 16]. Therefore, the Raman spectroscopy can be successfully used for the identification of various polymeric phases of C_{60} and for the *in situ* high-pressure studies of phase transformations in the fullerene-related materials. Our experimental Raman studies of the pressure dependence of the 2D(T) polymeric phase of C_{60} [17, 18] have revealed prominent irreversible changes in the Raman spectra of this material near 20 GPa, most probably related to its further polymerization in accordance with the theoretical prediction by Okada *et al.* [12]. The well-structured phonon spectrum of the new high-pressure phase in the 2D(T) polymer provides strong indications that the fullerene molecular cage is retained and that the new phase may be related to a three-dimensional network of the C_{60} molecules. The Raman data [17, 18] disagree with the results of recent high-pressure X-ray studies of the 2D(T) polymer: this fact shows that the material undergoes an irreversible amorphization in the pressure region between 10 and 20 GPa [19]. It is also interesting to note that contrary to the 2D(T) polymer, the 2D(R) polymeric phase of C_{60} transforms to a new phase at approximately 15 GPa; this phase is characterized by very diffused Raman bands that are most probably related to the random covalent bonding between molecules belonging to adjacent 2D polymeric sheets [20].

In this work, we present a detailed photoluminescence and Raman study of the intramolecular phonon modes and electronic spectrum behavior of the 2D(T) polymeric phase of C_{60} at high pressure. Our motivation was to study, in detail, the properties and stability of the material in both 2D(T) and high-pressure induced phases, to obtain quantitative data on the pressure dependence of the phonon and electron spectra, and to compare them with those of pristine C_{60}.

2. EXPERIMENTAL

Two-dimensional polymerized C_{60} was obtained by subjecting 99.99% pure C_{60} powder to the pressure 2.2 GPa at the temperature about 820 K [21]. An X-ray analysis of the samples from the same batch after the high-pressure and high-temperature treatment confirmed that the crystal structure of the polymer is tetragonal (the space group $P4_2mmc$), while a Raman analysis showed the typical spectrum of the 2D(T) polymer of C_{60} with no detectable inclusions of orthorhombic or rhombohedral phases [14, 21, 22].

Raman spectra were recorded using a triple monochromator (Dilor XY-500) equipped with a CCD liquid-nitrogen-cooled detector system. The spectra were taken in the back-scattering geometry using a micro-Raman system comprising an Olympus microscope equipped with objectives of 100× and 20× magnification and the respective spatial resolutions of approximately 1.7 μm and 8 μm. The spectral width of the system was about 5 cm^{-1}. The 514.5-nm line of an Ar^+ laser was used for excitation. The laser power was kept lower than 20 mW measured directly before the high-pressure cell, in order to avoid the destruction of polymeric bonds caused by laser heating effects and related changes in the phonon spectrum and the crystal structure [23–25]. The photoluminescence spectra were recorded using a single monochromator Jobin Yvon THR-1000 equipped with a CCD liquid-nitrogen-cooled detector system. The spectral width of the system was approximately 0.5 meV. The 488-nm line of an Ar^+ laser was used for excitation of the luminescence spectra. The laser power was kept at 2 mW measured directly in front of the high-pressure cell.

Measurements of the Raman and photoluminescence spectra at high pressures were carried out using the diamond anvil cell of Mao–Bell type [26]. A 4 : 1 methanol–ethanol mixture was used as the pressure transmitting medium, and the ruby fluorescence technique was used for pressure calibration [27]. The samples used in the present study had dimensions of 100 μm and were selected from the batch material for their intense, clear, and spatially uniform Raman response, typical of the 2D(T) polymeric phase [14]. The band frequencies in the Raman and photoluminescence spectra were obtained by fitting Lorentzian peak functions to the experimental peaks after the background subtraction.

3. RESULTS AND DISCUSSION

3.1. Phase Transitions

The Raman spectra of the 2D(T) polymer of C_{60} at various pressures up to 27.5 GPa and room temperature, in the frequency region 200–2050 cm^{-1}, are illustrated in Fig. 1a. In this figure, the spectra were recorded upon pressure increase; the spectral region around the strong triple-degenerate T_{2g} mode of dia-

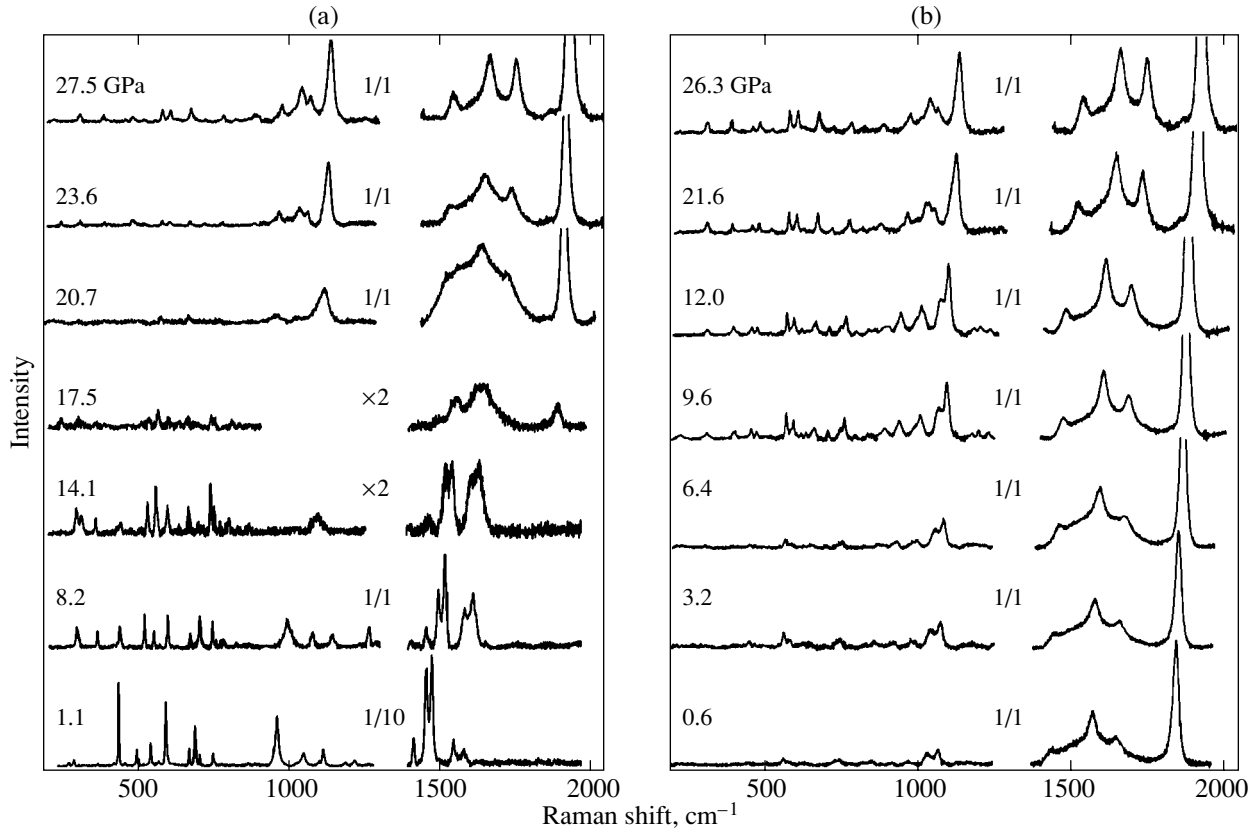


Fig. 1. Raman spectra of the 2D(T) polymer of C_{60} at 300 K and various pressures, recorded for (a) increasing and (b) decreasing pressure runs. The numbers 1/*x* indicate the relative scale of the spectra.

mond appearing at 1332 cm^{-1} at ambient pressure [28] is omitted. The initial spectrum at 1.1 GPa represents a typical Raman spectrum of the 2D(T) polymeric phase and is identical with the spectra reported earlier [14, 17, 18]. Lowering the molecular symmetry from I_h in pristine C_{60} to D_{2h} in the 2D(T) polymer results in the splitting of the degenerate icosahedral intramolecular modes and in the activation of initially silent modes [14, 29, 30]. Moreover, although the 2D(T) phase retains the inversion center of the pristine C_{60} molecule, we cannot discard the possibility that imperfections in the crystal structure of the polymer and/or the natural ^{13}C substitution may facilitate the appearance of some ungerade modes in its Raman spectrum [30]. For these reasons, the Raman spectrum of the 2D(T) polymer is richer in structure than that of pristine C_{60} [31].

As can be clearly seen from Fig. 1a, the Raman peaks of the 2D(T) polymer remain narrow and well resolved for pressures up to 14 GPa, showing the homogeneity and stability of the samples used. We note that, as recently shown [30], the pressure dependence of the Raman modes of the 2D(T) polymer is fully reversible up to 12 GPa. For pressures $P > 14$ GPa, the Raman peak bandwidths of the polymer increase gradually and the intensities of peaks decrease considerably. In addition, the peak broadening is accompanied by a gradual enhancement of the background (not shown in Fig. 1, because the Raman spectra are presented after the background subtraction). Because the fluorescence from the 2D(T) polymer of C_{60} appears in another energy region, this background is most probably related to the enhancement of strain and inhomogeneity within the sample induced at higher pressure.

The drastic changes in the Raman spectrum of the 2D(T) polymer are first observed at $P > 20$ GPa, where new distinct peaks appear in the spectrum and their intensities increase with a further increase in pressure. On the contrary, some of the initial Raman peaks of the polymer disappear above this critical pressure. At $P \geq 20$ GPa, the Raman spectrum of the material is significantly different from the initial one at lower pressure; the observed changes can be attributed to the transition of the polymer to a new high-pressure phase. From Fig. 1a, it is clear that, even for an applied pressure as high as 27.5 GPa, the Raman spectrum of the high-pressure phase is well resolved with relatively narrow peaks. Moreover, the frequency positions of the majority of the peaks in the new phase can be tracked back to the peaks observed in the initial 2D(T) polymeric phase of C_{60}. This is a first experimental indication that the C_{60} molecular cages are retained at pressures higher than 20 GPa, as the Raman peaks in the high-pressure

phase have their origin on intramolecular cage vibrations.

Figure 1b shows the Raman spectra of the material upon pressure release. The decrease in pressure from 27.5 GPa to ambient conditions results in the gradual shift of the Raman peaks to lower energies. The release of pressure does not lead to any observable changes in the Raman intensity distribution, and the high-pressure phase remains stable down to the ambient pressure. The bottom spectrum in Fig. 1b was recorded at 0.6 GPa, while the sample was recovered in air after disassembling the diamond anvil cell and tested again by means of micro-Raman probing. In Fig. 2, the Raman spectrum of the 2D(T) polymer recorded at ambient conditions (Fig. 2a) is illustrated in comparison with that of the high-pressure phase of the recovered material (Fig. 2b). The spectra in Fig. 2 were recorded outside the diamond anvil cell, and it was therefore attainable to also measure the spectrum of each material in the frequency region of the T_{2g} mode of diamond. The Raman spectrum of the high-pressure phase at ambient conditions is quite different from that of the initial 2D(T) polymeric phase of C_{60}. The peak positions for the 2D(T) polymer and those for the recovered high-pressure phase are shown in table at ambient conditions.

The recovered high-pressure phase of the 2D(T) polymer exhibits a metastable behavior. More specifically, after a time period of several days from the moment of its exposure to air, the recovered sample was detonated upon laser irradiation using the power 0.4 mW (objective: 100×) on the laser spot at the surface of the sample. The detonation is a rapid explosive process that was probably initiated by the thermal energy deposited by the probing laser beam; it results in a rapid relaxation of the built-in strain in the sample. We note that, under these conditions, the laser power density on the sample is higher than the density reaching the sample inside the high-pressure cell due to the different optical systems used (different magnification factors and laser spot diameters). In addition, thermal dissipation conditions are different in the two cases, because the sample is surrounded by the pressure-transmitting medium in the cell. After the detonation, two phases were identified among the pieces of the recovered sample, characterized by their completely different Raman spectra presented in Figs. 2c and 2d. In Fig. 2c, the spectrum of the main part of the detonated sample is illustrated. This Raman spectrum is similar to that expected from a mixture consisting of dimers and monomers of C_{60} [1, 2, 14]. The presence of this phase in the detonated sample definitely proves that the C_{60} molecular cages are retained in the high-pressure phase of the 2D(T) polymer. Finally, in Fig. 2d, the Raman spectrum of the phase that is a minority among the pieces of the detonated sample is given. As can be seen, the Raman spectrum of this phase is rather weak, consisting of two relatively broad peaks at 1342 and 1591 cm^{-1} (see table). We have recorded the spectrum of this phase after the sample detonation, but because the spectrum is too weak, we are not sure whether this phase was also present in the sample before its detonation, or even in the initial 2D(T) polymer before the application of pressure. Fullerite phases having Raman spectra similar to that in Fig. 2d have been observed in C_{60} treated at the pressure 12.5 GPa and temperature higher than 700°C [32], as well as in C_{60} treated at the pressure 9.5 GPa and temperature higher than 1500°C [33]. These phases were characterized by X-ray and microhardness studies as disordered carbon phases having high density and hardness [32–34] and were attributed to the breakdown of C_{60} molecular cages and the formation of a cross-linked structure of graphite-like layers [33]. Indeed, the Raman spectra of these carbon phases, as well as the spectrum presented in Fig. 2d, are similar to that of the amorphous carbon containing a significant amount of sp^2 bonded carbon atoms [35] and to those of the microcrystalline graphite or diamond-like carbon

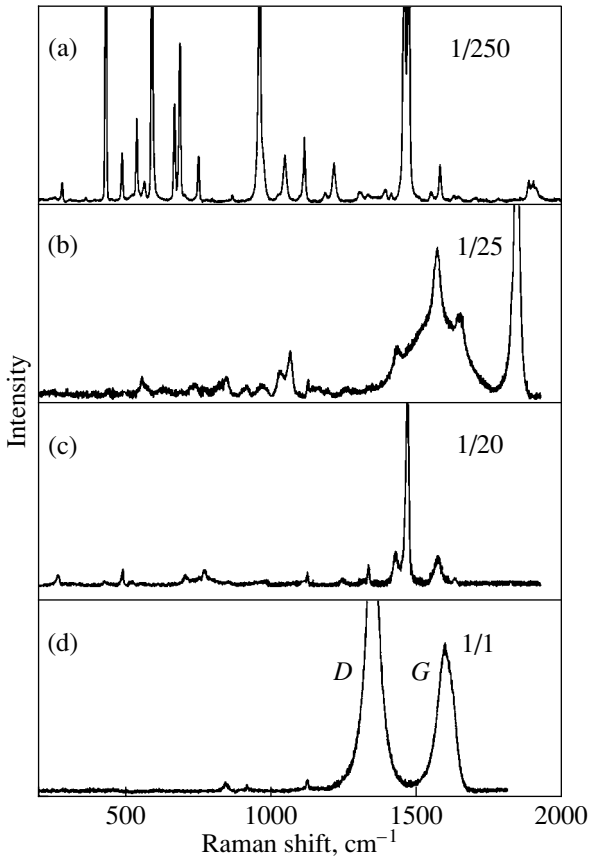


Fig. 2. Raman spectra of the initial 2D(T) polymer and the recovered high-pressure phase after pressure release, at ambient conditions. The numbers 1/*x* indicate the relative scale of the spectra. (a) The initial 2D(T) polymeric phase. (b) The high-pressure phase of the polymer. (c) The main component among the pieces of the detonated sample identified as a mixture of the C_{60} monomer and dimer. (d) The diamond-like carbon phase identified among the pieces of the detonated sample.

Phonon frequencies, pressure coefficients, and the Grüneisen parameters for the initial 2D(T) polymeric phase of C_{60}, the high-pressure phase, and pristine C_{60}. The phonon frequencies for the diamond-like carbon and dimeric C_{60} phases observed after the sample detonation are also included

Mode[a]	2D(T) polymeric C_{60}			High-pressure phase				Pristine C_{60}			Dimer C_{60}	Diamond-like phase
	ω_i, cm^{-1}	$\partial\omega_i/\partial P$, cm^{-1} GPa[b]	γ_i	ω_i, cm^{-1}	$\partial\omega_i/\partial P$, cm^{-1} GPa[b]		γ_i	ω_i, cm^{-1}	$\partial\omega_i/\partial P$, cm^{-1} GPa	γ_i	ω_i, cm^{-1}	ω_i, cm^{-1}
		$P < 4$ GPa			$P <$ 10 GPa	$P >$ 10 GPa		0.4 < P < 2.4 GPa				
$H_g(1)$	259	5.8	0.78	–	–		–	272	3.2	0.165	266	–
$H_g(1)$	282	2.3	0.264	297	0.3		0.047	294	2.5	0.119	–	–
$F_{2u}(1)$	363	–0.2	–0.019	391	–0.2		–0.024	345	2.9	0.118	–	–
$H_g(2)$	416	–0.1	–0.009	–	–		–	389	–0.2	–0.007	–	–
$H_g(2)$	432	0.6	0.049	442	0.6		0.064	435	2.4	0.077	427	–
$H_g(2)$	456	0.3	0.023	459	0.8		0.079	454	1.4	0.043	–	–
$A_g(1)$	487	4.5	0.322	–	–		–	495	4.2	0.119	489	–
$F_{1u}(1)$	–	–	–	–	–		–	522	1.4	0.027	523	–
$F_{2g}(1)$	536	1.4	0.091	540	0.6		0.052	–	–	–	–	–
$F_{1g}(1)$	563	1.4	0.087	554	1.3	0.8	0.111	563	0.8	0.02	–	–
$F_{1g}(1)$	588	0.8	0.047	571	1.8	0.9	0.148	–	–	–	–	–
$F_{1g}(1)$	–	–	–	634	2.7	0.9	0.201	624	1.5	0.034	–	–
$H_g(3)$	666	0.7	0.036	–	–		–	–	–	–	–	–
$H_g(3)$	683	2.3	0.118	688	1.4	1.0	0.096	–	–	–	–	–
$H_g(3)$	–	–	–	–	–		–	710	–0.8	–0.016	704	–
$H_g(4)$	748	–0.7	–0.033	738	1.9	1.3	0.121	729	–2.9	–0.056	–	–
$H_g(4)$	751[c]	–	–	–	–		–	755	–4.1	–0.078	–	–
$H_g(4)$	772	–1.2	–0.054	769	2.1	1.6	0.128	772	–2.7	–0.049	768	–
$H_g(4)$	773[c]	–	–	–	–		–	–	–	–	–	–
$H_g(4)$	–	–	–	826	2.4		0.137	–	–	–	–	–
$H_u(4)$	861	–0.6	–0.024	877	1.6		0.086	–	–	–	847	841
$H_u(4)$	–	–	–	902	2.1		0.109	–	–	–	–	–
$H_u(4)$	–	–	–	915	2.2		0.113	–	–	–	–	915
$G_g(2)$	955	4.5	0.164	961	3.0		0.147	–	–	–	956	–
$G_g(2)$	–	–	–	972	3.7		0.179	–	–	–	–	–
$F_{2u}(4)$	1041	4.2	0.141	1029	3.8		0.174	–	–	–	–	–
$F_{2u}(4)$	–	–	–	1064	2.8		0.124	–	–	–	–	–
$H_g(5)$	1107	4.8	0.151	–	–		–	–	–	–	–	–
$G_g(3)$	1178	6.7	0.198	–	–		–	–	–	–	–	–
$F_{2g}(3)$	1206	7.6	0.22	–	–		–	–	–	–	–	–
$H_g(6)$	–	–	–	–	–		–	–	–	–	1239	–
$F_{2u}(5)$	–	–	–	–	–		–	–	–	–	1328	*D*1342
$H_g(7)$	1403	6.6	0.164	–	–		–	1422	9.8	0.096	1420	–
$A_g(2)$	1448	6.1	0.147	1430	4.3		0.14	1467	5.5	0.053	1461	–
$F_{1g}(3)$	1464	7.6	0.181	–	–		–	–	–	–	–	–
$F_{2g}(4)$	1541	5.1	0.115	1509	3.9		0.119	–	–	–	–	–
$H_g(8)$	1572	5.9	0.131	1567	3.7		0.111	1570	4.8	0.043	1566	*G*1591
$G_g(6)$	1623	4.7	0.1	1647	4.1		0.117	–	–	–	1624	–
$G_g(6)$	–	–	–	1842	3.5		0.089	–	–	–	–	–

[a] The mode assignment refers to the irreducible representations of the icosahedral C_{60} molecule [40] and follows that in [14]; it is given here only for the initial phase of the 2D(T) polymer and the "dimeric" C_{60} phase.

[b] Data taken from [30].

[c] Frequency value at P = 6 GPa.

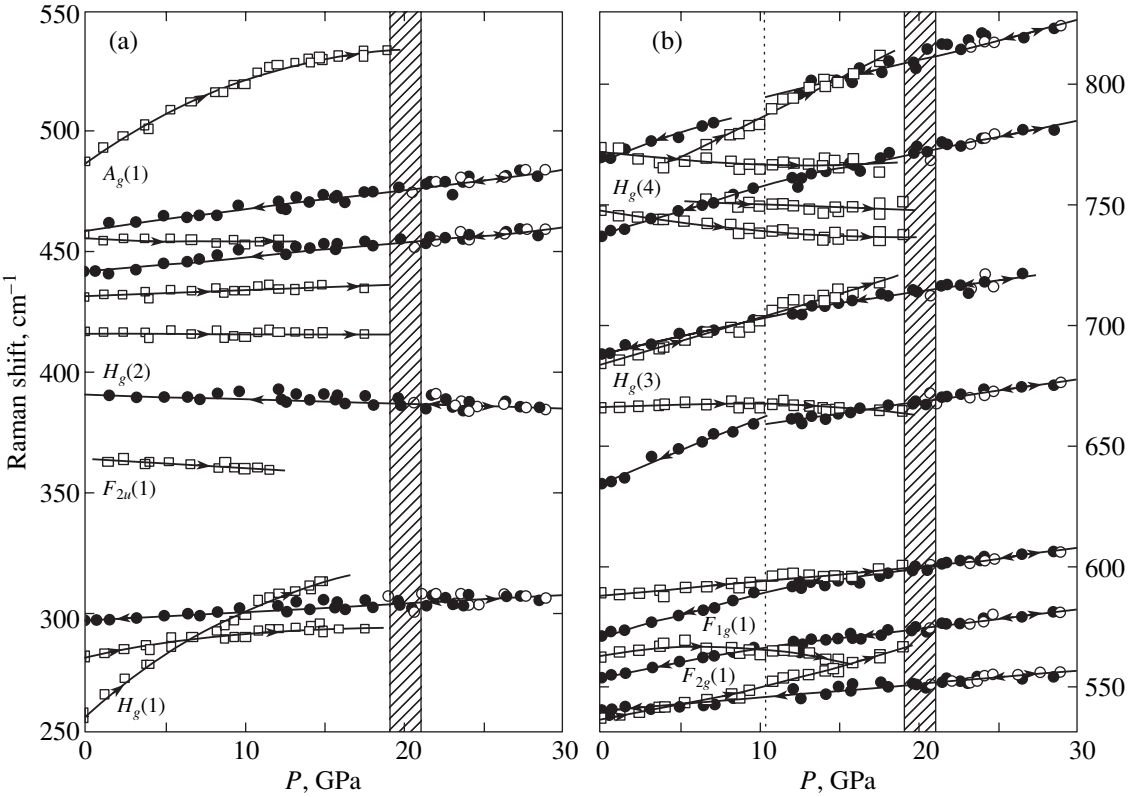


Fig. 3. The pressure dependence of the Raman modes of the 2D(T) polymeric phase of C_{60} in the frequency regions of 250–550 cm^{-1} (a) and 530–830 cm^{-1} (b). Squares and circles represent data taken for the 2D(T) polymer in the initial and high-pressure phases, respectively. The open (solid) symbols denote data taken for increasing (decreasing) pressure runs. The shaded area near 20 GPa denotes the pressure range of the phase transformation. The dotted vertical line near 10.4 GPa in Fig. 3b denotes the pressure where changes in the slope of the pressure dependence occur during pressure decrease.

films, mostly consisting of sp^3 hybridized carbon atoms [36, 37]. We therefore characterize the phase that is the minority among the pieces of the detonated sample as the diamond-like phase, and we ascribe the respective Raman peaks of this phase at 1342 cm^{-1} and 1591 cm^{-1} (Fig. 2d and table) to the *D* (diamond) and *G* (graphite) peaks of the microcrystalline graphite [38].

The obtained experimental data provide a strong indication that the 2D(T) polymer of C_{60} undergoes a phase transition above 20 GPa. The transformation occurs via an intermediate state having a rather diffuse Raman spectrum, which characterizes a highly disordered pretransitional state of the material at a pressure near 4 GPa. The fact that the prominent Raman peaks of the high-pressure phase are related to the retention of the C_{60} cages in this phase is an indication that the new phase of the polymer can be related to a three-dimensional (3D) polymerization of C_{60}. The observed peculiarities in the pretransitional pressure range also support the assumption of a further pressure-induced polymerization, which is a solid-state chemical reaction rather than a structural phase transformation. The Raman spectrum of the high-pressure phase is dominated by a very strong Raman peak around 1842 cm^{-1}, which cannot be related to any internal vibrational mode of the C_{60} molecular cage. The strong Raman peaks ranged near 1600–1900 cm^{-1} in some chemical compounds of carbon are related to the stretching vibrations of isolated double C=C bonds [39]. In analogy to that, the strong peak at 1842 cm^{-1} can be attributed to the destruction of a number of double C=C bonds during further polymerization of the 2D(T) polymer and to the appearance of some of the remaining ones as isolated C=C bonds in the 3D network of the C_{60} polymeric material. A more detailed analysis of the phonon modes and their pressure dependence in the initial 2D(T) polymer and in the high-pressure phase are discussed in the next subsection.

3.2. Phonon Modes

The pressure dependence of the Raman modes of the 2D(T) polymer of C_{60} in the initial phase (squares) and the high-pressure phase (circles) is shown in Figs. 3 and 4. The open (solid) symbols denote data taken for increasing (decreasing) pressure runs. Solid lines are drawn to guide the eye, and arrows indicate the pressure increase or decrease. In these figures, the

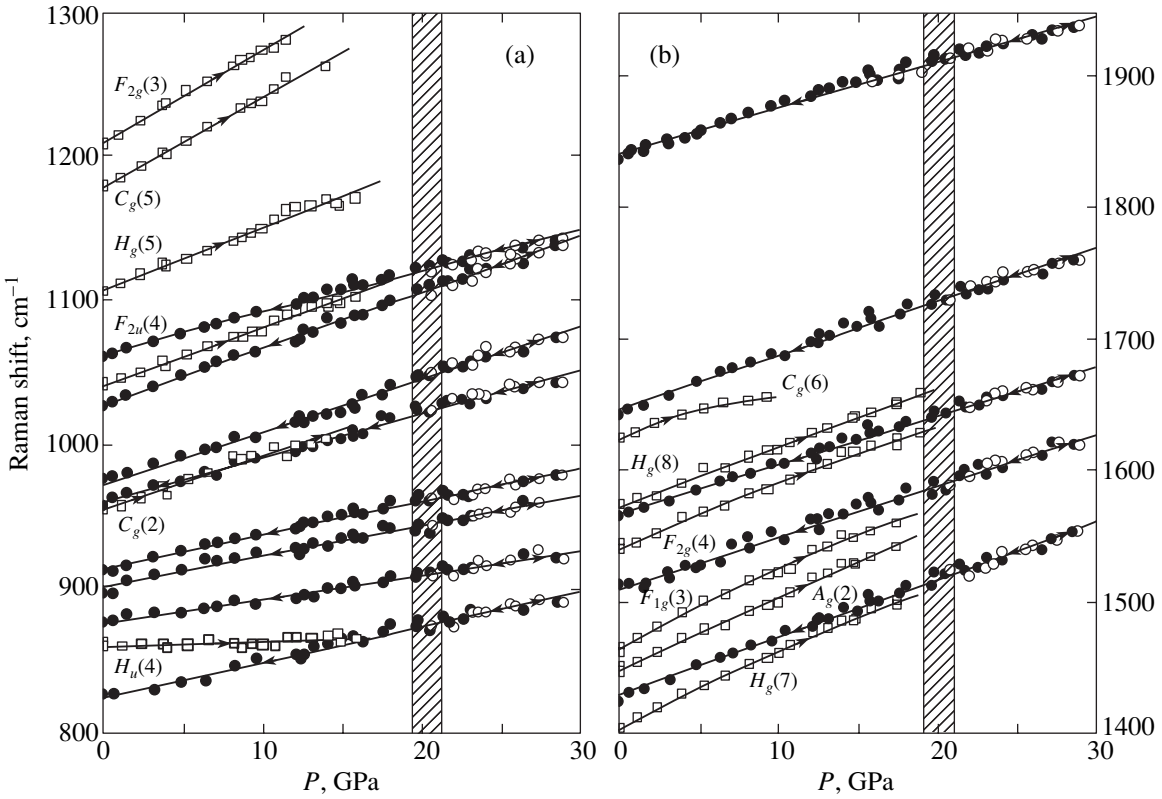


Fig. 4. The same as in Fig. 3, but for C_{60} in the frequency regions of 800–1300 cm^{-1} (a) and 1400–1950 cm^{-1} (b).

mode assignment refers to the irreducible representations of the parent C_{60} molecule (I_h symmetry) [40], following the notation in [14] in general, and is given here only for the initial 2D(T) phase of the polymer. Table contains a compilation of the mode assignment (given for the initial 2D(T) polymeric and the "dimeric" C_{60} phases), phonon frequencies ω_i, pressure coefficients $\partial\omega_i/\partial P$, and the corresponding Grüneisen parameters γ_i that in the present work are defined for the initial 2D(T) polymeric and high-pressure phases. The Grüneisen parameters

$$\gamma_i = -\frac{\partial\omega_i/\omega_i}{\partial V/V} = \frac{B_0}{\omega_i^0}\frac{\partial\omega_i}{\partial P}$$

were calculated using the experimental data of the pressure coefficients $\partial\omega_i/\partial P$ for the phonon modes in both phases of the 2D(T) polymer. The bulk modulus B_0 = 34.8 GPa for the initial 2D(T) polymeric phase was taken from [19]. Because the experimental data are absent for this material, we have also used the theoretical value of the bulk modulus B_0 = 47 GPa [12] for the calculation of the Grüneisen parameters in the high-pressure phase. We note that the values of γ_i for the high-pressure phase are only an estimate because the real value of B_0 can differ from the theoretically predicted one. The appropriate data of the phonon mode frequencies of pristine C_{60} and their pressure coefficients, and Grüneisen parameters calculated using the bulk modulus B_0 = 14.4 GPa [41] are included in table for comparison [42]. The last two columns in table contain the phonon mode frequencies for the two phases (diamond-like and dimeric C_{60}) observed after the sample detonation at ambient conditions.

As can be seen from Figs. 3 and 4, all the Raman peaks of the initial 2D(T) phase disappear in the pressure range $16 < P < 20$ GPa, while the Raman peaks related to the high-pressure phase gradually appear and gain in intensity above 20 GPa (the shaded area in Figs. 3 and 4 indicates the pressure range of the transformation). It is also clear that the majority of the Raman modes of the high-pressure phase are related to those of the 2D(T) polymer, showing that they originate from the C_{60} molecular cage vibrations. The nature of some phonon modes in the initial phase of the 2D(T) polymer of C_{60}, in particular, the Raman peak near 1040 cm^{-1}, is related to the covalent intermolecular bonding within the 2D polymeric layers [1, 14, 21]. More specifically, the peak near 1040 cm^{-1} is associated with the vibrations of the sp^3-like coordinated carbon atoms; the much lower frequency of this peak compared to that of the T_{2g} mode of diamond [28] can be attributed to the different lengths of the sp^3-like bonds in the 2D(T) polymer (1.64 Å) and diamond (1.54 Å). In the recovered high-pressure phase, this mode

appears to have two components with the frequencies 1029 and 1064 cm^{-1}. Assuming that the high-pressure phase is related to the formation of a 3D polymeric phase of C_{60} proposed by Okada *et al.* [12], we can associate these two Raman peaks with the existence of two types of sp^3-like coordinated carbon atoms with slightly different bond lengths.

Another important feature of the phonon spectrum of the high-pressure phase is the drastic changes in the region of the $A_g(2)$ pentagonal-pinch (PP) mode with respect to pristine C_{60} and its 2D(T) polymeric phase. The PP mode in pristine C_{60} is related to the in-phase stretching vibration of the five double C=C bonds originating in the vertices of each pentagon in the fullerene molecular cage. The frequency of the PP mode in the polymeric fullerenes decreases as the number of the polymeric covalent bonds per each molecular cage increases. Thus, the PP mode in the Raman spectrum of the dimeric fullerene C_{120} is disposed around 1462 cm^{-1} (from 1469 cm^{-1} in pristine C_{60}) and is then softened to 1457 cm^{-1} in 1D(O), to 1449 cm^{-1} in 2D(T), and further down to 1406 cm^{-1} in the 2D(R) polymer [14]. The additional increase in the number of sp^3-like coordinated carbon atoms in the 3D polymer results in more drastic changes in the PP-mode region. Namely, five strong peaks appear in the Raman spectrum of the high-pressure phase, with the most intense of them located near 1842 cm^{-1}. The breakdown of a large number of double C= C bonds in the high-pressure phase therefore leads to quenching of the PP mode; as a result, a number of new Raman peaks appear that are possibly related to the stretching vibrations of the remaining isolated double C=C bonds. It is well known that the stretching vibrations of the isolated double C=C bonds are Raman-active and their frequencies range up to 2000 cm^{-1} [39].

The pressure dependence of the phonon frequencies of the 2D(T) polymer shows a linear behavior for almost all modes and is absolutely reversible for pressures up to 12 GPa [30]. Nevertheless, two modes, $H_g(1)$ and $A_g(1)$, demonstrate a strong sublinear pressure dependence. The $A_g(1)$ mode is a breathing mode of the fullerene molecular cage and is associated with radial displacements in the atomic motions. To a large extent, the $H_g(1)$ mode is also related to the radial displacements of the carbon atoms. These two modes are therefore characterized by out-of-plane displacements of carbon atoms, and in our opinion, their sublinear pressure dependence can be associated with the high anisotropy related to the van der Waals intermolecular bonding of adjacent 2D polymeric layers and the covalent intermolecular bonding within the layers. Such a behavior is typical of the 2D polymeric phases of C_{60} and was also recently observed in the 2D(R) polymer [20]. In addition, the $A_g(1)$ mode completely disappears at $P \geq 20$ GPa and is not present in the high-pressure phase. Such a behavior can be the result of the 3D polymeric bonding in the high-pressure phase, which quenches the breathing vibration of the fullerene molecular cage.

In the high-pressure phase, the frequencies of all the observed modes increase with increasing the pressure, except for the peak at 391 cm^{-1}, which shows a small negative pressure slope (see table). The pressure coefficients $\partial\omega_i/\partial P$ of the Raman modes in the high-pressure phase range from –0.2 to 4.1 cm^{-1}/GPa, and the pressure slopes in the initial phase of the 2D(T) polymer range from –1.2 to 7.6 cm^{-1}/GPa. At the same time, the pressure coefficients of the Raman modes in pristine C_{60} range from –4.1 to 9.8 cm^{-1}/GPa. These data agree with the fact that the polymerized materials become harder as the degree of polymerization increases [1, 12, 41]. It is interesting to note that the pressure coefficients of the Raman peaks at 1029 and 1064 cm^{-1}, associated with the sp^3-like coordinated carbon atoms, are comparable to that of the T_{2g} mode of the crystalline diamond (3.8, 2.8, and approximately 2.7 cm^{-1}/GPa, respectively) [43]. Finally, it is important to note that several Raman modes of the high-pressure phase, located in the frequency region of 550–800 cm^{-1}, reveal changes in their pressure slopes to higher values as the pressure decreases below 10 GPa (see table and the dotted line in Fig. 3b). These changes in the pressure slopes can be related to the theoretically predicted relaxation of the tetragonal lattice parameters in the high-pressure phase after the pressure release. As shown in [12], the lattice parameter a of the high-pressure phase at normal conditions is enlarged with respect to that of the initial 2D(T) polymer from about 0.3 to 9.4 Å. We also think that the relaxation of the lattice parameter in the recovered high-pressure phase is responsible for the softening of the 1040 cm^{-1} mode in the initial 2D(T) polymer to 1029 cm^{-1} in the new high-pressure phase (the low-frequency split component).

In Fig. 5, we show the correlation of the Grüneisen parameters γ_i to the phonon mode frequencies ω_i. The respective data for pristine C_{60}, for the initial phase of the 2D(T) polymer of C_{60}, and for its high-pressure phase are represented by open triangles, circles, and squares. The single solid star indicates the Grüneisen parameter of the T_{2g} mode of diamond, which is as large as 0.895. The values of γ_i for the three materials investigated vary between –0.078 and 0.78; in general, they behave similarly to frequency, exhibiting two maxima near 600 and 1300 cm^{-1} and two minima near 400 and 750 cm^{-1}. Both minima are related to soft intramolecular modes and are characterized by negative values of γ_i, indicating the possible instability of the C_{60} molecular cage with respect to the atomic displacements related to these modes. We note that the polymerization of pristine C_{60} leads to a deformation of the fullerene molecular cage, which, in particular, results in the reduction of the number of modes with negative values of γ_i. This

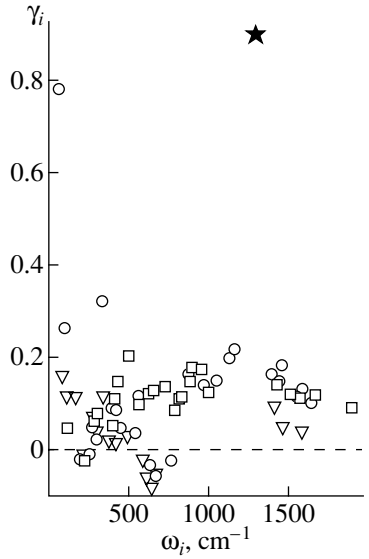


Fig. 5. Grüneisen parameters for the various modes of pristine C_{60} (triangles), the initial 2D(T) polymeric phase (circles), and its high-pressure phase (squares). The solid star indicates the Grüneisen parameter of the T_{2g} mode of diamond at 1332 cm^{-1}.

reduction becomes more significant in the fullerenes with a higher degree of polymerization. Another interesting observation is that the maximum in the frequency dependence of the Grüneisen parameters around the T_{2g} mode of diamond at 1332 cm^{-1} is more pronounced for the polymerized fullerenes. This implies, in particular, that the appropriate phonon modes of the polymeric fullerenes involve atomic motions of the sp^3-like coordinated carbon atoms as in the case of the T_{2g} mode of diamond.

We finally consider the very large values of γ_i for the $H_g(1)$ and $A_g(1)$ modes in the 2D(T) polymer. These modes are greatly affected by the van der Waals interaction between the 2D polymeric layers due to the out-of-plane nature of the relevant atomic motions. In addition, the deformation of the fullerene molecular cage related to the additional 3D polymeric bonding also contributes to the large values of γ_i for these two modes.

3.3. Photoluminescence and Electronic Spectrum

The photoluminescence spectra of the 2D(T) polymeric phase of C_{60}, for pressures up to 3 GPa and room temperature, recorded for various increasing and decreasing pressure runs, are shown in Figs. 6a and 6b, respectively. The inset in Fig. 6b shows the photoluminescence spectrum of the 2D(T) polymer in comparison with the photoluminescence spectra of pristine C_{60} at room temperature and at 10 K [44]. The fluorescence intensity in the 2D(T) polymer is noticeably higher than in pristine C_{60}. This is related to the fact that the lowest excited singlet state of the C_{60} molecule has the $^1T_{1g}$

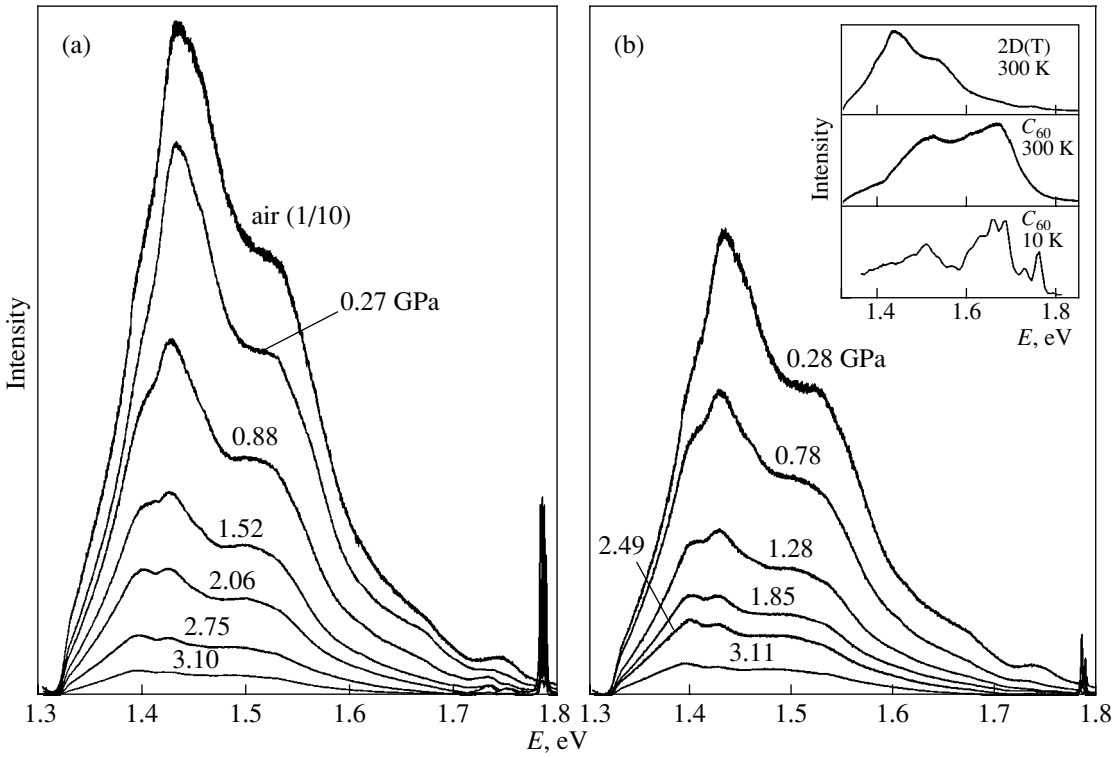


Fig. 6. Photoluminescence spectra of the 2D(T) polymer at 300 K and various pressures for the increasing (a) and decreasing (b) pressure runs. Inset: photoluminescence spectra of pristine C_{60} at 10 and 300 K in comparison with that of the 2D(T) polymer at room temperature.

symmetry and optical transitions to this state are dipole-forbidden [45, 46]. In solutions, the fluorescence of C_{60} is related to the vibronically assisted Herzberg–Teller transitions [47], whereas, in the solid state, the additional increase in the fluorescence quantum yield is related to the violation of the symmetry restrictions by the presence of crystal impurities and defects. The polymerization reduces the molecular symmetry from I_h in pristine C_{60} to D_{2h} in the 2D(T) polymer and eliminates the symmetry restrictions leading to a substantial increase in the photoluminescence quantum yield.

As can be seen from the inset in Fig. 6b, the photoluminescence spectrum of the 2D(T) polymer differs from that of pristine C_{60} in both the intensity distribution and the onset of luminescence. The well-structured low-temperature photoluminescence spectrum of the high-quality single crystals of C_{60} is mainly related to excitons localized on defects or impurity levels [44, 48]. On the contrary, the photoluminescence spectrum of C_{60} at room temperature consists of two broad bands and is related to the vibronically assisted transitions from the lowest singlet $^1T_{1g}$ level (the first band at about 1.665 eV). The most intense bands in the photoluminescence spectrum of the 2D(T) polymer located near 1.533 and 1.435 eV are related to the fluorescence of the host 2D(T) polymer, while the weak shoulders at higher energies (near 1.748 and 1.661 eV) originate from the impurity phase of 2D(R) that is present in the 2D polymeric samples [49]. We note that the very sharp lines in Figs. 6a and 6b near 1.785 eV are related to the R_1 and R_2 luminescence bands of the ruby chips used for pressure calibration [27].

The increase in pressure results in a gradual shift of the photoluminescence spectrum to lower energies, lowering and redistributing the fluorescence intensity. These changes are absolutely reversible in the pressure range investigated ($P \le 4$ GPa), as can be seen in Fig. 6. The integrated intensity of the photoluminescence spectrum rapidly decreases with the increase in pressure and recovers its strength nearly to its initial value after the total pressure release.

The pressure dependence of the band positions in the photoluminescence spectrum of the 2D(T) polymer is shown in Fig. 7. The open (closed) triangles, squares, circles, and diamonds are related to the bands near 1.748, 1.661, 1.533, and 1.435 eV for the upstroke (downstroke) pressure runs, respectively. The open hexagons show the pressure dependence of the first peak in the fluorescence spectrum of pristine C_{60} [44] and are included in the figure for comparison. The pressure-induced shift $\partial E/\partial P$ for the photoluminescence bands of the 2D(T) polymer varies from –9 to –17 meV/GPa, whereas the corresponding value for pristine C_{60} is equal to –78 meV/GPa. The large difference in $\partial E/\partial P$ leads to the intersection of the pressure dependences for the 2D(T) polymer (circles) and pristine C_{60} (hexagons) near 1.9 GPa, which is close to 2.2 GPa used in the HPHT treatment procedure for the production of the 2D(T) polymer. This implies that the fluorescence onset in the 2D(T) polymer is shifted to lower energies mainly because of the decrease in the intermolecular distances caused by the formation of the polymer. On the contrary, the deformation of the C_{60} cage in the polymer, which leads to the lowering of molecular symmetry, does not significantly affect the shift of the electronic spectrum.

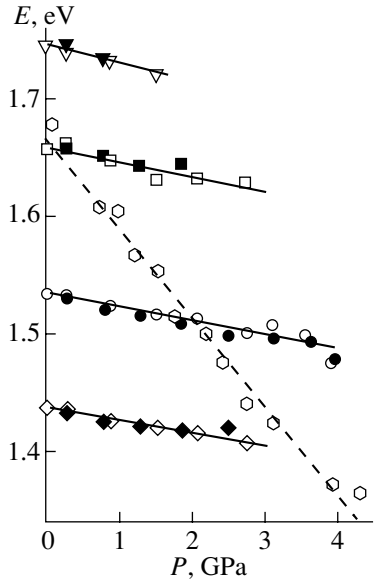


Fig. 7. Pressure dependence of the fluorescence bands in the photoluminescence spectra of the 2D(T) polymer and pristine C_{60}. Open (closed) triangles, squares, circles, and diamonds show various bands of the 2D(T) polymer for the increasing (decreasing) pressure runs. Open hexagons show the first fluorescence band of pristine C_{60} for an upstroke pressure run.

We can calculate the deformation potential

$$D = \frac{\partial E_g}{\partial \ln(V_0/V)} = -B_0 \frac{\partial E_g}{\partial P},$$

where E_g is the direct gap and B_0 is the bulk modulus of the material, using the experimental data concerning the pressure-induced shift of the photoluminescence bands, $\partial E/\partial P$, obtained for both materials. Taking into account that the position of the first band in the photoluminescence spectrum coincides with the direct gap (in the case of vibronically assisted transitions, it differs from the phonon energy), we obtain D = 0.42 eV and D = 1.09 eV for the 2D(T) polymer and pristine C_{60}, respectively. It is interesting to note that a noticeable increase in the bulk modulus in the 2D(T) polymer compared to pristine C_{60} does not result in a similar increase in the deformation potential due to the lowering of the pressure-induced shift of the electronic spectrum.

4. CONCLUSIONS

The Raman scattering data under high hydrostatic pressure show that an irreversible transformation occurs in the 2D(T) polymeric phase of C_{60} above 20 GPa. The new phase is preceded by a pretransitional state characterized by diffuse Raman peaks. The spectrum of the high-pressure phase remains intense and well resolved at pressures as high as 30 GPa. The phonon modes of the high-pressure phase, especially in the high-energy region, are noticeably different from those of the initial 2D(T) polymer; nevertheless, they can be tracked back to the phonon modes related to the fullerene molecular cage. The recovered high-pressure phase is metastable and detonates under laser irradiation. The main part of the detonated sample is a mixture of monomeric and dimeric C_{60}, showing that the fullerene molecular cages are retained in the high-pressure phase. The high-pressure phase seems to be related to further creation of covalent bonds between molecules belonging to the adjacent polymeric layers in accordance with the theoretically predicted 3D polymerization of the 2D(T) C_{60} polymer at 20.2 GPa [12]. Our Raman experiments reveal that the 3D polymeric C_{60} resulting from the application of high pressure on the 2D(T) polymer is not related to the previously observed ultrahard fullerite phases [1, 8, 9]. The electronic spectrum of the 2D(T) polymer is noticeably different from that of pristine C_{60}. This difference is related both to the deformation of the fullerene molecular cage caused by the polymerization of material and to the decrease in the in-plane intermolecular distances in the 2D(T) polymer.

ACKNOWLEDGMENTS

The support by the General Secretariat for Research and Technology, Greece, is gratefully acknowledged. One of the authors (K.P.M.) acknowledges the support by the Russian State Research Program "Physical Properties of Carbon-Based Nanostructures and Development of New Electronic Devices." Another author (B.S.) acknowledges the support from the Swedish Research Councils for Natural Sciences (NFR) and Engineering Sciences (TFR).

The authors thank A. Soldatov and T. Wägberg for help with preparation and characterization of the 2D(T) polymer of C_{60}.

REFERENCES

1. B. Sundqvist, Adv. Phys. **48**, 1 (1999).
2. A. M. Rao, P. Zhou, K.-A. Wang, *et al.*, Science **259**, 955 (1993).
3. P. W. Stephens, G. Bortel, G. Faigel, *et al.*, Nature (London) **370**, 636 (1994).
4. K. Prassides, K. Vavekis, K. Kordatos, *et al.*, J. Am. Chem. Soc. **119**, 834 (1997).
5. Y. Iwasa, T. Arima, R. M. Fleming, *et al.*, Science **264**, 1570 (1994).
6. M. Nunez-Regueiro, L. Marques, J.-L. Hodeau, *et al.*, Phys. Rev. Lett. **74**, 278 (1995).
7. V. V. Brazhkin, A. G. Lyapin, and S. V. Popova, Pis'ma Zh. Éksp. Teor. Fiz. **64**, 755 (1996) [JETP Lett. **64**, 802 (1996)].
8. V. D. Blank, M. Yu. Popov, S. G. Buga, *et al.*, Phys. Lett. A **188**, 281 (1994).
9. V. D. Blank, S. G. Buga, G. A. Dubitsky, *et al.*, Carbon **36**, 319 (1998).
10. L. Marques, M. Mezouar, J.-L. Hodeau, *et al.*, Science **283**, 1720 (1999).
11. L. A. Chernozatonskii, N. R. Serebryanaya, and B. N. Mavrin, Chem. Phys. Lett. **316**, 199 (2000).
12. S. Okada, S. Saito, and A. Oshiyama, Phys. Rev. Lett. **83**, 1986 (1999).
13. E. Burgos, E. Halac, R. Weht, *et al.*, Phys. Rev. Lett. **85**, 2328 (2000).
14. V. A. Davydov, L. S. Kashevarova, A. V. Rakhmanina, *et al.*, Phys. Rev. B **61**, 11936 (2000).
15. G. A. Kourouklis, S. Ves, and K. P. Meletov, Physica B (Amsterdam) **265**, 214 (1999).
16. J. Arvanitidis, K. Papagelis, I. Tsilika, *et al.*, Physica B (Amsterdam) **265**, 234 (1999).
17. K. P. Meletov, J. Arvanitidis, I. Tsilika, *et al.*, Phys. Rev. B **63**, 054106 (2001).
18. K. P. Meletov, S. Assimopoulos, I. Tsilika, *et al.*, Chem. Phys. Lett. **341**, 435 (2001).
19. J. M. Leger, J. Haines, V. A. Davydov, and V. Agafonov, Solid State Commun. **121**, 241 (2002).
20. K. P. Meletov, J. Arvanitidis, G. A. Kourouklis, *et al.*, submitted to Chem. Phys. Lett. (2002).
21. T. Wågberg, A. Soldatov, and B. Sundqvist, unpublished.
22. R. Moret, P. Launois, T. Wågberg, *et al.*, Eur. J. Phys. B **15**, 253 (2000).
23. K. P. Meletov, E. Liarokapis, J. Arvanitidis, *et al.*, Chem. Phys. Lett. **290**, 125 (1998).
24. P.-A. Persson, P. Jacobsson, S. Stafstrom, *et al.*, Europhys. Lett. **49**, 631 (2000).
25. P.-A. Persson, U. Edlund, P. Jacobsson, *et al.*, Chem. Phys. Lett. **258**, 540 (1996).
26. A. Jayaraman, Rev. Sci. Instr. **57**, 1013 (1986).
27. D. Barnett, S. Block, and G. J. Piermarini, Rev. Sci. Instr. **44**, 1 (1973).
28. S. A. Solin and A. K. Ramdas, Phys. Rev. B **1**, 1687 (1970).
29. V. A. Davydov, L. S. Kashevarova, A. V. Rakhmanina, *et al.*, Phys. Rev. B **58**, 14786 (1998).
30. J. Arvanitidis, K. P. Meletov, K. Papagelis, *et al.*, J. Chem. Phys. **114**, 9099 (2001).
31. K. P. Meletov, D. Christofilos, S. Ves, *et al.*, Phys. Rev. B **52**, 10090 (1995).
32. V. V. Brazhkin, A. G. Lyapin, S. V. Popova, *et al.*, Phys. Rev. B **56**, 11465 (1997).
33. V. D. Blank, S. G. Buga, N. R. Serebryanaya, *et al.*, Carbon **36**, 665 (1998).
34. V. V. Brazhkin, A. G. Lyapin, S. V. Popova, *et al.*, J. Appl. Phys. **84**, 219 (1998).

35. M. Weiler, S. Sattel, T. Giessen, *et al.*, Phys. Rev. B **53**, 1594 (1996).
36. R. J. Nemanich and S. A. Solin, Phys. Rev. B **20**, 392 (1979).
37. R. E. Shroder, R. J. Nemanich, and J. T. Glass, Phys. Rev. B **41**, 3738 (1990).
38. M. S. Dresselhaus, M. A. Pimenta, P. C. Eklund, *et al.*, in *Raman Scattering in Material Science*, Ed. by W. H. Weber and R. Merlin (Springer-Verlag, Berlin, 2000), p. 314.
39. D. A. Long, *Raman Spectroscopy* (McGraw-Hill, London, 1976), p. 158.
40. M. C. Martin, X. Du, J. Kwon, and L. Mihaly, Phys. Rev. B **50**, 173 (1994).
41. J. Haines and J. M. Leger, Solid State Commun. **90**, 361 (1994).
42. K. P. Meletov, G. Kourouklis, D. Christofilos, *et al.*, Zh. Éksp. Teor. Fiz. **108**, 1456 (1995) [JETP **81**, 798 (1995)].
43. A. F. Goncharov, I. N. Makarenko, and S. M. Stishov, Pis'ma Zh. Éksp. Teor. Fiz. **41**, 150 (1985) [JETP Lett. **41**, 184 (1985)].
44. K. P. Meletov and V. D. Negriĭ, Pis'ma Zh. Éksp. Teor. Fiz. **68**, 234 (1998) [JETP Lett. **68**, 248 (1998)].
45. F. Negri, G. Orlandi, and F. Zerbetto, Chem. Phys. Lett. **144**, 31 (1988).
46. F. Negri, G. Orlandi, and F. Zerbetto, J. Chem. Phys. **97**, 6496 (1992).
47. G. Herzberg and E. Teller, Z. Phys. Chem., Abt. B **21**, 410 (1933).
48. W. Guss, J. Feldman, E. O. Gobel, *et al.*, Phys. Rev. Lett. **72**, 2644 (1994).
49. K. P. Meletov, S. Assimopoulos, and G. A. Kourouklis, unpublished.

Journal of Experimental and Theoretical Physics, Vol. 95, No. 4, 2002, pp. 748–752.
Translated from Zhurnal Éksperimental'noĭ i Teoreticheskoĭ Fiziki, Vol. 122, No. 4, 2002, pp. 863–868.

SOLIDS
Structure

Phase Transformations in Perovskites $TbBaCo_{2-x}Fe_xO_{5+\gamma}$

I. O. Troyanchuk[a, *], A. N. Chobot[a], D. D. Khalyavin[a], R. Szymczak[b], and H. Szymczak[b]

[a] *Institute of Solid-State and Semiconductor Physics, Belarussian Academy of Sciences, ul. Brovki 17, Minsk, 220072 Belarus*

[b] *Institute of Physics, Polish Academy of Sciences, 02-668 Warsaw, Poland*

**e-mail: troyan@ifttp.bas-net.by*

Received May 6, 2002

Abstract—The transport, magnetic, and elastic properties of $TbBaCo_{2-x}Fe_xO_{5+\gamma}$ are investigated. It is shown that these compounds exhibit first-order metal–insulator and antiferromagnet–weak ferromagnet transitions in the orthorhombic phase ($x < 0.12$), while these transitions are not observed in the tetragonal phase ($x > 0.12$). In the concentration range corresponding to the orthorhombic phase, doping with iron stabilizes the weakly ferromagnetic phase. However, the tetragonal phase is antiferromagnetic. Oxygen vacancies are assumed to be ordered in the orthorhombic case and disordered in the tetragonal phase. An analysis of Young's modulus, magnetostriction, and effects of pressure and substitution of the O^{18} oxygen isotopes for O^{16} indicates a weak correlation between magnetic transformations and the crystal lattice. © *2002 MAIK "Nauka/Interperiodica".*

1. INTRODUCTION

The interest in compounds of the type $LnBaCo_2O_{5+\gamma}$ (Ln stands for a lanthanide) is due to a clearly manifested relation between the magnetic state and the electrical conductivity as well as the existence of transitions of the type metal–insulator and charge or orbital ordering. These compounds were synthesized recently [1–4] and immediately attracted attention of many researchers [5–9]. It was found as a result of X-ray and electron-microscopic studies that Ln and Ba ions are ordered in such a way that the planes occupied by Ba and Ln ions alternate in the [100] direction. The structure is perovskite-like and is a derivative of a structure of the type $YBa(CuFe)O_5$ [10]. The oxygen content may vary from 5 to 5.6 depending on the method of obtainment [11]. The samples of $LnBaCo_2O_{5.5}$ (Ln = Gd, Tb, Y, or Ho) exhibit the following sequence of transitions upon heating: spontaneous magnetization emerges at T_i and vanishes at T_N, and a metal–insulator transition takes place at T_{MI}. At T_i, a jump in the electrical conductivity and a strong magnetoresistive effect are observed [2]. The nature of the phase transformations is disputable. It was proposed [6, 7] that, at T_i and T_{MI}, the spin state of cobalt ions changes. Further investigations are required to explain the mechanism of these transformations. It is known that the properties of manganites can be modified considerably by substituting other ions for the manganese ions [12]. This work is devoted to investigation of cobaltite samples in which some of the cobalt ions are replaced by iron ions. The choice of iron ions was dictated by the fact that the properties of disordered perovskites $La_{1-x}Sr_x(Co_{1-y}Fe_y)O_{3-\gamma}$ and $SrCo_{1-x}Fe_xO_{3-\gamma}$ are quite sensitive to this type of substitution [13].

2. EXPERIMENT

The samples were synthesized from a mixture of high-purity oxides and carbonates Tb_6O_{11}, $BaCO_3$, Co_3O_4, and Fe_2O_3 weighed in the required proportion. After preliminary annealing at 900°C, the pellets were thoroughly crushed in an agate mortar. The synthesis was carried out in air at 1150°C. The oxygen concentration was optimized by cooling at a rate of 100°C/h. The X-ray diffraction analysis was carried out on a DRON-3 diffractometer with the K_α radiation of Co. The oxygen concentration was monitored with the help of chemical and thermogravimeric analyses. Magnetostriction was measured by strain gauges. The magnetic measurements were made on a commercial Foner-type vibrational magnetometer. The electrical conductivity was measured by the standard four-probe method. The contacts were formed by ultrasonic deposition of indium. The elastic properties were measured by the resonance method in the acoustic frequency range on cylindrical samples having a length of 40–50 mm and a diameter of 5 mm. The substitution of O^{18} for O^{16} isotopes was carried out for 48 h at 1100°C with a cooling rate of 100°C/h. A sample subjected to thermal treatment under the same conditions in an O^{16} flow was used for comparison of properties.

3. RESULTS AND DISCUSSION

X-ray phase analysis revealed that the samples of $TbBaCo_{2-x}Fe_xO_{5+\gamma}$ contain only one phase up to $x = 1$. Compounds with $x \le 0.11$ are orthorhombically distorted (space group *Pmmm*), while the same samples are tetragonal (space group *P*4/*mmm*) for $x > 0.12$. A sample with $x = 0.12$ is a mixture of the tetragonal and orthorhombic phases. According to the results of chem-

ical and thermogravimetric analyses, the oxygen content in samples gradually decreases upon a substitution of iron ions for cobalt ions. The oxygen content in a sample with $x = 1$ corresponds to $\gamma = 0$. The results of our recent Mössbauer experiments on this system with the Fe^{57} isotope indicate that iron ions are in the trivalent state. It was observed in [10] that iron ions in the $YbaCuFeO_5$ system are also trivalent, while copper ions are in the bivalent state. For this reason, we assume that the substitution of Fe^{3+} ions for Co^{3+} ions leads to a transition of a part of Co^{3+} ions to the bivalent state and to the emergence of oxygen vacancies. Apparently the configuration $Fe^{3+} + Co^{2+}$ in the system $TbBaCo_{2-x}Fe_xO_{5+\gamma}$ is thermodynamically more stable than $Fe^{3+} + Co^{3+}$.

Figure 1 shows the temperature dependence of magnetization measured for samples with different iron contents. The sample with $x = 0$ exhibits spontaneous magnetization in the temperature range 250–280 K. The hysteresis for measurements in the heating and cooling regimes does not exceed two degrees; the substitution of iron ions for cobalt ions clearly increases the temperature range of spontaneous magnetization, and a considerable temperature hysteresis appears. For a sample with $x = 0.1$, it was found that spontaneous magnetization emerges and vanishes at $T_i = 195$ K and $T_N = 302$ K, respectively, upon heating and at $T_i = 160$ K and $T_N = 302$ K, respectively, upon cooling. The maximum value of spontaneous magnetization for samples with $x = 0$ and $x = 0.1$ is approximately the same and corresponds to $(0.15–0.17)\mu_B$ (per cobalt ion). Compounds with $x \geq 0.15$ do not exhibit spontaneous magnetization in the temperature range 5–350 K.

In the vicinity of temperatures T_i, T_N, and T_{MI}, the temperature dependence of electrical conductivity exhibits anomalies (Fig. 2). The electrical conductivity is doubled jumpwise in the vicinity of T_i, while near T_{MI} it changes by an order of magnitude. In the region of T_N, the anomalous behavior is manifested weakly. The substitution of Fe ions for Co ions in the orthorhombic phase leads to an insignificant variation in the electrical conductivity anomalies at T_i and T_{MI}.

In the vicinity of the transition T_i, a strong magnetoresistive effect is observed for the compound with $x = 0$, the magnitude of the effect decreasing gradually upon the substitution of iron ions for cobalt ions (Fig. 2). In the vicinity of T_N, the magnetoresistance peak is pronounced much more weakly.

Figure 3 shows the results of measurements of magnetization in compounds with $x = 0$ in the heating regime. It can be seen that an applied pressure shifts the temperatures T_i and T_N by 2–3 degrees. An effect of approximately the same magnitude was observed for a sample in which O^{16} atoms were replaced by the O^{18} isotope.

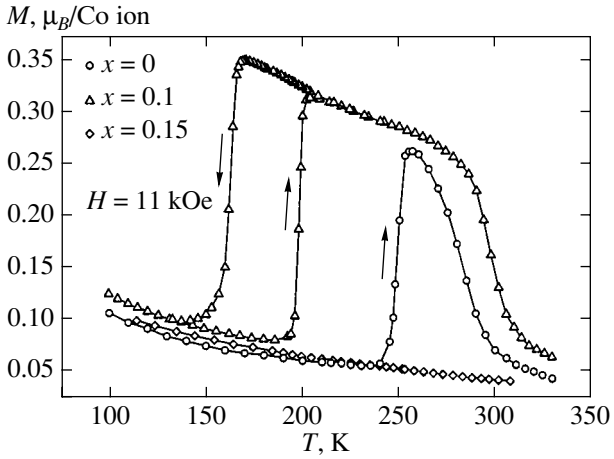


Fig. 1. Temperature dependence of magnetization for $TbBaCo_{2-x}Fe_xO_{5+\gamma}$ compounds.

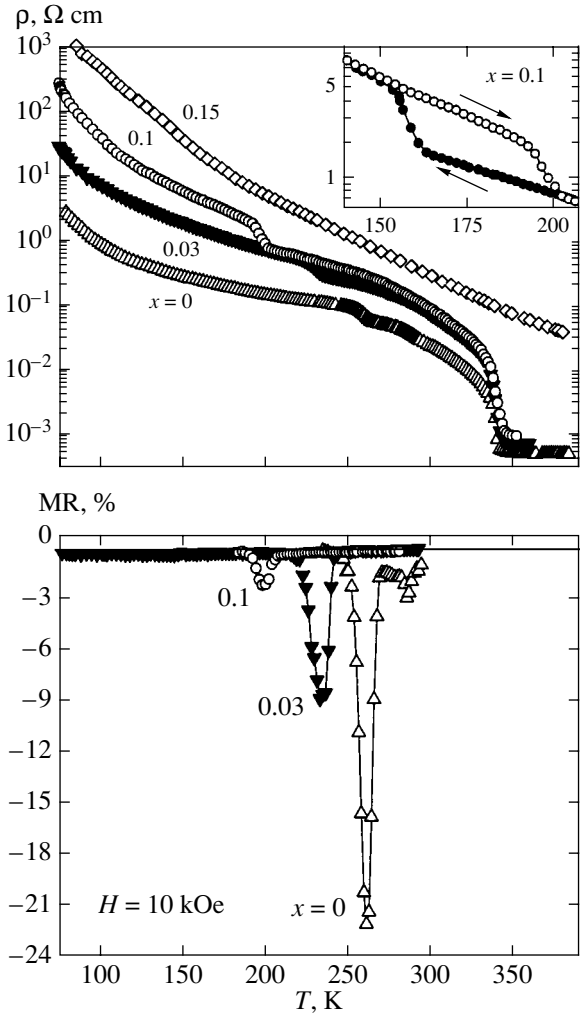


Fig. 2. Temperature dependences of electrical conductivity and magnetoresistance for $TbBaCo_{2-x}Fe_xO_{5+\gamma}$ compounds.

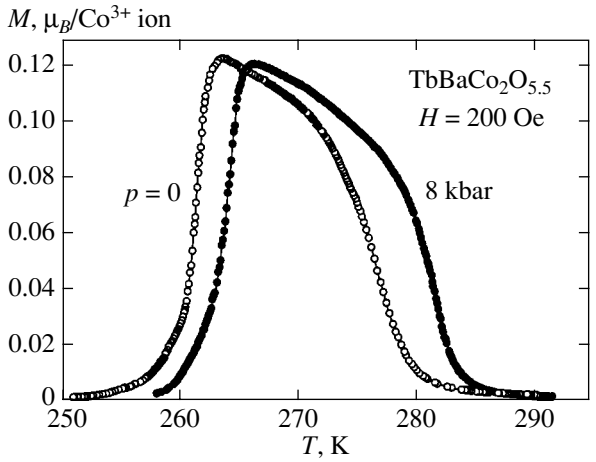


Fig. 3. Temperature dependences of magnetization measured at $x = 0$ and pressures $P = 0$ and 8 kbar.

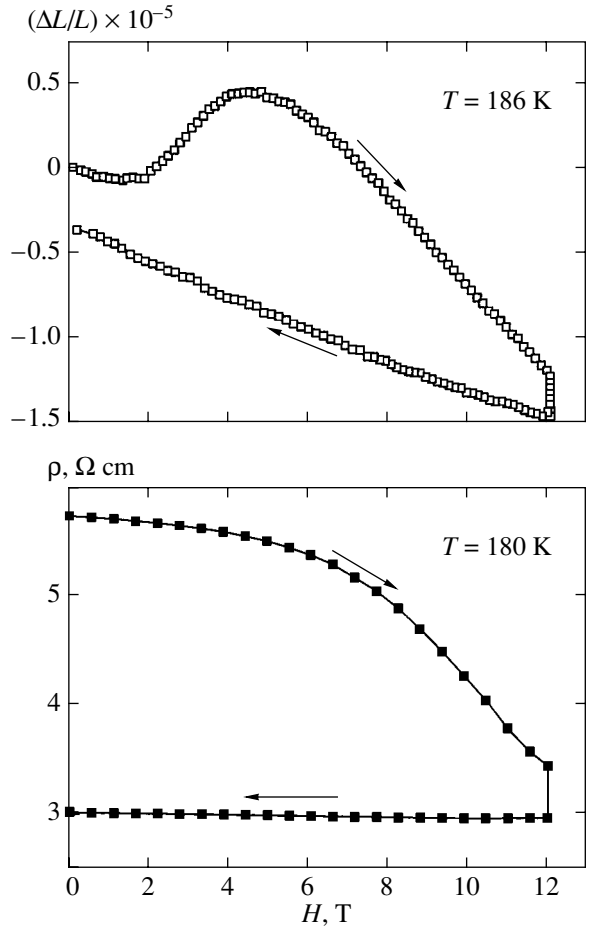


Fig. 4. Magnetic field dependences of magnetostriction and electrical conductivity for $x = 0.1$.

Figure 4 presents the results of magnetostriction measurements for a sample with $x = 0.1$ in the temperature range in which the magnetic ground state depends on the magnetic past history. At first, the sample was cooled to 80 K and then heated to 186 K, i.e., below the temperature corresponding to the beginning of transition to the spontaneous magnetization state. The external magnetic field induces a transition to the state with spontaneous magnetization, which is accompanied by an increase in the sample length. However, the change in the linear size is relatively small (of the order of 10^{-5}). The transition is irreversible in temperature as expected, since the measurements were made in the temperature range of phase bistability. The electrical conductivity was measured as a function of the magnetic field at $T = 180$ K (Fig. 4). The magnetic transition leads to a decrease in resistance. In the vicinity of the transition, a strong dependence of the resistance on the delay time was observed. It should be noted that strong relaxation effects were also observed in the case of metamagnetic transitions from the charge-ordered antiferromagnetic phase to the ferromagnetic phase in manganites [14].

It was shown in [2, 3] that the magnetic transition at T_i is noticeably displaced to the low-temperature region upon the application of an external field. This effect is responsible for the large magnetoresistance. However, the substitution of iron ions for cobalt ions leads to stabilization of the boundaries of different magnetic phases (Fig. 5). It was found that an external magnetic field of 5 T shifts the high-temperature phase boundary only by 6 degrees, which is almost an order of magnitude smaller than in the case of compounds that are not doped with iron. The *H–T* magnetic phase diagram for compounds with $x = 0.11$ in fields up to 5 T is presented in Fig. 5.

The results of analysis of Young's modulus are shown in Fig. 6. The temperature dependence of the square of resonance frequency for a sample with $x = 0$ exhibits a clearly manifested anomaly at the temperature of transition from the insulator state to the metal state in the vicinity of $T_{MI} = 345$ K. A temperature hysteresis is observed in the vicinity of 4 K, which coincides with the results obtained from electrophysical measurements. Near T_i, there exists a weak anomaly incompatible with the crystal-structure phase transformation. The absence of crystal-structure phase transformations at T_i was also confirmed by the results of low-temperature X-ray diffraction studies.

Let us discuss the experimental results obtained from an analysis of phase transition in the $TbBaCo_{2-x}Fe_xO_{5+\gamma}$ system.

At T_i, a magnetic moment appears, which is equal to $0.15–0.17\mu_B$ per Co^{3+} ion in the high-temperature phase. This value is an order of magnitude smaller than expected for ferromagnetic ordering of Co^{3+} ions in the intermediate spin state. The magnetic moment for the intermediate spin state of a Co^{3+} ion must be equal approximately to $2\mu_B$ per Co^{3+} ion. It is well known that the superexchange interaction Co^{3+} (intermediate spin state)–O–Fe^{3+} is strong and negative [13, 15]. An

increase in T_N upon the substitution of iron for cobalt is in accord with this fact. This means that the magnetic structure in the phase with spontaneous magnetization is mainly antiferromagnetic. In our opinion, such a behavior of magnetic properties can be expected in the case of a Morin-type transition, i.e., a transition from the antiferromagnetic to weakly ferromagnetic Dzyaloshinski–Moriya state A similar transition was observed in $DyFeO_3$ [16]. It should be noted, however, that the magnetic moment of classical weak ferromagnets of the type of orthoferrites, orthochromites, and orthomanganites is smaller by a factor of 2–3 as a rule.

In the tetragonal phase, spontaneous magnetization disappears in spite of the fact that the weakly ferromagnetic phase is gradually stabilized upon an increase in the iron concentration in the concentration range corresponding to the orthorhombic phase. In our opinion, this means that the magnetic structure in the tetragonal phase changes. It was established earlier with the help of electron microscopy that oxygen vacancies in orthorhombic $TbBaCo_2O_{5.5}$ are ordered and form a superstructure of the $2 \times 1 \times 1$ type [4]. The tetragonal phase of $TbBaCo_2O_{5.35}$ is also characterized by vacancy ordering and a superstructure of the $3 \times 3 \times 1$ type [17]. In the $3 \times 3 \times 1$ phase, the weakly ferromagnetic state is stable in the entire temperature range of magnetic ordering. For this reason, we assume that oxygen vacancies in the tetragonal phase with an iron content $x > 0.12$ are disordered, which leads to disappearance of the weakly ferromagnetic state.

The transition at T_i is associated with electrical conductivity. The magnetoresistance is due to the shift of T_i under the action of a magnetic field. The absence of strong anomalies in elastic properties and unit cell parameters and a weak magnetostriction indicate that neither the crystal structure nor the orbital state of cobalt ions changes at T_i. In all probability, a relatively narrow 3*d* band in the metal phase above T_{MI}, which is formed predominantly by t_{2g} orbitals of cobalt, and the broad 2*p* band of oxygen overlap. We believe that the conductivity of this class of compounds below T_{MI} is due to excitation of charge carriers from the broad 2*p* band to the closely spaced 3*d* band. At T_i, the energy gap between the bands decreases due to the emergence of spontaneous magnetization.

The small change in the electrical conductivity and the small magnetoresistance at T_N are unexpected since clearly manifested anomalies in magnetotransport properties were observed at T_i. This behavior sharply contradicts the properties of magnetoresistive manganites, which exhibit a strong magnetoresistive effect during magnetic ordering [14]. For this reason, we assume that a well-defined short-range magnetic order is preserved in the high-temperature phase at temperatures above T_N. The results of our heat capacity measurements support this hypothesis [18]. We observed that the heat capacity changes insignificantly at T_N.

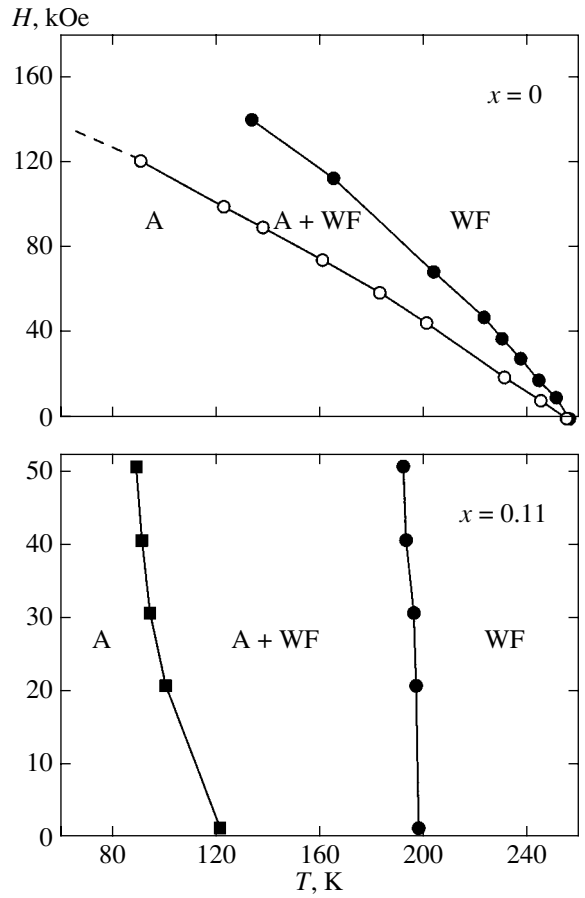


Fig. 5. Magnetic phase diagrams for $x = 0$ and 0.11 (WF denotes weak ferromagnet and A antiferromagnet).

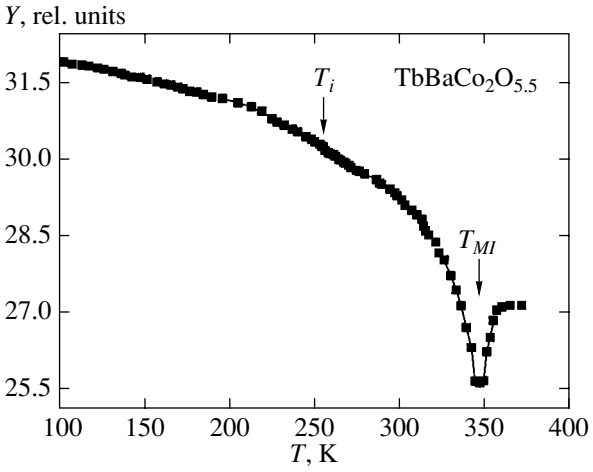


Fig. 6. Young's modulus as a function of temperature for a $TbBaCo_2O_{5.5}$ sample

4. CONCLUSIONS

It has been established that the substitution of iron ions for cobalt ions is an effective method of controlling the properties of compounds of the $TbBaCo_2O_{5+\gamma}$ type,

which considerably extends the temperature range of spontaneous magnetization, so that the Néel temperature becomes higher than room temperature. The transition at T_i is in all probability a Morin-type (antiferromagnet–weak ferromagnet) transition. The electrical conductivity jump near T_i is due to the emergence of spontaneous magnetization, reducing the energy gap between the 3*d* and 2*p* bands. In contrast to magnetoresistive manganites, the family of cobaltites of the $TbBaCo_2O_{5+\gamma}$ type does not exhibit a clearly manifested correlation between the magnetic state and the crystal lattice.

ACKNOWLEDGMENTS

This study was partly financed by the Belarussian Foundation for Basic Research (project no. F00-223) and the Polish Committee of Science (grant no. 5P03B01620).

REFERENCES

1. C. Martin, A. Maignan, D. Pelloquin, *et al.*, Appl. Phys. Lett. **71**, 1421 (1997).
2. I. O. Troyanchuk, N. V. Kasper, D. D. Khalyavin, *et al.*, Phys. Rev. Lett. **80**, 3380 (1998).
3. I. O. Troyanchuk, N. V. Kasper, D. D. Khalyavin, *et al.*, Phys. Rev. B **58**, 2418 (1998).
4. A. Maignan, C. Martin, D. Pelloquin, *et al.*, J. Solid State Chem. **142**, 247 (1999).
5. T. Vogt, P. M. Woodward, P. Karen, *et al.*, Phys. Rev. Lett. **84**, 2969 (2000).
6. Y. Moritomo, M. Takeo, X. J. Liu, *et al.*, Phys. Rev. B **58**, R13334 (1998).
7. D. Akahoshi and Y. Ueda, J. Phys. Soc. Jpn. **68**, 736 (1999).
8. E. Suard, F. Fauth, and V. Caignaert, Physica B (Amsterdam) **276–278**, 254 (2000).
9. E. Suard, F. Fauth, V. Caignaert, *et al.*, Phys. Rev. B **61**, R11871 (2000).
10. W. Zhou, C. T. Lin, and W. Y. Liang, Adv. Mater. **5**, 735 (1993).
11. I. O. Troyanchuk, N. V. Kasper, D. D. Khalyavin, *et al.*, J. Phys.: Condens. Matter **10**, 6381 (1998).
12. C. Martin, A. Maignan, F. Damay, *et al.*, J. Magn. Magn. Mater. **202**, 11 (1999).
13. V. G. Sathe, S. K. Paranjpe, V. Siruguri, *et al.*, J. Phys.: Condens. Matter **10**, 4045 (1998).
14. B. Raveau, C. Martin, and A. Maignan, J. Alloy Comp. **275–277**, 461 (1998).
15. M. A. Senaris-Rodrigez and J. B. Goodenough, J. Solid State Chem. **118**, 323 (1995).
16. K. P. Belov, A. K. Zvezdin, A. M. Kadomtseva, *et al.*, *Orientational Transitions in Rare Earth Magnetics* (Nauka, Moscow, 1979).
17. D. Akahoshi and Y. Ueda, J. Solid State Chem. **156**, 355 (2001).
18. N. V. Kasper, I. O. Troyanchuk, D. D. Khalyavin, *et al.*, Phys. Status Solidi B **215**, 697 (1999).

Translated by N. Wadhwa

Journal of Experimental and Theoretical Physics, Vol. 95, No. 4, 2002, pp. 753–761.
Translated from Zhurnal Éksperimental'noĭ i Teoreticheskoĭ Fiziki, Vol. 122, No. 4, 2002, pp. 869–878.

SOLIDS
Electronic Properties

Tunneling Magnetoresistance of Phase-Separated Manganites

A. O. Sboychakov[a], A. L. Rakhmanov[a], K. I. Kugel'[a], M. Yu. Kagan[b, *], and I. V. Brodsky[b]

[a]*Institute for Theoretical and Applied Electrodynamics, Russian Academy of Sciences, Moscow, 125412 Russia*

[b]*Kapitza Institute for Physical Problems, Russian Academy of Sciences, Moscow, 117334 Russia*

**e-mail: kagan@kapitza.ras.ru*

Received March 20, 2002

Abstract—A simple model of phase separation is used to study the magnetoresistance of manganites in the nonmetallic state. It is assumed that the phase separation corresponds to the emergence of small ferromagnetic metallic droplets (ferrons) in a nonconducting antiferromagnetic or paramagnetic medium, with the metallic phase concentration being far from the percolation threshold. The charge transfer is accomplished by way of electron jumps between droplets. The magnetoresistance in such a system is defined both by the variation of the volume of the metal phase and by the dependence of the probabilities of electron transitions on the magnitude of the magnetic field. It is demonstrated that, in the region of low magnetic fields, the magnetoresistance is quadratic with respect to the field and decreases with temperature by the T^{-n} law, where n takes values from 1 to 5 depending on the correlation between the parameters. In the high-field limit, the magnetoresistance increases abruptly with the volume of the metal phase. The crossover of the field dependence from quadratic to a stronger one may be accompanied by the emergence of a platean in the magnetoresistance. The correlation between the obtained results and the available experimental data is discussed. © *2002 MAIK "Nauka/Interperiodica".*

1. INTRODUCTION

The question of the nature of colossal magnetoresistance is one of the key questions in the physics of strongly correlated electron systems such as manganites. The investigations performed in recent years have revealed that manganites are characterized by a tendency for phase separation with the formation of, for example, conducting ferromagnetic and nonconducting antiferromagnetic or paramagnetic regions [1–3]. Apparently, the tendency for phase separation is closely related to the mechanism of colossal magnetoresistance. The phase separation may be due to various reasons and may be characterized by different spatial scales. One of the most important types of phase separation for transport phenomena in manganites is the electron separation which was first suggested by Nagaev [4] in application to magnetic semiconductors. In this case, the current carriers are concentrated in small (of the order of several lattice constants) ferromagnetic metal droplets (ferrons). Regions with a low carrier concentration turn out to be nonconducting. In the limit of strong Coulomb interaction, each ferron in the ground state contains one charge carrier located in the potential well of ferromagnetically ordered local spins.

A model which enables one to estimate the ferron size was treated in [5, 6]. According to this model, the ferron radius R (the ferron is assumed to be spherical) is determined by way of minimizing the energy,

$$E \sim t\left(\frac{\pi d}{R}\right)^2 + \frac{4\pi}{3} JzS^2\left(\frac{R}{d}\right)^3,$$

where the first term is the energy of the ground state of an electron in a spherical potential well of radius R, and the second term corresponds to the Heisenberg energy of interaction of local spins S. The minimization of the energy over the ferron radius leads to the estimate of $R \propto d(\pi t/JzS^2)^{1/5}$, where t is the amplitude of electron jump to the neighboring site, J is the constant of antiferromagnetic exchange interaction between atoms, z is the number of nearest neighbors, and d is the lattice spacing.

The number of charge carriers is proportional to the concentration x of the bivalent doping element. The critical concentration $x_c(d/R)^3 \sim 0.15$, at which ferrons start to overlap and large metal clusters arise in the system, was estimated in [6]. We will treat a situation in which x is much lower than x_c. The charge transfer in such a system may occur both due to the ferron motion and as a result of the electron tunneling from one ferron to another. The efficiency of the former mechanism is low because of the fairly large effective mass of a ferron. In addition, the ferrons in manganites are often pinned on impurities [1]. Therefore, it is natural to

assume that the conductivity is largely defined by electron jumps between ferrons.

Within the assumptions formulated above, we have thoroughly analyzed the field and temperature dependences of the magnetoresistance associated with the tunneling of charge carriers from ferron to ferron. It is demonstrated that, in low fields, the magnetoresistance increases as the square of the field; in high fields, it increases exponentially [3, 7] or even faster as the percolation threshold is approached. The temperature dependence of magnetoresistance is likewise nontrivial. In low fields, this dependence obeys the $1/T^n$ law, where n runs from 1 to 5 depending on the parameters of the problem. The obtained results may be important from the standpoint of interpreting the results of recent experiments in studying the magnetoresistance of complex magnetic oxides in the nonmetallic region of temperatures and concentrations [2, 8, 9].

2. DESCRIPTION OF THE MODEL

We will treat a system consisting of N metal droplets (ferrons) located in an antiferromagnetic matrix in an external uniform magnetic field **H**. As was demonstrated in [7], with not-too-low concentrations of ferrons, their number may be assumed to be constant and equal to the number N_e of current carriers arising during doping. We will assume that the ferrons are spherical in shape. We will designate the spin of the magnetic atom of droplet as S, and the number of magnetic atoms in a droplet as $N_p = (4\pi/3)(R/d)^3$, where R is the ferron radius. Then, the magnetic moment of a ferron is $M_0 = \mu_B g S N_p$, where μ_B is a Bohr magneton and g is the gyromagnetic ratio. Because the value of the number SN_p is rather high, we will describe the magnetic moment of a droplet in the classical manner; i.e., we will assume that the vector $\mathbf{M}_0$ has a certain direction in space, determined by the angles θ and ϕ.

At zero temperature, each ferron contains one charge carrier, and its magnetic moment is directed along the magnetic field. We will assume that, at nonzero temperatures, all spins S in a droplet are parallel with one another; however, their direction may differ from that of the external field. Let the magnetic field be directed along the z axis. In the approximation of low density, we will ignore the interactions of ferrons with one another. Then, the energy of a droplet containing one electron may be represented as [6]

$$E_{1,\sigma}(\theta) = t\frac{\pi^2 d^2}{R^2} + \frac{4\pi}{3}JzS^2\frac{R^3}{d^3} - \mu_B g S\frac{4\pi}{3}\frac{R^3}{d^3}H\cos\theta + U_{\mathrm{el},\sigma}(\mathbf{H}). \quad (1)$$

The first term in this expression is the electron energy in a potential well formed by ferromagnetically ordered local spins S, which is counted off from the bottom of the conduction band. The second term is the energy of exchange interaction of the magnetic atoms of the droplet. The third term is the variation of the droplet energy in the magnetic field (Zeeman energy). The last term $U_{\mathrm{el},\sigma}(\mathbf{H})$ is the energy of interaction of a conduction electron in a ferron with the effective magnetic field.

Note that, in the concentration range of $x \sim x_{\mathrm{opt}} \sim 0.3$ that is optimal for the observation of colossal magnetoresistance, but at temperatures above the Curie point T_C, the system is again (as in the case of $x < x_c$) a paramagnetic insulator with suppressed ferromagnetism but with strong ferromagnetic fluctuations of small radius. As was demonstrated in [3, 10, 11], these fluctuations may be treated as ferromagnetic droplets; a special term of "temperature ferrons" was introduced for these fluctuations. In formula (1), one must change from energy to free energy; in the second term, the energy of antiferromagnetic interaction of local spins $(4\pi/3)JzS^2R^3/d^3$ must be replaced by the corresponding entropy contribution $(4\pi/3)T\ln(2S+1)R^3/d^3$ [3]. As a result, the estimation of the optimal radius of a temperature ferron gives $R \sim d(\pi t/T\ln(2S+1))^{1/5}$. Therefore, in view of this substitution, all of the results obtained below for the low-temperature region with $T < T_C$ and $x < x_c$ may be used in the high-temperature region with $T > T_C$ and $x \sim x_{\mathrm{opt}}$ as well.

In expression (1) for energy, one must take into account the fact that the electron in the droplet interacts both with the external field **H** and with a field developed by magnetic atoms of a ferron. We will designate the effective number of atoms with which a charge carrier interacts as z' ($z' \le N_p$). Then, we can introduce the effective magnetic field in which the electron is located [12],

$$\mathbf{H}_{\mathrm{eff}}(\theta) = \mathbf{B}(\mathbf{H}) + \mathbf{n}(\theta)H_{\mathrm{mol}}, \quad (2)$$

where $H_{\mathrm{mol}} = \bar{J}/\mu_B = 2J_{\mathrm{fer}}z'S/g\mu_B$ is the molecular magnetic field acting on the electron from the ferromagnetically ordered local spins, J_{fer} is the respective exchange integral defining the Curie temperature, and **n** is a unit vector directed along the magnetic moment of the ferron. The quantity $\mathbf{B}(\mathbf{H})$ in formula (2) is the magnetic field induction inside the droplet in view of, in particular, its demagnetization factor.

The steady states of an electron are states with a certain projection of spin $\sigma/2$ onto the direction of the effective field. We will use the notation of both $\sigma = \pm 1$ and $\sigma = \uparrow, \downarrow$. So, the energy $U_{\mathrm{el},\sigma}$ may be represented as

$$U_{\mathrm{el},\sigma} = -\mu_B\sigma H_{\mathrm{eff}}(\theta) = -\mu_B\sigma\sqrt{(\bar{J}/\mu_B)^2 + \mathbf{B}^2 + 2\bar{J}(\mathbf{n}\cdot\mathbf{B})/\mu_B}, \quad (3)$$

and, in the ground state, $\sigma = \uparrow$.

We will assume that the direction of the total magnetic moment of the droplet varies with time rather slowly, and the conventional thermodynamic treatment

of the problem is valid. Then, the ferron radius in the magnetic field is determined by minimizing expression (1), which gives

$$R(H) = R(0)(1 - bH\cos\theta)^{-1/5},$$
$$R(0) = d\left(\frac{\pi t}{2zJS^2}\right)^{1/5}, \quad b = \frac{\mu_B g}{zJS}, \tag{4}$$

where $R(0)$ is the ferron radius in the absence of a field. The estimation of the quantities b and $\bar{J}$ gives $1/b$, $\bar{J}/\mu_B \sim 100$ T. In the region of not-too-high fields, where $bH \ll 1$, we can expand the field dependences of the droplet radius and energy into a series and restrict ourselves to linear terms. Then, for the energy given by Eq. (1), we have

$$E_{1,\sigma}(\theta) = E_1^{(0)}\left(1 - \frac{2}{5}bH\cos\theta\right) - \sigma\bar{J} - \sigma\mu_B(\mathbf{n}\cdot\mathbf{B}(\mathbf{H})) + O(H^2). \tag{5}$$

In this expression,

$$E_1^{(0)} = \frac{5\pi}{3}(\pi t)^{3/5}(2zS^2J)^{2/5} \tag{6}$$

is the ferron energy at $H = 0$ disregarding the splitting of the electron energy in the effective field. Note that $(2/5)E_1^{(0)}b = M_0 = \mu_B g S N_p$ is the magnetic moment of the ferron. Because $SN_p \gg 1$ and $\bar{J}/\mu_B \sim 100$ T, we will, in what follows, ignore the term $\sigma\mu_B(\mathbf{n}\cdot\mathbf{B})$ in formula (5); i.e., we will assume that $U_{\mathrm{el},\sigma} = -\sigma\bar{J}$, and the direction of the effective field coincides with the direction of the magnetic moment of the ferron.

As a result of electron tunneling to the neighboring droplets, empty ferrons and ferrons with more than one electron are formed. In addition, part of the electrons in one-electron channels may be in excited states. As will be shown below, for the characteristic values of the parameters, the ground state energy $E_1^{(0)} \sim 1.5$ eV $\gg kT$. One can readily understand that, in zero field, the distance between the ground and second electron levels in a ferron is of the same order. On the other hand, the same estimates indicate that the splitting of the first electron level in a molecular field $\mu_B H_{\mathrm{mol}} \sim 0.01$–$0.03$ eV. Therefore, in what follows, we take into account only the first excited level, which, in the range of parameters of the model being treated, corresponds to the ground state of an electron but with an oppositely directed spin $\sigma = -1$. Because $SN_p \gg 1$, one can assume that tunneling transitions of electrons do not affect the direction of large magnetic moments of droplets.

Empty droplets and droplets with two electrons decay during a time of the order of characteristic magnon times τ_0. Then, they may be treated as stable only if the time of electron tunneling τ is much shorter than τ_0. Because the number of particles does not vary, the times of electron tunneling to a empty droplet and of electron escape from a two-electron droplet are equal to one another, and the condition $\tau \ll \tau_0$ is valid if $\exp(-V/2kT)/8\pi l^3 n_e \ll 1$, where V is the energy of Coulomb repulsion of two electrons on a single droplet, l is the characteristic tunneling length, and n_e is the number of ferrons per unit volume [7]. In meeting the latter condition, one can assume that the radii of one-electron, empty, and two-electron droplets are equal to one another and defined by relation (4).

The energy of an empty droplet, defined by the second and third terms in formula (1), is

$$E_0(\theta) = E_0^{(0)}\left(1 - \frac{2}{5}bH\cos\theta\right) + O(H^2),$$
$$E_0^{(0)} = \frac{2}{5}E_1^{(0)}. \tag{7}$$

In a two-electron droplet, the electrons must have opposite spin projections onto the direction of the magnetic moment. The energy of such a ferron may be represented as

$$E_2(\theta) = E_2^{(0)}\left(1 - \frac{2}{5}bH\cos\theta\right) + V(0)\frac{1}{5}bH\cos\theta + O(H^2), \tag{8}$$

where $E_2^{(0)} = 2E_1^{(0)} - E_0^{(0)} + V(0)$ and the Coulomb energy is $V(H) = e^2/\epsilon R(H) \approx V(0)(1 - (1/5)bH\cos\theta)$. For estimation, we will assume the value of permittivity $\epsilon = 10$ and $R(0) = 2d$. From this, we find that $V(0) \sim 0.2$ eV $\gg \mu_B H_{\mathrm{mol}}$. We will treat the temperature range $kT \ll V(0)$. It is in this case that one can ignore the probability of formation of ferrons with three or more electrons.

In principle, our model describes the electron tunneling between potential wells with due regard for the Coulomb repulsion of electrons inside the wells. In this sense, it is similar to the Hubbard model with a random arrangement of centers. As was mentioned above, we treat the case of one electron per center with a strong Coulomb repulsion. In this case, there are upper and lower Hubbard subbands separated by a Coulomb gap. The lower subband is filled, and the upper one is empty. Then, the conductivity is determined by electron jumps from the lower to upper subband. Because the value of temperature is assumed to be much less than that of the Coulomb gap, the number of electrons involved in the process of charge transfer is much less than the total number of electrons which formed ferrons. Therefore, one can assume that an electron, during jumps, passes from the upper state of the filled band to the bottom of the empty band. The activation energy in such a process is fixed and is of the order of the Coulomb gap; therefore, transitions between subbands may be replaced by

transitions between levels. Naturally, if the numbers of electrons and ferrons were significantly different, such a simplified approach would need to be refined. In particular, one would have to take into account the finiteness of the bandwidth and the Fermi statistics of the carriers (at least, in the region of fairly low temperatures).

Therefore, the system contains $N_{1,\sigma}$ droplets with an electron having the $\sigma/2$ spin projection onto the effective field direction, N_0 empty ferrons, and $N_2 = N_0$ ferrons with two electrons. We will find the average values of N_0, N_2, and $N_{1,\sigma}$. The number of possible states with given quantities $\{N_a\}$ (where the subscript a runs through the values $\{0, (1, \uparrow), (1, \downarrow), 2\}$) is $N_e!/(N_0!N_{1,\uparrow}!N_{1,\downarrow}!N_2!)$. Then, the partition function of the system may be represented as

$$\mathcal{Z} = N_e! \sum_{\{N_a\}} \delta_{N_0}^{N_2} \delta_{N_{1,\uparrow}+N_{1,\downarrow}}^{N_e-2N_2} \times \prod_a \left[\frac{1}{N_a!}\left(\int d\Omega \exp\left(-\frac{E_a(\theta)}{kT}\right)\right)^{N_a}\right], \quad (9)$$

where the Kronecker symbols take into account the fact that, in our model, $N_0 = N_2$ and $N_{1,\uparrow} + N_{1,\downarrow} + 2N_2 = N_e$. The summation with respect to N_0, N_2 is performed from 0 to $N_e/2$, and that with respect to $N_{1,\sigma}$, from 0 to $N_e - 2N_2$. Because $N_e \gg 1$, the sum with respect to N_2 may be replaced by an integral using the Stirling formula for the factorials appearing in Eq. (9). The resultant integral is readily calculated by the saddle-point method. After that, we integrate with respect to solid angle to derive, for the average numbers $\bar{N}_2$ and $\bar{N}_{1,\sigma}$,

$$\bar{N}_2 = N_e \frac{\exp(-V(0)/2kT)}{2\cosh(\bar{J}/kT)} \times \left[\frac{\alpha_1^2}{\alpha_0\alpha_2}\frac{\sinh(\alpha_0 H)\sinh(\alpha_2 H)}{\sinh^2(\alpha_1 H)}\right]^{1/2}, \quad (10)$$

$$\bar{N}_{1,\sigma} = (N_e - 2\bar{N}_2)\frac{e^{\sigma\bar{J}/kT}}{2\cosh(\bar{J}/kT)}, \quad (11)$$

where the following notation is used:

$$\alpha_0 = \frac{2E_0^{(0)}b}{5kT}, \quad \alpha_1 = \frac{2E_1^{(0)}b}{5kT}, \quad \alpha_2 = \frac{2(2E_1^{(0)} - E_0^{(0)} + V(0)/2)b}{5kT}. \quad (12)$$

It follows from Eq. (10) that, in weak fields, the first correction to $\bar{N}_2(0)$ increases as the square of the field, and in fairly strong fields, $\bar{N}_2 \propto \exp(bV(0)H/10kT)$. Note that expressions (10) and (11) agree with the results obtained in [7] in the absence of a magnetic field and disregarding the electron spin, that is, for $H = 0$ and $\bar{J} = 0$.

For a ferron in the state a, the probability density for the direction of the magnetic moment within the solid angle $d\Omega$ may be written as

$$P_0(\theta) = \frac{\exp(\alpha_0 H\cos\theta)}{\int d\Omega' \exp(\alpha_0 H\cos\theta')}, \quad P_{1,\sigma}(\theta) = \frac{\exp(\alpha_1 H\cos\theta)}{\int d\Omega' \exp(\alpha_1 H\cos\theta')}, \quad P_2(\theta) = \frac{\exp(\alpha_2 H\cos\theta)}{\int d\Omega' \exp(\alpha_2 H\cos\theta')}. \quad (13)$$

The values of probability for droplets containing different numbers of electrons differ, because the magnetic field dependences of their energies (5), (7), and (8) are different. Note that the probabilities $P_{1,\sigma}$ are independent of the sign of σ, because we ignored the term $\mathbf{B}(\mathbf{H})$ in formula (2) for the effective field acting on an electron.

3. MAGNETORESISTANCE

Within our model, the conductivity is defined by the following tunneling processes.

(i) An electron tunnels from a ferron in the state 1, σ_1 onto a droplet containing an electron with the spin projection $\sigma_2/2$. As a result, a empty ferron and a ferron with two electrons arise. This process may be schematically represented as $(1\sigma_1, 1\sigma_2) \longrightarrow (0, 2)$.

(ii) A process inverse to the previous one, i.e., $(2, 0) \longrightarrow (1\sigma_1, 1\sigma_2)$.

(iii) The ferron with two electrons and that containing an electron with the spin projection $\sigma_1/2$ exchange electrons: $(2, 1\sigma_1) \longrightarrow (1\sigma_2, 2)$.

(iv) The one-electron ferron transfers its electron to the empty ferron: $(0, 1\sigma_1) \longrightarrow (1\sigma_2, 0)$.

Note that, in the last two cases, σ_1, generally speaking, is not equal to σ_2. The contribution to the current density by each tunneling process may be represented as [7, 13]

$$i_{(a', b'; a, b)} = \frac{e}{V}\left\langle \sum_{i=0}^{N_a}\sum_{j=0}^{N_b} \frac{r_{ij}\cos\gamma_{ij}}{\tau_{i,j}(a', b'; a, b)}\right\rangle, \quad (14)$$

where r_{ij} and γ_{ij} denote the distance between the ferrons and the angle between the vector $\mathbf{r}_{ij}$ and the direction of the electric field, respectively. The summation with respect to i and j is performed over the ferrons from which and to which an electron may tunnel, and the symbol $\langle\ldots\rangle$ denotes both thermal and volume averaging. The quantity $\tau_{ij}(\mathbf{f}, \mathbf{i})$ is the time of tunneling from

the initial **i** = (*a*, *b*) to the final **f** = (*a*', *b*') state. It may be represented as

$$\tau_{ij}(\mathbf{f}, \mathbf{i}) = \omega_0^{-1} \exp\left(\frac{r_{ij}}{l} - \frac{e(\mathbf{E}\cdot\mathbf{r}_{ij})}{kT}\right) W^{-1}(\mathbf{f}, \mathbf{i}), \quad (15)$$

where l and ω_0^{-1} denote the tunneling length and the period of electron motion in the potential well, respectively. The quantities $W^{\mathbf{f},\mathbf{i}}$ by definition are independent of the distance between ferrons. Because, as a result of tunneling transitions, the energy of the system, generally speaking, varies, the probability W depends on the difference between the energies before and after the tunneling. In addition, the probability of tunneling must depend on the relative orientation of the magnetic moments of ferrons.

In order to find the quantity $W(\mathbf{f}, \mathbf{i})$, we will use the detailed balance relation for each tunneling process in the absence of an electric field. A ferron is characterized by the number and state of electrons it contains, as well as by the direction of the magnetic moment. Therefore, its state is defined by the set of quantities (a, **n**), where a numbers the electron states and runs through the values {0, (1, ↑), (1, ↓), 2}, and **n** is a unit vector directed along the magnetic moment of the ferron. The number of ferrons in the state (a, **n**) per unit volume is $\bar{N}_a P_a(\mathbf{n})\, d\Omega/V$. We assumed that, during tunneling transitions, the direction of the magnetic moments of droplets is retained. Then, the number of transitions from the state (a_1, $\mathbf{n}_1$, a_2, $\mathbf{n}_2$) to the state (a_1', $\mathbf{n}_1$, a_2', $\mathbf{n}_2$) per unit time in an element of volume dV is

$$\frac{1}{V^2}\frac{\bar{N}_{a_1}P_{a_1}(\theta_1)\bar{N}_{a_2}P_{a_2}(\theta_2)d\Omega_1 d\Omega_2 dV}{\tau(a_1', a_2'; a_1, a_2)}. \quad (16)$$

In equilibrium, this number must be equal to the number of inverse transitions. We can use equality (15) for the tunneling time, as well as the expressions for $\bar{N}_a$ and $P_a(\theta)$, and write the balance relation in the form

$$\frac{W(a_1', a_2'; a_1, a_2)}{\exp[(E_{a_1}(\theta_1) + E_{a_2}(\theta_2))/kT]} = \frac{W(a_1, a_2; a_1', a_2')}{\exp[(E_{a_1'}(\theta_1) + E_{a_2'}(\theta_2))/kT]}. \quad (17)$$

We will follow the thermodynamics and assume that the probability $W(\mathbf{f}, \mathbf{i})$ exponentially depends on the energy difference between the initial and final states. Then,

$$W(a_1', a_2'; a_1, a_2) = f^{\mathbf{f},\mathbf{i}}(\cos\nu) \times \exp\left\{-\frac{E_{a_1'}(\theta_1) + E_{a_2'}(\theta_2) - E_{a_1}(\theta_1) - E_{a_2}(\theta_2)}{2kT}\right\}, \quad (18)$$

where ν is the angle between the directions of the magnetic moments of ferrons,

$$\cos\nu = \cos\theta_1\cos\theta_2 + \sin\theta_1\sin\theta_2\cos(\phi_1 - \phi_2). \quad (19)$$

The functions $f^{\mathbf{f},\mathbf{i}}$ define the dependence of the tunneling probability on the relative orientation of the magnetic moments of droplets. These probabilities must coincide for the direct and inverse processes. We will treat the process (ii). If an electron with the spin projection σ/2 onto the direction of the magnetic moment of a droplet 1 enters a empty ferron 2, its energy in the effective field varies by the quantity $-\sigma\bar{J}\cos\nu + \sigma\bar{J}$. Then, for f, we have

$$f^{1\sigma_1, 1\sigma_2; 0, 2}(\cos\nu) = \sum_\sigma q_\sigma \exp\left(-\frac{\sigma\bar{J}(1 - \cos\nu)}{kT}\right), \quad (20)$$

where q_σ is the probability that an outgoing electron will have the spin projection σ/2. In view of normalization, it is equal to $e^{\sigma\bar{J}/kT}/2\cosh(\bar{J}/kT)$. So, the final expression for $f^{1\sigma_1, 1\sigma_2; 0, 2}$ has the form

$$f^{1\sigma_1, 1\sigma_2; 0, 2}(\cos\nu) = \frac{\cosh(\bar{J}\cos\nu/kT)}{\cosh(\bar{J}/kT)}. \quad (21)$$

The function $f^{0, 2; 1\sigma_1, 1\sigma_2}$ for the process (i) inverse to the process treated above must be given by the same expression. The same method may be used to demonstrate that formula (21) for the function $f^{\mathbf{f},\mathbf{i}}$ is valid for all of the tunneling processes listed above.

We will now use expressions (15) and (18) to perform the volume averaging of the current density given by Eq. (14). Then, an approximation linear with respect to the field E gives

$$i_{(a', b'; a, b)} = \frac{32\pi e^2 l^5 \omega_0}{kT}\bar{n}_a\bar{n}_b\langle W(a', b'; a, b)\rangle E,$$

where $n_a = N_a/V$, and the averaging is performed over the directions of the moments of ferrons, distributed with the probability density $P_a(\theta_1)P_b(\theta_2)$. On summing up the contributions from all tunneling processes, we derive for the conductivity $\sigma = i/E$ the following expression:

$$\sigma(H) = \sigma_0\frac{j}{4\sinh j}\int\frac{d\Omega_1}{4\pi}\int\frac{d\Omega_2}{4\pi}\cosh(j\cos\nu) \times \exp\left(\frac{1}{2}\alpha_1 H\cos\theta_1\right)\exp\left(\frac{1}{2}\alpha_1 H\cos\theta_2\right) \times\left(\exp\left(\frac{1}{2}\alpha_0 H\cos\theta_1\right)\exp\left(\frac{1}{2}\alpha_0 H\cos\theta_2\right)\right. \times\sqrt{\frac{\alpha_0\sinh(\alpha_2 H)}{\alpha_2\sinh(\alpha_0 H)}} \quad (22)$$

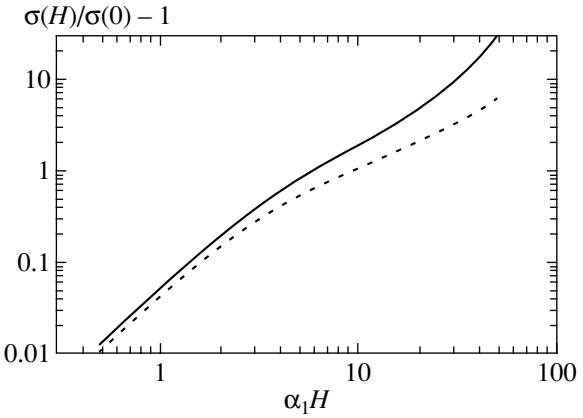


Fig. 1. The magnetoresistance as a function of magnetic field at $j = 2.25$, $V(0)/E_1^{(0)} = 0.2$ (solid curve) and $j = 1.875$, $V(0)/E_1^{(0)} = 0.1$ (dashed curve).

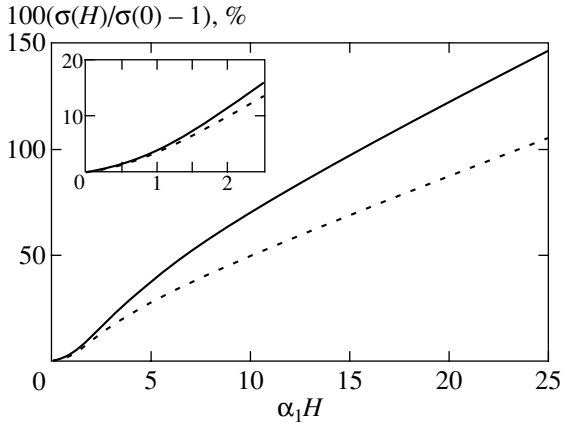


Fig. 2. The magnetoresistance as a function of magnetic field at $j = 1.5$ (solid curve) and $j = 1.0$ (dashed curve). In both cases, $V(0)/E_1^{(0)} = 0.07$.

$$+ 2\exp\left(\frac{1}{2}\alpha_0 H\cos\theta_1\right)\exp\left(\frac{1}{2}\alpha_2 H\cos\theta_2\right)$$
$$+ \exp\left(\frac{1}{2}\alpha_2 H\cos\theta_1\right)\exp\left(\frac{1}{2}\alpha_2 H\cos\theta_2\right)$$
$$\times\sqrt{\frac{\alpha_2\sinh(\alpha_0 H)}{\alpha_0\sinh(\alpha_2 H)}}\Bigg)\left[\frac{\sinh(\alpha_1 H)}{\alpha_1 H}\right]^{-2},$$

where $j = \bar{J}/kT$, and σ_0 is the conductivity in the absence of a field,

$$\sigma_0 = \frac{128\pi e^2 l^5 \omega_0 n_e^2 \cosh^2(j/2)\sinh j}{j\cosh^3 jkT}\exp\left(-\frac{V(0)}{2kT}\right), \qquad n_e = N_e/V. \tag{23}$$

If we ignore the splitting of the electron energy due to the interaction of its spin with the molecular field, i.e., assume that $j = 0$, we will arrive at the expression for conductivity σ_0 obtained in [7].

In what follows, we will be interested in the magnetoresistance $\mathrm{MR}(H) = \sigma(H)/\sigma_0 - 1$. The integrals in formula (22) cannot be calculated analytically in the general case. However, we can obtain asymptotics for the magnetoresistance at $H \longrightarrow 0$ and $\alpha_0 H \gg 1$. In the limit of low fields, the expansion of the conductivity in terms of H proceeds over even powers of the field. We restrict ourselves to the quadratic term to derive

$$\mathrm{MR}(H) = \frac{H^2}{24}\left((\alpha_1+\alpha_2)^2 + (\alpha_1+\alpha_0)^2 - 8\alpha_1^2\right) \propto H^2/T^2. \tag{24}$$

Note that the coefficient of H^2 in this expression is independent of $\bar{J}$. Therefore, the dependence of the magnetoresistance on the relative orientation of the magnetic moments of ferrons is observed only in higher orders with respect to the field.

In rather strong fields, when $\alpha_0 H \gg 1$ (but still $bH \ll 1$), the magnetoresistance exponentially depends on the field, in accordance with the results of [3, 7],

$$\mathrm{MR}(H) = j\cot j\left(\frac{(\alpha_0/\alpha_2)^{1/4}}{1+\alpha_0/\alpha_1} + \frac{(\alpha_2/\alpha_0)^{1/4}}{1+\alpha_2/\alpha_1}\right)^2 \times\exp\left(\frac{bV(0)}{10kT}H\right) - 1. \tag{25}$$

Figures 1 and 2 give the field dependences of magnetoresistance calculated by formula (22) for different values of the parameters α_a and j. In Fig. 1, in particular, one can see that, in the region of weak fields, the magnetoresistance is quadratic with respect to the field, and then the MR(H) dependence becomes linear; in strong fields, the magnetoresistance starts to increase exponentially. No exponential increase is observed in the curves in Fig. 2 since they correspond to the case of higher temperatures and to a lower value of Coulomb energy.

4. THE EFFECT OF INCREASE IN THE VOLUME OF THE METALLIC PHASE

In the model treated above, the variation of conductivity in the magnetic field was due to the field dependence of the probabilities of tunneling transitions and occupation numbers $\bar{N}_2$ and $\bar{N}_{1,\sigma}$. Along with these mechanisms, the magnetoresistance may be directly related to the increase in the size of ferrons. Indeed, the following formula for the conductivity of a insulator-metal mixture is valid away from the percolation threshold [14]:

$$\sigma = \sigma_d(1+x), \tag{26}$$

where σ_d is the conductivity of the insulator matrix and x is the volume fraction of the metal phase. Here, we have assumed that the conductivity of the insulator is much lower than that of the metal.

In our case, the quantity x may be represented as

$$x = \frac{4\pi}{3V}\sum_a \bar{N}_a \langle R^3\rangle_a. \quad (27)$$

The symbol $\langle\ldots\rangle_a$ in this expression implies the averaging over the directions of the magnetic moment with the probability density P_a. Because the numbers $\bar{N}_2$ and $\bar{N}_0$ are exponentially small compared with $\bar{N}_{1,\uparrow} + \bar{N}_{1,\downarrow} \approx N_e$, and the values of probability $P_{1,\sigma}$ do not depend on the electron spin projection, one can write, instead of Eq. (27),

$$x = \frac{4\pi}{3} n_e \langle R^3\rangle_1 = x_0 \left\langle \frac{R^3(\mathbf{H})}{R^3(0)} \right\rangle_1, \quad (28)$$

where x_0 is the volume fraction of ferrons in the absence of a field. In the case being treated, this quantity must be less than the critical value, $x_c \approx 0.15$, at which a percolation transition occurs.

We use equality (4) for the ferron radius to find that the contribution to magnetoresistance due to the mechanism being treated is described by the formula

$$\begin{aligned} \mathrm{MR}(H) &= x(H) - x_0 \\ &= x_0 \frac{3b}{5}\left(H\cot(\alpha_1 H) - \frac{1}{\alpha_1}\right). \end{aligned} \quad (29)$$

In the case of low fields, the magnetoresistance is quadratic with respect to the field,

$$\mathrm{MR}(H) = x_0 \frac{b}{5}\alpha_1 H^2 \propto H^2/T. \quad (30)$$

Note that, unlike Eq. (24), this expression decreases with increasing temperature as $1/T$.

In the case of fairly high fields, when $\alpha_1 H \gg 1$ (but still $bH \ll 1$), the magnetoresistance given by Eq. (29) depends linearly on the field,

$$\mathrm{MR}(H) = x_0 \frac{3b}{5} H. \quad (31)$$

5. INCLUSION OF MAGNETIC ANISOTROPY

In the expression for ferron energy in the previous sections, we ignored the contribution related to the magnetic anisotropy and the droplet shape effect and assumed the droplet to be spherical. We will now consider the influence of these effects on the behavior of magnetoresistance. To simplify calculations, we will assume that the droplet size does not vary with increasing field. Obviously, this is admissible only in the region of weak fields. In addition, we will assume that the magnetic anisotropy is uniaxial and that ferrons have the shape of an ellipsoid of revolution with the axis parallel to the anisotropy axis. Note that, apparently, ferrons in layered manganites are ellipsoidal [15, 16]. Then, the additional contribution to the droplet energy, associated with anisotropy, is [14]

$$U_\alpha = -M_0 H_a \cos^2\psi, \quad (32)$$

where ψ is the angle between the easy magnetic axis (easy axis) and magnetic moment, and H_a is the anisotropy field which is made up of the crystalline anisotropy field proper and of the form anisotropy. The latter contribution may be estimated as $\pi m(1 - 3\mathcal{N})$, where m is the magnetic moment of unit volume and $\mathcal{N}$ is the corresponding demagnetizing factor. Note that, with even a slight deviation of the ferron shape from spherical, this term may reach values which exceed significantly the crystalline anisotropy fields typical of ferromagnets. For example, with the ellipsoid semiaxis ratio of two and with the values of parameters typical of manganites, the shape contribution to the anisotropy field turns out to be of the order of 0.5–0.7 T.

Let the easy axis direction be determined by the angles β and ϕ_β; in so doing, one can assume, without loss of generality, that the azimuthal angle ϕ_β is zero. Then,

$$\cos\psi = \cos\theta\cos\beta + \sin\theta\sin\beta\cos\phi. \quad (33)$$

We perform calculations analogous to those made in Sections 2 and 3 to derive

$$\sigma(H) \propto \frac{\int d\Omega_1 \int d\Omega_2 \exp\{\alpha_1(H\cos\theta_1 + H_a\cos^2\psi_1)\}\cosh(j\cos\nu)\exp\{\alpha_1(H\cos\theta_2 + H_a\cos^2\psi_2)\}}{\left(\int d\Omega \exp\{\alpha_1(H\cos\theta + H_a\cos^2\psi(\theta,\phi))\}\right)^2} \quad (34)$$

for conductivity.

Analytical results in this case can be obtained only in the region of low fields and fairly high temperatures, when j and $\alpha_1 H_a < 1$. We expand Eq. (34) in terms of j, H_a, and H accurate within quadratic terms to derive the following expression for magnetoresistance:

$$\begin{aligned} \mathrm{MR}(H) &= \frac{2}{225}\left(\cos^2\beta - \frac{1}{3}\right)\alpha_1^3 j^2 H^2 H_a \\ &\times \left(1 + \frac{2\alpha_1 H_a}{7}\right) \propto H^2/T^5. \end{aligned} \quad (35)$$

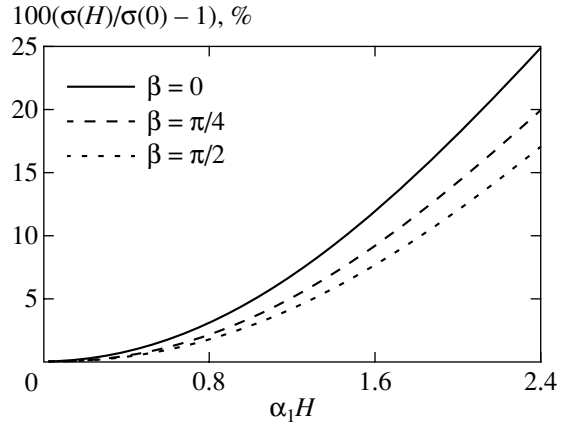


Fig. 3. The magnetoresistance as a function of magnetic field for different angles β between the easy axis and magnetic field. The parameters are $j = 1.5$, $\alpha_1 H_a = 0.8$, and $V(0)/E_1^{(0)} = 0.1$.

It follows from this relation that, in the region of low fields, the magnetoresistance is proportional to T^{-5}. It is interesting to note that, depending on the orientation of the magnetic field relative to the easy axis, the magnetoresistance may change its sign: it is positive at small angles β and negative at large angles.

It is possible to demonstrate that, if we consistently take into account the effects associated both with the increase in the ferron size and with the presence of anisotropy, the magnetoresistance in weak fields is given by the sum of expressions (24), (30), and (35).

Figure 3 gives the results of numerical calculation of the dependence of magnetoresistance on the magnetic field in the region of weak fields at different angles β between the easy axis and magnetic field.

6. DISCUSSION

Thus, in all of the cases treated above, the magnetoresistance is quadratic with respect to the field in the region of low fields, which agrees with experiment [2, 8, 9]. However, the temperature dependences of the coefficient of H^2 in equalities (24), (30), and (35) vary and are determined by the relationship between the parameters of the problem.

For numerical estimation, we will use the values of the parameters characteristic of manganites, namely, $t = 0.3$ eV, $J \approx J_{fer} = 1.5 \times 10^{-3}$ eV, $z = 6$, $g = 2$, $S = 2$, and $d = 0.4$ nm. We have assumed that the constants of ferromagnetic and antiferromagnetic interaction are close to one another in magnitude, because the Neél and Curie points in manganites are usually of the same order of magnitude. Note that this assumption agrees with the results of [17]. Further, let $z' = 10$ and the permittivity $\epsilon = 10$. Then, we derive from formulas (4) and (5) that $E_1^{(0)} = 1.5$ eV, $V(0) = 0.2$ eV, $1/b = 100$ T, and $\bar{J} = 200$ K. It follows from the first two estimates that the Coulomb energy may be ignored in the determination of α_2 compared with other terms. As a result, $\alpha_2 = (8/5)\alpha_1$ and $\alpha_0 = (2/5)\alpha_1$, and, because $\alpha_1 kT = M_0$ is the magnetic moment of a droplet, formulas (24), (35), and (30) may be rewritten as

$$\mathrm{MR}_1(H) = \frac{3}{100}\frac{M_0^2 H^2}{(kT)^2}, \tag{24'}$$

$$\mathrm{MR}_2(H) = \frac{2}{225}\frac{M_0^3 \bar{J}^2 H_a H^2}{(kT)^5} \times\left(\cos^2\beta - \frac{1}{3}\right)\left(1 + \frac{2M_0 H_a}{7kT}\right), \tag{35'}$$

$$\mathrm{MR}_3(H) = x_0 \frac{b}{5}\frac{M_0 H^2}{kT}. \tag{30'}$$

It should be recalled that the magnetic moment of a ferron is defined as $M_0 = \mu_B g S N_p$, where $N_p = (4\pi/3)(R/d)^3 \approx 30$ is the number of magnetic atoms in a droplet. We use the obtained values of b, $\bar{J}$, M_0, and $x_0 < 0.1$ to find that, at $T = 100$ K, the coefficient of H^2 in formula (30') is two orders of magnitude less than the respective coefficients in expressions (24') and (35'), which, in turn, are of the order of 10^{-1} T^{-2}. Therefore, the effect associated with the increase in the volume of the metallic phase may be ignored.

The magnetoresistance in formula (35') depends on the angle between the magnetic field and direction of the easy axis. Therefore, this expression is valid, strictly speaking, only for single crystals. However, as was demonstrated, the main contribution to H_a is made by the form anisotropy. If we assume that the directions of ellipsoid axes of different ferrons do not coincide and are distributed by some law $Q(\beta)$ (which depends, generally speaking, on the magnetic field), we can write, instead of Eq. (35'),

$$\mathrm{MR}_2(H) = \gamma\frac{2}{225}\frac{M_0^3 \bar{J}^2 H_\alpha H^2}{(kT)^5}\left(1 + \frac{2M_0 H_a}{7kT}\right), \tag{36}$$

where $\gamma = \int d\beta \sin\beta Q(\beta)(\cos^2\beta - 1/3)$ and may be both positive and negative. In the case of $Q(\beta) = 1$, we have $\gamma = 0$. However, one can expect that $\gamma > 0$, because it is preferable that the long axes of elliptic magnetic droplets line up along the field. Therefore, in the region of weak fields, we will finally derive

$$\mathrm{MR}(H) = \frac{3}{100}\frac{M_0^2 H^2}{(kT)^2}\left(1 + \gamma\frac{8}{27}\frac{M_0 \bar{J}^2 H_a}{(kT)^3}\right), \tag{37}$$

where we ignored the second term in parentheses in formula (36).

Depending on the relationships between the parameters in the temperature region being treated, the second term in parentheses in formula (37) may be both greater and less than unity. Therefore, the magnetoresistance may decrease with increasing temperature both by the $1/T^2$ law and by the $1/T^5$ law. Note that the behavior of magnetoresistance, which obeyed the H^2/T^5 law, was observed in a wide temperature range for the $(La_{1-x}Pr_x)_{0.7}Ca_{0.3}MnO_3$ system [8]. In order of magnitude, the value of magnetoresistance agrees with the foregoing estimate.

We will now treat the region of strong fields. The estimation of the Coulomb energy reveals that the factor with bH in the exponent in formula (25) is of the order of unity at $T \sim 100$ K. Therefore, prior to exponential growth, the magnetoresistance curve has a plateau. In this region, we combine expressions (25) and (31) to derive

$$\mathrm{MR}(H) = \frac{\bar{J}^2}{3(kT)^2} + bH\left(\frac{V(0)}{10kT} + \frac{3x_0}{5}\right), \qquad (38)$$

where we have taken into account the fact that, proceeding from the estimates of α_a and $\bar{J}$, the preexponential factor in formula (25) may be approximately taken to be equal to $1 + j^2/3$.

The region of fields treated by us is restricted to values of the order of 10 to 20 T. This is due to the fact that we ignored the nonlinear dependence of the droplet radius on the magnetic field. In addition, a percolation transition of the system into the metallic state is possible in the region of high fields. Therefore, the exponential growth of magnetoresistance in the case of high fields arises at low temperatures and/or at a rather high value of Coulomb energy, when the exponent in formula (25) becomes of the order of unity. Then, at $j > 1$, formula (25) will yield

$$\mathrm{MR}(H) \approx \frac{\bar{J}}{kT}\exp\left(\frac{V(0)bH}{10kT}\right). \qquad (39)$$

Note that the magnetoresistance in the case being treated may amount to several hundred percent even in the nonmetallic phase.

ACKNOWLEDGMENTS

We are grateful to N.A. Babushkina, Ya.M. Blanter, P. Woelfle, A.B. Granovskiĭ, P. Fulde, and D.I. Khomskiĭ for valuable discussions.

This study received financial support from the Russian Foundation for Basic Research (project nos. 00-15-96570, 00-02-16255, and 02-02-16708), CRDF (grant no. RP2-2355-MO-02), INTAS (grant no. 01-2008), and NWO (grant no. 047-008-017), as well as a stipend of the President of the Russian Federation (grant no. 00-15-9654).

REFERENCES

1. E. L. Nagaev, Phys. Rep. **346**, 387 (2001).
2. E. Dagotto, T. Hotta, and A. Moreo, Phys. Rep. **344**, 1 (2001).
3. M. Yu. Kagan and K. I. Kugel', Usp. Fiz. Nauk **171**, 577 (2001) [Phys. Usp. **44**, 553 (2001)].
4. É. L. Nagaev, Pis'ma Zh. Éksp. Teor. Fiz. **6**, 484 (1967) [JETP Lett. **6**, 18 (1967)].
5. M. Yu. Kagan, K. I. Kugel', and D. I. Khomskiĭ, Zh. Éksp. Teor. Fiz. **120**, 470 (2001) [JETP **93**, 415 (2001)].
6. M. Yu. Kagan, D. I. Khomskii, and M. V. Mostovoy, Eur. Phys. J. B **12**, 217 (1999).
7. A. L. Rakhmanov, K. I. Kugel, Ya. M. Blanter, and M. Yu. Kagan, Phys. Rev. B **63**, 174424 (2001).
8. N. A. Babushkina *et al.*, private communication.
9. M. B. Salamon and M. Jaime, Rev. Mod. Phys. **73**, 583 (2001).
10. M. A. Krivoglaz, Usp. Fiz. Nauk **111**, 617 (1973) [Sov. Phys. Usp. **16**, 856 (1973)].
11. P. Horsch, J. Jaklic, and F. Mack, Phys. Rev. B **59**, R14149 (1999).
12. J. S. Smart, *Effective Field Theories of Magnetism* (Saunders, London, 1966; Mir, Moscow, 1968).
13. N. F. Mott and E. A. Davis, *Electronic Processes in Non-Crystalline Materials* (Clarendon, Oxford, 1979; Mir, Moscow, 1982).
14. L. D. Landau and E. M. Lifshitz, *Course of Theoretical Physics*, Vol. 8: *Electrodynamics of Continuous Media* (Nauka, Moscow, 1982; Pergamon, New York, 1984).
15. M. Yu. Kagan, M. V. Mostovoy, and D. I. Khomskii, Physica B (Amsterdam) **284–288**, 1209 (2000).
16. M. Hennion, F. Moussa, G. Biotteau, *et al.*, Phys. Rev. Lett. **81**, 1957 (1998).
17. Yu. A. Izyumov and Yu. I. Skryabin, Usp. Fiz. Nauk **171**, 121 (2001) [Phys. Usp. **44**, 109 (2001)].

Translated by H. Bronstein

Journal of Experimental and Theoretical Physics, Vol. 95, No. 4, 2002, pp. 762–767.
Translated from Zhurnal Éksperimental'noĭ i Teoreticheskoĭ Fiziki, Vol. 122, No. 4, 2002, pp. 879–885.

SOLIDS
Electronic Properties

Spin-Current-Induced Classical and Quantum Effects in the Dynamics of a Mesoscopic Magnet

A. K. Zvezdin* and K. A. Zvezdin

Institute of General Physics, Russian Academy of Sciences, ul. Vavilova 38, Moscow, 119991 Russia
**e-mail: zvezdin@nnet.org*

Received April 22, 2002

Abstract—The dynamics of a high-spin quantum system with magnetic anisotropy of the easy plane type under the action of spin-polarized current permeating this system is considered. The spin-polarized current (with electron spins polarized along the hard magnetic axis of the system) induces the reorientation of the magnetic moment of the system from the easy plane to the hard magnetic axis. Analytical expressions describing characteristics of the reorientation process in the limiting cases of strong and weak dissipation are obtained. Under strong dissipation conditions, the reorientation is shown to have a threshold character with "soft" (continuous) displacement of the magnetic moment from the easy plane. Under weak dissipation, the reorientation occurs as a discrete process, that is, is accompanied by magnetic moment jumps and hysteresis as the spin current increases and decreases. At a fairly low temperature and weak damping, quantum effects arise in the system. The spin current induces excitations quasi-anionic in character, Bloch oscillations of magnetic moment precession, and tunneling between different precession quantum modes. These quantum effects, in particular, manifest themselves in the system under consideration by magnetic moment jumps and magnetic susceptibility peaks. *© 2002 MAIK "Nauka/Interperiodica".*

1. Recently, a new mechanism of the remagnetization of magnets has been suggested [1, 2] and experimentally substantiated [3]. The essence of the mechanism is as follows. Spin-polarized current passing through a magnetized sample creates a torque moment, which acts on the magnetic moment of the sample and can change its magnitude and direction. Current spin polarization arises when the current passes through a ferromagnetic metal; the polarization is retained along some length when the current goes out of the metal. This length depends on the properties of the surface and the material into which the polarized current passes (this length can be 1 μm or larger for some metals and semiconductors). The spin-polarized current is proportional to the flux of angular momentum, and a change in this current (called spin current in the problem under consideration) in passing through a magnetized sample determines the torque moment value that acts on sample spins. This new effect attracts much attention in nanoelectronics and nanomagnetism, because it opens up a possibility for creating advanced devices. Various aspects of the effect under consideration were discussed in [4–17].

In this work, we study the dynamics of the remagnetization of a mesoscopic magnetic sample with rhombic anisotropy under the spin current action. Attention is focused on new spin-polarized-current-induced quantum effects, namely, Bloch oscillations and Zener macroscopic tunneling.

2. Consider the dynamics of the magnetic moment of a magnetic molecule (nanocluster) or nanoparticle with easy plane-type magnetic anisotropy[1] situated between two planar contacts, at least one of which is ferromagnetic. It is assumed that the gap between contacts F_1 and F_2 is sufficiently narrow for tunnel current to pass between them. The tunnel current from a ferromagnet is known to be spin-polarized. Let the degree of polarization of electrons be p. For the Fe, Co, Ni, and Cd ferromagnets, p equals 0.40, 0.35, 0.23, and 0.14, respectively [18]. Of special interest are so-called ferromagnetic halfmetals with $p \approx 1$. These are $LaMnO_3$, CrO_2, Fe_3O_4, and Heisler alloys. Instead of a dielectric, a metallic nonmagnetic interlayer with embedded nanoparticles or rare-earth metal ions (such as Ho, Tb, etc.)[2] can be used. As far as nanoparticles (magnetic molecules or rare-earth metal ions) are concerned, it is assumed that, in the ground state, they possess a magnetic moment and magnetic anisotropy of the easy plane type but with some fairly small azimuthal anisotropy in the easy plane.

Another topology of experiments can be based on the use of conducting organic molecules, which are grown by self-assembly methods and are in contact with a nonmagnetic or ferromagnetic metal (for instance, Au or Ni).[3] Such a contact is formed as a result of chemisorption of so-called thiol groups at the

[1] More exactly, we consider rhombic anisotropy; namely, it is assumed that weak uniaxial anisotropy exists in the easy plane.
[2] See [31], where the quantum dynamics of Ho^{3+} ions in $Li(Y,Ho)F_4$ crystals was studied.
[3] See [32] on this matter.

organic molecule end that is in contact with a metallic electrode. Examples of such organic molecules with thiol terminal groups are benzene-thiol molecules. A magnetic element (a magnetic nanocluster or an ion with easy-plane-type anisotropy[4]) is attached to the other organic molecule end chemically or by van der Waals forces. Used as the second contact can be a magnetic metallic (for instance, nickel) needle of an atomic-force microscope. Such a configuration of experiments is currently typical of nanophysics and nanoelectronics.

For definiteness, let the hard magnetic axis of the molecule coincide with the *z* axis.

In the pioneering work by Slonczewski [1], the following generalization of the Landau–Lifshits equation to the situation when a magnetic particle is under the action of spin-polarized current was suggested:

$$\frac{d\mathbf{n}}{dt} = \left[[\mathbf{n}\times\gamma\mathbf{H}_{\text{eff}}] - \alpha\dot{\mathbf{n}} + \frac{I_e g}{eS}[\mathbf{n}_e\times\mathbf{n}]\right], \qquad (1)$$

where $\mathbf{n}$ is the unit vector directed along the total magnetic moment of the particle; $\mathbf{n}_e$ is the unit vector directed along the mean spin of the electronic current; γ is the gyrom agnetic ratio; α is the dimensionless damping constant; I_e is the total current permeating the particle;

$$g = \frac{1}{-4 + (1+p)^3(3+\mathbf{n}_e\cdot\mathbf{n})/4p^{3/2}} > 0; \qquad (2)$$

S is the total spin of the magnetic particle; p is the spin polarization of the current defined as

$$p = \frac{\rho_\uparrow - \rho_\downarrow}{\rho_\uparrow + \rho_\downarrow}, \qquad (3)$$

where $\rho_\uparrow(\rho_\downarrow)$ is the density of electrons with spins directed along the z axis (in the opposite direction); and $\mathbf{H}_{\text{eff}} = \mathbf{H} + \mathbf{H}_A$, where $\mathbf{H}$ is the external field and $\mathbf{H}_A$ is the anisotropy field.

3. Equation (1) in spherical coordinates θ, φ, where polar angle θ is counted from the z axis and the azimuthal angle φ, from the x axis, which coincides with the easy axis, has the form

$$\dot{\theta}\sin\theta + \alpha\dot{\varphi}\sin^2\theta = -\frac{\gamma}{M}\frac{\partial E}{\partial\varphi} + \frac{Ig}{Se}\sin^2\theta, \qquad (4)$$

$$\alpha\dot{\theta} - \dot{\varphi}\sin\theta = -\frac{\gamma}{M}\frac{\partial E}{\partial\theta}, \qquad (5)$$

where

$$E = K_1\cos^2\theta + K_2\sin^2\theta\sin^2\varphi - MH\cos\theta, \qquad (6)$$
$$0 < K_2 \ll K_1.$$

[4] Currently, a substantial number of nanoclusters with such anisotropy have been synthesized and studied [33]; also see [34], where Langmuir–Blodgett films with Mn_{12} magnetic molecules were prepared.

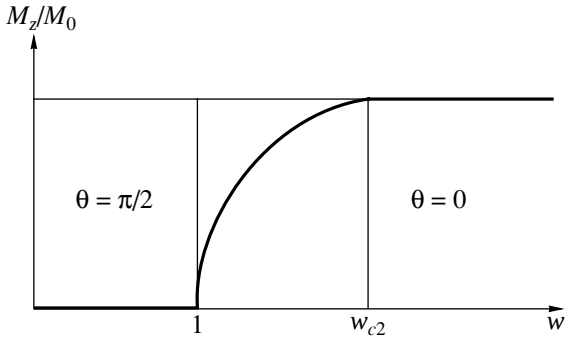


Fig. 1. Reduced magnetic moment M_z/M_0 as a function of $w = j/j_{c1}$, where $j_{c1} = \omega_2/2$, for the strong damping case with $\alpha\omega_1 \gg \omega_p$.

Substituting (6) into (4) and (5) yields

$$\dot{\theta} + \alpha\dot{\varphi}\sin\theta = -\frac{\omega_2}{2}\sin\theta\sin2\varphi + j\sin\theta, \qquad (7)$$

$$\alpha\dot{\theta} - \dot{\varphi}\sin\theta = (\omega_1 - \omega_2\sin^2\varphi)\sin\theta\cos\theta - \omega_H\sin\theta, \qquad (8)$$

where

$$\omega_1 = \gamma\frac{2K_1}{M}, \quad \omega_2 = \gamma\frac{2K_2}{M}, \quad \omega_H = \gamma H, \quad j = \frac{gI}{Se}. \qquad (9)$$

The ground state of the system (at $H = 0$) corresponds to $\theta = \pi/2$ and $\pi = 0, \pi$. It is easy to see that the resonance frequency (the ferromagnetic resonance frequency) at $\alpha = 0$ is $\omega_p = \sqrt{\omega_1\omega_2}$. Damping in the system is characterized by the frequency $\omega_d = \alpha\omega_1$. It is natural to consider two limiting cases, those of strong ($\omega_d \gg \omega_p$) and weak ($\omega_d \ll \omega_p$) damping.

4. The case of strong damping. If $\omega_d \gg \omega_p$, (7) and (8) take the form

$$\alpha\dot{\varphi} + \frac{\omega_2}{2}\sin2\varphi = j, \qquad (10)$$

$$\cos\theta = \frac{\omega_H}{\omega_1} - \frac{1}{\omega_1}\dot{\varphi}. \qquad (11)$$

Here, the condition $\omega_2 \ll \omega_1$ is also used. Equation (10) with the initial condition $\varphi(0) = 0$ has the solution

$$\varphi = \begin{cases} 2\arctan\left[\left(\dfrac{4j^2-\omega_2^2}{4j^2}\right)^{1/2}\tan\dfrac{\pi t}{T} - \dfrac{\omega_2}{2j}\right], & j \geq \dfrac{\omega_2}{\alpha}, \\ \dfrac{1}{2}\arcsin\left(\dfrac{2j}{\omega_2}\right), & j \leq \dfrac{\omega_2}{\alpha}. \end{cases} \qquad (12)$$

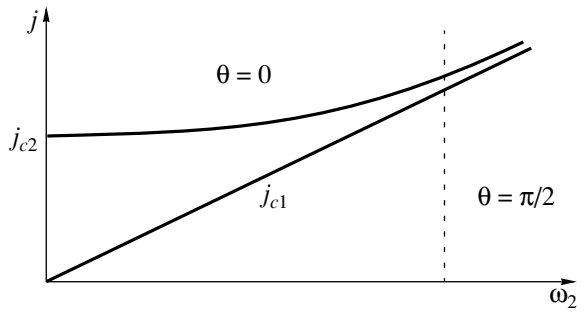


Fig. 2. Phase j–ω_2 diagram for the strong damping case.

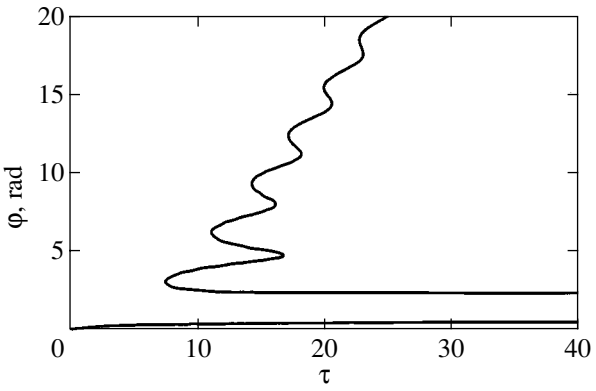


Fig. 3. Dependence $\varphi(\tau)$ at $j = 0.15\omega_1$, that is, in the region of the existence of forbidden regions for angle φ.

Here, T is the period of spin-current-induced precession,

$$T = \frac{4\pi\alpha}{\sqrt{4j^2 - \omega_2^2}}. \quad (13)$$

Magnetization $M_z = M_0\cos\theta$ is determined by (11). Clearly, magnetization oscillates if $j \geq \omega_2/\alpha$. At $H = 0$, its period T average value is

$$\langle M_z \rangle = \begin{cases} 0, & j \leq j_{c1}, \\ \dfrac{M_0}{2\alpha\omega_1}(4j^2 - \omega_2^2)^{1/2}, & j_{c1} \leq j \leq j_{c2}, \\ M_0, & j \geq j_{c2}. \end{cases} \quad (14)$$

Equation (14) describes the curve of spin-current-induced particle magnetization (Fig. 1). The critical spin currents are $j_{c1} = \omega_2/2$ and $j_{c2} = (\omega_2^2/4 + \alpha^2\omega_1^2)^{1/2}$ (at $H = 0$). Obviously, the first spin current value can be obtained from formulas (12) and (13), and the second, from the condition $M_z/M_0 = \cos\theta \leq 1$.

The phase diagram of the system under consideration is shown in Fig. 2. External field H shifts the upper boundary of the phase diagram in an obvious way.

5. The case of weak damping ($\alpha\omega_1 \ll \omega_p$). Equations (7) and (8) then take the form

$$\dot{\theta} = -\frac{\omega_2}{2}\sin\theta\sin 2\varphi + j\sin\theta, \quad (15)$$

$$\dot{\varphi} = -\omega_1\cos\theta + \omega_H. \quad (16)$$

Equations (15) and (16) have the first integral

$$\ln\frac{\sin\theta}{(\tan(\theta/2))^{\omega_H/\omega_1}} - U(\varphi) = \text{const}, \quad (17)$$

where

$$U(\varphi) = \frac{\omega_2}{4\omega_1}(1 - \cos 2\varphi) - \frac{j}{\omega_1}\varphi.$$

At $H = 0$, (17) gives

$$\sin\theta = \exp\left(\frac{\omega_2}{2\omega_1}\sin^2\varphi - \frac{j}{\omega_1}\varphi\right). \quad (18)$$

Substituting (18) into (16) yields

$$\dot{\varphi} = -\omega_1\left(1 - \exp\left(\frac{\omega_2}{\omega_1}\sin^2\varphi - \frac{2j}{\omega_1}\varphi\right)\right)^{1/2}. \quad (19)$$

Here, the initial condition is $\varphi(t = 0) = 0$. The integration of (19) results in

$$\tau = \int_0^{\varphi} \frac{du}{\left(1 - \exp\left(\frac{\omega_2}{\omega_1}\sin^2 u - \frac{2j}{\omega_1}u\right)\right)^{1/2}}, \quad (20)$$

where $\tau = -\omega_1 t$.

Equations (16) and (18–20) determine the time τ dependence of magnetization $M_z = M_0\cos\theta$ in the parametric form

$$\cos\theta = \left(1 - \exp\left(\frac{\omega_2}{\omega_1}\sin^2\varphi - \frac{2j}{\omega_1}\varphi\right)\right)^{1/2}, \quad (21)$$

$$\tau = \int_0^{\varphi} \frac{du}{\cos\theta(u)}. \quad (22)$$

This parametric form makes it possible to comparatively easily numerically construct the dependences $\cos\theta(\tau)$ and $\varphi(\tau)$. The characteristic curves are shown in Figs. 3–5. Their analysis leads us to the following conclusions about the dynamics of the spin-current-induced reorientation of the magnetic moment of a nanoparticle from the easy plane to the z axis.

(1) There exists a critical spin current value $j_1 = 0.3623\omega_1$ below which the $\varphi(\tau)$ dependence has several branches. Between these branches, there is a region ("gap") of angle φ values that are forbidden for magnetic moment orientations. Figure 3 illustrates this sit-

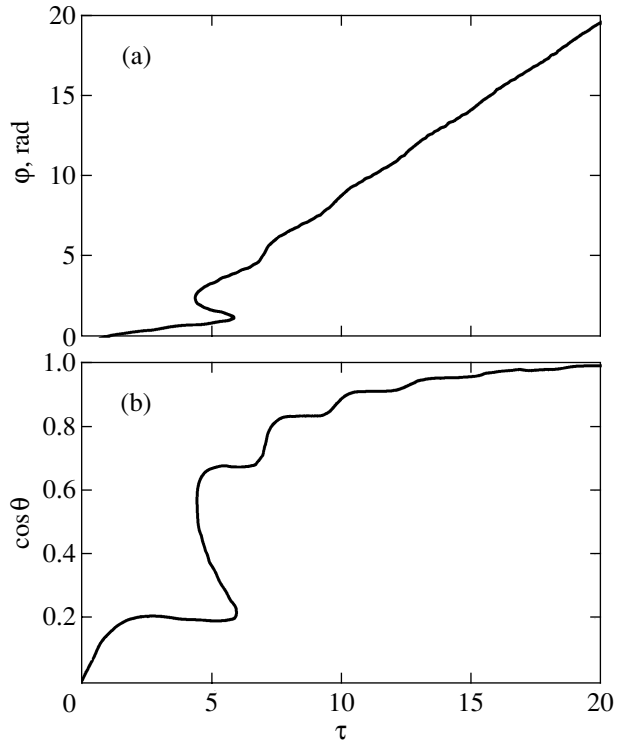


Fig. 4. Dependences $\varphi(\tau)$ and $\cos\theta(\tau)$ at $j = 0.45\omega_1$, that is, in the $j_1 < j < j_2 \sim 0.65\omega_1$ region, when there are no forbidden regions for angle φ, but the $\varphi(\tau)$ and $\theta(\tau)$ dependences remain ambiguous.

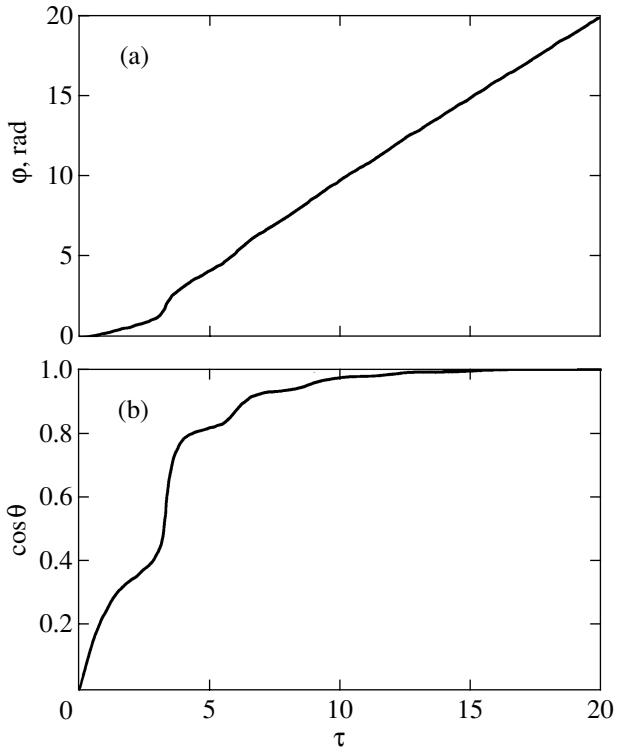


Fig. 5. Dependences $\varphi(\tau)$ and $\cos\theta(\tau)$ at $j = 0.75\omega_1$, that is, in the $j > j_2$ region, where the dynamics of spin-current-induced **m** reorientations from the easy plane to the hard magnetic axis is smooth, although nonmonotonic.

uation. In this figure, the lower branch corresponds to the localized state (finite motion) of the magnetic moment, and the upper branch, to precession with a gradual approach to axis *z*. The figure also shows that precession is not continuous and is accompanied by jumps and hystereses. This means that the forward and backward $\varphi(\tau)$ trends do not coincide, which also directly follows from the observation that $\varphi(\tau)$ is an ambiguous dependence.

(2) At $j_1 < j < j_2 \approx 0.65\omega_1$, there are no forbidden φ angle regions, but the $\varphi(\tau)$ dependence remains ambiguous. The $\varphi(\tau)$ and $\theta(\tau)$ angles therefore experience jump changes, and the magnetic moment dynamics exhibits hystereses (Figs. 4a, 4b).

(3) At $j > j_2$, the dynamics of spin-current-induced magnetic moment reorientation from the easy plane to axis *z* is smooth (Figs. 5a and 5b).

6. Quantum effects should be taken into account in the dynamics of the system at fairly low temperatures for small-sized nanoparticles. Consider this problem under the condition that the angle of departure of the magnetic moment out of the easy plane, that is, $\psi = \pi/2 - \theta$, is a small parameter. It is then natural to consider the case of weak dissipation, when quantum effects are most noticeable. At $\psi \ll 1$, (16) gives

$$\psi = \frac{\omega_H - \dot{\varphi}}{\omega_1}. \tag{23}$$

Substituting (23) into (15) yields

$$\ddot{\varphi} + \frac{1}{2}\omega_p^2 \sin 2\varphi = j\omega_1 + \dot{\omega}_H. \tag{24}$$

The Lagrange and Hamilton functions of (24) have the form

$$L = \frac{J\dot{\varphi}^2}{2} - U(\varphi), \tag{25}$$

$$H = \frac{P_\varphi^2}{2J} + U(\varphi), \tag{26}$$

where $J = M/2\gamma\omega_1$, and generalized momentum P_φ and potential energy $U(\varphi)$ are given by

$$P_\varphi = \frac{\partial L}{\partial \dot{\varphi}} = J\dot{\varphi} \tag{27}$$

and

$$U(\varphi) = K_2 \sin^2\varphi - \frac{M}{\gamma}\left(j + \frac{\dot{\omega}_H}{\omega_1}\right)\varphi, \tag{28}$$

respectively. Note that magnetization $M_z = M_0\cos\theta$ can be expressed via P_φ as follows:

$$M_z = \frac{M^2}{2K_1} - 2\gamma P_\varphi. \quad (29)$$

Equations (24)–(28) are isomorphic to the corresponding equations describing the dynamics of the Josephson transition in the classical approximation.

In 1963, Anderson [19] suggested that macroscopic quantum effects can be described as the dynamics of the Josephson transition if the charge and the phase, which are the generalized momentum and coordinate of the system, respectively, are treated as operators. Similar quantities in our problem are magnetization $M_z = M_0\cos\theta$ and angular variable φ or, according to (29), P_φ and φ. The commutation relation for P_φ and φ has the form

$$[P_\varphi, \varphi] = i\hbar. \quad (30)$$

Taking (26)–(30) into account, we can write the Schrödinger equation

$$i\hbar\dot{\Psi} = \left(\frac{P_\varphi^2}{2J} + \frac{1}{2}K_2(1-\cos 2\varphi) - F\varphi\right)\Psi, \quad (31)$$

where $\Psi(\varphi, t)$ is the wave function describing precession of the total spin of the nanoparticle,

$$F = \frac{M}{\gamma}\left(j + \frac{\dot{\omega}_H}{\omega_1}\right). \quad (32)$$

Equation (31) was thoroughly studied in [23] (also see [24, 25]) for the problem of the dynamics of the quasi-classical spin under the action of a magnetic field with a linear time dependence. In this problem, the F quantity [Eq. (32)] plays the role of the rate of field variations with time.

The Schrödinger equation with "washboard"-type potential energy as in (31) has been studied earlier, for example, in the problem of motion of the electron in a crystal under the action of a uniform constant electric field [20, 21] and in the problem of weak superconductivity [22].

Equation (31) is similar to the Schrödinger equation for a particle moving in a periodic potential under the action of "force" F. We refer the reader to [23–25] for details and only give the results of the analysis performed there. The system has a band spectrum characteristic of electrons in crystals, and the wave functions are Bloch functions determined by continuous parameter m (quasi-spin). The corresponding excitations are quasi-particles with a fractional (or continuous) spin number and are very close to anions in their properties [26–28].

The widths of the lower zeroth and first allowed bands are $2\hbar^2/J$ and $6\hbar^2/J$, respectively. The lower forbidden band equals K_2, and the higher forbidden bands rapidly decrease in magnitude.

Spin current (and field $\dot{H}$) excite coherent quantum oscillations of the precession motion of the magnetic moment. For instance, the magnetic moment with a zero mean quasi-spin value at the initial time shifts to the Brillouin zone boundary, is reflected from it, then propagates to the left Brillouin zone boundary, is again reflected, and so on. This process is called Bloch oscillations. Their frequency in this case is

$$f_{\text{Bloch}} = \left(j + \frac{\dot{\omega}_H}{\omega_1}\right)S. \quad (33)$$

Each reflection from the Brillouin zone boundary manifests itself by a jump of angle θ, that is, of M_z.

If spin current (or external magnetic field) has a harmonic component of frequency f, resonances at frequencies $f_r = rf_{\text{Bloch}}$ are possible, where r is a rational number. Such resonances are called Stark resonances.

When spin current increases to

$$\frac{M}{\gamma}\left(j + \frac{\dot{\omega}_H}{\omega_1}\right) > K_2, \quad (34)$$

the Zener tunnel effect between neighboring allowed bands arises [29, 30]. In particular, the probability of the tunnel transition per unit time between two lower bands is given by

$$g_{01} = f_{\text{Bloch}}\exp(-\beta), \quad (35)$$

where $\beta = \pi K_2^2/2\hbar^2(\dot{\omega}_H + j\omega_1)$.

The $|\psi\rangle$ wave function of the process should then be treated as the superposition of two amplitudes, that is, $|\psi\rangle = c_1|0\rangle + c_2|1\rangle$, where $|0\rangle$ and $|1\rangle$ are the wave functions that describe Bloch oscillations in the zeroth and first allowed zones.

7. Consider several numerical estimates. Suppose that nanoparticles are a thin-layer element of thickness $d \sim 1$ nm. Set $K_1 \sim 10^6$ erg/cm^3 and $M_s = 0.8\times 10^3$ Gs cm^3. We then have $\omega_1 = 2\gamma K_1/M_s \sim 10^{11}$ s^{-1} and $\omega_2 \approx 0.01\omega_1 = 10^9$ s^{-1}; this is the case of weak damping with the critical values $j_{c1} = 3.5\times 10^{10}$ s^{-1} and $j_{c2} = 6.5\times 10^{10}$ s^{-1}, which correspond to the current densities through the structure $I_1 = 3.5\times 10^8$ A/cm^2 and $I_2 = 6.5\times 10^8$ A/cm^2, respectively.

The Bloch oscillation frequency estimated by (33) is 10^8 Hz ($f_{\text{Bloch}} \sim gI/e$) at an about 10^4-A/cm^2 current density and a 1-nm nanoparticle diameter. For such nanoparticles, the Zener breakdown begins at current densities of the order of 10^7–10^8 A/cm^2.

8. To summarize, we obtained analytic expressions that characterized the reorientation of the magnetic moment of nanoparticles with rhombic anisotropy from the easy plane to the hard magnetic axis under the action of spin current. At low temperature and weak damping, spin current was shown to induce coherent quantum effects in nanoparticles, namely, Bloch oscil-

lations of precession magnetization motion and tunnel transitions between different precession quantum modes.

ACKNOWLEDGMENTS

One of us (A.K.Z.) wishes to express sincere thanks to B. Barbara and R. Giraud for discussions and warm hospitality at Laboratoire L. Néel (Grenoble). This work was financially supported by the Russian Foundation for Basic Research (project no. 02-02-17389), the Physical Properties of Nanostructures Program, and INTAS (grant no. 99-01839).

REFERENCES

1. J. C. Slonczewski, J. Magn. Magn. Mater. **159**, L1 (1996).
2. L. Berger, Phys. Rev. B **54**, 9353 (1996).
3. M. V. Tsoi, A. G. M. Jansen, J. Bass, *et al.*, Phys. Rev. Lett. **80**, 4281 (1998).
4. L. Berger, J. Appl. Phys. **49**, 2156 (1978).
5. L. Berger, J. Appl. Phys. **55**, 1954 (1984).
6. P. P. Frutas and L. Berger, J. Appl. Phys. **57**, 1266 (1985).
7. C.-Y. Hung and L. Berger, J. Appl. Phys. **63**, 4276 (1988).
8. J. C. Slonczewski, J. Magn. Magn. Mater. **195**, L261 (1999).
9. Ya. B. Bazaliy, B. A. Jones, and Shou-Cheng Zhang, cond-mat/0009034.
10. J. E. Wegrowe *et al.*, Europhys. Lett. **45**, 626 (1999).
11. E. B. Myers *et al.*, Science **285**, 867 (1999).
12. J. A. Katine *et al.*, Phys. Rev. Lett. **84**, 3149 (2000).
13. J. Z. Sun, J. Magn. Magn. Mater. **202**, 157 (1999).
14. C. Heede, P. E. Zilberman, and R. J. Elliott, cond-mat/0005064.
15. Ya. B. Bazaliy *et al.*, Phys. Rev. B **57**, R3213 (1998).
16. J. Z. Sun, Phys. Rev. B **62**, 570 (2000).
17. L. Berger, J. Appl. Phys. **81**, 4880 (1997).
18. W. P. Pratt *et al.*, Phys. Rev. Lett. **66**, 3060 (1991).
19. P. W. Anderson, in *Lectures on the Many-Body Problem*, Ed. by E. Caianielo (Academic, New York, 1964), Vol. 2, p. 113.
20. G. H. Wannier, Phys. Rev. **117**, 432 (1960).
21. F. Bloch, Phys. Rev. Lett. **137**, A787 (1965); **166**, 415 (1968).
22. K. K. Likharev, *Introduction to the Dynamics of Josephson Junctions* (Nauka, Moscow, 1985).
23. A. K. Zvezdin, Kratk. Soobshch. Fiz. (Bull. Lebedev Phys. Inst. Russ. Acad. Sci.), No. 12, 13 (1999).
24. A. K. Zvezdin, Kratk. Soobshch. Fiz. (Bull. Lebedev Phys. Inst. Russ. Acad. Sci.), No. 4 (2000).
25. A. K. Zvezdin, Kratk. Soobshch. Fiz. (Bull. Lebedev Phys. Inst. Russ. Acad. Sci.), No. 3, 37 (2001).
26. J. M. Leinaas and J. Myrheim, Nuovo Cimento **37**, 1 (1977).
27. F. Wilczek, Phys. Rev. Lett. **48**, 1144 (1982); **49**, 957 (1982).
28. S. Rao, hep-th/9209066.
29. C. Zener, Proc. R. Soc. London, Ser. A **145**, 523 (1934).
30. L. V. Keldysh, Zh. Éksp. Teor. Fiz. **33**, 994 (1958) [Sov. Phys. JETP **6**, 763 (1958)].
31. R. Giraud, W. Wernsdorfer, A. Tkachuk, *et al.*, Phys. Rev. Lett. **87**, 057203 (2001).
32. M. A. Reed *et al.*, Science **278**, 253 (1997); J. K. Gimzewski and C. Joachim, Science **283**, 1683 (1999); J. Chen *et al.*, Science **286**, 1550 (1999); C. P. Collier *et al.*, Science **289**, 1172 (2000); E. G. Emberley and G. Kirczenow, cond-mat/0201344.
33. W. Wernsdorfer, private communication.
34. M. Clemente-Leon, H. Soyer, E. Coronado, *et al.*, Angew. Chem. Int. Ed. Engl. **37**, 2842 (1998).

Translated by V. Sipachev

Journal of Experimental and Theoretical Physics, Vol. 95, No. 4, 2002, pp. 768–776.
Translated from Zhurnal Éksperimental'noĭ i Teoreticheskoĭ Fiziki, Vol. 122, No. 4, 2002, pp. 886–896.

SOLIDS
Electronic Properties

The Instability of the Front of Magnetization Reversal in Anisotropic Superconductors

A. L. Rakhmanov[a], L. M. Fisher[b, *], V. A. Yampol'skiĭ[c], M. Baziljevich[d], and T. H. Johansen[d]

[a]*Institute of Theoretical and Applied Electrodynamics, IVTAN (Institute of High Temperatures) Scientific Association, Russian Academy of Sciences, Moscow, 127412 Russia*

[b]*State Scientific Center–State Unitary Enterprise All-Russia Electrotechnical Institute, Moscow, 111250 Russia*

[c]*Usikov Institute of Radiophysics and Electronics, National Academy of Sciences of Ukraine, Kharkov, 61085 Ukraine*

[d]*Department of Physics, University of Oslo, Oslo, Norway*

**e-mail: fisher@vei.ru*

Received April 23, 2002

Abstract—The mechanism of instability of the front of magnetization reversal, i.e., of the boundary separating the regions of existence of vortices with mutually opposite orientations of the magnetic flux (vortex and antivortex regions) in type II superconductors, is suggested. The instability is associated with the anisotropy of flow of a vortex fluid, caused by planar defects in the *ab* plane of high-temperature superconducting single crystals of the 1–2–3 system. The anisotropy of the dynamic properties of vortex matter brings about a jump of the tangential component of the vortex velocity at the front of magnetization reversal; as is known from the classical hydrodynamics of viscous fluids, this leads to the turbulization of flow. It is demonstrated that the hydrodynamic approach to the description of vortex flow helps explain the emergence of a positive increment of the rise of vortex density fluctuations under conditions of a fairly strong anisotropy. The results of magnetooptical investigations of macroturbulence in the vortex system of Y-123 single crystal with a high density of the twinning boundaries lend qualitative support to the theoretical inferences. *© 2002 MAIK "Nauka/Interperiodica".*

1. INTRODUCTION

The dynamics of vortex matter in type II superconductors have been subjected to intensive studies since the late 1950s, starting with the pioneering work by A.A. Abrikosov. More recently, primary attention was given to hard superconductors, whose magnetic characteristics are defined by the presence of Abrikosov vortex pinning centers. The main features characterizing the nonuniform penetration of magnetic flux into such systems were revealed and studied; various theoretical models of electrodynamic processes in superconductors were suggested. The new outburst of activities in this field was initiated by the discovery of high-temperature superconductors. An important characteristic feature of newly discovered superconductors is their layered structure leading to a strong anisotropy of the current-carrying capacity. In addition, the existence of a superconducting state at relatively high temperatures made necessary a more consistent inclusion of thermal fluctuations and the study of diverse phase transitions in a vortex system. Many of the newly obtained results are described in large review papers [1, 2].

The use of high-resolution magnetooptical equipment made possible a detailed study of the dynamics of magnetic flux in superconductors. One of the most important characteristics revealed using this method is the fractal or turbulent structure of thermally activated flow of magnetic flux [3–6]. One of the most dramatic phenomena of this type is the macroturbulent instability of the magnetic flux on the boundary of the front of magnetization reversal that separates regions in which vortices are oppositely directed (vortices and antivortices) [3–5]. Note that the macroturbulence was revealed only in single-crystal samples of the 1–2–3 system.

The effect of macroturbulence essentially consists in the following. If an external magnetic field of opposite sign is applied to a superconductor with a trapped magnetic flux, a boundary arises in the sample separating the regions occupied by vortices and antivortices. Here and in what follows, for definiteness, we apply the term "antivortices" to the vortices whose direction coincides with that of the external magnetic field and the term "vortices" to the vortices which were present in the sample originally, prior to switching on a magnetic field of negative sign. In some range of magnetic fields and temperatures, the stationary plane interface between vortices and antivortices becomes unstable. A disordered magnetic flux flow arises at the front of magnetization reversal, which resembles a turbulent fluid flow. This process rapidly develops in time and is accompanied by the emergence of channels via which the antivortices penetrate into the region taken by the vortices. In other words, the front of magnetization reversal assumes a "fingerlike" shape. The annihilation of vortices and antivortices occurs at the front, and the process of macroturbulence soon ends in the complete disappearance of vortices. This pattern of penetration of

magnetic flux differs qualitatively from the steady-state slow motion of the front of magnetization reversal upon initial switching on of the magnetic field, when vortices of only one direction are present in the sample. Note that the characteristic times of instability development amount to seconds and more, and the arising spatial structures are macroscopic, i.e., contain a large number of single vortices.

Experiments performed to study macroturbulence usually involve the use of platelike single crystals placed in a transverse magnetic field; the magnetooptical image provides information about the distribution of the normal component of magnetic induction. The penetration of an electromagnetic field into a superconductor in such a geometry is of interest *per se* and was studied by numerous researchers (see, for example, [7, 8]). However, the turbulent behavior of magnetic flux is not a geometric effect. It was observed both in thin plates in a transverse field and in single crystals with a low demagnetizing factor. Frello *et al.* [5] reported the magnetooptical visualization of developed macroturbulence of vortex matter in an Nd-123 crystal $3.1 \times 2.5 \times 1.3$ mm^3 in size. In the magnetic field, this sample was divided into three magnetically unbound regions (each having smaller transverse dimensions) in which the turbulence developed independently. One must further take into account the fact that the instability is often observed under conditions of complete penetration of magnetic flux into the sample, when, as was demonstrated by Brandt [8], the difference in the distribution of induction in the cases of longitudinal and transverse geometry turns out to be not too significant.

Evidently, the macroturbulence cannot be understood either within the universally accepted model of critical state [9] or in the existing models of relaxation of magnetic flux in hard superconductors [10]. At the same time, this phenomenon is obviously similar to that of turbulence in hydrodynamics; therefore, its interpretation is of general physical interest.

In the helium temperature region, where the heat capacity of superconductors is low and the critical current density is high, the formation of macroscopic unsteady-state and inhomogeneous structures may be caused by instabilities in which thermal effects play an important part (see, for example, [11]). Therefore, Bass *et al.* [12] made an attempt at relating the nature of macroturbulence to the heat wave generated at the front of magnetization reversal due to the release of energy in the process of annihilation of vortices and antivortices. Unfortunately, it is unlikely that this mechanism may be responsible for macroturbulence. Indeed, the energy released during annihilation of a pair of vortices is of the order of $H_c^2 \xi^2$, where H_c is the thermodynamic critical field and ξ is the coherence length. If it is considered that the heat capacity of a superconductor at a temperature of 40 to 70 K is of the order of 10^6–10^7 erg/cm^3 K, $H_c \sim 10^3$ Oe, and $\xi \sim 10^{-7}$ cm, then, with the average magnetic induction in the sample of several hundred gauss, the sample will not be heated by more than 10^{-4} K if all vortices in it annihilate under adiabatic conditions. The heating will amount to only hundredths of a degree, even if we assume that all of the magnetic energy stored in the sample will be adiabatically released in this case. If we now recall that the macroturbulence develops slowly and covers a macroscopic fraction of the single crystal volume and that the sample is intensively cooled in the course of experiment, the assumption of the thermal nature of instability appears to be improbable.

Another possible reason for the emergence of macroturbulence was suggested by Vlasko-Vlasov *et al.* [13]. Their attention was drawn to the fact that the process of annihilation of a vortex–antivortex pair may be accompanied by the formation of spatial domains free of vortices (the so-called Meissner holes), and they assumed that the presence of such domains may cause instability. However, they did not treat the physical pattern of instability.

In this study, we gave our attention to the fact that the macroturbulence is observed only in single crystals of the 1–2–3 family. These samples, in contrast to other superconductors, are characterized by the clearly defined anisotropy of conduction in the *ab p*lane. The anisotropy in these materials is associated with the presence of two systems of mutually orthogonal twinning boundaries which are usually oriented at an angle of 45° to the single crystal boundaries (see, for example, [14]). In our opinion, this is where we can find a clue to understanding the nature of macroturbulence. Because of the presence of the twinning boundaries, the vortices and antivortices move under the effect of the Lorentz force in opposition to each other mainly along these "guiding" boundaries (the so-called guiding effect [15, 16]). As a result, the tangential component of vortex velocity suffers a discontinuity on the boundary separating the vortex and antivortex domains. It is known from the classical hydrodynamics [17] that the presence of tangential discontinuities of velocity brings about an instability of fluid flow.

Note that the electromagnetic instability of motion of a vortex system in the critical and resistive states in anisotropic hard superconductors was studied by Gurevich [18, 19]. However, his results cannot be directly used to explain the emergence of macroturbulence.

In this paper, we demonstrate that a purely hydrodynamic approach describing the motion of a system of vortices and antivortices with anisotropic viscosity makes it possible to understand the nature of the emergence of instability of the front of magnetization reversal, whose development leads to macroturbulence. In particular, the mechanism of macroturbulence caused by anisotropy enables one to interpret the special manifestations of instability in superconductors with different densities of twinning boundaries. A brief communi-

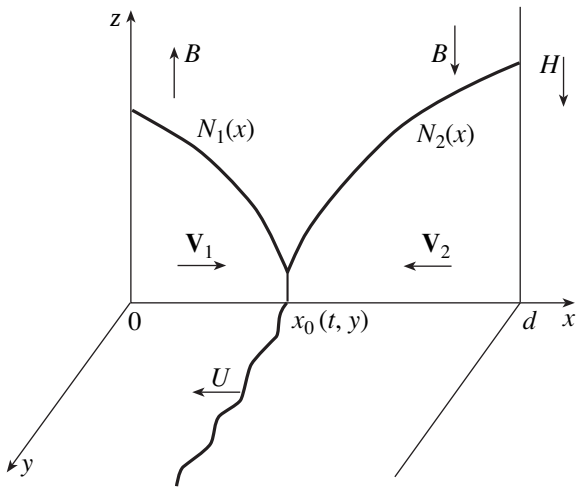


Fig. 1. The density distribution for vortices, $N_1(x)$, and antivortices, $N_2(x)$, in the right-hand half ($0 < x \le d$) of the plate.

cation reporting the results of our study was published previously [20].

2. FORMULATION OF THE PROBLEM AND BASIC EQUATIONS

We will treat an infinitely large superconducting plate of thickness $2d$, placed in an external magnetic field **H** oriented along the z axis in parallel with the sample surface. The x axis is directed normally to the plate, with the origin of coordinates $x = 0$ located at the plate center. We will assume that the magnetic field H first increased so that the magnetic flux in the form of vortices filled the sample and then decreased, passed zero, and assumed some negative value. In this case, the surface region of the plate on both sides is penetrated by vortices with the opposite direction of magnetic flux (antivortices), and three spatial domains arise in the superconductor. In the central part, vortices which originally entered the sample are located. In two peripheral regions arranged symmetrically relative to the median plane of the plate, antivortices are located. It is clear from the symmetry of the problem that it is sufficient to treat only one (for example, the right-hand, $0 < x < d$) half of the sample. The geometry of the problem is shown schematically in Fig. 1.

In hard superconductors, Abrikosov vortices are pinned on various lattice defects so as to result in a nonuniform distribution of magnetic induction. At zero temperature, this pattern turns out to be static, in spite of the Lorentz force exerted on the vortices by the neighboring vortices. At finite temperatures, owing to thermal activation, a slow flow of magnetic flux occurs because of the vortex depinning from the pinning centers. The vortex system flow may be treated within the macroscopic approach in terms of viscous flow of vortex fluid. Such a description was employed, for example, by Bass *et al.* [12]. In the situation being investigated, the magnetic flux flow is further promoted by the process of annihilation of vortices and antivortices, which occurs on the boundary $x = x_0$ separating the domains of their existence (see Fig. 1). This phenomenon leads to an additional entry of antivortices from the plate surface in an invariable external magnetic field H and reduction of the total number of vortices in the central part of the sample. As a result, the interface $x = x_0$ moves slowly at a velocity **U** into the bulk of the sample (Fig. 1).

We will designate the density of vortices and antivortices as $N_1(x)$ and $N_2(x)$, respectively. The relation between the vortex density $N_\alpha(x, y)$ ($\alpha = 1, 2$) and the magnetic induction $B(x, y)$ in the respective domain of superconductor is obvious:

$$N_\alpha(x, y) = \frac{s_\alpha B(x, y)}{\Phi_0}, \tag{1}$$

where Φ_0 is a magnetic flux quantum, $s_1 = 1$, and $s_2 = -1$. The density of vortices and antivortices must satisfy the continuity equation,

$$\frac{\partial N_\alpha}{\partial t} + \mathrm{div}(N_\alpha \mathbf{V}_\alpha) = 0, \tag{2}$$

where $\mathbf{V}_\alpha$ denotes the hydrodynamic velocity of vortices and antivortices.

The second equation for the functions N_α and $\mathbf{V}_\alpha$ is the dynamic equation for vortex fluid. Because we associate the instability of the front of magnetization reversal with the anisotropy of magnetic flux flow, we will assume the coefficient of viscosity to be the symmetric tensor of the second order η_{ik}. We will write the relation between the vortex velocity $V_{\alpha k}$ and the vortex-moving Lorentz force F_{Li} as

$$\eta_{ik} N_\alpha V_{\alpha k} = F_{Li}, \quad \mathbf{F}_L = \frac{1}{c}\mathbf{B} \times \mathbf{J}, \tag{3}$$

where the current density **J** is related to magnetic induction by the Maxwell equation

$$\mathbf{J} = \frac{c}{4\pi}\mathrm{curl}\mathbf{B}. \tag{4}$$

The phenomenological description of the dynamics of vortex matter using the set of hydrodynamic equations (2) and (3) is valid if the characteristic intervortex distance a is much less than all spatial scales of the problem, in particular, less than the characteristic scale of macroturbulence.

Equation (3) may be rewritten as follows:

$$V_{\alpha i} = -\frac{\Gamma \Phi_0^2}{4\pi}\gamma_{ik}\frac{\partial N_\alpha}{\partial x_k}, \tag{5}$$

where γ_{ik} is the dimensionless tensor of reciprocal viscosity ($\eta_{ik}^{-1} = \Gamma\gamma_{ik}$), and the coefficient Γ is selected so

that the principal values of the tensor γ_{ik} were equal to unity and ε, $0 < \varepsilon < 1$. The case of maximally strong anisotropy corresponds to $\varepsilon \longrightarrow 0$. Note that Γ is the reciprocal of viscosity and increases rapidly with temperature because of thermally activated depinning. In Eqs. (3) and (5), the inertial terms are omitted, because they are unimportant in the case of low velocities of interest to us [21].

For solving the problem, we must formulate the boundary conditions on the sample surface, as well as on the interface between the domains taken by vortices and antivortices. We will first discuss the conditions on the sample boundaries. We ignore the induction jump on the surface (which may be done in the case of fairly high values of H exceeding those of the lower critical field H_{c1}) and derive

$$N_2(d) = N_2(-d) = H/\Phi_0. \tag{6}$$

Because we treat only the right-hand part of the sample, we will replace the condition $N_2(-d) = H/\Phi_0$ by the requirement

$$V_\alpha(0) = 0, \tag{7}$$

which immediately follows from the symmetry of the problem and condition (6).

We will now turn to the conditions on the interface between the domains of existence of vortices and antivortices, which moves at some velocity toward the middle of the sample because of annihilation. The position of this interface, generally speaking, depends both on the time t and on the coordinate y. Indeed, it will be clear from our paper that, as the instability develops, the plane interface between vortices and antivortices starts winding and assumes the shape of a "washboard." Therefore, the interface is defined by the equation $x = x_0(y, t)$. Accordingly, the rate of its motion is defined by the formula

$$\begin{aligned} U_x &= \frac{\partial x_0}{\partial t}\frac{1}{\sqrt{1 + (\partial x_0/\partial y)^2}}, \\ U_y &= -\frac{\partial}{\partial t}\frac{\partial x_0/\partial y}{\sqrt{1 + (\partial x_0/\partial y)^2}}. \end{aligned} \tag{8}$$

Generally speaking, the velocity **U,** as well as $x_0(y, t)$, depends both on time and on coordinate y, $\mathbf{U} = \mathbf{U}(y, t)$.

The first one of the conditions on the interface consists in that the normal components of the vortex and antivortex flows in the frame of reference which moves along with the interface are equal in magnitude and opposite in sign,

$$N_1(\mathbf{V}_1 - \mathbf{U})_n + N_2(\mathbf{V}_2 - \mathbf{U})_n = 0. \tag{9}$$

Here, the subscript n indicates the normal vector component. The unit vector $\boldsymbol{\nu}$ of the normal to the interface is given by the obvious formulas

$$\begin{aligned} \nu_x &= \frac{1}{\sqrt{1 + (\partial x_0/\partial y)^2}}, \\ \nu_y &= \frac{\partial x_0/\partial y}{\sqrt{1 + (\partial x_0/\partial y)^2}}. \end{aligned} \tag{10}$$

The second boundary condition is defined by the rate of annihilation of vortices and antivortices. It is obvious that this rate goes to zero if the density of vortices or antivortices on the interface between them is zero. Then, we can follow the conventional approach to describing such kinetic processes and conveniently represent the rate of annihilation to be proportional to the product of vortices and antivortices,

$$N_1(\mathbf{V}_1 - \mathbf{U})_n = RN_1N_2. \tag{11}$$

A similar model of annihilation process was used by Bass *et al.* [12]. However, they assumed the annihilation parameter R to be proportional to the velocity difference $|V_1 - V_2|$. In this case, a macroscopic domain of coexistence of vortices and antivortices may arise in the sample; i.e., the vortex and antivortex domains may penetrate into each other. Unlike Bass *et al.* [12], we take into account the fact that the force of mutual attraction of vortices and antivortices in the vicinity of the interface is very strong. This force is determined by the current of individual vortices, which significantly exceeds the averaged macroscopic current J. Therefore, the relative velocity of vortices and antivortices in the region of their annihilation is many times higher than the hydrodynamic velocity V_α. As a result, the domain of coexistence of vortices and antivortices turns out to be of the same order of magnitude as the vortex lattice parameter a. In the hydrodynamic approximation, this scale is to be ignored. Therefore, we believe that the annihilation proceeds on the plane $x = x_0(y, t)$ rather than in a finite-size domain.

In principle, the annihilation parameter R may depend on the vortex velocity and density. However, the vortex densities on the interface must be treated as relatively low, and we can assume the right-hand part of equality (11) to be the first nonvanishing term of an expansion of a more general expression for the rate of annihilation in terms of vortex densities. The assumption of the smallness of the vortex velocity V_α compared with the "microscopic" velocity of annihilating vortices and antivortices enables one to ignore the dependence of R on the velocity V_α as well.

Finally, we assume that the average magnetic induction in the neighborhood of the interface is zero; i.e.,

$$N_1 = N_2 \tag{12}$$

at the point $x = x_0(y, t)$. One can readily demonstrate that this condition directly follows from Eq. (2) and relation (9) if it was valid at the moment of emergence of antivortices in the sample. In our case, $N_1 = N_2 = 0$ at the

moment when the decreasing external magnetic field assumed the value of $H = 0$.

The thus formulated equations and boundary conditions make possible, in principle, complete analysis of the behavior of a vortex system in time. The solution of the problem treated by us involves finding the undisturbed base distribution of vortex density $N_\alpha(x, t)$ and the velocity field $V_\alpha(x, t)$ and then testing for stability the linearized set of equations for the $\delta(N_\alpha(x, y, t))$ and $\delta(V_\alpha(x, y, t))$ fluctuations. The next section of the paper deals with analysis of the base profile.

3. BASE PROFILE OF VORTEX DENSITY DISTRIBUTION

At first glance, it would seem that the problem of finding the base profile is rather simple, because the position of the interface, its velocity, and the densities of vortices and antivortices, as well as their velocities, are independent of the coordinate *y*. In this case, however, one has to solve the nonlinear set of partial differential equations (2), (4), and (5). Especially complicated is the use of boundary conditions (9)–(12) on a moving interface whose position must be found self-consistently. One must further bear in mind that the problem set has no self-similar solutions with $U = \text{const} \neq 0$.

The first step will be the calculation of the base profile of another system in which the interface velocity is zero. The steady-state profile is realized, for example, in a superconducting plate through which a direct transport current passes which fills the entire cross section of the superconductor. In this case, vortices with opposite directions of magnetic flux penetrate on both sides of the plate. These vortices meet and annihilate on a stationary boundary in the middle of the plate. The problem is readily solved in this formulation. Indeed, we assume in formula (2) that $\partial N_1/\partial t = \partial N_2/\partial t = 0$ and find from relations (2)–(6) and (9)–(12) the profiles of the distributions $N_1(x)$ and $N_2(x)$ arranged symmetrically relative to the point x_0,

$$\begin{aligned} N_1(x) &= N_2(d)\left(\frac{x_0 + d/2r - x}{d + d/2r - x_0}\right)^{1/2}, \\ N_2(d) &= \frac{H}{\Phi_0}, \\ N_2(x) &= N_2(d)\left(\frac{x + d/2r - x_0}{d + d/2r - x_0}\right)^{1/2}, \\ r &= \frac{4\pi R d}{\Gamma\Phi_0^2\gamma_{xx}}. \end{aligned} \tag{13}$$

We will further assume for simplicity that the dimensionless parameter *r* is much greater than unity; i.e., we assume that the annihilation rate *R* and the coefficient of viscosity Γ^{-1} are fairly high. Then, the distributions given by Eqs. (13) demonstrate that the vortex density on the interface is relatively low, $N_1 = N_2 \sim N_2(d)r^{-1/2}$. At the same time, the spatial derivatives in the vicinity of the interface increase,

$$N_2' = -N_1' = \frac{N_1 r}{d}, \quad N_2'' = N_1'' = -N_1\left(\frac{r}{d}\right)^2. \tag{14}$$

One can readily demonstrate that the velocities $V_{\alpha x, y}(x = x_0)$ are relatively high, because they are proportional to $r^{1/2}$,

$$V_{1x, y} = \frac{\gamma_{xx, y}}{\gamma_{xx}} N_1 R \sim \frac{\Gamma\Phi_0}{4\pi d} H r^{1/2}. \tag{15}$$

It can easily be shown that relations (14) on the base profile are valid in all cases when the interface between the regions taken by vortices with opposite directions of magnetic flux remains stationary. In the case of interest to us, the number of vortices decreases in time, and the interface moves toward the middle of the sample. The inclusion of this motion in accordance with Eq. (2) and boundary condition (11) leads to changes in the expressions for vortex density derivatives,

$$\begin{aligned} N_1' &= -N_1\frac{r}{d} - \frac{4\pi}{\Gamma\Phi_0^2\gamma_{xx}}U, \\ N_2' &= N_1\frac{r}{d} - \frac{4\pi}{\Gamma\Phi_0^2\gamma_{xx}}U, \end{aligned} \tag{16}$$

$$\begin{aligned} N_1'' &= -N_1\left(\frac{r}{d}\right)^2 - \frac{8\pi r}{\Gamma\Phi_0^2\gamma_{xx}d}U, \\ N_2'' &= -N_1\left(\frac{r}{d}\right)^2 + \frac{8\pi r}{\Gamma\Phi_0^2\gamma_{xx}d}U. \end{aligned} \tag{17}$$

The velocity *U* of the front of magnetization reversal is lower than the hydrodynamic vortex velocity V_α. Therefore, in what follows, we will assume for simplicity that the condition $U \ll V_\alpha \sim N_1 R$ holds. This leads one to assume that the base profile differs little from steady-state.

4. STABILITY TEST FOR THE MAGNETIZATION REVERSAL FRONT

It is convenient to test the stability of the magnetization reversal front in the following dimensionless coordinates:

$$\begin{aligned} n_\alpha &= \frac{N_\alpha}{N_\alpha(x_0)}, \quad \tau = \frac{t}{t_0}, \quad t_0 = \frac{\Gamma\Phi_0^2\gamma_{xx}}{4\pi R^2 N_\alpha(x_0)}, \\ \xi &= \frac{x}{L}, \quad \zeta = \frac{y}{L}, \quad L = \frac{\Gamma\Phi_0^2\gamma_{xx}}{4\pi R} = \frac{d}{r}. \end{aligned} \tag{18}$$

The normalization to the time-dependent vortex density $N_\alpha(x_0(t))$ is permissible, because we assume that the instability develops over relatively short periods of time, and there is not enough time for marked variations of the base profiles to occur. We will represent the expression for the vortex density $n_\alpha(\xi, \tau)$ in the form of the sum of the unpurturbed term $n_{(0)\alpha}$ and the fluctuation term

$$n_\alpha = n_{(0)\alpha} + f_\alpha(\xi - \xi_0(\tau))\exp(\lambda\tau + ik\zeta). \quad (19)$$

The linearized boundary conditions must be written on the purturbed boundary,

$$\xi = \xi_0(\zeta, \tau) = \xi_0(\tau) + \delta\xi\exp(ik\zeta + \lambda\tau), \quad (20)$$

with the unit normal vector $\mathbf{\nu} = (1, -ik\delta\xi(\zeta, \tau))$. From boundary condition (12), it directly follows that

$$\delta\xi = \frac{f_1 - f_2}{2}. \quad (21)$$

We substitute Eq. (19) into expression (2) to derive the following ordinary differential equation with variable coefficients for the function $f_\alpha(\xi - \xi_0(\tau))$:

$$\begin{aligned} &f''_\alpha + f'_\alpha(u + 2\alpha ik + 2n'_{(0)\alpha}) \\ &-(\lambda + \beta k^2 - 2\alpha ikn'_{(0)\alpha} - n''_{(0)\alpha}) = 0, \end{aligned} \quad (22)$$

where $u = Ut_0/L$, $\alpha = \gamma_{xy}/\gamma_{xx}$, and $\beta = \gamma_{yy}/\gamma_{xx}$. We will assume that the perturbation of vortex density decays at a distance away from the front $\xi = \xi_0(\tau)$ that is much less than the sample thickness d. This makes possible the replacement of the variable coefficients $n'_{(0)\alpha}(\xi)$ and $n''_{(0)\alpha}(\xi)$ by their values n'_α and n''_α on the front of magnetization reversal. In dimensionless variables, one can readily derive

$$n'_1 = -1 - u, \quad n'_2 = 1 - u \quad (23)$$

and

$$n''_1 = -1 - 2u, \quad n''_2 = -1 + 2u \quad (24)$$

from expressions (16) and (17). The solution of Eqs. (22) defines the exponential behavior of the perturbations f_1 and f_2,

$$\begin{aligned} f_1(\xi - \xi_0) &= f_1\exp[p_1(\xi - \xi_0)], \\ f_2(\xi - \xi_0) &= f_2\exp[-p_2(\xi - \xi_0)]. \end{aligned} \quad (25)$$

The parameters p_1 and p_2, which are solutions to the respective quadratic characteristic equation, are

$$p_{1,2} = \pm 1 - \alpha ik + \frac{u}{2} \pm \Omega_{1,2}, \quad (26)$$

$$\Omega_{1,2} = \sqrt{\lambda + \varepsilon k^2 + \alpha iku + 2 \pm 3u}, \quad \mathrm{Re}\Omega_{1,2} > 0. \quad (27)$$

Here, the upper signs indicate coefficient p_1, and the lower signs, p_2. The substitution of expressions (19)–(27) into the boundary conditions given by Eqs. (9) and (11) produces the following set of two linear homogeneous algebraic equations for the f_1 and f_2 amplitudes:

$$\begin{aligned} &f_1[2\Omega_1 + 2\lambda + 2\alpha iku + u] \\ &+ f_2[-2\Omega_2 - 2\lambda - 2\alpha iku + u] = 0, \\ &f_1[2\Omega_1 + \lambda + \alpha ik(1 + u) + 2 - 2u] \\ &+ f_2[-\lambda - \alpha ik(1 + u) + 2 + 3u] = 0. \end{aligned} \quad (28)$$

We equate the determinant of the set to zero and ignore small terms of the order of u^2 to derive the dispersion equation defining the dependence of the increment of the rise of fluctuations λ on the wave number k,

$$\lambda = \Omega^2 - \epsilon\kappa^2 - 2i\sigma\kappa - 2, \quad (29)$$

where

$$\kappa = \frac{k|\alpha u|}{2}, \quad \epsilon = \frac{4\varepsilon}{(\alpha u)^2}, \quad \sigma = \mathrm{sgn}(\alpha u),$$

and Ω is a root with a positive real part, of the equation

$$\begin{aligned} &\Omega^4 + 3\Omega^3 + \Omega^2(-\epsilon\kappa^2) \\ &-\Omega(2\epsilon\kappa^2 + i\sigma\kappa + 4) - 3\sigma i\kappa = 0. \end{aligned} \quad (30)$$

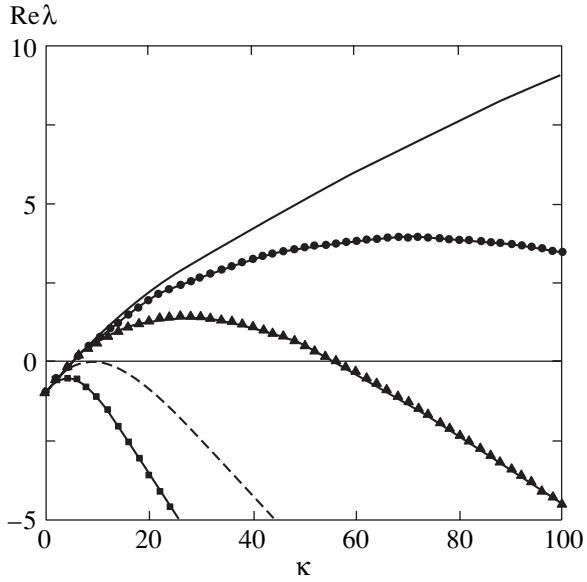


Fig. 2. The $\mathrm{Re}\lambda(\kappa)$ dependence for different values of the parameter ε: $\varepsilon = 0$ (solid line), $\varepsilon = 0.0015$ (circles), $\varepsilon = 0.005$ (triangles), $\varepsilon = \varepsilon_c = 0.019$ (dashed line), and $\varepsilon = 0.05$ (squares).

Figure 2 gives the increment $\mathrm{Re}\lambda$ as a function of dimensionless wave number κ for different values of the anisotropy parameter ϵ. One can see in the figure that, for fairly low values of $\epsilon < \epsilon_c \approx 0.019$, a region of wave numbers exists in which the increment is positive. This means that the plane front of magnetization reversal becomes unstable and the arising perturbation increases with time. The time scale of increase is

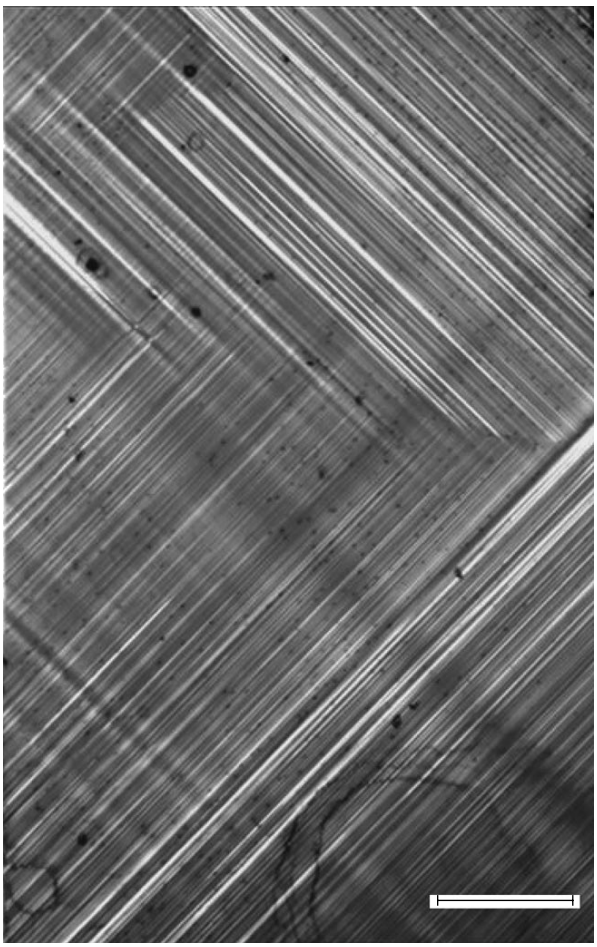


Fig. 3. The image of a surface region of a single crystal sample in polarized light. One can clearly see domains with mutually orthogonal directions of the twinning boundaries. The scale corresponds to 50 μm.

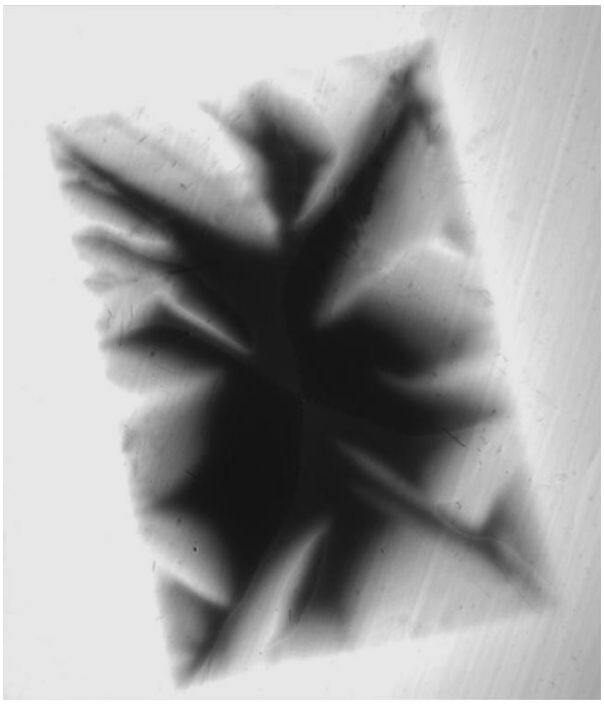


Fig. 4. The magnetooptical image of distribution of the normal component of magnetic induction B_n on the surface of a single crystal. The lighter colored areas correspond to the higher value of B_n. The external magnetic field H = 1000 Oe, T = 45 K.

defined by the maximal value of $\mathrm{Re}\lambda$ equal to λ_m which corresponds to the finite value of $\kappa = \kappa_m$. Accordingly, the spatial scale of instability is defined by the quantity $1/\kappa_m$. In the case of small values of ε, the value of κ_m significantly exceeds unity. One can readily see in Eq. (30) that

$$\lambda_m = 1/(4\sqrt{\epsilon}) - 2, \quad \kappa_m = 1/\sqrt{2}\epsilon^{3/4}. \tag{31}$$

Therefore, the instability arises if the anisotropy is high enough, i.e., $\epsilon < \epsilon_c$. This inequality may be rewritten in dimensionless variables as

$$\varepsilon < \epsilon_c\left[\frac{U\tan\theta}{2RN_1(x_0)}\right]^2. \tag{32}$$

Here, θ is the angle between the interface and the principal direction of the anisotropic viscosity tensor (the direction of the twinning boundaries). Now, one can understand why the macroturbulence is observed only in superconductors of the 1–2–3 system, which are characterized by a strong anisotropy in the *ab* plane due to the presence of the twinning boundaries.

The foregoing analysis is valid in the case of anisotropy with the first-order axis of symmetry. Such a symmetry corresponds to the twinning structure with a single preferred direction. In real crystals of the 1–2–3 system, domains usually exist with the characteristic dimension Δ of the order of 100 μm with the mutually orthogonal direction of the twinning boundaries, i.e., "tweed" structure (see Fig. 3). Our calculation results are valid if the characteristic value of the wave number k is fairly high, i.e., $1/k \ll \Delta$.

5. MAGNETOOPTICAL OBSERVATIONS AND DISCUSSION

The magnetooptical observations of macroturbulence were performed with an optimally doped YBCO single crystal containing a large number of twinning boundaries (see Fig. 3). The crystal was rectangular in the *ab* plane, and its maximal dimension was approximately 1 mm.

The magnetooptical pattern of distribution of the normal component of magnetic flux H = 1 kOe is given in Fig. 4. The single crystal was precooled in zero magnetic field to a temperature T = 45 K. The lighter colored areas correspond to higher values of magnetic induction. One can see in the figure that the magnetic

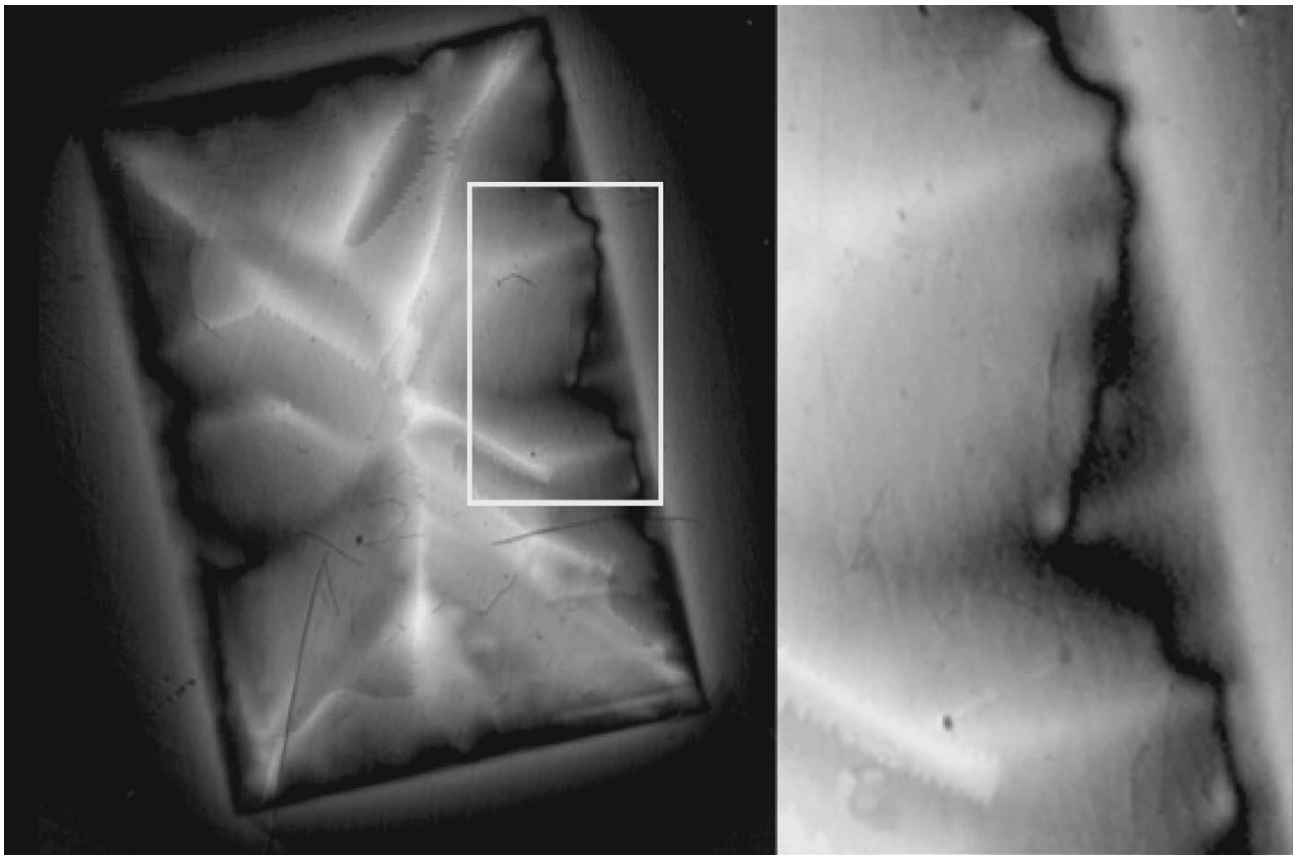


Fig. 5. The same as in Fig. 4 after the temperature rise to 67 K and subsequent switching off of the external magnetic field. Shown on the right is an enlarged area of the fragment outlined by the light colored rectangular contour.

flux penetration is anisotropic in the *ab* plane. The anisotropy axis mainly coincides with the direction of the twinning boundaries, which form an angle of approximately 45° relative to the sample boundaries. The dark areas in the middle part of the image are regions which the field does not penetrate because of the screening by superconducting currents. Note that this pattern of distribution of magnetic induction is quasi-equilibrium and well reproducible.

In the course of experiment, the sample temperature increased to 67 K. The temperature variation brought about a deeper penetration of the magnetic flux up to the middle of the sample. Then, the external magnetic field was switched off. As a result, a frozen magnetic flux remained in the sample. The pattern of its distribution is given in Fig. 5. The lines of the frozen magnetic flux must be closed. Therefore, in the neighborhood of the sample boundary, a domain arises in which the normal component of magnetic induction has a direction opposite to that of the trapped flux. Corresponding to this flux is the light "aura" around the sample in Fig. 5. One can see in the figure that a domain with a negative direction of the magnetic flux (antivortices) also exists in the sample in the vicinity of its edges. Therefore, in this geometry, no presence of an external magnetic field of opposite sign is required for the front of magnetization reversal to arise.

Clearly seen in Fig. 5 is a dark line passing along the sample boundary. This line, which separates two oppositely oriented magnetic domains (with vortices and antivortices), represents the zone of annihilation. The beginning of the development of macroturbulence shows up as the emergence of a meander on the front of magnetization reversal. The emergence of a meander is accompanied by the appearance of regions with a higher density of magnetic flux in the vicinity of the front of magnetization reversal. Such regions are clearly defined in the right-hand part of the figure, which shows an enlarged image of the region marked by a rectangular contour. Measurements with time resolution reveal that the shape of the front of magnetization reversal changes rapidly, which is accompanied by a marked redistribution of the magnetic flux. The switching on of an external magnetic field which has the direction opposite to that of the frozen magnetic flux brings about the advance of the front of magnetization reversal deep into the sample. In this case, the dynamic pattern of macroturbulence becomes even more impressive.

The macroturbulent behavior of magnetic flux in the investigated crystal is observed in a wide temperature range from 25 to 75 K. An increase in temperature makes for a more rapid development of turbulence. However, above 75 K, the macroturbulence disappears.

Consider the agreement between the results of theoretical treatment and the experimental data. Unfortunately, quantitative comparison of theory with experiment is difficult because of the several phenomenological parameters involved in theoretical treatment, for which no adequate information is available. The most important of such parameters is the annihilation parameter R. The estimation of this parameter is of interest *per se*. Nevertheless, we will discuss some features of the phenomenon of turbulence which find a natural qualitative explanation within the model developed by us.

One can readily explain, for example, the fact that the instability is observed only in some temperature range [5, 22]. As the temperature decreases, the viscosity of the vortex motion exponentially increases. Accordingly, the spatial scale of fluctuation L in formula (18) decreases and becomes equal to or less than the characteristic distance between the twinning boundaries. In this case, large regions of the front of magnetization reversal, which "accommodate" a large number of wavelengths, turn out to be located away from the twinning boundaries and cease to be affected by them. As a result, the anisotropy effectively weakens, and the instability disappears. On the other hand, when the temperature increases, the anisotropy weakens because of the thermal-activation motion of vortices. The instability disappears for the same reason. Note that, in the crystal with a high density of twinning boundaries investigated by us, the temperature window of existence of macroturbulence is significantly widened towards lower temperatures compared with other samples with weak twinning [5, 22]. This fact agrees with the suggested model as well.

The objective of this study is to reveal the main reason for the instability of the front of magnetization reversal, which we associate with the anisotropy of the dynamic properties of vortex matter. Note, however, that the selected model of anisotropic viscous flow of magnetic flux is simplified. It corresponds to the linear current–voltage characteristic observed in real superconductors only in the case of a very high current density. There is no doubt that such densities are not realized in experiments involving the investigation of macroturbulence. The crudeness of the model is reflected by the fact that the condition of emergence of instability given by Eq. (32) calls for an unrealistically strong anisotropy of viscosity. It is clear that a more consistent approach to studying the macroturbulence must be based on the use of an anisotropic nonlinear current–voltage characteristic. Preliminary analysis reveals that, in the case of theoretical treatment involving a power current–voltage characteristic with the exponent m, the part of the parameter ε is played by ε^m. In the case of high values of m of the order of 10, the condition of emergence of instability becomes much less rigid.

ACKNOWLEDGMENTS

We are grateful to G.I. Garas'ko for valuable discussions of our study. The study received financial support from INTAS (grant no. 02-2282), the Russian Foundation for Basic Research (project nos. 00-02-17145 and 00-02-18032), and the Research Council of Norway, as well as under the Russian federal target-oriented program on superconductivity (contract no. 40.012.1.1.11.46).

REFERENCES

1. G. Blatter, M. V. Feigel'man, V. B. Geshkenbein, *et al.*, Rev. Mod. Phys. **66**, 1125 (1994).
2. E. H. Brandt, Rep. Prog. Phys. **58**, 1465 (1995).
3. V. K. Vlasko-Vlasov, V. I. Nikitenko, A. A. Polyanskii, *et al.*, Physica C (Amsterdam) **222**, 361 (1994).
4. M. V. Indenbom, Th. Schuster, M. R. Koblischka, *et al.*, Physica C (Amsterdam) **209**, 259 (1993).
5. T. Frello, M. Baziljevich, T. H. Johansen, *et al.*, Phys. Rev. B **59**, R6639 (1999).
6. T. H. Johansen, M. Baziljevich, D. V. Shantsev, *et al.*, submitted to Phys. Rev. Lett. (2002).
7. A. V. Gurevich, Int. J. Mod. Phys. B **9**, 1045 (1995).
8. E. H. Brandt, Phys. Rev. Lett. **76**, 4030 (1996).
9. C. P. Bean, Phys. Rev. Lett. **8**, 250 (1962).
10. Y. Yeshurun, A. P. Malozemoff, and A. Shaulov, Rev. Mod. Phys. **68**, 911 (1996).
11. A. Vl. Gurevich, R. G. Mints, and A. L. Rakhmanov, *Physics of Composite Superconductors* (Nauka, Moscow, 1987).
12. F. Bass, B. Ya. Shapiro, I. Shapiro, and M. Shvartser, Phys. Rev. B **58**, 2878 (1998).
13. V. K. Vlasko-Vlasov, U. Welp, G. W. Crabtree, *et al.*, Phys. Rev. B **56**, 5622 (1997).
14. I. F. Voloshin, A. V. Kalinov, K. I. Kugel', *et al.*, Zh. Éksp. Teor. Fiz. **111**, 2158 (1997) [JETP **84**, 1177 (1997)].
15. A. K. Niessen and C. H. Weijsenfeld, J. Appl. Phys. **40**, 384 (1969).
16. H. Pastoriza, S. Candia, and G. Nieva, Phys. Rev. Lett. **83**, 1026 (1999).
17. L. D. Landau and E. M. Lifshitz, *Course of Theoretical Physics*, Vol. 6: *Fluid Mechanics* (Nauka, Moscow, 1976; Pergamon, New York, 1987).
18. A. V. Gurevich, Phys. Rev. Lett. **65**, 3197 (1990).
19. A. V. Gurevich, Phys. Rev. B **46**, 3638 (1992).
20. L. M. Fisher, P. E. Goa, M. Baziljevich, *et al.*, Phys. Rev. Lett. **87**, 247005 (2001).
21. E. B. Sonin, V. B. Geshkenbein, A. van Otterlo, and G. Blatter, Phys. Rev. B **57**, 575 (1998).
22. M. R. Koblischka, T. H. Johansen, M. Baziljevich, *et al.*, Europhys. Lett. **41**, 419 (1998).

Translated by H. Bronstein

Journal of Experimental and Theoretical Physics, Vol. 95, No. 4, 2002, pp. 777–787.
Translated from Zhurnal Éksperimental'noĭ i Teoreticheskoĭ Fiziki, Vol. 122, No. 4, 2002, pp. 897–908.

NONLINEAR PHYSICS

Linear and Nonlinear Vibrations and Waves in Optical or Acoustic Superlattices (Photonic or Phonon Crystals)

A. M. Kosevich* and M. A. Mamalui

Verkin Physicotechnical Institute of Low Temperatures, National Academy of Sciences of Ukraine, Kharkov, 61103 Ukraine

**e-mail: kosevich@ilt.kharkov.ua*

Received March 5, 2002

Abstract—A model of a dielectric or an elastic superlattice is proposed which describes quite simply the frequency spectrum of electromagnetic or acoustic waves. The frequency band spectrum of a one-dimensional lattice consists of minibands, which narrow down with increasing frequency (so that the forbidden bands in the spectrum broaden with increasing frequency). An elementary analysis of the spectrum of a one-dimensional lattice reveals the presence of many forbidden frequency bands in this case as well. It is shown that dynamic equations for superlattices can be generalized to the nonlinear case, leading to equations of the type of the nonlinear Schrödinger equation for the lattice. Soliton excitations are described and the particle-like dynamics of solitons is demonstrated. Local vibrations near point defects of different complexity in superlattices are studied and graphically illustrated. The existence of Bloch oscillations of a wave packet in a superlattice in a homogeneous external field is discussed. © *2002 MAIK "Nauka/Interperiodica".*

1. INTRODUCTION

The concept of photonic crystals was introduced to the scientific lexicon owing to book [1], in which the theory of dielectric superlattices was presented quite simply and in great detail. It is known that a homogeneous dielectric has a continuous spectrum of electromagnetic waves and vibrations: $\omega = ck$, where ω and k are the frequency and magnitude of the wave vector, respectively, and c is the velocity of light in a dielectric, which depends on the refractive index of the medium. A periodic structure composed of macroscopic layers (or blocks) of different dielectrics acquires new properties, which are manifested first of all in the properties of the spectrum of electromagnetic vibrations in the structure. It is well known from the study of the dynamics of a crystal lattice, which is a natural structure with a microscopic period, that the spectrum of any intrinsic vibrations of the crystal (mechanical, magnetic, and electromagnetic) always consists of bands due to a spatial periodicity of the crystal lattice. The widths of allowed frequency bands and forbidden bands (gaps) are determined by the reciprocal period of the lattice. In a crystal where the sound speed is s, these bandwidths in the spectrum of intrinsic mechanical vibrations (phonons) can be approximately estimated as $\delta\omega \sim s/a$, where a is the period of a crystal lattice (interatomic distance). Note that the width of the transparency region of a dielectric for electromagnetic waves can be estimated similarly as $\delta\omega \sim c/a$.

These estimates are related only to the lattice periodicity rather than to the period value. Therefore, they remain valid for a periodic structure with the macroscopic period d, which, by definition, greatly exceeds the interatomic distance ($d \gg a$). The latter means, in particular, that, in the transparency region of homogeneous dielectrics comprising a superlattice, a band spectrum of intrinsic electromagnetic oscillations (photons) appears, which resembles the spectrum of a crystal lattice, but with the bandwidths estimated as $\delta\omega \sim c/d \ll \delta\omega \sim c/a$. For this reason, the term "photonic crystals" was coined.

Similarly, a macroscopic periodic structure consisting of alternating elastic materials that differ in their elastic moduli and sound speeds is called a phonon crystal or an acoustic superlattice. The general theory of acoustics of layered media is presented in book [2], and useful references in this field can be found in one of the recent papers [3] devoted to the study of acoustic superlattices. The frequency spectrum of superlattices was studied in many papers. It is obvious that this spectrum is very complicated in the general case and includes a system of many eigenfrequency bands and gaps corresponding to forbidden frequencies. To describe these spectra qualitatively and to illustrate their main properties, it is useful to employ simple models that take into account these properties. In this paper, we propose a model of a superlattice that provides the analytic description of the high-frequency part of the superlattice spectrum and suggests a possible realization of an interesting acoustic superlattice. This model and its simplest possible applications were considered in paper [4].

1063-7761/02/9504-0777$22.00

We used the results of papers [5, 6], where the dispersion equations for one-dimensional dielectric and elastic superlattices were presented, which coincide with an accuracy of notation with the dispersion law for a quantum particle in the one-dimensional Kronig–Penni model [7]. We considered a limiting case of the dispersion equation, which was usually of no interest in the case of a quantum particle in a periodic field but proved to be a convenient and instructive model of the elastic superlattice. The frequency spectrum that we obtained is of interest because its high-frequency part has a system of narrowing allowed frequency bands in which the dispersion law can be calculated analytically with good accuracy. This is the dispersion law for size-quantized sound vibrations in individual layers (blocks), with discrete frequencies that are periodically blurred due to a weak interaction between adjacent blocks. The dispersion equations for the high-frequency spectrum can be readily generalized to two-dimensional and three-dimensional superlattices with a one-component wave field.

We describe a simplified dispersion law for each miniband by an effective equation of the envelope type, which coincides in its form with the one-dimensional Schrödinger equation in the lattice-point representation for the strong coupling model. We performed a trivial linear generalization leading to a discrete variant of the nonlinear Schrödinger equation (NSE) admitting soliton solutions. In one of the cases of nonlinear generalization, we obtained a completely integrable NSE for the lattice, whose solitons obey the Hamiltonian dynamics.

We studied vibrations localized in the vicinity of simplest defects in a one-dimensional superlattice, when the coupling parameter at one boundary (or at two to three boundaries) of adjacent blocks differs from that at all other boundaries (joints) of the blocks. In the presence of a local defect, both symmetric and antisymmetric independent vibrations relative to the defect center can occur because of the linearity of vibrational equations. In addition, an inphase or antiphase vibration of adjacent blocks can correspond to each type of the vibrations. Therefore, a great variety of local vibrations exist to which different equations for local frequencies correspond.

In the final section of the paper, we describe Bloch oscillations of a wave packet in a superlattice whose parameters weakly depend on the coordinate along the one-dimensional lattice (superlattice in homogeneous external field). It is noted that such oscillations can be observed experimentally.

2. DISPERSION EQUATION FOR A SIMPLEST SUPPERLATICE MODEL

Consider a one-dimensional lattice consisting of periodically arranged (along the x axis) layers of isotropic materials of two types [d_α are the thickness of layers ($\alpha = 1, 2$), c_α is the velocity of wave in the α layer, and the lattice period is $d = d_1 + d_2$]. The field of the elastic (electromagnetic) wave $\mathbf{u}(\mathbf{r}, t)$ propagating perpendicular to the layer plane is determined by a standard wave equation. In a system of isotropic blocks, the waves of two possible polarizations are independent, and we can restrict ourselves to analysis of dynamic equations for the scalar field $u^{(\alpha)}$

$$\frac{\partial^2 u^\alpha}{\partial t^2} - c_\alpha^2 \frac{\partial^2 u^\alpha}{\partial x^2} = 0, \quad \alpha = 1, 2. \tag{1}$$

The velocity of light in a dielectric is $c_\alpha = c/\sqrt{\varepsilon_\alpha}$ (c is velocity of light in vacuum, ε_α are dielectric constants of superlattice layers).[1] The velocity of a wave in an elastic medium is $c_\alpha = \sqrt{\mu_\alpha/\rho_\alpha}$ (μ_α and ρ_α are the elastic moduli and mass densities, respectively.

Let us focus our attention on an elastic lattice. In this case, Eqs. (1) should be solved using the boundary conditions according to which the displacements $u^{(\alpha)}$ and stresses $\sigma^\alpha = \mu_\alpha \partial u^{(\alpha)}/\partial x$ should be continuous at all boundaries of the blocks.

Each of the vibrational eigenmodes appearing in a periodic structure with a period d is characterized by the quasi-wave number k, and natural oscillations of the field in an elementary cell with the number n can be written in the form

$$u_n(x) = u_0(x - nd)e^{iknd}. \tag{2}$$

Dispersion relations for the structure under study were obtained by Rytov for both optical [5] and elastic [6] superlattices. They have the form

$$\cos kd = \cos q_1 d_1 \cos q_2 d_2 - \frac{1}{2}\left(\frac{q_1}{q_2} + \frac{q_2}{q_1}\right)\sin q_1 d_1 \sin q_2 d_2, \tag{3}$$

where $q_1 = \omega/c_1$, $q_2 = \omega/c_2$ (ω is frequency). As expected, these relations coincide with an accuracy to notation with those obtained by Kronig and Penni [7] for a quantum particle in a periodic potential.

Expression (3) gives the implicit dependence of the frequency on the quasi-wave number and allows us to describe readily the spectrum of long-wavelength vibrations ($kd \ll 1$), for which the sound spectrum is naturally obtained with averaged elastic moduli $\langle\mu\rangle$ and the density $\langle\rho\rangle$. It was shown in paper [6] that $\langle\rho\rangle d = \rho_1 d_1 + \rho_2 d_2$ and $d/\langle\mu\rangle = d_1/\mu_1 + d_2/\mu_2$. Based on such a representation of μ, which contains only ratios d_α/μ_α, it is interesting to consider a limiting case, which is of no interest from the physical point of view in the case of

[1] As shown in [4], the spectrum of this system consists of narrow bands, and the frequency dependence of ε_α can be neglected within each of these bands. Therefore, the dielectric constant corresponding to the relevant frequency can be assigned to each of the allowed bands.

the dynamics of a quantum particle in a periodic potential, but can demonstrate the most characteristic properties of the superlattice spectrum, when $d_2 \longrightarrow 0$ and $\mu_2 \longrightarrow 0$ for $d_2/\mu_2 = M = \text{const}$. In this case, $d_1 \longrightarrow d$, $k_2 d_2 = \omega d_2/c_2 = \sqrt{\rho_2 d_2}\,\omega\sqrt{d_2/\mu_2} \longrightarrow 0$. The, the dispersion law for the system is described by the equation

$$\cos kd = \cos z - Qz\sin z, \tag{4}$$

where $z = q_1 d = \omega d/c_1$ and $Q = \rho_2\mu_1 M/2\rho_1 d$. Note that Eq. (4) gives the dispersion law for an elastic superlattice consisting of periodic elastic blocks of length d with the parameters μ_1 and c_1. The normal stress at the boundaries of these blocks are continuous ($[\sigma]_-^+ = 0$, i.e., $[\partial u/\partial x]_-^+ = 0$), while elastic displacements exhibit a jump proportional to the value of stresses at the block boundaries:

$$[u]_-^+ = Q\left(\frac{\rho_1}{\rho_2}\right)\left(\frac{\partial u}{\partial x}\right). \tag{5}$$

A set of such boundary conditions for Q being fixed is used for the description of capillary effects in solids [8] or plane defects in crystals [9]. If the parameter Q is small, then the system under study represents a periodic sequence of elastic regions that are weakly connected with each other. A chain of piezoelectric layers connected by thin vacuum spaces can form such a superlattice. In this case, electromagnetic oscillations in vacuum spaces would provide the coupling between elastic vibrations in adjacent regions. A piezoelectric with a structure admitting both the construction of such a periodic system of layers and the independent propagation of one elastic electromagnetic wave of a certain polarization along the superlattice can be easily found.

Let us return to relation (4). The allowed vibrational frequencies of a continuous spectrum of the system under study can be qualitatively found by analyzing graphically Eq. (4), as shown in Fig. 1: if the expression $\cos z - Qz\sin z$ runs the values between ± 1, the roots of the equation run the values in the intervals shown on the abscissa.

Note that, as z increases, the allowed frequencies are localized within the narrowing intervals near the values $k_1 d = \pm m\pi$, where m is a large integer.[2] For the condition $m^2 Q \gg 1$, the dispersion law for the mth band can be readily found.

[2] The narrowing of bands with increasing frequency was mentioned earlier, in particular, in [3]. However, the statement [10] that the bandwidths do not change on average with increasing frequency contradicts the above affirmation.

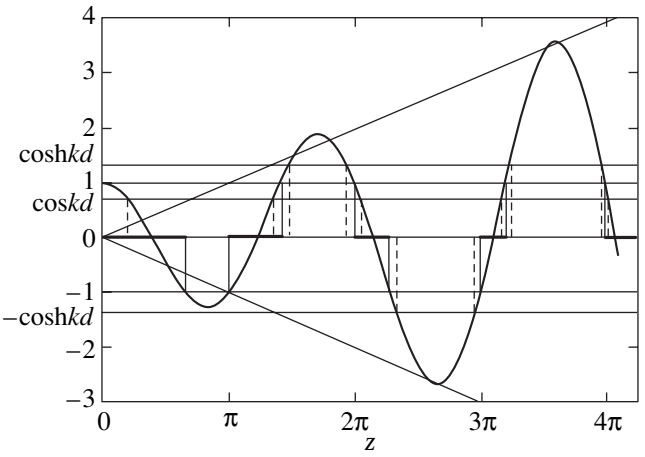


Fig. 1. Graphical solution of Eq. (4). If $\cos z - Qz\sin z$ runs the values between ± 1, the roots of equation determining bands of a continuous spectrum of superlattices run the values in the intervals indicated on the abscissa.

Indeed, near odd $m = 2p + 1$ (see the vicinity of $z = 3\pi$ in Fig. 1), we can write with sufficient accuracy ($c = c_1$)

$$\cos kd = -1 + Qm\pi(z - m\pi)$$
$$= -1 + \frac{m\pi d}{c}Q\left(\omega - \frac{m\pi c}{d}\right),$$

which yields

$$\omega = m\omega_0 + \frac{\Omega}{m}(1 + \cos kd), \tag{6}$$

where $\omega_0 = \pi c/d$ and $\Omega = c/(\pi Q d)$. Similarly, even near $m = 2p$ (see the vicinity of $z = 4\pi$ in Fig. 1), we can write

$$\cos kd = 1 - Qm\pi(z - m\pi)$$
$$= 1 - \frac{m\pi d}{c}Q\left(\omega - \frac{m\pi c}{d}\right),$$

which gives

$$\omega = m\omega_0 + \frac{\Omega}{m}(1 - \cos kd). \tag{7}$$

By combining (6) and (7), we obtain the dispersion laws for the mth band:

$$\omega = m\omega_0 + \frac{2\Omega}{m}\begin{cases}\sin^2(kd/2), & m = 2p,\\ \cos^2(kd/2), & m = 2p+1.\end{cases} \tag{8}$$

One can easily see that expressions (8) represent the size-quantization spectrum of phonons in a layer of thickness d, whose levels split into minibands due to a low "transparency" of the interface between layers. In paper [10], where the dispersion relation (3) was derived once more, an attempt was made to analyze the band spectrum of the superlattice. However, in the lim-

iting case, which is close to that considered in our paper, the analysis [10] is not satisfactory because it predicts that the widths of minibands do not change with increasing frequency.

Let us make some comments on the proposed realization of an elastic superlattice in the form of a periodic stack of piezoelectric layers separated by thin vacuum spaces. Upon normal incidence of a transverse electromagnetic wave from vacuum on a piezoelectric layer, the static piezoelectric effect does not "operate," and it is necessary to consider the relativistic electromechanical effect, which is proportional to a small parameter $(s/c)^2$. The point is that a vibrational mode of the piezoelectric of interest to us consists mainly of the elastic component with a small contribution [determined by the parameter$(s/c)^2$] of the electromagnetic component, while the interaction between the adjacent layers occurs through electromagnetic oscillations. For this reason, the coupling between elastic vibrations in adjacent piezoelectric layers is extremely weak. As a result, the widths of minibands, for which dispersion laws of the type (8) can be obtained, prove to be negligibly small. Therefore, such a superlattice should have a virtually discrete frequency spectrum of elastic waves, which is caused by the resonance propagation of electromagnetic waves through the piezoelectric and corresponds to the spectrum of size-quantized stationary elastic vibrations in an individual layer.

A real spread of discrete frequencies in such a superlattice will be caused not by the appearance of regular minibands but by processes of a different physical nature, namely, the uncertainty of the wave number in a superlattice of a finite length, the accuracy of its strict periodicity, and dissipative processes resulting in the attenuation of sound in a piezoelectric.

The model of a one-dimensional superlattice proposed above and the derivation the dispersion law (8) can be readily generalized to the two-dimensional and three-dimensional cases if the wave field under study has one component. The latter condition is simply satisfied for a two-dimensional superlattice consisting of blocks with a square cross section separated by two mutually perpendicular systems of boundary planes. The cross section of such a lattice represents a parquet of square regions of elastic media separated by planes (or by straight lines in the plane $z = 0$) $x = n_1 d$, $y = n_2 d$, where $n_1, n_2 = 0, \pm 1, \pm 2, \ldots$. If the elastic wave is polarized along the z axis, then the boundary conditions at two systems of boundary planes are formulated independently in the form of relations of the type (5), where the left-hand sides are $\omega = q_1 c$ or $\omega = q_2 c$. For example,

$$q_1 c = m_1 \omega_0 + \frac{2\Omega}{m_1} \sin^2 \frac{k_x d}{2}, \quad m_1 = 2p_1, \tag{9}$$

or

$$q_2 c = m_2 \omega_0 + \frac{2\Omega}{m_2} \cos^2 \frac{k_x d}{2}, \quad m_2 = 2p_2. \tag{10}$$

The dispersion law for the elastic wave in each block is obvious:

$$\omega^2 = (q_1^2 + q_2^2) c^2. \tag{11}$$

Therefore, in the limiting case $m_1 Q, m_2 Q \gg 1$, the dispersion laws for highly lying allowed bands have the form

(1) $m_1 = 2p_1$, $m_2 = 2p_2$,

$$\omega_{m1m2}(\mathbf{k}) = \omega_0 \sqrt{m_1^2 + m_2^2} + \frac{2\Omega}{\sqrt{m_1^2 + m_2^2}} \left[\sin^2 \frac{k_x d}{2} + \sin^2 \frac{k_y d}{2} \right]; \tag{12}$$

(2) $m_1 = 2p_1 + 1$, $m_2 = 2p_2 + 1$,

$$\omega_{m1m2}(\mathbf{k}) = \omega_0 \sqrt{m_1^2 + m_2^2} + \frac{2\Omega}{\sqrt{m_1^2 + m_2^2}} \left[\cos^2 \frac{k_x d}{2} + \cos^2 \frac{k_y d}{2} \right]; \tag{13}$$

(3) $m_1 = 2p_1$, $m_2 = 2p_2 + 1$,

$$\omega_{m1m2}(\mathbf{k}) = \omega_0 \sqrt{m_1^2 + m_2^2} + \frac{2\Omega}{\sqrt{m_1^2 + m_2^2}} \left[\sin^2 \frac{k_x d}{2} + \cos^2 \frac{k_y d}{2} \right]; \tag{14}$$

(4) $m_1 = 2p_1 + 1$, $m_2 = 2p_2$, similar to expression (14) in which the replacement $m_1 \longrightarrow m_2$ and $k_x \longrightarrow k_y$ is made.

The frequency spectrum described by expressions (12)–(14) is equivalent to the spectrum of a quantum particle in a periodic two-dimensional structure in the strong coupling model. The generalization of this result to a rectangular two-dimensional superlattice is obvious, namely, two periods d_x and d_y and two parameters Q_x and Q_y (or Ω_x and Ω_y) will appear.

The generalization of the above result to the case of a three-dimensional scalar wave field is also obvious (it coincides with the three-dimensional strong coupling model).

The main feature of the frequency spectrum of the one-dimensional superlattice is the presence of broad intervals of forbidden vibrational frequencies (gaps in a continuous spectrum). The existence of forbidden-frequency intervals ω in two-dimensional and three-dimensional superlattices is proved not so readily. Nevertheless, one can prove that such intervals are present at least in the two-dimensional lattice.

By choosing a two-dimensional rectangular coordinate net, we plot the quantity $q_1 d = (\omega/c) d$ along one of

the axes, by indicating the boundaries of the frequency intervals allowed by expressions (12)–(14). Let us perform a similar procedure with $q_2 d = (\omega/c)d$ for the second axis and then draw a system of straight lines parallel to the coordinate axes (see schematic Fig. 2). These straight lines separate rectangular or square regions in the plane in which points (q_1, q_2) corresponding to the allowed frequencies, according to (11), can be located. The broad arcs in Fig. 2 show the forbidden-frequency bands corresponding to gaps in a continuous spectrum.

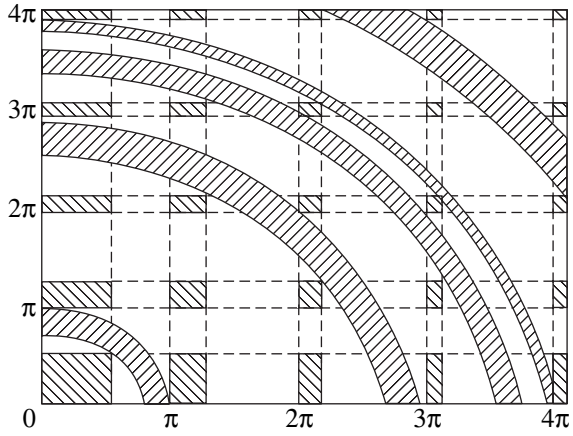


Fig. 2. Schematic view of the mutual arrangement of forbidden and allowed bands in a two-dimensional superlattice in the case of equal periods in two mutually perpendicular directions (on both axes the values $q_1 d = q_2 d = (\omega/c)d$ are plotted). The dashed rectangular and square regions, in which points (q_1, q_2) can be located, correspond to the allowed frequencies [see expression (11)]. The dashed arcs correspond to forbidden bands of the spectrum of a two-dimensional square superlattice.

3. SOLITARY WAVES IN A SUPERLATTICE

Let us return to Eq. (8) and note once more that this dispersion equation coincides with the dispersion law for a Bloch quasi-particle ($(\psi_m(z + d) = \psi_m(z)e^{ikd}$), which corresponds, in the strong coupling model, to the Schrödinger equation

$$i\frac{\partial \psi_n}{\partial t} = m\omega_0\psi_n \pm \frac{\Omega}{2m}(2\psi_n \pm \psi_{n+1} \pm \psi_{n-1}), \quad (15)$$

where the choice of the sign depends on the band number: the upper sign corresponds to $m = 2p$, while the lower sign corresponds to $m = 2p + 1$. From the point of view of studying the superlattice spectrum, Eq. (15) represents the equation for the envelope of the superlattice vibrations taken at discrete points. This equation describes the dynamics of a packet of waves corresponding to the allowed high frequencies. As usual, the order of the time derivative is lowered in equations for the envelope.

The discrete Schrödinger equation (15) can be treated as a dynamic equation for vibrations of a superlattice in the lattice-point representation. This allows us to generalize easily our studies to nonlinear excitations in the superlattice. By following the scheme of taking into account basic nonlinear effects in optics, we assume that the wave velocity c depends on the intensity of vibrations:

$$c = c_0 + \gamma|\psi|^2, \quad \gamma = \text{const}. \quad (16)$$

The dependence of c on $|\psi|^2$ should be taken into account only in the first term on the right-hand side of Eq. (8). Then, vibrations in the mth frequency band (for certainty, $m = 2p$) will obey the equation

$$i\frac{\partial \psi_n}{\partial t} = m\omega_0\psi_n - \frac{\Omega}{2m}(2\psi_n - \psi_{n+1} - \psi_{n-1}) + mg|\psi_n|^2\psi_n, \quad (17)$$

where $g = \pi\gamma/d$ is a new parameter.

An equation analogous to Eq. (17) is used in the nonlinear optics of the waveguide system [11]. Among the effects that are studied in nonlinear optics and are described by Eq. (17), the most interesting is the existence and dynamics of optical solitons. Therefore, Eq. (17) can be used for the description of soliton excitations in a superlattice. Unfortunately, Eq. (17) does not belong to the class of completely integrable equations and does not have exact soliton solutions. And only passage to the continuous limit, when Eq. (17) in finite differences is reduced to an ordinary NSE in partial derivatives, permits the construction of exact soliton solutions. It is an equation of this type that was used in paper [12] for studying nonlinear localized waves in periodic media. However, the approximation of narrow bands with the dispersion laws (8) used here does not allow the passage to such a limit.

Note, however, that other manifestations of nonlinearity can be observed in the superlattice under study. An important factor in the construction of the superlattice model proposed above is relation (5), containing the phenomenological parameter Q, which determines in turn the parameter Ω in (15). This parameter can depend on the intensity of vibrations. Let us assume that the dependence of the type (16) is valid for the parameter Q. We will use this assumption by preliminarily simplifying Eq. (15). Consider the case $m^2Q \gg 1$ by retaining on the right-hand side of this equation only the first terms of each type nonvanishing over the parameter $1/(m^2Q)$:

$$i\frac{\partial \psi_n}{\partial t} = m\omega_0\psi_n + \frac{\Omega}{2m}(\psi_{n+1} + \psi_{n-1}). \quad (18)$$

Then, we return to the assumption of the type (16):

$$\Omega = \Omega_0(1 + g_0|\psi_n|^2), \quad g_0 > 0. \quad (19)$$

As a result, not imposing any strict normalization requirements on the function ψ_n, we rewrite (18) in the form

$$i\frac{\partial\psi_n}{\partial t} = m\omega_0\psi_n + \frac{\Omega_0}{2m}(1+|\psi_n|^2)(\psi_{n+1}+\psi_{n-1}). \quad (20)$$

It is known that Eq. (20) for $\Omega_0 < 0$ is an exactly integrable version of the NSE for the lattice [13, 14].

Let us assume that $Q < 0$ and introduce a new dimensionless time $\tau = |\Omega_0|t/(2m)$. Then, Eq. (20) can be written in the standard form

$$i\frac{\partial\psi_n}{\partial\tau} = m\omega_*\psi_n - (\psi_{n+1}+\psi_{n-1})(1+|\psi_n|^2), \quad (21)$$

where $\omega_* = 2m^2\omega_0/\Omega_0 = 2\pi^2m^2|Q|$.

Consider the stationary solution of Eq. (21) of the type

$$\psi_n(\tau) = \phi_n^{\omega}(\tau)\exp\{ikn - i\omega t - i\Theta\}, \quad (22)$$

where $\phi_n^{\omega}(\tau)$ is a real function and Θ is a constant arbitrary phase. The function ϕ_n vanishing at infinity ($\phi_n = 0$ for $n \longrightarrow \pm\infty$) corresponds to soliton solutions.

If Eq. (21) has a stationary localized solution propagating along the chain at the velocity V, one can expect that, as in the case of the continuous NSE, its real amplitude will have the form

$$\phi_n(t) = \phi(n - Vt). \quad (23)$$

Taking this into account, the authors of paper [15] suggested the solution in the form

$$\phi(x) = A\,\mathrm{cn}[\beta(x - x_0, \kappa)], \quad (24)$$

where A = const and $\mathrm{cn}(z, \kappa)$ is the elliptic Jacobi cosine, whose parameter κ obeys the inequalities $0 < \kappa < 1$. For $\kappa \ll 1$, the elliptic cosine transforms to the trigonometric cosine, and Eq. (24) transforms to a solution of a linear equation. In the opposite limiting case $\kappa \longrightarrow 1$, the transformation $\mathrm{cn}(z, \kappa) \longrightarrow 1/\cosh z = \mathrm{sech} z$ takes place, and we obtain the solution

$$\psi_n(t) = \sinh\beta\,\mathrm{sech}[\beta(n - Vt - x_0)] \times \exp(ikn - i\omega t + i\theta), \quad (25)$$

where x_0 and θ are constants and the parameters β, V, ω, and k are related by two expressions

$$\omega = \omega_* - 2\cosh\beta\cos k, \quad (26)$$

$$V = \frac{2}{\beta}\sinh\beta\sin k. \quad (27)$$

Thus, for a certain sign of the parameter Q ($Q < 0$), solitons may propagate at a constant velocity in the superlattice. It was shown in paper [16] that the movement of a soliton described by solution (25) obeys the Hamiltonian dynamics. In particular, its movement is controlled by three integrals of motion: its mass (playing the role of the norm of the wave function)

$$N = \sum_n \ln(1+|\psi_n|^2); \quad (28)$$

its energy $\mathcal{H} = E + \Omega N$, where

$$E = -\sum_n(\psi_n\psi_{n+1}^* + \psi_n^*\psi_{n+1}); \quad (29)$$

and the total momentum of the soliton (which differs from the total quasi-momentum, which is not conserved[3])

$$S = -i\sum_n(\psi_n^*\psi_{n+1} - \psi_n\psi_{n+1}^*) = -i\sum_n\psi_n^*(\psi_{n+1} - \psi_{n-1}). \quad (30)$$

If the soliton solution has the form (25), then the integrals of motions are invariant with respect to continuous translations and, therefore, can be calculated by integration, by replacing sums in (28)–(30) by integrals ($\sum_n \ldots = \int dn \ldots$). Then, we obtain

$$N = 2\beta, \quad E = -4\sinh\beta\cos k, \quad S = 4\sinh\beta\sin k. \quad (31)$$

Therefore, three of the four parameters β, V, ω, and k are determined by fixed integrals of motion, while the fourth parameter (the quasi-wave number k) remains free. We see that the width of a soliton $\lambda = 1/\beta$ is determined only by the value of N, while its energy and velocity are periodic functions of k (as should be in a homogeneous periodic structure).

Of course, one should bear in mind that we discussed the properties of the exact soliton solution of the approximate equation obtained by simplifying the initial NSE (16) for the lattice, which is valid itself if the inequality $m^2|Q| \gg 1$ is satisfied. The perturbations caused by this approximation will distort and possibly destroy the soliton solution with time. However, the presence of a large parameter $m^2|Q| \gg 1$ for high-frequency bands ensures the smallness of these perturba-

[3] The operator of quasi-momentum is expressed in terms of the operator of translation by one period $\hat{T}$ of a one-dimensional chain as $\hat{q} = -i\ln\hat{T}$; therefore, the total quasi-momentum Q can be represented as the sum

$$Q = \sum_n\psi_n^*\hat{q}\psi_n = i\sum_n\sum_{m=1}^{\infty}\frac{(-1)^m}{m}(\psi_n^*\psi_{n+m} - \psi_n\psi_{n+m}^*).$$

This quantity is not an integral of motion in the anharmonic chain described by Eq. (21).

tions, and such solitons can exist, at least, as long-lived localized excitations of the superlattice.

4. LOCAL DEFECTS IN SUPERLATTICES

We studied the spectrum of natural vibrations of superlattices, which is characterized by a series of narrowing allowed frequency bands and has (in the general case) forbidden bands (gaps in a continuum spectrum). Consider now the possibility of the appearance of vibrational states in forbidden bands. Such vibrations correspond to solutions of the type $u_n \propto \exp(\mp\kappa n d)$ ($k = i\kappa$) or $u_n \propto (-1)^n \exp(\mp\kappa n d)$ ($k = i\kappa + \pi$) exponentially decreasing (increasing) with the number n. It is obvious that such states can have a physical meaning only on the coordinate semiaxis under the condition that a solution that vanishes at infinity is chosen, which reflects some boundary conditions at the coordinate origin.

For the solution of the first type ($k = i\kappa$), the frequency dependence of the parameter κ can be found from the relation

$$\cosh \kappa d = \cos z - Qz \sin z > 1, \tag{32}$$

while, for the solution of second type ($k = i\kappa + \pi$), it can be found from the relation

$$-\cosh \kappa d = \cos z - Qz \sin z < -1. \tag{32a}$$

The solutions of the first type correspond to frequencies in the intervals $(2p-1)\pi < z < 2p\pi$, and those of the second type, in the intervals $2p\pi < z < (2p+1)\pi$ (see Fig. 3).

It is obvious that such situations appear on the semiaxis due to a local distortion of a strict periodicity in the system, i.e., due to the presence of defects in the periodic ordering. The simplest defect is the difference of physical conditions at one of the boundaries of elastic blocks from those at other boundaries.

Let us assume that the boundary condition (5) at one of the joints is characterized by the parameter Q*, which differs from Q: $Q^* = Q + \xi Q$. We assume also that, at the boundary of joints with the number $n = 0$, condition (5) has the form

$$[u_{n=0}]_-^+ = \left(\frac{\partial u}{\partial x}\right) Q^* \left(\frac{\rho_1}{\rho_2}\right). \tag{33}$$

We are interested in the solution that is antisymmetric with respect to the choice of the sign of n and will represent the displacements u_n for positive n in the form

$$u_n(x, t) = w(x - nd)\exp\{-\kappa n d - i\omega t\}, \quad n \geq 0, \tag{34}$$

Then, the elastic stresses are

$$\sigma_n(x, t) = \mu_1 w_x'(x - nd)\exp\{-\kappa n d - i\omega t\}.$$

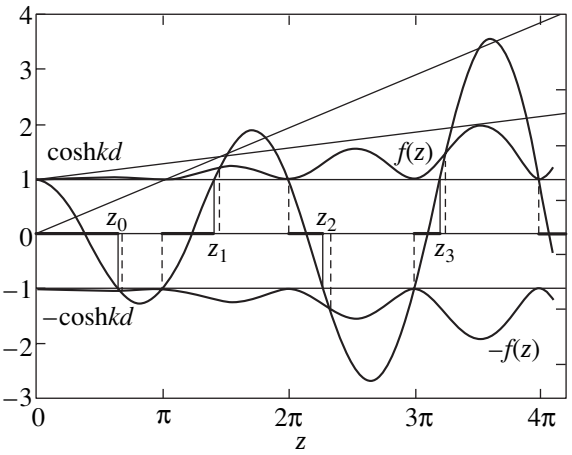


Fig. 3. Graphical determination of the roots of Eq. (38): roots z_0, z_2, … and z_1, z_3, … correspond to two types of vibrations (solution of the type $u_n \propto \exp(\mp\kappa n d)$ ($k = i\kappa$) and $u_n \propto (-1)^n \exp(\mp\kappa n d)$ ($k = i\kappa + \pi$), respectively). The frequencies of local vibrations corresponding to different signs of ξ are located in the alternating intervals between $z = 2p\pi$ and $z = (2p+1)\pi$ ($p = 0, 1, 2, \ldots$).

The function $w(x)$ can be conveniently written in the form

$$w(x) = a_0 \cos(qx - \theta), \tag{35}$$

where $q = \omega/c$ and $\theta = \text{const}$ is a constant phase.

The boundary conditions at $n = 0$ yield the relation

$$e^{-\kappa d} = \cos z - Q^* z \sin z, \quad z = qd. \tag{36}$$

By combining (36) with (32) and (32a), we obtain the equation

$$\sinh \kappa d = \xi Qz \sin z, \quad \xi = \frac{Q^* - Q}{Q}. \tag{37}$$

Equation (37), together with (32) and (32a), determines the frequencies of vibrations localized at such a defect. The local frequencies are determined by points of intersection of the plots of the right-hand sides of (32) and (32a) with the plot of the function $f(z) = \sqrt{1 + \sinh^2 z} = [1 + (\xi Qz)^2 \sin^2 z]^{1/2}$ determined by condition (37). Because $\kappa > 0$, the solutions are the frequencies (the values of z) determined from the equation[4]

$$\cos z - Qz \sin z = \operatorname{sgn}\{\xi Qz \sin z\}\sqrt{1 + (\xi Qz)^2 \sin^2 z}. \tag{38}$$

Figure 3 presents the method of graphical solution of Eq. (38). The curves corresponding to the left-hand and right-hand sides of formula (38) are shown, and

[4] Equation (37) for local frequencies is a limiting case of a more general relation obtained in [10].

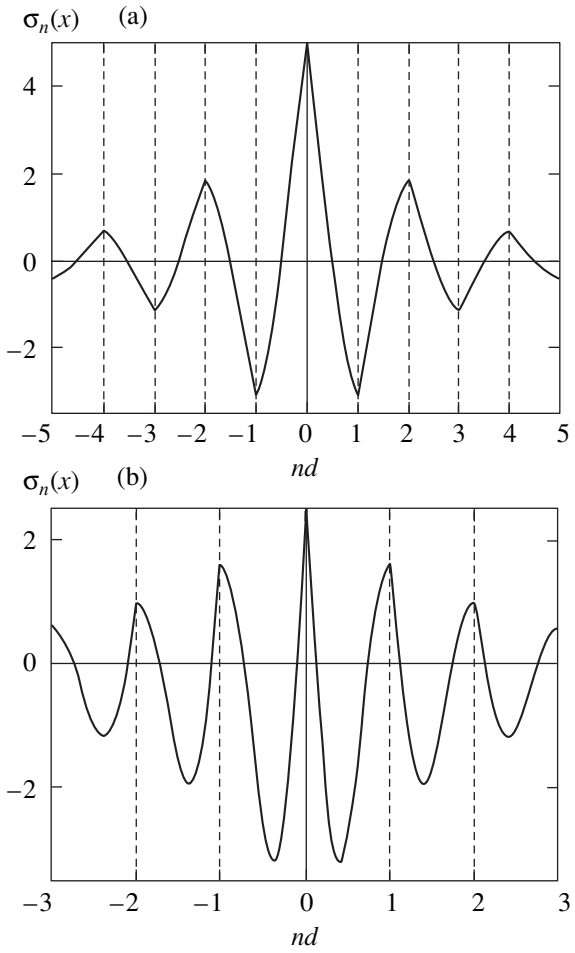


Fig. 4. Distribution of stresses near a superlattice defect localized at one joint (n = 0). Figures 4a and 4b show two types of localized states corresponding to the roots z_0, z_2, … and z_1, z_3, … in Fig. 3.

their points of intersection, which give the required frequencies, are indicated.

Note that the points of intersection exist for any sign of the parameter ξ. The frequencies of local vibrations corresponding to the opposite signs of ξ are located within the alternating intervals between $z = 2p\pi$ and $z = (2p + 1)\pi$ ($p = 0, 1, 2, \ldots$). The local vibrations for which $(2p - 1)\pi < z < 2p\pi$ (points z_1 and z_3 in Fig. 3) are described by the function that monotonically decreases with increasing number of the elementary cell, whereas the vibrations with frequencies $2p\pi < z < (2p + 1)\pi$ (points z_0 and z_2 in Fig. 3) are described by the function that is proportional to $(-1)^n\exp(-\kappa nd)$.

The displacement plots in both cases are presented in paper [4]. Figure 4 shows the calculated distributions of the corresponding stresses near a defect.

Thus, we studied a local vibration near a defect in the form of an antisymmetric solution to the main equation. A local vibration with an even eigenfunction for any sign of ξ cannot appear at a defect localized at one boundary. Let us assume that this is a joint with $n = 0$; in this case, $\sigma_0 = 0$ in the joint when the solution of $[u]^+_- = 0$ is even with respect to the choice of the sign. Therefore, excitation in the form of a standing wave with an even dependence on the coordinate x is insensitive to the value of the parameter Q at the joint $n = 0$ and does not differ from the vibration of a free boundary of the superlattice passing through this joint. The free boundary of the superlattice corresponds to the cut over the joint n = 0, which is equivalent to the condition $\sigma_0 = 0$, which is obtained in our model for $\xi = \infty$ ($Q^* = \infty$). In this case, only uniform vibrations (k = 0) are possible with frequencies $\omega = (c/d)\pi m$, $m = 0, 1, 2, \ldots$. Therefore, a localized wave cannot exist near a free end of the superlattice. This means that even vibrations cannot occur if a defect is localized at one joint. Such localized excitations appear when the parameter Q is changed (perturbed) at least at two adjacent joints.

Consider now just this situation. Let us assume that the same conditions (33) take place at both boundaries of a cell between $n = 0$ and $n = 1$. Then, bearing in mind the solution that is symmetric with respect to a "defect" cell, we write the displacements $u(x, t)$ for positive n in the form

$$\begin{gathered} u(x, t) = w_0(x)e^{i\omega t}, \\ 0 < x < d, \\ u(x, t) = w(x - nd)e^{-i\omega t}e^{\kappa nd}, \\ nd < x < (n + 1)d. \end{gathered} \tag{39}$$

It is reasonable to write the function $w_0(x)$ in the form

$$w_0(x) = a\cos q\left(x - \frac{d}{2}\right), \tag{40}$$

while the function $w(x)$ can be written in the form (35). Then, the boundary conditions for $n = 1$ and the condition of the continuity of stresses at joints for $n > 1$ give the relation

$$\frac{1}{2}[e^{-\kappa d} - 1] = \cos z - Q^* z\sin z, \quad z = qd. \tag{41}$$

By combining again (41) with (32) and (32a), we obtain the expression

$$e^{\kappa d} = 2\xi Qz\sin z - 1. \tag{42}$$

For solutions corresponding to $k = \pi + i\kappa$, when the signs of displacements in adjacent cells are opposite, relation (42) is replaced by

$$e^{\kappa d} = 2\xi Qz\sin z + 1. \tag{43}$$

By using expression (42) or (43), we can readily obtain the expression for $\cosh\kappa d$, which allows us to write

transcendental equations for determining local frequencies in these cases:

$$\cos z - Qz\sin z = \frac{1}{2}\left\{(2\xi Qz\sin z \mp 1) + \frac{1}{(2\xi Qz\sin z \mp 1)}\right\}, \quad (44)$$

where the minus sign corresponds to in-phase vibrations of the boundary points and the plus sign corresponds to antiphase vibrations of the boundary points. In the case of two adjacent defect boundaries, the localization of vibrations in the forbidden band can occur at any sign of the perturbation ξ.

The conditions for the appearance of local frequencies in the case of antisymmetric displacement functions can be analyzed in a similar way. Figure 5 shows the plots of displacements localized near a defect cell.

Finally, let us discuss briefly the possibility of the appearance of local frequencies near a more complicated distortion of the superlattice regularity, when three joints form a defect (for example, $n = 0$ and $n = \pm 1$). We assume that the previous boundary conditions are valid for each of the defect joints.

Consider a symmetric vibration near such a defect. It is obvious that, in the case of a symmetric solution, the displacements are described by a function that is continuous at the joint $n = 0$, which automatically means that stresses at this joint are zero. Therefore, the defect joint with $n = 0$ does not affect the properties of the vibration under study. The frequencies of local vibrations are determined by the equation

$$\cos z - Qz\sin z = \frac{1}{2}\left\{2\xi Qz\sin z + \frac{1}{2\xi Qz\sin z}\right\}. \quad (45)$$

One can see that equations for the frequencies of local vibrations near such a defect have a simpler form than in the case of two defect joints. It is important to note that, in the case of such a defect, not only the value but also the sign of perturbation is substantial: for $\xi Q > 0$, no local vibrations can appear. The latter is evident from the graphical solution of Eq. (45).

Note that, in all cases considered when a local vibration appeared and a defect was localized at more than one joint in the superlattice, two discrete levels appeared simultaneously in the gap, which, however, depended differently on the defect parameter ξ.

5. BLOCH OSCILLATIONS OF A WAVE PACKET IN A HOMOGENEOUS EXTERNAL FIELD

We mentioned above that the propagation of optical light pulses in a system of parallel waveguides can be described by a discrete Schrödinger equation of type (14) in the case of linear packets or of type (16) in the case of solitons. Based on these equations, the theory of motion of wave packets in one-dimensional systems with weakly varying parameters is developed. In particular, one can study Bloch oscillations of a wave packet as a quantum particle in a homogeneous external field. Bloch oscillations of an optical pulse were described and observed experimentally in papers [17, 18]. Therefore, it is worthwhile to discuss Bloch oscillations of a wave packet in superlattices.

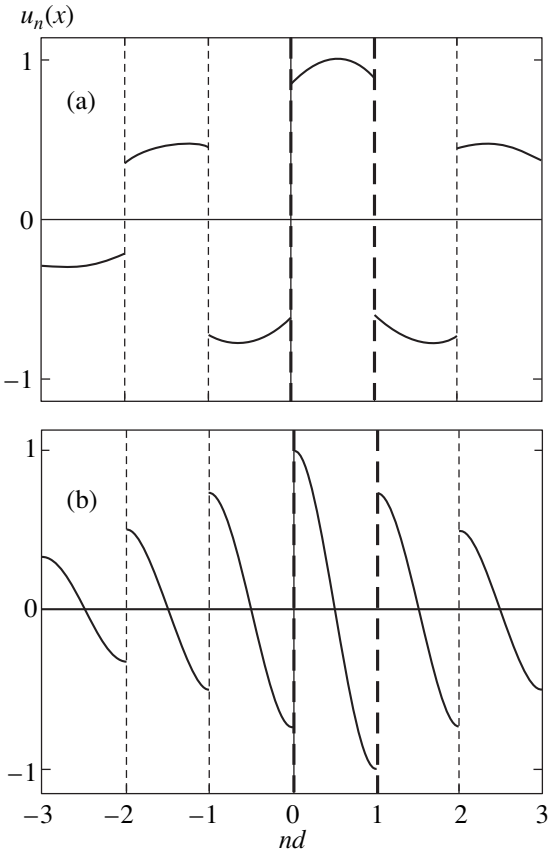


Fig. 5. Plots of solutions for the function $u_n(x)$ describing the coordinate dependence of displacements localized near a defect cell under the following conditions: (a) $u_0(0) = u_1(d)$ and (b) $u_0(0) = -u_1(d)$.

Let us assume that the physical characteristics of the elements of a superlattice weakly depend on the number n. For example, the velocity of a wave changes as $c = c_0 + \Delta c = c_0\{1 + \delta c/c_0\}$, where Δc is a small change in the wave velocity upon passing from one lattice element to another. In experiments [17, 18], such an inhomogeneity was produced by the temperature gradient determining the refractive index of a dielectric. Then, the parameter $m\omega_0$ in expression (8) and Eq. (15) should be replaced by $m\omega_0 + \eta nd$, where η is a small frequency gradient. As a result, in the case of $m = 2p$ (the case of $m = 2p + 1$ is analyzed similarly), we obtain the Schrödinger equation in a homogeneous field

$$i\frac{\partial \psi_n}{\partial t} = m\omega_0\psi_n + \eta nd\psi_n + \frac{\Omega}{2m}(2\psi_n - \psi_{n+1} - \psi_{n-1}), \quad (46)$$

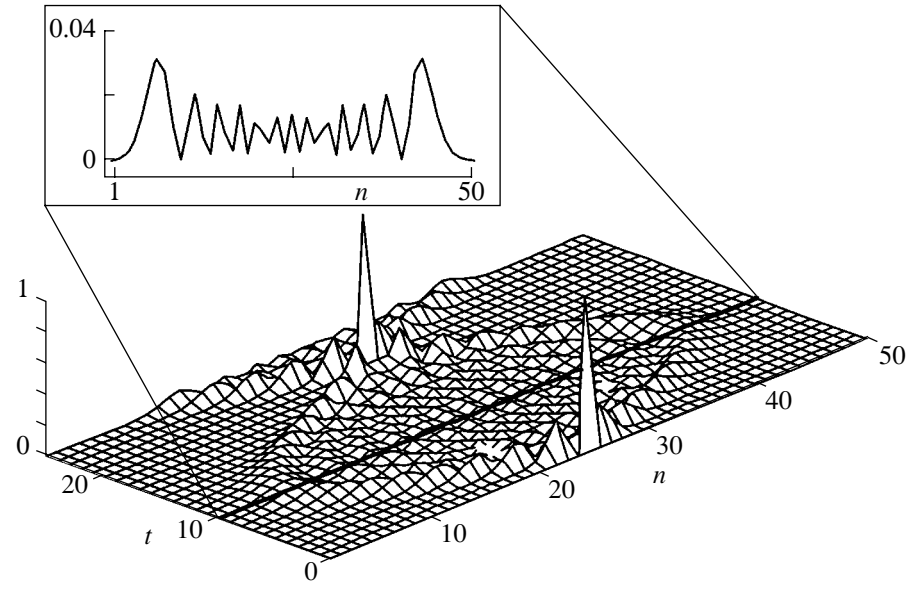


Fig. 6. Evolution of the wave packet undergoing Bloch oscillations in the case of a very initial ($t = 0$) packet in the x space.

and expression (8) transforms to the Hamiltonian of a particle under study, which depends on n and k.

Let us introduce the operator of the lattice-point number, which has the form $n = (i/d)\partial/\partial k$ in the one-band model.[5] Then, Eq. (46) in the k representation will take the following form, which is consistent with the above-mentioned Hamiltonian:

$$i\frac{\partial\psi(q)}{\partial t} = m\omega_0\psi(q) + i\eta\frac{\partial\psi(q)}{\partial q} + \frac{2\Omega}{m}\psi(q)\sin^2\frac{q}{2}, \quad (47)$$

where $q = kd$. A particle obeying Eqs. (46) or (47) experiences the so-called dynamic localization [19], which is manifested in peculiar oscillations with frequency that can be found as follows. The time and wave number in Eq. (47) for stationary states appear in a combination $q + \eta t$. Because the stationary eigenfunction in our case is a periodic function of q with the period 2π, it has an equidistant frequency spectrum with the characteristic frequency

$$\omega_B = \eta = m\Delta\omega_0 = m\omega_0\frac{\Delta c}{c} = m\frac{\pi c}{d}\frac{\Delta c}{c}. \quad (48)$$

This is the frequency of Bloch oscillations of a wave packet in superlattices.

The evolution of a wave packet undergoing Bloch oscillations is analyzed, for example, in reviews [21, 22]. The character of Bloch oscillations depends substantially on the wave packet width. A packet with a very narrow Fourier k spectrum (with well defined k) behaves as a particle with the Hamiltonian coinciding with (8). An external homogeneous field acting on the particle causes the dependence of the wave vector (quasi-momentum) of the particle on time

$$\frac{dk}{dt} = \eta. \quad (49)$$

According to formula (49), the motion of the packet center of gravity in a constant field is described by the expression

$$k = k_0 + \eta t. \quad (50)$$

By substituting (50) into the expression for the particle velocity, which follows from the dispersion law (8), we obtain the periodic time dependence of the particle velocity

$$v = \frac{\partial\omega}{\partial k} = \sin kd = \sin(k_0 + \eta t)d. \quad (51)$$

Expression (51) determines Bloch oscillations: the packet center of gravity oscillates with the frequency ω_B and the amplitude $A = \Omega d/(m\eta)$.

If the initial packet ($t = 0$) is very narrow in the x space (for example, it is localized at one lattice point with $n = 0$), its dynamics will be different. First, the excitation propagates to both sides of the packet center $n = 0$; however, after removal by a distance A, the excitation stops and then returns back, forming the initial pulse at the moment $t = 2\pi/\omega_B$, i.e., after the period of Bloch oscillations (see Fig. 6).

[5] The question about the role of interband transitions in the theory of Bloch oscillations was discussed in the literature [20]. In our "strong coupling" model, only one particle is considered. The appearance of interband interactions in a similar model of a nonlinear optical superlattice is considered in the analysis of the stability of gap self-localized vibrations (solitons) [12].

It seems that the estimate of the Bloch frequency $\omega_B \sim m(c/d)(\Delta c/c)$ and the amplitude $A \sim d(m^2 Q \Delta c/c)^{-1}$ allows their observation in acoustic superlattices.

ACKNOWLEDGMENTS

The authors thank V. Al'shits, A. Maier, and E. Ganapol'skii for useful discussions. This work was partially supported by INTAS, grant no. 167 of 1999.

REFERENCES

1. J. D. Joannopulos, R. D. Meade, and J. N. Win, *Photonic Crystals* (Prinston Univ. Press, Princeton, 1995).
2. L. M. Brekhovskikh, *Waves in Layered Media* (Nauka, Moscow, 1973; Academic, New York, 1980).
3. Pi-Gang Luan and Zhen Ye, Phys. Rev. E **63**, 66611 (2001).
4. A. M. Kosevich, Pis'ma Zh. Éksp. Teor. Fiz. **74**, 633 (2001) [JETP Lett. **74**, 559 (2001)].
5. S. M. Rytov, Zh. Éksp. Teor. Fiz. **29**, 605 (1955) [Sov. Phys. JETP **2**, 466 (1955)].
6. S. M. Rytov, Akust. Zh. **2**, 71 (1956) [Sov. Phys. Acoust. **2**, 68 (1956)].
7. R. de L. Kronig and W. G. Penni, Proc. R. Soc. London, Ser. A **130**, 499 (1931).
8. A. F. Andreev and Yu. A. Kosevich, Zh. Éksp. Teor. Fiz. **81**, 1435 (1981) [Sov. Phys. JETP **54**, 761 (1981)].
9. L. J. Pyrak-Nolte, J. Xu, and G. M. Haley, Phys. Rev. Lett. **68**, 3650 (1992); A. M. Kosevich and A. V. Tutov, Fiz. Nizk. Temp. **19**, 1273 (1993) [Low Temp. Phys. **19**, 905 (1993)]; A. M. Kosevich and A. V. Tutov, Phys. Lett. A **213**, 265 (1996).
10. A. Figotin and V. Gorentsveig, Phys. Rev. B **58**, 180 (1998).
11. A. Hasegawa, *Optical Solitons in Fibers* (Springer-Verlag, Berlin, 1989).
12. A. A. Sukhorukov and Yu. S. Kivshar, Phys. Rev. Lett. **87**, 083901 (2001).
13. M. A. Ablowitz and J. F. Ladik, J. Math. Phys. **17**, 1011 (1976).
14. L. D. Faddeev and L. A. Takhtajan, *Hamiltonian Methods in the Theory of Solitons* (Springer-Verlag, Berlin, 1987).
15. R. Schart and A. Bishop, Phys. Rev. A **43**, 6535 (1991).
16. A. M. Kosevich, Zh. Éksp. Teor. Fiz. **119**, 995 (2001) [JETP **92**, 866 (2001)].
17. N. Petsch, P. Dannberg, W. Elflein, *et al.*, Phys. Rev. Lett. **83**, 4752 (1999).
18. R. Morandotti, U. Perschel, J. S. Aitchison, *et al.*, Phys. Rev. Lett. **83**, 4756 (1999).
19. D. H. Dunlap and V. M. Kenkre, Phys. Lett. A **127**, 438 (1988).
20. M. Gluck, A. R. Kolovski, and H. J. Korsch, Phys. Lett. A **249**, 483 (1998); Phys. Rev. Lett. **82**, 1534 (1999).
21. A. M. Kosevich, Fiz. Nizk. Temp. **27**, 699 (2001) [Low Temp. Phys. **27**, 513 (2001)].
22. M. Gluck, A. R. Kolovski, and H. J. Korsch, submitted to Phys. Rep. (2001).

Translated by M. Sapozhnikov