Journal of Experimental and Theoretical Physics, Vol. 90, No. 5, 2000, pp. 725–732.
Translated from Zhurnal Éksperimental'noĭ i Teoreticheskoĭ Fiziki, Vol. 117, No. 5, 2000, pp. 835–843.

ATOMS, SPECTRA, RADIATION

Synchrotron Radiation of Electrons in Periodic Weakly-Focusing Magnetic Fields

O. E. Shishanin

Moscow State Industrial University, Moscow, 109280 Russia
e-mail: shisha-n@msiu.ru

Received April 20, 1999

Abstract—The dynamics of an electron in a weakly focusing accelerator with rectilinear gaps is studied by the method of averaging and by the perturbation method. The asymptotic solutions found are used to investigate the angular properties of the synchrotron radiation of electrons. It is shown that vertical betatron oscillations play an important role in the formation of the angular distributions of the radiation. © *2000 MAIK "Nauka/Interperiodica".*

1. INTRODUCTION

Interest in the problem posed in the title of this paper has not diminished with time, since a number of weakly focusing accelerators are still operating successfully in this country. The basic aspects of the theory of synchrotron radiation were expounded, specifically, in [1–3], and the corresponding experiments [1, 4–6] were performed first on accelerators with continuous axisymmetric and then discontinuous magnetic fields. These experiments have largely confirmed the results of the theory, but at the same time they have also revealed some new features. In this connection theoretical investigations have been performed of the properties of the radiation for an axisymmetric magnetic field [7–9]. These investigations have explained the change in the spectral-angular distributions of the radiation intensity as being due to betatron oscillations. The experiments subsequently performed at the Kharkov Physicotechnical Institute [10, 11] have already been analyzed on the basis of [8].

The existing machines contain, as a rule, rectilinear gaps with no magnetic field which have a definite effect on the properties of the radiation. In the present paper this question is studied for the example of a somewhat simplified model, where one periodicity cell of a closed orbit consists only of a steering magnet with a uniform field and is completely free of the gap field.

The particle dynamics in periodic magnetic fields is ordinarily described by the matrix method and using the betatron function [12]. However, the direct application of the results obtained to the problem of the radiation of an electron excessively complicates the computational procedure.

To simplify this difficult problem the calculations can proceed from the physical assumptions that the particle trajectory is continuous, stable, and smoothed. This makes it necessary to apply the methods of the theory of averaging and, consequently, to obtain continuous solutions to describe the revolution of an electron on an orbit. This device has been used by the present author to study the characteristics of synchrotron light in periodic magnetic fields with strong focusing [13, 14].

2. DESCRIPTION OF THE DYNAMICAL PROBLEM

Let a charged particle revolve in a magnetic system consisting of N periods, where one element of the system consists of a steering magnet of length $a = 2\pi R/N$ (R is the radius of curvature) and a free gap of length l. The length of the entire orbit will be

$$2\pi R + Nl = 2\pi R_0,$$

where R_0 is the so-called average radius, $R_0 = (1 + k)R$, and we shall assume the parameter $k = l/a$ to be small.

The magnetic-field gradient has the value $0 < n < 1$ for azimuthal angle

$$\varphi \in \left(0, \frac{a}{a+l}\frac{2\pi}{N}\right)$$

and zero for

$$\varphi \in \left(\frac{a}{a+l}\frac{2\pi}{N}, \frac{2\pi}{N}\right).$$

This step function can be represented, after expansion in a series, as $n(\tau) = f(\tau)n$, where the function or the operator $f(\tau)$ has the form

$$f(\tau) = \frac{1}{1+k}\left[1 + \frac{2(1+k)}{\pi}\sum_{\nu=1}^{\infty} g_\nu \cos[\nu(\tau - \tau_1)]\right],$$

$$\tau = N\varphi, \quad \tau_1 = \frac{\pi a}{a+l}, \quad g_\nu = \frac{\sin\nu\tau_1}{\nu}.$$

1063-7761/00/9005-0725$20.00

This operator is 1 on the section with the magnetic field and 0 in the free gap. It also "switches on" and "switches off" the magnetic field components H_r and H_z and the curvature and angular velocity near the steering magnets. We note that the angular velocity for an axisymmetric magnetic field has been found in [7, 8] in the form

$$\dot{\varphi} = \omega_0\left(1 - \frac{\rho}{R} + \frac{3-n}{2}\frac{\rho^2}{R^2} + \frac{n}{2}\frac{z^2}{R^2}\right),$$

where

$$\omega_0 = \frac{ceH_0}{E}, \quad H_0 = \frac{b}{R^n}, \quad b = \text{const}, \quad \rho = r - R.$$

On the basis of the remarks made above the equations for the betatron oscillations of an electron can be put into the form

$$\frac{d^2z}{d\tau^2} + \frac{(1+k)^2}{N^2} n f(\tau) z = 0, \tag{1}$$

$$\frac{d^2\rho}{d\tau^2} + \frac{(1+k)^2}{N^2}(1-n) f(\tau)\rho = 0. \tag{2}$$

These Hill equations give the correct oscillation frequencies. Their solutions will be found in the next two sections.

For what follows it is also necessary to determine the angular velocity of a particle moving in an accelerator with a discontinuous magnetic field. Certain assumptions concerning the radial rotation can be made in order to avoid singularities associated with the transition from circular to rectilinear motion and vice versa. This is also due to the fact that the amplitudes of the oscillations in the plane of the orbit are small compared with the radius and that these oscillations have only a weak effect on the properties of the radiation, in contrast, for example, to axial oscillations. In this connection it can be assumed that a particle moves along a circle with radius R_0, and the guiding magnetic field H_0 can be averaged over the entire length of the period. Then the transverse components of the magnetic field will assume the form

$$H_z = H_0\left[\frac{1}{1+k} - \frac{\rho}{R} n f(\tau)\right],$$

$$H_r = -H_0 \frac{z}{R} n f(\tau),$$

where $\rho = r - R_0$.

The angular velocity of the corresponding average motion can be determined by the expression

$$\dot{\varphi} = \frac{\omega_0}{1+k} \times \left[1 - \frac{\rho}{R_0} + \frac{3}{2}\frac{\rho^2}{R_0^2} + \frac{n}{R^2}\int (z\dot{z} - \rho\dot{\rho}) f(\tau) dt\right]. \tag{3}$$

3. APPLICATION OF THE AVERAGING METHOD

We shall consider equation (1) first. According to the averaging method of [15], we represent this equation in the standard form

$$\frac{dZ}{d\tau} = \varepsilon G Z,$$

where $\varepsilon = 1/N$,

$$Z = \begin{pmatrix} z \\ \frac{1}{\varepsilon}\frac{dz}{d\tau} \end{pmatrix}, \quad G = \begin{pmatrix} 0 & 1 \\ -g & 0 \end{pmatrix}, \quad g = (1+k)^2 n f(\tau).$$

We recall that the general theory of averaging studies the matrix equation

$$dZ/d\tau = \varepsilon Y(\tau, Z),$$

where the averaging operation

$$\langle Y(\tau, Z)\rangle = \lim_{T\to\infty}\frac{1}{T}\int_0^T Y(\tau, Z) d\tau$$

is introduced for the function $Y(\tau, Z)$ and the integrating operator is

$$\tilde{Y}(\tau, \xi) = \int [Y(\tau, \xi) - \langle Y(\tau, \xi)\rangle] d\tau.$$

Here the vector ξ satisfies the equation

$$\frac{d\xi}{d\tau} = \varepsilon\langle Y(\tau, \xi)\rangle,$$

and using this vector the first approximation is determined as

$$Z(\tau) = \xi(\tau) + \varepsilon\tilde{Y}(\tau, \xi).$$

This procedure was known up to the second approximation [15, 16]. However, the next corrections can also be obtained using the method of iterations. Specifically, taking

$$Z(\tau) = \xi + \sum_{i=1}^{3} \varepsilon^i Y_i,$$

we obtain

$$Y_1 = \tilde{Y}, \quad Y_2 = \widetilde{\frac{\partial Y}{\partial \xi} Y_1} - \int \frac{\partial Y_1}{\partial \xi} d\tau \langle Y\rangle,$$

$$Y_3 = \widetilde{\frac{\partial Y}{\partial \xi} Y_2} - \int \frac{\partial Y_1}{\partial \xi} d\tau \left\langle \frac{\partial Y}{\partial \xi} Y_1 \right\rangle - \int \frac{\partial Y_2}{\partial \xi} d\tau \langle Y \rangle.$$

For $Y(\tau, Z) = GZ$ we obtain

$$Z(\tau) = \left(1 + \sum_{i=1}^{3} \varepsilon^i G_i\right)\xi, \tag{4}$$

where

$$\frac{d\xi}{d\tau} = \varepsilon \left\langle G\left(1 + \sum_{i=1}^{3} \varepsilon^i G_i\right)\right\rangle \xi, \tag{5}$$

and the matrices have the form

$$G_1 = \tilde{G}, \quad G_2 = \widetilde{GG}_1 - \tilde{G}_1 \langle G \rangle,$$

$$G_3 = \widetilde{GG}_2 - \tilde{G}_1 \langle GG_1 \rangle - \tilde{G}_2 \langle G \rangle.$$

For our problem

$$\langle G \rangle = \begin{pmatrix} 0 & 1 \\ -n(1+k) & 0 \end{pmatrix}, \quad \tilde{G} = \begin{pmatrix} 0 & 0 \\ -\tilde{g} & 0 \end{pmatrix},$$

$$G_2 = \begin{pmatrix} -\tilde{\tilde{g}} & 0 \\ 0 & \tilde{\tilde{g}} \end{pmatrix}, \quad G_3 = \begin{pmatrix} 0 & 2\tilde{\tilde{g}} \\ \widetilde{g\tilde{g}} + n(1+k)\tilde{\tilde{g}} & 0 \end{pmatrix},$$

where, specifically,

$$g = n(1+k) + \frac{2n(1+k)^2}{\pi} \sum_{\nu=1}^{\infty} g_\nu \cos[\nu(\tau - \tau_1)],$$

$$\langle g \rangle = n(1+k),$$

$$\tilde{g} = \frac{2n(1+k)^2}{\pi} \sum_{\nu=1}^{\infty} \frac{g_\nu}{\nu} \sin[\nu(\tau - \tau_1)],$$

$$\langle \tilde{g} \rangle = 0,$$

$$\tilde{\tilde{g}} = -\frac{2n(1+k)^2}{\pi} \sum_{\nu=1}^{\infty} \frac{g_\nu}{\nu^2} \cos[\nu(\tau - \tau_1)].$$

An equation for ξ (here ξ is no longer a vector) can be found from equation (5):

$$\frac{d^2\xi}{d\tau^2} = -\frac{1}{N^2}\left[n(1+k) - \frac{1}{N^2}\langle g\tilde{\tilde{g}} \rangle\right]\xi,$$

where

$$\langle g\tilde{\tilde{g}} \rangle = -\frac{2n^2(1+k)^4}{\pi^2} \sum_{\nu=1}^{\infty} \frac{g_\nu^2}{\nu^2} = -\frac{\pi^2 n^2 k^2}{3}.$$

Denoting the expression in brackets as ν_z^2, we can write the average solution as

$$\xi = B\cos\left(\frac{\nu_z}{N}\tau + \delta\right),$$

where B and δ are, respectively, the amplitude and initial phase of the oscillations.

The frequency

$$\nu_z = \sqrt{n(1+k) + \frac{\pi^2 n^2 k^2}{3N^2}} \tag{6}$$

agrees, in the indicated approximation, with the well-known formula [17] for weak focusing

$$\cos\mu = \cos\left(\frac{2\pi}{N}\sqrt{n}\right) - \frac{l}{2}\frac{\sqrt{n}}{R}\sin\left(\frac{2\pi}{N}\sqrt{n}\right),$$

where $\mu = 2\pi\nu_z/N$ is the phase shift.

Substituting the matrices G_i into equation (4) we obtain

$$z = \left(1 - \frac{1}{N^2}\tilde{\tilde{g}}\right)\xi + \frac{2}{N^3}\tilde{\tilde{g}}\frac{d\xi}{d\varphi}.$$

This solution can be represented as

$$z = B\cos\left(\frac{\nu_z}{N}\tau + \delta\right)(1 + S_1) + B\sin\left(\frac{\nu_z}{N}\tau + \delta\right)\nu_z S_2, \tag{7}$$

where

$$S_1 = \frac{2n(1+k)^2}{\pi N^2} \sum_{\nu=1}^{\infty} \frac{g_\nu}{\nu^2} \cos[\nu(\tau - \tau_1)],$$

$$S_2 = \frac{4n(1+k)^2}{\pi N^3} \sum_{\nu=1}^{\infty} \frac{g_\nu}{\nu^3} \sin[\nu(\tau - \tau_1)].$$

The asymptotic form can be written, by analogy, for equation (2) also:

$$\rho = A\cos\left(\frac{\nu_\rho}{N}\tau + \chi\right)(1 + S_1) + A\sin\left(\frac{\nu_\rho}{N}\tau + \chi\right)\nu_\rho S_2, \tag{8}$$

where A and ρ are, respectively, the amplitude and phase of the radial oscillations,

$$\nu_\rho = \sqrt{(1-n)(1+k) + \frac{\pi^2(1-n)^2 k^2}{3N^2}}.$$

It is interesting to note that the series appearing in the solutions (7) and (8) can be summed [18] and expressed in terms of Bernoulli polynomials as follows:

$$\sum_{\nu=1}^{\infty}\frac{g_\nu}{\nu^2}\cos[\nu(\tau-\tau_1)] = \frac{\pi^3}{3}\left[B_3\left(\frac{\tau}{2\pi}\right)-B_3\left(\frac{\tau-2\tau_1}{2\pi}\right)\right],$$

$$\sum_{\nu=1}^{\infty}\frac{g_\nu}{\nu^3}\sin[\nu(\tau-\tau_1)] = \frac{\pi^4}{6}\left[B_4\left(\frac{\tau}{2\pi}\right)-B_4\left(\frac{\tau-2\tau_1}{2\pi}\right)\right],$$

$$B_3(x) = x^3-\frac{3}{2}x^2+\frac{1}{2}x,$$

$$B_4(x) = x^4-2x^3+x^2-\frac{1}{30},$$

where $0 \le x \le 2$.

4. PERTURBATION METHOD

The results found in the preceding section can also be compared with the expressions obtained below by combining the methods used in [15, 19] to solve the Mathieu equation.

For example, let us return to the Hill equation (1). We seek its solution, according to the Floquet theorem, in the form $z = \exp(i\gamma_z\tau)\varphi_z(\tau)$, where $\varphi_z(\tau) = \varphi_z(\tau + 2\pi)$. We set $\mathrm{Im}\gamma_z = 0$ in order to switch to the region of stability. Now we have a different differential equation for $\varphi_z(\tau)$:

$$\frac{d^2\varphi_z}{d\tau^2}+2i\gamma_z\frac{d\varphi_z}{d\tau}+\left[\frac{(1+k)^2 n f(\tau)}{N^2}-\gamma_z^2\right]\varphi_z = 0.$$

We set here

$$\varphi_z = \sum_{i=0}^{\infty}\frac{\varphi_i}{N^i} \text{ and } \gamma_z = \sum_{i=1}^{\infty}\frac{\gamma_i}{N^i}.$$

Substituting these expressions in the preceding equation and equating terms with like powers of the parameter $1/N$ to zero we arrive at the following sequence of simpler equations:

$$\ddot\varphi_0 = 0, \quad \ddot\varphi_1 + 2i\gamma_1\dot\varphi_0 = 0,$$

$$\ddot\varphi_2 = 2i(\gamma_1\dot\varphi_1+\gamma_2\dot\varphi_0)+[(1+k)^2 nf-\gamma_1^2]\varphi_0 = 0,$$

$$\ddot\varphi_3+2i(\gamma_1\dot\varphi_2+\gamma_2\dot\varphi_1+\gamma_3\dot\varphi_0)+[(1+k)^2 nf-\gamma_1^2]\varphi_1-2\gamma_1\gamma_2\varphi_0 = 0,$$

$$\ddot\varphi_4+2i(\gamma_1\dot\varphi_3+\gamma_2\dot\varphi_2+\gamma_3\dot\varphi_1+\gamma_4\dot\varphi_0)+[(1+k)^2 nf-\gamma_1^2]\varphi_2-2\gamma_1\gamma_2\varphi_1-2\gamma_1\gamma_3\varphi_0-\gamma_2^2\varphi_0 = 0.$$

The first two equations give $\varphi_0 = b$ and $\varphi_1 = b_1$, where b_0, b_1 = const. According to [15, 19], the next equations must be solved by illuminating the secular terms. We introduce the notation

$$C_2 = \sum_{\nu=1}^{\infty}\frac{g_\nu}{\nu^2}\cos[\nu(\tau-\tau_1)],$$

$$S_3 = \sum_{\nu=1}^{\infty}\frac{g_\nu}{\nu^3}\sin[\nu(\tau-\tau_1)].$$

We obtain from the third equation

$$\gamma_1 = \sqrt{(1+k)n} \text{ and } \varphi_2 = 2bn(1+k)^2 C_2/\pi.$$

Next we find that $\gamma_2 = 0$,

$$\varphi_3 = -\frac{4ibn\gamma_1(1+k)^2 S_3}{\pi}+\frac{2nb_1(1+k)^2 C_2}{\pi}.$$

The last equation gives

$$2\gamma_1\gamma_3 = \frac{2n^2(1+k)^4}{\pi^2}\sum_{\nu=1}^{\infty}\frac{g_\nu^2}{\nu^2},$$

whence we find $\gamma_3 = \pi^2 n^2 k^2/6\gamma_1$.

The frequency $\nu_z = \gamma N = \gamma_1 + \gamma_3/N^2$ is identical to the expression (6) found above. The general solution is constructed in accordance with the Floquet theory as

$$z = C\exp(i\gamma_z\tau)\varphi_z(\tau)+C^*\exp(-i\gamma_z\tau)\varphi_z^*(\tau).$$

Substituting $(B/2)\exp(i\delta)$ for the arbitrary constants $C(b + b_1/N)$, the asymptotic solution assumes the form

$$z = B\cos\left(\frac{\nu_z}{N}\tau+\delta\right)\left[1+\frac{2n(1+k)^2}{\pi N^2}C_2\right]+B\sin\left(\frac{\nu_z}{N}\tau+\delta\right)\gamma_1\frac{4n(1+k)^2}{\pi N^3}S_3$$

and is identical to equation (7).

We note that the last method is simpler than the averaging procedure. The expansion parameter in these asymptotic expansions is actually the quantity n/N^2, which, for example, for $n = 0.67$ and $N = 4$ will be approximately 0.04.

5. PROPERTIES OF SYNCHROTRON RADIATION

To study the synchrotron light itself we shall use Schwinger's semiclassical approach [20, 21], where the quantum recoil accompanying the emission of photons is taken into account but smaller corrections associated with the quantum motion of an electron in a magnetic field are neglected.

We shall give the direction of the radiation vector $\boldsymbol{\kappa} = w\mathbf{n}/c$ by the vector θ, where $\mathbf{n} = \{0, \sin\theta, \cos\theta\}$ is the spherical angle. The components of linear polarization of the radiation, which are characterized by two

orthogonal components of the electric field vector, can be determined by the vectors

$$\mathbf{e}_\sigma = \{1, 0, 0\} \text{ and } \mathbf{e}_\pi = \{0, \cos\theta, -\sin\theta\}.$$

Using the methods of [8], the components of the linear polarization of the radiation intensity can be represented in the first quantum approximation as

$$\frac{dW_\sigma}{d^3\kappa} = \frac{ce^2\nu'}{(2\pi)^3 R_0\nu}\left|\int dt\, v_x \exp\left[i\frac{\nu'}{\nu}(\omega t - \boldsymbol{\kappa}\cdot\mathbf{r})\right]\right|^2,$$

$$\frac{dW_\pi}{d^3\boldsymbol{\kappa}} = \frac{ce^2\nu'}{(2\pi)^3 R_0\nu} \tag{9}$$

$$\times\left|\int dt(v_y\cos\theta - v_z\sin\theta)\exp\left[i\frac{\nu'}{\nu}(\omega t - \boldsymbol{\kappa}\cdot\mathbf{r})\right]\right|^2,$$

where the radiation frequency is $\omega = \nu\omega_0$ and $\nu' = \nu(1 + \hbar\omega/E)$.

Here the phase is

$$\omega t - \boldsymbol{\kappa}\cdot\mathbf{r} = \frac{\nu}{1+k}$$

$$\times\left[\omega_0 t - \frac{\omega_0}{c}(R_0 + \rho)\sin\rho\sin\theta - \frac{\omega_0}{c}z\cos\theta\right].$$

Radiation in a prescribed direction, where, for example, a detector is located, is formed on a small section of the orbit with angle $\varphi \approx \omega_0 t/(1 + k)$ and near the plane of the orbit $\theta \sim \pi/2$ ($\cos\theta \sim m_0c^2/E$). We shall also assume τ, A/R_0, and B/R_0 to be expansion parameters.

From equation (7) we determine to within $1/N^2$

$$\frac{dz}{dt} = -B\nu_z\frac{\omega_0}{1+k}\sin\left(\frac{\nu_z}{N}\tau + \delta\right)$$

$$\times\left\{1 - \frac{2n(1+k)^2}{\pi N^2}\sum_{\nu=1}^{\infty}\frac{g_\nu}{\nu^2}\cos[\nu(\tau - \tau_1)]\right\}$$

$$-B\frac{2n\omega_0(1+k)}{\pi N}\cos\left(\frac{\nu_z}{N}\tau + \delta\right) \tag{10}$$

$$\times\sum_{\nu=1}^{\infty}\frac{g_\nu}{\nu}\sin[\nu(\tau - \tau_1)].$$

If only the main terms are taken into account in the components of the vector $\boldsymbol{\beta} = \mathbf{v}/c$

$$\left\{\frac{\dot\rho}{c}\cos\varphi - (R_0 + \rho)\sin\varphi\frac{\dot\varphi}{c}, \frac{\dot\rho}{c}\sin\varphi + (R_0 + \rho)\cos\varphi\frac{\dot\varphi}{c}, \frac{\dot z}{c}\right\},$$

then they can be represented as

$$\beta_x \approx -\varphi - \nu_\rho\frac{A}{R_0}\sin\chi\left[1 + \frac{\pi^2(1-n)k(1-k)}{3N^2(1+k)}\right]$$

$$+\frac{A}{R_0}\frac{\pi(1-n)k}{N}\cos\chi,$$

$$\beta_y \approx 1, \tag{11}$$

$$\beta_z \approx -\nu_z\frac{B}{R_0}\sin\delta\left[1 + \frac{\pi^2 nk(1-k)}{3N^2(1+k)}\right] + \frac{B}{R_0}\frac{\pi nk}{N}\cos\delta.$$

Here we also took account of the fact that

$$\sum_{\nu=1}^{\infty}\frac{g_\nu}{\nu}\sin(\nu\tau_1) = \frac{\pi^2 k}{2(1+k)^2},$$

$$\sum_{\nu=1}^{\infty}\frac{g_\nu}{\nu^2}\cos(\nu\tau_1) = -\frac{\pi^3 k(1-k)}{6(1+k)^3}.$$

From equation (3) we determine, by integrating, φ and therefore the quantity $\sin\varphi \approx \varphi - \varphi^3/6$. The parameter $R_0\omega_0/c$ can be found using the expression for the total velocity

$$v^2 = \dot\rho^2 + \dot z^2 + r^2\dot\varphi^2.$$

Making the substitution $\beta_x = -\omega_0 t$ to calculate the integrals appearing in equation (9), the expressions obtained will no longer depend on the radial oscillations. Moreover, averaging them over the initial phases δ of the axial oscillations, the spectral-angular formulas of the radiation can be represented in the form

$$\frac{dW_\sigma(\nu)}{d\Omega} = \frac{ce^2\nu}{2^{2/3}\pi^3 R_0^2(\nu')^{1/3}}\int_0^{2\pi} d\delta\left[\Phi'\left(\left(\frac{\nu'}{2}\right)^{2/3}\varepsilon_1\right)\right]^2,$$

$$\frac{dW_\pi(\nu)}{d\Omega} = \frac{ce^2\nu(\nu')^{1/3}}{2^{4/3}\pi^3 R_0^2}\int_0^{2\pi} d\delta\varepsilon_2\Phi^2\left(\left(\frac{\nu'}{2}\right)^{2/3}\varepsilon_1\right). \tag{12}$$

In these expressions

$$\varepsilon_1 = 1 - \beta^2 + \varepsilon_2,$$

$$\varepsilon_2 = \left[\cos\theta - \nu_\nu\frac{B}{R_0}\cos(\delta + \delta_0)\right]^2,$$

$$\nu_\nu = \nu_z\sqrt{1 + \frac{\pi^2 nk(2+k)}{3N^2(1+k)}},$$

$$\cos\delta_0 = \frac{\pi nk}{N\nu_\nu}.$$

Since radiation occurs only at the moment when the particle passes through a steering magnet, the total motion is more extended than a circular orbit, and this leads in equation (12) to R_0 instead of R.

The transition from the results obtained to an axisymmetric magnetic field [8] occurs for $k = 0$.

The quantum corrections in the formulas (12) are negligible. However, because the parameters $\cos\theta$,

$\nu_v B/R_0$, and m_0c^2/E are infinitesimals of the same order, the curves of the spectral-angular and angular distributions will differ strongly from the corresponding plots for a uniform magnetic field.

To perform the integration over δ in equation (12) and to use existing tables [22], different specifications can be used for the Airy function $V(x)$ and its derivative with the initial Fock conditions, when $V(0) = 0.629271$ and $V'(0) = -0.458745$. We note that $V(x) = \Phi(x) = \sqrt{\pi}\,\mathrm{Ai}(x)$.

Then the expressions (12) in the classical approximation will assume the form

$$\frac{dW_\sigma(\nu)}{d\Omega} = W_1\frac{1}{2\pi}\int_0^{2\pi} d\delta V'^2(x_1),$$
$$\frac{dW_\pi(\nu)}{d\Omega} = W_1\left(\frac{\nu}{2}\right)^{2/3}\frac{1}{2\pi}\int_0^{2\pi} d\delta\varepsilon_2 V^2(x_1), \tag{13}$$

where

$$W_1 = \frac{2^{1/3}ce^2\nu^{2/3}}{\pi^2R_0^2} \text{ and } x_1 = \varepsilon_1\left(\frac{\nu}{2}\right)^{2/3}.$$

The circular polarization can be described using the formula

$$\frac{dW_i(\nu)}{d\Omega} = W_1\frac{1}{2\pi}$$
$$\times\int_0^{2\pi} d\delta\left\{l_2V'(x_1) + l_3\left(\frac{\nu}{2}\right)^{1/3}\sqrt{\varepsilon_2}V(x_1)\right\}^2, \tag{14}$$

where we arrive at right-hand circular polarization ($i = 1$) for $l_2 = -l_3 = 1/\sqrt{2}$ and left-hand circular polarization ($i = -1$) for $l_2 = l_3 = 1/\sqrt{2}$.

For low frequencies (in the visible part of the spectrum) and near the plane of the orbit, where the parameter x_1 is much less than 1, the right-hand sides of equation (13) will assume, correspondingly, the form

$$W_1\left[c_1 + c_3x^2\left(1 + 2p + 4pg + \frac{3}{2}p^2\right)\right.$$
$$\left. + \frac{1}{3}c_1x^2[2 + 3p(2 + 3p)(1 + 4g) + 5p^3]\right],$$
$$W_1x\left[c_2(g + p) + c_3x(2p + 3p^2 + 2g + 10pg) + \frac{1}{2}c_1x^2\right.$$
$$\left.\times[2p + 6p^2 + 5p^3 + g(2 + 20p + 39p^2) + 8pg^2]\right], \tag{15}$$

where

$$x = \varepsilon(\nu/2)^{2/3}, \quad c_1 = V'^2(0),$$
$$c_2 = V^2(0), \quad c_3 = V(0)V'(0),$$
$$p = \frac{q^2}{2\varepsilon}, \quad g = \frac{\cos^2\theta}{\varepsilon}, \text{ and } q^2 = \frac{\nu_v^2B^2}{R_0^2}.$$

These formulas are convenient in the indicated region for estimating the degree of polarization

$$(dW_\sigma - dW_\pi)/(dW_\sigma + dW_\pi),$$

which for $\theta = \pi/2$ is no longer equal to 1 because of the presence of vertical oscillations.

Away from the plane of the orbit $\varepsilon \approx 1 - \beta^2 + \cos^2\theta$ increases and $p \ll 1$. In this case we obtain on the right-hand side in equation (13)

$$W_1\{V'^2 + 2x^2p[2xgU + (1 + 2g)VV']$$
$$+ \frac{1}{2}x^2p^2[12x^2g(1 + g)V^2 + x(3 + 24g + 16x^3g^2)U$$
$$+ 3(1 + 16x^3g + 24x^3g^2)VV']\},$$
$$W_1x\{(g + p)V^2 + 2xgp(2xU + 5VV') + 3xp^2VV'$$
$$+ \frac{1}{2}x^2gp^2[(39 + 16x^3g^2)U + 28xgV^2$$
$$+ 8x^2g(14 + 3g)VV']\}, \tag{16}$$

where $V = V(x)$, $V' = V'(x)$, and $U = xV^2 + V'^2$.

Since there are no betatron oscillations in a uniform magnetic field, $p = 0$ and the known spectral-angular distributions can be written as

$$\frac{dW_\sigma(\nu)}{d\Omega} = W_1V'^2(x),$$
$$\frac{dW_\pi(\nu)}{d\Omega} = W_1xgV^2(x). \tag{17}$$

Using the formulas presented above and additional terms, it is possible to calculate plots of the angular distributions (see figure) near the critical wavelength $\lambda = 50$ Å, taking, for example, the parameters for the Tomsk synchrotron [23]. In this case $n = 0.58$, $N = 4$, $l = 157$ cm, and $R = 423$ cm. Here, for electron energy $E = 1$ GeV the rms amplitude of the vertical oscillations B is about 2 mm. The effect of betatron oscillations on both radiation components can be seen in the figure. The degree of polarization in the plane of motion of the particles is 0.830 instead of 1.

In addition, the calculation performed with $\lambda = 35$ Å (for a given energy this corresponds to the extreme value $\nu \sim (m_0c^2/E)^3$) as shown in the presence of a maximum in the π component in the plane of the orbit instead of a minimum, a further decrease in the values for the σ component, and a decrease in the degree of polarization to 0.818. A similar analysis for wave-

lengths 1000 and 1500 Å has revealed that the role of vertical oscillations in the formation of radiation decreases when a shift into the long-wavelength region of the spectrum occurs. Here the degree of polarization is 0.953 and 0.963, respectively.

Approximately the same picture is also observed for angular distributions if summation over the entire spectrum is performed in the formulas (12).

Then we obtain without quantum corrections

$$\frac{dW_\sigma}{d\Omega} = \frac{7ce^2}{64\pi^2 R_0^2}\int_0^{2\pi} d\delta \frac{1}{\varepsilon_1^{5/2}},$$
$$\frac{dW_\pi}{d\Omega} = \frac{5ce^2}{64\pi^2 R_0^2}\int_0^{2\pi} d\delta \frac{\varepsilon_2}{\varepsilon_1^{7/2}}. \quad (18)$$

Methods for calculating the integrals appearing in equations (18) have been examined in [7]. In this case we introduce the additional notations

$$\varepsilon = \varepsilon_0 + \cos^2\theta, \quad \varepsilon_0 = 1-\beta^2, \quad p_1 = q^2/\varepsilon_0,$$
$$p_2 = q^2/\varepsilon, \quad g_1 = \cos^2\theta/\varepsilon_0,$$
$$f = \varepsilon_0/\varepsilon, \quad \Delta = (1+p_2)^2 - 4p_2 g,$$
$$2r^2 = 1-(1-p_2)/\sqrt{\Delta}.$$

The angular distributions of the components of linear polarization can be represented as

$$\frac{dW_\sigma}{d\Omega} = \frac{14W_2}{3\pi\varepsilon^{5/2}\Delta^{5/4}}$$
$$\times\left\{\left(3+p_1+16\frac{p_2 g}{\Delta}-\frac{2p_2}{\sqrt{\Delta}r^2}G_1\right)\mathbf{K}+\frac{2\sqrt{\Delta}}{f}G_1\mathbf{E}\right\},$$
$$\frac{dW_\pi}{d\Omega} = \frac{2W_2}{3\pi\varepsilon^{5/2}\Delta^{5/4}} \quad (19)$$
$$\times\left\{\left(G_2-\frac{p_2}{\sqrt{\Delta}r^2}G_3\right)\mathbf{K}+\frac{\sqrt{\Delta}}{f}G_3\mathbf{E}\right\},$$

where $\mathbf{K}(r)$ and $\mathbf{E}(r)$ are complete elliptic integrals,

$$W_2 = \frac{ce^2}{32\pi R_0^2}, \quad G_1 = p_1 - g_1 + \frac{2}{\Delta}[(p_2+f)^2 - g^2],$$
$$G_2 = p_1 + \frac{1}{\Delta}[8p_2(p_2+f) - 25p_2 g + 15g]$$
$$+\frac{8p_2 g}{\Delta^2}[9(p_2-g)^2 - 7f^2 + 2f(p_2+g)],$$
$$G_3 = 1 + 2(p_1 - g_1)$$
$$+\frac{1}{\Delta}[4(p_2^2 - f^2) + 3g(p_2 - 7f) - 7g^2]$$
$$+8\frac{gf}{\Delta^2}[7 - 9p_2^2 + 2p_2(g-f)].$$

To construct plots the values of the total elliptic integrals can be taken from the tables in [24].

However, if the integration over the angles θ and ψ is performed in equation (12), then the spectral formulas will be essentially identical to the corresponding expressions for a uniform field.

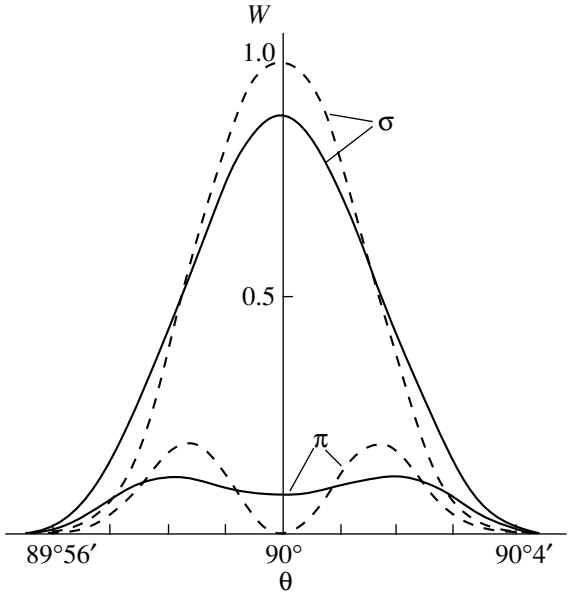
Additional analysis has also revealed the physical meaning of the parameter

$$\nu_\nu(B/R_0)\cos(\delta+\delta_0)$$

in ε_2, which is equal to $\beta_z|_{\tau=0}$ or the slope angle of the electron velocity vector. For any magnetic system it is sufficient to find formulas similar to equation (10) or (11) and to take them into account in the expressions (12). Such an approach makes it possible to return to the question of using the technique of β-functions, since plots of these functions are available for all recently constructed accelerators. Then, for the vertical oscillations

$$z = \sqrt{\frac{\beta_z A_z}{\pi}}\cos\left(\int\frac{ds}{\beta_z}+\delta_0\right),$$

where A is the emittance, β_z is the betatron function which depends on the length s of the orbit, and δ_0 is the initial phase. Here the emittance is the phase volume



Spectral-angular distribution of the components of linear polarization with radiation wavelength $\lambda = 50$ Å and energy $E = 1$ GeV. The dashed lines correspond to a uniform magnetic field.

occupied by the transverse motion of the particles in the beam.

Then, in equations (12), as shown in [25],

$$\varepsilon_2 = (\cos\theta - \alpha\cos\delta)^2,$$

$$\alpha = \sqrt{\frac{A_z}{\pi}}\left[\frac{1}{\sqrt{\beta_z}}\sqrt{1+\left(\frac{1}{2}\frac{d\beta_z}{ds}\right)^2}\right]_{\tau=0}.$$

Here the angle can be measured from the point whence radiation is extracted, and the derivative can be replaced by the tangent of the slope angle of the plot of the β_z function, where the dimensions along the both coordinate axes should be the same.

6. CONCLUSIONS

Since synchrotron radiation in accelerators is mainly incoherent, the results of the present paper are also applicable for a beam of particles. Nonetheless, the influence of the beam was taken into account here by averaging over the initial phases of the oscillations, since particle injection into the accelerator occurs during several revolutions, as well as by using in the computational formulas the rms amplitude of the oscillations. The latter parameter replaces the entire diversity of amplitudes from 0 up to maximum value in any transverse cross section of the beam. It follows from the expressions obtained that in order to improve the polarization, the amplitudes of the vertical oscillations of the particles must be decreased in the beam. Using the parameter α, the formulas obtained were extended to the magnetic systems of other accelerators and storage rings.

ACKNOWLEDGMENTS

I am deeply grateful to V.Ch. Zhukovskiĭ for helpful discussions.

REFERENCES

1. A. A. Sokolov and I. M. Ternov, *Synchrotron Radiation* (Akademie-Verlag, Berlin, 1968).
2. G. N. Kulipanov and A. N. Skrinskiĭ, Usp. Fiz. Nauk **122**, 369 (1977) [Sov. Phys. Usp. **20**, 559 (1977)].
3. I. M. Ternov, Usp. Fiz. Nauk **165**, 429 (1995) [Phys. Usp. **38**, 409 (1995)].
4. O. F. Kulikov, Tr. Fiz. Inst. Akad. Nauk SSSR **80**, 4 (1975).
5. A. A. Vorob'ev, M. M. Nikitin, and A. V. Kozhevnikov, At. Énerg. **29**, 389 (1970).
6. V. G. Bagrov and M. M. Nikitin, At. Énerg. **32**, 243 (1972).
7. V. Ch. Zhukovskiĭ and O. E. Shishanin, Izv. Vyssh. Uchebn. Zaved. Fiz., No. 3, 52 (1973).
8. V. Ch. Zhukovskiĭ and O. E. Shishanin, Zh. Éksp. Teor. Fiz. **61**, 1371 (1971).
9. V. Ch. Zhukovskiĭ and O. E. Shishanin, Opt. Spektrosk. **31**, 857 (1971).
10. I. S. Guk and P. I. Gladkikh, Zh. Tekh. Fiz. **50**, 890 (1980) [Sov. Phys. Tech. Phys. **25**, 537 (1980)].
11. I. S. Guk, P. I. Gladkikh, and S. G. Kononenko, Pis'ma Zh. Tekh. Fiz. **7**, 44 (1981) [Sov. Tech. Phys. Lett. **7**, 19 (1981)].
12. A. N. Lebedev and A. V. Shal'nov, *Foundations of Accelerator Physics and Engineering* (Énergoatomizdat, Moscow, 1991).
13. O. E. Shishanin, Zh. Éksp. Teor. Fiz. **103**, 1117 (1993) [JETP **76**, 547 (1993)].
14. O. E. Shishanin, Pis'ma Zh. Éksp. Teor. Fiz. **57**, 772 (1993) [JETP Lett. **57**, 785 (1993)].
15. N. N. Bogolyubov and Yu. A. Mitropol'skii, *Asymptotic Methods in the Theory of Nonlinear Oscillations* (Nauka, Moscow, 1974, 4th ed.; Gordon and Breach, New York, 1962).
16. A. A. Kolomensky and A. N. Lebedev, *Theory of Cyclic Accelerators* (Fizmatgiz, Moscow, 1962; North-Holland, Amsterdam, 1966).
17. H. Bruck, *Accelerateurs Circulaires de Particules* (Presses Universitaires de France, Paris, 1966; Atomizdat, Moscow, 1970).
18. A. P. Prudnikov, Yu. A. Brychkov, and O. I. Marichev, *Integrals and Series* (Nauka, Moscow, 1981; Gordon and Breach, New York, 1986), Vol. 1.
19. A. Nayfeh, *Introduction to Perturbation Techniques* (Wiley, New York, 1981; Mir, Moscow, 1984).
20. J. Schwinger, Proc. Natl. Acad. Sci. USA **40**, 132 (1954).
21. V. B. Berestetskiĭ, E. M. Lifshitz, and L. P. Pitaevskiĭ, *Quantum Electrodynamics* (Nauka, Moscow, 1989; Pergamon Press, Oxford, 1982).
22. G. D. Yakovleva, *Tables of Airy Functions and Their Derivatives* (Nauka, Moscow, 1969).
23. A. A. Vorob'ev, I. P. Chuchalin, and A. G. Vlasov, *1.5-GeV TPI Synchrotrin* (Atomizdat, Moscow, 1968).
24. E. Jahnke, F. Emde, and F. Losch, *Tables of Higher Functions* (McGraw-Hill, New York, 1960; Nauka, Moscow, 1977).
25. O. E. Shishanin, Zh. Tekh. Fiz. **68** (11), 133 (1998) [Tech. Phys. **43**, 1393 (1998)].

Translation was provided by AIP

Journal of Experimental and Theoretical Physics, Vol. 90, No. 5, 2000, pp. 733–740.
Translated from Zhurnal Éksperimental'noĭ i Teoreticheskoĭ Fiziki, Vol. 117, No. 5, 2000, pp. 844–852.

ATOMS, SPECTRA, RADIATION

Resonance Fluorescence Excited by Macroscopic Superposition in a Feedback Loop

D. B. Horoshko* and S. Ya. Kilin

Stepanov Institute of Physics, Belorussian Academy of Sciences, Minsk, 220072 Belarus
**e-mail: horoshko@ifanbel.bas-net.by*

Received June 25, 1999

Abstract—The resonance fluorescence of an individual atom excited by an optical field in a Yurke–Stoler state, consisting of a superposition of two coherent states with opposite phase, is studied. It is proposed that the decoherence of the field state be eliminated by means of electrooptic feedback [Phys. Rev. Lett. **78**, 840 (1997)]. The master equation for the density operator of the atom–field system is derived and an analytic solution is obtained for the case where the change in the field is adiabatically slow. It is shown that the interaction entangles the atomic and field states. A new effect is predicted: there are no Rabi oscillations of the dipole moment and of the atomic populations with the excitation method described. *© 2000 MAIK "Nauka/Interperiodica".*

1. INTRODUCTION

Resonance fluorescence is one of the most fundamental phenomena in modern optics and atomic physics. The history of the investigation of this phenomenon starts in 1931, when elastic resonance scattering of light by an atom was studied in [1]. In the 1960s the advent of lasers made it realistic to study inelastic scattering—resonance fluorescence itself. In a series of works [2] a semiclassical theory of resonance fluorescence was constructed and one of the most striking properties of this phenomenon was found—the triplet structure of the fluorescence spectrum, experimentally confirmed in the 1970s [3]. Further investigations in this field clarified the quantum features of resonance fluorescence of an individual atom excited by a coherent field. Phenomena such as antibunching of fluorescence photons [4, 5] and bunching of photons of the side components [6, 7] were predicted and observed experimentally. At the end of the 1980s, when nonclassical sources of light appeared, great interest developed in studying resonance fluorescence excited by fields different from a coherent laser field, for example, resonance interaction of an atom with a squeezed vacuum [8].

In the last few years, one of the most intently studied quantum states of light is a superposition of two coherent states with opposite phase [9]: $|\psi\rangle = N(|\alpha\rangle + e^{i\varphi}|-\alpha\rangle)$, where N is a normalization factor and φ is the relative phase. Such a state realizes the well-known "Schrödinger's cat" paradox [10], and the experimental attainment of this state [11] shows that it is possible for quantum superpositions to exist at a macroscopic level and makes it possible to investigate the process of their decoherence—the transformation into a statistical mixture in an interaction with the environment. The investigation of the process of decoherence of optical fields has become in the last few years one of the central problems of quantum optics. This is due to its fundamental significance for constructing a consistent quantum theory of measurement and the urgent need to eliminate decoherence in applications of quantum optics, such as quantum cryptography and the development of quantum computers (see review in [12]).

Thus, the analysis of the fundamental process of resonance fluorescence accompanying the excitation of an atom by a field in a superposition state is very interesting and topical. The existence of the interference part in the state of a field can result in fundamentally new behavior of an atom.

The main obstacle to realizing such an experiment is decoherence of the field state. To organize an atom-field interaction it is necessary either to use a stationary field source for exciting the atom or to place the atom in a cavity containing a field in a prescribed state. The first approach, without any modifications, is inapplicable for our case, since the superposition state in an open cavity undergoes rapid decoherence in connection with the escape of the field out of the cavity. The second approach for a field initially in a superposition state with $\varphi = \pi/2$—a positive Yurke–Stoler state (YSS) $|\alpha^+\rangle = (|\alpha\rangle + i|-\alpha\rangle)/\sqrt{2}$—has recently been examined in detail in [13]. In this case the field also undergoes rapid decoherence because of the interaction with the atom and in a time t_{decoh} it transforms into a statistical mixture. However, even for times $t \ll t_{\mathrm{decoh}}$ a number of interesting features can be observed in the behavior of the atom, for example, the suppression of Rabi oscillations of the dipole moment and exponential growth of the latter. Nonetheless, the problem of stationary resonance fluorescence excited by a quantum superposition remains unresolved.

1063-7761/00/9005-0733$20.00

In the present paper we propose one and possibly the only method for solving this problem [14]. The method is based on the first of the above-described approaches, which is modified taking into account the fact that coherent superpositions with $\varphi = \pm\pi/2$ (positive and negative YSS) in an open cavity can be protected from decoherence by the method, which we proposed recently in [15], of organizing electrooptic feedback between a detector continuously measuring the radiation escaping from the cavity and the field in the cavity. In Section 2 we shall show how this method makes it possible to produce a quasistationary source of a field in a YSS. Such a source can be used to excite resonance fluorescence of an individual atom; this is the content of Section 3. The general structure of the atom-field state will be analyzed in Section 4, and the quasistationary values of the atomic variables with the described method of excitation will be found in Section 5. The basic results of our investigation are discussed in the conclusions.

2. QUASISTATIONARY SOURCE OF SUPERPOSITION STATES

As is well known, the superposition state of a field in an open cavity is subject to rapid decoherence associated with losses on a half-transmitting mirror. This phenomenon is completely described by the master equation for the density matrix ρ_F of one cavity mode with line width κ:

$$\dot{\rho}_F = \frac{\kappa}{2}(2a\rho_F a^+ - a^+ a\rho_F - \rho_F a^+ a), \tag{1}$$

where a is an operator annihilating a photon of the given mode. The decoherence process can be represented most clearly by expanding the solution of equation (1) in an ensemble of quantum trajectories [16–18]:

$$\rho_F(t) = \sum_{n=0}^{\infty}\frac{1}{n!} \times \int_0^t dt_1 \ldots \int_0^t dt_n |\psi(t|t_1, \ldots, t_n)\rangle\langle\psi(t|t_1, \ldots, t_n)|. \tag{2}$$

Each trajectory $|\psi(t|t_1, \ldots, t_n)\rangle$ is, physically, a conditional unnormalized state of the field in a cavity, under the condition that the radiation escaping from the cavity is measured with a photodetector with 100% quantum efficiency and this detector detects precisely n counts at the times $t_1, t_2, \ldots, t_n$, and is determined by the expression

$$|\psi(t|t_1, \ldots, t_n)\rangle = \kappa^{n/2}\exp\left[-\frac{\kappa a^+ a(t-t_n)}{2}\right] \times a\exp\left[-\frac{\kappa a^+ a(t_n - t_{n-1})}{2}\right]a \ldots a\exp\left[-\frac{\kappa a^+ a t_1}{2}\right]|\psi(0)\rangle, \tag{3}$$

where $|\psi(0)\rangle$ is the initial state of the field. The expression (3) describes the continuous reduction of the field state in the process of quantum measurement consisting of two elementary processes: the operator $R = \sqrt{\kappa}\,a$ describes an abrupt reduction with the appearance of a count and the operator $W(t) = \exp(-\kappa a^+ a t/2)$ describes the continuous reduction between counts. When a count is present in the interval dt the increment of the density operator is $(d\rho_F)_1 = R\rho R^+ dt$ and in the absence of a count the increment is $(d\rho_F)_0 = dW\rho + \rho dW^+$, where dW is taken at $t = 0$ [19]. In the absence of information about detector counts, the unconditional evolution is given by the equation

$$d\rho_F = (d\rho_F)_1 + (d\rho_F)_0, \tag{4}$$

which is identical to equation (1).

Let us examine the dynamics of a field initially in a positive YSS $|\alpha_0^+\rangle = (|\alpha_0\rangle + i|-\alpha_0\rangle)/\sqrt{2}$. Using the identities $\exp(xa^+a)|\alpha\rangle = |\alpha e^x\rangle\exp\{-|\alpha|^2(1 - |e^{2x}|)/2\}$ and $a|\alpha\rangle = \alpha|\alpha\rangle$, which are satisfied for coherent states and automatically follow from the expansion of the coherent state in a Fock basis [17, 18], we find

$$\exp\left(-\frac{\kappa a^+ a\tau}{2}\right)\frac{1}{\sqrt{2}}(|\alpha\rangle \pm i|-\alpha\rangle) = \frac{1}{\sqrt{2}}\left[\left|\alpha\exp\left(-\frac{\kappa\tau}{2}\right)\right\rangle \pm i\left|-\alpha\exp\left(-\frac{\kappa\tau}{2}\right)\right\rangle\right] \times \exp\left\{-\frac{|\alpha|^2}{2}[1 - \exp(-\kappa\tau)]\right\}, \tag{5}$$

$$\sqrt{\kappa}a\frac{1}{\sqrt{2}}(|\alpha\rangle \pm i|-\alpha\rangle) = \sqrt{\frac{\kappa}{2}}\alpha(|\alpha\rangle \mp i|-\alpha\rangle), \tag{6}$$

i.e., the amplitude of the state will decrease exponentially between counts, $\alpha_t = \alpha_0\exp(-\kappa t/2)$, and at the moment of a count the state will change from a positive YSS $|\alpha_t^+\rangle = (|\alpha_t\rangle + i|-\alpha_t\rangle)/\sqrt{2}$ to a negative state $|\alpha_t^-\rangle = (|\alpha_t\rangle - i|-\alpha_t\rangle)/\sqrt{2}$ and vice versa. Thus, at the moment t the state of the field is either a positive or a negative YSS depending on whether the number of counts observed at the detector is even or odd. For sufficiently long times $t \gg t_{\text{decoh}}$ the even and odd number of counts are equally probable and averaging over the detector

indications gives the following density matrix of the field:

$$\rho_F(t) = \frac{1}{2}(|\alpha_t^+\rangle\langle\alpha_t^+| + |\alpha_t^-\rangle\langle\alpha_t^-|) = \frac{1}{2}(|\alpha_t\rangle\langle\alpha_t| + |-\alpha_t\rangle\langle-\alpha_t|). \quad (7)$$

The formula (7) describes the statistical mixture of states $|\alpha\rangle$ and $|-\alpha\rangle$ and does not contain interference terms characteristic of superposition. We can see that in this case decoherence appears as a result of a loss of information about the system [20]. The same result is obtained by solving equation (1) exactly. This gives an expression for the decoherence time $t_{\text{decoh}} = 1/(2\kappa|\alpha_0|^2)$ [9].

In the method which we developed to eliminate the above-described decoherence process [15], we proposed correcting the relative phase of the field state after each count via a feedback circuit. This procedure employs the fact that a negative YSS can be obtained from a positive YSS (and vice versa) by a simple phase shift of the field by π. This can be shown as follows. A phase shift of the coherent state by π can be described mathematically by the operator $\exp(i\pi a^+a)$, since $\exp(i\pi a^+a)|\alpha\rangle = |-\alpha\rangle$. From the physical standpoint, the operator $\exp(i\pi a^+a)$can be regarded as an evolution operator of the system under the action of the Hamiltonian $H = -\Delta\omega a^+a$, which contributes a detuning of the frequency in the time $t_{\text{int}} = \pi/\Delta\omega$, which should result in a phase shift by π. Applying the phase shift operator to a YSS gives

$$\exp(i\pi a^+a)\frac{1}{\sqrt{2}}(|\alpha_t\rangle \pm i|-\alpha_t\rangle) = \pm\frac{i}{\sqrt{2}}(|\alpha_t\rangle \mp i|-\alpha_t\rangle), \quad (8)$$

which is the required result. Thus, if another phase shift of the field by π, effectuated via a feedback circuit, follows after each count of the detector, then the state of the field in the cavity will remain a positive YSS with a decaying amplitude.

To describe such a feedback in the master-equation formalism (neglecting the time delay in the feedback circuit) the operator R must be replaced by the operator $\tilde{R} = \exp(i\pi a^+a)\sqrt{\kappa}\,a$, describing reduction and the phase shift which follow one another [15]. Feedback does not operate between counts, and therefore the operator $W(t)$ is not subject to modification. The evolution of the conventional field state will now be described by the following quantum trajectory:

$$|\psi(t|t_1, \ldots, t_n)\rangle = \kappa^{n/2}\exp\left[-\frac{\kappa a^+a(t-t_n)}{2}\right] \times \exp(i\pi a^+a)a\exp\left[-\frac{\kappa a^+a(t_n-t_{n-1})}{2}\right] \times \exp(i\pi a^+a)a\ldots\exp(i\pi a^+a)a\exp\left(-\frac{\kappa a^+at_1}{2}\right)|\alpha_0^+\rangle. \quad (9)$$

Using the relations (5), (6), and (8) it is easy to show that the state (9) is $|\alpha_t^+\rangle$ to within a numerical factor. Constructing an equation similar to equation (4) we find that the unconditional evolution is given by the equation

$$\dot{\rho}_F = \frac{\kappa}{2}[2\exp(i\pi a^+a)a\rho_F a^+\exp(-i\pi a^+a) - a^+a\rho_F - \rho_F a^+a] \equiv L_0\rho_F, \quad (10)$$

which is a modification of equation (1) for the case where feedback is present.

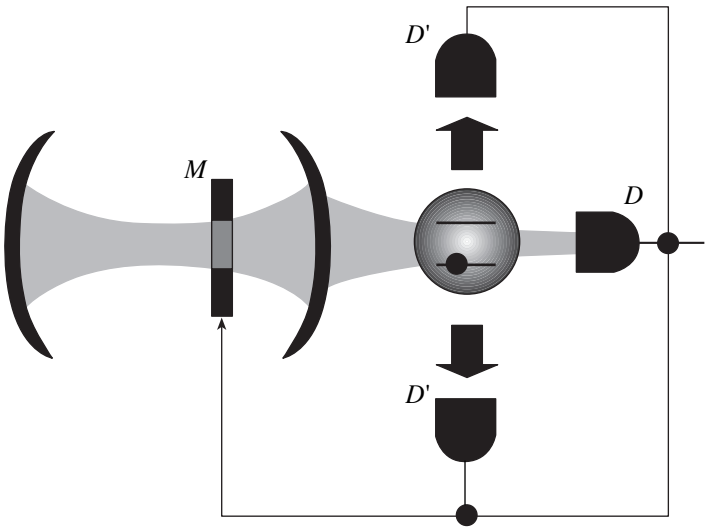
In summary, by preparing a YSS in a cavity and organizing feedback it is possible to produce a quasistationary source of such a state. On account of energy losses, the field state decays in a time $t_{\text{decay}} = \kappa^{-1}$, but this time interval may be sufficient for achieving a stationary state for the interaction of a field with other systems.

3. RESONANCE FLUORESCENCE IN A FEEDBACK LOOP

The above-described source of a superposition state can be used to excite the resonance fluorescence of an individual atom. In order that the field–atom interaction not destroy the structure of the field state, it is necessary to employ unidirectional action of the cavity field on the atom, such that the atom is irradiated with light leaking out of the cavity while the light scattered by the atom does not return into the cavity. In the scheme proposed the atom is placed between the cavity and a detector, which measures the outgoing radiation; we designate this detector as D. To follow the phase of the field in the cavity it is necessary to have complete control of the photons leaving the cavity, so that all of the light scattered by the atom, including the backscattered light, must be collected on photodetectors included into the same feedback circuit. We shall designate these detectors as D' (figure). Just as in the preceding section the feedback circuit shifts the phase of the field in the cavity by π after a count is recorded at any of these detectors.

A systematic quantum description of a unidirectional action of one quantum system on another has been developed in [21]. Application of this approach to the excitation of resonance fluorescence of an isolated atom has been studied in [22]. We shall examine first the basic processes occurring in this scheme in the absence of feedback. When the field is in resonance with a transition between the ground $|1\rangle$ and excited $|2\rangle$ states of an atom, the dynamics of the field in the cavity and of the atom is described by the following master equation for the density operator ρ of the atom–field system [22]:

$$\dot{\rho} = \frac{\kappa}{2}(2a\rho a^+ - a^+a\rho - \rho a^+a) + \frac{\gamma+\gamma'}{2}(2\sigma_-\rho\sigma_+ - \sigma_+\sigma_-\rho - \rho\sigma_+\sigma_-) + \sqrt{\kappa\gamma}(\sigma_-\rho a^+ + a\rho\sigma_+ - \rho a^+\sigma_- - a\sigma_+\rho), \quad (11)$$



The arrangement of the experimental apparatus for observing the resonance fluorescence of an individual atom excited by a quantum superposition. The field in the cavity is initially prepared in a superposition state. The radiation leaking out of the cavity excites the atom, which is in the ground state initially. The transmitted radiation is collected at the detector *D*, while the radiation scattered by the atom is collected at the detectors *D*'. The electric feedback circuit shifts, by means of a modulator *M*, the phase of the field in the cavity by π after each count on one of the detectors.

where γ and γ' are the rates of spontaneous emission of an atom into the solid angles covered by the detectors *D* and *D*', respectively; we shall designate the total rate of spontaneous emission as $\Gamma = \gamma + \gamma'$. The operators $\sigma_- = |1\rangle\langle 2|$ and $\sigma_+ = |2\rangle\langle 1|$ are lowering and raising operators of the atom, respectively.

The first term in equation (11) describes the decay of the field through a half-transmitting mirror, the second term describes spontaneous emission of the atom, and the third term describes the unidirectional action of the field on the atom, as described in the Kolobov–Sokolov approach [21]. In this description the parameter $\sqrt{\kappa\gamma}$ plays the role of the field–atom coupling constant. Averaging equation (11) over the states of the atom leads to equation (1) for the field, whose dynamics therefore does not depend on the atom, as should be for describing a unidirectional action of a field on an atom.

The solution of equation (11) can be expanded in an ensemble of quantum trajectories, which have the meaning of a conventional state of the field–atom system for given successive counts at the detectors *D* and *D*'. Each trajectory is formed by a combination of three elementary processes: reduction of the state when a count is recorded at the detector *D*, as described by the operator $R_1 = \sqrt{\kappa}\, a + \sqrt{\gamma}\,\sigma_-$; reduction of the state when a count is recorded at the detector *D*', described by the operator $R_2 = \sqrt{\gamma'}\,\sigma_-$; and, continuous evolution between counts, described by the operator

$$V(t) = \exp\{-(\kappa a^+ a + \Gamma\sigma_+\sigma_- + 2\sqrt{\kappa\gamma}\, a\sigma_+)t/2\}. \quad (12)$$

For example, for one photon at the detector *D* at the moment t_1 and one photon at the detector at *D*' at the moment $t_2 > t_1$, the trajectory will have the form

$$|\psi(t|t_1, t_2)\rangle = V(t - t_2)R_2V(t_2 - t_1)R_1V(t_1)|\psi(0)\rangle. \quad (13)$$

The equation obtained for the unconditional evolution of the density operator by averaging over such trajectories

$$d\rho = R_1\rho R_1^+ dt + R_2\rho R_2^+ dt + dV\rho + \rho dV^+, \quad (14)$$

is identical to equation (11). The equation (11) describes very well the known properties of resonance fluorescence—the triplet structure of its spectrum and the antibunching of scattered photons [22].

When feedback is present, a phase shift of the field in the cavity by π, described by the operator $\exp(i\pi a^+ a)$, will follow after each count. Just as in the preceding section, when the delay in the feedback circuit is neglected, the action of the feedback can be taken into account by replacing the operators R_1 and R_2 by the operators $\tilde{R}_1 = \exp(i\pi a^+ a)R_1$ and $\tilde{R}_2 = \exp(i\pi a^+ a)R_2$, which describe the reduction and phase shift which follow one another. The unconditional evolution of the atom–field system in the presence of feedback is described by the equation

$$d\rho = \tilde{R}_1\rho\tilde{R}_1^+ dt + \tilde{R}_2\rho\tilde{R}_2^+ dt + dV\rho + \rho dV^+, \quad (15)$$

or

$$\dot{\rho} = \frac{\kappa}{2}[2\exp(i\pi a^+ a)a\rho a^+ \exp(-i\pi a^+ a)$$

$$- a^+ a\rho - \rho a^+ a] + \frac{\gamma + \gamma'}{2}[2\exp(i\pi a^+ a)\sigma_-\rho\sigma_+$$

$$\times \exp(-i\pi a^+ a) - \sigma_+\sigma_-\rho - \rho\sigma_+\sigma_-]$$
$$+ \sqrt{\kappa\gamma}[\exp(i\pi a^+ a)\sigma_-\rho a^+ \exp(-i\pi a^+ a)$$
$$+ \exp(i\pi a^+ a) a\rho\sigma_+ \exp(-i\pi a^+ a)$$
$$- \rho a^+\sigma_- - a\sigma_+\rho] \equiv L\rho. \quad (16)$$

Finding the solution of equation (16) in the general case is a quite difficult problem. In the next section we shall find an analytic solution for the case where the initial state of the field is a positive YSS, while the atom is in the ground state.

4. ENTANGLEMENT OF ATOMIC AND FIELD STATES

The above-described scheme for exciting the resonance fluorescence of an isolated atom in the absence of losses possesses a definite internal symmetry. As a result, the solution of equation (16) with the initial condition

$$\rho(0) = |\alpha_0^+\rangle\langle\alpha_0^+| \otimes |1\rangle\langle 1| \quad (17)$$

can be immediately sought in the form

$$\rho(t) = C_{11}(t)|\alpha_t^+\rangle\langle\alpha_t^+| \otimes |1\rangle\langle 1|$$
$$+ C_{22}(t)|\alpha_t^-\rangle\langle\alpha_t^-| \otimes |2\rangle\langle 2| + C_{21}(t)|\alpha_t^-\rangle\langle\alpha_t^+| \otimes |2\rangle\langle 1| \quad (18)$$
$$+ C_{12}(t)|\alpha_t^+\rangle\langle\alpha_t^-| \otimes |1\rangle\langle 2|,$$

where $C_{ij}(t)$ are unknown functions, $C_{11} + C_{22} = 1$, and $C_{21} = C_{12}^*$.

The structure of equation (18) signifies a strong correlation or "entanglement" between the states $|\alpha_t^+\rangle$ and $|1\rangle$ on the one hand and the states $|\alpha_t^-\rangle$ and $|2\rangle$ on the other. The relation (18) can be proved by examining the transformation of a state of the system in the presence of three elementary processes described by the operators $\tilde{R}_1$, $\tilde{R}_2$, and $V(t)$.

Let us examine first the evolution operator $V(t)$ (12) between counts. It can be rewritten in the form

$$V(t) = \exp\left(-\frac{\kappa a^+ a t}{2}\right)\exp\left(-\frac{\Gamma\sigma_+\sigma_- t}{2}\right)$$
$$\times\left\{1 - a\sigma_+\sqrt{\kappa\gamma}\int_0^t \exp\left[-\frac{(\Gamma-\kappa)\tau}{2}\right]d\tau\right\}. \quad (19)$$

The operator $a\sigma_+$ converts the state $|\alpha_t^+\rangle|1\rangle$ into $|\alpha_t^-\rangle|2\rangle$ in accordance with equation (6). The operator $\sigma_+\sigma_-$ is diagonal in the basis $|1\rangle$ and $|2\rangle$, and therefore it changes only the numerical factors in front of these vectors. The operator $\exp\{-\kappa a^+ a t/2\}$ leads to exponential decay of the field amplitude, $\alpha_{t+\tau} = \alpha_t \exp(-\kappa\tau/2)$, in accordance with equation (5). Therefore the operator (19) converts the state $|\alpha_t^+\rangle|1\rangle$ into a state of the form

$$c_1(\tau)|\alpha_{t+\tau}^+\rangle|1\rangle + c_2(\tau)|\alpha_{t+\tau}^-\rangle|2\rangle, \quad (20)$$

where $c_1(\tau)$ and $c_2(\tau)$ are functions of time. It can be shown similarly that the operator (19) converts the state $|\alpha_t^-\rangle|2\rangle$ into the state $|\alpha_{t+\tau}^-\rangle|2\rangle$ to within a numerical factor.

On the basis of the relations (6) and (8) it can be concluded that the operators $\tilde{R}_1$ and $\tilde{R}_2$ do not change the structure of the state (20) and influence only the numerical factors.

In summary, we find that a quantum trajectory, i.e., the conventional state of the system at the moment t, under the condition that a sequence of counts is observed at the detector D at the moments $t_1, t_2, \ldots, t_n$ and a sequence of counts is observed at the detector D' at the moments $t_1', t_2', \ldots, t_m'$, has the form

$$|\psi(t|t_1, \ldots, t_n; t_1', \ldots, t_m')\rangle$$
$$= c_1(t|t_1, \ldots, t_n; t_1', \ldots, t_m')|\alpha_t^+\rangle|1\rangle \quad (21)$$
$$+ c_2(t|t_1, \ldots, t_n; t_1', \ldots, t_m')|\alpha_t^-\rangle|2\rangle.$$

An unconditional state can be obtained by averaging over all possible indications of the detectors similarly to the expression (2) with the difference that the integration and summation must be performed over the counts of both detectors. Such averaging leads to equation (18).

The mathematical relations presented above have a clear physical explanation. When a photon emerges from the cavity, the field state changes abruptly from a positive to a negative YSS. Next, the photon enters one of the detectors, irrespective of whether or not it has been scattered by an atom, and the feedback circuit, shifting the phase of the field by π, restores a positive YSS in the cavity. When a photon is scattered by an atom, the photon is delayed by a time of the order Γ^{-1}; equation (16) takes this delay into account. During this time the atom is in an excited state, and the field is in a negative YSS. Thus, the conventional atom–field state of the system has the form (21) for any number of counts.

The entanglement of the atomic and field states in the above-described excitation method is itself a remarkable phenomenon, which could be employed in various applications of quantum information theory—quantum computation, quantum teleportation, and quantum cryptography (see the review in [12]). We note that for the conventional excitation of resonance fluorescence using coherent laser light, the states of the

field and atom remain statistically independent, which makes the semiclassical description applicable.

5. QUASISTATIONARY REGIME

To find the equations for the coefficients $C_{ij}(t)$ the expression (18) must be substituted in equation (16). We find the temporal evolution of the field vectors from equation (5)

$$|\alpha_t^{\pm}\rangle = \exp\left(-\frac{\kappa a^{+} a t}{2}\right)|\alpha_0^{\pm}\rangle \times \exp\left\{\frac{1}{2}|\alpha_0|^2[1-\exp(-\kappa t)]\right\}, \quad (22)$$

whence direct differentiation using equation (6) gives

$$\frac{d}{dt}|\alpha_t^{\pm}\rangle\langle\alpha_t^{\pm}| = L_0|\alpha_t^{\pm}\rangle\langle\alpha_t^{\pm}|, \quad (23)$$

$$\frac{d}{dt}|\alpha_t^{\pm}\rangle\langle\alpha_t^{\mp}| = (L_0 + 2\kappa|\alpha_t|^2)|\alpha_t^{\pm}\rangle\langle\alpha_t^{\mp}|, \quad (24)$$

where L_0 is the Liouville field evolution operator determined by equation (10).

Now, substituting the expression (18) into equation (16) and equating the coefficients in front of the identical atomic matrix elements gives the equations

$$\dot{C}_{22} = -\Gamma C_{22} - \sqrt{\kappa\gamma}(\alpha_t C_{12} + \alpha_t^* C_{21}), \quad (25)$$

$$\dot{C}_{21} = -\left(\frac{\Gamma}{2} + 2\kappa|\alpha_t|^2\right)C_{21} - \sqrt{\kappa\gamma}\alpha_t(C_{11} + C_{22}). \quad (26)$$

The physical meaning of the coefficients C_{22} and C_{21} is determined by the relations $\langle\sigma_+(t)\sigma_-(t)\rangle = C_{22}(t)$ and $\langle\sigma_-(t)\rangle = C_{21}(t)\exp\{-2|\alpha_t|^2\}$, i.e., C_{22} is the population of the upper level of the atom and C_{21} is, to within a factor, the dipole moment. Thus equations (25) and (26) are the analogs of the conventional Bloch equations for the atomic variables, which hold for excitation of resonance fluorescence by coherent light. The equation (25) is identical to the corresponding Bloch equation, and equation (26) differs from the corresponding equation in two respects. In the first place, the relaxation rate of the dipole moment includes an additional term proportional to the field intensity. In the second place, this equation contains not the difference of the populations of the upper and lower levels but rather their sum, which is equal to 1. In consequence, equations (25) and (26) do not have oscillating solutions, i.e., the so-called Rabi oscillations do not occur in the above-described method of excitation.

In what follows we shall be interested primarily in the case $\kappa \ll \Gamma$, i.e., a regime in which the field essentially does not change in times required for the atom to reach a stationary state. In such a regime the time dependence of the field amplitude α_t in equations (20) and (26) can be neglected, and these equations can be solved with a fixed amplitude α. For the initial condition (17), i.e., $C_{22}(0) = C_{21}(0) = 0$, the solutions of equations (25) and (26) have the form

$$C_{21}(t) = C_{21}^{ss}\left\{1 - \exp\left[-\left(\frac{\Gamma}{2} + 2\kappa|\alpha|^2\right)t\right]\right\}, \quad (27)$$

$$C_{22}(t) = C_{22}^{ss}\left[1 - \exp(-\Gamma t) - \frac{2\Gamma}{\Gamma - 4\kappa|\alpha|^2} \times \left\{\exp\left[-\left(\frac{\Gamma}{2} + 2\kappa|\alpha|^2\right)t\right] - \exp(-\Gamma t)\right\}\right], \quad (28)$$

where the quasistationary values of C_{21}^{ss} and C_{22}^{ss} are given by the expressions

$$C_{21}^{ss} = -\frac{\sqrt{\kappa\gamma}\alpha}{\Gamma/2 + 2\kappa|\alpha|^2}, \quad (29)$$

$$C_{22}^{ss} = \frac{\gamma}{2}\frac{4\kappa|\alpha|^2}{\Gamma + 4\kappa|\alpha|^2}. \quad (30)$$

For $4\kappa|\alpha|^2 = \Gamma$ equation (28) assumes the form

$$C_{22}(t) = \frac{\gamma}{2\Gamma}[1 - e^{-\Gamma t} - \Gamma t e^{-\Gamma t}]. \quad (31)$$

Using equations (18), (25), and (26) it is also possible to calculate the spectrum of the scattered radiation, which on account of the absence of oscillating solutions of equations (25) and (26) will possess a singlet and not a triplet structure, characteristic for the resonance fluorescence spectrum under excitation by a sufficiently intense coherent field. However, measurement of the fluorescence spectrum in the scheme under consideration is problematic, since the measurement procedure itself can interfere with the use of scattered light in the feedback circuit. The correlation function $G^{(2)}(\tau)$ of the intensity of the scattered light is of much greater practical interest. This function can be obtained by analyzing the photocurrent of the detector D' without additionally complicating the measurement scheme. This function is determined in terms of the stationary density matrix ρ_{ss} as follows:

$$G^{(2)}(\tau) = \mathrm{Tr}\{\sigma_+\sigma_- e^{L\tau}\sigma_-\rho_{ss}\sigma_+\} = \mathrm{Tr}\{|2\rangle\langle 2|e^{L\tau}C_{22}^{ss}|\alpha_t^-\rangle\langle\alpha_t^-| \otimes |1\rangle\langle 1|\} = C_{22}^{ss}C_{22}(\tau), \quad (32)$$

where L is the Liouville operator determined by equation (16). The derivation of equation (32) took account of the fact that equations (17)–(31) are invariant under the substitution $|\alpha_t^{\pm}\rangle \longrightarrow |\alpha_t^{\mp}\rangle$, i.e., the solution of equation (16) with the initial condition $|\alpha_0^-\rangle|1\rangle$ has the

form (18) with the positive YSS replaced by a negative YSS and vice versa.

The equations (32) and (28) show that the intensity correlation function reaches a stationary value in times of the order of Γ^{-1} with no oscillations for any values of the Rabi frequency $\Omega = \sqrt{\kappa\gamma}\,\alpha$. This behavior differs strongly from the case of resonance fluorescence excited by a coherent state, when for sufficiently large values of $\Omega > \Gamma/4$ the intensity correlation function of the scattered light oscillates with frequency Ω.

6. DISCUSSION

In the present paper, the resonance fluorescence of an individual atom excited by a field in a superposition state was examined, and the laws of the behavior of an atom in a quasistationary regime were found for the first time. The relations obtained lead to the conclusion that certain familiar properties of resonance fluorescence, specifically, oscillations of the dipole moment and of the populations of the atomic levels, are related with the use of a coherent or a nearly coherent state of the field for excitation of the atom. It was found that when a superposition of two coherent states is used for this purpose, the oscillations of the dipole moment and populations do not occur for arbitrarily large values of the field intensity. This can be observed by measuring the spectrum of the intensity of the scattered light. It should be noted that the absence of oscillations of the dipole moment was predicted in [13], where the interaction of an isolated atom in a cavity with the field of this cavity, initially in a superposition state, was examined. The populations of the atomic levels in this problem undergo the familiar Rabi oscillations for sufficiently high field intensities. However, the field state in this approach is destroyed because of induced emission of the atom into a cavity mode at precisely the rate determined by the Rabi frequency. Thus, the present paper extends [13] to the case of a quasistationary regime.

The absence of Rabi oscillations in the excitation method described above can be explained quite simply. According to equation (26), the rate of decay of the off-diagonal density matrix element acquires an additional term $2\kappa|\alpha|^2$. This quantity is the photon flux incident on the atom; it is identical to the decoherence rate of the superposition state interacting with the environment. At the same time it is well known [23] that as the time T_2 (the coherence time of the off-diagonal element) decreases, for example, as a result of collisional dephasing, Rabi oscillations are suppressed for a given value of the field amplitude, though as the latter increases, oscillations could be restored. In our case the dephasing rate increases quadratically with increasing field amplitude, while the Rabi frequency is related with the latter linearly, so that for any values of the field intensity the oscillations of the atomic variables are overdamped. The described "dephasing via the field" can be related with the property of a superposition state and with the property of an excitation scheme employing feedback, since $2\kappa|\alpha|^2$ is the rate of change of the phase of the cavity field (the number of switchings per second). This duality is due to the fact that it is the Yurke–Stoler superposition state that is the "eigenstate" of such a feedback loop [15].

The scheme, studied in the present paper, for resonance fluorescence of an individual atom is in our view an interesting "Gedanken experiment," making it possible to understand more deeply the quantum aspects of the interaction of light and matter. At the same time it seems to us that the predictions of the theory can be checked experimentally now or in the near future. To assess this possibility we shall find the limits of applicability of the two basic approximations on which the above description is based: the unit quantum efficiency of detectors and the absence of time delay in the feedback circuit. As shown in [15], for quantum efficiency of a detector in the feedback loop $\eta < 1$ the YSS undergoes decoherence at the rate $2\kappa(1-\eta)|\alpha|^2$, i.e., decoherence can be neglected for $\Gamma \gg 2\kappa(1-\eta)|\alpha|^2$, which is achievable for modern detectors and high-Q microwave and optical cavities. Thus, for $\Gamma \sim 10^8$ Hz, $\kappa \sim 10^6$ Hz, $\eta \sim 0.9$ (optical range) or $\kappa \sim 10^4$ Hz, $\eta \sim 0.1$ (microwave range [11]) the inequality presented above holds well for an average number of photons $|\alpha|^2 \leq 100$. The delay in the feedback circuit must be less than the average time between two successive counts, i.e., $\kappa^{-1}|\alpha|^{-2}$, which is hundreds of nanoseconds for optical cavities with $\kappa \sim 10^6$ Hz and mesoscopic superpositions with $|\alpha|^2 \sim 10$, while the delay in the electric circuit usually does not exceed several tens of nanoseconds. This relation is even more optimistic in the microwave range. At the present time the rapid development of techniques for working with individual atoms and "brittle" quantum field states leaves no doubt that very soon it will be possible to investigate experimentally the characteristic features of resonance fluorescence of an individual atom excited by nonclassial light.

ACKNOWLEDGMENTS

We are grateful to the INTAS fund for supporting this investigation (project no. 96-0167).

REFERENCES

1. V. Weisskopf, Ann. Phys. (Leipzig) **9**, 23 (1931).
2. S. G. Rautian and I. I. Sobel'man, Zh. Éksp. Teor. Fiz. **41**, 456 (1961) [Sov. Phys. JETP **14**, 328 (1961)]; P. A. Apanasevich, Opt. Spektrosk. **7**, 64 (1963); B. R. Mollow, Phys. Rev. **188**, 1969 (1969).
3. F. Schuda, C. R. Stroud, Jr., and M. Hercher, J. Phys. B **7**, L198 (1974); F. Y. Wu, R. E. Grove, and S. Ezekiel, Phys. Rev. Lett. **35**, 1426 (1975).

4. H. J. Carmichael and D. F. Walls, J. Phys. B **9**, 1199 (1976); H. J. Kimble and L. Mandel, Phys. Rev. A **13**, 2123 (1976).
5. H. J. Kimble, M. Dagenais, and L. Mandel, Phys. Rev. Lett. **39**, 691 (1977).
6. P. A. Apanasevich and S. Ya. Kilin, Phys. Lett. A **62**, 83 (1977).
7. A. Aspect, G. Roger, S. Reynaud, *et al.*, Phys. Rev. Lett. **45**, 617 (1980).
8. C. W. Gardiner, Phys. Rev. Lett. **56**, 1917 (1986).
9. V. Buzek and P. L. Knight, Prog. Opt. **34**, 1 (1995).
10. E. Schrödinger, Naturwissenschaften **23**, 807 (1935); **23**, 823 (1935); **23**, 844 (1935).
11. M. Brune, E. Hagley, J. Dreyer, *et al.*, Phys. Rev. Lett. **77**, 4887 (1996).
12. S. Ya. Kilin, Usp. Fiz. Nauk **169**, 507 (1999).
13. S. Ya. Kilin and V. N. Shatokhin, Phys. Rev. Lett. **76**, 1051 (1996).
14. D. B. Horoshko and S. Ya. Kilin, in *ICONO'98 Technical Digest, 1998*, p. 39; in *EQEC'98 Technical Digest, 1998*, p. 68.
15. D. B. Horoshko and S. Ya. Kilin, Phys. Rev. Lett. **78**, 840 (1997).
16. M. D. Srinivas and E. B. Davies, Opt. Acta **28**, 981 (1981).
17. S. Ya. Kilin, *Quantum Optics: Fields and Their Detection* (Navuka i Tékhnika, Minsk, 1990).
18. H. Carmichael, *An Open System Approach to Quantum Optics* (Springer, Berlin, 1993).
19. M. Ueda, N. Imoto, and T. Ogawa, Phys. Rev. A **41**, 3891 (1990); D. B. Khoroshko and S. Ya. Kilin, Opt. Spektrosk. **82**, 913 (1997) [Opt. Spectrosc. **82**, 838 (1997)].
20. D. B. Horoshko and S. Ya. Kilin, Opt. Express **2**, 347 (1998).
21. M. I. Kolobov and I. V. Sokolov, Opt. Spektrosk. **62**, 112 (1987) [Opt. Spectrosc. **62**, 69 (1987)].
22. C. W. Gardiner, Phys. Rev. Lett. **70**, 2269 (1993); H. J. Carmichael, Phys. Rev. Lett. **70**, 2273 (1993).
23. W. Vogel and D.-G. Welsch, in *Lectures on Quantum Optics* (Akademie, Berlin, 1994), p. 362.

Translation was provided by AIP

Journal of Experimental and Theoretical Physics, Vol. 90, No. 5, 2000, pp. 741–752.
Translated from Zhurnal Éksperimental'noĭ i Teoreticheskoĭ Fiziki, Vol. 117, No. 5, 2000, pp. 853–866.

ATOMS, SPECTRA, RADIATION

Electron–Ion Collision Integral in a Strong Magnetic Field

S. A. Koryagin

Institute of Applied Physics, Russian Academy of Sciences, Nizhni Novgorod, 603600 Russia
e-mail: koryagin@appl.sci-nnov.ru

Received October 6, 1999

Abstract—Boltzmann's collision integral is extended to the case of helical (Larmor) particle trajectories in a magnetic field of arbitrary strength. The main characteristics of collisions of electrons with positively charged ions in strong magnetic fields, where the Larmor radius of electrons becomes less than the characteristic impact parameter of close collisions in the absence of a magnetic field (Landau's parameter), are investigated. The differential scattering cross section and the corresponding electron–ion collision integral in strong fields are found. The transport collision frequencies are calculated, and the results are compared with the similar quantities for weaker magnetic fields. © *2000 MAIK "Nauka/Interperiodica".*

1. INTRODUCTION

It is well known that strong magnetic fields substantially change the state of atoms and molecules (see the review [1]) and the collisions between charged particles.[1] For example, the ground state of a stationary hydrogen atom contracts in a direction across the magnetic field and acquires the form of a needle oriented parallel to the field, when the energy interval between the Landau levels exceeds 13.6 eV. This relationship obtains in magnetic fields $B \gtrsim 2.35 \times 10^9$ G. An approximate calculation of the ground-state energy can be found in [3]. In [4, 5] a numerical method was presented for performing an accurate calculation of the energy levels and wave functions of the hydrogen atom in arbitrary fields. The translational motion of an atom across the field results in further transformation of the electron cloud (see, for example, [6] and the references cited there), since a uniform electric field, which shifts the electronic cloud relative to the nucleus, appears in the coordinate system of the atom.

For free states of an electron (as well as the Rydberg states and excitons in semiconductors) the effect of a magnetic field is substantially manifested in weaker fields. In the classical method of describing collisions of an electron moving with a definite velocity v and a stationary ion with charge number Z, the magnetic field becomes strong when the electron cyclotron period $2\pi/\omega_B$ is less than the duration r_s/v of close collisions in the absence of a field. Here

$$r_s = Ze^2/mv^2$$

is equal to the distance r at which the electron kinetic energy $mv^2/2$ is of the order of its potential energy Ze^2/r in the field of an ion;

$$\omega_B = eB/mc$$

is the electron cyclotron frequency; m and $e > 0$ are the electron mass and charge; and, c is the speed of light. The strong-field condition

$$Z\frac{B}{B_{\rm cl}}\left(\frac{v}{c}\right)^{-3} \gg 1, \qquad (1)$$

where the constant

$$B_{\rm cl} = m^2c^4/e^3 = 6.05 \times 10^{15}\ \text{G},$$

depends on the particle velocity v. For this reason, fast and slow particles for which a given field is weak or strong, respectively, are singled out in a fixed magnetic field.

In weak magnetic fields, where the condition opposite to (1) is satisfied, collisions with impact parameters p_r less than the Larmor radius $r_B(v) = v/\omega_B$ occur just as without a magnetic field: the well-known hyperbolic trajectory connects the helical lines of Larmor rotation of an electron before and after a collision. For distant collisions with impact parameters $p_r \gtrsim r_B$ the condition for the magnetic drift approximation are satisfied and the pitch angle of the electron remains essentially unchanged. Consequently, the Larmor radius becomes the natural maximum impact parameter $p_{\max}$ in the Coulomb logarithm $L = \ln(p_{\max}/r_s)$ for collision frequencies that characterize the rate of transformation of the energy of motion of an electron in a direction parallel and transverse to the magnetic field.

A distinguishing feature of investigations of collisions in weak fields is the use of an approximate method in which the action of the Coulomb center is treated as a weak perturbation to the initial cyclotron rotation of an electron and the main transport frequencies are proportional to the squared charges of the colliding particles. In [7] this method was used to calculate the relaxation times of an electron gas in a magnetic field with an azimuthally symmetric particle velocity

[1] New results concerning the configuration of the ground state of the hydrogen molecule in a strong magnetic field were presented in [2].

distribution. In [8] the method was extended to plasma with an arbitrary particle distribution function. In [9] the pair-collision integral was obtained for an ideal magnetized plasma where the maximum impact parameter in the form of the Debye radius r_D is introduced to take account of collective effects (electrostatic screening). In [10, 11] the transport collision frequencies, appearing in the high-frequency permittivity of a plasma, as well as the rate of energy transfer between the electronic and ionic components in plasma (see also [12]) were investigated. In [13] the Lenard–Balescu collision integral was extended to the case of a magnetized plasma. In [14–16] the changes in the parameters of test particles in collisions with plasma particles were calculated.

The quantum calculation of Coulomb collisions in weak fields is based on the Born approximation, where the particle velocity is quite high and the distance r_s is shorter than the de Broglie wavelength $\lambda_B = \hbar/mv$. The quantum-mechanical transition probabilities between Landau levels in Coulomb collisions were calculated in [17] and then in [18] in a simpler form. A relativistic quantum-mechanical calculation of the transition probabilities was performed in [19] as well as in [20]. The corresponding transport coefficients in a quantized plasma were investigated in [21] and [22, 23]. Expressions for the bremsstrahlung absorption coefficients off the cyclotron line can be found in [24].

In strong magnetic fields (1) the character of the collisions changes as compared with a zero magnetic field at all impact parameters. In this case, the collisions of an electron with positively and negatively charged particles exhibit qualitatively different properties. In an effective collision the electron velocity near the scattering center is substantially different from the initial velocity. This characteristic property of effective collisions precludes the use of perturbation methods, which are used for weaker fields, in this case.

The case of strong magnetic fields (1) was realized in [25] under laboratory conditions in an experiment with an electron plasma. In [26] the basic properties of interelectronic collisions were investigated and a corresponding collision integral was obtained. As a result of electron–electron collisions over a time equal to the stopping time of a thermal electron in a unmagnetized plasma, a particle distribution of the form

$$f(\mathbf{v}) = f_{\parallel}(v_{\parallel}) f_{\perp}(v_{\perp}) \tag{2}$$

is established. Such rapid relaxation is due to head-on collisions of electrons with close velocities, in which the particles exchange the given components of the velocity, along a magnetic field. However, isotropization of the distribution and transformation of energy between the electronic degrees of freedom in a direction parallel and transverse to the magnetic field are substantially slower, since as a result of collisions the particles approach only to distances of the order of $r_s \gg r_B$ and the total transverse electron energy is conserved, to a high degree of accuracy, in collisions [27, 28].

Under astrophysical conditions the case of strong fields is characteristic for photospheres of magnetic white dwarfs, where the magnetic fields attain magnitudes of 10^9 G and the temperature $T \sim 10^4$ K is not too high. A qualitative analysis of electron-proton collisions [29, 30] has shown that a large change in the electron pitch angle occurs primarily in collisions with impact parameters $p_h \lesssim p_{\mathrm{cr}}$, where

$$p_{\mathrm{cr}} = \left(\frac{mc^2}{B^2}\right)^{1/3} = r_e\left(\frac{B}{B_{\mathrm{cl}}}\right)^{-2/3}, \text{ and}$$
$$r_e = e^2/mc^2 = 2.82 \times 10^{-13} \text{ cm} \tag{3}$$

is the classical electron radius. In these collisions the characteristic time over which the Coulomb field changes during the passage of an electron near an ion is less than the cyclotron period. The distance p_{cr} is on the one hand less than the distance r_s and on the other hand greater than the Larmor radius. For effective collisions the particles move along quasi-bound trajectories, i.e., the direction of the electron velocity along the magnetic field changes repeatedly during a collision, and a proton plays the role of an "electrostatic" trap.

For an ideal plasma the distance $r_s \gg p_{\mathrm{cr}}$ is much shorter than the distance between particles, so that effective collisions can be treated as pair collisions. In Section 2 the general form of the pair-collision integral in a magnetic field is found in Boltzmann's form for an arbitrary electron distribution function (including a function which is asymmetric with respect to the directions of the velocities transverse to the magnetic field). For this, the corresponding impact parameters are determined for helical (Larmor) particle trajectories in a magnetic field of arbitrary magnitude. In Section 3 a logarithmically accurate value of the parameter p_{cr} is found and the role of distant collisions $p_h \gtrsim p_{\mathrm{cr}}$ in the change in the parameters of an electron in strong magnetic fields is investigated. In Section 4 the differential scattering cross section and the collision integral for an electron colliding with positive ions in a strong magnetic field (1) are found. The basic transport collision frequencies of an electron with positive and negative ions in strong and weak fields are compared in the conclusions.

2. COLLISION INTEGRAL IN BOLTZMANN'S FORM IN A MAGNETIC FIELD

We shall consider a spatially uniform magnetized plasma consisting of electrons and stationary ions of one kind with density n_i. To obtain the collision integral we shall employ the well-known kinetic method, which is described in, for example, [31] for the case of an isotropic medium. We shall consider electron–ion and electron–electron collision integrals

$$I_{st}^{(ei)}(\mathbf{v}) = \frac{1}{m}\int \frac{\partial U_{ei}}{\partial \mathbf{r}} \frac{\partial f_{ei}}{\partial \mathbf{v}}\bigg|_{t' = t} d\mathbf{r}, \tag{4}$$

$$I_{st}^{(ee)}(\mathbf{v}) = \frac{2}{m}\int \frac{\partial U_{ee}}{\partial \mathbf{r}}\frac{\partial f_{ee}}{\partial \mathbf{v}_{\rm rel}}\bigg|_{\substack{t'=t \\ \mathbf{v}_{cm}=\mathbf{v}-\mathbf{v}_{\rm rel}/2}} d\mathbf{v}_{\rm rel}d\mathbf{r}, \qquad (5)$$

where $U_{ei}(\mathbf{r})$ and $U_{ee}(\mathbf{r})$ are, respectively, the electron–ion and electron–electron interaction potentials and the functions $f_{ee}(\mathbf{r}, \mathbf{v}_{\rm rel}, \mathbf{v}_{\rm cm}, t'; t)$ and $f_{ei}(\mathbf{r}, \mathbf{v}, t'; t)$ approximate the two-particle electron–electron and electron–ion distribution functions.[2] Since the ion distribution is spatially uniform, the function f_{ei} depends only on the electron velocity $\mathbf{v}$ and the electron position $\mathbf{r}$ relative to an ion. The arguments of the function f_{ee} are the relative position of the electrons in a pair $\mathbf{r} = \mathbf{r}_1 - \mathbf{r}_2$, the relative velocity of the electrons $\mathbf{v}_{\rm rel} = \mathbf{v}_1 - \mathbf{v}_2$, and the center-of-mass velocity $\mathbf{v}_{\rm cm} = (\mathbf{v}_1 + \mathbf{v}_2)/2$, where $\mathbf{r}_1$, $\mathbf{v}_1$ and $\mathbf{r}_2$, $\mathbf{v}_2$ are the coordinates and velocities, respectively, of the first and second electrons in a pair. It should be noted that in a collision of two electrons in a magnetic field their motion and the motion of the center of mass are independent. The motion of the center of mass is a cyclotron rotation, described by the equation

$$\dot{\mathbf{v}}_{\rm cm} = \boldsymbol{\omega}_B \times \mathbf{v}_{\rm cm},$$

and the relative motion

$$\dot{\mathbf{v}}_{\rm rel} = \boldsymbol{\omega}_B \times \mathbf{v}_{\rm rel} - \nabla(2U_{ee})/m$$

is a scattering of a particle with mass m in the potential $2U_{ee}$ of a stationary center. This effect makes it possible to assume the function f_{ee} to be uniform with respect to the center-of-mass coordinate.

Boltzmann's collision integral models the change in the single-particle electron distribution function $f(\mathbf{v}, t)$ as a result of instantaneous pair collisions, whose duration $t_{\rm col}$ is short compared with the characteristic time $t_{\rm rel}$ of the change occurring in the distribution f as a result of collisions. To calculate the collision integral in a magnetic field it is necessary to take into account in an appropriate manner the quasiperiodic change in the distribution f as a result of the cyclotron rotation of the electrons. In the present section the collision integrals are studied for cyclotron periods $2\pi/\omega_B$ which are short compared with the mean free time $t_{\rm rel}$ and for an arbitrary ratio of the collision duration $t_{\rm col}$ and the quantity ω_B^{-1}.

To find the general form of the collision integral we seek the corresponding expressions for the functions f_{ei} and f_{ee} in terms of f on the basis of the kinetic equations

$$\frac{\partial f_{ei}}{\partial t'} + \mathrm{div}_{\mathbf{r}}(\mathbf{v} f_{ei}) + \boldsymbol{\omega}_B \times \mathbf{v}\frac{\partial f_{ei}}{\partial \mathbf{v}} - \frac{1}{m}\frac{\partial U_{ei}}{\partial \mathbf{r}}\frac{\partial f_{ei}}{\partial \mathbf{v}} = 0, \qquad (6)$$

$$\frac{\partial f_{ee}}{\partial t'} + \mathrm{div}_{\mathbf{r}}(\mathbf{v}_{\rm rel} f_{ee}) + \boldsymbol{\omega}_B \times \mathbf{v}_{\rm rel}\frac{\partial f_{ee}}{\partial \mathbf{v}_{\rm rel}} - \frac{2}{m}\frac{\partial U_{ee}}{\partial \mathbf{r}}\frac{\partial f_{ee}}{\partial \mathbf{v}_{\rm rel}} + \boldsymbol{\omega}_B \times \mathbf{v}_{cm}\frac{\partial f_{ee}}{\partial \mathbf{v}_{\rm cm}} = 0, \qquad (7)$$

supplemented by the boundary conditions

$$f_{ei}(\mathbf{r}, \mathbf{v}, t'; t)\big|_{\substack{zv_z<0 \\ |z|\to\infty}} = n_i f(\hat{F}(t-t')\mathbf{v}, t), \qquad (8)$$

$$f_{ee}(\mathbf{r}, \mathbf{v}_{\rm rel}, \mathbf{v}_{\rm cm}, t'; t)\big|_{\substack{zv_{{\rm rel}\,z}<0 \\ |z|\to\infty}} = f(\hat{F}(t-t')(\mathbf{v}_{\rm cm} + \mathbf{v}_{\rm rel}/2), t) \times f(\hat{F}(t-t')(\mathbf{v}_{\rm cm} - \mathbf{v}_{\rm rel}/2), t). \qquad (9)$$

Here $\boldsymbol{\omega}_B = (eB/mc)\mathbf{z}^\circ$ and $\hat{F}(\Delta t)$ is a linear operator which rotates an arbitrary vector $\mathbf{a}$ by an angle $\omega_B\Delta t$ around the axis $\mathbf{z}^\circ$

$$\hat{F}(\Delta t)\mathbf{a} = \mathbf{z}^\circ(\mathbf{z}^\circ \cdot \mathbf{a}) + \mathbf{z}^\circ \times \mathbf{a}\sin(\omega_B\Delta t) + [\mathbf{z}^\circ \times \mathbf{a}] \times \mathbf{z}^\circ\cos(\omega_B\Delta t).$$

In the limit $\omega_B \longrightarrow 0$ the operator $\hat{F}$ becomes the unit operator and the expressions (8) and (9) become identical with the well-known condition for the particles to be uncorrelated before a collision. Then the conditions (8) and (9) and the corresponding solutions of equations (6) and (7) are stationary with respect to the argument t'. In a magnetic field the distribution f changes on time intervals $\Delta t \ll t_{\rm rel}$ primarily as a result of the cyclotron rotation of the electrons. Consequently, for arbitrary times t and t' we have

$$f(\mathbf{v}, t') = f(\hat{F}(t-t')\mathbf{v}, t).$$

The last relation is used in the boundary conditions (8) and (9) to separate formally the fast quasiperiodic variations of f with the cyclotron frequency and the slower collisional relaxation of f on time intervals of the order of $t_{\rm rel}$. Ultimately, the dependence of the functions f_{ei} and f_{ee} on the argument t' is periodic and reflects rapid variations of the two-particle distributions, while the dependence on the parameter t corresponds to a slow collisional relaxation of the electron distribution, which is negligible in instantaneous-collisions models. Actually, the functions f_{ei} and f_{ee} approximate to an adequate degree of accuracy the two-particle distributions at an arbitrary time t_0 with $t' = t_0$ and the parameters $|t - t_0| \ll t_{\rm rel}$.

In the present paper we shall study the collision integrals for stationary axisymmetric particle interaction potentials. In this case the periodic change in the distributions f_{ei} and f_{ee} is especially simple. The solutions of equations (6) and (7) with the boundary conditions (8) and (9) possess a symmetry, which can be convention-

[2] It should be underscored that the time dependence of the functions f_{ei} and f_{ee} is described by the argument t' and the parameter t. The advantages of using two variables and their meaning will be specially discussed below (see equations (6), (7) and (8), (9)).

ally called a synchronous rotation in coordinate and velocity space:

$$f_{ei}(\mathbf{r}, \mathbf{v}, t'; t) = f_{ei}(\hat{F}(\Delta t')\mathbf{r}, \hat{F}(\Delta t')\mathbf{v}, t' + \Delta t'; t), \quad (10)$$

$$f_{ee}(\mathbf{r}, \mathbf{v}_{\rm rel}, \mathbf{v}_{\rm cm}, t'; t) = f_{ee}(\hat{F}(\Delta t')\mathbf{r}, \hat{F}(\Delta t')\mathbf{v}_{\rm rel}, \hat{F}(\Delta t')\mathbf{v}_{\rm cm}, t' + \Delta t'; t), \quad (11)$$

where $\Delta t'$ is an arbitrary time shift. The distributions (10) and (11) become identical when particles with the same initial rotation phase before a collision arrive at definite points of the phase space at corresponding moments in time. Indeed, let $\mathbf{r}(t')$ be the trajectory of a particle in a collision with a stationary center. Then the function $\tilde{\mathbf{r}}(t) = \hat{F}(\Delta t')\,\mathbf{r}(t' - \Delta t')$ corresponds to a different particle trajectory. Far from a scattering center the particle velocities on both trajectories are the same:

$$\dot{\tilde{\mathbf{r}}}(t') = \hat{F}(\Delta t')\dot{\mathbf{r}}(t' - \Delta t') = \hat{F}(\Delta t')\hat{F}(-\Delta t')\dot{\mathbf{r}}(t') = \dot{\mathbf{r}}(t').$$

Consequently, the distribution functions on these trajectories are also the same, specifically, they are equal at the time t_1' at the point $\mathbf{r}(t_1')$, $\dot{\mathbf{r}}(t_1')$ and the at $t_2' = t_1' + \Delta t'$ at the point $\tilde{\mathbf{r}}(t_2') = \hat{F}(\Delta t')\mathbf{r}(t_1')$, $\dot{\tilde{\mathbf{r}}}(t_2') = \hat{F}(\Delta t')\dot{\mathbf{r}}(t_1')$, which corresponds to the expressions (10) and (11).

The symmetry property (10) and (11) established above imposes additional differential relations on the derivatives of the functions f_{ei} and f_{ee}:

$$\frac{\partial f_{ei}}{\partial t'} + \boldsymbol{\omega}_B \times \mathbf{r}\frac{\partial f_{ei}}{\partial \mathbf{r}} + \boldsymbol{\omega}_B \times \mathbf{v}\frac{\partial f_{ei}}{\partial \mathbf{v}} = 0,$$

$$\frac{\partial f_{ee}}{\partial t'} + \boldsymbol{\omega}_B \times \mathbf{r}\frac{\partial f_{ee}}{\partial \mathbf{r}} + \boldsymbol{\omega}_B \times \mathbf{v}_{\rm rel}\frac{\partial f_{ee}}{\partial \mathbf{v}_{\rm rel}} + \boldsymbol{\omega}_B \times \mathbf{v}_{\rm cm}\frac{\partial f_{ee}}{\partial \mathbf{v}_{\rm cm}} = 0.$$

These relations make it possible to eliminate the time derivatives $\partial f_{ei}/\partial t'$ and $\partial f_{ee}/\partial t'$ from the kinetic equations (6) and (7) and to express the integrands in the collision integrals (4) and (5) in the equivalent divergence form

$$\frac{1}{m}\frac{\partial U_{ei}}{\partial \mathbf{r}}\frac{\partial f_{ei}}{\partial \mathbf{v}} = \mathrm{div}_{\mathbf{r}}((\mathbf{v} - \boldsymbol{\omega}_B \times \mathbf{r})\, f_{ei}),$$

$$\frac{2}{m}\frac{\partial U_{ee}}{\partial \mathbf{r}}\frac{\partial f_{ee}}{\partial \mathbf{v}_{\rm rel}} = \mathrm{div}_{\mathbf{r}}((\mathbf{v}_{\rm rel} - \boldsymbol{\omega}_B \times \mathbf{r})\, f_{ee}).$$

Then the initial collision integrals (4) and (5) reduce to the corresponding integrals over distant closed surfaces of arbitrary shape:

$$I_{st}^{(ei)}(\mathbf{v}) = \oint f_{ei}((\mathbf{v} - \boldsymbol{\omega}_B \times \mathbf{r}) \cdot d\mathbf{S}), \quad (12)$$

$$I_{st}^{(ee)}(\mathbf{v}) = \int d\mathbf{v}_{\rm rel} \times \oint f_{ee}((\mathbf{v}_{\rm rel} - \boldsymbol{\omega}_B \times \mathbf{r}) \cdot d\mathbf{S})\big|_{\mathbf{v}_{\rm cm} = \mathbf{v} - \mathbf{v}_{\rm rel}/2}. \quad (13)$$

If the surfaces of integration were to rotate with the cyclotron frequency relative to the symmetry axes of the potentials, then the quantities $\boldsymbol{\omega}_B \times \mathbf{r}$ in the expressions (12) and (13) would represent the instantaneous velocities of points on these surfaces. Consequently, the collision integrals in the form (12) and (13) can be interpreted as fluxes of particles with the velocities $\mathbf{v}$ and $\mathbf{v}_{\rm rel}$ through the surfaces rotating with the cyclotron frequency ω_B.

Now we shall reduce the expressions (12) and (13) to the standard form, where the functions f_{ei} and f_{ee} are expressed in terms of the single-particle distribution function, and common variables, which do not depend on the choice of surfaces, are used for the integration. For rectilinear particle trajectories in the absence of a magnetic field this is achieved by introducing the well-known impact parameter $\mathbf{p}_r(\mathbf{r}) = \mathbf{r} - \mathbf{v}(\mathbf{r} \cdot \mathbf{v})/v^2$ (in the case of electron–electron collisions the velocity $\mathbf{v}$ here and below in this section is $\mathbf{v}_{\rm rel}$). The scattering cross section is the Jacobian of the transformation $\mathbf{p}_r(\mathbf{n}')$, where $\mathbf{n}' = \mathbf{v}'/v$ is the direction of the particle velocity after a collision. Since in a magnetic field the direction $\mathbf{n}'$ varies continuously because of Larmor rotation, the difference of the states of a particle before and after a collision must be characterized by the change in the stationary parameters of its free motion: the pitch angle θ, the initial rotation phase ϕ_0, and the modulus of the velocity v. Then an elastic collision is characterized by the change $\Delta\theta$ in the pitch angle and the change $\Delta\phi$ in the initial rotation phase. The pair of variables p_h, φ (Fig. 1), which give the position of the helical line of a particle,

$$x = p_h\cos\phi_B + r_B\cos(\omega_B z/v_z + \phi_B + \varphi),$$
$$y = p_h\sin\phi_B + r_B\sin(\omega_B z/v_z + \phi_B + \varphi) \quad (14)$$

before or after a collision relative to a scattering center, is the analog of the impact parameter $\mathbf{p}_r$ in a magnetic field (compare [16]). The parameter p_h is the distance between the axis of the helical line and the symmetry axis of the potential, i.e., the modulus of the radius vector of the guiding center of rotation

$$\mathbf{R}_{B_\perp} = \mathbf{r} + \boldsymbol{\omega}_B \times \mathbf{v}/\omega_B^2 - \mathbf{z}^\circ(\mathbf{r} \cdot \mathbf{z}^\circ).$$

The angle ϕ_B is the azimuthal angle of the vector $\mathbf{R}_{B_\perp}$. The parameter φ is the angle between the vectors $\mathbf{r}_B = -\boldsymbol{\omega}_B \times \mathbf{v}/\omega_B^2$ and $\mathbf{R}_{B_\perp}$ at an arbitrary point $\mathbf{r}$ of the undisturbed trajectory, increased by the phase difference of the Larmor rotation $(-\omega_B z/v_z)$ at $z = 0$ and at the point $\mathbf{r}$.

Since the interaction potential is axisymmetric, the change $\Delta\theta$ in the pitch angle and the change $\Delta\phi$ in the rotation phase are functions of four arguments: $\Delta\theta(p_h, \varphi, \theta, v)$ and $\Delta\phi(p_h, \varphi, \theta, v)$.

In the limit $\omega_B \longrightarrow 0$, when the characteristic size of the collision region $L_{\rm col} \ll r_B$ and the standard impact

parameter $\mathbf{p}_r$ can be used, the quantities (p_h, φ) and $\mathbf{p}_r$ are related as follows:

$$\mathbf{p}_r = (r_B - p_h)\mathbf{n}_1 + v_z r_B(\varphi - \pi)\mathbf{n}_2/v.$$

Here $\mathbf{n}_1 \uparrow\uparrow \mathbf{v} \times \boldsymbol{\omega}_B$ and $\mathbf{n}_2 \uparrow\uparrow \mathbf{v} \times [\mathbf{v} \times \boldsymbol{\omega}_B]$ are mutually orthogonal unit vectors.

It should be noted that the expressions (14) are valid if the interaction potential decreases sufficiently rapidly with distance (for example, for a screened Coulomb potential $U \propto \exp(-r/r_D)/r$), while an ordinary Coulomb field changes the longitudinal velocity of the particle and the pitch of the helical line at arbitrarily large distances from the scattering center. In the latter case, for

$$r \gg \max[p_h, |Z|e^2/mv_z^2]$$

the trajectory is determined by the same expressions (14), if $\omega_B z/v_z$ is replaced by

$$(\omega_B z/v_z)(1 - \ln(|z|/l)Ze^2/mv_z^2|z|),$$

where l is a constant (compare [32]). This modification changes the definition of the parameter φ, but it has no effect on the final form of the collision integral. Consequently, to simplify the exposition we shall assume the Coulomb potential to be screened.

Having determined the method for parameterizing the trajectories, we express the functions f_{ei} and f_{ee} in the collision integrals (12) and (13) in terms of the single-particle distribution function f and the shifts $\Delta\theta$ and $\Delta\phi$. We divide the surface of integration S by the plane $z = 0$ into two parts, S_1 and S_2, for which, respectively, $zv_z < 0$ and $zv_z > 0$. The particles on the surface S_1 have still not collided, so that for them, in accordance with the boundary conditions (8) and (9),

$$f_{ei}(\mathbf{r}, \mathbf{v}, t; t)\big|_{S_1} = n_i f(\mathbf{v}, t), \tag{15}$$

$$\begin{aligned} &f_{ee}(\mathbf{r}, \mathbf{v}_{\rm rel}, \mathbf{v}_{\rm cm}, t; t)\big|_{S_1} \\ &= f(\mathbf{v}_{\rm cm} + \mathbf{v}_{\rm rel}/2, t) f(\mathbf{v}_{\rm cm} - \mathbf{v}_{\rm rel}/2, t). \end{aligned} \tag{16}$$

The particles which have collided lie on the surface S_2. Using the constancy of the distributions along the particle trajectories and the symmetry of the trajectories under time reversal and specular reflection relative to an arbitrary plane passing through the symmetry axis of the potential, we find the distributions on the surface S_2:

$$f_{ei}(\mathbf{r}, \mathbf{v}, t; t)\big|_{S_2} = n_i f(\hat{L}(-\Delta\theta_{ei}, -\Delta\phi_{ei})\mathbf{v}, t), \tag{17}$$

$$\begin{aligned} &f_{ee}(\mathbf{r}, \mathbf{v}_{\rm rel}, \mathbf{v}_{\rm cm}, t; t)\big|_{S_2} \\ &= f(\mathbf{v}_{\rm cm} + \hat{L}(-\Delta\theta_{ee}, -\Delta\phi_{ee})\mathbf{v}_{\rm rel}/2, t) \\ &\times f(\mathbf{v}_{\rm cm} - \hat{L}(-\Delta\theta_{ee}, -\Delta\phi_{ee})\mathbf{v}_{\rm rel}/2, t), \end{aligned} \tag{18}$$

where

$$\Delta\theta_{ei(ee)} = \Delta\theta_{ei(ee)}(p_h(\mathbf{r}, \mathbf{v}), -\varphi(\mathbf{r}, \mathbf{v}), \pi - \theta, v),$$

$$\Delta\phi_{ei(ee)} = \Delta\phi_{ei(ee)}(p_h(\mathbf{r}, \mathbf{v}), -\varphi(\mathbf{r}, \mathbf{v}), \pi - \theta, v)$$

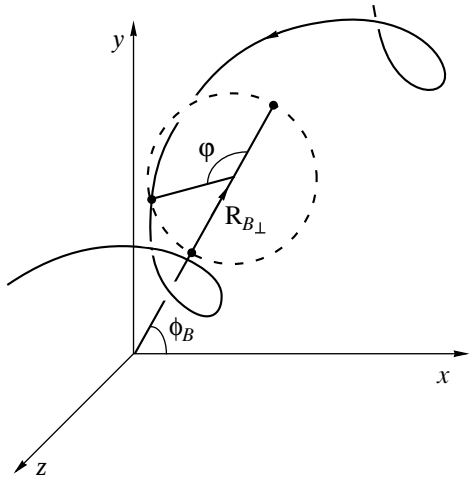


Fig. 1. Parameterization of the helical trajectory of an electron far from a scattering center.

are the changes in the pitch angle and phase of a particle in a collision in the corresponding potentials U_{ei} and $2U_{ee}$ with the parameters of the trajectory before a collision $p_h' = p_h$, $\varphi' = -\varphi$, $\theta' = \pi - \theta$, and $v' = v$, and the parameters $p_h(\mathbf{r}, \mathbf{v})$, $\varphi(\mathbf{r}, \mathbf{v})$, θ, and v refer to the point of integration $\mathbf{r}$ on the surface S_2. The operator $\hat{L}(\delta\theta, \delta\phi)$ establishes a correspondence between an arbitrary vector $\mathbf{a}$ with the spherical coordinates (a, θ, ϕ) and the vector $\mathbf{a}' = (a, \theta + \delta\theta, \phi + \delta\phi)$.

The parameters (p_h, φ) introduced above are also convenient from the standpoint of the integration in equations (12) and (13) over an arbitrary surface S, since the differential

$$(\mathbf{v} - \boldsymbol{\omega}_B \times \mathbf{r}) \cdot d\mathbf{S} = \pm|v_z p_h dp_h d\varphi|. \tag{19}$$

The plus and minus signs in equation (19) refer, respectively, to particles leaving and entering the volume. This relation can be established by a direct calculation of the Jacobian of the transformation $(p_h(\mathbf{r}), \varphi(\mathbf{r}))$, where $\mathbf{r} \in S$.

Using the expressions (15)–(19) the integrals (12) and (13) can be finally put into the form

$$\begin{aligned} I_{st}^{(ei)}(\mathbf{v}, t) &= n_i|v_z| \int_{-\pi}^{\pi} d\varphi \int_0^{\infty} dp_h p_h \\ &\times [f(\hat{L}(-\Delta\theta_{ei}, -\Delta\phi_{ei})\mathbf{v}, t) - f(\mathbf{v}, t)], \end{aligned} \tag{20}$$

$$\begin{aligned} I_{st}^{(ee)}(\mathbf{v}, t) &= \int d\mathbf{v}_{\rm rel}|v_{{\rm rel}\,z}| \int_{-\pi}^{\pi} d\varphi \int_0^{\infty} dp_h p_h \\ &\times [f(\mathbf{v}_{\rm cm} + \hat{L}(-\Delta\theta_{ee}, -\Delta\phi_{ee})\mathbf{v}_{\rm rel}/2, t) \\ &\times f(\mathbf{v}_{\rm cm} - \hat{L}(-\Delta\theta_{ee}, -\Delta\phi_{ee})\mathbf{v}_{\rm rel}/2, t) \\ &- f(\mathbf{v}, t) f(\mathbf{v} - \mathbf{v}_{\rm rel}, t)]\big|_{\mathbf{v}_{\rm cm} = \mathbf{v} - \mathbf{v}_{\rm rel}/2}. \end{aligned} \tag{21}$$

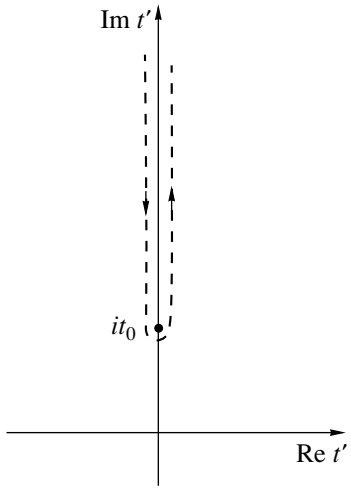


Fig. 2. Deformed contour of integration for calculating J.

The definition of the differential scattering cross section in the form

$$\sigma_{ei(ee)}(\theta', \phi'; \theta, \phi, v) = \sum_{\substack{\Delta\theta_{ei(ee)}(p_h, -\varphi, \pi-\theta, v) = \theta-\theta' \\ \Delta\phi_{ei(ee)}(p_h, -\varphi, \pi-\theta, v) = \phi-\phi'}} \left| \frac{v}{v_z p} \times \frac{\partial(\cos(\theta - \Delta\theta_{eiee}), \Delta\phi_{eiee})}{\partial(p_h, \varphi)} \right|^{-1} \quad (22)$$

makes it possible to write the expressions (20) and (21) in the standard Boltzmann form

$$I_{st}^{(ei)}(\mathbf{v}, t) = n_i v \oint d\mathbf{n}' \sigma_{ei}(\mathbf{n}'; \mathbf{v})[f(v\mathbf{n}', t) - f(\mathbf{v}, t)], \quad (23)$$

$$I_{st}^{(ee)}(\mathbf{v}, t) = \int d\mathbf{v}_{\rm rel} \oint d\mathbf{n}' v_{\rm rel} \sigma_{ee}(\mathbf{n}'; \mathbf{v}_{\rm rel}) \times [f(\mathbf{v}_{\rm cm} + v_{\rm rel}\mathbf{n}'/2, t) f(\mathbf{v}_{\rm cm} - v_{\rm rel}\mathbf{n}'/2, t) - f(\mathbf{v}, t) f(\mathbf{v} - \mathbf{v}_{\rm rel}, t)]|_{\mathbf{v}_{\rm cm} = \mathbf{v} - \mathbf{v}_{\rm rel}/2}, \quad (24)$$

which is identical to the analogous expression in an isotropic medium.

3. CHARACTERISTIC FEATURES OF ELECTRON COLLISIONS WITH POSITIVE IONS IN STRONG MAGNETIC FIELDS

It is shown in [29, 30] that in strong magnetic fields the change in the electron pitch angle is determined by close collisions with impact parameters

$$p_h \leqslant p_{\rm cr} \sim (mc^2/B^2)^{1/3}.$$

However, the Coulomb force near an ion is much stronger than the Lorentz force on an unperturbed Larmor trajectory, so that the pitch angle changes substantially even with larger impact parameters differing from $(mc^2/B^2)^{1/3}$ by a logarithmic factor.

In the present section the dimensionless lengths, velocities, and times, normalized to

$$L_u = (Zmc^2/B^2)^{1/3}, \quad V_u = \omega_B L_u, \quad T_u = \omega_B^{-1},$$

respectively, will be used. In this normalization the electron equations of motion

$$\ddot{\mathbf{r}} = [\mathbf{z}^\circ, \dot{\mathbf{r}}] - \mathbf{r}/r^3$$

contain no numerical parameters.

To calculate p_{cr} we shall find the change in the transverse velocity of an electron with respect to the corresponding "unperturbed" particle trajectory $\mathbf{r}(t)$. We shall assume that on a given trajectory the Coulomb field changes only the component v_z of the electron velocity

$$v_z^2 = v_{0\|}^2 + \frac{2}{\sqrt{p_h^2 + z^2}},$$

where $v_{0\|} > 0$ is the initial electron velocity along the magnetic field. In this approximation the displacement of the guiding center of rotation, as a result of electric and polarization drifts, as well as the change in the electron rotation frequency in the presence of a nonuniform electric field are neglected. All indicated effects are important only for $p_h \sim 1$. Then, after a collision the components of the electron velocity in the xy plane have the form

$$v_x(t) = \mathrm{Re}[v_{0\perp}\exp(-i\phi_0 - it)(1 - J\exp(i\tilde{\varphi}))],$$

$$v_y(t) = -\mathrm{Im}[v_{0\perp}\exp(-i\phi_0 - it)(1 - J\exp(i\tilde{\varphi}))],$$

where

$$J = p_h \int_{-\infty}^{\infty} \frac{\exp(it')}{(p_h^2 + z^2(t'))^{3/2}} dt',$$

$v_{0\perp}$ is the modulus of the initial particle velocity in the xy plane, the constant $\tilde{\varphi}$ is uniquely related with the collision parameter φ, and the function $z(t')$ is determined by the relation

$$t'(z) = \frac{p_h^{3/2}}{\sqrt{2}} \int_0^{z/p_h} \frac{dw}{\sqrt{v_{0\|}^2 p_h/2 + (1 + w^2)^{-1/2}}}.$$

The variable t' and the corresponding function $z(t')$ can be assumed to be complex. We deform the contour of integration along the real axis t' (Fig. 2) into a contour directly encompassing the cut $(it_0, +i\infty)$. Confining our attention to the case $v_{0\|}^2 p_h \ll 1$, we find in the region that is important for the integration

$$t'(z) \approx it_0 + \frac{2^{7/4}}{5}\exp\left(\frac{i\pi}{8}\right)p_h^{3/2}\left(\frac{z}{p_h} - i\right)^{5/4},$$

$$(p_h^2 + z^2)^{3/2} \approx \left(\frac{25}{2}\right)^{3/5} \exp\left(\frac{i3\pi}{5}\right) p_h^{6/5} (t' - it_0)^{6/5},$$

where

$$t_0(p_h) = \frac{p_h^{3/2}}{\sqrt{2}} \int_0^1 (1 - w^2)^{1/4} dw$$

$$= \sqrt{\frac{\pi}{8}} \frac{\Gamma(5/4)}{\Gamma(7/4)} p_h^{3/2} = \frac{\Gamma^2(1/4)}{12\sqrt{\pi}} p_h^{3/2} \approx 0.62 p_h^{3/2}$$

(see formula 2.5.3.1 in [33]), and $\Gamma(x)$ is the gamma function. Then

$$J = \frac{2^{8/5}\pi}{5^{6/5} p_h^{1/5}} \exp(-t_0) \frac{i}{2\pi} \int_C \frac{\exp(-t)}{(-t)^{-6/5}} dt$$

$$= \frac{2^{8/5}\pi}{5^{6/5} p_h^{1/5}} \frac{\exp(-t_0)}{\Gamma(6/5)} \approx \frac{1.50}{p_h^{1/5}} \exp(-t_0), \quad (25)$$

where the contour C passes in a clockwise direction around the positive semiaxis $(0, +\infty)$ and the corresponding integral along this contour is identical to the Hankel integral [34].

We shall determine the desired parameter p_{cr} as the value of the parameter p_h for which the quantity (25) is the initial electron velocity v_0:

$$p_{cr} = \left(\frac{12\sqrt{\pi}}{\Gamma^2(1/4)} \ln\frac{1}{v_0}\right)^{2/3} \approx 1.38 \ln^{2/3}\frac{1}{v_0}, \quad (26)$$

or in dimensional form

$$p_{cr} = r_e Z^{1/3} \left(\frac{12\sqrt{\pi}}{\Gamma^2(1/4)} \frac{B_{cl}}{B}\right)^{2/3} \ln^{2/3}\left(\left(\frac{ZB}{B_{cl}}\right)^{1/3} \left(\frac{c}{v_0}\right)\right). \quad (27)$$

The expressions (26) and (27) are logarithmically accurate values of the parameter p_{cr}.

We shall now consider a change in the initial rotation phase of an electron. In close collisions it is of the order of π. Consequently, the corresponding contribution of close trajectories to the transport frequencies and the collision integral are characterized by cross sections of the order of $\sigma_0 = \pi p_{cr}^2$. This contribution will be considered in the next section.

In each distant collision the change in the initial phase is $\Delta\phi \ll 1$, but its magnitude decreases quite slowly—according to a power law—with increasing impact parameter. The shift $\Delta\phi$ is related with the difference of the electron rotation frequency in the presence of the nonuniform electric field of the scattering center and the cyclotron frequency: a variable component of the force, proportional to the gradient of the field and the rotation radius of the particle, which produces an additional centripetal acceleration, is present on the particle trajectory.[3] The small shift of the instantaneous rotation frequency ω can be found as done in [26]:

$$\Delta\omega \equiv \omega - \omega_B = \frac{1}{2}\Delta_\perp U, \quad (28)$$

where

$$\Delta_\perp = \partial^2/\partial x^2 + \partial^2/\partial y^2$$

is the Laplace operator in a plane orthogonal to the magnetic field. The shift of the initial rotation phase as a result of collisions has regular and diffusion components which are characterized by the cross sections

$$\sigma_r = \int_{-\pi}^{\pi} d\varphi \int_{p_{cr}}^{\infty} dp_h p_h \Delta\phi,$$

$$\sigma_d = \int_{-\pi}^{\pi} d\varphi \int_{p_{cr}}^{\infty} dp_h p_h (\Delta\phi)^2.$$

Assuming the trajectories of the guiding centers to be rectilinear and using the expression (28), we find the phase shift in a distant collision

$$\Delta\phi = -\frac{1}{2}\int_{-\infty}^{\infty} \frac{dz}{\sqrt{v_{0\parallel}^2 + 2/\sqrt{p_h^2 + z^2}}} \left(\Delta_\perp \frac{1}{r}\right)_{r = \sqrt{p_h^2 + z^2}}.$$

Then the cross section σ_r can be written as

$$\sigma_r = -\pi \int_{p_{cr}}^{\infty} dp_h$$

$$\times \int_{-\infty}^{\infty} dz \frac{p_h}{\sqrt{v_{0\parallel}^2 + 2/\sqrt{p_h^2 + z^2}}} \left(\Delta_\perp \frac{1}{r}\right)_{r = \sqrt{p_h^2 + z^2}} = \sigma_{r1} - \sigma_{r2},$$

where

$$\sigma_{r1} = -\frac{\pi}{v_{0\parallel}} \int_0^{\infty} d\eta$$

$$\times \int_{-\infty}^{\infty} d\zeta \frac{\eta}{\sqrt{1 + 1/\sqrt{\eta^2 + \zeta^2}}} \left(\Delta_\perp \frac{1}{r}\right)_{r = \sqrt{\eta^2 + \zeta^2}}, \quad (29)$$

$$\sigma_{r2} = -\frac{\pi}{v_{0\parallel}} \int_0^{p_{cr} v_{0\parallel}^2/2} d\eta \int_{-\infty}^{+\infty} d\zeta$$

$$\times \frac{\eta}{\sqrt{1 + 1/\sqrt{\eta^2 + \zeta^2}}} \left(\Delta_\perp \frac{1}{r}\right)_{r = \sqrt{\eta^2 + \zeta^2}}.$$

[3] In the magnetic drift approximation (averaging method of [35]) it can be shown that the adiabatic inclusion of an additional centripetal acceleration changes the effective rotation radius while preserving the oscillatory velocity of the particle.

The component σ_{r2} can be found exactly in the limit $p_{\rm cr}v_{0\|}^2/2 \ll 1$:

$$\sigma_{r2} = \frac{4\sqrt{\pi}}{5}\Gamma^2(3/4)\sqrt{p_{\rm cr}} = \frac{2\pi}{15}\frac{\sigma_0}{\ln(1/v_0)} \ll \sigma_0 = \pi p_{\rm cr}^2.$$

The component σ_{r1} should be especially noted, since in a strict analysis the integral (29) does not converge: in the region $\zeta \gg 1$ and $\eta \gg 1$ the integrand in equation (29) slowly approaches 0 as $1/(\zeta^2 + \eta^2)$. Ultimately, the result of the integration depends strongly on the method of passage to the limits with respect to ζ and η. Even though there is an additional indefiniteness in the cross section σ_{r1}, we note that this cross section increases as the velocity $v_{0\|}$ decreases and it exceeds the cross section σ_0 for close collisions with low velocities $v_{0\|}$. However, the value of σ_{r1} is not fundamental for calculating the collision integral (see also [14]). It follows from the expression (29) that

$$\sigma_{r1} \propto v_{0\|}^{-1}$$

and the regular shift of the rotation frequency

$$d\phi_0/dt = n_i v_{0\|}\sigma_{r1}$$

does not depend on the particle velocity. Consequently, the regular collisional change in the initial phase can be reduced to a replacement of the magnetic field B by the corresponding effective value

$$B_{\rm eff} = B + (mc/e)d\phi_0/dt \approx B.$$

We shall now find the cross section σ_d using the same approximation as in the calculation of σ_{r2}:

$$\sigma_d \approx \frac{\pi}{8}v_{0\|}^2 \int\limits_{p_{\rm cr}v_{0\|}^2/2}^{\infty} \frac{d\eta}{\eta^2} \times \left(\int\limits_{-\infty}^{\infty} d\zeta(1+\zeta^2)^{1/4}\frac{\partial^2}{\partial\zeta^2}\frac{1}{\sqrt{1+\zeta^2}}\right)^2 = \frac{2\Gamma^4(3/4)}{25p_{\rm cr}} = \frac{\pi^2}{450}\frac{\sigma_0}{\ln^2(1/v_0)}.$$

As one can see, the cross section $\sigma_d \ll \sigma_0 = \pi p_{\rm cr}^2$ so that the effect of distant collisions on particle diffusion along the initial rotation phases is negligible compared with the analogous contribution of close collisions.

Thus, as a result of the specific form of the regular change in the initial rotation phase in distant collisions and the smallness of the cross section σ_d compared with σ_0, it can be concluded that only close collisions determine the form of the electron–ion collision integral in strong magnetic fields (1).

4. DIFFERENTIAL SCATTERING CROSS SECTION IN A STRONG MAGNETIC FIELD

A general expression for the collision integral in a magnetic field for an arbitrary axisymmetric interaction potential was obtained in Section 2. For a specific potential it is necessary to perform a corresponding calculation of the differential scattering cross section. This calculation can be performed either analytically or numerically. A large number of collisions must be taken into account in order to find the differential scattering cross section numerically. Moreover, under conditions where the drift velocity is large compared with the Larmor rotational velocity the required accuracy for the numerical values of the pitch angle and phase can be attained only by highly accurate calculations of the trajectories.

There is no known general analytical solution of the problem of the motion of an electron in a uniform magnetic field and a Coulomb electric field. Approximate analytic methods of the type "instantaneous impact" [36] or the magnetic drift approximation are inapplicable in strong magnetic fields (1). Thus, an assumption based on a characteristic property of collisions is required in order to calculate the scattering cross section analytically. Since an electron in an effective collision moves along a quasi-bound trajectory with multiple passes near the ion,[4] it is natural to infer that the result is equivalent to several collisions with a single pass, i.e., the distribution of the particles scattered by an individual ion does not depend on the distribution of the particles in the incident flux.

Let us consider a spatially uniform flux of electrons with unit density and a fixed pitch angle θ_0, initial phase ϕ_0, and velocity v_0:

$$f_t(\mathbf{v}, t) = \frac{1}{v_0^2}\delta(v - v_0) \times \delta(\phi - \phi_0 - \omega_B t)\delta(\cos\theta - \cos\theta_0), \quad (30)$$

where $\delta(x)$ is a delta function. We shall find the rate of appearance of particles with pitch angle θ and phase ϕ which are different from the corresponding parameters θ_0 and $\phi_0 + \omega_B t$ in the incident flux:

$$\int\limits_0^{\infty} dv\, v^2\left(\frac{\partial f_t}{\partial t}\right)_{\substack{\theta\neq\theta_0\\ \phi\neq\phi_0+\omega_B t}} = n_i v_0 \sigma_{ei}(\theta_0, \phi_0 + \omega_B t; \theta, \phi, v_0). \quad (31)$$

The total number of close collisions per unit time for the incident flux $n_i v_0|\cos\theta_0|\sigma_0$ must be identical to the expression (31) integrated over all directions (θ, ϕ) with

[4] See the qualitative analysis and discussion of numerical calculations of such trajectories in [30].

the exception of the direction of the velocity in the incident flux (θ_0, $\phi_0 + \omega_B t$):

$$\oint\limits_{\substack{\theta \neq \theta_0 \\ \phi \neq \phi_0 + \omega_B t}} d\phi d\theta \sin\theta \times \int\limits_0^\infty dv\, v^2 \left(\frac{\partial f_t}{\partial t}\right)_{\substack{\theta \neq \theta_0 \\ \phi \neq \phi_0 + \omega_B t}} \tag{32}$$

$$= n_i v_0 |\cos\theta_0| \sigma_0.$$

In accordance with the assumption made above concerning the character of the scattering by a single center, the ratio $\sigma_{ei}(\theta_0, \phi_0 + \omega_B t; \theta, \phi, v_0)/(|\cos\theta_0|\sigma_0)$ of the expressions (31) and (32) should not depend on θ_0 and ϕ_0. Therefore the differential cross section σ_{ei} has the form of the product $\sigma_{ei}(\theta'; \theta) = \sigma_0|\cos\theta'|w(\theta)$, where the function $w(\theta)$ satisfies the normalization

$$2\pi \int\limits_0^\pi d\theta \sin\theta w(\theta) = 1.$$

The condition for the conservation of the total number of particles

$$\int\limits_0^\infty dv\, v^3 \int\limits_0^{2\pi} d\phi \int\limits_0^\pi d\theta \sin\theta \int\limits_0^{2\pi} d\phi' \int\limits_0^\pi d\theta' \sin\theta' \sigma_{ei}(\theta'; \theta)$$
$$\times [f_t(v, \theta', \phi', t) - f_t(v, \theta, \phi, t)]$$

$$= (2\pi)^2 v_0 \int\limits_0^\pi d\theta \sin\theta [\sigma_{ei}(\theta_0; \theta) - \sigma_{ei}(\theta; \theta_0)]$$

$$= 2\pi v_0 \sigma_0 (|\cos\theta_0| - 2\pi w(\theta_0)) = 0$$

uniquely determines the factor $w(\theta) = |\cos\theta|/2\pi$ and the differential cross section

$$\sigma_{ei}(\theta'; \theta, v) = \frac{\sigma_0(v)}{2\pi} |\cos\theta' \cos\theta|, \tag{33}$$

where

$$\sigma_0(v) = \pi p_{cr}^2(v) = \pi \left(\frac{12\sqrt{\pi}}{\Gamma^2(1/4)}\right)^{4/3}$$
$$\times \left(\frac{Zmc^2}{B^2}\right)^{2/3} \ln^{4/3}\left(\frac{Ze^3 B}{m^2 c v^3}\right)^{1/3}.$$

The scattering cross section (33) admits a simple quantum-mechanical interpretation. An electron occupies with equal probability an arbitrary Landau level with energy less than the initial energy of the electron. Here the nth Landau level is treated as a collection of states with definite energy $mv_\perp^2/2 = (n + 1/2)\hbar\omega_B$. Indeed, the rate of change of the level populations

$$\frac{2\pi\hbar\omega_B}{m} \frac{\partial}{\partial t} \int\limits_{-\infty}^\infty f_t(\mathbf{v}, t) dv_z$$
$$= n_i |\cos\theta_0| \sigma_0(v_0) \frac{\hbar\omega_B}{m v_0} \times \begin{cases} 1, & \text{if } v_\perp \leq v_0 \\ 0, & \text{if } v_\perp > v_0, \end{cases} \tag{34}$$

is the same for all $v_\perp \leq v_0$ for scattering of the flux (30). Having occupied a level n, the electron occupies with equal probability one of the states with a fixed momentum $p_\parallel$ in the direction of the magnetic field in the interval from $p_\parallel(n)$ to $p_\parallel(n) + \Delta p_\parallel(n)$. If the initial electron energy $mv_0^2/2$ falls into the range from E_0 to $E_0 + \Delta E_0$, then the corresponding interval is

$$\Delta p_\parallel(n) = m\Delta E_0 / |p_\parallel(n)| \propto |\cos\theta|^{-1}.$$

Since the number Δk of states in the interval $\Delta p_\parallel$ is proportional to the length of this interval, the conditional probability $1/\Delta k$ for an electron to occupy one of these states is proportional to $|\cos\theta|$, which the factor $|\cos\theta|$ in the expression (33) describes.

It should be noted that the electron–ion collisions considered above result in isotropization of the electron distribution, since the collision integral (23) with the cross section (33) is zero only for an isotropic distribution function.

5. DISCUSSION

In the present paper we found the pair-collision integral in Boltzmann's form (20)–(24) in a magnetic field and we obtained the cross section (33) for electron–ion collisions in a strong magnetic field (1) for positively charged ions. The results obtained are limited by the possibility of the classical method of describing the motion of particles in a collision, when the de Broglie wavelength $\lambda_B = \hbar/mv$ is less than the characteristic lengths in the problem. The Larmor radius of an electron is such a smallest scale, so that the results obtained are valid in the region $\lambda_B < v/\omega_B$, where the quantization of the transverse motion of an electron is unimportant:

$$\hbar\omega_B < mv^2.$$

In the "magnetic field–electron energy" parameter plane (Fig. 3) the classical approximation is valid below the line $r_B = \lambda_B$ and to the left of the vertical straight line $r_s = \lambda_B$. In this region the sectors of strong (*1*) and weak (*2*) magnetic fields are separated by the line $r_B = r_s$. The Born approximation is valid in the region (*3*) to the right of the straight line $r_s = \lambda_B$, including for $E < \hbar\omega_B$, where the electron can be scattered only backwards, remaining in the lowest Landau level. A special investigation and calculation of Coulomb collisions is required in the sector (*4*).

It should be noted that the spatial nonuniformity of the plasma leads to the appearance of additional terms

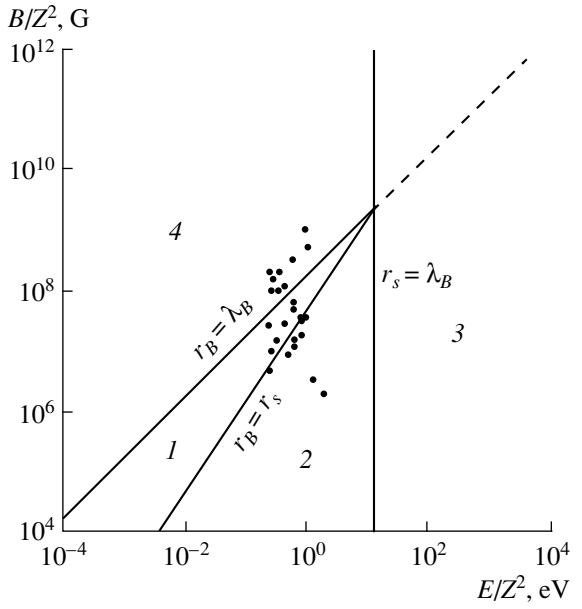


Fig. 3. Regions of applicability of various approximations for calculating collisions. $E = mv^2/2$—electron energy. Above the straight line $r_B = \lambda_B$, where $E < \hbar\omega_B$ and the electron occupies the lowest Landau level, the quantity E refers to the energy $m v_\parallel^2/2$ of longitudinal motion. The dots show the corresponding parameters for thermal electrons in the photospheres of magnetic white dwarfs [37].

in the collision integral. These terms describe the diffusion of the guiding center of rotation because of collisions and are proportional to the spatial derivatives of the particle distribution function. In [38] an electron–ion collision integral in a weak magnetic field was obtained. This integral takes account of the nonuniformity of the particle density (see also [39]). In a strong magnetic field (1) the spatial diffusion is due primarily to electric drift of the guiding center at distant collisions $p_h \gtrsim r_s$, so that it can be calculated using the results of [38]. Compared with [38], only the Coulomb logarithm changes in the collision integral. This problem requires additional calculations.

The scattering cross section (33) found above makes it possible to calculate the relaxation of the electron distribution as a result of elastic electron–ion collisions[5] with other factors, for example, electron–electron collisions or electromagnetic radiation and absorption, being negligible. Thus, in a quasineutral plasma a distribution of the form (2) is established as a result of electron–electron collisions. In the general case this distribution is nonisotropic and possesses a directed hydrodynamic velocity. The electron–ion collisions lead to isotropization of this distribution, as a result of which Maxwell's distribution is established over a characteristic mean free time $(n_i v_T \sigma_0(v_T))^{-1}$ of a particle with thermal velocity v_T.

We shall now present the main transport frequencies characterizing the relaxation. The change in the velocity $u_\parallel \ll v_T$ directed along the magnetic field for the distribution

$$f(\mathbf{v}) = \frac{n_e}{(2\pi)^{3/2} v_T^3} \exp(-(\mathbf{v} - u_\parallel \mathbf{z}^\circ)^2/(2v_T^2)),$$

where $n_e = n_i$ is the electron density, is given by the frequency

$$\nu_\parallel = 2\sqrt{2\pi} n_i v_T p_{\mathrm{cr}}^2 = 9.52 n_i v_T \left(\frac{Zmc^2}{B^2}\right)^{2/3} \ln^{4/3}\left(\frac{Ze^3 B}{m^2 c v_T^3}\right)^{1/3}, \quad (35)$$

so that $du_\parallel/dt = -\nu_\parallel u_\parallel$. A similar change in the modulus of the hydrodynamic velocity in a direction transverse to the magnetic field is given by the frequency

$$\nu_\perp = \sqrt{2\pi} n_i v_T p_{\mathrm{cr}}^2 = 4.76 n_i v_T \left(\frac{Zmc^2}{B^2}\right)^{2/3} \ln^{4/3}\left(\frac{Ze^3 B}{m^2 c v_T^3}\right)^{1/3}. \quad (36)$$

Relaxation of the anisotropy of the temperatures for the distribution

$$f(\mathbf{v}) = \frac{n_e}{(2\pi)^{3/2} v_{T_\perp}^2 v_{T_\parallel}} \exp\left(-\frac{v_\parallel^2}{2v_{T_\parallel}^2} - \frac{v_\perp^2}{2v_{T_\perp}^2}\right)$$

is characterized by the rate

$$dv_{T_\perp}^2/dt = \nu_{\mathrm{eff}}(v_{T_\parallel}^2 - v_{T_\perp}^2),$$

where for any ratio of $v_{T_\perp}$ and $v_{T_\parallel}$ the frequency

$$\nu_{\mathrm{eff}} = \sqrt{\frac{\pi}{2}} n_i v_{T_\parallel} p_{\mathrm{cr}}^2 = 2.38 n_i v_{T_\parallel} \left(\frac{Zmc^2}{B^2}\right)^{2/3} \ln^{4/3}\left(\frac{Ze^3 B}{m^2 c v_T^3}\right)^{1/3}. \quad (37)$$

It should be noted that estimates of $\nu_\parallel$ and ν_{eff} for electron-proton collisions have been found in [29, 30]:

$$\nu_\parallel \sim 0.88 n_i v_T \left(\frac{mc^2}{B^2}\right)^{2/3}, \quad (38)$$

$$\nu_{\mathrm{eff}} \sim 0.37 n_i v_T \left(\frac{mc^2}{B^2}\right)^{2/3}. \quad (39)$$

The numerical factor in equation (38) was determined from the condition that in the region $\ln(r_B/r_s) = 1$ the

[5] Energy transfer between electrons and ions requires a special investigation. Distant collisions, whose duration is shorter than an ion cyclotron period, as well as close collisions as a result of the long residence time of an electron near an ion on a quasi-bound orbit could be important for this process.

quantity (38) is equal to the corresponding frequency in weak magnetic fields (see (41) and the discussion below). Adjustment of the analogous coefficient in equation (39) gives a smooth transition from the frequency (39) to the corresponding frequency ν_{eff} (see equation (40)) in weak magnetic fields. This transition occurs at $\ln(r_B/r_s) = 3/4$.

As already noted, in strong magnetic fields the collisions of electrons with positive and negative ions possess completely different properties (the latter can be calculated using directly the results of the investigation of electron–electron collisions [26]), while in weak fields their properties are largely similar (the transport frequencies depend on the squared ion charge). We shall trace the reasons for the difference in collisions with different ions for the transport frequencies (35), (36), and (37). Using the collision integral from [12] or [13] in weak magnetic fields, it is easy to find the frequency

$$\nu_{\text{eff}} = \frac{8}{15}\sqrt{2\pi}n_i v_T\left(\frac{Ze^2}{mv_T^2}\right)^2 \ln\frac{r_B}{r_s}. \qquad (40)$$

At the boundary of the regions of the parameters for weak and strong fields this frequency is equal in order of magnitude to the corresponding frequencies for positive ions (37) and negative ions (compare [27]). For subsequent motion in the strong-field region, the frequency ν_{eff} of collisions with negative ions decreases exponentially, $\ln\nu_{\text{eff}} \sim (r_s/r_B)^{2/5}$ [27], and satisfies the expression (37) for positive ions.

The transport frequency $\nu_\parallel$ in weak magnetic fields has the form [10]

$$\nu_\parallel = \frac{4}{3}\sqrt{2\pi}n_i v_T\left(\frac{Ze^2}{mv_T^2}\right)^2\left(\ln\frac{r_B}{r_s} + \frac{3}{2}\ln\frac{r_D}{r_B}\right). \qquad (41)$$

The corresponding frequency in strong fields for negative ions, calculated by analogy to electron–electron collisions [26], is

$$\nu_\parallel = 4\sqrt{2\pi}n_i v_T\left(\frac{Ze^2}{mv_T^2}\right)^2 \ln\frac{r_D}{r_s}. \qquad (42)$$

At the boundary of the regions of the parameters for strong and weak magnetic fields $r_B = r_s$, the frequencies (41) and (42) differ by a factor of 2,[6] while the frequencies (41) and (35) differ by a large logarithmic factor $\ln(r_D/r_s)$. This difference is due to the fact that the second logarithmic factor in expression (41) is of the same nature as the corresponding factor in expression (42). It is determined by collisions of thermal particles with small longitudinal velocities

$$\sqrt{|Z|e^2/(mr_D)} \ll v_\parallel \ll \min[v_T, \sqrt{|Z|e^2/(mr_B(v_T))}],$$

for which the distance

$$r_s(v_\parallel) = \frac{|Z|e^2}{mv_\parallel^2}$$

is much greater than their Larmor radius $r_B(v_T) \ll r_D$ but less than the Debye radius r_D. The indicated collisions with impact parameters $p_h < 2r_s(v_\parallel)$ are head-on collisions of electrons with negative ions, for which the direction of the longitudinal velocity of the electron changes while the modulus of this velocity is conserved. The factor of two difference between the frequencies (41) and (42) is due to the fact that the perturbation method does not give a correct description of such collisions. Similar collisions with positive ions do not lead to any resulting change in the longitudinal velocity: the electron simply passes by the ion with acceleration. This makes it possible to conclude that in weak magnetic fields the second logarithmic factor in the expression (41) is present in the frequency of collisions with negative ions and must be multiplied by two, but it does not occur in the corresponding frequency for positive ions. The analysis of the frequency $\nu_\parallel$ also shows that in weak magnetic fields perturbation methods must be used with great care for particles with small components of the velocity along the magnetic field, especially when Fourier transforms of the interaction potentials are used and the region where the perturbation method breaks down is difficult to single out explicitly. In very weak magnetic fields, where $r_B(v_T) \gg r_D$, the effects determined by slow particles become negligible because the number of such particles is relatively small: their velocities in a direction transverse to the magnetic field are limited from above by the quantity $\omega_B r_D \ll v_T$.

The frequency $\nu_\parallel$ ordinarily also characterizes bremsstrahlung absorption of electromagnetic waves polarized in the direction of the magnetic field. For this it is necessary to calculate the contribution to $\nu_\parallel$ only from collisions whose duration is shorter than the period of oscillations of the electromagnetic wave. For collisions of longer duration the change in the electron velocity occurs without absorption of a wave with a fixed frequency. In strong magnetic fields, for a fixed radiation frequency $\omega \lesssim \omega_B$ there exist many collisions with impact parameters $p_h \gtrsim p_{\text{cr}}$ in which the longitudinal velocity essentially does not change as a result of collisions, but the duration of the collisions is shorter than the period of the oscillations of the electromagnetic wave. Radiation absorption occurs in these collisions and it can be substantial in a definite frequency range. The calculation of the inverse bremsstrahlung coefficients in strong fields is the subject of further investigations.

ACKNOWLEDGMENTS

I thank V.V. Zheleznyakov and A.V. Serber for their careful attention to this work and for helpful remarks during the preparation of the article.

[6] A similar and more general relation for the electron–electron collision integral has been given in [26].

This work was supported by the Russian Foundation for Basic Research (project no. 99-02-18244), the Council on State Support of Leading Scientific Schools (project no. 00-15-96674), and the Commission of the Russian Academy of Sciences on Work with Youth (project "Interaction of intense electromagnetic and neutrino radiation with plasma in superstrong magnetic and gravitational fields of compact cosmic objects: white dwarfs, neutron stars, and black holes").

REFERENCES

1. M. A. Liberman and B. Johansson, Usp. Fiz. Nauk **165**, 121 (1995) [Phys. Usp. **38**, 117 (1995)].
2. Yu. P. Kravchenko and M. A. Liberman, Phys. Rev. A **57**, 3403 (1998).
3. L. D. Landau and E. M. Lifshitz, *Quantum Mechanics: Non-Relativistic Theory* (Nauka, Moscow, 1989, 4th ed.; Pergamon Press, Oxford, 1977, 3rd ed.), Parag. 112.
4. Yu. P. Kravchenko, M. A. Liberman, and B. Johansson, Phys. Rev. Lett. **77**, 619 (1996).
5. Yu. P. Kravchenko, M. A. Liberman, and B. Johansson, Phys. Rev. A **54**, 287 (1996).
6. A. Y. Potekhin, J. Phys. B **31**, 49 (1998).
7. E. M. Lifshitz, Zh. Éksp. Teor. Fiz. **7**, 390 (1937).
8. S. T. Belyaev, in *Plasma Physics and Thermonuclear Research* (Akad. Nauk SSSR, Moscow, 1958; Pergamon Press, Oxford, 1959), Vol. 3, p. 50–65.
9. V. P. Silin, Zh. Éksp. Teor. Fiz. **38**, 1771 (1960) [Sov. Phys. JETP **11**, 1277 (1960)].
10. V. P. Silin, Zh. Éksp. Teor. Fiz. **41**, 861 (1961) [Sov. Phys. JETP **14**, 617 (1962)].
11. V. P. Silin, Zh. Éksp. Teor. Fiz. **43**, 1813 (1962) [Sov. Phys. JETP **16**, 1281 (1963)].
12. V. P. Silin, in *Introduction to Kinetic Theory of Gases* (Nauka, Moscow, 1971), Chap. 10, p. 276.
13. N. Rostoker, Phys. Fluids **3**, 922 (1960).
14. N. Rostoker and M. N. Rosenbluth, Phys. Fluids **3**, 1 (1960).
15. D. Montgomery, G. Joyce, and L. Turner, Phys. Fluids **17**, 2201 (1974).
16. D. K. Geller and J. C. Weisheit, Phys. Plasmas **4**, 4258 (1997).
17. J. Ventura, Phys. Rev. C **8**, 3021 (1973).
18. G. G. Pavlov and D. G. Yakovlev, Zh. Éksp. Teor. Fiz. **70**, 753 (1976) [Sov. Phys. JETP **43**, 389 (1976)].
19. S. H. Langer, Phys. Rev. D **23**, 328 (1981).
20. M. C. Storey and D. B. Melrose, Aust. J. Phys. **40**, 89 (1987).
21. L. Hernquist, Astrophys. J., Suppl. Ser. **56**, 325 (1984).
22. A. Y. Potekhin, Astron. Astrophys. **306**, 999 (1996).
23. A. Y. Potekhin and D. G. Yakovlev, Astron. Astrophys. **314**, 341 (1996).
24. G. G. Pavlov and A. N. Panov, Zh. Éksp. Teor. Fiz. **71**, 572 (1976) [Sov. Phys. JETP **44**, 300 (1976)].
25. A. W. Hyatt, C. F. Driscoll, and J. H. Malmberg, Phys. Rev. Lett. **59**, 2975 (1987).
26. T. M. O'Neil, Phys. Fluids **26**, 2128 (1983).
27. T. M. O'Neil and P. G. Hjorth, Phys. Fluids **28**, 3241 (1985).
28. M. E. Glinsky, T. M. O'Neil, M. N. Rosenbluth, *et al.*, Phys. Fluids B **4**, 1156 (1992).
29. V. V. Zheleznyakov, in *Radiation in Astrophysical Plasma* (Yanus-K, Moscow, 1997), Parag. 13.1, p. 318.
30. V. V. Zheleznyakov, S. A. Koryagin, and A. V. Serber, Pis'ma Astron. Zh. **25**, 513 (1999) [Astron. Lett. **25**, 437 (1999)].
31. E. M. Lifshitz and L. P. Pitaevskiĭ, *Physical Kinetics* (Nauka, Moscow, 1979; Pergamon Press, Oxford, 1981), Parag. 16.
32. L. D. Landau and E. M. Lifshitz, *Quantum Mechanics: Non-Relativistic Theory* (Nauka, Moscow, 1989, 4th ed.; Pergamon Press, Oxford, 1977, 3rd ed.), Parag. 135.
33. A. P. Prudnikov, Yu. A. Brychkov, and O. I. Marichev, *Integrals and Series* (Nauka, Moscow, 1981; Gordon and Breach, New York, 1986), Vol. 1.
34. *Handbook of Mathematical Functions*, Ed. by M. Abramowitz and I. A. Stegun (Dover, New York, 1971; Nauka, Moscow, 1979).
35. A. I. Morozov and L. S. Solov'ev, in *Reviews of Plasma Physics*, Ed. by M. A. Leontovich (Gosatomizdat, Moscow, 1963; Consultants Bureau, New York, 1966), Vol. 2.
36. G. J. Pert, J. Phys. B **12**, 2755 (1979).
37. G. D. Schmidt, Lect. Notes Phys. **328**, 305 (1989).
38. S. T. Belyaev, in *Plasma Physics and Thermonuclear Research* (Akad. Nauk SSSR, Moscow, 1958; Pergamon Press, Oxford, 1959), Vol. 3, p. 66–85.
39. E. M. Lifshitz and L. P. Pitaevskiĭ, *Physical Kinetics* (Nauka, Moscow, 1979; Pergamon Press, Oxford, 1981), Parag. 60.

Translation was provided by AIP

Journal of Experimental and Theoretical Physics, Vol. 90, No. 5, 2000, pp. 753–768.
Translated from Zhurnal Éksperimental'noĭ i Teoreticheskoĭ Fiziki, Vol. 117, No. 5, 2000, pp. 867–884.

ATOMS, SPECTRA, RADIATION

Scattering of Relativistic Electrons by a Focused Laser Pulse

N. B. Narozhny* and M. S. Fofanov**

Moscow State Engineering Physics Institute (Technical University), Moscow, 115409 Russia
**e-mail: narozhny@pc1k32.mephi.ru*
***e-mail:fofanov@pc1k32.mephi.ru*

Received November 26, 1999

Abstract—The problem of the motion of a classical relativistic electron in a focused high-intensity laser pulse is solved. A new three-dimensional model of the electromagnetic field, which is an exact solution of Maxwell's equations, is proposed to describe a stationary laser beam. An extension of the model is proposed. This extension describes a laser pulse of finite duration and is an approximate solution of Maxwell's equations. The equations for the average motion of an electron in the field of a laser pulse, described by our model, are derived assuming weak spatial and temporal nonuniformities of the field. It is shown that, to a first approximation in the parameters of the nonuniformities, the average (ponderomotive) force acting on a particle is described by the gradient of the ponderomotive potential, but it loses its potential character even in second order. It is found that the three-dimensional ponderomotive potential is asymmetric. The trajectories of relativistic electrons moving in a laser field are obtained and the cross sections for scattering of electrons by a stationary laser beam are calculated. It is shown that reflection of electrons from the laser pulse and the surfing effect are present in the model studied. It is found that for certain impact parameters of the incident electrons the asymmetic ponderomotive potential can manifest itself effectively as an attractive potential. It is also shown that even in the case of a symmetric potential the scattering cross section contains singularities, known as rainbow scattering. The results are applicable for fields characterized by large (compared to 1) values of the dimensionless parameter $\eta^2 = e^2\langle \mathbf{E}^2\rangle/m^2\omega^2$ and arbitrary electron energies. © *2000 MAIK "Nauka/Interperiodica".*

1. INTRODUCTION

The present paper is concerned with the solution of the problem of the motion of a classical relativistic electron in a rapidly oscillating electromagnetic field with amplitude varying slowly in space and time.

In nonrelativistic mechanics this problem can be solved by representing the motion of an electron as a displacement along a smooth trajectory with simultaneous rapid oscillations around it. After averaging over the rapid oscillations the equations for the smooth trajectory assumes the form of Newton's equations with a potential force [1, 2], said to be a ponderomotive force (see, for example, [3]). The ponderomotive potential corresponding to this force is the average kinetic energy of the oscillatory motion [1].

The concept of a ponderomotive potential is widely used to explain many effects in atomic and plasma physics. Specifically, ponderomotive scattering plays a large role in the formation of the spectrum and the angular distribution of electrons in the phenomenon of above-threshold ionization [3, 4] and a variety of nonlinear effects (for example, self-focusing), arising when a laser beam propagates in plasma [5]. Inelastic scattering of nonrelativistic electrons (the surfing effect and reflection) by a ponderomotive potential produced by a comparatively high-power laser pulse in vacuum was first observed experimentally in 1986 [6].

The analysis of the ponderomotive effect in nonrelativistic mechanics assumes that, besides the electron energy being small, the dimensionless parameter

$$\eta^2 = \frac{e^2\langle \mathbf{E}^2\rangle}{m^2\omega^2}, \tag{1}$$

which characterizes the field intensity and determines the energy of the oscillatory motion of an electron [7], is small.[1] Compact optical-frequency lasers with pulse intensity I greater than 10^{18} W/cm^2, which corresponds to $\eta \gtrsim 1$, have now been developed, and experiments on collisions of ultrarelativistic electron beams with laser pulses of this power are being conducted [8, 9]. Consequently, the problem of the interaction of relativistic particles with laser fields of relativistic intensity ($\eta \gtrsim 1$) is now topical.

Kibble [10] was the first to examine the interaction of relativistic electrons with a nonuniform laser field. He obtained equations for the average motion of a relativistic electron in such a field and predicted that electrons will be reflected from the center of the focal region and the surfing effect. Although Kibble's work

[1] We employ the system of units in which the speed of light $c = 1$.

[10] gives a clear picture of the mechanism leading to ponderomotive scattering, it assumed that the intensity of the laser field is small $\eta \ll 1$, which limits the possibility of using this work to describe experiments with laser pulses of relativistic intensity.

The interaction of relativistic electrons with a focused laser beam for $\eta \gtrsim 1$ was first studied in [11, 12]. In these works a field with a two-dimensional configuration was used as a model of the laser pulse. Such a field is a solution of Maxwell's equations, and in principle it can be obtained by using a long cylindrical lens to focus a plane wave. The equations obtained in [11, 12] for the average motion describe qualitatively correctly the characteristic features of the motion of a relativistic electron in a focused pulse, but, of course, they cannot serve to describe a real laser experiment.

In the present paper we propose a new three-dimensional model of a field. This model is a generalization of the model employed in [11, 12]. Our model is an exact solution of Maxwell's equations, it is characterized by the parameters R and L, which can be interpreted as, respectively, the focusing radius and the diffraction length, and it can serve to describe a stationary laser beam. In this model the field configuration is determined by giving certain coordinate functions which satisfy second-order partial differential equations. The choice of various solutions of these equations corresponds to various physical models. Specifically, they can be chosen so that the field corresponds to a superposition of monochromatic plane waves which have the same frequencies and whose wave vectors lies inside a cone with a small aperture angle. Such a field is very close to a model studied in the monograph [13] and obtained by applying the Huygens–Fresnel principle to the problem of Fraunhofer diffraction of a spherical wave by a circular opening. Another choice of solutions describes the Gaussian beams which are widely used in optics [14, 15]. Here, if an ordinary Gaussian envelope can be used to describe a weakly focused field, we propose a generalization of this envelope which is also applicable for a laser beam focused to the diffraction limit. Just as in [11, 12], the proposed model can be generalized to a laser pulse with a finite duration τ. In this case, however, it will be an approximate solution of Maxwell's equations. The next section of this paper is concerned with a discussion of the field model.

The equations of the average motion of an electron in the field of a laser pulse, as described by our model, are derived in Section 3. The derivation employs the standard method of separating the motion into smooth and rapidly oscillating components [1], so that the equations are derived under the assumption that the characteristic field dimensions R and L are much greater than the characteristic wavelength, and the pulse duration is much greater than the reciprocal of the frequency:

$$\omega R \gg 1, \quad \omega L \gg 1, \quad \omega\tau \gg 1. \tag{2}$$

It is also assumed that

$$\omega\tau \gtrsim \omega R. \tag{3}$$

The expression for the ponderomotive force is obtained up to second-order infinitesimals in the parameters

$$\Delta = 1/\omega R, \quad \tilde{\Delta} = 1/\omega\tau. \tag{4}$$

We note, however, that the method which we propose for deriving the equations of motion makes it possible, in principle, to obtain an expression for the ponderomotive force in a stationary laser beam with a prescribed accuracy. This procedure is meaningless for a pulsed field, since this field itself is an approximate solution of Maxwell's equations.

It is shown to a first approximation in the parameters Δ and $\tilde{\Delta}$ that the ponderomotive force is determined by the gradient of the ponderomotive potential, but it loses its potential character even in second order. The most important result of this section is the discovery that the three-dimensional ponderomotive potential is asymmetric, which is the reason for a variety of interesting and, in principle, observable effects which are discussed in the next section.

In Section 4 the derived equations of motion are used to solve the problem of the scattering of relativistic electrons by a laser field. We confine our attention only to a collision geometry in which an electron in the initial state moves in a direction perpendicular to the direction of propagation of the laser beam. The results of a numerical solution of the equations of average motion, including calculations of the trajectories of electrons and a calculation of the cross section for scattering of electrons by a stationary laser beam, are presented. It is shown that for certain impact parameters of the incident electrons an asymmetric ponderomotive potential can effectively manifest itself as an attractive potential. It is also shown that even in the case of a symmetric potential the scattering cross section contains singularities which are known as rainbow scattering [16].

The conditions for applicability of the method developed and the relation between the equations obtained for the average motion and Kibble's equations are discussed in Section 5. A solution of the equations of the field model corresponding to Gaussian beams is presented in the Appendix.

2. FIELD MODEL

We shall consider first a field which is a superposition of monochromatic plane waves with the same frequencies and wave vectors lying inside a cone with aperture angle 2Δ. Choosing a coordinate system with the z axis oriented along the axis of the cone, the vector potential of such a field can be represented in the form

$$\mathbf{A}(\mathbf{r}, t) = \int d^3k \mathbf{A}(\mathbf{k}) \exp[i(\mathbf{k}\cdot\mathbf{r} - \omega t)], \quad A_0 = 0, \tag{5}$$

where $\omega = |\mathbf{k}|$ and

$$\mathbf{A}(\mathbf{k}) = \frac{1}{\pi\Delta^2\omega^2}\mathbf{a}(\vartheta, \alpha)\theta(\Delta^2 - \vartheta^2)\delta(\omega - \omega_0), \qquad \mathbf{k}\cdot\mathbf{a}(\vartheta, \alpha) = 0. \tag{6}$$

Here α is the angle between the projection of the vector $\mathbf{k}$ on a plane perpendicular to the z axis and the x axis, ϑ is the angle between the vector $\mathbf{k}$ and the z axis, and $\theta(x)$ is the Heaviside step function. As $\Delta \longrightarrow 0$ the field obtained passes, as should happen, into the field of a monochromatic plane wave:

$$\mathbf{A}(\mathbf{r}, t) \longrightarrow \mathbf{b}\exp[-i\omega_0(t - z)], \qquad \mathbf{b} = \frac{1}{2\pi}\int_{-\pi}^{\pi} d\alpha\,\mathbf{a}(0, \alpha). \tag{7}$$

We introduce two orthogonal unit vectors:

$$\mathbf{l} = \frac{\mathbf{n}\times\mathbf{e}_z}{\sin\vartheta}, \quad \mathbf{l}\times\mathbf{n}, \tag{8}$$

which form a basis in a plane perpendicular to the vector $\mathbf{k} = \omega\mathbf{n}$, and we expand the vector $\mathbf{a}(\vartheta, \alpha)$ in terms of these unit vectors:

$$\mathbf{a}(\vartheta, \alpha) = a_1(\vartheta, \alpha)\mathbf{l} + a_2(\vartheta, \alpha)\mathbf{l}\times\mathbf{n}. \tag{9}$$

It is easy to see that the Cartesian components of the vector $\mathbf{a}(\vartheta, \alpha)$ can be expressed in terms of the functions a_1 and a_2 as follows:

$$\begin{aligned} a_x &= a_1(\vartheta, \alpha)\sin\alpha - a_2(\vartheta, \alpha)\cos\vartheta\cos\alpha, \\ a_y &= -a_1(\vartheta, \alpha)\cos\alpha - a_2(\vartheta, \alpha)\cos\vartheta\sin\alpha, \\ a_z &= a_2(\vartheta, \alpha)\sin\vartheta. \end{aligned} \tag{10}$$

We shall assume that a_i depend only on the angle α (it is easy to see that they cannot be set equal to constants, since in this case the vector potential of the field vanishes in the limit $\Delta \longrightarrow 0$):

$$a_i = a_i(\alpha). \tag{11}$$

The vector potential (5) and (6) can then be represented in the form (we shall designate the frequency of the field (5) by the letter ω without an index)

$$\begin{aligned} \mathbf{A}(\mathbf{r}, t) = \frac{\exp(i\varphi)}{2\pi} &\times \int_{-\pi}^{\pi} d\alpha\left\{\left[a_1(\alpha)\sin\alpha - a_2(\alpha)\cos\alpha\left(1 - i\Delta^2\frac{\partial}{\partial\chi}\right)\right]\mathbf{e}_x\right. \\ &- \left[a_1(\alpha)\cos\alpha + a_2(\alpha)\sin\alpha\left(1 - i\Delta^2\frac{\partial}{\partial\chi}\right)\right]\mathbf{e}_y \\ &\left. - i\Delta a_2(\alpha)\frac{\partial}{\partial\nu}\mathbf{e}_z\right\}G(\nu, \chi; \Delta), \end{aligned} \tag{12}$$

where

$$\varphi = \omega(t - z), \quad \nu = \omega r\Delta\cos(\alpha - \phi), \quad \chi = \omega z\Delta^2, \quad r = \sqrt{x^2 + y^2}, \quad \cos\phi = \frac{x}{r}, \quad \sin\phi = \frac{y}{r}, \tag{13}$$

and the function $G(\nu, \chi; \Delta)$ is determined by the relation

$$G(\nu, \chi; \Delta) = 2\int_0^1 du\frac{\sin\Delta u}{\Delta} \times \exp\left\{i\nu\frac{\sin\Delta u}{\Delta} - 2i\chi\frac{\sin^2\left(\frac{\Delta u}{2}\right)}{\Delta^2}\right\}. \tag{14}$$

We have for the field strengths

$$\begin{aligned} \mathbf{E} &= i\omega\mathbf{A}, \\ \mathbf{H} &= \frac{i\omega}{2\pi}\exp(-i\varphi) \\ &\times \int_{-\pi}^{\pi} d\alpha\left\{\left[a_1(\alpha)\cos\alpha\left(1 - i\Delta^2\frac{\partial}{\partial\chi}\right) + a_2(\alpha)\sin\alpha\right]\mathbf{e}_x\right. \\ &+ \left[a_1(\alpha)\sin\alpha\left(1 - i\Delta^2\frac{\partial}{\partial\chi}\right) - a_2(\alpha)\cos\alpha\right]\mathbf{e}_y \\ &\left. + i\Delta a_1(\alpha)\frac{\partial}{\partial\nu}\mathbf{e}_z\right\}G(\nu, \chi; \Delta). \end{aligned} \tag{15}$$

We are interested in the case of weak nonuniformity of the laser field, so that the field can be assumed to be close to that of a plane wave. It is obvious that for this the variance in the directions of the wave vectors $\mathbf{k}$ of the monochromatic waves comprising the superposition (5) must be small, i.e.,

$$\Delta \ll 1. \tag{16}$$

The properties of integrals of the type (12) with the function $G(\nu, \chi; 0)$ determined by the relation (14) are well-known in connection with the problem of Fraunhofer diffraction by a circular opening (see, for example, [13]). Specifically, it is known that the field (13) is maximum at the point $\nu = 0$ and $\chi = 0$, and as ν and χ increase, it oscillates with a gradual decrease of the amplitude. The first zeros of the field amplitude arise for $\nu \sim 1$ and $\chi \sim 1$. Thus, the range of variation of the parameters ν and χ

$$|\nu| \lesssim 1, \quad |\chi| \lesssim 1$$

gives the region of space where the field differs most strongly from zero. It is convenient to characterize the dimensions of this region by the parameters

$$R \equiv \frac{1}{\omega\Delta}, \quad L \equiv \frac{1}{\omega\Delta^2}. \tag{17}$$

The parameter R is the focusing radius of the field, and L is sometimes called the diffraction length. In the language of these parameters the condition (16) can be written as

$$\omega R \gg 1, \quad L = \omega R^2 \gg R. \tag{18}$$

This means that the focusing radius must be much greater than the wavelength, and the diffraction length must be much greater than the focusing radius.

It is important to note that the potential (12) is the exact solution of Maxwell's equations for any function $G(\nu, \chi; \Delta)$ satisfying the equation

$$2G'_\chi - i\Delta^2 G''_{\chi\chi} - iG''_{\nu\nu} = 0. \tag{19}$$

Of course, the function (14) is not the only solution of this equation. Consequently, the representation (12) gives an entire class of exact solutions of Maxwell's equations in empty space which, physically, do not generally speaking reduce to the initial model (5) and (6). Of course, in order to be able to treat the field determined by the solution of equation (19) as a model of a focused laser beam, the following conditions must be satisfied:

$$\lim_{\Delta \to 0} G(0, 0; \Delta) = 1, \quad \lim_{\substack{\nu \to \infty \\ \chi \to \infty}} G(\nu, \chi; \Delta) = 0. \tag{20}$$

We shall show below that the Gaussian beams widely used in the literature also can be represented in the form (12).

We shall now examine the question of the polarization of the field (12) and (15). Strictly speaking, the concept of polarization is ordinarily used for fields which possess the transversality property. As we can see from equations (12) and (15), our field is not transverse. Nonetheless, it can always be represented as a combination of two fields for which either the vector **E** or the vector **H** lies in a plane perpendicular to the direction of propagation of the wave. Just as in [13], we shall call such waves, respectively, E and H polarized[2] or simply E and H waves. It is evident from equation (15) that to obtain an expression for E or H polarized waves the function $a_2(\alpha)$ or $a_1(\alpha)$, respectively, must be set to zero in equation (9).

In view of the conditions (18), there exists in the field (12) a spatial region

$$r \ll R, \quad |z| \ll L,$$

which can be called a plane-wave zone, since inside this region the properties of the field (12) are very close to the properties of the field of a plane wave with the potential (7). We shall attribute the polarization of a wave in the plane-wave zone to the field (15) as a whole and, in this sense, we shall talk about, for example, a linearly or circularly polarized E or H wave.

Let $a_i(\alpha)$ be linear functions of $\sin\alpha$ and $\cos\alpha$

$$a_i(\alpha) = c_i \cos\alpha + d_i \sin\alpha, \quad i = 1, 2, \tag{21}$$

where c_i and d_i are constants. After the expressions (10) with $a_i(\alpha)$ from equation (21) are substituted into the second of equations (7) we obtain for the amplitude of the wave in the plane-wave zone

$$b_x = \frac{d_1 - c_2}{2}, \quad b_y = -\frac{c_1 + d_2}{2}. \tag{22}$$

Hence it is evident that for an E wave linearly polarized along the x axis the coefficients c_i and d_i must be chosen in the form

$$d_1 = 2b, \quad c_1 = c_2 = d_2 = 0,$$

and for a circularly polarized E wave in the form

$$c_1 = \mp 2ib, \quad d_1 = 2b, \quad c_2 = d_2 = 0,$$

and so on.

It will be convenient to use in what follows the expressions for the field strengths (15) written in terms of the functions

$$F_1(\xi, \chi; \Delta) = \frac{1}{2\pi}\int_{-\pi}^{\pi} d\alpha G(\xi\cos\alpha, \chi; \Delta),$$
$$F_2(\xi, \chi; \Delta) = \frac{1}{2\pi}\int_{-\pi}^{\pi} d\alpha \cos(2\alpha) G(\xi\cos\alpha, \chi; \Delta), \tag{23}$$

where $\xi = r/R$. For example, in the case of circular polarization we obtain for an E wave

$$\mathbf{H}^e = \pm\omega b \exp(-i\varphi)\left\{\left(1 - i\Delta^2\frac{\partial}{\partial\chi}\right)\right.$$
$$\times [F_1(\mathbf{e}_x \pm i\mathbf{e}_y) + F_2 \exp(\pm 2i\phi)(\mathbf{e}_x \mp i\mathbf{e}_y)]$$
$$\left. + 2i\Delta \exp(\pm i\phi)\frac{\partial F_1}{\partial\xi}\mathbf{e}_z\right\}, \tag{24}$$

$$\mathbf{E}^e = i\omega b \exp(-i\varphi)$$
$$\times \{F_1(\mathbf{e}_x \pm i\mathbf{e}_y) - F_2 \exp(\pm 2i\phi)(\mathbf{e}_x \mp i\mathbf{e}_y)\}.$$

The expressions for the field strengths of a circularly polarized H wave can be calculated according to the formulas

$$\mathbf{E}^h = \pm i\mathbf{H}^e, \quad \mathbf{H}^h = \mp i\mathbf{E}^e. \tag{25}$$

[2] In waveguide optics E and H polarized waves are called, respectively, H and E type waves. See, for example, [17].

The field strengths (24) and (25) are once again exact solutions of Maxwell's equations, if the function F_1 satisfies the equation

$$2i\frac{\partial F_1}{\partial \chi} + \Delta^2\frac{\partial^2 F_1}{\partial \chi^2} + \frac{1}{\xi}\frac{\partial}{\partial \xi}\left(\xi\frac{\partial F_1}{\partial \xi}\right) = 0, \tag{26}$$

and the function F_2 is expressed in terms of F_1 as follows:

$$F_2 = F_1 - \frac{2}{\xi^2}\int d\xi \xi F_1. \tag{27}$$

These relations follow from the definition of the functions F_1 and F_2 (23) in terms of the function G and equation (19). In accordance with the conditions (20), the functions F_1 and F_2 must approach zero as $\xi, |\chi| \longrightarrow \infty$ and, in addition, they must satisfy the conditions

$$\lim_{\Delta \to 0} F_1(0, 0; \Delta) = 1, \quad \lim_{\Delta \to 0} F_2(0, 0; \Delta) = 0. \tag{28}$$

We note that one of the solutions of equations (26) and (27) for $\Delta \ll 1$ can be written in the form (see Appendix)

$$F_1 = (1 + 2i\chi)^{-2}\left\{1 - \frac{\xi^2}{1 + 2i\chi}\right\}\exp\left\{-\frac{\xi^2}{1 + 2i\chi}\right\},$$
$$F_2 = -\xi^2(1 + 2i\chi)^{-3}\exp\left\{-\frac{\xi^2}{1 + 2i\chi}\right\}. \tag{29}$$

Waves with envelopes of this type are ordinarily called Gaussian beams (compare [14, 15]).

The field which we have considered thus far can be used as a model of a stationary laser beam. Proceeding by analogy to [11], equations can be obtained for the field which models a short laser pulse. For this, the following substitutions must be made in equations (24) and (25):

$$\exp(-i\varphi) \longrightarrow if'(\varphi), \quad \Delta\exp(-i\varphi) \longrightarrow \Delta f(\varphi), \tag{30}$$

where the function $f(\varphi)$ contains, besides an ordinary phase factor, the temporal envelope $g(\varphi/\omega\tau)$ of the pulse

$$f(\varphi) = g(\varphi/\omega\tau)\exp(-i\varphi). \tag{31}$$

The envelope g is assumed to be 1 at the center of the pulse, $g(0) = 1$, and exponentially decreasing for $|\varphi| \gg \omega\tau$. Here τ is the duration of the laser pulse in the laboratory coordinate system. We assume that

$$\tau \gtrsim R. \tag{32}$$

Such a field is no longer an exact solution of Maxwell's equations. But, as can be easily check directly, it will satisfy Maxwell's equations to terms of order $1/(\omega R)^2$ and $(1/\omega\tau)(1/\omega R)$, inclusively.

In closing this section we note that the dimensionless field-intensity parameter determined by the relation (1) can be written for circularly polarized E and H waves for $\Delta \ll 1$ in the form

$$\eta^2 = \frac{e^2b^2}{m^2}g^2\left(\frac{\varphi}{\omega\tau}\right)\{|F_1|^2 + |F_2|^2\}. \tag{33}$$

As should be the case, in the field of a focused laser pulse the parameter η depends on the coordinates and time and reaches its maximum value η_0 at the center of the pulse. By virtue of the relations (28) and the properties of the temporal envelope $g(\varphi/\omega\tau)$

$$\eta_0^2 = \frac{e^2b^2}{m^2}. \tag{34}$$

3. EQUATIONS OF AVERAGE MOTION

We write the equations of the motion of an electron in the form

$$\frac{d\pi^\mu}{ds} = \frac{e}{m}F^\mu_\nu\pi^\nu, \quad m\frac{dx^\mu}{ds} = \pi^\mu, \quad \pi^2 = m^2, \tag{35}$$

where π is the kinetic 4-momentum of an electron and s is the proper time of the electron. The tensor F^μ_ν for the present model of the field is a function of the phase $\varphi = (kx)$ of the wave, where the 4-vector k^μ in the laboratory system is

$$k^\mu = (\omega, 0, 0, \omega)$$

and the spatial coordinates x, y, and z

$$F^\mu_\nu = F^\mu_\nu\left(\varphi; \frac{x}{R}, \frac{y}{R}, \frac{z}{kR^2}\right). \tag{36}$$

Just as for a monochromatic plane wave (see, for example, [7]), it is convenient to switch in the equations of motion (35) from the proper time s to the variable φ, which we shall call the "phase time." The relation

$$\frac{d\varphi}{ds} = \frac{k\pi}{m} \tag{37}$$

follows trivially from equations (35). Using this relation we rewrite equation (35) in the form

$$\frac{d\pi^\mu}{d\varphi} = \frac{e}{k\pi}F^\mu_\nu\pi^\nu, \quad \pi^2 = m^2, \tag{38}$$

$$\frac{dx^\mu}{d\varphi} = \frac{\pi^\mu}{k\pi}. \tag{39}$$

We shall seek the solutions of these equations in the form

$$\begin{cases} \pi = q(\varphi) + \pi', \quad x = x^{(0)}(\varphi) + x', \\ \pi' = \sum\limits_{n=1}^{\infty}(\pi^{(n)}(\varphi)\cos(n\varphi) + \tilde{\pi}^{(n)}(\varphi)\sin(n\varphi)), \\ x' = \sum\limits_{n=1}^{\infty}(x^{(n)}(\varphi)\cos(n\varphi) + \tilde{x}^{(n)}(\varphi)\sin(n\varphi)), \end{cases} \tag{40}$$

where $q(\varphi)$, $x^{(0)}(\varphi)$, $\pi^{(n)}(\varphi)$, $\tilde{\pi}^{(n)}(\varphi)$, $x^{(n)}(\varphi)$, and $\tilde{x}^{(n)}(\varphi)$ are slowly varying functions of φ.

A search for a solution in this form corresponds to separation of the electron motion into a regular displacement along a smooth trajectory and rapid oscillations around the trajectory with frequencies which are a multiple of the frequency of the external field. The form of the smooth electron trajectory $x^{(0)}(\varphi)$, just as the quantity $q(\varphi)$, can be obtained by averaging, respectively, the functions $x(\varphi)$ and $\pi(\varphi)$ over the rapid oscillations. Consequently, we shall call $x^{(0)}(\varphi)$ the trajectory of the average motion, and $q(\varphi)$ the average kinetic momentum. However, it should be kept in mind that the proposed method makes it possible to obtain a solution of the equations of motion (38) and (39) in the form (40) with a prescribed accuracy without using any averaging procedures. Of course, this assertion makes sense if the external field is an exact solution of Maxwell's equations.

It goes without saying that a representation of the solution of the equations of motion (38) and (39) in the form (40) is justified only if the problem contains two substantially different time scales. Specifically, the time of flight of an electron over distances of the order of the dimensions of the field nonuniformities, which in our case are determined by the focusing radius R and the diffraction length L, should be much greater than the period of the wave. For this, of course, the conditions (18) must be satisfied. For a relativistic particle, however, certain additional conditions must be satisfied. We shall assume that these conditions are satisfied in the following calculations; we shall establish their specific form later.

We note that for the field of a monochromatic plane wave the solutions of the equations of motion can also be represented in the form (40). Then the quantity q and all coefficients in the oscillating parts of the functions $\pi(\varphi)$ and $x(\varphi)$ are constants, and only one harmonic will be present in them for a circularly polarized wave and only two harmonics will be present for a linearly polarized wave (see [7]).

Using the representation (40) we write the electromagnetic field tensor (36) in the form

$$F = F\left(\varphi;\ \rho_x + \frac{x'}{R},\ \rho_y + \frac{y'}{R},\ \zeta + \frac{z'}{\omega R^2}\right), \tag{41}$$

where we have introduced the notation

$$\frac{x^{(0)}}{R} = \rho_x, \quad \frac{y^{(0)}}{R} = \rho_y, \quad \frac{z^{(0)}}{\omega R^2} = \zeta. \tag{42}$$

We shall assume that the amplitude of the particle oscillations is the same in order of magnitude as the wavelength of the external field (the corresponding restriction on the parameter η_0 is discussed in Section 5). As a result, the following estimates, which calculations confirm, hold:

$$\frac{x'}{R} \sim \frac{\lambda}{R} \sim \Delta, \quad \frac{z'}{\omega R^2} \sim \frac{\lambda}{\omega R^2} \sim \Delta^2. \tag{43}$$

Consequently, for $\Delta \ll 1$ the amplitude of the intensity of the laser field changes very little during the oscillations of particles around the trajectory of the average motion and the electromagnetic field tensor can be expanded in a series:

$$F\left(\varphi;\ \rho_x + \frac{x'}{R},\ \rho_y + \frac{y'}{R},\ \zeta + \frac{z'}{\omega R^2}\right) \approx F(\varphi;\ \rho,\ \psi,\ \zeta) + F'_{\rho_x}\frac{x'}{R} + F'_{\rho_y}\frac{y'}{R} + F'_{\zeta}\frac{z'}{\omega R^2} + \frac{1}{2!}F''_{\rho_x\rho_x}\left(\frac{x'}{R}\right)^2 + \ldots. \tag{44}$$

Here we have introduced the notation

$$\frac{\rho_x}{\rho} = \cos\psi, \quad \frac{\rho_y}{\rho} = \sin\psi, \quad \rho = \sqrt{\rho_x^2 + \rho_y^2} = \frac{r^{(0)}}{R}. \tag{45}$$

Substituting the expansion (44) into equation (38), using the representation (40), and equating the coefficients of $\cos(n\varphi)$ and $\sin(n\varphi)$ for each n on the right- and left-hand sides of equations (38), we obtain a system of equations for determining the coefficient functions in equations (40). We shall illustrate the proposed method for a laser pulse modeled by a circularly polarized E wave (24) with the temporal envelope (31). As we have already noted, such a field is an approximate solution of Maxwell's equations, so that for this case it makes sense to retain in the equations of motion (38) and (39) terms of order no higher than second in the parameters Δ and $\tilde{\Delta}$. For conciseness, we shall obtain the equations of the average motion taking into account first-order terms only, and we shall present the second-order terms without derivation.

Substituting into the equation for the kinetic momentum (38) the expressions for the real parts of the field intensities of a circularly polarized E wave (24) with the temporal envelope (31) and retaining only the terms which are first-order in the parameters Δ and $\tilde{\Delta}$ we obtain the equations

$$\frac{d\pi_x}{d\varphi} = -ebg\,\mathrm{Im}(e^{-i\varphi}A_-) - ebg'\tilde{\Delta}\,\mathrm{Re}(e^{-i\varphi}A_-) \mp ebg\Delta\frac{\pi_y}{\pi_-}\mathrm{Im}(e^{-i\varphi}K) - ebg\frac{x'}{R}\mathrm{Im}\left\{e^{-i\varphi}\left[\cos\psi\frac{\partial A_-}{\partial\rho} \pm i\sin\psi\frac{2F_2}{\rho}e^{\pm 2i\psi}\right]\right\}$$

$$-ebg\frac{y'}{R}\mathrm{Im}\left\{e^{-i\varphi}\left[\sin\psi\frac{\partial A_-}{\partial\rho}\mp i\cos\psi\frac{2F_2}{\rho}e^{\pm 2i\psi}\right]\right\},$$

$$\frac{d\pi_y}{d\varphi} = \mp ebg\mathrm{Re}(e^{-i\varphi}A_+)\pm ebg'\tilde{\Delta}\mathrm{Im}(e^{-i\varphi}A_+) \quad (46)$$

$$\pm ebg\Delta\frac{\pi_x}{\pi_-}\mathrm{Im}(e^{-i\varphi}K)$$

$$\mp ebg\frac{x'}{R}\mathrm{Re}\left\{e^{-i\varphi}\left[\cos\psi\frac{\partial A_+}{\partial\rho}\mp i\sin\psi\frac{2F_2}{\rho}e^{\pm 2i\psi}\right]\right\}$$

$$\mp ebg\frac{y'}{R}\mathrm{Re}\left\{e^{-i\varphi}\left[\sin\psi\frac{\partial A_+}{\partial\rho}\pm i\cos\psi\frac{2F_2}{\rho}e^{\pm 2i\psi}\right]\right\},$$

$$\frac{d\pi_-}{d\varphi} = 0.$$

Here $F_i = F_i(\rho, \zeta, \Delta)$, and

$$A_+ = F_1 + F_2e^{\pm 2i\psi}, \quad A_- = F_1 - F_2e^{\pm 2i\psi},$$
$$K = 2\frac{\partial F_1}{\partial\rho}e^{\pm i\psi}. \quad (47)$$

Equating on both sides of equations (46) and (39) terms which do not contain rapidly oscillating functions, we obtain equations of the average motion which can be written in the form

$$\frac{dq_x}{d\varphi} = \mp\frac{ebg\Delta}{2q_-}\mathrm{Im}\{(\pi_y^{(1)} - i\tilde{\pi}_y^{(1)})K\}$$

$$-\frac{ebg}{2R}\mathrm{Im}\left\{(x^{(1)} - i\tilde{x}^{(1)})\left[\cos\psi\frac{\partial A_-}{\partial\rho}\pm i\sin\psi\frac{2F_2}{\rho}e^{\pm 2i\psi}\right]\right\}$$

$$-\frac{ebg}{2R}\mathrm{Im}\left\{(y^{(1)} - i\tilde{y}^{(1)})\left[\sin\psi\frac{\partial A_-}{\partial\rho}\mp i\cos\psi\frac{2F_2}{\rho}e^{\pm 2i\psi}\right]\right\},$$

$$\frac{dq_y}{d\varphi} = \pm\frac{ebg\Delta}{2q_-}\mathrm{Im}\{(\pi_x^{(1)} - i\tilde{\pi}_x^{(1)})K\} \quad (48)$$

$$\mp\frac{ebg}{2R}\mathrm{Re}\left\{(x^{(1)} - i\tilde{x}^{(1)})\left[\cos\psi\frac{\partial A_+}{\partial\rho}\mp i\sin\psi\frac{2F_2}{\rho}e^{\pm 2i\psi}\right]\right\}$$

$$\mp\frac{ebg}{2R}\mathrm{Re}\left\{(y^{(1)} - i\tilde{y}^{(1)})\left[\sin\psi\frac{\partial A_+}{\partial\rho}\pm i\cos\psi\frac{2F_2}{\rho}e^{\pm 2i\psi}\right]\right\}$$

$$\frac{dx^{i(0)}}{d\varphi} = \frac{q^i}{\omega q_-}. \quad (49)$$

Here we employed the fact that to the accuracy which we require, as follows from the third of equations (46),

$$\pi_- = q_- = \text{const}. \quad (50)$$

As one can see from equations (48), the coefficient functions entering there must be known with accuracy up to zero order in the parameter Δ. The expressions for the coefficient functions $\pi^{(1)}$ and $\tilde{\pi}^{(1)}$ can be easily obtained by equating the coefficients of $\cos\varphi$ and $\sin\varphi$ on the right- and left-hand sides of equations (46). The derivatives of $\pi^{(1)}$ and $\tilde{\pi}^{(1)}$ must be dropped, since they are at least first-order infinitesimals in the parameter Δ. As a result, we have

$$\pi_x^{(1)} - i\tilde{\pi}_x^{(1)} = -ebgA_-^*,$$
$$\pi_y^{(1)} - i\tilde{\pi}_y^{(1)} = \pm iebgA_+^*. \quad (51)$$

The expressions for the coefficient functions $x^{(1)}$, $\tilde{x}^{(1)}$ and $y^{(1)}$, $\tilde{y}^{(1)}$ can be obtained similarly after substituting the expansions (40) into equation (39)

$$x^{(1)} - i\tilde{x}^{(1)} = i\frac{ebg}{\omega q_-}A_-^*,$$
$$y^{(1)} - i\tilde{y}^{(1)} = \pm\frac{ebg}{\omega q_-}A_+^*. \quad (52)$$

We note that, compared with the functions found, all other coefficient functions in the expansions (40) are higher-order infinitesimals, in complete agreement with the fact that, as already noted, only one harmonic is present in the solutions of the equations of motion in the limit of a circularly polarized plane wave.

Substituting the expressions (51) and (52) into equations (48), we obtain finally for the equation of the average motion in a circularly polarized field, up to terms of order Δ,

$$\frac{dq_x}{d\varphi} = -\frac{m}{\omega q_-}\frac{\partial U}{\partial x^{(0)}} = -\Delta\frac{m}{q_-}\frac{\partial U}{\partial\rho}\cos\psi,$$
$$\frac{dq_y}{d\varphi} = -\frac{m}{\omega q_-}\frac{\partial U}{\partial y^{(0)}} = -\Delta\frac{m}{q_-}\frac{\partial U}{\partial\rho}\sin\psi, \quad (53)$$

$$\frac{dx^{(0)}}{d\varphi} = \frac{q_x}{\omega q_-}, \quad \frac{dy^{(0)}}{d\varphi} = \frac{q_y}{\omega q_-}, \quad \frac{dz^{(0)}}{d\varphi} = \frac{q_z}{\omega q_-}, \quad (54)$$

where U, given by

$$U = \frac{e^2b^2}{2m}g^2\left(\frac{\varphi}{\omega\tau}\right)\left(|F_1|^2 + |F_2|^2\right) = \frac{m\eta^2}{2} \quad (55)$$

can be called the ponderomotive potential by analogy to the nonrelativistic case.

It is simplest to determine the quantity q_z appearing in equation (54) by averaging the z component of the kinetic momentum

$$q_z = \langle \pi_z \rangle = \frac{m^2 + \langle \pi_\perp^2 \rangle}{2q_-} - \frac{q_-}{2} = \frac{m_*^2 + \mathbf{q}_\perp^2}{2q_-} - \frac{q_-}{2}, \quad (56)$$

where we have introduced the quantity m_*

$$m_*^2 = m^2 + \langle \boldsymbol{\pi}_\perp'^2 \rangle, \quad (57)$$

which can be called the effective mass of an electron in an external field. Using the relation (51) it is easy to show that the effective mass is related with the ponderomotive potential as

$$m_*^2 = m^2 + 2mU = m^2(1 + \eta^2). \quad (58)$$

Differentiating the expression (56) for q_z with respect to φ and using the equations of motion (53) and (54) and the expression for the ponderomotive potential (55), we obtain

$$\frac{dq_z}{d\varphi} = \frac{m}{q_-}\frac{\partial}{\partial \varphi} U\left(\frac{\varphi}{\omega\tau}, \mathbf{r}^{(0)}(\varphi)\right) = \frac{e^2b^2}{2q_-}\frac{d}{d\varphi}\left[g^2\left(\frac{\varphi}{\omega\tau}\right)\right](|F_1|^2 + |F_2|^2). \quad (59)$$

It is evident from the relation (59) that in a stationary beam, where $g(\varphi/\omega\tau) = 1$, the z component of the average kinetic momentum and by virtue of the relation (50) the average energy $q_0 = q_- + q_z$ are conserved in our approximation. In a pulsed field these quantities are not conserved. This is the so-called surfing effect.

The calculations show that in the approximation under study the equations of the average motion (53) and (54), together with the relation (56), retain their form in an external field (15) with any polarization and differ only by the form of the ponderomotive potential. Specifically, the ponderomotive potential for a circularly polarized H wave is identical to the expression (55), and for a circularly polarized mixed-type wave, obtained as a result of a superposition of E and H waves, it can be qualitatively different. For example, if the coefficients in equation (21) are chosen in the form

$$a_1 = 2b\sin\alpha, \quad a_2 = \mp 2ib\sin\alpha, \quad (60)$$

then the field (15), according to the criteria formulated in the preceding section, remains circularly polarized but the ponderomotive potential for it will assume the form

$$U = \frac{e^2b^2}{2m} g^2\left(\frac{\varphi}{\omega\tau}\right)\left\{|F_1 - F_2|^2 \cos^2\psi + |F_1 + F_2|^2 \sin^2\psi \mp i(F_1F_2^* - F_1^*F_2)\sin 2\psi\right\}. \quad (61)$$

An important feature of the potential (61) is its asymmetry, i.e., the dependence on the azimuthal angle ψ. The ponderomotive potential for linearly or elliptically polarized E or H waves possesses the same property. The fact that the expression (61) is substantially different from equation (55) (although the field determined by the coefficients (60) is a linear superposition of E and H waves) is, of course, due to the nonlinearity of the equations of motion (38).

We shall now present the equations of the average motion, up to terms of second order, inclusively, with respect to the parameters Δ and $\tilde{\Delta}$, for a circularly polarized E wave. In cylindrical coordinates they have the form

$$\dot{q}_\rho - q_\psi\dot{\psi} = f_\rho, \quad \dot{q}_\psi + q_\rho\dot{\psi} = f_\psi, \quad \dot{q}_z = f_z,$$
$$\dot{\rho} = \frac{\Delta}{q_-}q_\rho, \quad \rho\dot{\psi} = \frac{\Delta}{q_-}q_\psi, \quad \dot{\zeta} = \frac{\Delta^2}{q_-}q_z, \quad (62)$$

where an overdot indicates differentiation with respect to the phase time φ,

$$f_\rho = -\Delta\frac{m}{q_-}\frac{\partial U}{\partial \rho} \mp \Delta^2\frac{q_\psi}{2q_-^2}V,$$
$$f_\psi = \pm 2\Delta\tilde{\Delta}\frac{g'}{g}\frac{m}{q_-}\frac{\partial W}{\partial \rho} \pm \Delta^2\frac{q_\rho}{2q_-^2}V, \quad (63)$$
$$f_z = \frac{m}{q_-}\frac{\partial U}{\partial \varphi} - \Delta^2\frac{m}{q_-}\frac{\partial U}{\partial \zeta} \pm 4\Delta\tilde{\Delta}\frac{g'}{g}\frac{mq_\psi}{q_-^2}\frac{\partial W}{\partial \rho} \mp \Delta^2\frac{mq_\psi q_\rho}{q_-^3}\left(\frac{1}{\rho}\frac{\partial W}{\partial \rho} - \frac{\partial^2 W}{\partial \rho^2}\right),$$

and W and V are determined by the following relations:

$$W = \frac{e^2b^2g^2}{2m}(|F_1|^2 - |F_2|^2),$$
$$V = e^2b^2g^2\left\{\left(2 + \rho\frac{\partial}{\partial \rho}\right)(\Upsilon_{1\rho}'^2 + \Upsilon_{2\rho}'^2) - 2(\Phi_2\Upsilon_{1\zeta}' - \Upsilon_1\Phi_{2\zeta}' + \Upsilon_2\Phi_{1\zeta}' - \Phi_1\Upsilon_{2\zeta}')\right\}. \quad (64)$$

The functions Φ and Υ in equation (64) are linear combinations of F_1 and F_2:

$$\Phi = F_1 + F_2 = \Phi_1 + i\Phi_2,$$
$$\Upsilon = F_1 - F_2 = \Upsilon_1 + i\Upsilon_2. \quad (65)$$

The components of the vector $\mathbf{f}$ in equations (62) are, to within a constant factor, the phase-time-averaged spatial components of the four-vector of the force acting on an electron in the field of a laser pulse. We shall call this force the relativistic ponderomotive force. The formulas (63) give its expansion with respect

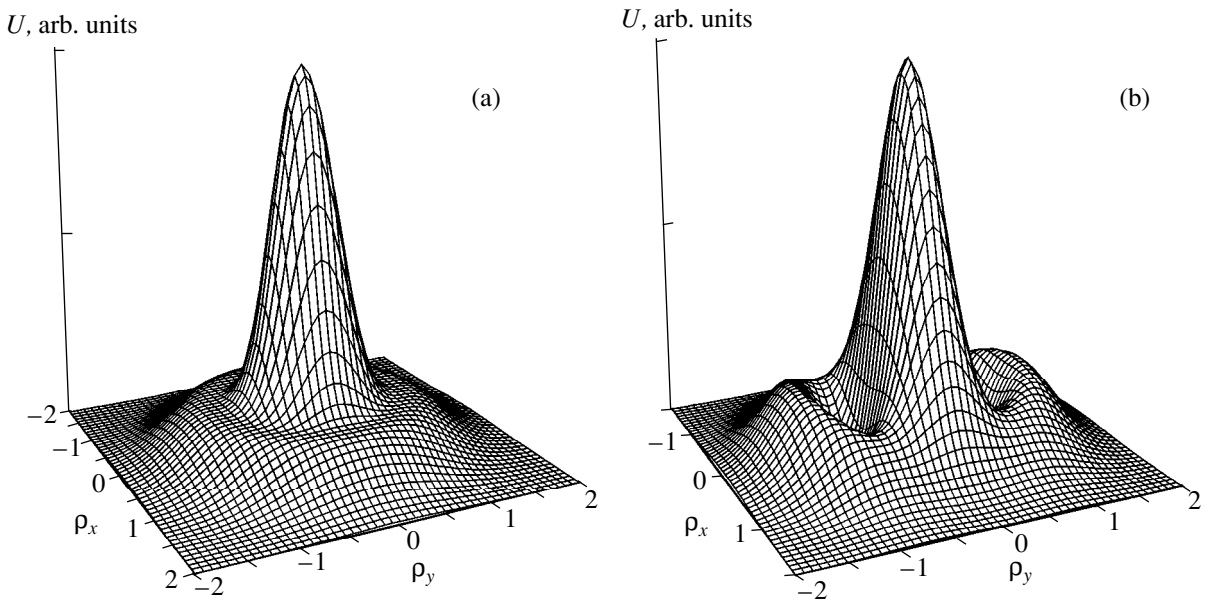


Fig. 1. The ponderomotive potential (a) in a circularly polarized *E* or *H* wave and (b) a wave of mixed type.

to the parameters Δ and $\tilde{\Delta}$. To a first approximation the relativistic ponderomotive force is determined by the gradient of the ponderomotive potential. For the symmetric potential (55), of course, it possesses only a radial component. For an asymmetric potential, for example, determined by the expression (61), a nonzero azimuthal component f_ψ appears even in the first approximation. In the second approximation the ponderomotive force is not of a potential character, as one can see from equations (63). The azimuthal component appears even for a circularly polarized *E* wave and, in addition, a longitudinal force component arises. The presence of a longitudinal component of the relativistic ponderomotive force was first discovered in [11] (see also [18]). In an *E* wave the longitudinal component arises because of averaging of the corresponding component of the Lorentz force. In the general case, the average longitudinal component of the electric field also contributes to it.

4. SCATTERING OF AN ELECTRON BY A LASER PULSE

In this section we present the results of a numerical solution of the equations of the average motion (53) and (54) and calculations of the cross section for scattering of electrons in the field of a focused laser pulse (15). As one can see from equations (53) and (54), the trajectory of an electron is determined, to a first approximation, entirely by the form of the ponderomotive potential. We shall consider both the symmetric (55) and asymmetric (61) potentials. As noted in the preceding section, the first case pertains to circularly polarized *E* and *H* waves, and therefore they are the same in the present approximation of the trajectory of an electron colliding with such waves. The second case determines the trajectory of an electron in the field of a circularly polarized wave of a mixed type. However, the specific form of the dependence of the asymmetric potential on the variable ψ has no effect on the qualitative features of the electron motion, and for this reason the qualitative results obtained for this case pertain equally to linearly polarized *E* and *H* waves.

The numerical calculations were performed for Gaussian beams (29). Plots of the potentials (55) and (61) in the $z = 0$ plane with the functions F_1 and F_2 determined by equations (29) are shown in Figs. 1a, 1b, respectively. For a collision of electrons with a pulsed field the temporal envelope of the pulse was chosen to be of the form

$$g\left(\frac{\varphi}{\omega\tau}\right) = \exp\left[-\left(\frac{\varphi}{\omega\tau}\right)^2\right]. \qquad (66)$$

The collision geometry was chosen so that the electrons are incident on the laser beam in a direction perpendicular to the direction of propagation of the beam.

Figure 2 shows the trajectory of electrons in the field of a stationary laser beam with the ponderomotive potential (61). For the collision geometry under consideration here, the initial longitudinal momentum $q_z = 0$. Since to a first approximation the longitudinal component of the average kinetic momentum is conserved (see equation (59)), the electron trajectory is flat to the same accuracy. Figure 2 shows electron trajectories intersecting a laser beam in the $z = 0$ plane, i.e., in a plane passing through the focal center. All electrons in this figure have the same initial energy but different impact parameters. For clarity, the level lines of the ponderomotive potential (61) are shown.

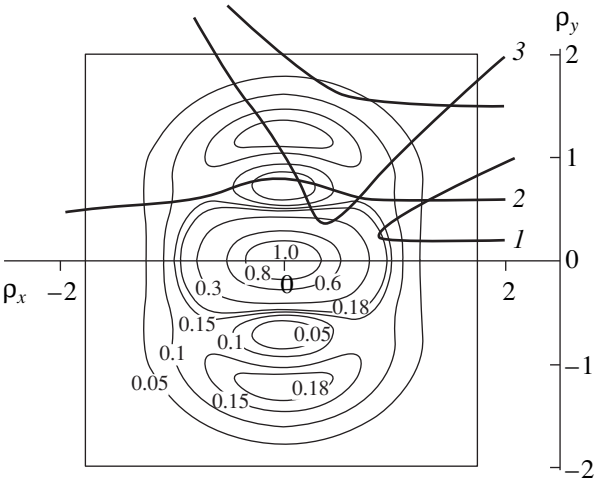


Fig. 2. Trajectories of the average motion of electrons (gamma factor $\gamma = 1.2$) in a circularly polarized field of mixed type with intensity $\eta_0 = 1$ and various values of the impact parameter of the particles.

It is evident from Fig. 2 that for small impact parameters, when the incident electron falls into a region near the maximum of the ponderomotive potential, particles are reflected from the laser beam (trajectory *1*), as first predicted by Kibble [10] and then observed experimentally in [6]. Naturally, this effect occurs only for particles with a comparatively low energy. As the energy increases, electrons with any impact parameter will be able to overcome the potential barrier and the reflection effect will not be observed.

The condition under which reflection will occur when an electron is scattered by a symmetric potential, can be obtained in an explicit form. As one can easily see from the equations of motion (62) written in cylindrical coordinates, to a first approximation, besides the energy, the projection of the angular momentum on the *z* axis is also conserved:

$$M_z = \rho q_\psi = \frac{p}{R}\left|q_\rho^{in}\right| = \text{const}, \qquad (67)$$

where p is the impact parameter and q_ρ^{in} is the initial value of the radial component of the electron momentum. It should be kept in mind that because of the possible asymmetry of the ponderomotive potential, positive (corresponding to $q_\psi^{in} > 0$) and negative (corresponding to $q_\psi^{in} < 0$) impact parameters should be distinguished. Then, we have from the energy conservation $q_0 = q_- + q_z$, using equations (56) and (67),

$$\frac{q_0^2 - q_z^2 - m^2}{2m} = \frac{q_\rho^2}{2m} + \frac{M_z^2}{2m\rho^2} + U(\rho) = \text{const}. \qquad (68)$$

(It is interesting to note that the second part of the equation (68) is identical to the expression for the energy of a nonrelativistic particle in a central field; see, for example, [1].) It is obvious that the most stringent conditions will be the conditions for a particle to pass by with zero impact parameter or with $M_z = 0$. Then it follows from the relation (68) that all particles with initial momentum

$$(q_\rho^{in})^2 > 2mU_{\max}$$

will overcome the potential barrier. Therefore the condition for reflection during scattering of electrons by a symmetric potential can be written as

$$\gamma = q_0/m < \sqrt{1 + \eta_0^2} \qquad (69)$$

and is identical to the corresponding condition obtained in [12] for the two-dimensional field model. For an asymmetric potential, because angular momentum is not conserved, the equations of motion will have to be solved completely in order to obtain the reflection condition in an explicit form. Such a solution cannot be obtained analytically. However, as we shall see below, the corresponding restriction on the energy, though not identical to the condition (69), is of the same order of magnitude.

It should be noted that for an asymmetric potential an electron with certain impact parameters is deflected in the direction of the focus (trajectory *2* in Fig. 2), i.e., for such values of p the potential is effectively attractive. The existence of such trajectories is easy to understand. It is clear from Fig. 1b that an asymmetric potential contains, besides the main maximum, a series of additional maxima at the periphery, though these maxima are not as high. Consequently, for certain impact parameters an electron could feel more strongly not the main maximum but rather the additional maximum, which, in addition, could lie to the right of the main maximum. This is the reason that an electron is deflected in the "wrong" direction.

We note that all trajectories shown in Fig. 2, except for the trajectory *3*, correspond to particles incident on a laser beam along the *x* axis. The trajectory *3* refers to an electron incident at an angle $\pi/4$ to the *x* axis and characterizes another feature of an asymmetric potential: the fact that even a particle with a zero impact parameter can be scattered by a nonzero angle. This is due to the presence of a nonzero azimuthal component of the ponderomotive force

$$f_\psi = -\Delta\frac{m}{q_-}\frac{1}{\rho}\frac{\partial U}{\partial \psi}.$$

We note that for $z = 0$ for the potential (61), $f_\psi = 0$ at $\psi = 0$ and $\pi/2$. Consequently, trajectories of the type *3* are not observed if the incident particle initially moves along the *x*- or *y*-axis.

Figures 3 and 4 show the scattering angle χ_{sc} as a function of the impact parameter p for different energies and angles of incidence of the electrons for symmetric (55) (dashed line) and asymmetric (61) (solid

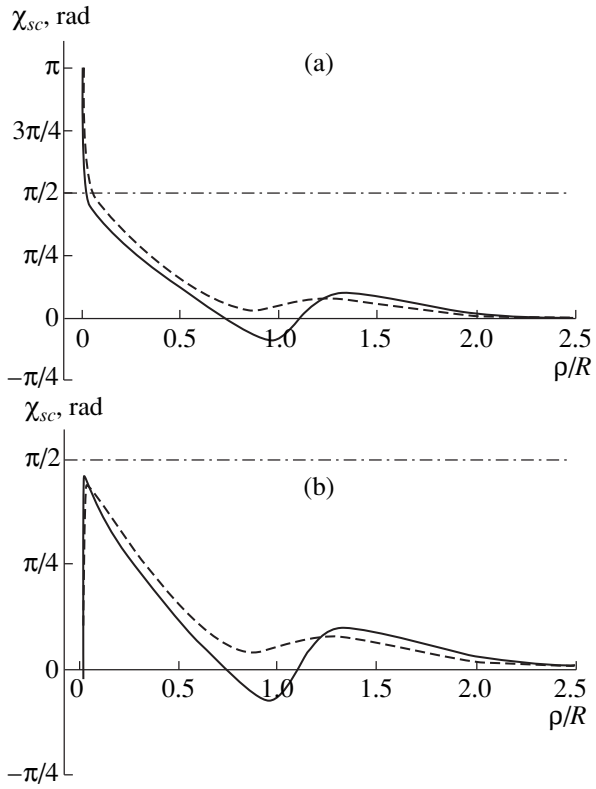


Fig. 3. The scattering angle χ_{sc} as a function of the impact parameter of an electron in asymmetric (solid line) and symmetric (dashed line) potentials with incidence along the x axis: $\eta_0 = 1$, $\gamma = 1.4$ (a) and 1.42 (b).

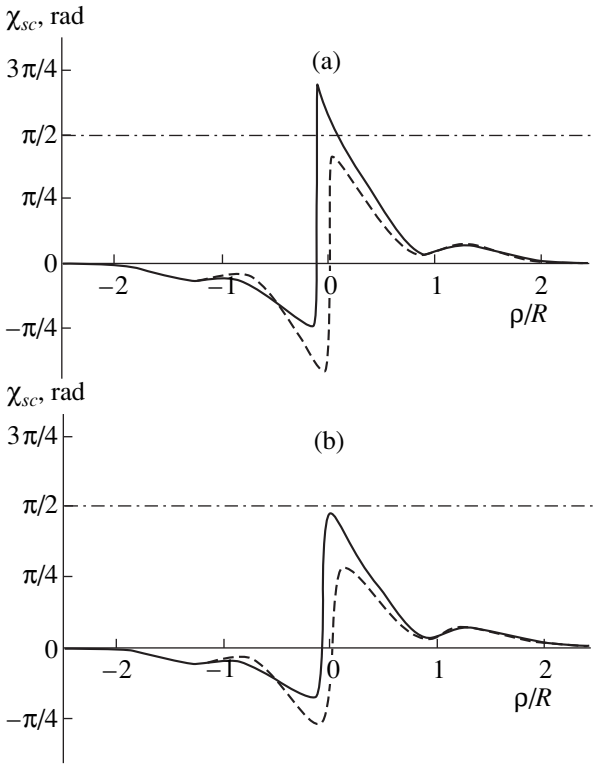


Fig. 4. The scattering angle χ_{sc} as a function of the electron impact parameter in asymmetric (solid line) and symmetric (dashed line) potentials with incidence at angle $\psi_0 = \pi/4$ to the x axis: $\eta_0 = 1$, $\gamma = 1.42$ (a) and 1.48 (b).

line) potentials. Figure 3a corresponds to electrons incident along the x axis with energy somewhat less than ($\eta_0 = 1$, $\gamma = 1.40$) and Fig. 3b corresponds to electrons with energy somewhat greater than ($\eta_0 = 1$, $\gamma = 1.42$) the critical value for a symmetric potential

$$\gamma = \sqrt{1 + \eta_0^2} = \sqrt{2}.$$

For lower energies the scattering angle in both cases for some values of p is greater than $\pi/2$. This attests to the existence of reflection of an electron from the laser focus, and for higher energies the scattering angle for $p = 0$ is zero and does not reach $\pi/2$ for any value of the impact parameter. This indicates the absence of reflection. Hence it follows that the reflection criterion for an asymmetric potential (61) with a zero angle of incidence is essentially identical to the potential (69). As the impact parameter increases, the scattering angle changes nonmonotonically, which is a consequence of the complicated structure of the potential. In addition, for an asymmetric potential the scattering angle for a definite range of impact parameters assumes negative values. As follows from the preceding discussion, this effect can be explained by reflection from the additional maxima.

Figure 4 shows χ_{sc} as a function of p for electrons incident on a beam with an asymmetric potential along the line $\psi_0 = \pi/4$ for the same value $\eta_0 = 1$, and $\gamma = 1.42$ (Fig. 4a) and 1.48 (Fig. 4b). In contrast to Fig. 3, the values of χ_{sc} for positive and negative impact parameters are presented. It is evident from the plots that for $\gamma = 1.42$, in contrast to the case of incidence along the x axis, reflection for an asymmetric potential occurs, while for $\gamma = 1.48$ there is no reflection. This means that for an asymmetric potential the reflection criterion depends on ψ_0, but this dependence is quite weak. We call attention to the fact that the curves $\chi_{sc}(p)$ are symmetric with respect to the point $p = 0$ for a ponderomotive potential (55), while for the potential (61) there is no such symmetry. Moreover, the maximum scattering angle for an asymmetric potential corresponds to not a zero but rather a negative impact parameter. This is explained by the existence, already discussed above, of an azimuthal component of the ponderomotive force. A particle incident on an asymmetric potential with zero impact parameter p (i.e., moving along the trajectory *3*, Fig. 2) does not reach the maximum of the potential. This happens for a particle with negative p, which is reflected in Fig. 4. We also note that for incidence at an angle $\psi_0 = \pi/4$ scattering of particles with positive impact parameters by negative angles is not observed.

The computed functions $\chi_{sc}(p)$ make it possible to find the effective cross sections for scattering of elec-

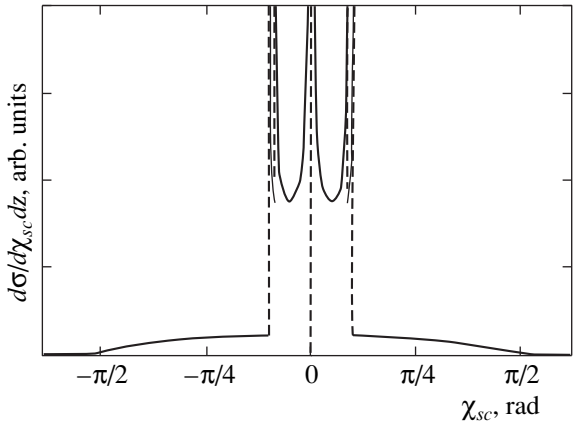


Fig. 5. The differential scattering cross section for electrons in a laser beam with an asymmetric ponderomotive potential as a function of the angle χ_{sc}: $\psi_0 = 0$, $\gamma = 1.4$, and $\eta_0 = 1$.

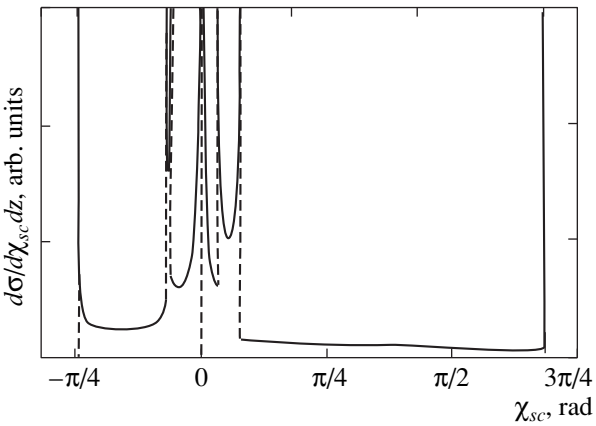


Fig. 6. The differential scattering cross section for electrons in a laser beam with an asymmetric ponderomotive potential as a function of the angle χ_{sc}: $\psi_0 = \pi/4$, $\gamma = 1.42$, and $\eta_0 = 1$.

trons by a laser field. If the collision of a beam of electrons with a laser field occurs in a region close to the focal plane, and the diameter of the beam cross section is small compared with the diffraction length L, then it can be assumed that the target (laser field) possesses cylindrical symmetry. In this case it is possible to introduce a differential scattering cross section calculated per unit length:

$$\frac{d\sigma}{d\chi_{sc}dz} = \sum_i \left|\frac{dp_i(\chi_{sc})}{d\chi_{sc}}\right|, \tag{70}$$

where the summation extends over all branches of the function $p(\chi_{sc})$.

Figure 5 shows the differential scattering cross section for electrons in a laser field with an asymmetric potential (61) as a function of the scattering angle χ_{sc}. The parameters $\psi_0 = 0$, $\eta_0 = 1$, and $\gamma = 1.4$ correspond to the function $\chi_{sc}(p)$ shown in Fig. 3a. It is evident from Fig. 5 that the cross section is singular for definite values of the scattering angle. These singularities of the cross section (with the exception of the trivial singularity at zero, due to the fact that the scattering by small angles corresponds to infinite impact parameters) are related with the presence of extrema on the curve $\chi_{sc}(p)$ and are called rainbow scattering (see, for example, [16]). The cross section corresponding to the case shown in Fig. 3b, when there is no reflection, differs from the case shown in Fig. 5 only by the presence of rainbow scattering with one additional value of χ_{sc}, which corresponds to the maximum scattering angle.

As one can easily see from Figs. 3a, 3b, rainbow scattering also occurs for the potential (55). Small differences occur because scattering by negative angles is impossible for a symmetric potential. But these differences do not greatly change the overall picture, and therefore we do not present here the corresponding dependences. The dependence of the scattering cross section on the parameter ψ_0 strongly distinguishes the case of an asymmetric potential. This is easy to see by comparing Figs. 5 and 6. We note that besides "deformation" the cross section shown in Fig. 6 differs from Fig. 5 by the presence of rainbow scattering at the left- and right-hand boundaries. This is explained by the fact that for negative impact parameters for the case shown in Fig. 6 reflection does not occur at all, while for positive impact parameters the maximum deflection angle, though exceeding $\pi/2$, is less than π (see Fig. 4a), i.e., backscattering is impossible.

We shall now examine a collision of an electron with a laser pulse of duration τ. In this case the longitudinal component of the electron momentum q_z is not conserved and the trajectory of an electron incident on the laser pulse in a direction perpendicular to the propagation of the pulse is not planar. As follows from equation (59), the electron is displaced in the positive or negative direction along the z axis from the direction of incidence, depending on whether the electron in a collision falls on the leading or trailing edge of the laser pulse. In this case the electron energy also is not conserved. Since in the first approximation the quantity q_- once again remains an integral of the motion, the change in energy is given by the same formula (59). Consequently, if a particle interacts with the leading edge of the focus, the energy of the particle increases, and if it interacts with the trailing edge, the particle energy decreases. This is the so-called surfing effect [6, 10].

The total change in energy can be calculated according to the formula

$$\delta q_0 = \frac{m}{q_-}\int_{-\infty}^{\infty} d\varphi \frac{\partial}{\partial\varphi} U\left(\frac{\varphi}{\omega\tau}, \mathbf{r}^0(\varphi)\right). \tag{71}$$

It can be positive as well as negative, depending on the intervals of phase time that the particle spent on the leading and trailing edges of the laser pulse. Figure 7 shows the electron energy versus the delay $\delta\varphi$. Here the

delay is the interval of phase time between the moment when an electron is located at the origin, if the electron has not interacted with the laser pulse, and the moment when the focal center of the laser pulse is located at the same point. For large positive (negative) values of $\delta\varphi$, the electron energy does not change. This is due to the fact that the electron crosses the region of the focus before (after) the arrival of the laser pulse, not interacting with it. As $\delta\varphi$ decreases, the electron starts to feel the leading edge of the laser pulse. During the motion the particle acquires a potential energy, which converts into kinetic energy when the particle leaves the light pulse. On account of this, the particle energy increases after the passage of the pulse. As $\delta\varphi$ decreases further, the reverse process occurs: the electron is on the trailing edge of the laser pulse and is slowed down by the pulse. The nonmonotonic dependence of the energy for positive as well as negative values of $\delta\varphi$ can be explained by the complicated structure of the potential.

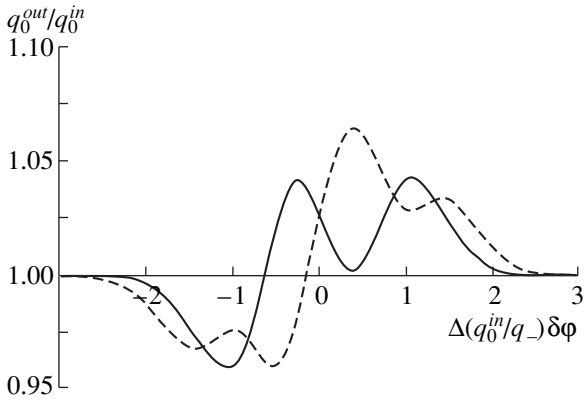


Fig. 7. The electron energy after interaction with a short laser pulse as a function of the delay $\delta\varphi$. The solid line corresponds to an asymmetric and the dashed line to a symmetric ponderomotive potential: $\psi_0 = 0$, $\eta = 1$, $\gamma = 1.2$, $p/R = 0.5$, and $\tau/R = 1$.

5. CONDITIONS OF APPLICABILITY OF THE METHOD AND THE KIBBLE EQUATIONS

We shall show that, to a first approximation, the equations which we have obtained for the average motion are formally identical to the Kibble equations. For this, it is sufficient to show that the equations for q_z and q_0 can be written in a form similar to equation (53). We recall that since

$$q_- = q_0 - q_z = \text{const}$$

both equations have the same form (see equation (59)),

$$\frac{dq_z}{d\varphi} = \frac{dq_0}{d\varphi} = \frac{m}{q_-}\frac{\partial}{\partial\varphi}U\left(\frac{\varphi}{\omega\tau}, \mathbf{r}^{(0)}(\varphi)\right). \tag{72}$$

Just as in the derivation of equations (54), it is easy to show that the time $t^{(0)}$ averaged over φ is determined by the equation

$$\frac{dt^{(0)}}{d\varphi} = \frac{q_0}{\omega q_-}. \tag{73}$$

Then, it is obvious that

$$\varphi = \omega(t^{(0)} - z^{(0)}). \tag{74}$$

Substituting this expression into the equation for the ponderomotive potential on the right-hand sides of equations (72), we rewrite these equations in the form

$$\begin{aligned}\frac{dq_z}{d\varphi} &= -\frac{m}{\omega q_-}\frac{\partial}{\partial z^{(0)}}U(x^{(0)}),\\ \frac{dq_0}{d\varphi} &= \frac{m}{\omega q_-}\frac{\partial}{\partial t^{(0)}}U(x^{(0)}).\end{aligned} \tag{75}$$

Switching now to differentiation with respect to the proper time s, using the relation (37), which to a first approximation is

$$d\varphi/ds = \omega q_-/m,$$

we rewrite the complete system of equations for the average motion (54), (53), (75), and (73) in the form

$$\frac{dq^\mu}{ds} = \frac{\partial}{\partial x_\mu^{(0)}}U(x^{(0)}), \quad m\frac{dx_\mu^{(0)}}{ds} = q_\mu, \tag{76}$$

which is identical to the Kibble equations in [10]. However, the derivation of the Kibble equations in [10] made substantial use of, besides the assumption that the nonuniformity of the field is weak, the condition that the field is weak or, in other words,

$$\eta \ll 1.$$

In our method the solution of the equations of motion in the field of a monochromatic plane wave of arbitrary intensity was used as the zeroth approximation, and only the parameters of the spatial Δ and temporal $\tilde{\Delta}$ nonuniformities were assumed to be small.

We shall now discuss in greater detail the conditions of applicability of equations (53), (54), or (76) obtained in the present paper. They are determined by two requirements. In the first place, the phase time for an electron to traverse distances of the order of the characteristic dimensions of the field nonuniformities must be much greater than the period of the oscillations. In the second place, the amplitude of the transverse particle oscillations must be much less than the focal radius R, and the amplitude of the longitudinal oscillations must be much less than the diffraction length L.

The phase times φ_R and φ_L, in which an electron traverses distances of the order of the radius R of the focus and the diffraction length L, respectively, can be easily estimated from equations (54) as

$$\varphi_R \sim \frac{\omega R q_-}{q_\perp}, \quad \varphi_L \sim \frac{(\omega R)^2 q_-}{|q_z|},$$

where $q_\perp$ is the absolute magnitude of the component of the average momentum perpendicular to the direction of propagation of the laser pulse. Then the first condition can be written in the form

$$\Delta\frac{q_\perp}{q_-} \ll 1, \quad \Delta^2\frac{|q_z|}{q_-} \ll 1. \tag{77}$$

The amplitudes of the transverse and longitudinal oscillations can be estimated using the solutions of the equations of motion of an electron in the field of a monochromatic plane wave (see, for example, [7]). They are of the order of $m\eta_0/\omega q_-$ and $m\eta_0 q_\perp/\omega q_-^2$, respectively. (We note that the first estimate follows immediately from equations (52).) It follows from these estimates and equations (77) that the second condition can be written as

$$\eta_0 \ll \frac{q_-}{m\Delta} \tag{78}$$

and is therefore the limit on the intensity of the laser field (compare with equation (17) of [12]).

It is easy to see that the conditions (77) and (78) depend, generally speaking, on the collision geometry. Since the parameter Δ is small, the conditions (77) are satisfied for any value of the transverse $\mathbf{q}_\perp$ and longitudinal q_z momenta of the electron, if $q_z < 0$. Then

$$q_- \gtrsim m, \quad |q_z|/q_- < 1, \quad q_\perp/q_- < 1.$$

For $q_z > 0$ it is possible to have $q_- \ll m_*$. This clearly occurs for a collision geometry in which the initial electron momentum is oriented almost in the direction of propagation of the laser pulse (parallel motion), and the electron itself is ultrarelativistic. This case requires a separate analysis.

For values of q_- which are not small, the condition on the field intensity (78) is quite soft, and by virtue of equation (2) it is possible to study quite strong fields, including fields with $\eta_0 \gg 1$. This circumstance distinguishes our equations for the average motion from the Kibble equations, even though they are formally identical in the first approximation. In this connection we call attention to [19], where an endeavor is made to derive an expression for the ponderomotive force by averaging the Lagrangian of a relativistic particle in a nonuniform electromagnetic field of arbitrary intensity. The equations obtained in [19] for the average motion are identical to our equations (and hence for $\eta_0 \ll 1$ they are also identical with the Kibble equations) only in the nonrelativistic limit, where the average kinetic momentum of a particle is small compared to the particle mass. This result is explained by the inconsistent and therefore incorrect method of averaging employed in [19]. In terms of our paper, in averaging over the phase time we assumed the coefficients of the function in the expansions (40) to be functions of the laboratory time t.

We shall now consider the possibility of using our equations for the problem of parallel motion of an ultrarelativistic electron. Let

$$q_z > 0, \quad q_z \gg m_*, q_\perp.$$

Under these conditions the quantity q_- is related with q_z as follows:

$$q_- \approx \frac{m_*^2 + q_\perp^2}{2|q_z|} \ll |q_z|. \tag{79}$$

Substituting this expression for q_- into equations (77) we obtain the condition of applicability of our method to the problem of parallel motion, which can be written as

$$\sqrt{m_*^2 + q_\perp^2} \ll |q_z| \ll \frac{1}{\Delta}\sqrt{m_*^2 + q_\perp^2}. \tag{80}$$

The relations (78) and (79) give a limit on the field

$$\eta_0 \ll \frac{m_*^2 + q_\perp^2}{m\Delta|q_z|}. \tag{81}$$

We note that the problem of parallel motion without the limits (80) and (81) was studied in [20] using computer simulation.

In conclusion we note that equations (76) are explicitly covariant, thereby proving that our method is covariant. This is not surprising, since we use an invariant averaging procedure and consequently

$$q_\mu = \langle \pi_\mu \rangle, \quad x_\mu^{(0)} = \langle x_\mu \rangle$$

are 4-vectors, and the ponderomotive force is determined (see equation (58)) in terms of the invariant effective mass $m_*^2 = q^2$ and therefore it too is a relativistic invariant. However, it should be remembered that the meaning of the parameters Δ and $\tilde{\Delta}$, which in the laboratory coordinate system are, respectively, the ratios of the wave length of the field to the focusing radius and the period of the field to the pulse duration, generally speaking, will change in an arbitrary coordinate system.

ACKNOWLEDGMENTS

We are grateful to N.B. Delone and the participants of his seminar for a discussion of the results obtained in this work. We are especially grateful to S.P. Goreslavskiĭ and M.V. Fedorov for many helpful remarks and suggestions.

This work was partially supported by the Russian Foundation for Basic Research (project no. 97-02-16973) and the Ministry of Education of the Russian Federation (project no. 97-5.2-117).

APPENDIX

Let us consider the first of equations (26). The substitution $\sigma = \xi^2$ changes the equation into the form

$$2i\frac{\partial F_1}{\partial \chi} + \Delta^2\frac{\partial^2 F_1}{\partial \chi^2} + 4\frac{\partial F_1}{\partial \sigma} + 4\sigma\frac{\partial^2 F_1}{\partial \sigma^2} = 0. \quad \text{(A.1)}$$

We seek the solution of this equation in the form

$$F_1(\sigma, \chi; \Delta) = \int_{-\infty}^{\infty} d\chi' f(\chi, \chi')h_1(\sigma, \chi'). \quad \text{(A.2)}$$

Substituting the expression (A.2) into equation (A.1) gives

$$\int_{-\infty}^{\infty} d\chi'\left\{ f(\chi, \chi')\left[4\frac{\partial h_1(\sigma, \chi')}{\partial \sigma} + 4\sigma\frac{\partial^2 h_1(\sigma, \chi')}{\partial \sigma^2}\right]\right. \quad \text{(A.3)}$$
$$\left. + \left[2i\frac{\partial f(\chi, \chi')}{\partial \chi} + \Delta^2\frac{\partial^2 f(\chi, \chi')}{\partial \chi^2}\right]h_1(\sigma, \chi')\right\} = 0.$$

We now require that the function $f(\chi, \chi')$ satisfy the equation

$$2i\frac{\partial f(\chi, \chi')}{\partial \chi} + \Delta^2\frac{\partial^2 f(\chi, \chi')}{\partial \chi^2} = -2i\frac{\partial f(\chi, \chi')}{\partial \chi'}. \quad \text{(A.4)}$$

Then, assuming $f(\chi, \chi')h_1(\sigma, \chi') \longrightarrow 0$ as $\chi' \longrightarrow \pm\infty$ and integrating equation (A.3) by parts gives

$$\int_{-\infty}^{\infty} d\chi' f(\chi, \chi')\left\{4\frac{\partial h_1(\sigma, \chi')}{\partial \sigma}\right. \quad \text{(A.5)}$$
$$\left. + 4\sigma\frac{\partial^2 h_1(\sigma, \chi')}{\partial \sigma^2} + 2i\frac{\partial h_1(\sigma, \chi')}{\partial \chi'}\right\} = 0,$$

and in order that the integral (A.2) be the solution of equation (A.1) the function $h_1(\sigma, \chi')$ must satisfy equation (A.1) without the term with the second derivative with respect to χ.

We consider equation (A.4) first. Let $f(\chi, \chi') = f(\mu, \chi')$, where $\mu = \chi - \chi'$. Then equation (A.4) can be rewritten as

$$2i\frac{\partial f(\mu, \chi')}{\partial \chi'} = -\Delta^2\frac{\partial^2 f(\mu, \chi')}{\partial \mu^2}. \quad \text{(A.6)}$$

This equation can be easily solved by separation of variables. We choose its solution in the form

$$f(\mu, \chi') = \frac{1}{2i\Delta}\exp[i(\mu/\Delta - \chi'/2)]. \quad \text{(A.7)}$$

The choice of a normalization factor is justified below.

First, we consider the equation for $h_1(\sigma, \chi')$:

$$4\frac{\partial h_1(\sigma, \chi')}{\partial \sigma} + 4\sigma\frac{\partial^2 h_1(\sigma, \chi')}{\partial \sigma^2} + 2i\frac{\partial h_1(\sigma, \chi')}{\partial \chi'} = 0. \quad \text{(A.8)}$$

We seek its solution in the form

$$h_1(\sigma, \chi') = [u_1(\chi') + \sigma u_2(\chi')]\exp[-v^2(\chi')\sigma].$$

Substituting this expression into equation (A.8) and equating the coefficients of like powers of σ, we obtain the following equations for the functions u_1, u_2, and v:

$$iv' = v^3, \quad iu_2' = 6u_2v^2, \quad 4u_2 - 4u_1v^2 + 2iu_1' = 0.$$

These equations can be solved trivially, and we have for the function h_1 the result

$$h_1(\sigma, \chi') = v^4(\chi')[1 - \sigma v^2(\chi')]e^{-v^2(\chi')\sigma}, \quad \text{(A.9)}$$

where

$$v(\chi') = (1 + 2i\chi')^{-1/2}. \quad \text{(A.10)}$$

We note that as $\chi' \longrightarrow \pm\infty$ the function (A.9) does indeed vanish, which justifies the method of solution proposed above.

Substituting the expressions (A.7) and (A.9) into equation (A.2), we obtain a particular solution of equation (A.1) in the form

$$F_1(\sigma, \chi; \Delta) = \frac{1}{2i\Delta}e^{-i\chi/2} \quad \text{(A.11)}$$
$$\times \int_{-\infty}^{\infty} d\mu e^{i\mu/\Delta}e^{i\mu/2}h_1(\sigma, \chi - \mu).$$

It is obvious that for $\Delta \ll 1$ the neighborhood of the point $\mu = 0$ makes the main contribution to the integral in equation (A.11). Therefore, to calculate the asymptotic form of the function $F_1(\sigma, \chi; \Delta)$ for $\Delta \ll 1$ it is convenient to represent the expression (A.11) in the form

$$F_1(\sigma, \chi; \Delta) = \frac{1}{2i\Delta}e^{-i\chi/2}$$
$$\times\left\{\int_0^{\infty} d\mu e^{i\mu/\Delta}e^{i\mu/2}h_1(\sigma, \chi - \mu)\right. \quad \text{(A.12)}$$
$$\left. + \int_0^{\infty} d\mu e^{-i\mu/\Delta}e^{-i\mu/2}h_1(\sigma, \chi + \mu)\right\},$$

after which the standard formula for estimating the asymptotic behavior of the Fourier integrals [21] can be applied to each of the integrals appearing on the right-hand side of equation (A.12). The result is

$$F_1 = e^{-i\chi/2}(1 + 2i\chi)^{-2}\left\{1 - \frac{\sigma}{1 + 2i\chi}\right\}\exp\left\{-\frac{\sigma}{1 + 2i\chi}\right\}, \quad \text{(A.13)}$$
$$\sigma = \xi^2.$$

We can see that the expression (A.13) differs from the first of equations (29) only by a phase factor. However, since the variables χ and z are related (see equation (13)), the presence of the phase factor reduces to redefinition of the phase of the field intensities (24)

$$\varphi = \omega(t - z) \longrightarrow \omega t - \omega(1 + \Delta^2/2)z,$$

which is of no consequence in our approximation. For this reason, the phase factor in equation (A.13) should be dropped, and the asymptotic expression for the function (A.11) for $\Delta \ll 1$ is identical to the first of equations (29).

Using the relation (27), we obtain for the function $F_2(\xi, \chi; \Delta)$ a representation that differs from equation (A.11) by the function h_1 being replaced by h_2, which is identical to the expression on the right-hand side of the second of equations (29). It is easy to see that the corresponding asymptotic formula is identical to the function F_2 from equation (29).

In conclusion, we note that, of course, equation (29) for F_1 can be obtained directly by solving the first of equations (26) with the term containing Δ^2 being dropped, just as we did in order to find the function h_1. Nonetheless, we presented a detailed derivation of equation (A.11), since it (together with the corresponding expression for F_2) serves as an extension of the Gaussian envelope to the case $\Delta \sim 1$ and therefore is of interest in itself.

REFERENCES

1. L. D. Landau and E. M. Lifshitz, *Mechanics* (Nauka, Moscow, 1973; Pergamon Press, Oxford, 1976).
2. A. V. Gaponov and M. A. Miller, Zh. Éksp. Teor. Fiz. **34**, 242 (1958).
3. M. V. Fedorov, *Electron in a Strong Optical Field* (Nauka, Moscow, 1991).
4. N. B. Delone and M. V. Fedorov, Usp. Fiz. Nauk **158**, 215 (1989) [Sov. Phys. Usp. **32**, 500 (1989)].
5. M. Tabak, J. Hammer, M. E. Glinsky, *et al.*, Phys. Plasmas **1**, 1626 (1994).
6. P. H. Bucksbaum, M. Bashkansky, and T. J. McIlrath, Phys. Rev. Lett. **58**, 349 (1987).
7. L. D. Landau and E. M. Lifshitz, *The Classical Theory of Fields* (Nauka, Moscow, 1973; Pergamon Press, Oxford, 1975).
8. C. Bula, K. T. McDonald, E. J. Prebys, *et al.*, Phys. Rev. Lett. **76**, 3116 (1996).
9. D. L. Burke, R. C. Field, G. Horton-Smith, *et al.*, Phys. Rev. Lett. **79**, 1626 (1997).
10. T. W. B. Kibble, Phys. Rev. **150**, 1060 (1966).
11. S. P. Goreslavskiĭ, N. B. Narozhny, and V. P. Yakovlev, Laser Phys. **1**, 670 (1991).
12. S. P. Goreslavskiĭ, N. B. Narozhny, O. V. Shcherbachev, and V. P. Yakovlev, Laser Phys. **3**, 418 (1993).
13. M. Born and E. Wolf, *Principles of Optics* (Pergamon Press, Oxford, 1969; Mir, Moscow, 1973).
14. G. S. Landsberg, *Optics* (Nauka, Moscow, 1976).
15. N. V. Karlov, *Lecturs in Quantum Electronics* (Nauka, Moscow, 1988).
16. N. F. Mott and H. S. W. Massey, *The Theory of Atomic Collisions* (Clarendon Press, Oxford, 1965; Mir, Moscow, 1969).
17. L. D. Landau and E. M. Lifshitz, *Electrodynamics of Continuous Media* (Nauka, Moscow, 1982; Pergamon, Oxford, 1984).
18. D. R. Bituk and M. V. Fedorov, Zh. Éksp. Teor. Fiz. **116**, 1198 (1999) [JETP **89**, 640 (1999)].
19. D. Bauer, P. Mulser, and W.-H. Steeb, Phys. Rev. Lett. **75**, 4622 (1995).
20. F. V. Hartemann, S. N. Fochs, G. P. Le Sage, *et al.*, Phys. Rev. E **51**, 4833 (1995).
21. A. Erdelyi, *Asymptotic Expansions* (Dover, New York, 1956; Gostekhizdat, Moscow, 1962).

Translation was provided by AIP

Journal of Experimental and Theoretical Physics, Vol. 90, No. 5, 2000, pp. 769–777.
From Zhurnal Éksperimental'noĭ i Teoreticheskoĭ Fiziki, Vol. 117, No. 5, 2000, pp. 885–894.

ATOMS, SPECTRA, RADIATION

Towards the Fluorescence Resonance Energy Transfer (FRET) Scanning Near-Field Optical Microscopy: Investigation of Nanolocal FRET Processes and FRET Probe Microscope¶

S. K. Sekatskiĭ*, *, G. T. Shubeita*, M. Chergui*, G. Dietler*, B. N. Mironov**, D. A. Lapshin**, and V. S. Letokhov****

**Institut de Physique de la Matiére Condensée, Université de Lausanne, CH-1015, Lausanne-Dorigny, Switzerland*

***Institute of Spectroscopy, Russian Academy of Sciences, Troitsk, Moscow oblast, 142092 Russia*

****e-mail: sekats@lls.isan.troitsk.ru*

Received November 30, 1999

Abstract—The fluorescence resonance energy-transfer (FRET) process is investigated between donor dye molecules deposited on the sample surface and acceptor dye molecules deposited on the tips of scanning near-field and atomic force microscopes. The FRET process was observed only when the tip acquired contact with the sample and took place in regions of sizes of only a few tens of nanometers with only a few thousands (or even hundreds) of molecules involved. The dependence of the FRET intensity on the tip-sample acting force is recorded and interpreted. In relation to the obtained results, the construction of a previously proposed one-atom FRET SNOM is described. © *2000 MAIK "Nauka/Interperiodica".*

1. INTRODUCTION

Scanning near-field optical microscopy (SNOM) is a valuable research tool for imaging and investigating different samples with a subwavelength spatial resolution. The spatial resolution of SNOM is usually limited by the size of the aperture for light transmission and ranges from 50 to 100 nm, although a 20 nm resolution has been demonstrated [1, 2]. Further improvement of the resolution seems problematic for the "classical" SNOM configurations because the number of photons "seeping" through an aperture is rapidly decreasing with the decrease of the aperture size. A number of new approaches have been proposed recently to improve the resolution, such as the molecular exciton-based SNOM [3], apertureless SNOM [4], and SNOM using Fluorescence Resonant Energy Transfer (FRET) between a single fluorescence center of the tip and the sample under study [5].

In the latter case, the idea is based on the fact that when the distance between donor and acceptor molecules becomes smaller than the characteristic radius of a resonant energy transfer R_0 (which for typical donor–acceptor pairs ranges within 2–6 nm [6]), the probability of a dipole–dipole energy transfer between these molecules is close to unity (see, for instance, papers [6, 7] for a review). One should prepare the tip containing a single fluorescent center in the apex and scan it in close proximity to the sample surface (the relative distance should be smaller than R_0). If the donor fluorescent centers of the imaging tip are excited and the fluorescence of the acceptor centers of the sample is monitored (or vice versa), the spatial resolution will be governed not by the aperture size of the microscope but by the value of R_0. An analysis shows that not only the spatial resolution, but the sensitivity as well can be improved when using these FRET SNOMs [5, 8], which, of course, would be very important for the subsequent progress in the field.

The applicability of SNOM to detect a single molecule fluorescence is well established at present (see, e.g., recent reviews [9, 10] and references therein) and the possibility of the nondestructive scanning of the SNOM tips in the close proximity of the sample surface (in the contact mode) has been demonstrated [11, 12]. In this paper, we present the first experimental evidence of the applicability of FRET phenomena for near-field optical microscopy: a nanolocal resonant energy transfer process has been observed between two different dyes. One of them (the donor) has been deposited onto the glass sample surface and other (the acceptor) has been deposited onto the surface of a SNOM tip (sharpened optical fiber) or a standard AFM silicon nanotip. The FRET process has been realized only when the tip acquires a contact with the sample, i.e., in the regions with the sizes of only a few tens of nanometers, and it involves only thousands (or even hundreds) of dye molecules.

¶ This article was submitted by the authors in English.

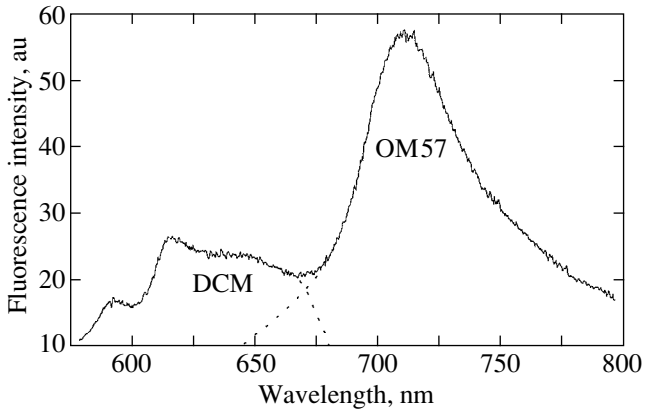


Fig. 1. Fluorescence spectrum of codeposited submonolayers of DCM and OM57 dyes.

A part of these results has been briefly discussed earlier in the letter [13]. A recent paper by Vickery and Dunn [14], where first images obtained with a FRET SNOM (without an analysis of the signal as a function of the tip-sample acting force) should also be mentioned in relation with the described problem.

2. SELECTION OF A DONOR–ACCEPTOR PAIR

A careful selection of a donor and acceptor dye molecule pair was necessary for the experiments described. When the laser excitation radiation wavelength is fixed (we have selected the 488 nm line of a cw argon ion laser), the donor molecules to be used should efficiently absorb this laser radiation and reemit light with a sufficiently large Stokes shift and a high quantum efficiency. The acceptor molecules to be used should efficiently absorb the photons reemitted by the donor, (i.e., good overlapping of the corresponding fluorescence and absorption spectra is required) and should also exhibit a high fluorescence yield with a large red shift with respect to the donor fluorescence. In addition, their direct excitation by the laser radiation should be minimal in order to diminish the background fluorescence and facilitate the observation of a nanolocal FRET phenomenon.

DCM dye molecules (4-dicyanomethilene-2-methyl-6-(p-dimethylaminoatyryl)-4H-pyran, number LC 6500 in Spectra Physics GmbH catalogue [15]) have been selected as donors because of their excellent fluorescent properties (the fluorescence quantum yield in solutions is close to unity, the absorption cross section value σ at the 488 nm wavelength is 6×10^{-17} cm^2) and high photostability.

Different dyes have been tested as acceptors. The best results have been obtained when using 1-butyl-3,3-dimethyl-2-[5-(1-butyl-3,3-dimethyl-3H-benz[e]indolin-2-yliden)-1,3-pentadienyl]-3H-benz[e]indolium perchlorate molecules (OM57 dye, Al'pha Akonis Company, Moscow): their absorption spectrum corresponds well to the fluorescence spectrum of DCM, and their absorption at the 488 nm wavelength is at least three orders of magnitude smaller than at the maximum; these molecules also have a reasonable fluorescence quantum yield (no smaller than 0.3) and photostability.

In Fig. 1, we present the spectrum of fluorescence of the two dyes, DCM and OM57, codeposited onto the same glass slide with the surface concentrations 3×10^{13} cm^{-2}. Such a concentration corresponds to a submonolayer coating: as known for Rhodamine dyes, one monolayer coating corresponds to the surface concentration $\sim 10^{14}$ cm^{-2} [16]. It is clear from this figure that under such conditions, the fluorescence of OM57 molecules (the spectral range 650–800 nm) is even more prominent than that of DCM (the spectral range 550–700 nm), keeping in mind that OM57 molecules do not absorb the excitation wavelength (the fluorescence spectrum of OM57 molecules, deposited in the same concentration but without DCM molecules on a glass slide, was orders of magnitude less intense and barely exceeded the noise level). Thus, this figure can be regarded as a demonstration of the dipole–dipole resonant energy transfer process between DCM and OM57 dye molecules on the surface.

For some other pairs of donor and acceptor molecules (DCM–DTDCI, DCM–HITCI, see [15] for the description of these dyes), the FRET process has been also observed but was not so prominent and the donor and acceptor fluorescence spectra were not so well resolved as for the DCM–OM57 pair. This is why we selected this particular pair of dyes for the subsequent experiments.

The characteristic radius R_0 of the resonance dipole–dipole energy transfer for this pair can be calculated using the well known relation [6, 7]

$$R_0 = \left(\frac{3}{4\pi}\int \frac{c^4}{\omega^4 n^4} F(\omega)\sigma(\omega)d\omega\right)^{1/6}, \qquad (1)$$

where $F(\omega)$ is the normalized fluorescence line shape of the donor and $\sigma(\omega)$ is the optical absorption cross section of the acceptor. From (1), it is easy to see that such a radius has a relatively slight dependence on the spectral overlapping integral (inverse sixth power only); calculations show that it ranges between 3 and 4 nm for all "reasonably overlapping" dye pairs, including DCM–OM57 (compare with the data given in [6]).

3. SCANNING NEAR-FIELD OPTICAL MICROSCOPE GEOMETRY

3.1. Experimental Equipment and Procedures

Different experimental schemes have been implemented for the demonstration of FRET phenomena in scanning probe microscopy. We start our discussion of the experimental results with the SNOM-based scheme where more quantitative results have been obtained.

The scheme of the experiment performed using the photon scanning tunneling (PSTM) version of SNOM [1, 2] is shown in Fig. 2. Two different homemade shear force-based SNOMs and homemade electronic units to

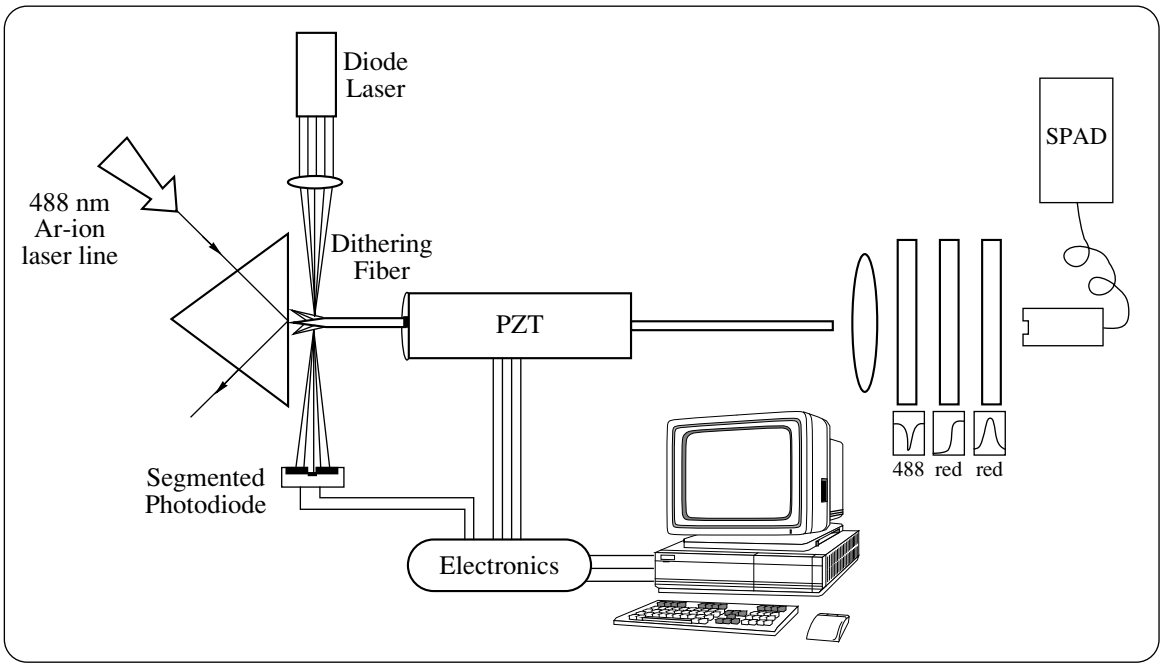


Fig. 2. Scheme of the PSTM-based FRET experiment.

control its performance have been used. In most of the experiments, the standard optical detection method to measure the amplitude of the lateral tip vibrations [1, 2] was utilized, but some critical experiments were performed using a non-optical shear force detection method similar to that described in [17]. We used SNOM tips with the curvature radius 100–200 nm prepared by the usual etching procedure of optical fibers in concentrated HF solutions [1, 2] or commercially available from Nanonics Supertips (Israel). Additional details about the construction of the microscopes used in the experiments can be found in [13, 18, 19].

To deposit the donor dye molecules onto the sample surface, we used the spin coating technique [20] as well as the simple method of spreading a small droplet (10 μl) of the dye solution onto the surface with the subsequent drying in air. No essential difference of the results has been observed.

Acceptor molecules have been deposited onto the surface of the SNOM tip. In this case, it seems essentially more difficult to elaborate the procedure to deposit the dye layers with a known and well-controllable concentration of the molecules: the complex shape of the tip, its very small size, and fragility make it impossible to use such well-established methods as the spin coating technique or long-time deposition of the sample inside a dye solution with the subsequent elimination of the excess solution by a tissue paper [21]. Thus the so-called withdrawal, or "dipping" approach seems to be the most suitable: the sharpened fiber tip is rapidly dipped into and extracted out of the dye solution; the thickness of the deposited film (and thus the surface concentration of dye molecules) is governed by fluid dynamics. The results of our experiments did not contradict the data known about the concentration dependence for flat surfaces [22, 23], and we are planning to finally clarify this point in future experiments.

The 488 nm spectral line of a cw argon ion laser was focused onto a glass prism surface under the conditions of total internal reflection (spot sizes were of the order 300 μm and the laser irradiation intensity I was 15 W/cm^2). Light coming out of the opposite side of the sharpened fiber was detected by a single photon avalanche diode (SPAD; EG & G, Canada, noise level 80 s^{-1}) after passing through a number of filters to suppress stray light and select the light originating from fluorescence of the acceptor molecules. The set of filters included a holographic notch filter for the 488 nm line, a red glass filter with the absorption edge of 660 or 695 nm, and interference filters centered at 750 nm with a width of 70 or 40 nm. It can be seen from Fig. 1 that such a set of filters enables the separation of the fluorescence coming from OM57 molecules from that coming from DCM molecules.

3.2. Results

In Fig. 3, we present the dependence of the fluorescence signal recorded by the EG & G SPAD for the case of DCM and OM57 molecules deposited with the surface concentrations 3×10^{13} cm^{-2} on the voltage driving the piezotube in the z-direction. Each point on this graph is a result of an average of 10–20 counts during one second each. The moment of contact between the tip and the sample was determined by the beginning of

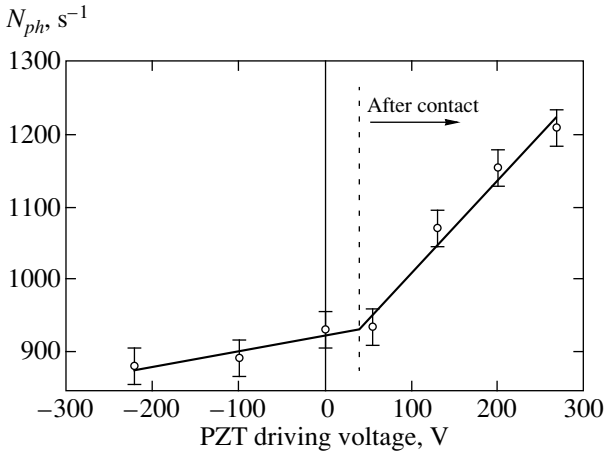


Fig. 3. The acceptor fluorescence signal dependence on the acting force recorded during the PSTM-based FRET experiment.

a decrease of the tip dithering amplitude as observed on an oscilloscope.

The distance between the tip and the sample, Δz, when out of contact, can be easily calculated as a function of the potential difference ΔU using the known calibration data for the driving piezo, $\Delta z = \zeta \Delta U$, where, $\zeta = 9.5$ nm V. After acquiring the contact, it is more reasonable to speak about the change in the force acting between the tip and the sample rather than about the change of a relative distance; an increase of the voltage tends to push the tip (rigidly fixed on the piezo) more strongly against the sample. The acting force F can be calculated using the spring constant k of the sharpened fiber by an obvious relation $F = k\zeta\Delta U$, and an action of this force leads to the flexural bending [12, 24, 25] and deformation of the tip.

From Fig. 3, it is easy to see that after acquiring the contact, the acceptor fluorescence signal starts to increase rapidly as the acting force increases. This effect has been well reproduced during at least a few tens of the cycles contact–out of contact measurements, but an overall slow decrease in the signal due to the photodegradation of the dyes was noticed.

A number of control experiments have been performed using the same tip-sample configuration but with the donor and acceptor dyes (either or one of or both them) absent. None of these control experiments revealed a behavior analogous to that presented in Fig. 3; only a very slow change in the fluorescence signal as a function of the driving voltage was usually observed and the contact point did not correspond to any peculiarities in the fluorescence signal. Of course, the absolute value of the recorded signal was smaller.

Similar results were obtained when we used another mode of SNOM operation, namely, the illumination mode SNOM (see Fig. 4) instead of the PSTM version described above. Donor dye molecules were deposited onto a thin glass slide surface. The SNOM tip, covered

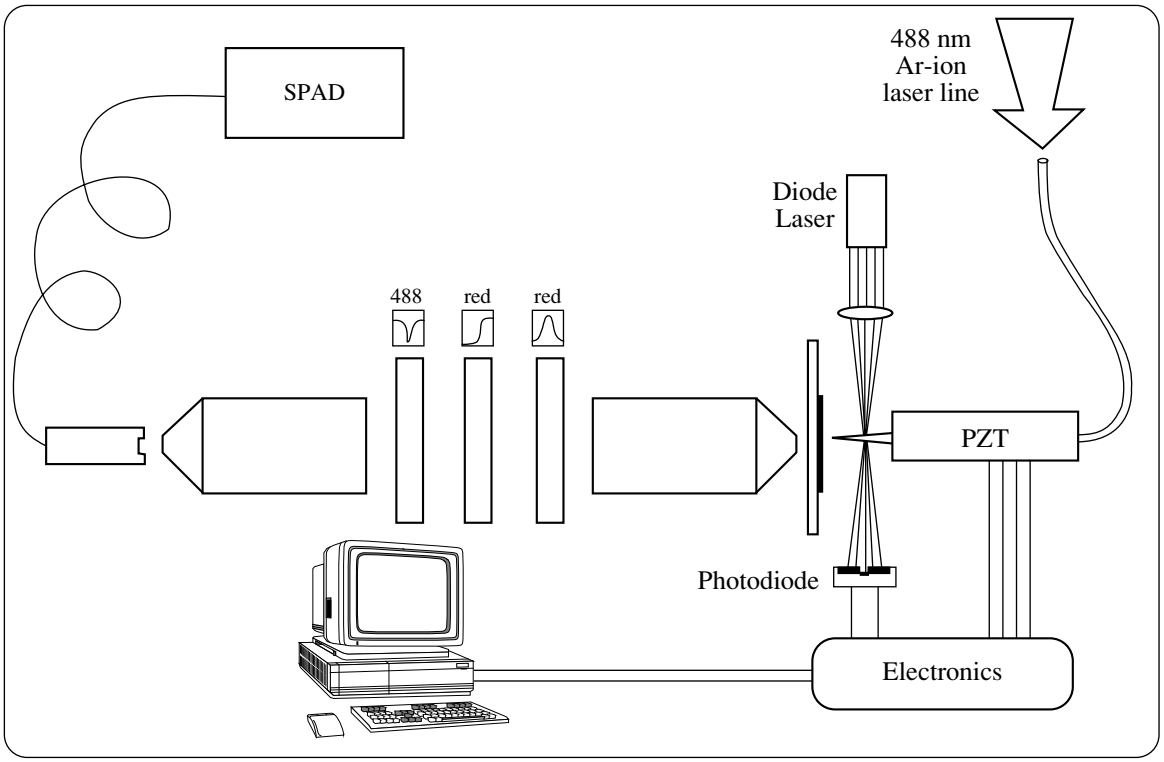


Fig. 4. Scheme of the illumination mode SNOM-based FRET experiment.

with a OM57 acceptor molecule layer, was used as a light source. The same detector and combination of filters were used. As in the previous case, the OM57 fluorescence signal drastically increased after the acquisition of the contact only when both dye layers were present.

3.3. Discussion

Thus, the acceptor fluorescence signal behavior presented in Fig. 3 is definitely due to the presence of both dyes and should be regarded as a demonstration of FRET phenomena in scanning-probe microscopy. The increase of the fluorescence signal as a result of the increase of the acting force was due to the corresponding increase of the contact surface and thus, of a number of molecules involved in the energy transfer process.

Semiquantitatively, the experimental data can be described as follows. Experimental measurements of the spring constant k for the glass fiber tips [11] as well as calculations based on the mechanical properties of the flexural bending of a glass cone [24] show that for a tip with the curvature radius 100 nm, the spring constant should be of the order 500–1000 N/m. This means that for an equivalent displacement of the piezo, $\zeta\Delta U$, (maximum value attains 1.9 μm) the acting force value should range within 10^{-4}–10^{-3} N. (Note that a similar range of forces was used in the recent SNOM experiments using normal dithering of a tip [12, 24]. Under the action of such a force, the tip will exhibit flexural bending [24, 25] and elastic deformation. Both these processes will result in an increase of the contact surface. For a rough estimate of the elastic deformation, one can use the known Hertzian expression to describe the contact radius r_c of a sphere pressed against a flat sample surface as a function of the acting force F (see, for example, [18], where the problem of elastic deformations in AFM has been specially investigated:

$$r_c = \left(\frac{3(1-\nu^2)Fr}{4E}\right)^{1/3}. \qquad (2)$$

Here r is the curvature radius of the tip, $E = 7 \times 10^{10}$ N/m^2 and $\nu = 0.25$ are typical Young modulus and Poisson ratio for glass. For $F = 10^{-4}$ N, r = 100 nm, this expression gives $r_c = 46$ nm, which corresponds to $N_1 \sim 2000$ molecules in the "FRET active" contact area for the surface concentration 3×10^{13} cm^{-2}.

An absolute value of the fluorescence signal recorded for the sharpest tips used was equal to N_{ph} = 80–100 s^{-1} (with the signal to noise ratio of the order of unity). Knowing this value, we can estimate the number of molecules N_2 contributing to the measured signal using the simple relation

$$N_2 = \frac{N_{ph}h\nu}{I\sigma\eta\Phi}, \qquad (3)$$

where $h\nu = 4.07 \times 10^{-19}$ J is the photon energy, Φ is the fluorescence quantum yield of the acceptor molecule, and η is an overall efficiency of the photon collection and detection for our experimental system. The latter can be estimated as follows. The efficiency of the fluorescence photon collection by a sharpened fiber for a geometry similar to ours was reported to be $2–5 \times 10^{-3}$ [26]. We estimate the efficiency of the detection of photons coming out of the fiber about 0.1–0.05 (this value is mainly due to the registration only within a rather narrow spectral band of the total acceptor fluorescence because of the strong filtering, see above) and thus, the overall efficiency of the detection is $1–5 \times 10^{-4}$. This means that $N_2 \sim$ 300–1500 acceptor molecules contribute to the measured signal (we assume $\Phi = 0.3$). Both N_1 and N_2 values are in reasonable coincidence with each other, which strengthens our conclusions about an observation of the "nanolocal" FRET phenomenon with only hundreds to thousands of molecules involved.

4. ATOMIC FORCE MICROSCOPE GEOMETRY

The above results were qualitatively confirmed in another series of experiments performed with the same DCM–OM57 dye pair. OM57 acceptor molecules were deposited onto the surface of a silicon tip of a standard AFM cantilever (NT–MDT, Moscow, the force constant 0.12 N/m, the curvature radius of the tip 10–20 nm). DCM donor molecules were deposited onto a thin glass slide surface and a 488 nm laser line was focused onto this surface by a 40× microobjective after a reflection from a selective mirror at the angle of 45° (see Fig. 5). Contact between the AFM tip and the sample as well as the acting force was controlled by monitoring the reflection of a focused diode laser radiation from the opposite side of the cantilever, as is typical in the usual contact mode AFM. The same driving piezo and electronic control unit as in the SNOM-based experiment described above were used. Fluorescent light was collected using the same 40 x microobjective, and after the passage through the selective mirror without reflection, it was refocused onto the entrance slit of a CCD-camera-equipped monochromator. The same set of filters as described earlier (except for an interference filter centered at 750 nm) was used. Light intensity was essentially higher, ~600 W/cm^2, and as a result, the photodegradation was more prominent. Nevertheless, it was possible to observe the difference between the "contact" and "noncontact" fluorescence during a number of the cycles contact-out of contact measurements (analogous to the SNOM experiments described above).

No quantitative information has been collected in these series of experiments, but in Fig. 6, we present two fluorescence spectra recorded when in deep contact (1) and out of contact (2). The acting force, estimated for the "deep contact" case in the same manner as discussed above for the SNOM case, was equal to ~10^{-6} N. It can be clearly seen that the signal obtained while in

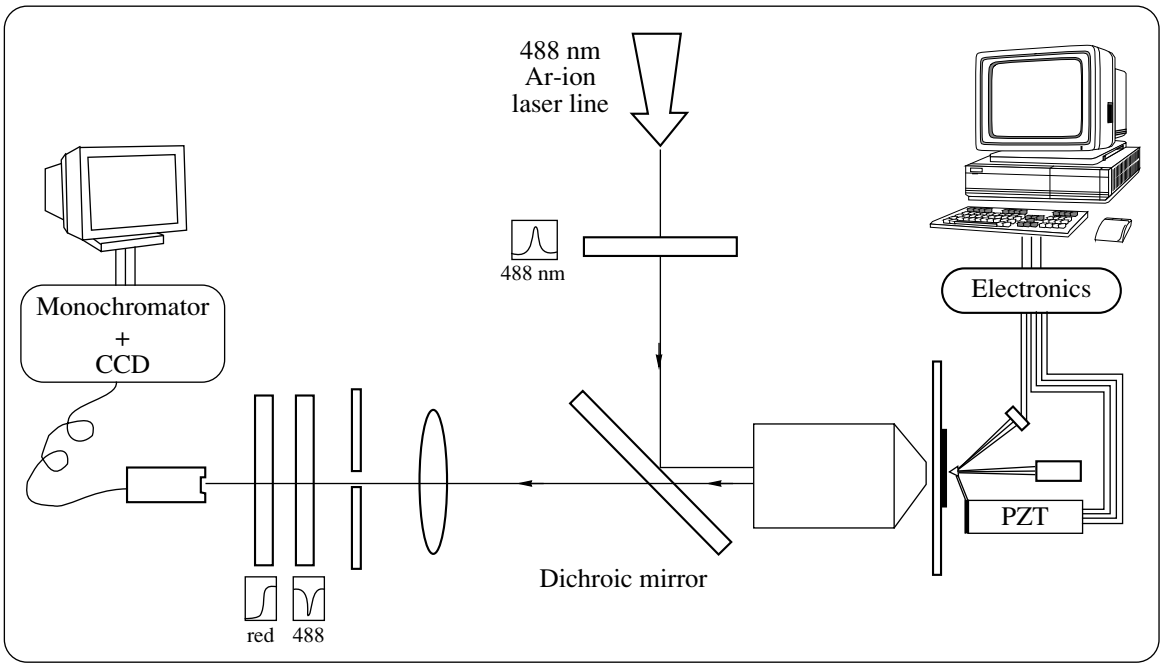


Fig. 5. Scheme of the AFM-based FRET experiment.

contact was 10–20% larger than that obtained out of contact. This effect was not observed in control experiments, when both or one dye were missing, which is an additional indication of the observation of the AFM- (or SNOM-) related nanolocal FRET phenomenon.

5. FLUORESCENCE RESONANCE ENERGY TRANSFER PROBE MICROSCOPE

The results presented in this work should be treated as a demonstration of the FRET phenomenon in scanning near-field optical microscopy and atomic force microscopy techniques. Of course, the difference between SNOM and AFM is rather arbitrary in this context: for example, silicon AFM levers similar to those described in Section 4 are optically semitransparent and optical near-field microscopy can be implemented to measure the light passing through these levers [27]. An important point is that we succeeded in recording a nanolocal "contact dependent" fluorescence signal corresponding to a few thousands or even hundreds of FRET-active molecules. Relying on the obtained experimental data. Sensitivity, noise level, etc., we believe that using a slightly modified approach (for example, that based on an illumination-geometry SNOM, where single fluorescent molecules were indeed observed recently [9, 10]), it will be possible to observe the FRET phenomenon at the level of only one single molecule.

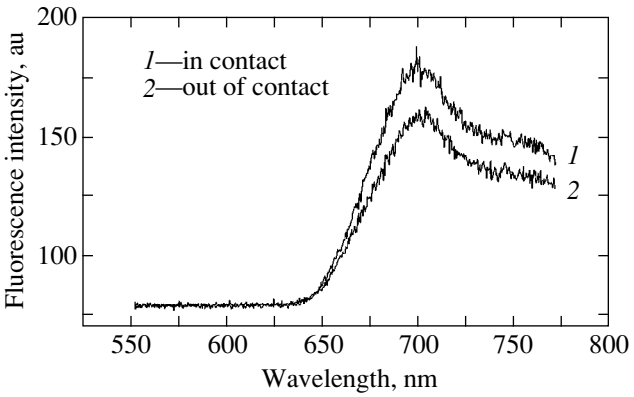


Fig. 6. Fluorescence signals recorded during the AFM-based FRET experiment for the tip in contact (curve *1*) and out of contact (curve *2*) with the sample.

Thus, based on the experimental results already obtained, we believe that it is now the time to briefly discuss a FRET probe Microscope that is under construction in our laboratory (Troitsk). The particular scheme of this microscope is presented in Fig. 7. The principal element of the FRET probe is a small polystyrene bead stained with a dye. Such microspheres are commercially available in a broad range of sizes (10 μm–20 nm) and dyes and are routinely used in the optical fluorescent microscopy as markers (see, e.g., [28]). Preliminary results indicating FRET processes involving some dye doped beads commercially available from Molecular Probes, Oregon, have been recently obtained in our laboratory (Lausanne: unpublished), but one should be sure that the dye molecules at the surface are not covered by protective layers.

To provide the possibility of the probe manipulation, we are elaborating the method to fix it to the apex of a microcapillary, as indicated in Fig. 7. Fixing may

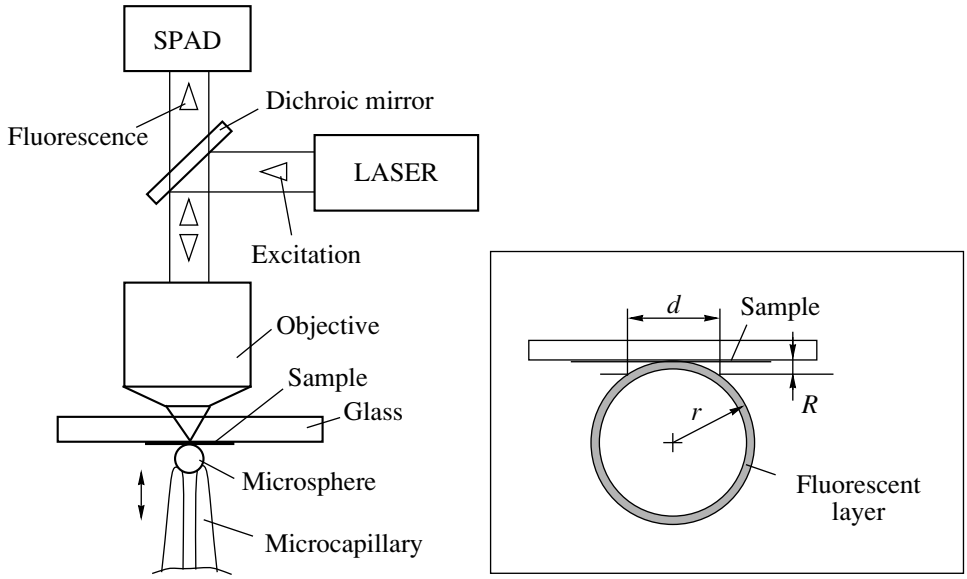


Fig. 7. Scheme of a FRET probe microscope. Only the upper part of the capillary is shown. The glass slide with the sample is supposed to be moved by a scanner of the microscope. The electronic part (not shown) is similar to that of a scanning probe microscope.

be possible relying on the van-der-Waals forces or a specific chemical binding, and a number of microcapillaries with the diameter ranging from 100 nm to a few microns is now commercially available.

The optical scheme of the FRET probe microscope is similar to that of the confocal fluorescent microscope. The laser beam is reflected from a dichroic mirror and is focused on the sample. Let us assume that the sample contains donor molecules and the bead contains acceptor dye molecules. If the donor–acceptor pair is chosen as described above, mainly the donor molecules are excited due to the FRET process. The light from fluorescent acceptor molecules is collected by the same objective that is used for illumination of the sample. The light coming through the dichroic mirror is detected by a SPAD or a PMT.

The possibility to modulate the probe-sample distance is implemented in the microscope construction. This enables one to improve the sensitivity and to remove the background signals caused by "tails" of donor fluorescence and the direct excitation of acceptor molecules. The modulation allows the distance-dependent part of the signal to be extracted, because the FRET between the probe and the sample is possible only when the bead comes in contact with the sample. The modulation in the range of several tens of nanometers will suffice because the Förster radius does not exceed several nanometers [6], and such a modulation can be realized based on the usual shear force feedback [1, 2] or normal tip vibrations that we have recently realized [12, 24].

A simple estimation of the lateral resolution of the microscope can be made from elementary geometrical considerations. If we model the sphere as touching the surface without elastic deformation (see insert in Fig. 7), the FRET is possible for acceptor molecules located between the surface and the imaginary plane at the distance R_0 (Förster radius) from the surface, one should now distinguish between two different situations: i) there is only one acceptor molecule inside this area, and ii) there are several molecules inside it. In the former case, the resolution is governed by the Förster radius [5]. Note that this is exactly the case for the commercially available beads with the diameter $2r = 20$ nm (Molecular Probes, Oregon). Such a bead contains $N \sim 180$ molecules of the dye distributed in the volume of the sphere (not the surface) [28], and therefore, we can find

$$n \approx (1/2)N(R_0/r)^3 \approx 0.7$$

acceptor molecules in the FRET-active area (we take the Förster radius R_0 equal to 2 nm). Indeed, it is not necessary to use the smallest available spheres with the diameter 20 nm to attain such a resolution. Similar numbers can be obtained for much large beads: applying the same relation for $2r = 1$ μm and $N = 1.3 \times 10^7$ [28] gives $n = 0.4$.

In the latter case, the resolution is determined by the diameter d of the interaction zone. Taking into account that the radius of the sphere $r \gg R_0$, it is easy to find that $d \approx 2\sqrt{2rR_0}$, which corresponds to the resolution $d \approx 12$ nm for the beads with $2r = 20$ nm. An important issue is the number N of interacting molecules in the bead. The area of the bead surface inside the diameter d is defined by the expression

$$2\pi r^2\left(\sqrt{1-\frac{d^2}{4r^2}}\right).$$

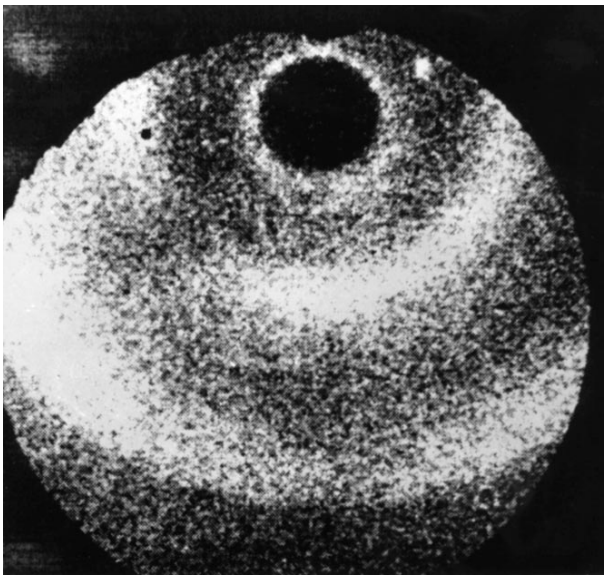


Fig. 8. The field emission image of the borosilicate glass microcapillary with the inner diameter of 0.5 μm. The image of the central hole is shifted to the upper part of the figure. The thickness of walls looks exaggerated due to the peculiarities of image formation in the field emission microscope.

For an estimate, an approximation gives simply $S = \pi d^2/4$. If the surface density of dye molecules is $n = 10^{14}$ cm^{-2} (beads having dye molecules on the surface can be easily prepared in the laboratory), we have $N = nS \approx 270$ molecules, which is large enough to avoid the photostability problem, which is crucial at the single-molecule level.

We emphasize that the resolution 10–20 nm can be achieved without the subwavelength aperture that is a principal element of the standard near-field microscope. Complex boundary conditions for the near-field at the probe apex of SNOM considerably complicate the analysis of SNOM images which can lead to a number of artifacts (see, e.g., [29]). The FRET probe microscope is free from these drawbacks because of the absence of the aperture and the physical clarity of the interaction.

6. CONCLUSIONS

In this paper, we have presented experimental results concerning an observation of nanolocal FRET processes for the usually used SNOM and AFM geometries; we then discussed the FRET probe microscope currently under construction in Troitsk. In addition, practical elaboration of this FRET microscope is especially timely because we have at our disposal the already finished one-atom fluorescent tips made of LiF : F_2 crystal fragments, where only one effective and very photostable fluorescent center in the tip apex region (F_2 aggregate center, which is the specific defect of LiF crystalline lattice) has been observed using the laser selective photoelectron projection microscopy technique [8, 30]. We used the same technique to observe the apex of the nanocapillary currently explored in the FRET probe microscope under construction (see Fig. 8), and will use it to control the fixing of a dye-saturated bead on the capillary apex, as described in Section 5.

The practical realization of the FRET probe microscope makes it possible to drastically improve the spatial resolution and the sensitivity of scanning near-field optical microscopy, thereby opening new prospects in the field [5, 8]. The FRET SNOM will be very useful when working not only at the single-molecule level, but also with hundreds or thousands of molecules involved (exactly as reported here), because this approach, in any case, improves the resolution and sensitivity of SNOM and enlarges the number of possible experimental schemes of the microscope to be used.

ACKNOWLEDGMENTS

This research was supported by the Swiss National Science Foundation (grants nos. 20-50427.97, 20-57133.99 and 21-50805.97) and Russian Department of Physics and Technology (grant no. 3-1).

REFERENCES

1. *Near Field Optics*, Ed. by D. W. Pohl and D. Courjon (Kluwer, Dordrecht, 1993).
2. M. A. Paesler and P. J. Moyer, *Near Field Optics: Theory, Instrumentations, and Applications* (Wiley, New York, 1996).
3. R. Kopelman and W. Tan, Appl. Spectrosc. Rev. **29**, 39 (1994).
4. F. Zenhausern, M. P. O'Boyle, and H. K. Wickramasinghe, Appl. Phys. Lett. **65**, 1623 (1994).
5. S. K. Sekatskiĭ and V. S. Letokhov, Pis'ma Zh. Éksp. Teor. Fiz. **63**, 323 (1996) [JETP Lett. **63**, 319 (1996)].
6. P. Wu and L. Brandt, Anal. Biochem. **218**, 1 (1994).
7. V. M. Agranovitch and M. D. Galanin, *Electron Excitation Energy Transfer in Condensed Matter* (North-Holland, Amsterdam, 1982).
8. S. K. Sekatskiĭ and V. S. Letokhov, Appl. Phys. B **63**, 525 (1996).
9. W. E. Moerner and M. Orrit, Science **283**, 1670 (1999).
10. S. Weiss, Science **283**, 1676 (1999).
11. C. E. Talley, G. A. Cooksey, and R. C. Dunn, Appl. Phys. Lett. **69**, 3809 (1996).
12. D. A. Lapshin, S. K. Sekatskiĭ, V. S. Letokhov, and V. N. Reshetov, Pis'ma Zh. Éksp. Teor. Fiz. **67**, 243 (1998) [JETP Lett. **67**, 263 (1998)].
13. G. T. Shubeita, S. K. Sekatskiĭ, M. Chergui, *et al.*, Appl. Phys. Lett. **74**, 3453 (1999).
14. S. A. Vickery and R. C. Dunn, Biophys. J. **76**, 1812 (1999).
15. U. Brackmann, *Lambdachrome Laser Dyes* (Lambda Physics, Göttingen, 1997).
16. D. C. Nguen, R. E. Muenchausen, R. A. Keller, and N. S. Nogar, Opt. Commun. **60**, 111 (1986).

17. A. Drabenstedt, J. Wrachtrup, and C. von Borczyskowski, Appl. Phys. Lett. **68**, 3497 (1996).
18. M. Heuberger, G. Dietler, and L. Schlapbach, Nanotechnology **5**, 12 (1994).
19. D. A. Lapshin, Zh. Tekh. Fiz. **68** (9), 51 (1998) [Tech. Phys. **43**, 1055 (1998)].
20. M. Lieberherr, C. Fattinger, and W. Lukosz, Surf. Sci. **189–190**, 954 (1987).
21. H. Gerischer and F. Willig, Top. Curr. Chem. **61**, 31 (1976).
22. S. Garoff, R. B. Stephens, C. D. Hanson, and G. K. Sorensen, Opt. Commun. **41**, 257 (1982).
23. C.-C. Yang, J. E. Josefowicz, and L. Alexandru, Thin Solid Films **74**, 117 (1980).
24. D. A. Lapshin, V. N. Reshetov, S. K. Sekatskiĭ, and V. S. Letokhov, Ultramicroscopy **76**, 13 (1999).
25. P. K. Wei and W. S. Fann, J. Appl. Phys. **84**, 4655 (1998).
26. M. Ohtsu, J. Lightwave Technol. **13**, 1200 (1995).
27. H. U. Danzenbrink, A. Castiaux, C. Girard, *et al.*, Ultramicroscopy **71**, 371 (1998).
28. R. P. Haugland, *Handbook of Fluorescent Probes and Research Chemicals* (Molecular Probes, Eugene, Oregon, 1997).
29. B. Hecht, H. Bielefeldt, Y. Innouye, *et al.*, J. Appl. Phys. **81**, 2492 (1997).
30. S. K. Sekatskiĭ and V. S. Letokhov, Pis'ma Zh. Éksp. Teor. Fiz. **65**, 435 (1997) [JETP Lett. **65**, 465 (1997)].

Journal of Experimental and Theoretical Physics, Vol. 90, No. 5, 2000, pp. 778–787.
Translated from Zhurnal Éksperimental'noĭ i Teoreticheskoĭ Fiziki, Vol. 117, No. 5, 2000, pp. 895–905.

ATOMS, SPECTRA, RADIATION

Tunneling Limit in the Theory of Photoelectron Rescattering by the Parent Ion

S. P. Goreslavskiĭ* and S. V. Popruzhenko

Moscow Engineering Physics Institute (Technical University), Moscow, 115409 Russia
**e-mail: grslv@theor.mephi.msk.su*

Received December 1, 1999

Abstract—A generalized Keldysh model is used to obtain simple analytical expressions for the energy and angular distributions of photoelectrons rescattered by the parent ion. The dependence of the form, absolute magnitude, and interference structure of the distributions on the parameters of the field and atom is investigated. It is shown that even though the semiclassical three-step rescattering model determines correctly the position of the boundaries beyond which the distributions decay rapidly, the model itself is inapplicable near them. *© 2000 MAIK "Nauka/Interperiodica".*

1. INTRODUCTION

Investigations of the nonlinear ionization of atoms in strong laser fields have revealed a variety of physical phenomena which are determined by the interaction of the freed electron with the parent atom (see the reviews [1, 2]). One channel of this interaction, ordinarily termed elastic rescattering, leads to the fact that a small fraction of the electrons injected into the continuum and accelerated by the laser field acquires high energies as a result of interacting with the parent ion. As a result, the electron spectrum detected in a fixed direction consists of two sections with substantially different properties. The initial segment up to energies of the order of $2U_p$ ($U_p = F^2/4\omega^2$ is the average oscillatory energy of an electron in a linearly polarized field with amplitude F and frequency ω) is explained by the mechanism of direct above-threshold ionization. Here the spectrum decays rapidly as the electron energy increases. Next follows an extended section with small decay (high-energy plateau) which is produced as a result of rescattering. The plateau terminates clearly in a distinct boundary, beyond which exponential decay occurs once again. The structure and details of the electron distributions on the plateau are seen in greatest relief and most clearly in the data from a series of experiments performed in the tunneling regime [3–5].

The theory of electronic distributions in the region of the plateau is based primarily on three approaches: numerical integration of the Schrödinger equation [6], a semiclassical three-step model [3–5, 7, 8], and a generalized Keldysh model [9–13]. Numerical methods are ineffective in the tunneling regime. The three-step model describes separately the freeing of an electron from an atom, the subsequent motion of the electron along a classical trajectory returning to the parent ion, and finally the act of scattering by an ion. This model has successfully explained the position of the top boundary of the plateau and the maximum angle of emergence of electrons, but the reason for the discrepancies with quantum calculations in the description of the angular distributions has remained unclear [8], and there is no clear picture of its accuracy and conditions of applicability.

The generalized Keldysh model arises as a result of an iterative solution of the quantum equations of motion of an electron in laser and atomic fields, when the Keldysh approximation serves as the zeroth approximation and the interaction of the freed electron with the atomic fragment is taken into account as a perturbation [9–13]. This substantially quantum approach makes it possible, specifically, to calculate the interference structure in the distributions of the rescattered electrons, which is in principle impossible to do on the basis of the three-step model. A complicating feature of this approach is that the rescattering amplitude has a very complicated form (a five-fold integral). A model of the atom as a zero-radius potential permits reducing this expression to a single integral of an infinite sum of products of Bessel functions [11]. An attempt to construct the rescattering amplitude, after calculating three of the five integrals by the saddle-point method, has been made in [10]. This calculation confirmed that the classical trajectories returning to the parent ion play an important role in the formation of the rescattering amplitude. However, the computational scheme adopted led to a singularity in the remaining double integral, and the final results turned out to be sensitive to the value of the regularization parameter, which is the initial width of the ionized wave packet.

The objective of the present work is to obtain, in the tunneling regime, for the spectral-angular distribution of the rescattered electrons quite simple formulas from which the dependences on the parameters of the field and the atom would be evident and the mechanism of

1063-7761/00/9005-0778$20.00 © 2000 MAIK "Nauka/Interperiodica"

quantum interference in the distributions of the rescattered electrons would be understandable, as done previously for direct ionization [14, 15]. Looking ahead, we note that the results will be presented in a form closely related with the three-step model, which will make it possible to establish the status and region of applicability of the latter.

A general expression for the rescattering amplitude is obtained in the next section. Section 3 is devoted to a derivation of the approximate formulas for the tunneling limit. Finally, the physical consequences are discussed in section 4. Some of the results presented have been reported in preliminary publications [13, 16, 17].

2. FORMULATION OF THE MODEL AND THE RESCATTERING AMPLITUDE

The wave function of an atomic electron interacting with a laser field satisfies the Schrödinger equation (atomic units are employed)

$$i\frac{\partial\Psi(\mathbf{r},t)}{\partial t} = \left\{\frac{1}{2}\left(\mathbf{p}+\frac{1}{c}\mathbf{A}(t)\right)^2 + U(\mathbf{r})\right\}\Psi(\mathbf{r},t), \quad (1)$$

where $\mathbf{A}(t)$ is the vector potential of the field in the dipole approximation. Before the field is switched on, in the limit $t \longrightarrow -\infty$, the atom is in a bound state

$$\varphi_0(\mathbf{r},t) = \varphi_0(r)e^{iIt}$$

with ionization potential I. We seek the solution in the form of a sum of the initial state and a new, unknown function

$$\Psi(\mathbf{r},t) = \varphi_0(\mathbf{r},t) + \chi(\mathbf{r},t).$$

In contrast to the phenomenological approach of [18], this formulation transforms equation (1) in a natural manner into a Schrödinger equation with a source. Switching to the momentum representation and using an exponential substitution to eliminate the diagonal matrix elements of the kinetic energy

$$\varepsilon_{\mathbf{p}}(t) = (\mathbf{p}+\mathbf{A}(t)/c)^2/2$$

are equivalent to writing the desired function in the form of the expansion

$$\chi(\mathbf{r},t) = \sum_{\mathbf{p}}\chi(\mathbf{p},t)\Psi_{\mathbf{p}}(\mathbf{r},t)$$

over the Volkov states

$$\Psi_{\mathbf{p}}(\mathbf{r},t) = \exp\{i\mathbf{p}\mathbf{r} - iS_V(\mathbf{p},t)\}$$

with the phase

$$S_V(\mathbf{p},t) = \int_{-\infty}^{t} d\tau\varepsilon_{\mathbf{p}}(\tau).$$

The amplitudes in the expansion satisfy the initial condition

$$\chi(\mathbf{p}, t \longrightarrow -\infty) = 0$$

and the equation

$$i\dot{\chi}(\mathbf{p},t) = V(\mathbf{p},t)\varphi_0(\mathbf{p})\Psi^*_{\mathbf{p}}(0,t)\exp\{iIt\} + \sum_{\mathbf{k}}U(\mathbf{p}-\mathbf{k})\Psi^*_{\mathbf{p}}(0,t)\Psi_{\mathbf{k}}(0,t)\chi(\mathbf{k},t), \quad (2)$$

where $\varphi_0(\mathbf{p})$ and $U(\mathbf{q})$ are the Fourier transforms of the initial state and the atomic potential, respectively, and

$$V(\mathbf{p},t) = \mathbf{p}\cdot\mathbf{A}/c + \mathbf{A}^2/2c^2.$$

The equation (2) can be solved by iteration

$$\chi(\mathbf{p},t) = \chi^{(0)}(\mathbf{p},t) + \chi^{(1)}(\mathbf{p},t) + \ldots$$

provided that the integral term is small.

The solution in the zeroth approximation is obtained in the form of an integral of the source:

$$\chi^{(0)}(\mathbf{p},t) = -i\varphi_0(\mathbf{p}) \times \int_{-\infty}^{t} dt_1 V(\mathbf{p},t_1)\exp\{iIt_1 + iS_V(\mathbf{p},t_1)\}, \quad (3)$$

and its limiting value $\chi^{(0)}(\mathbf{p}, +\infty)$ is identical to the ionization amplitude used in practical calculations of the energy and angular distributions in the Keldysh model [14, 19].

Performing the next iteration we find

$$i\chi^{(1)}(\mathbf{p},\infty) = \int_{-\infty}^{\infty} dt\sum_{\mathbf{k}}U(\mathbf{p}-\mathbf{k})\Psi^*_{\mathbf{p}}(0,t)\Psi_{\mathbf{k}}(0,t)\chi^{(0)}(\mathbf{k},t). \quad (4)$$

The expression (4) has a simple meaning: it is the scattering amplitude, calculated in first-order perturbation theory in the atomic potential, for the ionization-produced coherent packet of the Volkov states $\chi^{(0)}(\mathbf{r},t)$ into the final Volkov state $\Psi_{\mathbf{p}}(\mathbf{r},t)$. Taking an individual Volkov wave as the initial state in equation (4), we obtain the Born amplitude for induced bremsstrahlung [20].

The coordinate function in the zeroth approximation

$$\chi^{(0)}(\mathbf{r},t) = \sum_{\mathbf{k}}\chi^{(0)}(\mathbf{k},t)\Psi_{\mathbf{k}}(\mathbf{r},t)$$

describes a spreading electron cloud. The norm of this function grows linearly with time [21]. This makes it possible to interpret in the limit $t \longrightarrow \infty$ the quantity

$$dW = t^{-1}\left|\chi^{(0)}(\mathbf{p},t)\right|^2 d^3p/(2\pi)^3$$

as the ionization probability per unit time. Taking into account the first iteration, the following substitution must be made in this probability:

$$\left|\chi^{(0)}(\mathbf{p}, +\infty)\right|^2 \longrightarrow \left|\chi^{(0)}(\mathbf{p}, +\infty) + \chi^{(1)}(\mathbf{p}, +\infty)\right|^2.$$

As noted in the Introduction (see also [11]), for final states with the energies

$$\varepsilon = p^2/2 < 2U_p$$

direct above-threshold ionization predominates, i.e.,

$$\left|\chi^{(0)}\right| \gg \left|\chi^{(1)}\right|,$$

and for $\varepsilon > 2U_p$ the relation between the amplitudes is reversed. In each of the indicated regions it is sufficient to retain only the main contribution. Then the desired ionization probability in states with high energies is determined by the rescattering amplitude (4). We shall not study the transitional region, where

$$\left|\chi^{(0)}\right| \approx \left|\chi^{(1)}\right|$$

and interference of the two terms is possible.

The integrand in the amplitudes $\chi^{(0,1)}$ has the form

$$B^{(0,1)}(t)\exp\{i\lambda t\},$$

where

$$B^{(0,1)}(t + 2\pi/\omega) = B^{(0,1)}(t), \quad \lambda = p^2/2 + U_p + I.$$

A standard Fourier series expansion of the periodic factor gives the ionization probability in the form

$$dW = 2\pi\sum_n \left|B_n^{(0,1)}\right|^2 \delta(\lambda - n\omega)\frac{d^3p}{(2\pi)^3}.$$

In the tunneling regime

$$F < F_a = (2I)^{3/2}, \quad \omega < I, \quad \gamma = \omega(2I)^{1/2}/F < 1, \quad (5)$$

the energy of a laser photon is the smallest energy scale in the problem, and replacing in the probability the sum over n by an integral we obtain the momentum (spectral-angular) distribution, which describes the envelope of the above-threshold peaks of direct ionization and ionization with rescattering (indices d and r, respectively)

$$dW_{d,r} = 2\pi\left|B^{(0,1)}(\mathbf{p})\right|^2 \frac{d^3p}{\omega(2\pi)^3}. \quad (6)$$

The Fourier coefficients in equation (6) are denoted as $B^{(0,1)}_{n=\lambda/\omega} = B^{(0,1)}(\mathbf{p})$.

We also note that with time the amplitudes $Wt \ll 1$ satisfying the zero initial condition remain small as long as $\chi^{(0)}(\mathbf{p}, t)$, making the integral term in equation (2) small and justifying the use of an iteration procedure.

3. RESCATTERING AMPLITUDE IN THE TUNNELING LIMIT

The conditions (5) make it possible to calculate the amplitudes (3) and (4) by the saddle-point method. If a linearly polarized field $F\sin(\omega t)$ acts along the x axis, then the equation for the saddle-points of the integral (3) is

$$2I + k_\perp^2 + v_x^2(t_s) = 0, \quad (7)$$

where

$$v_x(t) = k_x + p_F\cos(\omega t), \quad p_F = F/\omega.$$

In the tunneling limit the complex roots t_s of equation (7) lie near the real axis. Neglecting the contributions of order $\gamma^2 \ll 1$ and higher we obtain

$$\mathrm{Im}\, t_s = \gamma/\omega.$$

This time, interpreted as the transit time of an electron through the barrier, is a small fraction of the optical period [19]. To the same accuracy $\mathrm{Re}\, t_s$ is identical to the real moment in time t_0 when the electron kinetic energy reaches its minimum value. For momenta $|k_x| < p_F$ which are important in the direct-ionization spectrum, the projection of the velocity on the direction of the field vanishes at the time t_0:

$$v_x(t_0) = k_x + p_F\cos(\omega t_0) = 0. \quad (8)$$

The contribution of an individual saddle point, calculated in the approximation described, is

$$\delta\chi^{(0)}(\mathbf{k}, t_0) = C(F)\frac{2\pi i}{\sqrt{F(t_0)}} \times \exp\left\{-\frac{F_a}{2F(t_0)}\left(\frac{2}{3} + \frac{k_\perp^2}{2I}\right) + i(It_0 + S_V(\mathbf{k}, t_0))\right\}. \quad (9)$$

Here $k_\perp$ is the momentum in a plane perpendicular to the x axis, $F(t) = |F\sin(\omega t)|$ is the magnitude of the field, and the function $t_0 = t_0(k_x)$ found from equation (8) must be substituted. The factor

$$C(F) = 1$$

corresponds to ionization from a well with zero radius, and

$$C(F) = 2\sqrt{2}F_a/F$$

gives the correct static limit for a Coulomb field [22, 23].

The partial amplitude (9) looks as if the quantum transition into a given state of the continuum occurs instantaneously, at the moment of the minimum kinetic energy, which depends on the initial momentum k_x. The integral in equation (3) is equal to the sum of the contributions of all points t_0 falling within the limits of

integration. It is convenient to represent this sum over real stationary points as an integral:

$$\chi^{(0)}(\mathbf{k}, t) = \int_{-\infty}^{t} dt_0 F(t_0)\delta(k_x - p_F\cos(\omega t_0))\delta\chi^{(0)}(\mathbf{k}, t_0). \quad (10)$$

The amplitude for direct ionization in equation (6) is related with the partial contribution (9) by the relation

$$B^{(0)}(\mathbf{p}) = \frac{\omega}{2\pi}(\delta\chi^{(0)}(\mathbf{p}, t_{0-}) + \delta\chi^{(0)}(\mathbf{p}, t_{0+})),$$

since for a given k_x there are two stationary points t_{0-} and t_{0+} in one optical period. The presence of two terms gives rise to interference in the direct-ionization distributions [14, 15]. Keeping in mind the subsequent comparison with rescattering, we underscore that for direct ionization $t_{0\pm}$ belong to neighboring optical half-periods with opposite orientation of the electric field.

The five-fold integral in the rescattering amplitude (4) can be calculated as follows. We substitute the expression (10) into equation (4), and using the δ function we calculate the integral over k_x. Since the amplitude $\delta\chi^{(0)}(\mathbf{k}, t)$ decreases rapidly for

$$k_\perp > (F/F_a)^{1/2}\sqrt{2I},$$

the $\mathbf{k}_\perp$ in the argument of the atomic potential can be neglected, setting

$$q = |\mathbf{p} - \mathbf{k}| \approx |\mathbf{p} - \mathbf{k}(t_0)|,$$

where the vector $\mathbf{k}(t_0)$ possesses a single projection

$$k_x(t_0) = -p_F\cos(\omega t_0).$$

(A more radical but somewhat less accurate approximation $q \approx p$ is also possible, since the amplitudes $\chi^{(0)}(\mathbf{k}, t)$ are large only for $k^2/2 \ll 2U_p$ and we are interested in rescattering into states with high energies $\varepsilon = p^2/2 > 2U_p$.) After this simplification the integral over $\mathbf{k}_\perp$ becomes Gaussian, and a calculation of the integral leads to the appearance of a complex transverse width of a spreading wave packet in the denominator of the integrand (see equation (11) below). In the remaining double integral we change the order of integration and switch to the dimensionless times

$$\varphi_0 = \omega t_0, \quad \varphi_1 = \omega t_1.$$

It is easy to verify that the integrand in the outer integral over infinite limits in the variable φ_0 possesses the necessary property of periodicity, ensuring the correct form of the law of conservation of energy during ionization (see discussion preceding equation (5)). Under the transformations described above, the amplitude $B^{(1)}$ in equation (6) becomes

$$B^{(1)}(\mathbf{p}) = \int_0^{2\pi} d\varphi_0 \int_{\varphi_0}^{\infty} d\varphi_1 \frac{C(F)U(q)\sqrt{F(\varphi_0)}}{(2\pi)^2\omega\left(\Delta_0^2 + i\frac{\varphi_1 - \varphi_0}{\omega}\right)} \times \exp\left\{-\frac{F_a}{3F(\varphi_0)} + iS(\mathbf{p}, \varphi_0, \varphi_1)\right\}. \quad (11)$$

Here the phase is determined by the relation

$$\omega S(\mathbf{p}, \varphi_0, \varphi_1) = I\varphi_0 + \int_{-\infty}^{\varphi_1} d\varphi'\varepsilon_{\mathbf{p}}(\varphi') - \int_{\varphi_0}^{\varphi_1} d\varphi'\varepsilon_{\mathbf{k}(\varphi_0)}(\varphi') \quad (12)$$

and $\Delta_0 = \sqrt{F_a/2IF(\varphi_0)}$ is the transverse width of the electron wave packet at the moment of ionization [21]. Since the phase (12) is proportional to the large parameter

$$z_F = 4U_p/\omega \gg 1,$$

we shall calculate the double integral (11) by the saddle-point method. The conditions of stationariness

$$\partial S/\partial\varphi_0 = 0, \quad \partial S/\partial\varphi_1 = 0$$

yield the equations

$$(\varphi_1 - \varphi_0)\cos\varphi_0 - \sin\varphi_1 + \sin\varphi_0 = 0, \quad (13)$$

$$\varepsilon_{\mathbf{p}}(\varphi_1) = \varepsilon_{\mathbf{k}(\varphi_0)}(\varphi_1). \quad (14)$$

The solution of the system (13) and (14) is the point $(\varphi_0(\varepsilon, \theta), \varphi_1(\varepsilon, \theta))$, whose position depends on the final electron energy $\varepsilon = p^2/2$ and angle of emergence θ, measured from the direction of the field. Depending on (ε, θ), the solutions can be both complex and real. The latter are of greatest interest, since their contribution does not contain an exponential smallness additional to with the tunneling exponential present in the integrand. This is why a plateau arises in the energy spectrum. The implicit functional relations between the four real parameters determined by equations (13) and (14) have been studied in detail in the three-step rescattering model [8]. We shall use these results, appropriately reformulated, to find the stationary points. In the three-step model equation (13) signifies that the freed electron, having left the atom at the time φ_0 with zero initial velocity and moving subsequently in a laser field, returns to the origin of coordinates at the time φ_1. At the moment it returns the electron is elastically scattered in accordance with the law of conservation of energy (14). The transformations

$$(\varphi_0, \varphi_1) \longrightarrow (\varphi_0 + 2\pi, \varphi_1 + 2\pi),$$
$$(\varphi_0, \varphi_1, \theta) \longrightarrow (\varphi_0 + \pi, \varphi_1 + \pi, \pi - \theta)$$

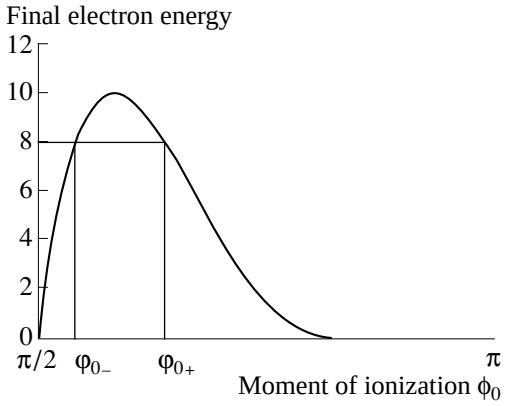


Fig. 1. Final energy of the rescattered electron as a function of the moment of ionization. The curve $\varepsilon(\varphi_0, \theta = \pi)/U_p$ was calculated using equations (13) and (14) in the first-return approximation. The form of the curve for $\theta \neq \pi$ is similar but the height of the maximum decreases monotonically with increasing angle: $\varepsilon_{cl}(\theta) < \varepsilon_{cl}(\pi) = 10U_p$. The contributions of the points φ_{0-} and φ_{0+} interfere in the rescattering amplitude in the state ($\varepsilon = 8U_p$, $\theta = \pi$).

do not change the form of the equations, and therefore it is sufficient to analyze the solutions on one half-period of the laser field, for example, $0 < \varphi_0 < \pi$. For moments of ionization in the interval $0 < \varphi_0 < \pi/2$ returns are impossible, and for $\pi/2 < \varphi_0 < \pi$ (depending on the value of φ_0) one to several returns (right up to an infinite number) are possible. The contribution from the later returns decreases rapidly because of the transverse spreading of the wave packet, and in what follows we shall consider only the first return. For electrons which are freed in the time interval $\pi/2 < \varphi_0 < \pi$ and possess a finite energy $p^2/2 > 2U_p$, the directions of emergence fall into the range $\pi/2 < \theta < \pi$. The instantaneous electron velocity appearing in the energy conservation law (14) rotates by the angle $\pi/2 < \theta_0 < \pi$, i.e., the electron is scattered backward with respect to the velocity which it possessed immediately prior to scattering. The angles θ and θ_0 are related by the relation [8]

$$(p/p_F)\cos\theta + \cos\varphi_1 = (\cos\varphi_1 - \cos\varphi_0)\cos\theta_0. \tag{15}$$

Rescattering in the interval $0 < \theta < \pi/2$ occurs on an adjoining half-period with the opposite orientation of the field. The energy of the rescattered electron, considered as a function of the moment of ionization, possesses an isolated maximum on the segment $\pi/2 < \varphi_0 < \pi$ (see Fig. 1). The height of the maximum $\varepsilon = \varepsilon_{cl}(\theta)$ is the upper limit of the spectrum, as predicted by the semiclassical model for electrons rescattered by the angle θ. It follows from the existence of a maximum in the function $\varepsilon = \varepsilon(\varphi_0, \theta)$ that the inverse function $\varphi_0 = \varphi_0(\varepsilon, \theta)$ required to calculate the amplitude (11) is two-valued. In other words, on a quarter period there are two moments in time φ_{0-} and φ_{0+} (see Fig. 1) which correspond to rescattering in a direction θ with energy ε. According to equation (13), their respective moments of return are φ_{1-} and φ_{1+}. To calculate the rescattering amplitude by the saddle-point method the phase in equation (12) must be expanded in a Taylor series around each of the points (φ_{0-}, φ_{1-}) and (φ_{0+}, φ_{1+}), a two-dimensional Gaussian integral must be calculated, and the results must be summed. Summing the contributions of the indicated stationary points engenders interference in the rescattering amplitude [13]

$$dW_r = (w_- + w_+ - 2\sqrt{w_- w_+}\sin(z_F s_{+-}))d^3p. \tag{16}$$

The contribution of an individual stationary point has the form (we omit the indices $\pm$)

$$w(\varphi_1, \varphi_0) = \frac{2IC^2(F)U^2(q)\omega^3\sin^2\varphi_0}{(2\pi)^4 F^2 F_a \Delta_\perp^2(\varphi_1, \varphi_0)|D|} \times \exp\left(-\frac{2F_a}{2F(\varphi_0)}\right). \tag{17}$$

The following notation has been introduced in equations (16) and (17):

$$s = S/z_F,$$

where the phase S is determined by equation (12) without the term $I\varphi_0$;

$$s_{+-} = s_+ - s_-$$

is the difference of the reduced phases at the stationary points;

$$\Delta_\perp^2(\varphi_1, \varphi_0) = \Delta_0^2 + (\varphi_1 - \varphi_0)^2/(\omega\Delta_0)^2$$

is the squared transverse width of the wave packet of a freed electron at the moment of return [21];

$$D = s_{00}s_{11} - (s_{01})^2$$

is the determinant of the matrix of second derivatives of the phase with respect to φ_0 and φ_1 at the expansion point. The sign of D is different for the two branches of stationary points, and therefore the interference term in equation (20) contains a sine and not a cosine, as in [13, 16].

The procedure for summing the contributions of independent stationary points is inapplicable if the final state (ε, θ) lies near the classical boundary. It is evident from Fig. 1 that as $\varepsilon \longrightarrow \varepsilon_{cl}(\theta)$ the points φ_{0-} and φ_{0+} approach from different sides the point $\varphi_{0m}(\theta)$ where the function $\varepsilon = \varepsilon(\varphi_0, \theta)$ reaches a maximum, and it cannot be assumed that they are isolated. A manifestation of this circumstance is that the distribution (17) becomes infinite in the limit $\varepsilon \longrightarrow \varepsilon_{cl}(\theta)$, since $D \longrightarrow 0$ in the denominator. At the boundary itself $D = 0$, and this equality is equivalent to the condition for an extremum

$$\partial\varepsilon(\theta, \varphi_0)/\partial\varphi_0 = 0.$$

Indeed, writing equations (13) and (14) in the form

$$s_0(\varphi_0, \varphi_1(\varphi_0)) = 0, \quad s_1(\theta, \varphi_0, \varepsilon(\theta, \varphi_0), \varphi_1(\varphi_0)) = 0,$$

differentiating with respect to φ_0 with constant θ, and eliminating $d\varphi_1/d\varphi_0$, it is easy to see that

$$\partial\varepsilon(\theta, \varphi_0)/\partial\varphi_0 \propto D$$

(here and below the partial derivatives of the reduced phase s with respect to the times φ_0 and φ_1 are denoted by $\partial s/\partial\varphi_0 \equiv s_0$, $\partial^2 s/\partial\varphi_0\partial\varphi_1 \equiv s_{01}$, and so on).

Near the boundary the method for calculating the integral must be modified as follows. Since we are interested in the final state (θ, ε) close to the boundary, let us consider the state $(\theta, \varepsilon_{cl}(\theta))$ that lies on the boundary and to which the stationary points $\varphi_{0+} = \varphi_{0-} = \varphi_{0m}(\theta)$ and $\varphi_{1+} = \varphi_{1-} = \varphi_{1m}(\theta)$ correspond in the plane of integration variables. We expand the phase (12) with respect to the deviations $\delta\varphi_0 = \varphi_0 - \varphi_{0m}$ and $\delta\varphi_1 = \varphi_1 - \varphi_{1m}$ in a Taylor series as follows:

$$s = s(\varphi_{0m}, \varphi_{1m}) + s_1\delta\varphi_1 + \frac{1}{2}[s_{00}(\delta\varphi_0)^2 + 2s_{01}\delta\varphi_0\delta\varphi_1 + s_{11}(\delta\varphi_1)^2] + \frac{1}{6}s_{111}(\delta\varphi_1)^3. \quad (18)$$

In equation (18) the energy is a free parameter, and therefore at the chosen point of expansion $s_1 \neq 0$, while the ε-independent derivative $s_0 = 0$ is absent in equation (18). The substitutions

$$\xi = \delta\varphi_0 + s_{01}\delta\varphi_1/s_{00}, \quad \eta = \delta\varphi_1$$

reduce the quadratic form in equation (18) to the diagonal form

$$s_{00}\left(\xi^2 + D\eta^2/s_{00}^2\right).$$

It is evident that when calculating the amplitude (11) under the conditions $D \longrightarrow 0$, the integral over the variable η diverges. To regularize the integral it is sufficient to take into account the next term in the expansion, specifically, $s_{\eta\eta\eta}\eta^3$. An additional investigation shows that we can set $s_{\eta\eta\eta} \approx s_{111}$ in the region important for the integral, as done in equation (18). Using the expansion (18), the integral (11) can be calculated analytically, and the spectral-angular distribution of the photoelectrons near the classical boundary can be written as [17]

$$dW_r = \frac{I\omega^3 C^2(F)U^2(p)\sin^2\varphi_{0m}\exp(-2F_a/3F(\varphi_{0m}))}{4\pi^3 F_a F^{4/3}\Delta_\perp^2(\varphi_{0m}, \varphi_{1m})|s_{00}||s_{111}/2|^{2/3}} \times \mathrm{Ai}^2\left\{z_F^{2/3}Y(\varepsilon, \theta)\right\}d^3p. \quad (19)$$

The smooth function of order one

$$Y(\varepsilon, \theta) = (2/|\varepsilon_{111}|)^{1/3}[s_1 - D^2/(2|s_{111}|s_{00}^2)] \quad (20)$$

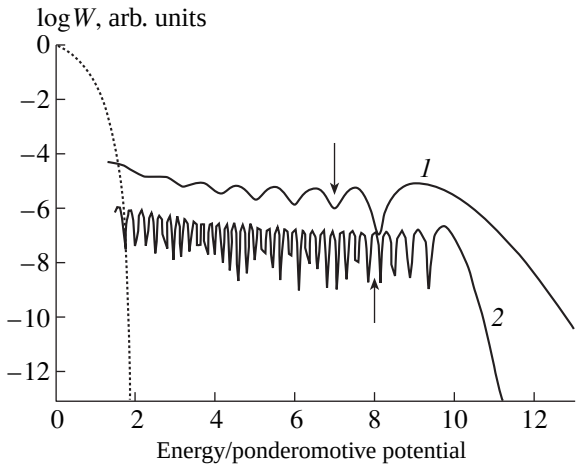


Fig. 2. Rescattering spectrum in the direction of polarization of the field, calculated using equations (16) and (19) for ionization of helium (I = 24.6 eV) by Ti:Sa-laser radiation ($\hbar\omega$ = 1.58 eV). Curve *1*: intensity 2×10^{14} W/cm^2 (γ = 1, z_F = 30, F/F_a = 0.03). Curve *2*: intensity 10^{15} W/cm^2 (γ = 0.46, z_F = 148, F/F_a = 0.07). The curves are normalized to unit value of the direct-ionization probability (dotted line) in a state with zero energy. The arrows mark the locations where the distributions (16) and (19) match.

vanishes on the classical boundary and assumes positive and negative values, respectively, for $\varepsilon > \varepsilon_{cl}(\theta)$ and $\varepsilon < \varepsilon_{cl}(\theta)$. It follows from the properties of the Airy function $\mathrm{Ai}(x)$ that the distribution (19) has no singularities at the classical boundary and describes the transition from monotonic decay beyond the limits of the classical region to an interference structure with minima and maxima inside this region.

The distributions (16) and (19) match fairly well in the classically allowed region. In the first place, replacing $\mathrm{Ai}(x)$ in equation (19) by its asymptotic form (the term s_1 in equation (20) must be neglected in so doing) and averaging over the oscillations ($\sin^2\varphi_0 \longrightarrow 1/2$) gives a result which is literally identical to the distribution (16), in which the sign-alternating term must be dropped and the contributions $w_- = w_+$ must be taken at the same point $(\varphi_{0m}, \varphi_{1m})$. In the second place, superposing the spectra tabulated according to the complete formulas (16) and (19) shows that on a small section (marked by arrows in Fig. 2) the interference spikes in both curves coincide to graphical accuracy. An estimate [17] shows that a match obtains near the energy

$$\varepsilon_*(\theta) \approx \varepsilon_{cl}(\theta) - (40\text{–}60)z_F^{-2/3}U_p.$$

The existence of a match is an additional argument in favor of the approximation (18).

The small term $\gamma^2/2\sin\varphi_0$, arising from the differentiation of the term $I\varphi_0$ in the phase (12) and equal to the ratio of the width of the oscillating potential barrier to the amplitude of the electron oscillations in the laser field, was dropped in equation (13). Just as in the problem of the generation of high harmonics [24], this con-

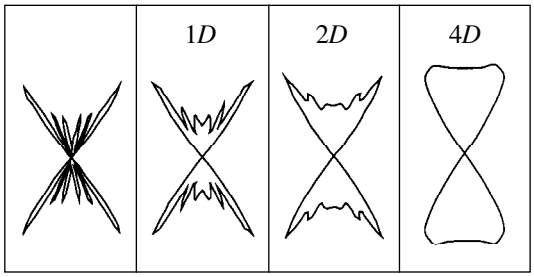


Fig. 3. Angular distributions of electrons with energy $\varepsilon = 7U_p$, calculated from equation (19) and averaged over a Gaussian intensity distribution at the laser focus in accordance with equation (23). The curves are normalized to a unit value at the maximum. The peak intensity at the center of the focus is 3.5×10^{14} W/cm^2, which corresponds to $\gamma = 0.77$ and $z_F = 51$ for a He atom and a Ti:Sa laser. The first figure shows the unaveraged distribution. The second figure shows the result of one-dimensional averaging (for example, over the temporal nonuniformity of the laser pulse), and so on.

tribution shifts the classical boundary of the spectrum into the high-energy range by an amount of the order of the ionization potential.

4. DISCUSSION OF THE RESULTS

The spectral-angular distribution of the rescattered photoelectrons (16), (17), and (19) agrees with the numerical results of model, but precise, quantum calculations [11, 12] and describes these experiments completely satisfactorily [3–5]. Our approximate expressions show explicitly the dependence of the characteristics of the distribution on the parameters of the atom and the laser.

4.1. Interference Structure of the Spectrum

Figure 2 shows the electron energy spectrum, calculated using equations (16) and (19), in the direction of the field. Just as in [11, 12], the interference spikes increase in size as the plateau boundary is approached, so that the last maximum, beyond which the distribution decays into the classically inaccessible region, has the largest width and the greatest height.

A change in the laser intensity changes the number of spikes within the plateau. Calculations show that the function $s_{+-}(\varepsilon/U_p, \theta = 0)$ decreases monotonically, almost linearly, from the value $s_{+-}(2, 0) = 1.33$ to 0 at $\varepsilon = 10U_p$. Hence it follows that within the plateau there are approximately $z_F/4$ interference maxima, whose width $\Delta\varepsilon \approx 8\omega$ on most of the plateau is independent of the intensity. A different characteristic scale is obtained from equation (19) for several interference maxima adjoining the classical boundary. Thus, the distance between the greatest maximum at the end of the spectrum and the preceding dip is

$$\Delta\varepsilon \approx 4U_p^{1/3}\omega^{2/3}.$$

In estimates of other dimensions of the interference pattern in this part of the spectrum, only the numerical factor will change.

A decrease of the laser intensity decreases not only the number of spikes but also the degree of modulation of the interference pattern (see curve *1* in Fig. 2). The reason is that for a large value of the ratio F_a/F the interfering amplitudes differ more strongly from one another because of the tunneling exponential. The interference relief structure remains only at the end of the plateau, where w_+ and w_- are always close.

The heights of the maxima in the spectrum (Fig. 2) and in the angular distribution (Fig. 3) vary monotonically. Similar plots in [11, 12] show that the individual maxima destroy this monotonicity. The nonmonotonicity is due to the contribution of the subsequent returns [25], which were neglected in our calculations.

4.2. The Atomic Potential and the Form of the Spectrum

Thus far, the form of the scattering potential has not been specified. The question of the choice of a potential does not arise in the single-electron problem (hydrogen atom); the potential in the rescattering amplitude (4) must be the same as the potential used to find the initial wave function in the direct-ionization amplitude. However, if the model with one active electron is applied to a multielectron atom, it is natural to assume that the freed electron moves in the average field of the atomic core. To calculate the initial state of the external electron in this complex potential, large distances, where the nuclear charge is screened essentially to 1, are substantial. At the same time, the energy of the electrons returning to the parent ion is of the order of $U_p \gg I$, and their scattering into a state on the plateau is determined by the structure of the average potential at short distances. For this reason, it is physically entirely justified to use different potentials to calculate the direct ionization and rescattering.

The experimentally observed behavior in the electron distributions on the plateau [3–5] is described well by the expressions (16), (17), and (19) with a potential of the form

$$U(q) = 4\pi\{Z - n(q)\}/q^2, \qquad (21)$$

where the electron density for helium is

$$n(q) = 1/(1 + q^2/16)^2,$$

and for a singly charged ion of a complex atom

$$n(q) = (Z-1)/\left[1 + (0.48qZ^{-1/3})^2\right]$$

(see [26]). For rescattering into a state on the plateau the transferred momentum is determined primarily by the magnitude p of the final momentum (see the discussion preceding equation (11)). For this reason, the factor $U^2(q)$ in equations (17) and (19) strongly influences the energy spectrum and has virtually no effect on the

form of the angular distribution. The small slope of the plateau in the direction of high energies [4] is due to the Coulomb factor q^{-2} in equation (21). A zero-radius scattering potential gives a horizontal plateau [11]. On most of the plateau the transferred momentum is much greater than the atomic momentum and the effective charge in equation (21) is close to the nuclear charge Z, i.e., for otherwise equal conditions, rescattering by heavy atoms is more effective [4, 13].

The plateau level with respect to the direct-ionization spectrum moves lower as the laser intensity increases [4] (see Fig. 2). Since

$$U^2(q) \propto \varepsilon^{-2} \propto U_p^{-2}$$

and at the moment of return

$$\Delta_\perp^2 \propto \Delta_0^{-2} \propto F,$$

the field dependence in equation (17) can be written in the form

$$\frac{dW_r}{d^3p} \propto C^2(F)F^{-7}\exp\left(-\frac{2F_a}{3F}\right),$$

while for direct ionization

$$\frac{dW_d}{d^3p} \propto C^2(F)F^{-1}\exp\left(-\frac{2F_a}{3F}\right).$$

The ratio of these probability densities decreases with increasing field as F^{-6}. Multiplying the density by the corresponding phase volume for rescattering $p_F^3 \propto F^3$ and for direct ionization $(p_\perp^2 p_x)_{eff} \propto F^{5/2}$, we find that the fraction of the electrons rescattered into all states of the plateau varies with the field as

$$\mu = W_r/W_d \propto F^{-5.5}.$$

The experimentally measured dependence has the form [4]

$$\mu_{\exp} \propto F^{-5}.$$

We note that μ does not depend on the Coulomb correction $C(F)$, but it depends strongly on the form of the scattering potential. For scattering by a zero-radius potential $\mu \propto F^{-1.5}$. The Landau–Dykhne method [27] gives the correct position of the plateau boundary but it does not permit discussing the relative height of the plateau.

Taking account of the Coulomb factor $C(F)$ for scattering by the hydrogen atom ($Z = 1$, $n(p) = 0$) the total rescattering probability is

$$W_r = \lambda I\gamma^6\exp(-2F_a/3F),$$

where λ is a number of the order of 1, arising in the integration over dimensionless variables.

4.3. Angular Distributions

The first diagram in Fig. 3 for electrons with energy $\varepsilon = 7U_p$ (see [17] for a more detailed discussion) demonstrates the typical, for the center of a plateau, angular distribution where the side maximum predominates. As ε increases, the cone angle into which scattering occurs and the number of interference spikes within the cone angle decrease. In this part of the plateau the angular distribution is described by equation (19) and is determined primarily by an Airy function. In the region $\theta \le 30°$, $\varepsilon \ge 8U_p$ and for $z_F^{2/3} > 7.7$ its argument is approximated well by the expression

$$x(\varepsilon, \theta) \equiv z_F^{2/3}Y(\varepsilon, \theta) \approx 0.13z_F^{2/3}\left(\frac{\varepsilon}{U_p} - 10 + 7.8\theta^2\right), \quad (22)$$

which makes it possible to find easily in an explicit form the position of the characteristic points and the critical values of the energy, which determine the qualitative change in the form of the angular distribution. Specifically, the positions of the large side maximum $\theta_{\max}(9U_p) = 17.5°$ and $\theta_{\max}(8U_p) = 27°$ predicted by equation (22) with $z_F = 146$ do not differ much from the numerical results $\theta_{\max}(9U_p) = 20°$ and $\theta_{\max}(8U_p) = 30°$ [12]. At the end of the plateau the angular diagram consists of a single lobe of width $\approx z_F^{-1/3}$, extending in the direction of the field.

As ε decreases below $7U_p$ compared with Fig. 3, the scattering cone expands and the number of interference spikes increases. In these angular diagrams the distribution (19) describes a large outer maximum and its immediate vicinity. The rest of the distribution must be calculated using the formula (16). The spikes near $\theta = 0, \pi$ are lower than near the maximum angle and, therefore, averaging the distribution over the interference oscillations (actually dropping the last term in equation (16)) gives a smooth angular distribution with a minimum in the direction of the field. For energies close to $8U_p$ the ratio of the height of this minimum to the height of the main side maximum, estimated using equations (19) and (22), is $\approx z_F^{-1/3}$ [17]. Hence it follows that the minimum in the average angular distribution at $\theta = 0, \pi$ is a deep dip for $z_F^{1/3} \gg 1$ and a small decrease for $z_F^{1/3}$ of the order of 1. The angular distribution of rescattered electrons with a deep dip in the direction of the field was first predicted on the basis of a three-step model [8]. It was concluded that the model underestimates the rescattering by angles close to 0 and π. However, as shown below in Subsection *4.4*, the three-step model adequately describes precisely this range of angles, but it is inapplicable near the large side maximum and therefore cannot correctly predict its height.

A dip is not seen in the angular distributions measured in the tunneling regime on the plateau [4]. This is due to the nonuniformity of the laser radiation. In a field with envelope $F(\mathbf{r}, t)$ the contribution of the space-

time point to the measured density of the momentum distribution is proportional to $w(\varepsilon, \theta, F(\mathbf{r}, t))d^3rdt$, where $w(\varepsilon, \theta, F)$ is determined in equation (16) or (19). Integrating this expression successively over separate variables, it is possible to trace how a superposition of the distributions corresponding to different laser field intensities averages and smears the interference pattern. This is especially easily done for a Gaussian field profile in space and time, for which the result of the n-fold integration ($1 \le n \le 4$) has the form

$$\langle w(\varepsilon, \theta)\rangle_n \propto \int u^{n-1} w(\varepsilon, \theta, F(u))du, \qquad (23)$$

where

$$F(u) = F_0 \exp(-u^2/2)$$

and F_0 is the peak field in the focused laser pulse. It should be underscored that in the integral (23) the electron energy ε is fixed, and the dimensionless combination $\varepsilon/U_p(u)$ appearing in the momentum distribution depends on the variable of integration. Figure 3 shows the evolution of the angular distribution of electrons with $\varepsilon = 7U_p(0)$ in subsequent averaging. The last diagram attests to the fact that for field and atom parameters which are standard for a modern experiment complete averaging over the volume of the focus and over the time completely smears the interference pattern and the expected minimum of the angular distribution in the direction of polarization. This diagram agrees qualitatively with the measurements performed in [4].

4.4. Relation with the Three-Step Model

We now switch in the probability (16) from quantities characterizing the final state (p, θ, β) to the variables (φ_0, θ_0, β) used in the three-step model [8]. The equations (13)–(15) give the relation between them. The azimuthal angle β in both sets of variables is the same. We have for the differentials of the variables

$$d\varphi_0 d\Omega_0 = Jdpd\Omega, \qquad (24)$$

where the Jacobian of the transformation is

$$J = \frac{\partial(\cos\theta_0, \varphi_0)}{\partial(\cos\theta, p)} = \frac{(\partial\cos\theta_0/\partial\cos\theta)_{\varphi_0}}{(\partial p/\partial\varphi_0)_\theta} = \frac{p^2\sin\varphi_0}{p_F^3|D|}. \qquad (25)$$

Dropping the interference term in equation (16) and using equation (25) as well as the fact that w_- and w_+ are determined as functions of φ_0, respectively, in the intervals $\pi/2 < \varphi_0 < \varphi_m(\theta)$ and $\varphi_m(\theta) < \varphi_0 < \pi$, we obtain in the new variables

$$dW_r = \frac{\sigma(q)}{\pi\Delta_\perp^2(\varphi_1, \varphi_0)} W_{st}(F(\varphi_0))\frac{dt_0 d\Omega_0}{T}. \qquad (26)$$

Here $\pi/2 < \varphi_0 < \pi$ and the following notations have been introduced:

$$\sigma(q) = (U(q)/2\pi)^2$$

is the Born differential scattering cross section, $q = 2p_F|\cos\varphi_1 - \cos\varphi_0|\sin(\theta_0/2)$;

$$W_{st}(F) = IC^2(F)(F/F_a)\exp(-2F_a/3F)$$

is the probability of ionization by a static field; and,

$$F(\varphi_0) = F|\sin\varphi_0|, \quad T = 2\pi/\omega.$$

The formula (26) determines the probability per unit time that at the moment of return $W_{st}(\varphi_0)dt_0$ electrons freed in a time dt_0 are scattered into the solid angle $d\Omega_0$; this corresponds precisely to the assertion of the three-step model. Thus the three-step model and the approximate quantum calculation by the saddle-point method using isolated stationary points give the same (to within the interference term) distribution of rescattered electrons, only written in different variables.

It follows from the equivalence of these two approaches that their regions of applicability are the same and therefore the three-step model is inapplicable near the classical boundary. In the variables of the three-step model (26), the distribution itself has no singularities. The transformation of variables

$$(\varphi_0, \theta_0) \longrightarrow (p, \theta)$$

introduces a singularity, and therefore it is of a universal character: it is conserved for any form of the ionization-time distribution and of the scattering cross section as well as when the effect of the Coulomb field of the ion on the classical electron trajectory in the continuum is taken into account.

Ordinarily, in the three-step model [3–5, 8] the transition to measured quantities is done numerically. The distribution given on a discrete grid (φ_0, θ_0) is projected onto the grid in the (p, θ) plane. The finite sizes of the grid cells mask the singularity of the transformation. The singularity is nonetheless manifested in the fact that as the size of the cells in the plane of the final states (p, θ) decreases, the distribution increases near the classical boundary while remaining unchanged in the interior region of the phase space. Apparently, this computational instability explains the maximum in the energy spectrum in Fig. 3 in [3] as well as the exceedingly large magnitude of the angular distribution near the limiting rescattering angle [8].

5. CONCLUSIONS

The semiclassical analysis, presented above, of the electronic distributions in the region of the high-energy plateau makes it possible to describe in an explicit form the dependence of the distributions on the parameters of the laser field and the atom. The theory preserves the freedom of choice of the specific form of the potential of the parent ion and contains no other free parameters.

The results presented agree well, not only qualitatively but also quantitatively, with measurements in the tunneling regime and with the numerical calculations of the interference structure of electronic distributions performed using a generalized Keldysh model.

ACKNOWLEDGMENTS

We are grateful to the participants of N.B. Delone's seminar at the Institute of General Physics of the Russian Academy of Sciences for many stimulating discussions.

This work was supported by the Russian Foundation for Basic Research (project no. 99-02-17810) and the Competitive Center for the Fundamental Natural Sciences of the Ministry of Education of the Russian Federation (project no. 97-5.2-117).

REFERENCES

1. L. F. DiMauro and P. Agostini, Adv. At., Mol., Opt. Phys. **35**, 79 (1995).
2. N. B. Delone and V. P. Kraĭnov, Usp. Fiz. Nauk **168**, 531 (1998) [Phys. Usp. **41**, 469 (1998)].
3. L. F. DiMauro, K. C. Kulander, and P. Agostini, in *Super-Intense Laser-Atom Physics IV*, Ed. by H. G. Muller and M. V. Fedorov (Kluwer, Dordrecht, 1995), p. 97.
4. B. Sheehy, B. Walker, R. Lafon, *et al.*, in *Proceedings of the 7th International Conference on Multiphoton Processes, Garmisch-Partenkirchen, 1996*, Ed. by P. Lombropolus and H. Walther (Inst. of Physics Pub., Boston, 1997), No. 154, p. 106.
5. B. Sheehy, R. Lafon, M. Widmer, *et al.*, Phys. Rev. A **58**, 3942 (1998).
6. B. Yang, K. J. Schafer, B. Walker, *et al.*, Phys. Rev. Lett. **71**, 3770 (1993).
7. P. B. Corkum, Phys. Rev. Lett. **71**, 1994 (1993).
8. G. G. Paulus, W. Becker, W. Nicklich, *et al.*, J. Phys. B **27**, L703 (1994).
9. D. F. Zaretskiĭ and É. A. Nersesov, Zh. Éksp. Teor. Fiz. **103**, 1192 (1993) [JETP **76**, 583 (1993)].
10. M. Lewenstain, K. C. Kulander, K. J. Schafer, *et al.*, Phys. Rev. A **51**, 1495 (1995).
11. A. Lohr, M. Kleber, R. Kopold, *et al.*, Phys. Rev. A **55**, R4003 (1997).
12. D. B. Milosevic and F. Ehlotzky, Phys. Rev. A **58**, 3124 (1998).
13. S. P. Goreslavskii and S. V. Popruzhenko, Phys. Lett. A **249**, 477 (1998).
14. A. I. Perelomov, V. S. Popov, and M. V. Terent'ev, Zh. Éksp. Teor. Fiz. **50**, 1393 (1966) [Sov. Phys. JETP **23**, 924 (1966)].
15. G. G. Paulus, F. Zacher, H. Walther, *et al.*, Phys. Rev. Lett. **80**, 484 (1998).
16. S. P. Goreslavskiĭ and S. V. Popruzhenko, Pis'ma Zh. Éksp. Teor. Fiz. **68**, 858 (1998) [JETP Lett. **68**, 902 (1998)].
17. S. P. Goreslavskiĭ and S. V. Popruzhenko, J. Phys. B **32**, L533 (1999).
18. H. G. Muller, Comments At. Mol. Phys. **24**, 355 (1990).
19. L. V. Keldysh, Zh. Éksp. Teor. Fiz. **47**, 1945 (1964) [Sov. Phys. JETP **20**, 1307 (1964)].
20. F. V. Bunkin and M. V. Fedorov, Zh. Éksp. Teor. Fiz. **49**, 1215 (1965) [Sov. Phys. JETP **22**, 844 (1966)].
21. S. P. Goreslavskiĭ and S. V. Popruzhenko, Laser Phys. **8**, 1013 (1998).
22. A. M. Perelomov, V. S. Popov, and M. V. Terent'ev, Zh. Éksp. Teor. Fiz. **52**, 514 (1967) [Sov. Phys. JETP **25**, 336 (1967)].
23. V. P. Kraĭnov and B. Shokri, Zh. Éksp. Teor. Fiz. **107**, 1180 (1995) [JETP **80**, 657 (1995)].
24. M. Lewenstain, Ph. Balcou, M. Yu. Ivanov, *et al.*, Phys. Rev. A **49**, 2117 (1994).
25. R. Kopold, W. Becker, and M. Kleber, submitted to Optics Communications (1999).
26. W. Brandt and M. Kitagawa, Phys. Rev. B **25**, 5631 (1982).
27. M. B. Smirnov and V. P. Kraĭnov, J. Phys. B **31**, L519 (1998).

Translation was provided by AIP

Journal of Experimental and Theoretical Physics, Vol. 90, No. 5, 2000, pp. 788–793.
Translated from Zhurnal Éksperimental'noĭ i Teoreticheskoĭ Fiziki, Vol. 117, No. 5, 2000, pp. 906–912.

ATOMS, SPECTRA, RADIATION

Polarization Bremsstrahlung of a Fast Charged Particle on a Thomas–Fermi Atom

V. A. Astapenko*, L. A. Bureeva, and V. S. Lisitsa*****

**Moscow Physicotechnical Institute, Dolgoprudnyĭ, Moscow oblast, 141700 Russia*
e-mail: astval@hotmail.com

***Scientific Council on Spectroscopy, Russian Academy of Sciences, Moscow, 117924 Russia*

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

Received December 16, 1999

Abstract—An universal description of the polarization bremsstrahlung of a fast charged particle on a multielectron atom ($Z \gg 1$, Z is the nuclear charge) is obtained using the local electron density method and the Thomas–Fermi statistical model. It is shown that the cross section of the process can be represented in the form $d\sigma^{PB}(\omega) = Z^2 d\tilde{\sigma}^{PB}(\nu)$, where the function $d\tilde{\sigma}^{PB}(\nu)$ exhibits approximate scaling with respect to the parameter $\omega/Z = \nu$, and the corresponding R factor (ratio of the cross sections in the polarization and ordinary channels) is greater than 1 in a wide range of frequencies and reaches its maximum value at frequencies $\omega \approx Z\,\mathrm{Ry}$. It is demonstrated that in the frequency range $p_a c < \hbar\omega < \gamma^2 p_a c$ (γ is the relativistic factor, p_a is the characteristic momentum of the atomic electrons, and c is the speed of light) the angular distribution of the polarization bremsstrahlung of a relativistic charged particle undergoes narrowing due to the compensation, by the momentum of the emitted photon, of the momentum transfer from the incident particle to the atomic core. © *2000 MAIK "Nauka/Interperiodica".*

1. INTRODUCTION

The local electron density method, which is widely used to calculate atomic characteristics [1], has been used to investigate polarization bremsstrahlung and its interference with ordinary bremsstrahlung in the case of strong inelastic scattering of quasiclassical electrons with moderate energies by ions with a core [2]. This is situation is of current interest for radiation processes occurring in plasma with the participation of plasma electrons with thermal energies. The advantages of the local electron density method are due to the universal, for all atoms, character of the results which it yields. It is of interest to use this approach to calculate the polarization bremsstrahlung of a fast, including relativistic, charged particle on a multielectron system.

As is well known [3], a variety of features appear in the bremsstrahlung of a relativistic electron (positron) in the ordinary (static) channel: the directional pattern of the radiation becomes narrower as the energy of the incident particle increases, and a density effect, which suppresses the static bremsstrahlung at low frequencies, arises in the radiation in a medium. As first shown in [4, 5], polarization bremsstrahlung is modified in a special manner for relativistic particles: its spectral cross section increases logarithmically with increasing energy of the incident particle, and for not too short wavelengths ($\lambda > R_a$, λ is the wavelength of the radiation and R_a is the characteristic size of the atom) the directional pattern retains its dipole character.

The purpose of the present paper is to investigate, using the local electron density method for a Thomas–Fermi atom, the frequency and angular dependences of the polarization bremsstrahlung cross section of a fast charged particle.

2. POLARIZABILITY OF A THOMAS–FERMI ATOM IN THE LOCAL ELECTRON DENSITY APPROXIMATION

As is well known [6], the main characteristic determining the bremsstrahlung cross section in the polarization channel is the nondipolar dynamic polarizability of the atom $\alpha(\omega, q)$. The systematic quantum-mechanical calculation of this quantity for a multielectron atom is a very complicated computational problem that must be solved anew for each specific target [7].

In this connection it is helpful to use simple universal models suitable for estimating the polarization bremsstrahlung cross section and determining the general qualitative characteristics of this process.

One such model is the local electron density method (or local plasma frequency method), first proposed by Brandt and Lundquist for calculating the photoelectric absorption cross section of multielectron atoms [1].

In this model the expression for the dynamic polarizability of an atom has the form

$$\alpha(\omega) = \int_0^\infty \frac{\omega_p^2(r) r^2 dr}{\omega_p^2(r) - \omega^2 - i0}. \quad (1)$$

Here $\omega_p(r) = \sqrt{4\pi e^2 n(r)/m}$ is the local plasma frequency, which depends on the local electron density $n(r)$ of the electronic core.

The Thomas–Fermi approximation will be used below for the function $n(r)$. This approximation gives [8]

$$n(r) = n_{TF}(r) = Z^2 f\left(\frac{r}{r_{TF}}\right), \qquad f(x) = \frac{32}{9\pi^3}\left(\frac{\chi(x)}{x}\right)^{3/2}. \quad (2)$$

Here $r_{TF} = ba_0/Z^{1/3}$ is the Thomas–Fermi radius ($b = (9\pi^2/128)^{1/3} \simeq 0.8853$, a_0 is the Bohr radius, and Z is the nuclear charge), and $\chi(x)$ is the Thomas–Fermi function.

The expression (1) can be transformed into a form that reveals the similarity (scaling) law with respect to the parameter $\nu = \hbar\omega/2\,\mathrm{Ry}\,Z$ (Ry = 13.6 eV):

$$\alpha(\nu) = r_{TF}^3 \beta(\nu) = \frac{b^3 a_0^3}{Z}\beta(\nu). \quad (1a)$$

Here a dimensionless complex function $\beta(\nu)$ (the reduced polarizability of a Thomas–Fermi atom) has been introduced. The imaginary part of this function is (the prime denotes differentiation with respect to the argument x)

$$\mathrm{Im}\{\beta(\nu)\} = \pi \frac{f(x_\nu) x_\nu^2}{|f'(x_\nu)|}, \quad (3a)$$

and the real part can be calculated from the Kramers–Kronig relation:

$$\mathrm{Re}\{\beta(\nu)\} = \frac{2}{\pi}\int_0^\infty \frac{[\mathrm{Im}\{\beta(\nu)\} - \mathrm{Im}\{\beta(\tilde\nu)\}]\tilde\nu d\tilde\nu}{\nu^2 - \tilde\nu^2}. \quad (3b)$$

In equations (3a) and (3b) the quantity x_ν is determined by the solution of the equation

$$4\pi f(x) = \nu^2. \quad (4)$$

The dependence $g(\nu)$ of the ratio of the modulus of the function $\beta(\nu)$, calculated using equations (3a) and (3b), to the modulus of its high-frequency limit ($\beta_\infty(\nu) = -\nu^{-2}$) is presented in Fig. 1. The figure also shows for a KrI atom the corresponding dependence reconstructed from the data of [7], where the dynamic polarizability of an atom was calculated quantum mechanically on the basis of the random-phase with exchange approximation.

It is evident that the function $g(\nu) = \nu^2|\beta(\nu)|$ for a Thomas–Fermi atom is maximum for small values of the dimensionless parameter $\nu \approx 0.5$ and approaches its high-frequency limit for $\nu > 10$.

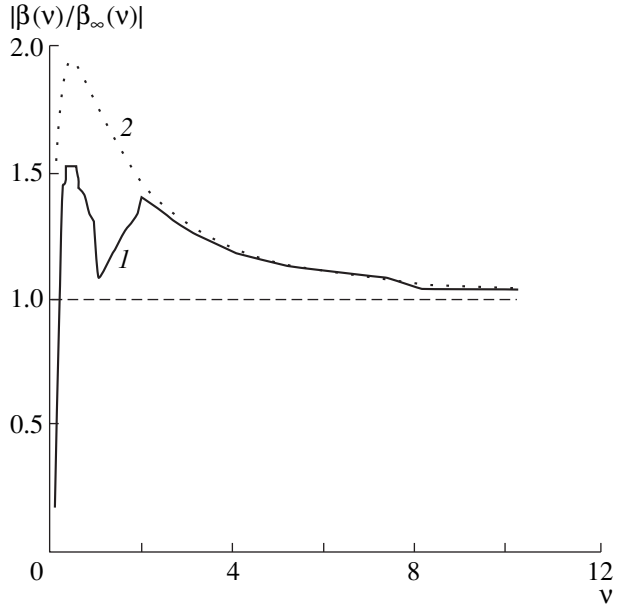


Fig. 1. Dynamical polarizabilities, normalized to their high-frequency limit, as a function of the dimensionless frequency $\nu = \hbar\omega/2Z\,\mathrm{Ry}$ for a KrI atom according to data from [7] (curve *1*) and for a Thomas–Fermi atom in the local electron density model (curve *2*).

However, it should be kept in mind that the approximation used becomes inadequate at low frequencies $\nu < 0.1$, since, on the one hand, the local plasma frequency approximation for the polarizability (1) of the atom works poorly, and on the other hand the peripheral regions of the atom, where the statistical model is inapplicable, contribute to the polarizability at these frequencies. Indeed, a calculation using equation (4) gives $x_{0.1} = 3.4$, $x_1 = 0.64$, and $x_{10} = 0.053$, while the region of validity of the statistical model in the variable x is determined by the inequality $Z^{-2/3} \ll x \ll Z^{1/3}$.

Consequently, the analysis below will be confined to the range $\nu > 0.1$. We note that for $Z \approx 50$ this corresponds to photon energies $\hbar\omega > 130$ eV, which is much higher than the ionization potential of the outer electron shell of a neutral atom, so that the electronic core can be assumed to be "unfrozen." In addition, as one can see from the values of x_ν presented above, the inequality $x_\nu \le 3.4$ holds in this frequency range. According to [9], the limiting reduced radius of a neutral atom as calculated in the Thomas–Fermi–Dirac model (with exchange) is approximated well by the formula $x_0 = 4Z^{0.4}$. Thus, in our case ($Z \gg 1$) $x_\nu \ll x_0$ and the results of the further analysis are essentially independent of the refinements of the initial statistical Thomas–Fermi model; they are also valid for ions with a sufficiently low ionization density, if $x_0(Z_i/Z) \gg x_\nu$, which is confirmed by the calculation performed.

As one can see from Fig. 1, good agreement between the modulus of the dynamic polarizability of a Thomas–Fermi atom, calculated in the local electron density approximation, and the results of the quantum-mechanical calculations [7] obtains for the dimensionless frequency $\nu > 2$. Both approaches give the same values of the frequency at the maximum of the function $g(\nu)$: $\nu = 0.5$ or $\hbar\omega_{max} = 490$ eV, so that $\hbar\omega_{max} \gg I_p(\mathrm{KrI}) = 14$ eV and the electronic core of KrI can be assumed to "unfrozen." Near the maximum of the function $g(\nu)$ the difference is less than 30% in the case at hand.

As shown in [2], the static polarizabilities in the statistical model on the basis of the local electron density approximation agree satisfactorily with the results of more systematic calculations [10], performed in a statistical model of the atom, and with the experimental results.

3. POLARIZATION BREMSSTRAHLUNG CROSS SECTION

The spectral polarization bremsstrahlung cross section of a fast electron on an atom is described in the Born approximation by the expression [6, Chap. 5] (hereafter, we employ the relativistic system of units, $\hbar = c = m_e = 1$)

$$\frac{d\sigma^{PB}}{d\omega} = \frac{\omega^5}{(2\pi)^3 v} \times \int d\Omega_{\mathbf{n}} d\mathbf{q} |\alpha(\omega, \mathbf{q} + \mathbf{k})|^2 [\mathbf{n} \times \mathbf{A}(q)]^2 \delta(\omega + \mathbf{q} \cdot \mathbf{v}). \quad (5)$$

Here $d\Omega_{\mathbf{n}}$ is the solid angle in the direction of emission of the photon, $\mathbf{k}$ and ω are, respectively, the wave vector and frequency of the bremsstrahlung photon, $\mathbf{q} = \mathbf{p}_f - \mathbf{p}_i$ is the change in the momentum of the incident particle, and $\mathbf{A}(q)$ is the space–time Fourier transform of the vector potential of the electromagnetic field of the incident particle, which in the axial gauge ($A_0 = 0$) is given by the expression

$$\mathbf{A}(q^0, \mathbf{q}) = \frac{4\pi e}{q^0} \frac{\mathbf{q} - \mathbf{v}q^0}{\mathbf{q}^2 - (q^0)^2}. \quad (6)$$

We note that the formula (5) is of a classical character: it does not contain Planck's constant, and it can be obtained on the basis of classical calculations for a uniformly moving charge.

In what follows we shall calculate the bremsstrahlung cross section in the Born–Bethe approximation, in which one can set

$$\alpha(\omega, q) = \alpha(\omega)\theta(p_a - q), \quad (7)$$

where $\theta(x)$ is the unit step function.

We shall use the Thomas–Fermi momentum $p_a = Z^{1/3}/(ba_0)$ as the characteristic atomic momentum.

The integral in equation (5) can be calculated analytically in the Born–Bethe approximation (7). The result, however, remains unwieldy. Consequently, we present here a general formula which explicitly singles out two characteristic frequency ranges ($\omega < p_a v$ and $\omega > p_a v$) for the spectral polarization bremsstrahlung cross section in terms of a single integral over the magnitude of the transferred momentum:

$$\frac{d\sigma^{PB}}{d\omega} = \frac{4\omega^3}{v^2}|\alpha(\omega)|^2 \times \left\{ \theta\left(\frac{p_a v}{1+v} - \omega\right)[H_1(\omega, p_a - \omega) + H_2(\omega)] + \theta\left(\omega - \frac{p_a v}{1+v}\right) H_1\left(\omega, \frac{\omega}{v}\right) \right\}, \quad (8)$$

$$H_1(\omega, q_{\min}) = \int_{q_{\min}}^{p_a} G_1(q, \omega) \frac{q dq}{(q^2 - \omega^2)^2},$$

$$H_2(\omega) = \int_{\omega/v}^{p_a - \omega} G_2(q, \omega) \frac{q dq}{(q^2 - \omega^2)^2}.$$

Here

$$G_1 = \frac{p_a^2 - (q-\omega)^2}{2\omega q}\left[\omega^2 v^2 + q^2 - \frac{5}{2}\omega^2 + \frac{\omega^4}{2q^2 v^2}\right] - \frac{1}{3}\left[\left(\frac{p_a^2 - (q-\omega)^2}{2\omega q}\right)^3 + 1\right]\left[q^2 - \frac{5}{2}\omega^2 + \frac{3\omega^4}{2q^2 v^2}\right], \quad (8a)$$

$$G_2 = 2\omega^2\left(v^2 - \frac{5}{3}\right) + \frac{4}{3}q^2.$$

The formula (8) in the frequency range $\omega < p_a v$, where the first term in braces contributes to the cross section, reduces in the "large logarithm" approximation to the well-known expression for the spectral polarization bremsstrahlung cross section of a relativistic incident electron [4, 5] (see also [5, Chap. 5]):

$$\frac{d\sigma^{PB}}{d\omega} = \frac{16\omega^3|\alpha(\omega)|^2}{3v^2} \ln\frac{2\gamma p_a v}{\omega(1+v)}, \quad \omega < p_a v, \quad (9)$$

where $\gamma = (1 - v^2)^{-1/2}$ is the relativistic factor.

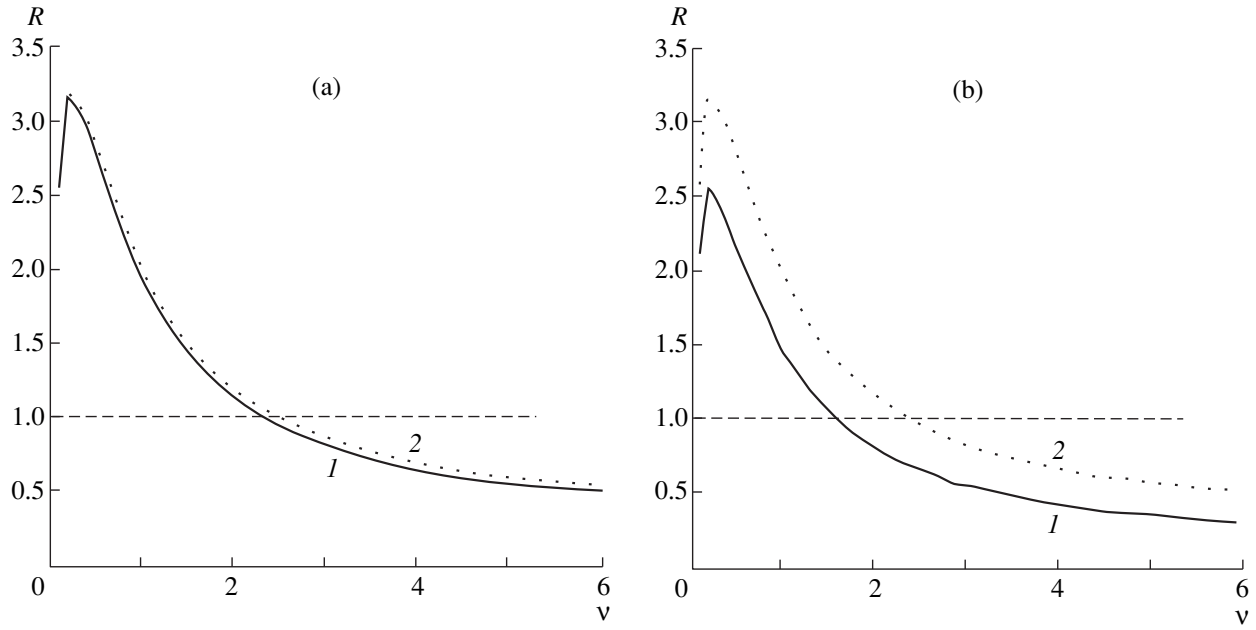


Fig. 2. Curves of the R factor versus the reduced frequency ν, calculated for bremsstrahlung of a fast electron ($\gamma = (1 - v^2)^{-1/2}$—relativistic factor) on a Thomas–Fermi atom with nuclear charge Z: (a) $\gamma = 10$; $Z =$ (*1*) 60, $Z =$ (*2*) 30; (b) $Z = 60$; $\gamma =$ (*1*) 3, $\gamma =$ (*2*) 10.

Switching in equation (9) to dimensionless variables using equation (1a) and the definition of the Thomas–Fermi radius, we obtain the following expression for the spectral polarization bremsstrahlung cross section:

$$d\sigma^{PB}(\nu) = \frac{16Z^2b^6}{3v^2}\left|\nu^2\beta(\nu)\right|^2\frac{d\nu}{\nu} \times \ln\frac{2\gamma v}{\nu_0(1+v)Z^{2/3}} = Z^2 d\tilde{\sigma}^{PB}(\nu). \quad (10)$$

In equation (10) we have introduced the function $d\tilde{\sigma}^{PB}(\nu)$, which can be naturally called the reduced cross section for the process, since in the case of the polarization bremsstrahlung of a fast incident particle, considered here, an approximate scaling in the parameter ω/Z holds for it, while the dependence on the nuclear charge is only logarithmic.

The spectral cross section for ordinary (static) bremsstrahlung taking account of the screening of the nuclear field [3] in the case of weakly inelastic scattering of an electron is given by the expression

$$d\sigma^{OB}(\omega) = \frac{16Z^2}{3v^2}\frac{d\omega}{\omega}\ln\frac{v}{p_a}, \quad \omega < p_a v. \quad (11)$$

The ratio of the cross sections determined by equations (10) and (11) makes it possible to find the R factor in the frequency range under study ($\omega < p_a v$) and in the relativistic limit ($v \simeq 1$):

$$R(\nu, Z, \gamma) \equiv \frac{d\sigma^{PB}}{d\sigma^{OB}} = b^6\left|\nu^2\beta(\nu)\right|^2\frac{\ln\dfrac{137\gamma}{\nu Z^{2/3}}}{\ln\dfrac{137}{Z^{1/3}}}, \quad \nu < \frac{137}{Z^{2/3}}. \quad (12)$$

The computational results for the R factor as a function of the dimensionless frequency ν are presented in Fig. 2 for various values of Z and γ in the range $\nu < 137/Z^{2/3}$.

We note that photon energies $\hbar\omega < 14$ keV correspond to the given values of the dimensionless frequency ν (for nuclear charge $Z \approx 50$).

It is evident that the value of the R factor of a relativistic incident electron is of the order of 1 in a wide frequency range, and it reaches its maximum value of the order of 2.5–3 for $\omega \approx Z\mathrm{Ry}$ ($\nu = 0.5$). The "sublogarithmic" effect of the nuclear charge on the scaling with respect to Z is vanishingly small, and the effect of the relativistic factor is stronger.

We note that the interference of the polarization and static bremsstrahlung channels in the case of a relativistic charged particle is small because of the differences in the directional patterns of the radiation: the standard channel gives sharply directed radiation in a cone with an angle of the order of $1/\gamma$ [3], while the angular distri-

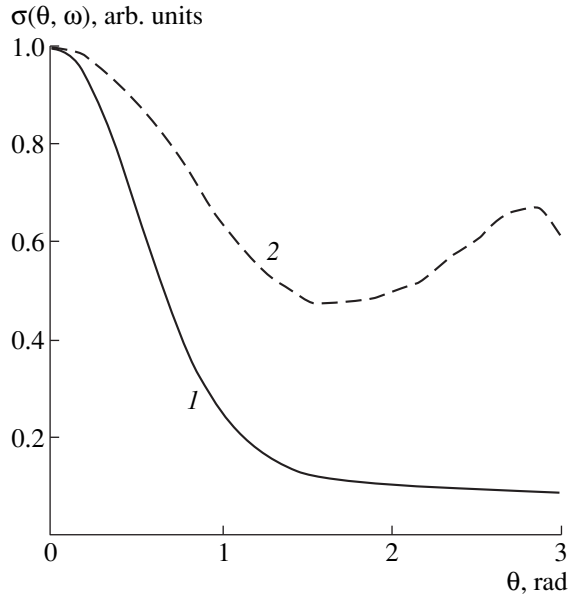


Fig. 3. Normalized angular dependences of the cross section of polarization bremsstrahlung of a relativistic electron ($\gamma = 2.3$) on a Thomas–Fermi atom ($Z = 30$) for various photon energies: $\hbar\omega$ = (*1*) 50; (*2*) 5 keV.

bution of the polarization bremsstrahlung for frequencies $\omega < p_a v$ is of a dipole character [4, 5].

In the present case of weakly inelastic scattering of a charged Born particle in the frequency range $\omega < p_a v$, small scattering angles, where the influence of the penetration of the incident particle into the electron core of the atom is small, make the main contribution to bremsstrahlung. Here lies the substantial difference from the situation studied in [2], where bremsstrahlung of electrons with threshold energies is accompanied by penetration of the plasma electron into the ion core, the penetration being all the stronger the higher the frequency of the emitted bremsstrahlung photon.

As a result, the frequency dependences of the polarization bremsstrahlung cross section are different for different degrees of inelasticity of the scattering of the incident electron. For the process studied in the present work, the frequency maximum of the bremsstrahlung cross section is shifted substantially into the region of high frequencies and decreases more slowly with increasing bremsstrahlung frequency than the corresponding spectral dependences for emission of threshold-energy photons [2].

In the frequency range $\omega > p_a v$ the law of conservation of energy–momentum makes it necessary for the incident charged particle to penetrate into the electronic core of the target. For this reason, remission of a virtual photon of the intrinsic field of the electron itself into a photon on atomic electrons loses it coherent character. As a result, the spectral polarization bremsstrahlung cross section is suppressed compared with the cross section for standard bremsstrahlung.

As first noted in [5], in the relativistic case for frequencies $p_a < \omega < \gamma^2 p_a$ there exists a range of emission angles $\vartheta \le \sqrt{p_a/\omega}$ where compensation of the momentum transfer from the incident particle to the target is possible as a result of the momentum of the bremsstrahlung photon, so that bremsstrahlung remains coherent in the polarization channel.

This range of angles is determined by the inequality $|\mathbf{q}_{\min} + \mathbf{k}| \le p_a$, which can be rewritten in the form

$$\sin\frac{\vartheta}{2} \le \frac{1}{2}\sqrt{\frac{p_a^2 v}{\omega^2} - (1 - v)^2}$$

(since $\mathbf{q}_{\min} = -\omega\mathbf{v}/v^2$ by virtue of the law of conservation of energy–momentum). Hence follows in the relativistic limit for $p_a < \omega < \gamma^2 p_a$ the fact that the bremsstrahlung in the polarization channel is directed, and the characteristic maximum angle of emission is determined by the quantity $\sqrt{p_a/\omega}$ and not by the reciprocal of the relativistic factor, as in the case of ordinary bremsstrahlung.

The indicated narrowing of the directional pattern of the bremsstrahlung in the polarization channel as the energy of the bremsstrahlung photon increases from 5 to 50 keV for an incident electron with a relativistic factor $\gamma = 2.3$ and a Thomas–Fermi atom with $Z = 30$ is shown in Fig. 3. In the high-frequency range $\omega > p_a v$ the dimensionless frequency ν satisfies the inequality $\nu > 10$ (we assume that $Z \ge 30$), and as one can see from Fig. 1 the reduced polarizability of the Thomas–Fermi atom is close to its high-frequency limit: $\beta(\nu) \approx \beta_\infty(\nu) = -\nu^{-2}$. The frequency dependence of the polarization bremsstrahlung cross section in this case is determined primarily by the integral over the angular variables and over the magnitude q of the momentum transfer in equation (5).

The formula (9) is no longer valid in the frequency range $\omega > p_a v$, and the general expression (8) must be used to determine the polarization bremsstrahlung cross section. In this case the second term in braces in equation (8) makes a contribution. Analysis shows that the factor $(p_a/\omega)^2$, determining the smallness of the contribution of the polarization channel to the total spectral bremsstrahlung cross section, appears in the expression for the spectral cross section. However, the spectral-angular cross section for polarization bremsstrahlung in the range of photon emission angles $\gamma^{-1} < \vartheta \le \sqrt{p_a/\omega}$ is larger than the corresponding cross section in the static channel.

The model developed also makes it possible to calculate the cross section for induced bremsstrahlung and absorption of photons. Such calculations, including the polarization channel and its interference with the ordinary (static) channel, have been performed and analyzed recently [11] for thermal-energy electrons scattered by ions in a near-resonance laser field. It follows

from these calculations, specifically, that the magnitude of the corresponding cross sections in the ultraviolet spectral range studied is very small even with a resonant external field. The effect will be even smaller in the nonresonant case considered in this paper. Indeed, as is well known [12], the cross section for induced bremsstrahlung and absorption in the standard (static) channel decreases as the inverse of the fourth power of the frequency of the external radiation. Since the ratio of the cross section in the polarization channel to the static channel does not exceed a value of the order of 1 in the entire frequency range, it is obvious that the bremsstrahlung cross section in the high-frequency range considered here ($\hbar\omega > 100$ eV) will be extremely small. In addition, a direct photoelectric absorption channel is open in this case. This channel will prevail over absorption in the bremsstrahlung channel.

4. CONCLUSIONS

A universal description of polarization bremsstrahlung of a fast charged Born particle on a multielectron atom ($Z \gg 1$) in the bremsstrahlung photon energy range $\hbar\omega > 100$ eV was given on the basis of the local electron density method and the Thomas–Fermi model.

It was shown that the R factor, which determines the relative contribution of the polarization channel to the total bremsstrahlung cross section, possesses approximate scaling in the parameter ω/Z and reaches its maximum value $R_{max}(\gamma) = 2.5$–3, which increases logarithmically with increasing energy of the incident particle, at frequencies $\omega_{max} \approx Z\mathrm{Ry}$.

The decrease in the R factor with increasing energy of the emitted photon in the low-frequency range $\omega < p_a v$ is most strikingly expressed right up to frequencies of the order of $20Z\mathrm{Ry}$, where the modulus of the polarizability of the Thomas–Fermi model decreases, reaching its high-frequency asymptotic value.

In the spectral interval $10Z\,\mathrm{Ry} < \omega < p_a v$ the decrease in the R factor and in the intensity of the polarization bremsstrahlung is weakly logarithmic and is due to the decrease in the maximum impact parameter.

In the high-frequency range $\omega > p_a v$ the frequency variation of the intensity of the polarization bremsstrahlung is determined primarily by kinematic factors and by the breakdown of the coherence of the reradiation of a virtual photon into a real photon on the atomic electrons. The decay of the spectral intensity follows a power law. At the same time, the directional pattern of the radiation in the polarization channel becomes narrower, so that $\vartheta \leq \sqrt{p_a/\omega}$, and in the frequency range $p_a < \omega < \gamma^2 p_a$ there exists a range of bremsstrahlung angles, $\gamma^{-1} < \vartheta \leq \sqrt{p_a/\omega}$, where the polarization mechanism prevails over the standard (static) emission mechanism.

ACKNOWLEDGMENTS

This work was supported by the Russian Foundation for Basic Research (project no. 98-02-16763) and by the International Scientific-Technical Center (grant no. 856-98).

REFERENCES

1. W. Brandt and S. Lundquist, Phys. Rev. A **139**, 612 (1965).
2. V. A. Astapenko, L. A. Bureeva, and V. S. Lisitsa, Zh. Éksp. Teor. Fiz. **117**, 496 (2000) [JETP **90**, 434 (2000)].
3. V. B. Berestetskii, E. M. Lifshitz, and L. P. Pitaevskii, *Quantum Electrodynamics* (Nauka, Moscow, 1980; Pergamon Press, Oxford, 1982).
4. M. Ya. Amus'ya, M. Yu. Kuchiev, A. V. Korol', and A. V. Solov'ev, Zh. Éksp. Teor. Fiz. **88**, 383 (1985) [Sov. Phys. JETP **61**, 224 (1985)].
5. V. A. Astapenko, V. M. Buĭmistrov, Yu. A. Krotov, *et al.*, Zh. Éksp. Teor. Fiz. **88**, 1560 (1985) [Sov. Phys. JETP **61**, 930 (1985)].
6. *Polarization Bremsstrahlung*, Ed. by V. N. Tsytovich and I. M. Oiringel' (Nauka, Moscow, 1987; Plenum, New York, 1992).
7. A. V. Korol', A. G. Lyalin, O. I. Obolenskiĭ, and A. V. Solov'ev, Zh. Éksp. Teor. Fiz. **114**, 458 (1998) [JETP **87**, 251 (1998)].
8. L. D. Landau and E. M. Lifshitz, *Quantum Mechanics: Non-Relativistic Theory* (Nauka, Moscow, 1989, 4th ed.; Pergamon Press, Oxford, 1977, 3rd ed.).
9. A. D. Ulantsev and V. P. Shevel'ko, Opt. Spektrosk. **65**, 1003 (1988) [Opt. Spectrosc. **65**, 590 (1988)].
10. A. V. Vinogradov and V. P. Shevel'ko, Tr. Fiz. Inst. Akad. Nauk SSSR **119**, 158 (1980).
11. V. A. Astapenko, Laser Phys. **9**, 1032 (1999).
12. F. V. Bunkin, A. E. Kazakov, and M. V. Fedorov, Usp. Fiz. Nauk **107**, 559 (1972) [Sov. Phys. Usp. **15**, 416 (1972)].

Translation was provided by AIP

Journal of Experimental and Theoretical Physics, Vol. 90, No. 5, 2000, pp. 794–804.
Translated from Zhurnal Éksperimental'noĭ i Teoreticheskoĭ Fiziki, Vol. 117, No. 5, 2000, pp. 913–925.

ATOMS, SPECTRA, RADIATION

Phase Control of the Degree of Ionization of Rydberg Atoms by a Strong Laser Field

N. P. Poluéktov and M. V. Fedorov*
General Physics Institute, Russian Academy of Sciences, Moscow, 117942 Russia
**e-mail: fedorov@gon.ran.gpi.ru*
Received January 12, 2000

Abstract—It is shown that the rate and degree of ionization of an atom can be controlled by preparing a specific superposition of Rydberg states. This effect is due to the Raman interaction of Rydberg levels via resonant lower-lying levels of the atom. Depending on the initial population of an atom interacting with a strong laser field, the atom either ionizes completely or exhibits an extremely high degree of stability (in this case the probability of the atom remaining in a bound state when the laser pulse is switched off is close to 1). It is shown that states which are stable against ionization decay or, conversely, decay rapidly into the continuum can be prepared using double laser pulses with a controllable time delay. *© 2000 MAIK "Nauka/Interperiodica".*

1. INTRODUCTION

It is well known that stabilization of an atom is possible in a strong laser field [1–10]. Here stabilization means saturation (at a level less than 1) or even a decrease in the ionization probability of an atom with increasing field intensity. Two types of stabilization are distinguished: stabilization of unexcited atoms by the Kramers–Henneberger mechanism, which appears in relatively intense fields ($>10^{14}$ W/cm^2) [7], and interference stabilization of initially excited (Rydberg) atoms in weaker fields. The physical mechanism leading to interference stabilization is coherent repopulation of Rydberg levels as a result of Raman-type transitions in the course of ionization. Such a repopulation can occur as a result of virtual transitions via the continuum (Λ-type transitions) and via lower-lying resonance levels (V-type transitions) (see Fig. 1). Correspondingly, depending on the predominant transition channel, Λ- and V-type interference stabilization are distinguished. The characteristic radiation intensities at which interference stabilization appears vary over wide limits, which depend on the configuration of the quantum system (10^{10}–10^{13} W/cm^2 for Λ stabilization and 10^7–10^{10} W/cm^2 for V stabilization).

If the temporal profile of the laser pulse is square, then it is best to use the formalism of quasienergies and quasistationary states to describe an atom in the field of the laser wave. The corresponding analysis shows that the stability of various quasistationary states (for a quantum system using V-type transitions) against ionization decay is different and different quasienergy levels possess different decay times. The weight of the quasistationary states with long and short decay times and different phase relations between the initial amplitudes of population of the coherent Rydberg state could be different. In this connection it is reasonable to infer that depending on the initial population of the atomic levels, an atom can decay into a continuum within longer or shorter times (or, if the interaction time with the laser field is fixed, depending on the initial population, a different fraction of the atomic population enters the continuum). The objective of the present work is to check this conjecture and to formulate the conditions under which it is possible to observe the dependence of the emergence of an electron on the initial atomic population distribution. The influence of the phase properties of the initial state of an atom on the ionization of V and Λ systems was investigated. It was shown that even though the decay time of various quasienergy levels in a Λ system is the same, for short times (shorter than the Kepler period of revolution of an electron on a classical

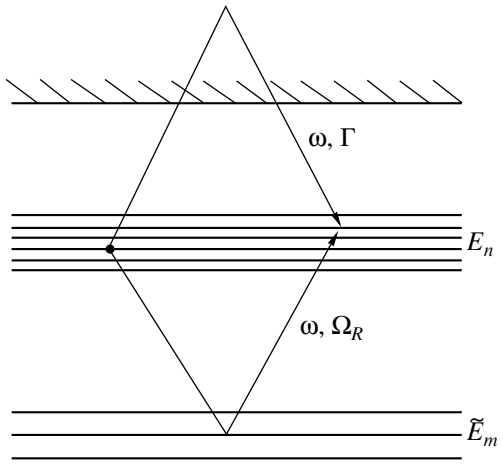


Fig. 1. Scheme of Raman-type transitions via the continuum (Λ type) and lower-lying resonance states (V type), resulting in coherent repopulation of the Rydberg atom giving rise to interference stabilization.

orbit) the initial population of an atom strongly influences the ionization probability of the atom.

We note that this is the first time that the problem considered here has ever been formulated. In other works [1–6, 8–10] concerning interference stabilization only one atomic level was initially populated. An attempt to investigate the dependence of the ionization on the initial atomic population distribution (on the basis of a Λ system) was made in [11], but a comprehensive phase analysis was not made there. An important and absolutely new result obtained in the present work is a prescription for preparing a special coherent state of an atom and for controlling the ionization of this state. We propose that the atom be exposed to two successive laser pulses. Depending on the time interval between the pulses, the electron can either completely leave the atom and enter the continuum or it can be trapped, to a high degree, in the atomic levels.

In Section 2 all calculations are performed on the basis of a model that takes into account only Λ-type transitions. In Section 3 similar calculations are performed for a model that takes into account *V*-type transitions. Section 4 is devoted to a discussion of the possibility of observing experimentally the dependence of the emergence of an electron on the initial state of the electron. The basic results of this work are summarized in the Conclusions section. The atomic system of units ($e = m_e = \hbar = 1$) is used throughout.

2. COHERENT POPULATION AND IONIZATION OF A Λ SYSTEM

Let the atom be initially excited in such a way that the Rydberg levels near a certain level E_{n_0} are coherently populated. Let the frequency of the field, which ensures single-photon emergence into the continuum ($\omega > |E_{n_0}|$), be such that there are no resonances with lower-lying levels. This is a typical scheme for Λ-type interference stabilization.

The evolution of an atom in the field of an electromagnetic wave is determined by the nonstationary Schrödinger equation

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

Here $\hat{H}_0$ is the atomic Hamiltonian and $\hat{V}(t) = -\mathbf{F}\cdot\mathbf{r}\cos(\omega t)$ is the operator describing the interaction of the atom and the field (F is the amplitude of the laser field).

To solve equation (1) we expand the wave function of the atom in a field in terms of the eigenfunctions of a free atom. We shall assume the polarization of the field to be linear, and we shall also assume that the angular momentum quantum number for all Rydberg levels which effectively interact because of the laser field is the same (for simplicity, we set $l = 0$) and that transitions to levels with other orbital angular momenta can be neglected. We shall also neglect free-free atomic transitions and we shall take into account only the interaction of levels with the states of the continuum with $l = 1$. The applicability, quality, and possibility of going beyond the framework of these approximations has been discussed in detail in [11].

We shall write the wave function $\Psi(t)$ of the atom in the form of an expansion:

$$\Psi(t) = \sum_n a_n(t)\varphi_n(\mathbf{r}) + \int dE\, a_E(t)\varphi_E(\mathbf{r}), \qquad (2)$$

where φ_n and φ_E are the wave functions of the discrete spectrum and the continuum, respectively (n is the principal quantum number and $E > 0$ is the energy in the continuum).

The probability of ionization of an atom during the interaction with the pulse is

$$w_{\text{ion}} = 1 - \sum_n |a_n|^2, \qquad (3)$$

where the functions a_n are taken at the moment the laser pulse is switched off.

Using the expansion (2), the Schrödinger equation (1) can be reduced to a system of equations for the probability amplitudes $a_n(t)$ and $a_E(t)$. Using the procedure of adiabatic elimination of the continuum [11, 12], the functions $a_E(t)$ can be expressed in terms of $a_n(t)$ and the equations for the latter can be written as

$$i\dot{a}_n = E_n a_n - i\sum_{n'}\frac{\Gamma_{n,n'}}{2}a_{n'}. \qquad (4)$$

Here

$$\Gamma_{n,n'} = 2\pi\frac{V_{n,E}}{2}\frac{V_{E,n'}}{2}$$

is the ionization-widths tensor, $V_{\alpha\beta} = -\mathbf{F}\cdot\mathbf{r}_{\alpha,\beta}$, and $\mathbf{r}_{\alpha,\beta} = \langle\alpha|\mathbf{r}|\beta\rangle$ are the dipole moment matrix elements.

If the amplitude F of the field does not depend on the time, then the system of equations (4) possesses stationary solutions $a_n(t) = b_n\exp(-i\gamma t)$, which determine the complex quasienergies γ_j of an atom in a field and the corresponding quasistationary wave functions ψ_j. The equations for the quasienergies and the constants b_n follow directly from equation (4):

$$(\gamma - E_n)b_n = -i\sum_{n'}\frac{\Gamma_{n,n'}}{2}b_{n'}. \qquad (5)$$

In terms of the quasienergies, the solution of the initial problem is determined by a superposition of the quasienergy states. Specifically, the probability ampli-

tudes $a_n(t)$ in such an approach can be represented in the form

$$a_n = \sum_j C_{n,j}\exp(-i\gamma_j t). \tag{6}$$

The expansion coefficients $C_{n,j}$ satisfy, in the first place, the same equations as the constants b_n (5) and, in the second place, the initial conditions

$$\sum_j C_{n,j} = a_n(0) \tag{7}$$

($a_n(0)$ are the probability amplitudes of initial population of the levels).

We shall assume everywhere below that the matrix elements $V_{n,E}$ and the components of the ionization-widths tensor $\Gamma_{n,n'}$ do not depend on the index n: $V_{n,E} \approx V_{n_0,E}$ and $\Gamma_{n,n'} \approx \Gamma_{n_0,n_0} \equiv \Gamma$, which is valid for $n \gg 1$ [13].

The simplest model in which interference stabilization can occur is a model in which there are two discrete levels plus the continuum. The complex quasienergies of such a system have the form

$$\gamma_\pm = E_1 + \frac{\Delta}{2} - i\frac{\Gamma}{2} \pm i\beta, \tag{8}$$

where E_1 and E_2 are the energies of the discrete levels, $\Delta = E_2 - E_1$ is the distance between the levels, and

$$\beta = \sqrt{\left(\frac{\Gamma}{2}\right)^2 - \left(\frac{\Delta}{2}\right)^2}.$$

For a two-level system the strong-field criterion is $\Gamma \gg \Delta$. In this limiting case the widths of the two quasienergy levels are substantially different: $\mathrm{Im}[\gamma_-] = -\Gamma$ and $\mathrm{Im}[\gamma_+] = -\Delta^2/4\Gamma$. The smallness of the level width γ_+ for $\Gamma \gg \Delta$ corresponds to interference stabilization.

The solution of the initial problem for a two-level model gives

$$\begin{aligned}
a_1(t) &= \left[\frac{i(\Gamma/2)^2}{2\beta(\Delta/2 + i\beta)}a_1(0) - \frac{\Gamma/2}{2\beta}a_2(0)\right]\exp(-i\gamma_+ t) \\
&+ \left[\frac{-i(\Gamma/2)^2}{2\beta(\Delta/2 - i\beta)}a_1(0) + \frac{\Gamma/2}{2\beta}a_2(0)\right]\exp(-i\gamma_- t), \\
a_2(t) &= \left[\frac{-\Gamma/2}{2\beta}a_1(0) - i\frac{\Delta/2 + i\beta}{2\beta}a_2(0)\right]\exp(-i\gamma_+ t) \\
&+ \left[\frac{\Gamma/2}{2\beta}a_1(0) + i\frac{\Delta/2 - i\beta}{2\beta}a_2(0)\right]\exp(-i\gamma_- t).
\end{aligned} \tag{9}$$

The probability amplitudes—$a_1(0)$ and $a_2(0)$—of the initial coherent population satisfy the normalization condition $|a_1(0)|^2 + |a_2(0)|^2 = 1$. In the strong-field limit ($\Gamma \gg \Delta$) and at sufficiently long times ($\Gamma t > 1$) the formulas (9) take the extremely simple form

$$\begin{aligned}
a_1(t) &= \frac{1}{2}[a_1(0) - a_2(0)]\exp(-i\gamma_+ t), \\
a_2(t) &= -\frac{1}{2}[a_1(0) - a_2(0)]\exp(-i\gamma_+ t).
\end{aligned} \tag{10}$$

In equation (10) we neglected the exponentially small terms proportional to the factor $\exp(-i\gamma_- t) \sim \exp(-\Gamma t)$.

The formulas (10) are interesting in that the probability amplitudes for an atom to occupy discrete levels (after the laser field is switched off) are proportional to the difference of the initial amplitudes ($a_1^{(0)} - a_2^{(0)}$). Thus, depending on the value of this difference, the population of the discrete spectrum could be substantially different. Correspondingly, the ionization probability, determined by the expression

$$w_{\mathrm{ion}} = 1 - \frac{1}{2}\left|a_1^{(0)} - a_2^{(0)}\right|^2 \exp[-(\Gamma - 2\beta)t] \tag{11}$$

depends strongly on the initial population of the atom.

As the field increases, the argument of the exponential in equation (11) approaches zero (see equation (8)) and, hence, the exponential itself can be made as close to 1 as desired. If it is assumed that $a_1^{(0)}$ and $a_2^{(0)}$ are equal to one another in absolute magnitude ($\left|a_1^{(0)}\right| = \left|a_2^{(0)}\right| = 1/\sqrt{2}$) and only their phases are different, then the ionization probability of an atom is either 1 (if $a_1^{(0)}$ and $a_2^{(0)}$ are in-phase) or it is close to 0 (in the opposite case, when the difference of the phases of $a_1^{(0)}$ and $a_2^{(0)}$ is π)! Thus, controlling the phase difference between the amplitudes of the initially populated atomic levels makes it possible either to attain complete ionization of the atom or to forbid an electron to leave the atom into the continuum.

Everything said above is valid only for a two-level system, and it is not claimed that the description of a Rydberg atom is adequate. The model of an atom with an infinite number of equidistant energy levels with the same oscillator strengths is found to be much closer to reality. In such a model $E_n \approx E_{n_0} + (n - n_0)\Delta$, where $\Delta = n_0^{-3}$ (in the atomic system of units), and the system (5) can be rewritten as

$$i\dot{a}_n = (n - n_0)\Delta a_n - i\frac{\Gamma}{2}\sum_{m=-\infty}^{\infty} a_m. \tag{12}$$

(In equation (12), energy is measured from the level E_{n_0}.)

The solution of the initial-value problem for the system (12) gives

$$a_n(t) = \left[\sum_k (\gamma - \Delta k)^{-2}\right]^{-1} \times \sum_\gamma \frac{\exp(-i\gamma t)}{\gamma - n\Delta} \sum_m \frac{a_m(0)}{\gamma - m\Delta}, \quad (13)$$

where γ are the complex quasienergies levels determined as roots of the equation

$$1 + i\frac{\Gamma}{2}\sum_m \frac{1}{\gamma - m\Delta} = 0. \quad (14)$$

In the limit $\Gamma \gg \Delta$ the solutions of equation (14) have the asymptotic form [1]

$$\gamma = \left(s + \frac{1}{2}\right)\Delta - i\frac{2\Delta}{\pi^2\Gamma}, \quad (15)$$

where $s = 0, \pm 1, \pm 2, \ldots$.

Substituting the expression (13) into equation (3) gives the following expression for the ionization probability of an atom when the laser pulse is switched off:

$$W_{\text{ion}} = 1 - \exp(-\Gamma_{\text{str}} t)\left|\sum_k (\gamma - k\Delta)^{-2}\right|^{-2} \times \sum_n \sum_{s_1, s_2} \exp[-i(s_1 - s_2)\Delta t]\{[(s_1 - n + 1/2)\Delta - i\Gamma_{\text{str}}/2] \times [(s_2 - n + 1/2)\Delta + i\Gamma_{\text{str}}/2]\}^{-1} \sum_{m_1, m_2} a_{m_1}(0) a^*_{m_2}(0) \times \{[(s_1 - m_1 + 1/2)\Delta - i\Gamma_{\text{str}}/2] \times [(s_2 - m_2 + 1/2)\Delta + i\Gamma_{\text{str}}/2]\}^{-1}, \quad (16)$$

where $\Gamma_{\text{str}} = 4\Delta/\pi^2\Gamma$ is the ionization width of the quasienergy levels of the atom in the field of an intense electromagnetic wave [1]. In the strong-field limit the width Γ_{str} is small ($\Gamma_{\text{str}} \ll \Delta$), which once again corresponds to interference stabilization. Therefore, in a strong field the expression on the right-hand side of equation (16) can be expanded in powers of $\Gamma_{\text{str}}/\Delta$. In the first approximation this expansion has the form

$$W_{\text{ion}} = 1 - \exp(-\Gamma_{\text{str}} t) \times \left\{1 + i\frac{\Gamma_{\text{str}}}{\Delta} \sum_{m_1 \neq m_2} a_{m_1}(0) a^*_{m_2}(0) \frac{\exp[-i(m_1 - m_2)t\Delta]}{m_1 - m_2}\right\}. \quad (17)$$

We call attention to the fact that the formula (17) is identical to the analogous formula from [11, p. 190], if the latter is written to the same accuracy as equation (17).

It is easy to see that if only one atomic level is populated at the moment the laser field is switched on, the second term in braces in equation (17) vanishes. In such a situation the ionization probability as a function of the duration of the laser pulse follows a simple exponential law (with decay constant Γ_{str}). However, as follows from equation (17), a large deviation from this law will occur if several Rydberg levels are initially populated.

To simplify the analysis and make it clearer we switch from the expression for the ionization probability to a formula for the ionization rate. For this we differentiate the expression (17) with respect to time

$$\dot{W}_{\text{ion}} = \Gamma_{\text{str}} \exp(-\Gamma_{\text{str}} t)\left|\sum_m a_m(0) \exp(-imt\Delta)\right|^2. \quad (18)$$

It follows from equation (18) that the course of the ionization process depends strongly on the initial atomic population. Indeed, if initially N Rydberg levels are populated, then the ionization rate (at the time $t = 0$) will vary from 0 (when $\sum_m a_m(0) = 0$) to $N\Gamma_{\text{str}}$ (when the phases $a_m(0)$ do not depend on m). Correspondingly, at short times (much shorter than the Kepler period $T_K = 2\pi\Delta^{-1}$) the initial population of the Rydberg atom strongly influences the degree of ionization of the atom (which in the present case can be calculated as the product of the ionization rate by the duration of the laser pulse).

This influence vanishes at longer interaction times. Integrating the expression (18) over time in one Kepler period we obtain

$$W_{\text{ion}}\Big|_t^{t+T_K} \approx \Gamma_{\text{str}} T_K \exp(-\Gamma_{\text{str}} t). \quad (19)$$

As follows from equation (19), on the average over one Kepler period a fraction of the atomic population that ionizes does not depend on the initial population of the atom. This is also confirmed by numerical calculations (see Fig. 2): the ionization curve has a "step" form with "steps" of identical height. Thus, in a quantum system where population redistribution on the quantum levels is due to Raman transitions via the continuum (Λ type), at times comparable to a Kepler period, it is impossible to control the emergence of an electron into the continuum by varying the initial population of the atom. This is probably due to the fact that in a Λ system the ionization widths of all quasienergy levels of an atom in a field are identical, i.e., a single parameter determines the time scale on which ionization occurs, and an electron emerges into the continuum, on the average, at the same rate.

We note that if the nonequidistant nature of the atomic spectrum, as is characteristic for a Rydberg atom, is taken into account, then the "step" character of the curve $W_{\text{ion}}(t)$ will be destroyed for large values of t (see Fig. 3). For times at which the nonequidistant nature of the atomic spectrum becomes important, the

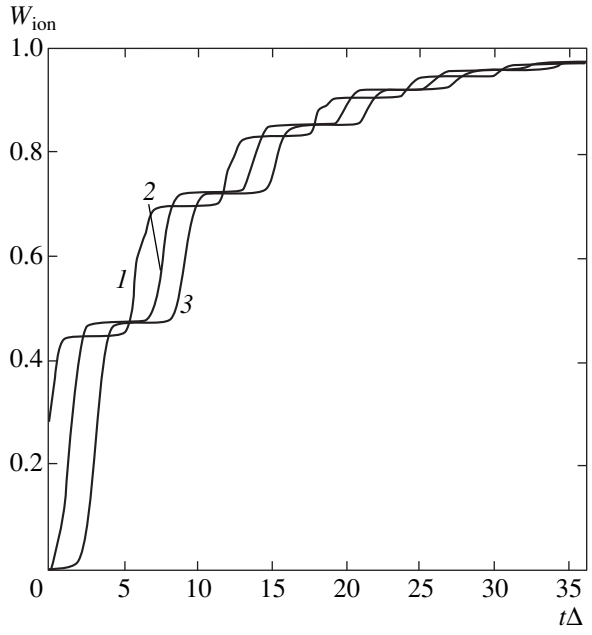


Fig. 2. Ionization probability of a Λ system (equidistant spectrum), whose initial population is a Gaussian distribution ($a_n(0) = A\exp\{-(n-n_0)^2/\sigma^2 + in\theta\}$), as a function of the duration of the laser pulse. θ = (*1*) 0, (*2*) π/2, (*3*) π. Values of the remaining parameters: Γ/Δ = 4, σ = 2.0. The computational scheme contains 17 levels.

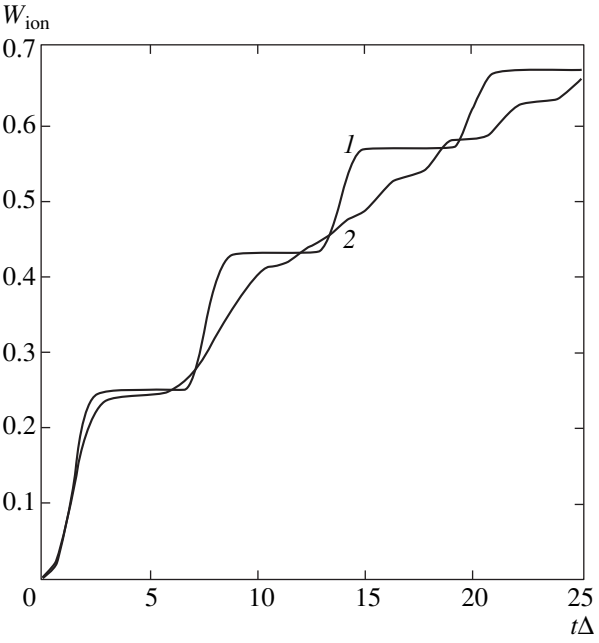


Fig. 3. Ionization probabilities for a Λ system, whose initial population is a Gaussian distribution ($a_n(0) = A\exp\{-(n-n_0)^2/\sigma^2 + in\theta\}$), as a function of the duration of the laser pulse. The values of the parameters chosen for the calculations are: Γ/Δ = 9, σ = 2.0, θ = π/2. The curves *1* and *2* correspond to the equidistant spectrum and the real spectrum of a Rydberg atom, respectively; $n_0 = 60$. The computational scheme contains 17 levels.

ionization probability curve passes into a curve corresponding to exponential decay with rate Γ_{str}.

3. COHERENT POPULATION AND THE IONIZATION OF A *V* SYSTEM

For a two-level system the widths of the quasienergy levels were different and, as a result, the ionization process depended strongly on the initial atomic population distribution. Is it possible for a situation where different quasienergy levels possess different widths to arise in a system that adequately describes a real atom (in the simplest case, a system consisting of an infinite number of equidistant levels coupled with the field by different matrix elements)?

We shall show that such a situation occurs when the Rydberg levels interact owing to Raman transitions via the lower-lying resonance levels. Just as in the preceding section, we assume that the orbital angular momentum quantum number of the Rydberg levels is fixed. For definiteness we shall assume that initially the Rydberg levels of the *s* series ($l = m = 0$), coupled by the field with the *p* continuum and a collection of low-lying resonant *p* levels, are populated. For such a system the wave function of the atomic electron can be written in the form

$$\Psi(t) = \sum_m \tilde{a}_m(t)\tilde{\varphi}_m(\mathbf{r}) + \sum_n a_n(t)\varphi_n(\mathbf{r}) + \int dE\, a_E(t)\varphi_E(\mathbf{r}). \tag{20}$$

The formula (20) differs from equation (2) by the presence of the additional term

$$\sum_m \tilde{a}_m(t)\tilde{\varphi}_m(r)$$

($\tilde{\varphi}_m$ and $\tilde{a}_m$ are, respectively, the wave functions of the lower-lying levels and the expansion coefficients corresponding to these levels). The formula (3) for the ionization probability of the atom will also change similarly. It assumes the form

$$W_{\text{ion}} = 1 - \sum_n |a_n|^2 - \sum_m |\tilde{a}_m|^2. \tag{21}$$

Here, just as in equation (3), the values of the functions a_n and $\tilde{a}_m$ are taken at the moment the laser pulse is switched off.

The equations determining a_n and $\tilde{a}_m$ in the adiabatic elimination of the continuum approximation have the form

$$\begin{aligned} i\dot{\tilde{a}}_m &= \tilde{E}_m\tilde{a}_m + \sum_{n'} \Omega_{m,n'}e^{i\omega t}a_{n'}, \\ i\dot{a}_n &= \sum_{m'} \Omega_{n,m_1}e^{-i\omega t}\tilde{a}_{m'} + E_n a_n - i\sum_{n'} \frac{\Gamma_{n,n'}}{2}a_{n'}, \end{aligned} \tag{22}$$

where $\Omega_{\alpha,\beta} = \tilde{V}_{\alpha,\beta}/2$ are the components of the Rabi-frequency tensor.

If $|E_n| \ll \omega$, where E_n is the characteristic energy of the initially populated Rydberg levels, then, as can be easily shown [10], the spacing between neighboring lower-lying resonance levels can be greater than all other frequency parameters in the problem. Thus, if the initially populated Rydberg levels are in resonance with a lower-lying state (i.e., when $|E_n - \tilde{E}_{m_0} - \omega| \sim \Delta$, where Δ is the spacing introduced in equation (8) between neighboring Rydberg energy levels and $\tilde{E}_{m_0}$ is the lower-lying energy level closest to resonance), then in equation (22) only one term (the term corresponding to the index m_0) need be retained in the sum over m'.

Once again, we shall employ the approximation of index-independent matrix elements ($\Gamma_{n,n'} \approx \Gamma$, $\Omega_{n,m_0} \approx \Omega_R$) and an equidistant level spectrum E_n ($E_n \approx E_{n_0} + (n - n_0)\Delta$). Assuming, as we did in the preceding section, that the temporal profile of the laser pulse is square, we shall solve the system (22) by the method of eigenvalues (quasienergies) and the corresponding eigenvectors, i.e., we seek the solution of equation (22) in the form

$$\tilde{a}_{m_0} = \sum_j \tilde{C}_{m_0,j}\exp(-i\gamma_j t + i\omega t),$$
$$a_n = \sum_j C_{n,j}\exp(-i\gamma_j t). \tag{23}$$

The formulas (23) are a modification of the relation (6) for an atomic model corresponding to *V*-type transitions. Substituting the expressions (23) into equation (22) easily gives the general form of the solution of the system of equations (23):

$$\tilde{a}_{m_0}(t) = -\Omega_R \sum_{\gamma_j} \frac{B(\gamma_j)}{A(\gamma_j)}\exp(-i\gamma_j t + i\omega t).$$
$$a_n(t) = \sum_j \frac{-\Omega_R^2 + i(\Gamma/2)(\gamma_j - \delta)}{[\gamma_j - (n - n_0)\Delta]A(\gamma_j)} B(\gamma_j)\exp(-i\gamma_j t), \tag{24}$$
$$A(\gamma) = 1 + i\frac{\Gamma}{2}\sum_{n'}\frac{1}{\gamma - (n' - n_0)\Delta}$$
$$-\left[(\gamma - \delta)i\frac{\Gamma}{2} - \Omega_R^2\right]\sum_{n'}\frac{1}{[\gamma - (n' - n_0)\Delta]^2},$$
$$B(\gamma) = \sum_{n'}\frac{a_{n'}(0)}{\gamma - (n' - n_0)\Delta}.$$

In equation (24), just as in equation (13), $a_n(0)$ are the initial probability amplitudes for finding an atom in the *n*th Rydberg level (it is assumed that initially the population of the lower-lying resonance level is zero), and $\delta = E_{m_0} + \omega - E_{n_0}$ is the detuning from resonance. The quasienergies γ are determined from the characteristic equation

$$(\gamma - \delta)\left\{1 + i\frac{\Gamma}{2}\sum_{n'}\frac{1}{\gamma - (n' - n_0)\Delta}\right\}$$
$$-\Omega_R^2\sum_{n'}\frac{1}{\gamma - (n' - n_0)\Delta} = 0. \tag{25}$$

Let us analyze the role of the parameters appearing in equation (25). As we have shown in [8], when the frequency is much greater than the ionization threshold ($\omega \gg |E_n|$)—the condition formulated above—in the field range where nonlinear effects can be observed, the Rabi frequency Ω_R is much larger than all other frequency parameters—Δ and Γ—in equation (25). Then an approximate solution of the equation can be sought by expanding in powers of Ω_R^{-1}. In the first nonvanishing approximation the real and imaginary parts of γ_j have the form

$$\mathrm{Re}[\gamma_j] = \left(1 + \frac{1}{2}\right)\Delta - \left(\frac{\Delta}{\pi\Omega_R}\right)^2\left[\left(j + \frac{1}{2}\right)\Delta - \delta\right],$$
$$\mathrm{Im}[\gamma_j] = -\frac{\Gamma}{2}\left(\frac{\Delta^2}{\pi\Omega_R^2}\right)^2\left[j + \frac{1}{2} - \frac{\delta}{\Delta}\right]^2. \tag{26}$$

As follows from equation (26), because of resonance mixing of Rydberg levels with the lower-lying level $\tilde{E}_{m_0}$, ionization is strongly suppressed ($|\mathrm{Im}[\gamma_j]| \ll \Gamma$), and a large number of levels described by the inequality

$$\left|j + \frac{1}{2} - \frac{\delta}{\Delta}\right| < \pi\left(\frac{\Omega_R}{\Delta}\right)^2 \tag{27}$$

are drawn into resonance.

When the condition (27) no longer holds, the resonance interaction of the Rydberg levels with the lower-lying level $\tilde{E}_{m_0}$ is negligible, and the quasienergies are now determined not by the formulas (26) but rather by the relations obtained by solving the problem of the evolution of a Λ-type system (see equations (14) and (15)).

As follows from equation (26), the imaginary part of individual quasienergies can be made to vanish by varying the detuning δ. For an equidistant spectrum this situation is realized when the detuning δ from resonance satisfies

$$\delta = (s + 1/2)\Delta, \tag{28}$$

where s is an arbitrary integer. In this case the quasienergy level γ_s is equal to the detuning δ, and its imaginary part vanishes exactly (this is easily shown by substituting the expression (26) (with $j = s$) into equation (25)). Thus, when the condition (28) is satisfied, the quasienergy spectrum contains an absolutely stable level whose population is "locked": an electron cannot leave a level γ_s into the continuum, irrespective of the duration of the laser pulse.

Without loss of generality of the exposition we can remain with the case $\delta = \Delta/2$, since adding to the detuning δ an integer multiple of the value of Δ can be reduced to the same case by renumbering the infinite system of Rydberg levels. The ionization probability can be calculated using equations (24) only numerically, and an analytic solution of the problem can be sought only in the limiting case of long-duration laser pulses. If a laser field interacts with the atom for a period of time long enough so that all quasienergy levels with the exception of the level with zero width can be emptied, then in equations (24) the summation over j can be removed, retaining only one term corresponding to the stable level $\gamma = \Delta/2$, since all other terms will be exponentially small. The criterion for a long pulse duration for which such an approach is valid is

$$|\mathrm{Im}[\gamma \neq \Delta/2]|t > 1, \tag{29}$$

where t is the time during which the pulse interacts with the atom. Using equation (26) the latter formula can be written as

$$\Gamma\left(\frac{\Delta^2}{\pi\Omega_R^2}\right)^2 j^2 t > 1, \quad j = \pm 1, \pm 2, \ldots \tag{30}$$

To estimate the pulse duration, satisfying the condition (30), the parameters Γ and Δ can be set equal to one another. In addition, if the inequality (30) holds for $|j| = 1$, then it clearly holds for other values of j. Thus the criterion for a long pulse duration will be

$$t > \left(\frac{\Omega_R}{\Delta}\right)^4 T_K. \tag{31}$$

Here T_K is the classical period of revolution, introduced in the preceding section, of an electron on a Kepler orbit.

The condition (31) is very stringent: the pulse duration must be several times greater than the Kepler period. Assuming the inequality (31) to hold, an approximate (neglecting contributions from all quasienergies except $\gamma = \Delta/2$) calculation of the ionization probability gives

$$W_{\mathrm{ion}} = 1 - \frac{(\Omega_R/\Delta)^2}{1 + \pi^2(\Omega_R/\Delta)^2}\left|\sum_n \frac{a_n(0)}{n - n_0 - 1/2}\right|^2. \tag{32}$$

It is easy to see that when only one Rydberg level is populated initially, $a_n(0) = \delta_{n,k}$ ($\delta_{n,k}$ is the Kronecker delta function), the formula (32) is identical to the formula obtained in [6]. However, in the more general case where several atomic levels are coherently populated initially, the formula (32) is much richer.

If the initial distribution $\{a_n(0)\}$ is fixed, then as the field increases the ionization probability decreases, reaching in the limit of a laser field of infinite intensity the value

$$1 - \frac{1}{\pi^2}\left|\sum_n \frac{a_n(0)}{n - n_0 - 1/2}\right|^2 \geq 0,$$

i.e., as the field intensity increases, the atom becomes increasingly more stable with respect to ionization decay. The ionization probability (32) is maximum and becomes equal to 1 if $a_n(0)$ is an "even" function of the difference $n - n_0$ (i.e., $a_{n_0-(n-n_0)} = a_{n_0+(n-n_0)+1}$, $n = n_0$, $n_0 \pm 1$, $n_0 \pm 2$ …). Then the sum presented in equation (32) vanishes exactly.

What is the minimum value of the ionization probability (32) and for what initial distribution is it obtained? It is obvious that if the population amplitudes of the Rydberg levels are initially chosen so that when the laser field is switched on they form a vector corresponding to quasienergy $\gamma = \Delta/2$, then the population of the atom remains unchanged (because the quasienergy level does not have the decay width) when the laser field is switched off, and therefore the ionization probability of the atom will be zero in this case. Thus, the problem of searching for the most stable state of an atom reduces to finding the state vector corresponding to the quasienergy level $\gamma = \Delta/2$. It is easy to calculate this vector. Taking account of the normalization, it is given by

$$\begin{aligned} \tilde{a}_{m_0} &= -[1 + \pi^2(\Omega_R/\Delta)^2]^{-1/2} e^{-i\omega t}, \\ a_n &= \frac{\Omega_R/\Delta}{[1 + \pi^2(\Omega_R/\Delta)^2]^{1/2}} \frac{1}{n - n_0 - 1/2}. \end{aligned} \tag{33}$$

We recall that everywhere in the discussion above we solved the problem of the ionization of a V-type quantum system in which only the upper levels E_n were populated initially. The state (33) will satisfy this condition in the limit of a high-intensity field, $\Omega_R \gg \Delta$, when in equation (33) $\tilde{a}_{m_0}$ approaches zero. Such a state gives a minimum ionization probability (32) on the class of states studied on the basis of this work. The higher the intensity of the laser field, the smaller the emergence of an electron from this state is.

Such a state can be represented explicitly as

$$\begin{aligned} \tilde{a}_{m_0} &= 0, \\ a_n &= \frac{1}{\pi} \frac{1}{n - n_0 - 1/2}, \end{aligned} \tag{34}$$

and the formula for the probability calculated according to the state (34) has the form

$$W_{\text{ion}} = 1 - \frac{\pi^2(\Omega_R/\Delta)^2}{1+\pi^2(\Omega_R/\Delta)^2} = \frac{1}{1+\pi^2(\Omega_R/\Delta)^2}. \quad (35)$$

As follows from equation (35), even in comparatively weak fields, for which $\Omega_R \sim \Delta$, the ionization probability does not exceed 10% and drops rapidly with increasing field (inversely as the laser intensity).

The formulas (32) and (35) are valid only when the condition (31) is satisfied, i.e., in the approximation of a laser pulse with infinite duration. The problem was solved numerically for finite interaction times with the electromagnetic field. Figure 4 displays the dependence of the ionization probability of an atom which initially is in the state (34) on the laser field strength (expressed in dimensionless units, $V = F/\omega^{5/3}$) for different durations of the laser pulse. As one can see in the figure, the probability curve reaches its asymptotic limit (35) all the earlier the longer the field-atom interaction time, but even for short pulses with $t = T_K/6$ the asymptotic limit is reached in fields $V < 1$.

Figure 5 also shows the probability of ionization from the state (34), but as a function of the detuning δ. As one can see from the figure, in a strong field the ionization probability is insensitive to a change in the detuning. This result corresponds to the conclusions drawn in [10], we showed that in the strong-field limit the ionization of a Rydberg atom becomes increasingly less sensitive to the detuning.

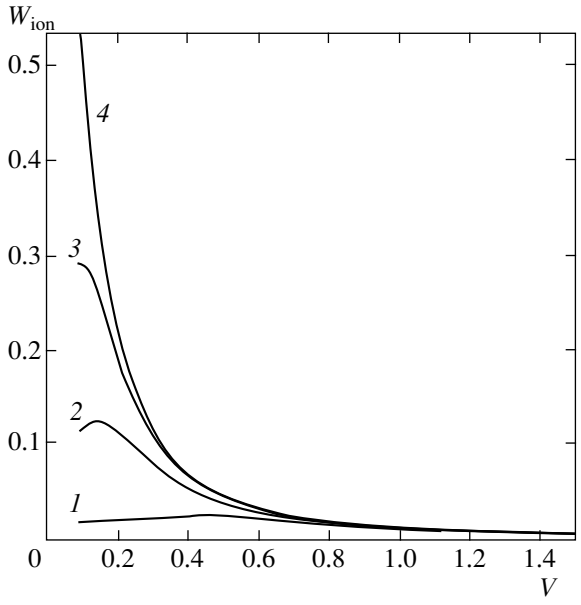


Fig. 4. Ionization probability of an V system (equidistant spectrum) initially in an "absolutely stable" state (34) as a function of the dimensionless parameter $V = F/\omega^{5/3} = (\Gamma/\Delta)^{1/2}$. Pulse duration t = (*1*) T_K, (*2*) $5T_K$, (*3*) $15T_K$, (*4*) $t \longrightarrow \infty$. The values of all other parameters are $\delta = \Delta/2$ and $\Omega_R^2/\Gamma = 9$. The computational scheme contains 18 levels.

4. ANALYSIS OF THE RESULTS

As follows from the two preceding sections, the dependence of the ionization probability on the initial population distribution on Rydberg levels in the limit of long pulse duration ($>T_K$) is manifested only in the case of a system in which a resonance interaction (V-type transitions) is taken into account. Therefore a possible experiment to observe the indicated effect at such times necessarily requires examining atoms and laser fields forming a scheme of V-type transitions (i.e., the spectrum of the atom and the frequency of the laser field must be such that the populated Rydberg levels initially are in a resonance state with respect to any lower-lying atomic level).

All results presented in the preceding section for a V-type system were obtained on the basis of an equidistant model of an atomic spectrum. However, it is easy to show that the conclusions drawn in the preceding section remain valid when the nonequidistant nature of the atomic spectrum and the dependence of the oscillator strengths on the principal quantum number n are taken into account. For this it is sufficient to show that for an arbitrary nonequidistant spectrum there always exists a detuning δ such that one of the levels in the quasienergy spectrum will possess zero width. Two

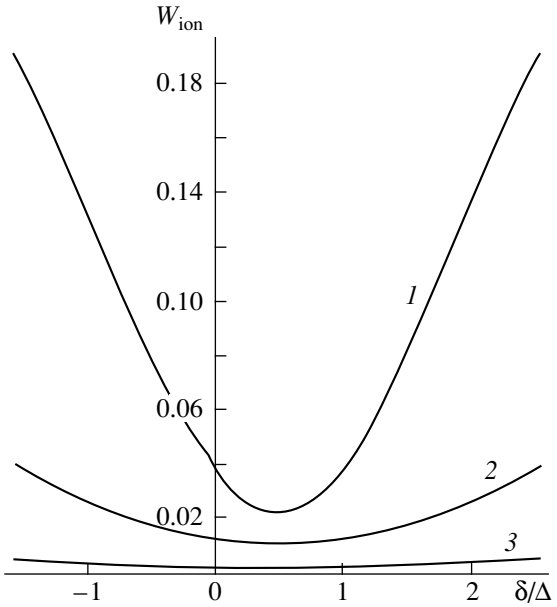


Fig. 5. Ionization probability of a V system (equidistant spectrum) initially in an "absolutely stable" state (34) as a function of the detuning δ. V = (*1*) 0.3, (*2*) 1.0, (*3*) 3.0. The values of all other parameters are $t = T_K$, $\Omega_R^2/\Gamma = 9$. The computational scheme contains 18 levels.

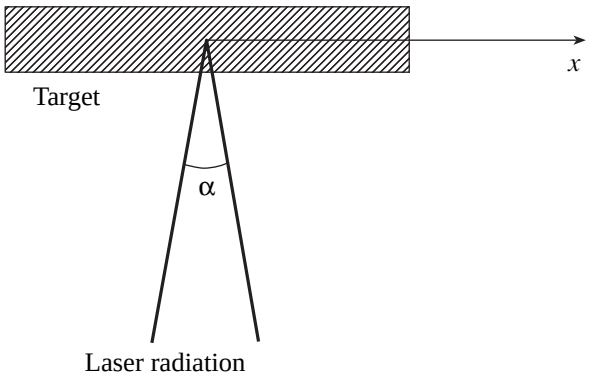


Fig. 6. Diagram of an experiment designed to observe the controllable ionization of a Rydberg atom by double laser pulses.

questions remain. How can such coherent states whose ionization is either virtually complete or, conversely, almost zero be produced? Can the emergence of an electron be controlled by regulating the initial state of the atom?

The technology for selectively exciting a distinguished Rydberg level is now well developed. If an atom with one excited level is prepared, then there is no difficulty in "smearing" the atomic population over neighboring Rydberg levels. It is sufficient to expose the atom to an intense laser field and then, on account of Raman transitions through the continuum (Λ type) or lower-lying resonance state (V-type), the atomic population becomes redistributed over levels neighboring the initially excited levels. Photoionization from the coherent state produced can be observed by exposing the atom to a second laser pulse.

However, let us follow in greater detail the state produced by the first pulse. We require that the parameters of the pulse—the frequency, intensity, duration, and polarization—be the same as for the second pulse. We also require that the frequency permit the presence of a stable state in the quasienergy spectrum (equation (28) in the case of an equidistant spectrum). If the duration of both pulses is sufficiently long, then there is enough time for all quasienergy states with a nonzero ionization width to decay over the interaction time with the first pulse. This means that by the moment when the atom starts to interact with the second laser pulse the distribution of the probability amplitudes for population of the Rydberg levels, more precisely, their absolute values, is identical to the analogous distribution for an absolutely stable state.

Does this mean that ionization of the atom by the second pulse is impossible? A careful analysis shows that this is not the case. It is necessary to take account of the fact that the phases of the amplitudes a_n for expansion of the wave function of an atomic electron over the spectrum of a free atom (when the first pulse is switched off) evolved differently in time. If a distribution of the atomic population $\{a_n\}$ was present when the pulse was switched off, then at the moment the second pulse is switched on this distribution will be different:

$$\begin{aligned} a_n(0) &= a_n \exp(-iE_n\tau) \\ &= \{a_n \exp[-i(E_n - E_{n_0})\tau]\} \exp(-iE_{n_0}\tau). \end{aligned} \tag{36}$$

We designated the distribution of probability amplitudes at the moment the second pulse is switched on in accordance with the notations used in the preceding sections: τ is the time delay between the two laser pulses and E_{n_0} is the energy level populated up to the start of the interaction.

For an equidistant spectrum $E_n - E_{n_0} = (n - n_0)\Delta$. It is obvious that if the delay time τ between the two laser pulses is a multiple of the Kepler period $T_K = 2\pi\Delta^{-1}$, then to within an irrelevant general phase factor the distribution (36) will transform into $\{a_n\}$, and at the moment the second pulse is switched on the atom will be in an absolutely stable state, the probability of ionization from which is zero. However, the situation will be qualitatively different for other delay times. Moreover, it is easy to see that for τ equal to a half-integer of the Kepler period T_K, it will be completely opposite - in this case the atom ionizes completely. It is easy to arrive at this conclusion by noting that if the distribution $\{a_n\}$ were "odd" (i.e., $a_{n_0-k} = -a_{n_0+k+1}$, $k = 0, 1, 2, \ldots$), which is valid for an absolutely stable state (see preceding section), then the distribution (36) for $\tau = \pi/\Delta$ becomes "even" ($a_{n_0-k} = a_{n_0+k+1}$), and for such states the ionization probability (32) becomes 1. Thus, the probability of ionization of an atom by a second pulse as a function of the delay time varies from 0 ($\tau = 0$, T_K, $2T_K$, …) to 1 ($\tau = T_K/2$, $3T_K/2$, …); it is easy to see that the ionization probability is T_K-periodic. (We note that the periodicity of the structure of this function for a real Rydberg atom on long times will break down because the atomic spectrum is naturally nonequidistant.)

The technique for producing a pair of pulses separated by a controllable delay time is now quite well developed. An experiment for checking the results of the theory expounded here can employ the scheme shown in Fig. 6, similar to the scheme proposed in [14]. Two wide laser beams, propagating in directions making an angle α with one another, converge on a target consisting of a rarefied gas of Rydberg atoms. If it is assumed that at the point $x = 0$ the time delay between the pulses is zero, then at an arbitrary point the time delay is given by the formula

$$\tau = \frac{x}{c} 2\sin\frac{\alpha}{2} \approx \frac{x}{c}\alpha. \tag{37}$$

Thus, if the spatial ionization pattern is photographed after both pulses interact with the atom, then a periodic structure with period cT_K/α will be obtained. Figure 7a shows the dependence of the ionization prob-

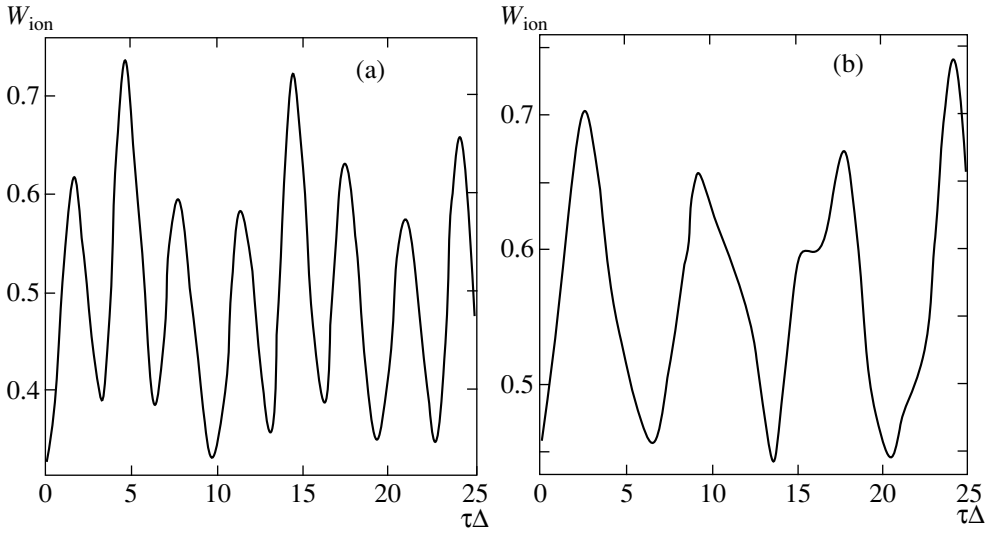


Fig. 7. Probability of ionization of a Rydberg atom by successive laser pulses as a function of the delay between the pulses: (a) $n_0 = 5$, $\tilde{m}_0 = 2$, $\omega \approx 4 \times 10^{15}$ s^{-1}, $F = 1 \times 10^8$ V/cm ($I = 3 \times 10^{13}$ W/cm^2), $t = 15T_K = 0.3$ ps; the computational scheme contains 6 levels; (b) $n_0 = 25$, $\tilde{m}_0 = 10$, $\omega \approx 1.7 \times 10^{14}$ s^{-1}, $F = 6 \times 10^5$ V/cm ($I = 1 \times 10^9$ W/cm^2), $t = 50T_K = 120$ ps; the computational scheme contains 18 levels.

ability on the delay time between the two pulses for the case at the limit of applicability of the theory expounded in this paper: the principal quantum number of the initially populated level is relatively small, $n_0 = 5$. Nonetheless, even for such a weakly excited state the dependence of the decay probability on τ is appreciable (although this dependence is not periodic because of the strongly nonequidistant nature of the atomic spectrum for small principal quantum numbers). Switching in the plot from the time coordinate to the spatial coordinate x (assuming $\alpha = 5°$), the characteristic scale of the structure obtained will be 10^{-2} cm. The fact that the ionization probability is different from zero at $\tau = 0$ can be explained by the fact that the first pulse "washes out" a definite fraction of the atomic population.

The conditions of an experiment capable of revealing the dependence of the emergence of an electron on the initial population of the atom which are consistent with the assumptions and the approximations employed in this work can be chosen as follows: $n_0 = 25$, $\tilde{m}_0 = 10$, $\omega \approx 1.7 \times 10^{14}$ s^{-1} (more accurately, the frequency must be determined taking into account the characteristic features of the spectrum of a specific atom), $F = 6 \times 10^5$ V/cm ($I = 1 \times 10^9$ W/cm^2), $t = 100$ ps (t is the duration of each pulse), and $\alpha = 5°$. In this case the experiment should give a picture close to the one shown in Fig. 7b (the period of the spatial structure corresponding to the curve presented in Fig. 7b is of the order of 1 cm).

5. CONCLUSIONS

Summarizing the results obtained in this work, there is no question that the ionization process for a Rydberg atom is extremely sensitive to the initial state of the atom. However, for two different transition schemes, Λ and V types, between Rydberg levels the effect of the initial population of the atom on the ionization process is substantially different.

For a Λ-type system a coherent initial population of the atom results either in suppression (right up to complete) of the initial ionization rate or, conversely, to a several-fold intensification of the rate. Nonetheless, the ionization rate is a strongly oscillating function of time and, as a result, on average over a Kepler period the same fraction of the atomic population, independent of the initial population of the atom, emerges into the continuum. Thus, for a Λ system the dependence of the ionization process on the initial coherent state of the atomic electron is observable only at times not exceeding the Kepler period $T_K = 2\pi\Delta^{-1}$.

The situation is completely different for a V system. The ionization probability is strongly dependent on the initial state of the atom, and this dependence is all the sharper the longer the interaction time with the laser pulse (this interaction time must at least exceed the Kepler period). Depending on the state the electron occupies at the moment the interaction of the atom with the laser field is switched on, the electron either completely emerges into the continuum ($W_{ion} = 1$) or, conversely, it remains completely in the bound state ($W_{ion} = 0$).

Such a difference in the behavior of the Λ and V systems can be explained by the fact that for a Λ system (containing an infinite number of equidistant energy levels) the widths of all quasienergy levels are the same and therefore the decay time of such a system is the same irrespective of its initial state. For a V system,

however, the widths are different for different quasienergy levels and, moreover, for a definite choice of field parameters the state with zero width, absolutely stable with respect to ionization decay, can be singled out. The closer the initial population of the atom is to this state, the more stable the atom is for an arbitrarily long interaction time with the laser pulse.

In the present paper the conditions of an experiment (see preceding section) in which the phenomenon described can be observed were formulated: the emergence of an electron depends on the initial population of the Rydberg atom. The results of the numerical experiment are presented in Fig. 7. It seems to us that the proposed scheme—improved and modified—for controlling the ionization of an atom using a laser field can be used in telecommunications and computer technology.

Of course, the dependence of the ionization of a Rydberg atom on the initial coherent population of the atom is manifested only when all conditions and assumptions stipulated in the preceding sections are satisfied. The fundamental condition, and possibly the one most difficult to satisfy, is the condition that the atom-field interaction is switched on and off instantaneously (the "instantaneousness" means that the characteristic time for switching the interaction on and off must be much shorter than the Kepler period). The degree to which the results of the problem are sensitive to the approximations is an independent problem, the solution of which is urgent not only for the present work but also for other problems concerning the interaction of a Rydberg atom with radiation. This question merits further investigations, but we are confident that they will not affect qualitatively the results of the present work.

ACKNOWLEDGMENTS

This work was supported by the Russian Foundation for Basic Research (project nos. 99-02-18034 and 97-02-71024).

REFERENCES

1. M. V. Fedorov and A. M. Movsesian, J. Phys. B **21**, L155 (1988).
2. L. D. Noordam, H. Stapelfeldt, and D. I. Duncan, Phys. Rev. Lett. **68**, 1496 (1992).
3. J. H. Hoogenraad, R. B. Vrijen, and L. D. Noordam, Phys. Rev. A **50**, 4133 (1994).
4. M. Yu. Ivanov, Phys. Rev. A **49**, 1165 (1994).
5. A. Wójcik and R. Parzinski, Phys. Rev. A **50**, 2475 (1994).
6. A. Wójcik and R. Parzinski, J. Opt. Soc. Am. B **12**, 369 (1995).
7. E. A. Volkova, A. M. Popov, O. V. Smirnova, and O. V. Tikhonova, Zh. Éksp. Teor. Fiz. **111**, 1194 (1997) [JETP **84**, 658 (1997)].
8. M. V. Fedorov and N. P. Poluektov, Laser Phys. **7**, 299 (1997).
9. M. V. Fedorov and N. P. Poluektov, Opt. Express **2**, 51 (1998).
10. N. P. Poluéktov and M. V. Fedorov, Zh. Éksp. Teor. Fiz. **114**, 821 (1998) [JETP **87**, 445 (1998)].
11. M. V. Fedorov, *Electron in a Strong Optical Field* (Nauka, Moscow, 1991).
12. M. V. Fedorov, Laser Phys. **3**, 219 (1993).
13. N. B. Delone, S. P. Goreslavsky, and V. P. Krainov, J. Phys. B **22**, 2941 (1989).
14. M. B. Campbell, T. J. Bensky, and R. R. Jones, Opt. Express **1**, 197 (1997).

Translation was provided by AIP

Journal of Experimental and Theoretical Physics, Vol. 90, No. 5, 2000, pp. 805–809.
Translated from Zhurnal Éksperimental'noĭ i Teoreticheskoĭ Fiziki, Vol. 117, No. 5, 2000, pp. 926–931.

PLASMA, GASES

Role of Excited States in Coherent Harmonic Bremsstrahlung in a Photoionized Plasma

V. P. Silin

Lebedev Physical Institute, Russian Academy of Sciences, Moscow, 117924 Russia
e-mail: silin@sci.lebedev.ru

Received October 20, 1999

Abstract—It is established for a photoionized plasma formed in the barrier-suppression ionization regime that the preliminary population of the excited states of the ionizing atoms plays an important role. It is established that in this case an anomalously strong (several orders of magnitude) increase occurs in the efficiency of generation of the harmonics of the pump radiation. It is shown that a relative decrease of the harmonics generation efficiency occurs with time as a result of collisions of the electrons produced by ionization. © *2000 MAIK "Nauka/Interperiodica".*

It has long been recognized, correctly, that the optical nonlinearity of plasma increases when the excited states of the atoms and ions become populated [1, 2]. This was first substantiated in a calculation of the cubic optical susceptibility of a gas of excited hydrogen atoms [2]. On the other hand, the concept of using a power series expansion in the components of the interacting fields to describe the nonlinearly optical interaction has very limited applicability for high-harmonics generation in plasma [3]. As a result, specifically, the answer to the question of the role of the excited states population in the generation of high harmonics has never been adequately worked out. In the present paper, a very large increase in the high-harmonics generation efficiency under the conditions of photoionization from excited states is predicted theoretically for harmonics generation accompanying the photoionization of a gas in the generation barrier-suppression regime. A quantitative description of this phenomenon is given for a hydrogen-like atom.

The theory of coherent harmonic bremsstrahlung of the pump field in plasma, employing the electron-ion collision integral in the Fokker–Planck–Landau form [3, 4], has made it possible to write down the following comparatively simple relation for the density $\delta\mathbf{j}$ of the electric current of the source of harmonics generation:

$$\frac{\partial \delta \mathbf{j}}{\partial t} = -\frac{4\pi e^2 e_i^2 n_i \Lambda}{m^2}\int d\mathbf{p}\frac{e\mathbf{V}}{V^3} f_e(\mathbf{p} - m\mathbf{u}_E(t)). \qquad (1)$$

Here e and m are the electron charge and mass, e_i and n_i are the ion charge and density, Λ is the Coulomb logarithm, $f(\mathbf{p})$ is the electron momentum distribution function, and $\mathbf{u}_E(t)$ is the oscillatory velocity of electrons in the pump field. Assuming the electric field $\mathbf{E} = (E_x, 0, 0)$ of the pump wave to be linearly polarized and monochromatic, $E_x = E\cos(\omega t - \phi_x)$, we have $\mathbf{u}_E = (u_{Ex}, 0, 0)$ and $u_{Ex}(t) = -V_E\sin(\omega t - \phi_x)$, where $V_E = (|e|E/m\omega)$.

We shall assume, in accordance with the ideas concerning the barrier-suppression ionization regime [5–7], that an electron is ejected from an atom freely when the pump energy flux density q exceeds the threshold value [6] $Q_{BSI} = (I_i^4\,[\text{eV}]/Z^2) \times 4 \times 10^9$ W cm^{-2}, where I_i is the ionization potential and Z is the nuclear charge of the atom. Then, in accordance with [8], the distribution function of ionized electrons in the coordinate system of their oscillations in the pump field corresponds to the electron momentum distribution in the atom

$$f_e(\mathbf{p}) = \frac{n_i}{4\pi}(mV_Z)^{-3}\sum_q F_q\left(\frac{\mathbf{p}}{mV_Z}\right), \qquad (2)$$

where the summation extends over the quantum states of the electrons in the atom and $V_Z = Ze^2/\hbar = ZV_a$ is the Coulomb unit of velocity [9].

In the present paper we shall demonstrate that the population of the high states of an atom influences the coherent harmonic bremsstrahlung. This influence can be seen for the example of the *ns* states of an electron in a hydrogen-like atom, where

$$F_{ns}\left(\left|\frac{\mathbf{p}}{mV_Z}\right|\right) = F_{ns}(\zeta) = \frac{8n\sin^2(2n\arctan n\zeta)}{\zeta^2(1+n^2\zeta^2)^2}. \qquad (3)$$

Then we obtain from the relation (1)

$$\frac{\partial \delta j_x}{\delta t} = \frac{\partial}{\partial t}\sum_{N=1}^{\infty}\sigma_{xx}^{(2N+1)}E\cos[(2N+1)(\omega t - \phi_x)], \quad (4)$$

where in the strong pump field limit

$$V_E \gg V_Z, \qquad (5)$$

we have

$$\sigma_{xx}^{(2N+1)} = \chi(\omega)\nu(V_Z)(2N+1)^{-3} \times R([2N+1]V_Z/V_E). \quad (6)$$

Here

$$\chi(\omega) = \frac{e^2 n_e}{m\omega^2}, \quad (7)$$

$$\nu(V_Z) = 6\Lambda(Z_i/Z)^2(I_Z/\hbar)Q_Z n_i, \quad (8)$$

$$R(X) = Z_i^{-1}\sum_q X^3 G_q(X), \quad (9)$$

$$G_q(X) = \frac{4}{\pi}\int_0^\infty \zeta^2 F_q(\zeta)\left(\frac{\sin X\zeta}{X\zeta} - \mathrm{Ci}(X\zeta)\right)d\zeta, \quad (10)$$

where Z_i is the ionization multiplicity of the ion, $I_Z = Z^2 I_H$, $O_Z = (4\pi/3)a_Z^3 = Z^{-3} \times 6.3 \times 10^{-25}$ cm^{-3}, $a_Z = \hbar^2/Ze^2m$ is the Coulomb unit of length [9], $I_H = me^4/2\hbar^2$ is the ionization potential of the hydrogen atom, and

$$\mathrm{Ci}(x) = \int_x^\infty \frac{dt}{t}\cos t$$

is the cosine integral [10].

It what follows we shall examine for the *ns* states of a hydrogen-like atom the function [8]

$$S_{ns}(X) = \frac{X^3}{16}G_{ns}(X) = \frac{2X^3 n}{\pi^2} \times \int_0^\infty d\zeta \frac{\sin^2(2n\arctan n\zeta)}{(1+n^2\zeta^2)^2}\left(\frac{\sin X\zeta}{X\zeta} - \mathrm{Ci}(X\zeta)\right), \quad (11)$$

characterizing, in accordance with equation (3), the dependence of the nonlinear partial conductivity (6) on the scaling parameter

$$X = (2N+1)V_Z/V_E. \quad (12)$$

First, we present here the following expressions for $n = 1, 2, 3$:

$$S_{1s}(X) = \frac{X^3}{4\pi}\left\{E_1(X) + \left(\frac{4}{3} + \frac{X}{3}\right)e^{-X}\right\},$$

$$S_{2s}(X) = \frac{X^3}{4\pi}\left\{E_1\left(\frac{X}{2}\right) + \left(\frac{26}{15} + \frac{11}{15}\left(\frac{X}{2}\right)\right.\right.$$

$$\left.\left. + \frac{1}{5}\left(\frac{X}{2}\right)^2 + \frac{1}{15}\left(\frac{X}{2}\right)^3\right)e^{-X/2}\right\}, \quad (13)$$

$$S_{3s}(X) = \frac{X^3}{4\pi}\left\{E_1\left(\frac{X}{3}\right) + \left(\frac{64}{35} + \frac{29}{35}\left(\frac{X}{3}\right) + \frac{82}{35}\left(\frac{X}{3}\right)^2\right.\right.$$

$$\left.\left. + \frac{26}{315}\left(\frac{X}{3}\right)^3 - \frac{4}{315}\left(\frac{X}{3}\right)^4 + \frac{2}{315}\left(\frac{X}{3}\right)^5\right)e^{-X/3}\right\},$$

where

$$E_1(X) = \int_X^\infty \frac{dt}{t}\exp(-t)$$

is the exponential integral [10]. We now consider certain asymptotic relations for the functions (11). In the limit

$$V_E \gg (2N+1)V_Z \quad (14)$$

the following asymptotic expression follows from the formula (11):

$$S_{ns}(X) = \frac{X^3}{4\pi}\left\{\ln\frac{n}{X} - C + 2 - \frac{2n}{4n^2-1}\right\}, \quad (15)$$

where $C = \gamma = 0.577$ is Euler's constant [10]. In the opposite limit,

$$V_Z \ll V_E \ll (2N+1)V_Z \quad (16)$$

we have from equation (11)

$$S_{ns}(x) \cong \frac{X^3}{4\pi(2n+1)!}\left(\frac{2X}{n}\right)^{2n-1}\exp\left(-\frac{X}{n}\right). \quad (17)$$

The formulas (15) and (17) indicate the existence of scaling associated with the variable

$$\xi = \frac{X}{n} = \frac{(2N+1)V_Z}{nV_E}. \quad (18)$$

This makes it possible to write

$$S_{ns}(X) \simeq \bar{S}_{ns}(\xi), \quad (19)$$

where in the case (15)

$$\bar{S}_{ns}(\xi) = \frac{n^3\xi^3}{4\pi}\left\{\ln\frac{1}{\xi} - C + 2 - \frac{2n}{4n^2-1}\right\}, \quad (20)$$

and in the case (17)

$$\bar{S}_{ns}(\xi) \simeq \frac{n^3\xi^3(2\xi)^{2n-1}}{4\pi(2n+1)!}\exp(-\xi). \quad (21)$$

For highly excited states ($n \gg 1$) the latter formula can be written, using Stirling's formula, in the form

$$\bar{S}_{ns}(\xi) = \frac{n^3(n+1)^{1/2}}{\pi^{3/2}} \times 2^{2n-3} \times \left(\frac{\xi}{2n+2}\right)^{2n+2} \exp(-\xi + 2n + 2). \quad (22)$$

The maximum of the expression (22) occurs at

$$\xi = \xi_{\max} = 2n + 2, \quad (23)$$

and then

$$\bar{S}_{ns}(\xi_{\max}) = \frac{n^3(n+1)^{1/2}}{2\pi^{3/2}} \times 2^{2n-2}. \quad (24)$$

The latter formula describes the very strong dependence on the principal quantum number n of the energy states of an electron in a hydrogen-like atom. For a hydrogen atom these states correspond to highly excited states of an electron.

We shall now use the relations obtained above for the characteristics of the harmonics generation efficiency as a function of the principal quantum number n. For this, we shall consider the case of a plane-wave pump field, which corresponds to

$$\phi_x = kz, \quad k = \frac{\omega}{c}\sqrt{1 - 4\pi\chi(\omega)}. \quad (25)$$

Substituting the expression (4) into Maxwell equations in plasma and using the relations (25) gives the harmonics field in the form of plane waves [3, 4]. Correspondingly, we obtain the following expression for the ratio of the energy flux density of the radiation of the $2N + 1$ harmonic to the energy flux density q of the pump:

$$\eta_{ns}^{(2N+1)} = \frac{q^{(2N+1)}}{q} = \left(\frac{\nu(V_Z)}{Z_i\omega}\right)^2 h^2(N)\bar{S}_{ns}^2\left(\frac{[2N+1]V_Z}{nV_E}\right), \quad (26)$$

where [8]

$$h(N) = \left[N\left(N + \frac{1}{2}\right)^2(N+1)\right]^{-1}. \quad (27)$$

The generation efficiency $\eta_{ns}^{(2N+1)}$ of various harmonics is characterized by the function

$$S_{ns}^2\left([2N+1]\frac{V_Z}{V_E}\right) = \bar{S}_{ns}^2\left([2N+1]\frac{V_Z}{nV_E}\right). \quad (28)$$

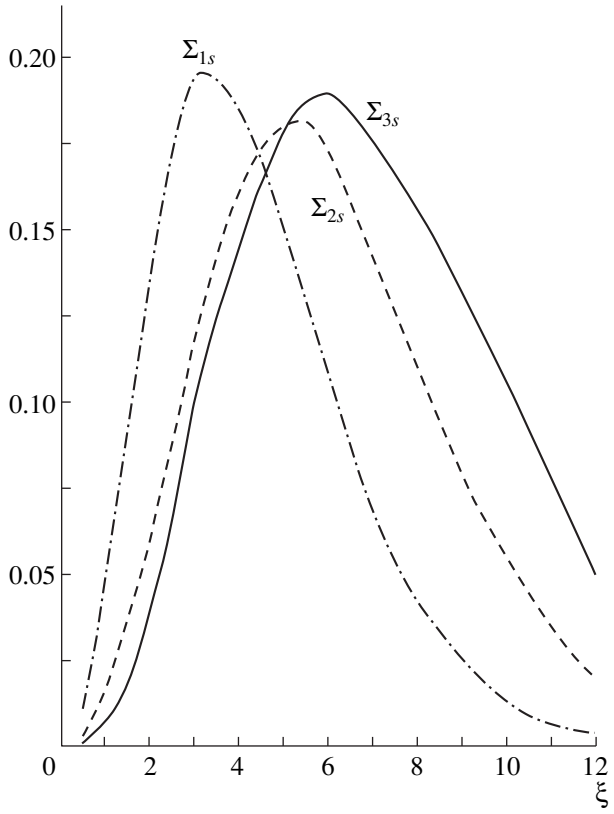
The approximate formulas (22) and (24) show the extremely pronounced dependence of (28) on the principal quantum number n of the excited state for $n \gg 1$. On the other hand, it is important to know the corresponding dependence for small principal quantum numbers also. We shall show that a sharp dependence, similar to equation (24), also holds for small values of n. For this we use the functions

$$\Sigma_{ns}(\xi) = \frac{\bar{S}_{ns}(\xi)}{n^3(n+1)^{1/2} \times 2^{2n-2}}, \quad (29)$$

which characterize, specifically, the difference of the maxima of the functions $\bar{S}_{ns}(\xi)$ from the scaling corresponding to the law (24). Three curves corresponding to the functions (29) for 1s, 2s, and 3s states are presented in the figure. The maximum values of these curves are approximately the same. This corresponds to the fact that the scaling (24) holds for small values also. We note that the corresponding maximum of these curves is approximately two times greater than the approximate value which can be obtained from the formula (24). The values of the ratios of the squared maximum values of the functions determined by the formulas (13) follow from the figure and equations (28) and (29):

$$\left(\frac{S_{2s,\max}}{S_{1s,\max}}\right)^2 \simeq 1300, \quad \left(\frac{S_{3s,\max}}{S_{1s,\max}}\right)^2 \simeq 350\,000. \quad (30)$$

Thus, ionization from the 2s state increases the maximum efficiency of harmonics generation as compared



Curves corresponding to the function (29); scaling variable $\xi = ([2N + 1]V_Z/nV_E)$.

with ionization from the 1*s* state by three orders of magnitude, and ionization from the 3*s* state increases the efficiency by 5.5 orders of magnitude. The scaling (24) corresponds to the ratio

$$\left(\frac{S_{ns,\max}}{S_{1s,\max}}\right)^2 \approx 0.008n^7 \times 2^{4n}. \tag{31}$$

The increase in the efficiency of harmonics generation by electrons freed from excited states is nonstationary. The time variation of this effect is due to collisions of electrons stripped from atoms. Since electron–ion collisions are suppressed in a strong pump field [3], the main collisions are electron–electron collisions, characterized by the electron–electron collision integral $J_{ee}[f_e, f_e]$ (see, for example, [11]). However, we shall not require the explicit form of the collision integral here. The temporal evolution of the electronic distribution function at short times of the order of the electron–electron collision time t_{ee} (appreciable heating of electrons has still not appeared) is described by the equation

$$\frac{\partial f_{ee}}{\partial t} = J[f_e, f_e]. \tag{32}$$

In our situation it must be assumed that initially, at $t = 0$, the electron distribution function corresponds to electrons arising from the *ns* states:

$$f_e(p,t) = \frac{2n_e n}{\pi^2 (mV_Z)^3} \frac{\sin^2\left(2n \arctan\left[\frac{np}{mV_Z}\right]\right)}{\left(\frac{p}{mV_Z}\right)^2 \left(1 + \left[\frac{np}{mV_Z}\right]^2\right)^2}. \tag{33}$$

The equation (32) possesses two conserved integrals. The first integral

$$\int d\mathbf{p}\, f_e = n_e = \text{const} \tag{34}$$

corresponds to conservation of the number of particles in collisions, and the second integral corresponds to conservation of the kinetic energy in collisions

$$\int d\mathbf{p}\left(\frac{p^2}{2m}\right) f_e = n_e\left(\frac{mV_Z^2}{2n^2}\right) = \text{const}. \tag{35}$$

Aside from this, in accordance with Boltzmann *H* theorem [11], the initial distribution (33) relaxes to Maxwell distribution, for which, according to equation (34), the electron density is known and the temperature is determined by the fact that in the final state the integral (35) is equal to $(3/2)n_e k_B T$, where k_B is Boltzmann constant. Thus, the electron temperature is

$$k_B T = (mV_Z^2/3n^2). \tag{36}$$

Since the temperature is inversely proportional to the squared principal quantum number, in the final Maxwellian state arising after collisional relaxation the comparatively cold electrons arising from the excited states generate harmonics more efficiently. This can be seen directly from the scaling of the maximum of the generation efficiency of high harmonics by Maxwellian electrons

$$\eta_{\max}^{(2N+1)} = 0.006(\nu_{ei}/N^4\omega)^2, \tag{37}$$

established in [12] (see also the review [13]). Indeed, since

$$\nu_{ei} = \frac{4\sqrt{2\pi}e^2 e_i^2 n_i \Lambda}{3m^{1/2}(k_B T)^{3/2}}, \tag{38}$$

we have according to equations (36) and (38)

$$\nu_{ei} \simeq 1.4n^3 \nu(V_Z). \tag{39}$$

Therefore, we obtain

$$\eta_{\max}^{(2N+1)} \simeq 0.012\left[\frac{\nu(V_Z)}{\omega}\right]^2 \frac{n^6}{N^8} \tag{40}$$

for the efficiency of high-harmonics generation by electrons from *ns* states after their distribution relaxes to a Maxwell distribution. According to the latter formula the efficiency of generation by 2*s* and 3*s* electrons is 64 and 729, respectively, times greater than the efficiency of generation by 1*s* electrons. However, even though the increase in the efficiency is comparatively large compared with the efficiency of generation by 1*s* electrons, the relaxation to the Maxwell distribution decreases the relative increase in the harmonics generation efficiency. Indeed, comparing the consequences of the formula (40) with the relations (30) shows that as a result of the relaxation of 2*s* and 3*s* states to a Maxwell distribution, the relative generation efficiency decreases approximately by a factor of 20 and 480, respectively.

A Maxwell distribution is established in the characteristic time

$$t_{ee} \approx \frac{3m^{1/2}(k_B T)^{3/2}}{4\sqrt{2\pi}e^4 n_e \Lambda} \approx \frac{1}{n^3}\left(\frac{10}{\Lambda}\right)\frac{Z^3}{Z_i}\frac{10^{-13}\ \text{s}}{n_i/10^{18}\ \text{cm}^{-3}}. \tag{41}$$

Consequently, for laser pump pulses with a shorter duration there is not enough time for the high relative efficiency corresponding to the distribution (33) to decrease. Conversely, for long times the relative efficiency can be described by the law (40), which conserves the relative increase in the harmonics generation by high *s*-state electrons, though this effect is not so strikingly expressed as for an electron momentum distribution corresponding to the intraatomic distribution.

To understand best the quantitative magnitude of the effects being discussed, we shall present some relations which are convenient for applications. Thus, the ratio of

the characteristic collision frequency and the pump field frequency can be written as

$$\frac{\nu(V_Z)}{\omega} = 5\times 10^{-4} \times\left\{\left(\frac{\Lambda}{10}\right)\left(\frac{Z_i^2}{Z^3}\right)\frac{n_i\ [18]}{\hbar\omega\ [\mathrm{eV}]}\right\} \equiv 5\times 10^{-4}\Pi, \tag{42}$$

where $\hbar\omega$ [eV] is the pump frequency and n_i [18] is the number density of ions in units of 10^{18} cm^{-3}. From equations (26) and (42) we obtain for the maximum generation efficiency of high harmonics $(2N+1) \gg 1$

$$\eta_{ns,\max}^{(2N+1)} \simeq 2\times 10^{-8}\frac{\Pi^2}{Z_i^2 N^8}\left(\frac{\bar{S}_{ns,\max}}{\bar{S}_{1s,\max}}\right)^2. \tag{43}$$

Hence, specifically, we have for the generation of the seventh ($N = 3$) and fifteenth ($N = 7$) harmonics with ionization of one electron from the 1*s* state

$$\eta_{3s,\max}^{(7)} \simeq 1.1\times 10^{-6}\left(\frac{\Pi}{Z_i}\right)^2,$$

$$\eta_{3s,\max}^{(15)} \simeq 1.2\times 10^{-9}\left(\frac{\Pi}{Z_i}\right)^2.$$

Rewriting equation (43) using the expression (24) we have

$$\eta_{ns,\max}^{(2N+1)} \simeq 5.3\times 10^{-7}\left(\frac{\Pi}{Z_i}\right)^2 N^{-8} n^7 2^{4(n-3)} \ll 1.$$

Hence, for example, we obtain for the 5*s* state

$$\eta_{ns,\max}^{(2N+1)} \simeq 10.6\left(\frac{\Pi}{Z_i}\right)^2 N^{-8}.$$

The last expression shows that our entire analysis is applicable only when the harmonics generation efficiency is much less than 1.

In summary, an anomalously striking increase (by several orders of magnitude) in the harmonics generations efficiency under conditions of ionization of high-energy electronic states was established for coherent harmonic bremsstrahlung by plasma electrons produced by ionization with powerful pump radiation in the barrier-suppression ionization regime. It was shown how the increase in the harmonics generation efficiency in the electron–electron collision time becomes less pronounced. However, the effect remains until the electrons are heated by the pump radiation and the memory of their effective intraatom temperature (36) is lost.

ACKNOWLEDGMENTS

In conclusion, I thank A.M. Zheltikov whose dissertation motivated me to write this paper.

This work was supported by the Russian Foundation for Basic Research (program of leading scientific schools, project 99-02-18075) and by INTAS (grant 97-0369).

REFERENCES

1. S. M. Gladkov, N. I. Koroteev, M. V. Rychev, and O. O. Shtentsel', Pis'ma Zh. Éksp. Teor. Fiz. **43**, 227 (1986) [JETP Lett. **43**, 287 (1986)].
2. S. M. Gladkov, M. V. Rychev, and O. O. Shtentsel', Opt. Spektrosk. **61**, 68 (1986) [Opt. Spectrosc. **61**, 3 (1986)].
3. V. P. Silin, Zh. Éksp. Teor. Fiz. **47**, 2254 (1964) [Sov. Phys. JETP **20**, 1510 (1964)].
4. V. P. Silin, Zh. Éksp. Teor. Fiz. **114**, 864 (1998) [JETP **87**, 468 (1998)].
5. H. A. Bethe, *Handbuch der Physik B1, 24/1: Quantenmechanik der Ein- und Zwei-Electronenprobleme* (Berlin, 1933; ONTI, Moscow, 1935).
6. S. Augst, D. D. Meyerhofer, D. Strinkland, and S. I. Chin, J. Opt. Soc. Am. B **8**, 858 (1991).
7. M. V. Fedorov and J. Peatross, Phys. Rev. A **52**, 504 (1995).
8. V. P. Silin, Pis'ma Zh. Éksp. Teor. Fiz. **69**, 486 (1999) [JETP Lett. **69**, 521 (1999)].
9. L. D. Landau and E. M. Lifshitz, *Quantum Mechanics: Non-Relativistic Theory* (Nauka, Moscow, 1989, 4th ed.; Pergamon Press, Oxford, 1977, 3rd ed.).
10. *Handbook of Mathematical Functions*, Ed. by M. Abramowitz and I. A. Stegun (Dover, New York, 1971; Nauka, Moscow, 1979).
11. V. P. Silin, *Introduction to Kinetic Theory of Gases* (Fiz. Inst. Ross. Akad. Nauk, Moscow, 1998).
12. K. N. Ovchinnikov and V. P. Silin, Kr. Soobshch. Fiz. **10**, 19 (1998).
13. V. P. Silin, Kvantovaya Élektron. (Moscow) **26**, 11 (1999).

Translation was provided by AIP

Journal of Experimental and Theoretical Physics, Vol. 90, No. 5, 2000, pp. 810–816.
Translated from Zhurnal Éksperimental'noĭ i Teoreticheskoĭ Fiziki, Vol. 117, No. 5, 2000, pp. 932–938.

PLASMA, GASES

Induced Scattering and Two-Photon Absorption of Alfvén Waves with Arbitrary Propagation Angles

V. N. Zirakashvili

Institute of Terrestrial Magnetism, Ionosphere, and Radiowave Propagation, Russian Academy of Sciences, Troitsk, Moscow oblast, 142092 Russia
e-mail: zirak@izmiran.rssi.ru

Received December 23, 1999

Abstract—An equation for the spectral energy density of collisionless Alfvén waves, propagating at arbitrary angles to the average magnetic field, is derived on the basis of the theory of weak turbulence. The main nonlinear processes for the case studied are induced scattering and two-photon absorption of Alfvén waves by thermal ions. An equation is derived for thermal particles which describes particle diffusion, accompanying these processes, in momentum space. The results are qualitatively different from previous results obtained by other authors for Alfvén waves propagating along the average magnetic field. *© 2000 MAIK "Nauka/Interperiodica".*

1. INTRODUCTION

It is now known that Alfvén waves are an important component of cosmic plasma. It is believed that most of the energy of interplanetary and interstellar magnetohydrodynamic (MHD) turbulence is contained in Alfvén waves. The main reason probably is that two other types of low-frequency waves—ion-sound and magnetosonic—are subjected to strong linear Landau damping [1, 2]. For Alfvén waves this damping is quite weak, and nonlinear effects must be taken into account in order to investigate dissipation of MHD turbulence. The main nonlinear processes for Alfvén waves are induced scattering and two-photon absorption of waves by thermal ions. The corresponding damping decrements have been obtained by many authors [3–9]. Nonetheless, a complete analysis of the above-indicated effects has still not been performed. Induced scattering of Alfvén waves was first studied in [3], but for $\beta \ll 1$. Here β is the squared ratio of the thermal velocity of ions $v_{Ti} = \sqrt{T_i/m_i}$ (T_i is the ion temperature in energy units and m_i is the ion mass) to the Alfvén velocity $v_a = B_0/\sqrt{4\pi\rho_0}$ (B_0 is the average plasma density). The case where waves propagate along the magnetic field was studied in [4–7]. Finally, Alfvén waves propagating at an angle with respect to the average magnetic field were studied in [8, 9], but the effect of the random components of the magnetic field, which correspond to magnetosonic waves, which are second-order in the amplitude of the Alfvén waves, was neglected. It will be shown below that taking these components into account strongly changes the nonlinear damping decrements. Specifically, the interaction of waves for which the projection of the wave vector on the direction of the average magnetic field has the same sign is found to be impossible.

2. BASIC EQUATIONS

The kinetic approach is the most systematic method in nonlinear plasma theory. In the theory of weak turbulence the method entails an expansion of the distribution function of thermal particles in powers of the random field and calculation of the nonlinear currents, which are then substituted into Maxwell equations. This procedure applied to Alfvén waves leads to very complicated expressions containing multiple series with Bessel functions (see, for example, [3]). For magnetized thermal particles these functions must be expanded in powers of their small argument. These calculations are very complicated. We shall use a different method, which seems to us to be more attractive from the physical standpoint also. We shall employ for magnetized thermal particles the guiding center equation in the zero-gyroradius approximation. This means that we shall neglect all effects associated with the finite gyroradius of thermal particles. The system of equations employed for the fully ionized plasma is [10]

$$\rho\left(\frac{\partial \mathbf{u}}{\partial t} + (\mathbf{u}\cdot\nabla)\cdot\mathbf{u}\right) = -\frac{1}{4\pi}[\mathbf{B}\times[\nabla\times\mathbf{B}]] - \nabla P_\perp - (\nabla\cdot\mathbf{b})\cdot\mathbf{b}(P_\parallel - P_\perp), \quad (1)$$

$$\frac{\partial F_\alpha}{\partial t} + (\mathbf{u}_E + v_\parallel \mathbf{b})\cdot\nabla F_\alpha$$

1063-7761/00/9005-0810$20.00 © 2000 MAIK "Nauka/Interperiodica"

$$-\frac{v_\perp}{2}\frac{\partial F_\alpha}{\partial v_\perp}(v_\parallel \nabla\cdot\mathbf{b} + \nabla\cdot\mathbf{u}_E + \mathbf{u}_E\cdot(\mathbf{b}\cdot\nabla)\cdot\mathbf{b}) + \frac{\partial F_\alpha}{\partial v_\parallel}\left(\frac{q_\alpha}{m_\alpha}E_\parallel + \frac{v_\perp^2}{2}\nabla\cdot\mathbf{b} + \mathbf{u}_E\cdot\left(v_\parallel(\mathbf{b}\cdot\nabla)\mathbf{b} + (\mathbf{u}_E\cdot\nabla)\mathbf{b} + \frac{\partial\mathbf{b}}{\partial t}\right)\right) = 0, \quad (2)$$

$$\frac{\partial\mathbf{B}}{\partial t} = [\nabla\times[\mathbf{u}\times\mathbf{B}]], \quad (3)$$

$$\nabla\cdot\mathbf{B} = 0, \quad (4)$$

$$\sum_\alpha q_\alpha\int d^3v F_\alpha = 0. \quad (5)$$

Here $\mathbf{b} = \mathbf{B}/B$ is a unit vector in the direction of the magnetic field $\mathbf{B}$, $\mathbf{u}_E = c[\mathbf{E}\times\mathbf{B}]/B^2$ is the electric-drift velocity, $E_\parallel$ is the component of the electric field $\mathbf{E}$ in the direction of the magnetic field, ρ and $\mathbf{u}$ are, respectively, the density and velocity of the plasma, and $F_\alpha(v_\parallel, v_\perp)$ is the distribution function of thermal particles of type α. The parallel and perpendicular pressure in equation (1) are determined by the formulas

$$P_\parallel = \sum_\alpha\int d^3v m_\alpha(v_\parallel - \mathbf{u}\cdot\mathbf{b})^2 F_\alpha, \quad (6)$$

$$P_\perp = \sum_\alpha\int d^3v \frac{m_\alpha v_\perp^2}{2}F_\alpha. \quad (7)$$

The equation for the guiding center (2) describes the kinetic effects of thermal particles. For a frozen-in magnetic field, which is the case considered here, the electric drift velocity $\mathbf{u}_E$ is equal to the plasma velocity component $\mathbf{u}$ perpendicular to the magnetic field.

3. SERIES EXPANSION IN POWERS OF THE VELOCITY OF THE MEDIUM

We shall assume that the average magnetic field $\mathbf{B}_0$ is directed along the z axis. We write all quantities in the form $\mathbf{B} = \mathbf{B}_0 + \delta\mathbf{B}$, $\rho = \rho_0 + \delta\rho$, and so on, and we expand them in Fourier series

$$\delta\mathbf{B} = \sum\mathbf{B}_k\exp\{i(\mathbf{k}\cdot\mathbf{r} - \omega t)\},$$
$$\delta\rho = \sum\delta\rho_k\exp\{i(\mathbf{k}\cdot\mathbf{r} - \omega t)\},$$

where $k = (\mathbf{k}, \omega)$. All Fourier components should be expanded in powers of the plasma velocity $\mathbf{u}_k$ in the Alfvén waves. The plasma velocity in an Alfvén wave is directed perpendicular to the average field and the wave vector. The polarization of the Alfvén waves is linear. Since in the linear theory of Alfvén waves $\delta\rho_k = 0$, $E_{\parallel k} = 0$, $B_{zk} = 0$, and $u_{zk} = 0$, these quantities should be viewed as being of order higher than first. The expansion of equation (2) in powers of the velocity of the medium makes it possible to determine $\delta F_{\alpha k}$ in second order. Substituting this quantity into the equation of quasi-neutrality (5) makes it possible to find $E_{\parallel k}$. As a result, we have

$$\delta F_{\alpha k} = \frac{B_{zk}}{B_0}\left[\frac{k_z}{v_\parallel k_z - \omega}\left(\frac{v_\perp^2}{2} - \frac{q_\alpha}{m_\alpha}\sigma_1(k)\right)\frac{\partial F_{0\alpha}}{\partial v_\parallel} - \frac{v_\perp}{2}\frac{\partial F_{0\alpha}}{\partial v_\perp}\right] + \frac{1}{2}\sum_{k = k' + k''}(\mathbf{u}_{k'}\cdot\mathbf{u}_{k''})\left[-\frac{k_z' k_z''}{2\omega'\omega''}v_\perp\frac{\partial F_{0\alpha}}{\partial v_\perp} + \frac{k_z}{v_\parallel k_z - \omega}\frac{\partial F_{0\alpha}}{\partial v_\parallel}\left(\frac{v_\perp^2}{2}\frac{k_z' k_z''}{\omega'\omega''} + \frac{v_\parallel}{k_z}\left(\frac{k_z'^2}{\omega'} + \frac{k_z''^2}{\omega''}\right) - 1 + \frac{q_\alpha}{m_\alpha}\sigma_0(k', k'')\right)\right], \quad (8)$$

$$E_{\parallel k} = ik_z\left[\frac{B_{zk}}{B_0}\sigma_1(k) - \sum_{k = k' + k''}\frac{1}{2}\sigma_0(k', k'')(\mathbf{u}_{k'}\cdot\mathbf{u}_{k''})\right]. \quad (9)$$

Here and below, for conciseness, we drop the averages of the products $(\mathbf{u}_{k'}\cdot\mathbf{u}_{k''})$, which must be subtracted from the corresponding product. The quantities σ_0 and σ_1 in equations (8) and (9) can be expressed in terms of the average distribution function $F_{0\alpha}$ of the thermal particles:

$$\sigma_0(k', k'') = \left[\sum_\alpha\frac{q_\alpha^2}{m_\alpha}\int\frac{d^3v}{v_\parallel(k_z' + k_z'') - \omega' - \omega''}\frac{\partial F_{0\alpha}}{\partial v_\parallel}\right]^{-1} \times\sum_\alpha q_\alpha\int\frac{d^3v}{v_\parallel(k_z' + k_z'') - \omega' - \omega''}\frac{\partial F_{0\alpha}}{\partial v_\parallel} \times\left(1 - \frac{v_\perp^2}{2}\frac{k_z' k_z''}{\omega'\omega''} - \frac{v_\parallel}{k_z' + k_z''}\left(\frac{k_z'^2}{\omega'} + \frac{k_z''^2}{\omega''}\right)\right), \quad (10)$$

$$\sigma_1(k) = \left[\sum_\alpha\frac{q_\alpha^2}{m_\alpha}\int\frac{d^3v}{v_\parallel k_z - \omega}\frac{\partial F_{0\alpha}}{\partial v_\parallel}\right]^{-1} \times\sum_\alpha q_\alpha\int\frac{d^3v}{v_\parallel k_z - \omega}\frac{v_\perp^2}{2}\frac{\partial F_{0\alpha}}{\partial v_\parallel}. \quad (11)$$

In what follows, we shall require an expression for the z component of the random magnetic field B_{zk}. To find this expression it is sufficient to expand the Fourier transform of equations (1) and (3) up to second order

and multiply the first equation by $\mathbf{k}_\perp$. Finally, simple calculations give

$$\frac{B_{zk}}{B_0} = \sum_{k=k'+k''} \frac{(\mathbf{u}_{k''}\cdot\mathbf{k}')(\mathbf{u}_{k'}\cdot\mathbf{k}'')}{v_a^2\Delta(k)}\left(\frac{k_z'k_z''c_a^2}{\omega'\omega''}-1\right)$$

$$-\frac{1}{2}\sum_{k=k'+k''}(\mathbf{u}_{k'}\cdot\mathbf{u}_{k''})\frac{k_z'k_z''}{\omega'\omega''}$$

$$-\sum_{k=k'+k''}\frac{(\mathbf{u}_{k'}\cdot\mathbf{u}_{k''})}{2v_a^2\Delta(k)}\left[\frac{k_z'k_z''}{\omega'\omega''}(\omega^2-c_a^2k_z^2)\right. \quad (12)$$

$$+k_\perp^2\sum_\alpha\frac{m_\alpha}{\rho_0}\int\frac{d^3v k_z}{v_\parallel k_z-\omega}\frac{v_\perp^2}{2}\frac{\partial F_{0\alpha}}{\partial v_\parallel}$$

$$\left.\times\left(\frac{v_\parallel}{k_z}\left(\frac{k_z'^2}{\omega'}+\frac{k_z''^2}{\omega''}\right)-1+\frac{q_\alpha}{m_\alpha}\sigma_2(k',k'')\right)\right].$$

Here the quantity σ_2 is determined by the expression

$$\sigma_2(k',k'') = \left[\sum_\alpha\frac{q_a^2}{m_\alpha}\int\frac{d^3v}{v_\parallel(k_z'+k_z'')-\omega'-\omega''}\frac{\partial F_{0\alpha}}{\partial v_\parallel}\right]^{-1}$$

$$\times\sum_\alpha q_\alpha\int\frac{d^3v}{v_\parallel(k_z'+k_z'')-\omega'-\omega''}\frac{\partial F_{0\alpha}}{\partial v_\parallel} \quad (13)$$

$$\times\left(1-\frac{v_\parallel}{k_z'+k_z''}\left(\frac{k_z'^2}{\omega'}+\frac{k_z''^2}{\omega''}\right)\right),$$

and the denominator

$$\Delta(k) = k_\perp^2D_0(k)+(c_a^2k_z^2-\omega^2)/v_a^2 \quad (14)$$

corresponds to the dispersion equation of magnetosonic waves $\Delta(k)=0$. The velocity

$$c_a = \sqrt{v_a^2-(P_{\parallel 0}-P_{\perp 0})/\rho_0}$$

is the group velocity of Alfvén waves in a plasma with anisotropic pressure. Here we are considering a plasma which is stable with respect to the development of the firehose instability. In this case the expression in the radicand is positive. The quantity $D_0(k)$ is determined by the expression

$$D_0(k) = 1+\frac{8\pi P_{\perp 0}}{B_0^2}+\frac{4\pi}{B_0^2}\sum_\alpha m_\alpha\int d^3v\frac{k_z}{v_\parallel k_z-\omega} \quad (15)$$

$$\times\left(\frac{v_\perp^2}{2}-\frac{q_\alpha}{m_\alpha}\sigma_1(k)\right)\frac{v_\perp^2}{2}\frac{\partial F_{0\alpha}}{\partial v_\parallel}.$$

4. CALCULATION OF THE DAMPING DECREMENTS

The standard method for calculating damping decrements assumes that a dynamic equation is obtained for the waves. For this, equations (1)–(3) must be expanded up to third order in the amplitude of the waves. Then the equation is averaged, using the random-phase approximation [11]. Here we shall employ a simpler method, which gives the same result. We derive an equation for the thermal particles and determine the scattering probability and the probability of two-photon absorption. In practice, a quasilinear equation must be derived from equation (2). The difference lies only in the fact that in equation (2) quantities of second order in the field amplitude stand in front of the derivatives of the velocity distribution function. The result is

$$\frac{\partial F_{0\alpha}}{\partial t} = \frac{\partial}{\partial v_\parallel}\sum_k\left|\frac{B_{zk}}{B_0}\left(\frac{v_\perp^2}{2}-\frac{q_\alpha}{m_\alpha}\sigma_1(k)\right)\right.$$

$$+\frac{1}{2}\sum_{k=k'+k''}(\mathbf{u}_{k'}\cdot\mathbf{u}_{k''})\left(\frac{v_\perp^2}{2}\frac{k_z'k_z''}{\omega'\omega''}+\frac{v_\parallel}{k_z}\left(\frac{k_z'^2}{\omega'}+\frac{k_z''^2}{\omega''}\right)\right. \quad (16)$$

$$\left.\left.-1+\frac{q_\alpha}{m_\alpha}\sigma_0(k',k'')\right)\right|^2\frac{k_z^2}{i(v_\parallel k_z-\omega)}\frac{\partial F_{0\alpha}}{\partial v_\parallel}.$$

Using the expression (12) and performing the averaging, we arrive at the equation

$$\frac{\partial F_{0\alpha}}{\partial t} = \frac{\partial}{\partial v_\parallel}\sum_{\mathbf{k}',\mathbf{k}''}W(\mathbf{k}')W(\mathbf{k}'')$$

$$\times[(k_z'+k_z'')^2w(\mathbf{k}',\mathbf{k}'',\omega(\mathbf{k}'),\omega(\mathbf{k}'')) \quad (17)$$

$$+(k_z'-k_z'')^2w(\mathbf{k}',-\mathbf{k}'',\omega(\mathbf{k}'),-\omega(\mathbf{k}''))]\frac{\partial F_{0\alpha}}{\partial v_\parallel}.$$

Here, to perform the averaging we employed the relation

$$\langle\mathbf{u}_k\mathbf{u}_{k'}\rangle = \delta(k+k')\frac{1}{2\rho_0} \quad (18)$$

$$\times[W(\mathbf{k})\delta(\omega-\omega(\mathbf{k}))+W(-\mathbf{k})\delta(\omega+\omega(\mathbf{k}))],$$

where $W(\mathbf{k})$ is the spectral density of Alfvén waves with the dispersion relation $\omega(\mathbf{k})=c_a|k_z|$. The first term in brackets on the right-hand side of equation (17) corresponds to two-photon absorption, and the second term corresponds to induced scattering. The two-photon absorption probability $w(\mathbf{k}',\mathbf{k}'',\omega(\mathbf{k}'),\omega(\mathbf{k}''))$ is given by the expression

$$w(\mathbf{k}',\mathbf{k}'',\omega(\mathbf{k}'),\omega(\mathbf{k}'')) = \frac{\pi}{4\rho_0^2v_a^4}$$

$$\times \delta(\omega(\mathbf{k}') + \omega(\mathbf{k}'') - v_\parallel(k_z' + k_z''))\Bigg|\frac{1}{\Delta(k' + k'')}$$

$$\times\left(\frac{v_\perp^2}{2} - \frac{q_\alpha}{m_\alpha}\sigma_1(k' + k'')\right)$$

$$\times\Bigg[2k_\perp' k_\perp'' \sin^2\varphi(\mathbf{k}_\perp', \mathbf{k}_\perp'')\left(c_a^2\frac{k_z' k_z''}{\omega'\omega''} - 1\right)$$

$$- \cos\varphi(\mathbf{k}_\perp', \mathbf{k}_\perp'')\left\{\frac{k_z' k_z''}{\omega'\omega''}((\omega' + \omega'')^2 - c_a^2(k_z' + k_z'')^2)\right. \quad (19)$$

$$\left. + (\mathbf{k}_\perp' + \mathbf{k}_\perp'')^2\sum_\alpha \frac{m_\alpha}{\rho_0}\int \frac{d^3v(k_z' + k_z'')}{v_\parallel(k_z' + k_z'') - \omega' - \omega''}\frac{v_\perp^2}{2}\frac{\partial F_{0\alpha}}{\partial v_\parallel}\right.$$

$$\left.\times\left(\frac{v_\parallel}{k_z' + k_z''}\left(\frac{k_z'^2}{\omega'} + \frac{k_z''^2}{\omega''}\right) - 1 + \frac{q_\alpha}{m_\alpha}\sigma_2(k', k'')\right)\right\}\Bigg]$$

$$+ v_a^2\cos\varphi(\mathbf{k}_\perp', \mathbf{k}_\perp'')$$

$$\times\left(\frac{v_\parallel}{k_z' + k_z''}\left(\frac{k_z'^2}{\omega'} + \frac{k_z''^2}{\omega''}\right) - 1 + \frac{q_\alpha}{m_\alpha}\sigma_2(k', k'')\right)\Bigg|^2_{\substack{\omega' = \omega(\mathbf{k}') \\ \omega'' = \omega(\mathbf{k}'')}}.$$

Here $\varphi(\mathbf{k}_\perp', \mathbf{k}_\perp'')$ is the angle between the vectors $\mathbf{k}_\perp'$ and $\mathbf{k}_\perp''$. The evolution equation for the energy density of Alfvén waves has the form

$$\frac{\partial W(\mathbf{k})}{\partial t} = -2\Gamma(\mathbf{k})W(\mathbf{k}), \quad (20)$$

and the damping decrement $\Gamma(\mathbf{k})$ can be expressed in terms of the distribution function of thermal particles and the probability of two-photon absorption [11, 12]:

$$\Gamma(\mathbf{k}) = -\omega(\mathbf{k})\sum_{\alpha,\,\mathbf{k}'} m_\alpha W(\mathbf{k}')$$

$$\times\int d^3v[(k_z + k_z')w(\mathbf{k}, \mathbf{k}', \omega(\mathbf{k}), \omega(\mathbf{k}')) \quad (21)$$

$$+ (k_z - k_z')w(\mathbf{k}, -\mathbf{k}', \omega(\mathbf{k}), -\omega(\mathbf{k}'))]\frac{\partial F_{0\alpha}}{\partial v_\parallel}.$$

It is easy to see from the expression for the probability (19) that waves with identical signs of k_z do not interact. This is not so if $\mathbf{k}_\perp'$ and $\mathbf{k}_\perp''$ are formally set to zero in the expression (19). The same result is obtained if B_{zk} is set to zero in equation (16). This means that the propagation of waves strictly along the field is a special case (see the discussion below).

The expression (21) for the decrement with the probability (19) can be converted to the simpler form

$$\Gamma(\mathbf{k}) = \frac{1}{4}\omega(\mathbf{k})\mathrm{Im}\sum_{\mathbf{k}',\,\pm}\frac{W(\mathbf{k}')}{B_0^2/4\pi}\Bigg[\left(1 - \frac{1}{c_a^2}\left(\frac{\omega \pm \omega'}{k_z \pm k_z'}\right)^2\right)^2$$

$$\times D_1(k \pm k')\cos^2\varphi(\mathbf{k}_\perp, \mathbf{k}_\perp') + \frac{(\mathbf{k}_\perp \pm \mathbf{k}_\perp')^{-2}}{\Delta(k \pm k')}$$

$$\times\left\{2k_\perp k_\perp'\left(1 - c_a^2\frac{k_z k_z'}{\omega\omega'}\right)\sin^2\varphi(\mathbf{k}_\perp, \mathbf{k}_\perp') \right. \quad (22)$$

$$\left.\pm \cos\varphi(\mathbf{k}_\perp, \mathbf{k}_\perp')\left(1 - \frac{1}{c_a^2}\left(\frac{\omega \pm \omega'}{k_z \pm k_z'}\right)^2\right)\right.$$

$$\left.\times\left(D_2(k \pm k')(\mathbf{k}_\perp \pm \mathbf{k}_\perp')^2 - \frac{c_a^2 k_z k_z'}{\omega\omega'}(k_z \pm k_z')^2\right)\right\}^2\Bigg]_{\substack{\omega = \omega(\mathbf{k}) \\ \omega' = \omega(\mathbf{k}')}}.$$

Here the quantities $D_1(k)$ and $D_2(k)$ are expressed in terms of the distribution function of the thermal particles:

$$D_1(k) = -v_a^2\sum_\alpha \frac{m_\alpha}{\rho_0}\int\frac{d^3v k_z}{v_\parallel k_z - \omega}\left(1 - \frac{q_\alpha}{m_\alpha}\sigma_3(k)\right)\frac{\partial F_{0\alpha}}{\partial v_\parallel}, \quad (23)$$

$$D_2(k) = -\sum_\alpha \frac{m_\alpha}{\rho_0}\int\frac{d^3v k_z}{v_\parallel k_z - \omega}\left(1 - \frac{q_\alpha}{m_\alpha}\sigma_3(k)\right)\frac{v_\perp^2}{2}\frac{\partial F_{0\alpha}}{\partial v_\parallel}, \quad (24)$$

and $\sigma_3(k)$ is determined by the expression

$$\sigma_3(k) = \left[\sum_\alpha \frac{q_\alpha^2}{m_\alpha}\int\frac{d^3v}{v_\parallel k_z - \omega}\frac{\partial F_{0\alpha}}{\partial v_\parallel}\right]^{-1}$$

$$\times\sum_\alpha q_\alpha\int\frac{d^3v}{v_\parallel k_z - \omega}\frac{\partial F_{0\alpha}}{\partial v_\parallel}. \quad (25)$$

It should be noted that the expression (22) for the decrement can be obtained directly by averaging the dynamical equation for the waves, but it requires more laborious calculations. Both methods lead to the same result, since it is known that three-wave processes, in which the resonance particles do not participate, are impossible for Alfvén waves [1]. The plus sign in the expression (22) corresponds to two-photon absorption, and the minus sign corresponds to induced scattering.

5. DAMPING DECREMENTS FOR PLASMA WITH A MAXWELL DISTRIBUTION

Strictly speaking, the particle collision integral, which describes the approach of the system to a state of

thermodynamic equilibrium, must be added to the right-hand side of equation (17). If the particle collision frequency is sufficiently high, the particle distribution function can be assumed to be Maxwellian (in this case, the collision frequency must be low enough so that the waves can be assumed to be collisionless). For simplicity, we shall assume the plasma to be a purely hydrogen plasma. In this case the nonlinear damping decrement (22) can be expressed in terms of the function $J_+(x)$ [1]:

$$J_+(x) = -i\sqrt{\frac{\pi}{2}}x\exp\left(-\frac{x^2}{2}\right) + \frac{x}{\sqrt{2\pi}}\int_{-\infty}^{\infty}\frac{dx'}{x-x'}\exp\left(-\frac{x'^2}{2}\right). \quad (26)$$

The integral in this expression is to be understood in the principal-value sense. Ions make the main contribution to the damping [9]. Then the dependence of the quantities $D_0(k)$, $D_1(k)$, and $D_2(k)$ on the wave vector and frequency reduces to a dependence on the quantity $x = \omega/v_{Ti}|k_z|$:

$$D_0(x) = 1 + 2\beta J_+(x) + \beta\frac{J_+^2(x)}{1 + T_i/T_e - J_+(x)}, \quad (27)$$

$$D_1(x) = \beta^{-1}\frac{T_i}{T_e}\frac{1 - J_+(x)}{1 + T_i/T_e - J_+(x)}, \quad (28)$$

$$D_2(x) = \left(1 + \frac{T_i}{T_e}\right)\frac{1 - J_+(x)}{1 + T_i/T_e - J_+(x)}. \quad (29)$$

Here it was assumed that ions and electrons can have different temperatures T_i and T_e, since temperature equalization occurs more slowly than Maxwellization of the ion and electron distributions separately. Nonetheless, the expression (22) is still quite complicated. Two extreme cases will be examined below: small and large β.

(1) $\beta \ll 1$. In this case induced scattering, described by the first term, and two-photon absorption, described by the second term, make the main contribution to the decrement (22). The imaginary part of the second term arises as a result of the denominator $\Delta(k)$. It is obvious that this term corresponds to excitation of a magnetosonic wave, which is absorbed by the thermal particles. The first term describes induced scattering of Alfvén waves, which can be interpreted as excitation of an ion-sound wave, also absorbed by the plasma particles. In this case scattering is of a differential character in k_z—waves with close absolute values of k_z interact:

$$\Gamma(k_z, \mathbf{k}_\perp) = \frac{\omega(\mathbf{k})}{4}\sum_{\mathbf{k}'}\theta(-k_z k_z')\frac{W(\mathbf{k}')}{B_0^2/4\pi}\left[\frac{v_a^2}{v_{Ti}^2}\frac{T_i^2}{T_e^2}\right. \times \left.\frac{\cos^2\varphi(\mathbf{k}_\perp, \mathbf{k}_\perp')\sqrt{\frac{\pi}{2}}x\exp\left(\frac{-x^2}{2}\right)}{\left[1 - \mathrm{Re}\,J_+(x) + \frac{T_i}{T_e}\right]^2 + \frac{\pi}{2}x^2\exp(-x^2)}\right|_{x = \frac{\omega(\mathbf{k}) - \omega(\mathbf{k}')}{v_{Ti}|k_z - k_z'|}} + \delta(4k_z k_z' + (\mathbf{k}_\perp + \mathbf{k}_\perp')^2)\frac{\pi}{(\mathbf{k}_\perp + \mathbf{k}_\perp')^2} \times\left(4k_\perp k_\perp'\sin^2\varphi(\mathbf{k}_\perp, \mathbf{k}_\perp') - (\mathbf{k}_\perp + \mathbf{k}_\perp')^2\cos\varphi(\mathbf{k}_\perp, \mathbf{k}_\perp')\right)^2\Bigg]. \quad (30)$$

Expanding the spectral energy density of Alfvén waves near $-k_z$ and switching from summation to integration, we obtain

$$\Gamma(k_z, \mathbf{k}_\perp) = \frac{v_a k_z^2}{B_0^2/4\pi}\int d^2k_\perp' \times\left[-\alpha_1\left(W(-k_z, \mathbf{k}_\perp') + k_z\frac{\partial}{\partial k_z}W(-k_z, \mathbf{k}_\perp')\right)\cos^2\varphi(\mathbf{k}_\perp, \mathbf{k}_\perp') + W\left(-\frac{(\mathbf{k}_\perp + \mathbf{k}_\perp')^2}{4k_z}, \mathbf{k}_\perp'\right)\frac{\pi k_z^{-2}}{16(\mathbf{k}_\perp + \mathbf{k}_\perp')^2} \times\left(4k_\perp k_\perp'\sin^2\varphi(\mathbf{k}_\perp, \mathbf{k}_\perp') - (\mathbf{k}_\perp + \mathbf{k}_\perp')^2\cos\varphi(\mathbf{k}_\perp, \mathbf{k}_\perp')\right)^2\right], \quad (31)$$

where α_1 is determined by the integral

$$\alpha_1 = \sqrt{\frac{\pi}{2}}\frac{T_i^2}{T_e^2}\int_{-\infty}^{\infty}dx \times\frac{x^2\exp(-x^2/2)}{[1 - \mathrm{Re}\,J_+(x) + T_i/T_e]^2 + (\pi/2)x^2\exp(-x^2)} = \pi. \quad (32)$$

The first term in the expression (31) corresponds to the result obtained in [3]. However, the numerical factor α_1 differs from the corresponding coefficient in [3], which

depends on the ratio T_i/T_e. The coefficient from [3] is obtained by setting $x = 0$ in the denominator in the integrand in the expression (32). In addition, the coefficient (32) contains an additional factor of 2. The relation obtained here is more accurate. It should be noted that the expression (31) can be obtained on the basis of ordinary magnetohydrodynamics.

(2) $\beta \gg 1$. In this case two-photon absorption, which once again is of a differential character in k_z, makes the main contribution to the damping:

$$\Gamma(k_z, \mathbf{k}_\perp) = \frac{\omega(\mathbf{k})}{4}\frac{v_{Ti}^2}{v_a^2}\sum_{\mathbf{k}'}\theta(-k_z k_z')\frac{W(\mathbf{k}')}{B_0^2/4\pi}$$

$$\times \cos^2\varphi(\mathbf{k}_\perp, \mathbf{k}_\perp')\mathrm{Im}\left\{\frac{1 - J_+(x)}{1 + T_i/T_e - J_+(x)}x^4 \right. \quad (33)$$

$$\left. \times \left[\frac{T_i}{T_e} + \frac{(1 + T_i/T_e)^2[1 - J_+(x)]}{J_+(x)[2 + 2T_i/T_e - J_+(x)]}\right]\right|_{x = \frac{\omega(\mathbf{k}) + \omega(\mathbf{k}')}{v_{Ti}|k_z + k_z'|}}\Bigg\}.$$

For a smooth spectral density of Alfvén waves, we replace summation by integration and obtain the damping decrement

$$\Gamma(k_z, \mathbf{k}_\perp) = \alpha_2\frac{v_{Ti}k_z^2}{B_0^2/4\pi} \quad (34)$$

$$\times \int d^2k_\perp' W(-k_z, \mathbf{k}_\perp')\cos^2\varphi(\mathbf{k}_\perp, \mathbf{k}_\perp'),$$

and the coefficient α_2 can be expressed in terms of the following integral:

$$\alpha_2 = \int_0^\infty x^2 dx \mathrm{Im}\left\{\frac{1 - J_+(x)}{1 + T_i/T_e - J_+(x)}\right. \quad (35)$$

$$\left. \times \left[\frac{T_i}{T_e} + \frac{(1 + T_i/T_e)^2[1 - J_+(x)]}{J_+(x)[2 + 2T_i/T_e - J_+(x)]}\right]\right\}.$$

Numerical integration shows that this quantity is essentially independent of the ratio T_i/T_e. The value obtained for $\alpha_2(T_i = T_e)$ is 2.25.

6. DISCUSSION

The main result of this work is a general expression for the decrement (22) with the formulas (14), (23)–(25) for the quantities appearing in it. Some particular cases have been examined previously. The result obtained is different from the results obtained in [8, 9], where Alfvén waves propagating at an angle with respect to the average magnetic field were also considered. The method employed in the present work differs in that it takes into account the components of the magnetic field which correspond to magnetosonic waves. Since the magnetosonic waves decay strongly in plasma with $\beta \gtrsim 1$ [2] (quasi-longitudinal and quasi-transverse propagation are exceptions), these quantities should be expressed in terms of the components of the velocity of Alfvén waves. Actually, we assume that they are second-order quantities, just as the longitudinal electric field $E_{\|k}$. This approach seems to us to be more consistent than the method employed in [8, 9], where the magnetic field components corresponding to magnetosonic waves were completely neglected. To reproduce the results of [8, 9] it is sufficient to neglect B_{zk} in equation (16). In the opinion of the author this is correct only for propagation of waves predominantly along the average magnetic field. In practice, the results of [8, 9] are the same as the results obtained in [4–7] for this case. The difference lies only in the fact that in a more rigorous kinetic analysis the waves propagating at small angles with respect to the field ($k_\perp/k_z \ll \sqrt{\omega/\Omega}$, Ω is the gyrofrequency of ions) are circularly polarized [1]. In the present work, we neglect this small region in the wavevector space. To obtain the condition of applicability of the small-angle approximation, when the results of [8, 9] are valid, it is sufficient to use an expression for the probability (19) and set $k_\perp'$ and $k_\perp''$ to zero. Then, one can see that everything depends on the ratio of the squares of these quantities and on $(\omega' + \omega'')^2/v_a^2 - (k_z' + k_z'')^2$, which are contained in $\Delta(k)$. For waves with the same signs of k_z, the next to last expression is zero, if the dispersion relation for Alfvén waves is used. In reality, there exists another nonlinear correction to the frequency, which is of the same order of magnitude as the nonlinear decrement of the waves. Finally, we obtain the condition of applicability of the method used in the present work:

$$\langle \delta B^2\rangle/B_0^2 \ll k_\perp^2/k_z^2. \quad (36)$$

Here

$$\langle \delta B^2\rangle = 4\pi\int d^3k W(\mathbf{k})$$

is the average squared random magnetic field of Alfvén waves, and k_z and $k_\perp$ are the values of the components of the wave vector of Alfvén waves, which contain most of the turbulence energy. The condition (36) signifies that most of the energy is not concentrated in waves propagating at small angles with respect to the average field. For the opposite relation, the results obtained in [8, 9] are correct. Specifically, waves with the same signs of k_z can interact.

Analysis of this problem shows that the above-mentioned effect arises as a result of terms containing $\nabla \cdot \mathbf{b} = (\mathbf{B} \cdot \nabla)B^{-1}$ in equation (2) and describing magnetic reflection of particles. If two Alfvén waves propagate at an angle with respect to the average field, then their magnetic pressure gives rise to additional plasma motions and second-order components of the magnetic field. For waves with the same signs of k_z, the compo-

nent B_{zk} arising has precisely the value required to compensate the change in the absolute value of the magnetic field (see the second term in the expression (12)), and as a result magnetic reflection vanishes and these waves do not interact. However, if the waves propagate along the average field, such compensation is impossible, since in this case $B_{zk} = 0$.

The second term in the expressions (22) and (31) corresponds to excitation of a magnetosonic wave. If the damping of magnetosonic waves is quite weak, their amplitude can become comparable to the amplitude of Alfvén waves, which does not correspond to the method employed. In plasma with $\beta \gtrsim 1$ the magnetosonic waves decay weakly for propagation at small angles with respect to the field. The linear damping decrement is determined by thermal ions and has the form [2]

$$\gamma = \sqrt{\frac{\pi}{2}} v_{Ti} \frac{k_\perp^2}{|k_z|}. \quad (37)$$

Comparing this decrement with the nonlinear decrement (34), we find that the condition of applicability of the method employed is the same as the condition (36). For $\beta \ll 1$ this approach is valid for even lower energy density of Alfvén waves. For

$$\frac{m_e T_i}{m_i T_e} \ll \beta \ll 1 \quad (37a)$$

the linear damping decrement of magnetosonic waves is determined by Landau damping on thermal electrons [1]:

$$\gamma = \sqrt{\frac{\pi}{8}} v_{Te} \frac{k_\perp^2}{|k_z|} \frac{m_e}{m_i}. \quad (38)$$

Comparing this decrement with the second term in the expression (31), we find that the condition of applicability of our method for the case (37a) is

$$\frac{\langle \delta B^2 \rangle}{B_0^2} \ll \sqrt{\frac{m_e T_e}{m_i T_i}} \beta \frac{k_\perp^2}{k_z^2}. \quad (39)$$

For the opposite relation the linear damping of magnetosonic waves is small compared with the nonlinear effects. Then, not only scattering processes [3] but also three-wave processes [1] with the participation of Alfvén and magnetosonic waves must be taken into account. Thus, it can be concluded that the approach considered in this paper is valid in plasma with $\beta \gtrsim 1$ when the condition (36) for the energy of Alfvén turbulence is satisfied. For $m_e T_i / m_i T_e \lesssim \beta \ll 1$ the approach is valid if the stronger condition (39) is satisfied.

ACKNOWLEDGMENTS

This work was supported by the Russian Foundation for Basic Research, the Federal program "Astronomiya," and the program of the Russian Academy of Sciences in support of young scientists.

REFERENCES

1. A. I. Akhiezer, I. A. Akhiezer, R. V. Polovin, A. G. Sitenko, and K. N. Stepanov, *Plasma Electrodynamics* (Nauka, Moscow, 1974; Pergamon Press, Oxford, 1975).
2. K. N. Stepanov, Zh. Éksp. Teor. Fiz. **34**, 1292 (1958) [Sov. Phys. JETP **7**, 892 (1958)].
3. M. A. Livshits and V. N. Tsytovich, Nucl. Fusion **10**, 241 (1970).
4. M. A. Lee and H. J. Völk, Astrophys. Space Sci. **24**, 31 (1973).
5. A. Achterberg, Astron. Astrophys. **98**, 161 (1981).
6. R. M. Kulsrud, Phys. Scr. **2** (1), 177 (1982).
7. J. A. Miller, Astrophys. J. **376**, 342 (1991).
8. A. Achterberg and R. D. Blandford, Mon. Not. R. Astron. Soc. **218**, 551 (1986).
9. V. N. Fedorenko, V. M. Ostryakov, A. N. Polyudov, and V. D. Shapiro, Fiz. Plazmy **16**, 443 (1990) [Sov. J. Plasma Phys. **16**, 253 (1990)].
10. M. N. Rosenbluth and N. Rostoker, Phys. Fluids **2**, 23 (1958).
11. V. N. Tsytovich, *Theory of Turbulent Plasma* (Atomizdat, Moscow, 1971; Plenum Press, New York, 1974).
12. V. N. Tsytovich and A. B. Shvartsburg, Zh. Éksp. Teor. Fiz. **49**, 797 (1965) [Sov. Phys. JETP **22**, 554 (1966)].

Translation was provided by AIP

Journal of Experimental and Theoretical Physics, Vol. 90, No. 5, 2000, pp. 817–822.
Translated from Zhurnal Éksperimental'noĭ i Teoreticheskoĭ Fiziki, Vol. 117, No. 5, 2000, pp. 939–946.

PLASMA, GASES

Structures of the Particles of the Condensed Dispersed Phase in Solid Fuel Combustion Products Plasma

A. A. Samaryan*, **, A. V. Chernyshev*, A. P. Nefedov*, O. F. Petrov*, Yu. M. Mikhaĭlov*, V. B. Mintsev***, and V. E. Fortov***

**Research Center for the Thermal Physics of Pulse Effects, Russian Academy of Sciences, Moscow, 127412 Russia*

***e-mail: ipdustpl@redline.ru*

****Institute of Chemical Physics, Russian Academy of Sciences, Chernogolovka, Moscow oblast, 142432 Russia*

Received January 17, 2000

Abstract—The results of experimental investigations of a type of dusty plasma which has been least studied—the plasma of solid fuel combustion products—were presented. Experiments to determine the parameters of the plasma of the combustion products of synthetic solid fuels with various compositions together with simultaneous diagnostics of the degree of ordering of the structures of the particles of the dispersed condensed phase were performed. The measurements showed that the charge composition of the plasma of the solid fuels combustion products depends strongly on the easily ionized alkali-metal impurities which are always present in synthetic fuel in one or another amount. An ordered arrangement of the particles of a condensed dispersed phase in structures that form in a boundary region between the high-temperature and condensation zones was observed for samples of aluminum-coated solid fuels with a low content of alkali-metal impurities. © *2000 MAIK "Nauka/Interperiodica".*

1. INTRODUCTION

The plasma of the combustion products of solid fuels is one of the least studied types of dusty plasma. A dusty plasma, or a plasma with macroscopic particles (such a plasma is often said to be aerosol, heterogeneous plasma or a plasma with a condensed dispersed phase), is characterized by the fact that the particles (which can vary in size from the hundredths of a micron up to several tens of microns) effectively interact with electrons and ions and therefore strongly influence the properties of the plasma.

Gas-discharge and thermal plasmas have been studied in greatest detail thus far. A thermal dusty plasma is formed from a heated neutral gas (1700–3000 K) at atmospheric pressure. The temperatures of the electrons, ions, and neutral particles are equal to one another, and the electron density lies in the range 10^9–10^{12} cm^{-3}. When particles are introduced into or appear in the plasma, for example, as a result of condensation, they can become charged as a result of electron and ion fluxes as well as thermal emission of electrons. Electron emission from particle surfaces can result in a positive electric charge, and the particles emitting electrons can increase the electron density in and the electric conductivity of the gas phase. In the limiting case of a nonionized gas, the presence of macroparticles completely determines the electrophysical properties of plasma. The conditions for the existence of a plasma with macroparticles can vary substantially. On account of the large charges which the particles can acquire (of the order of 10^2–10^3 elementary charges), the entire range of plasma states from a Debye plasma up to a strongly nonideal system of charged particles, depending on the density and sizes of the particles and the electron work function, as well as the densities of the electrons and ions present can be realized in plasma under typical conditions (T_g = 1700–3000 K, $n_e = 10^9$–10^{12} cm^{-3}).

The increased interest shown in the last few years in the properties of dusty plasma is due to practical problems (dusty plasma is a working object in a number of fields in technology, including power engineering, rocket building, plasma sputtering, and fusion) and fundamental studies in the field of nonideal-plasma physics and the electrophysics and electrodynamics of the combustion products of solid fuel. The discovery of plasma-crystalline structures [1–4] has stimulated further investigations of dusty plasma. A large part of these studies concerns plasma-dust formations in various types of discharges.

The thermodynamic properties of dusty plasma are largely determined by the interparticle interaction parameter, which is equal to the ratio of the interaction potential energy to the average thermal kinetic energy of the particles. Particle interactions are customarily described using either the single-component plasma model or a model with a screened (Debye) potential. The latter model is also known as the Yukawa model.

1063-7761/00/9005-0817$20.00

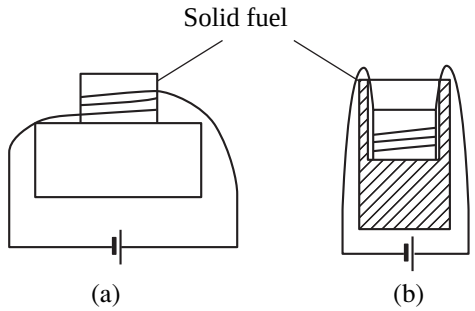


Fig. 1. Methods for securing the fuel samples: pellet on a stand (a); pellet in a holder (b).

In the single-component plasma model for a classical, quasi-neutral, spatially unbounded plasma the interaction parameter has the form

$$\gamma = \frac{(Ze)^2}{\langle r\rangle kT_g}, \quad (1)$$

where Ze is the charge of a macroparticle, T_g is the plasma temperature, $\langle r\rangle = (4\pi n_p/3)^{-1/3}$ is the average interparticle distance, and n_p is the particle density.

Screening of the charges of the macroparticles by the electrons and ions in the exterior plasma is taken into account in the Yukawa model, where the parameter

$$\Gamma = \frac{Z^2e^2}{\langle r\rangle kT_g}\exp\left(-\frac{\langle r\rangle}{\lambda_D}\right), \quad (2)$$

where λ_D is the Debye radius, is used to analyze the type of dust structure.

It is known from the single-component plasma model that short-range order appears in the system for $\gamma > 4$, and a single-component plasma crystallizes for $\gamma \simeq 170$ [5]. If screening effects play a large role in a dusty plasma, then two parameters—γ and $k = \langle r\rangle/\lambda_D$—are used to describe the structural transformation. Nonetheless, in a number of works [6–8] a single parameter, determined by the expression (2) and satisfactorily describing the properties of plasma for small values of Γ, is used to describe such a system. Short-range order is established in the system for $\Gamma > 1$ [9, 10].

Few experiments to study ordered structures in thermal plasma have been performed. In the early works the spatial distribution of the particles in the dispersed phase was investigated using electron microscopy of samples extracted by a probe introduced for a certain period of time into a plasma flow. Such photomicrographs were obtained in [11] in the flame of a synthetic aluminum-coated fuel. The arrangement of the particles in the photographs attested to the presence of an ordered structure of a polydispersed system of particles on the surface of the sampling probe. On this basis the authors concluded that structures of particles are present in the plasma.

In [12–14] the formation of ordered structures of macroparticles was observed in a weakly ionized, laminar, plasma flow at atmospheric pressure and temperature near 1700 K. The plasma consisted of positively charged micron-size CeO_2 particles, electrons emitted by the particles, and singly charged sodium ions. In these experiments the dusty plasma was produced by introducing particles into the inner flame of a two-flame Mekker-type burner; this gave a narrow range of variation of the plasma parameters and a relatively short plasma lifetime (of the order of 10 ms). As a result, crystallization could not be observed for a system of particles of a condensed dispersed phase in a propane-air flame.

The study of ordered structures of charged macroparticles in a plasma of the combustion products of synthetic solid fuels opens up greater possibilities. This is because the range of values of the plasma parameters is larger and the sizes and density of the particles of the condensed dispersed phase can be varied over a wide range. Our objective in the present work was to perform experimental investigations of the formation of ordered structures of particles of a condensed dispersed phase in the plasma of combustion products of synthetic solid fuel.

2. EXPERIMENTAL SETUP

The investigations were performed with samples of synthetic magnesium- and aluminum-coated solid fuel. The fuel samples were cylindrical with various lengths and diameters. Pellets whose size depended on the burn rate of the fuel and which were chosen so as to give an adequate time for performing the measurements (20–40 s) were fabricated to perform the measurements. Heaters consisting of nichrome wire (200 μm in diameter, 6 A heating current) were used for controlled and safe ignition of the fuel. The experimental procedure was as follows. Synchronized measurements of the main plasma parameters were started at the moment the fuel was ignited and structural diagnostics of the system of particles of the condensed dispersed phase was performed. The samples were secured using the schemes displayed in Fig. 1.

A diagnostics complex (see Fig. 2) which made it possible to determine the following parameters of the plasma formed by the combustion of solid fuel was used for the measurements: the temperature T_g of the gas phase, the sizes D_{32} and density n_p of the particles of the dispersed phase, the density n_a of the alkali-metal atoms, and the ion density n_i. A modified inversion method [17] was used to determine the temperature T_g in the high-temperature zone and a thermocouple was used to determine the temperature in the condensation zone thermocouple. The aperture transmittance [16] and spectral transmittance [15] methods were used to determine the particle sizes D_{32} and density n_p, respec-

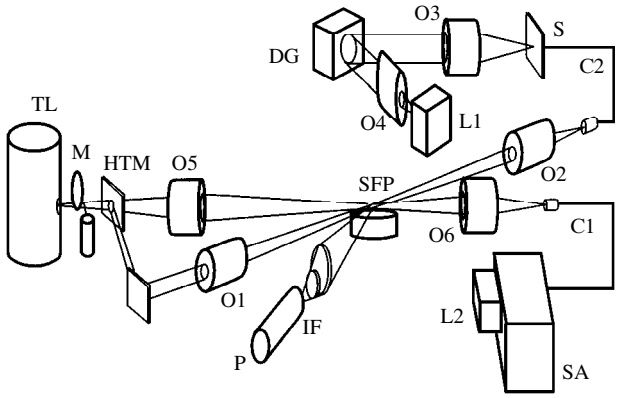


Fig. 2. Experimental setup: TL—tungsten reference lamp; M—modulator; HTM—half-transmitting mirror; O1—lens (F = 150 mm); O2–O6—objectives; C1—single-fiber light guide; C2—multifiber light guide; S—slit diaphragm; DG—diffraction grating; SA—spectral apparatus; IF—interference filter; P—photomultiplier; SFP—solid fuel pellet; L1, L2—CCD arrays.

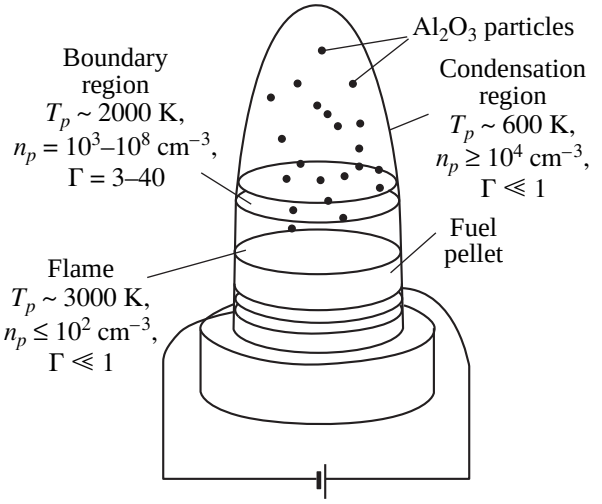


Fig. 3. Diagram of the main zones of the solid fuel combustion products plasma.

tively. The total absorption method was used to determine the density of alkali-metal atoms [17].

A single, continuously rotating, probe with a constant voltage $V_0 = -15$ V was used to measure the ion density n_i. As the probe passes through the plasma, an ion current flows onto the probe. A current pulse with a duration of about 2 ms corresponds to the n_i profile, and the current at the center of the pulse corresponds to the ion density in the measuring volume. Next, we determined n_i in the combustion-products plasma using calibration functions of the form

$$f_0(I_0) = I_0/n_{i0},$$

where I_0 is the current measured in a plasma seeded with an alkali metal with potential V_0 and n_{i0} is the ion density determined according to the saturation current.

The particle visualization method was used to observe and analyze the structures of the particles of the condensed dispersed phase in the solid fuel combustion products plasma. Illumination in a horizontal or vertical plane with a probe laser beam was used to visualize the particles. A cylindrical lens formed the argon-ion laser beam into a flat converging beam with a 100 μm thick and 3 mm wide waist in the combustion region. The horizontal probe beam could be moved over the height of the combustion zone, and the vertical probe beam could be moved over the height and radius. The light scattered by the particles was observed with a CDD camera at angles of 70° (in the case of a horizontal beam) and 90° (in the case of a vertical beam) with respect to the plane of the waist. The output signal from the video camera was recorded on a video tape recorder. The video images of the particles were digitally processed using a specially developed program and a binary correlation function $g(r)$ was constructed. Analysis of $g(r)$ makes it possible to describe the spatial structure and interparticle correlation of the particles.

3. MEASUREMENTS AND ANALYSIS OF THE RESULTS

The measurements were performed using nine fuel mixtures with different compositions (see Table 1). The aluminum-coated fuel produces a 10–30 mm high brownish-colored flame. A distinct condensation zone can be seen above the flame. The magnesium-coated fuel gives a strongly nonuniform combustion zone and numerous sparks. For the aluminum-coated fuel, the measurements were performed in characteristic regions of the plasma flame: in the high-temperature zone (in the flame), in a boundary region, and in the condensation zone (see Fig. 3); for the magnesium-coated fuel the measurements were performed at different heights of

Table 1. Composition of solid fuels

Sample no.	Al (Mg for samples 1 and 2)		NU, %	Binders, %
	%	D, μm		
1	30	1–30		
2	20	100		
3	40	30	20	40
4	10	100	50	P-2200-36, TMDN-4
5	20	30	40	P-2200-32, TMDN-6
6	10	1–10	40	50
7	10	1–30	40	50
8	20	30–90	36	PI-44
9	10	40	40	50

Table 2. Results of experimental measurements

Sample no.	D, μm	n_p, cm^{-3}	T, K	n_a, cm^{-3}	n_i, cm^{-3}
3	1.6	3.5×10^5	2105	2.72×10^{12}	2.54×10^{10}
4	0.3–0.5	1.8×10^5	2095	9.27×10^{11}	1.33×10^{10}
5	<0.2	8×10^4	2100	8.2×10^{10}	6.2×10^9
6	<0.2	1.2×10^5	2125	7.20×10^{11}	1.20×10^{10}
7	<0.2	2.5×10^6	1950	3.0×10^{10}	1.2×10^9
8	0.4	6.8×10^5	2090	4.5×10^{10}	3.4×10^9
9	0.5	8×10^5	2085	6.1×10^{10}	4.7×10^9

the plasma flow. The nonstationary nature of the combustion and the strong nonuniformity of the plasma region made it impossible to perform an accurate measurement of the parameters of the combustion-products plasma and structural diagnostics for magnesium-coated fuel. For this reason, experimental data are presented below only for aluminum-coated fuel with different compositions (samples 3–9 in Table 1).

The experimental measurements showed that the charge composition of the solid fuel combustion-products plasma depends strongly on the easily ionized alkali-metal impurities (Na and K), which are always present in one or another amount in the synthetic fuel and end up in the combustion products. The density n_a of alkali-metal atoms, the gas temperature $T \simeq T_p$, and the work function W_{te} of the thermal electrons from the surfaces of the particles determine the electrophysical properties of the thermal dusty plasma and have a large effect on the magnitude and sign of the charge of the

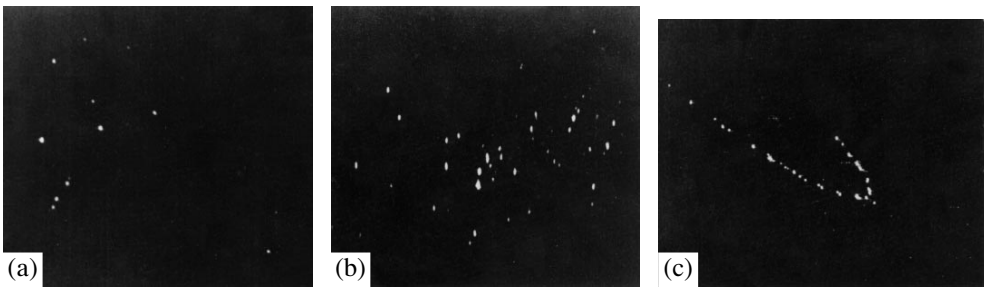


Fig. 4. Video images of particles in the high-temperature zone: fuel no. 7, T = 1950 K (a); fuel no. 8, T = 2090 K (b); fuel no. 9, T = 2085 K (c).

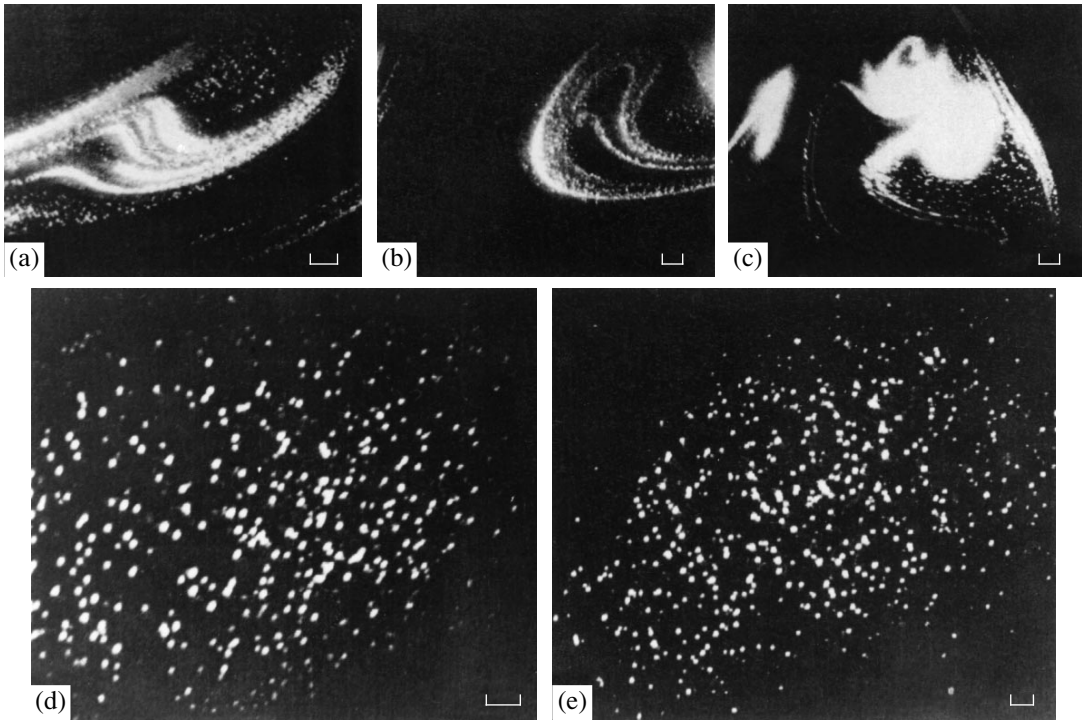


Fig. 5. Video images of particle structures in the condensation zone: fuel No. 7, h = 45 mm (a); fuel no. 8, h = 45 mm (b); fuel no. 9, h = 45 mm (c); fuel no. 7, h = 35 mm (d); fuel no. 9, h = 35 mm (e) (h—height above the solid-fuel pellet). The scale for images (a) (b) and (c) corresponds to 500 mm and for images (d) and (e) the scale is 100 mm.

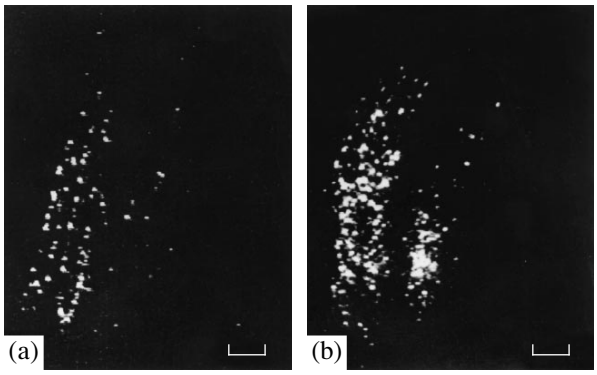


Fig. 6. Video images of particle structures in the boundary region of the condensation zone: fuel no. 7 (the scale indicated corresponds to 100 mm) (a), fuel no. 9 (the scale indicated corresponds to 300 mm) (b).

dust particles and therefore also on the formation conditions of ordered structures.

We shall employ the results of [10] and the results of diagnostics measurements (see Table 2) in order to analyze the value of the parameter Γ in a combustion-products plasma consisting of electrons, ions, and micron-size particles of a condensed dispersed phase. It is known that for given sizes and density of the particles of the condensed dispersed phase the value of the parameter Γ is determined by the screening of the particles by the plasma component formed as a result of the ionization of the alkali-metal impurity atoms. Numerical analysis shows that for the characteristic sizes of the particles of the condensed dispersed phase, which form in the combustion products of the solid fuel (see Table 2), structure formation requires that

$$\eta = \frac{n_a}{n_p} \ll 10^5. \quad (3)$$

For this reason, a solid fuel with small, natural alkali-metal impurities (mass fraction < 0.01%) must be used in order to observed ordered structures. This gives a density of easily ionized atoms in the combustion-products plasma not exceeding 10^9 cm^{-3} with the density of the particles of the condensed dispersed phase in the high-temperature zone greater than 10^4 cm^{-3}. The data in Table 2 show that in experiments with samples 3–6 $\eta \approx 10^6$–10^7, and the condition (3) is not satisfied, i.e., the particles were strongly screened and could form ordered structures. An estimate of the parameter Γ for the combustion products of fuel no. 3 gives a value of the order of 0.1, which shows that interparticle correlation is impossible in this system.

Figures 4–7 show video images of the particles obtained in experiments with fuels nos. 7–9. For these samples the parameter η was of the order of 10^4–10^5. For this reason it was of greatest interest to investigate these fuel samples. Figures 4 and 5 display typical video images of the particles in a flame and in the condensation zone of the combustion products. The extremely low densities (<10^2 cm^{-3}, which corresponds to an optical density $\tau < 0.01$) of the particles of the dis-

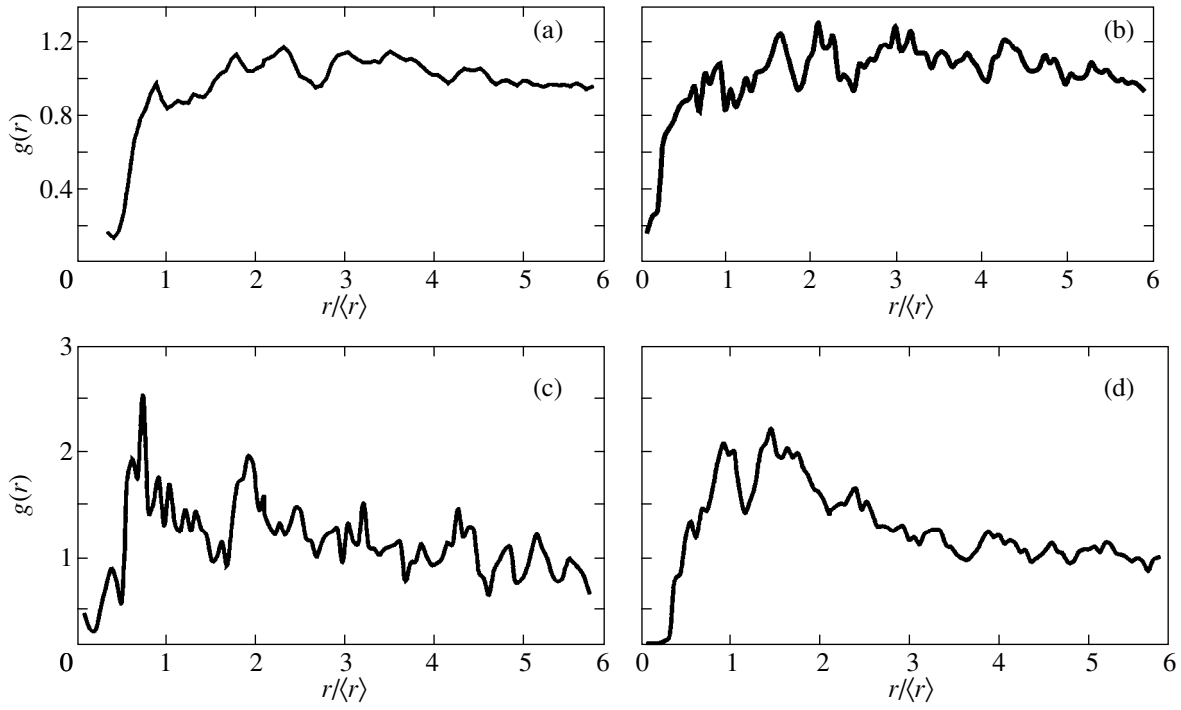


Fig. 7. Binary correlation functions $g(r)$ for the particle structures shown in Fig. 5d (a), Fig. 5e (b), Fig. 6a (c), and Fig. 6b (d).

persed phase in the flame are interesting. According to equation (2), under these conditions Γ will be close to zero and it is difficult to expect formation of ordered structures in this region. Chain structures can often be observed in the video images (see Fig. 4b), but the nature of these formations is still unclear. The particle density is much higher in the condensation zone (10^4 cm^{-3}), but the temperature of the medium is low—about 600 K ($\tau > 0.1$, $D \sim 0.4$ μm)—and therefore $\Gamma \ll 1$. Dust formations with different shapes, observed in the condensation zone, can be seen in Fig. 5. Analysis of the correlation functions did not show short-range order in the particle arrangement (Figs. 7a, 7b). Gas-dynamic processes probably determine the form and arrangement of the observed formations of particles of the condensed dispersed phase.

Analysis of small ($V < 10^{-4}$ cm^{-3}) structures of particles of the condensed dispersed phase, which form in the boundary region where the temperature is still high ($T \sim 1500$ K) and regions with high particle density appear (see Fig. 6), shows a different picture. Short-range order (see Figs. 7c, 7d) is observed in these structures when the particle density is sufficiently high ($n_p \sim 10^7$–10^8 cm^{-3}). The degree of correlation in the arrangement of the particles depends on the value of the parameter η. The highest correlation was observed in experiments with fuels nos. 7 and 9.

4. CONCLUSIONS

Experiments to determine the parameters of the plasma of combustion products of synthetic solid fuel with different compositions were performed. In most of the experiments the parameters of the solid fuel combustion-products plasma were such that the interparticle interaction parameter Γ was much less than 1. Ordered structures of particles of the condensed dispersed phase were not observed. The main obstacle for the formation of ordered structures was a large amount of alkali-metal impurities in the fuel samples and, in consequence, high electron and ion densities in the plasma. A low density of particles of the condensed dispersed phase in the high-temperature region was also an inhibiting factor. Ordered structures of particles of the condensed dispersed phase were also not observed in the condensation zone, where the particle density is quite high but the charge on the particles is low on account of the relatively low temperature of the medium.

An ordered arrangement of particles of the condensed dispersed phase in structures formed in the boundary region between the flame and the condensation zone was observed for samples of aluminum-coated solid fuel with a low alkali-metal impurity content. An estimate of the parameter Γ based on diagnostics measurements gives a value from 10 to 30 for various fuels. This corresponds to a "plasma liquid" type structure.

Our experiments show that formation of ordered structures of charged particles of a condensed dispersed phase is possible in a solid fuel combustion-products plasma. The main difficulties in performing further investigations on this object are a strong nonuniformity of the plasma parameters and the short lifetime of the structures formed. To overcome these difficulties, the scheme of the combustion of the solid fuel must be substantially updated and the existing diagnostics methods must be modified.

ACKNOWLEDGMENTS

This work was partially supported by the Russian Foundation for Basic Research (projects no. 98-02-16828 and no. 98-02-16825). The work of one of us (A. S.) was supported by INTAS YSF grant no. 99-214.

REFERENCES

1. J. H. Chu and Lin I, Phys. Rev. Lett. **72**, 4009 (1994).
2. H. Thomas, G. E. Morfill, V. Demmel, *et al.*, Phys. Rev. Lett. **73**, 652 (1994).
3. Y. Hayashi and K. Tachibana, Jpn. J. Appl. Phys. **33**, L804 (1994).
4. A. Melzer, T. Trottenberg, and A. Piel, Phys. Lett. A **191**, 301 (1994).
5. S. Ichimaru, Rev. Mod. Phys. **54**, 1017 (1982).
6. H. Ikezi, Phys. Fluids **29**, 1764 (1986).
7. A. Melzer, A. Homann, and A. Piel, Phys. Rev. E **53**, 2757 (1996).
8. M. A. Smith, H. U. Rahman, and U. Mohideen, in *Advances in Dusty Plasmas*, Ed. by P. K. Shukla, D. A. Mendis, and T. Desai (World Scientific, Singapore, 1997), p. 205.
9. V. E. Fortov and I. T. Yakubov, in *Nonideal Plasma* (Énergoatomizdat, Moscow, 1994), p. 302.
10. A. A. Samaryan, O. S. Vaulina, A. P. Nefedov, and O. F. Petrov, Fiz. Plazmy (in press) [Plasma Phys. Rep. (in press)].
11. G. S. Dragan, A. A. Mal'gota, and S. K. Protas, in *Proceedings of Scientific Conference on Methods of Fuel Preparation and Construction of Combustion Chambers for MDH Power Stations Operating on Coal* (Akad. Nauk SSSR, Moscow, 1984), p. 191.
12. V. E. Fortov, A. P. Nefedov, O. F. Petrov, *et al.*, Pis'ma Zh. Éksp. Teor. Fiz. **63**, 163 (1996) [JETP Lett. **63**, 187 (1996)].
13. V. E. Fortov, A. P. Nefedov, O. F. Petrov, *et al.*, Phys. Lett. A **219**, 89 (1996).
14. V. E. Fortov, A. P. Nefedov, O. F. Petrov, *et al.*, Phys. Rev. E **54**, R2236 (1996).
15. O. S. Vaulina, A. A. Samaryan, A. V. Chernyshev, *et al.*, Fiz. Plazmy **25**, 191 (1999) [Plasma Phys. Rep. **25**, 170 (1999)].
16. A. P. Nefedov, O. F. Petrov, and O. S. Vaulina, Appl. Opt. **36**, 1357 (1997).
17. A. A. Samaryan, A. V. Chernyshev, A. P. Nefedov, and O. F. Petrov, Fiz. Plazmy **25**, 988 (1999) [Plasma Phys. Rep. **25**, 912 (1999)].

Translation was provided by AIP

Journal of Experimental and Theoretical Physics, Vol. 90, No. 5, 2000, pp. 823–825.
Translated from Zhurnal Éksperimental'noĭ i Teoreticheskoĭ Fiziki, Vol. 117, No. 5, 2000, pp. 947–950.

SOLIDS
Structure

Ultrasoft X-ray Spectroscopy Investigation of the Model System Si–SiO_2

S. K. Kozhakhmetov
Karaganda State University, Karaganda, 470074 Kazakhstan
e-mail: kserik@kargu.krg.kz
Received May 12, 1999

Abstract—SiO_2 surface films with different thicknesses (ranging from 20 to 630 Å), grown on a crystal silicon substrate, have been investigated by the method of reflection and scattering of ultrashoft X-rays. It is shown on the basis of a simultaneous analysis of the Si$L_{2,3}$ reflection spectra and the scattering indicatrix that the critical angle θ_c for total external reflection for SiO_2 at $\lambda = 57$ Å lies in the range 4.5°–5°. The angular dependence of the thickness of the surface layer that forms the specular reflection is obtained. It is shown that the surface layer, whose thickness corresponds to the penetration depth of the radiation into the material with glancing angle close to the critical value θ_c, plays a large role in the formation of the anomalous scattering peak (Yoneda peak).

Ultrasoft X-ray spectroscopy is a promising experimental method for investigating the atomic and electronic structure of matter [1]. The small penetration depth of the radiation in matter (tens–hundreds of angstroms) and the local character of the photon absorption process in a specific atom in a multiatomic system make this method surface-sensitive [2, 3].

The purpose of the present work is to determine the depth of the near-surface layer of matter forming the reflected and scattered radiation in a two-phase system (surface layer + substrate) in the ultrasoft X-ray range. The system Si–SiO_2 with various thicknesses of the SiO_2 surface film, ranging from 20 to 630 Å, was chosen as such a system. This choice is due not only to the existence of details, differing in type and energy position, of the fine structure of the Si$L_{2,3}$ reflection spectra for crystalline Si and SiO_2 but also the well-known importance of this system in the production of various objects in the microelectronic industry.

All investigations were performed on an RSM-500 spectrometer-monochromator in a special camera attachment [4] using the bremsstrahlung and characteristic radiations from a tungsten anode. The energy resolution of the details of the spectra in the region of the Si$L_{2,3}$ ionization threshold was $\Delta E \simeq 0.3$ eV, and the angular resolution of the scattering indicatrix was 1.2°. The radiation was detected with a VEU-6 channel-type secondary-electron multiplier with a CsI photocathode. Wafers of a KDB (111) 4° silicon single crystal, which were 380 μm thick and 30 mm in diameter and were oxidized in a dry-oxygen atmosphere at $T = 1050$°C, were investigated. The thickness of the surface SiO_2 films grown was monitored using ellipsometry on an LÉF-2 apparatus.

Figure 1 shows the experimental Si$L_{2,3}$ reflection spectra for various SiO_2 thicknesses on a silicon substrate for various glancing angles θ_0 of the radiation. All systems were studied in a wide range of angles $4° \leq \theta_0 \leq 16°$, i.e., inside and outside the region of total external reflection of X-rays. The spectra presented are normalized to the maximum *B* (≃108.9 eV), with the exception of the spectrum for a system with a 20 Å thick surface film. We note that the fine structure of the reflection spectrum of crystalline Si is characterized by the presence of the details *a–g* (energy range 100–104 eV), while the spectrum of crystalline SiO_2 is characterized by the details *A–C* (105–120 eV). Thus, the presence of only one group of details in the spectrum corresponds to the case where the reflected radiation is formed by a single-phase system. The manifestation of details of both groups in the spectrum attests to reflection from a substrate + surface layer system. The fact that the contribution of the components of this system to the reflection is different should cause the corresponding group of details in the spectrum to dominate. It is natural to compare the thickness of the film for which details characteristic for a silicon spectrum, specifically, the bands *b* and *d–f*, appear in the reflection spectrum with the corresponding value of *d*—the penetration depth of the radiation in the material. Thus, for an Si–SiO_2 system with different thicknesses of the surface film, details characteristic of the substrate spectrum are present in the spectra for $\theta_0 = 4°$ right up to a thickness of 85 Å, and for glancing angles $\theta_0 = 8°$, 12°, and 16° they are present up to thicknesses 140, 190, and 260 Å, respectively. We also note the observed correlation of the spectra for different thicknesses of the surface layer but for different glancing angles. For example, the relative intensity distribution for the spectrum for a 85 Å thick film for $\theta_0 = 4°$ is virtually identical to that of the spectrum obtained for a 260 Å thick film for $\theta_0 = 16°$. This shows that the relative contributions of the surface

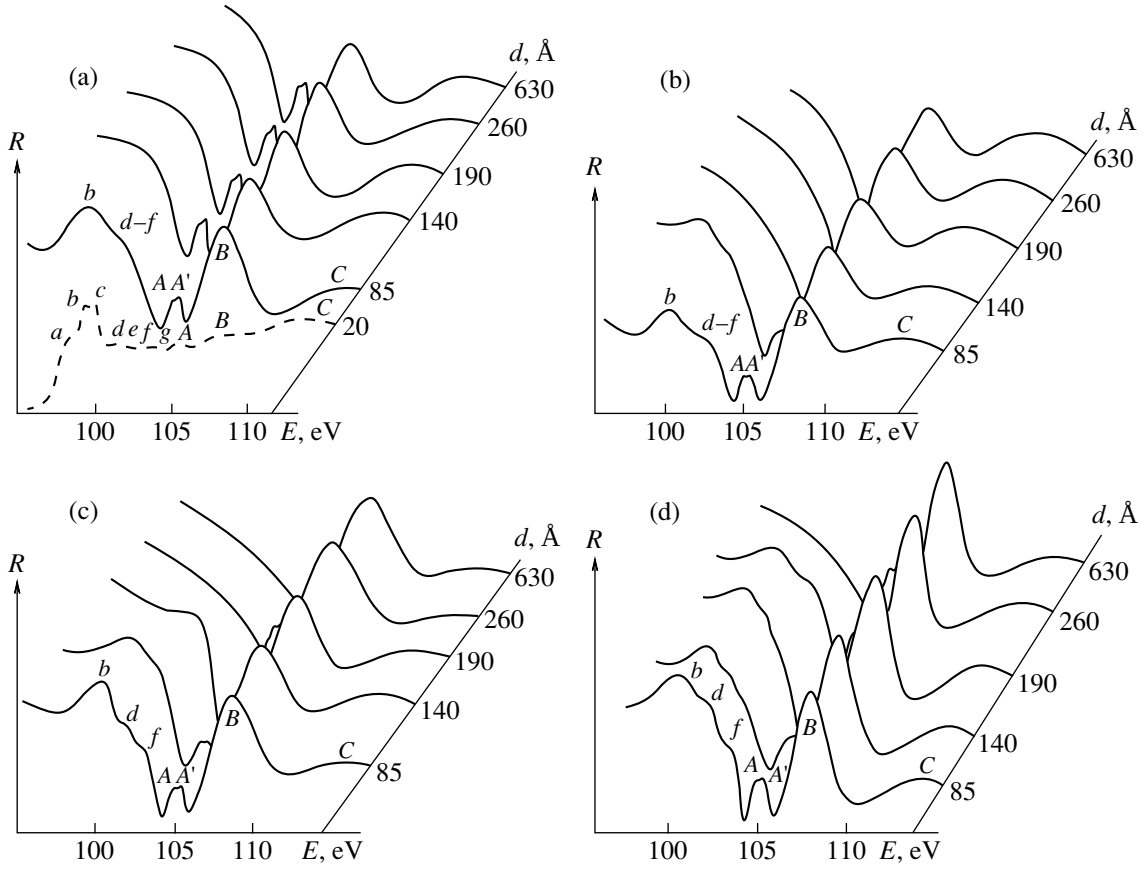


Fig. 1. Si$L_{2,3}$ reflection spectra for the system Si–SiO_2: θ_0 = (a) 4°; (b) 8°; (c) 12°; (d) 16°; *d* is thickness of the surface SiO_2 film.

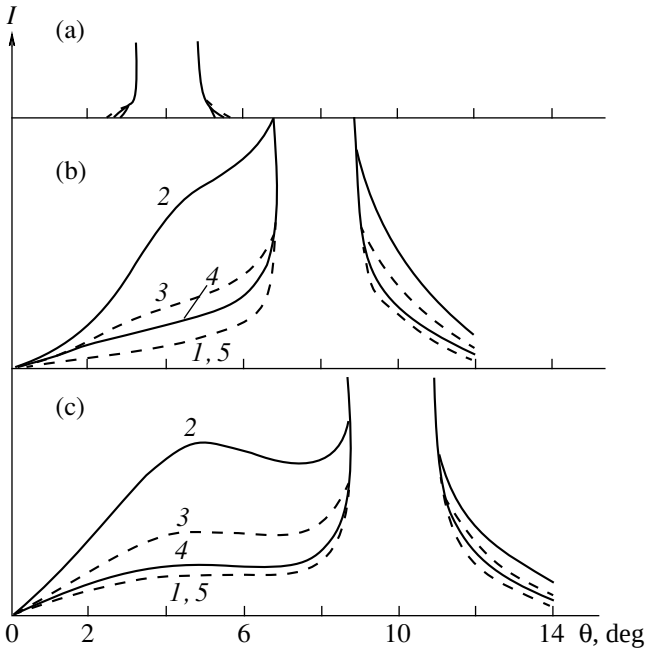


Fig. 2. Scattering indicatrix for the system Si–SiO_2 for λ = 57 Å for SiO_2 surface films with various thicknesses ((*1*) 20; (*2*) 85; (*3*) 140; (*4*) 190; (*5*) 630 Å): θ_0 = (a) 4°; (b) 8°; (c) 10°.

film and substrate to the reflected radiation are identical in these spectra.

The angular distribution of the scattered radiation (scattering indicatrix) from the surfaces of two-phase systems was also studied in this work. It is known [5, 6] that for glancing angles less than the critical angle θ_c, i.e., in the region of total external reflection, the scattering indicatrix carries information primarily about the roughness of the surface. In the range of angles $\theta_0 > \theta_c$, the characteristic features of the atomic spectra of the near-surface regions of materials can be judged from the form of the scattering indicatrix. For the samples investigated with different thicknesses of the surface film, it was found that in the region of total external reflection (θ_0 = 4°) the forms of the phase functions are essentially identical (Fig. 2a). Outside the region of total external reflection (θ_0 = 8°, 10°) an anomalous scattering peak appears in the scattering indicatrix (Yoneda peak) (Figs. 2b, 2c). In the figures presented, the detected radiation intensity was normalized to the intensity of the specularly reflected radiation. As one can see, the angular position of the Yoneda peak lies in the range 4.5°–5°, and the relative intensity is substantially different for different samples. The angular position of the anomalous scattering peak corresponds to

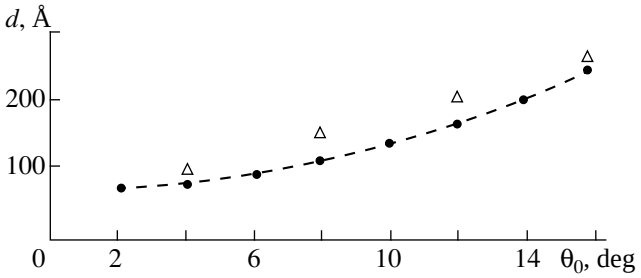


Fig. 3. Angular dependence of the formation depth of a specularly reflected beam in SiO_2: dots—calculation, triangles—experiment.

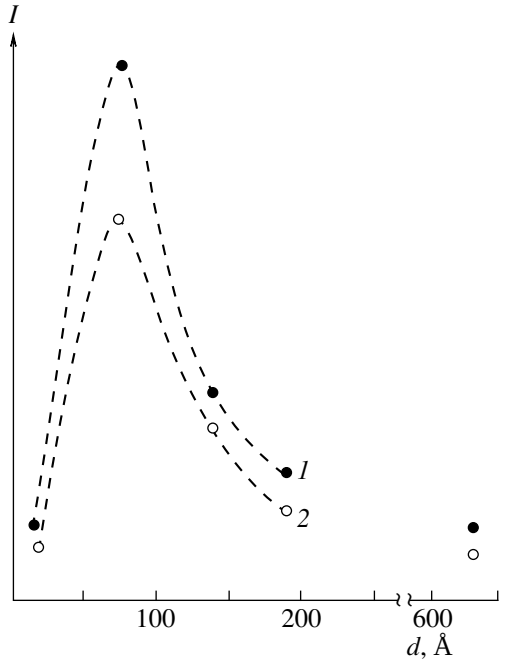


Fig. 4. Relative intensity of the anomalous scattering peak for SiO_2 surface films (λ = 57 Å) with different thicknesses: θ_0 = (*1*) 8°, (*2*) 10°.

the value of the critical angle θ_c [5], so that the critical angle for SiO_2 for λ = 57 Å can be assumed to fall within the indicated limits.

The angular dependence of the thickness of the surface layer forming the specularly reflected radiation in silicon dioxide can be constructed on the basis of the spectra obtained. This dependence is displayed in Fig. 3 together with the computed curve. The calculation was performed using the formula [7]

$$d = \frac{\lambda}{2\sqrt{2\pi}} \times [\sqrt{(\sin^2\theta_c - \sin^2\theta_0)^2 + \gamma^2} + \sin^2\theta_c - \sin^2\theta_0]^{-1/2},$$

where $\gamma = (\lambda/2\pi)\mu$, for the following values of the parameters: $\theta_c = 9°$, $\lambda = 117.4$ Å, and $\mu = 10^5$ cm^{-1}. The value of the critical angle was chosen on the basis of the fact that it is directly proportional to the wavelength of the radiation [8]. It is obvious that the theoretical and experimental curves agree well with one another in the entire range of comparison.

Figure 4 displays the dependence of the relative intensity of the Yoneda peak on the thickness of the SiO_2 film. As one can see, this dependence is distinctly nonmonotonic, and the maximum corresponds to a layer thickness of 85 Å. As inferred in [5], the largest relative intensity of the Yoneda peak can appear in cases where the vacuum-material boundary possesses a so-called transitional layer, i.e., the spatial structure of the material changes in a certain layer near the surface. The fact that the intensity of the Yoneda peak is greatest for d = 85 Å indicates that the largest structural changes occur in the layer forming this peak. On the other hand, the penetration depth of the radiation in the material for $\theta_0 = 4°–5°$ is close to d = 85 Å. It can be asserted on this basis that primarily a layer whose thickness corresponds to the penetration depth of the radiation in the material for $\theta_0 = \theta_c$ (in this case $\theta_c = 4.5°–5°$, d = 60–80 Å) influences the relative intensity of the anomalous scattering peak.

In summary, in the present work the thickness of the surface layer forming the specularly reflected radiation in silicon dioxide was determined for various glancing angles θ_0 (4°–16°) on the basis of experimental Si$L_{2,3}$ spectra for the system Si–SiO_2. The distinct nonmonotonic character established for the relative intensity of the Yoneda peak in the scattering indicatrix as a function of the thickness of the SiO_2 surface layer shows that the layer of matter corresponding to the radiation penetration depth in matter for $\theta_0 = \theta_c$ has the dominant effect on the form of this peak. This circumstance could make it possible to extract additional information when studying the characteristics of radiation scattered from the surface of various substances.

REFERENCES

1. T. M. Zimkina and V. A. Fomichev, *Ultrasoft X-ray Spectroscopy* (Leningr. Gos. Univ., Leningrad, 1971).
2. E. O. Filatova, A. S. Vinogradov, and T. M. Zimkina, Poverkhnost' **12**, 130 (1983).
3. A. S. Shulakov, E. O. Filatova, S. K. Kozhakhmetov, *et al.*, Fiz. Tverd. Tela (Leningrad) **32**, 2895 (1990) [Sov. Phys. Solid State **32**, 1682 (1990)].
4. I. A. Brytov, A. Ya. Grudskiĭ, E. A. Obolenskiĭ, *et al.*, Prib. Tekh. Éksp. **6**, 125 (1979).
5. A. V. Vinogradov and I. V. Kozhevnikov, Tr. Fiz. Inst. Akad. Nauk SSSR **196**, 18 (1989).
6. A. S. Shulakov, E. O. Filatova, S. K. Kozhakhmetov, *et al.*, Élektron. Tekh., Ser. 2: Poluprovodn. Prib. 212 (1991).
7. *Mirror X-ray Optics*, Ed. by A. V. Vinogradov (Mashinostroenie, Leningrad, 1989), p. 34.
8. A. V. Andreev, Usp. Fiz. Nauk **145**, 113 (1985) [Sov. Phys. Usp. **28**, 70 (1985)].

Translation was provided by AIP

Journal of Experimental and Theoretical Physics, Vol. 90, No. 5, 2000, pp. 826–833.
From Zhurnal Éksperimental'noĭ i Teoreticheskoĭ Fiziki, Vol. 117, No. 5, 2000, pp. 951–959.

SOLIDS
Structure

Coupling between Polarization, Orientation, and Chirality in Smectic Films Formed by Bent-Shaped Molecules¶

E. I. Kats*, ** and J. Lajzerowicz**

** Laboratoire de Spectrometrie Physique, Université Joseph-Fourier Grenoble 1, BP87-Saint-Martine D'Heres, Cedex, France*

*** Landau Institute for Theoretical Physics, Russian Academy of Science, Moscow, 117940 Russia*

Received January 18, 2000

Abstract—We formulate a simple Landau type model describing macroscopic behavior recently discovered in new smectic phases composed of achiral bent-shaped molecules. Films of such smectics exhibit three types of ordering related to dipole polarization, molecular tilt, and chirality. However, due to specific third-order coupling of the order parameters, these three symmetry-breaking types are not independent and this fact leads to specific domain structures that are observed in experiments. *© 2000 MAIK "Nauka/Interperiodica".*

1. INTRODUCTION

A variety of molecules form liquid crystalline phases (see, e.g., [1]). Many mesogen molecules have symmetries consistent with the formation of ferroelectric phases and nonzero dipole moments. Ferroelectric ordering is, however, extremely rare in positionally-disordered liquids or liquid crystals, and the surprising discoveries of ferroelectric [2] and antiferroelectric [3] behavior in liquid crystals are fundamental milestones in condensed matter physics. Since the discovery of ferroelectric liquid crystals, it has been usually assumed that ferroelectricity is possible only in the chiral smectic-*C** phase (formed by chiral molecules) that has the polar symmetry group C_2. In this case, polarization can be written as $\mathbf{P} = P\mathbf{n} \times \mathbf{n}_0$, where $\mathbf{n}$ is director [4] and $\mathbf{n}_0$ is the smectic layer normal.[1] The necessary conditions for the existence of nonzero polarization are a finite tilt angle (i.e., $\mathbf{n} \times \mathbf{n}_0 \neq 0$ or $\theta \neq 0$) and a molecular dipole perpendicular to the long axis of molecules. In racemic mixtures, which contain both enantiomers (that is, molecules that are mirror images of each other) in equal amounts, the electric polarization vanishes. Obviously, the electric polarization is directly connected to the molecular chirality in the Sm*C** ferroelectric liquid crystals.

However, there is no fundamental reason that non-chiral liquid crystals should not be ferroelectrics, since there is no unambiguous correspondence between the chirality of molecules and the existence of macroscopic ferroelectric properties or the structures they formed. The attempts of observation of ferroelectricity in non-chiral liquid crystals are, as a rule, centered around the synthesis and investigations of non-conventional liquid crystalline structures [5]. Recently, ferroelectric phases composed of achiral molecules were reported and investigated [6–8, 10–12]. In these papers, it was demonstrated that tilted smectic phases of achiral molecules show ferroelectric switching, and specific chiral domain structures. In [9], the bulk macroscopic properties of the lowest possible symmetry smectic phase (triclinic) were investigated and it was shown that such a system (though formed from achiral molecules) may possess ferroelectric and piezoelectric properties as well as macroscopic chirality. Due to polarity within smectic layers, such a smectic may have only integer strength of point-like defects in layers.

Note that in the above mentioned papers [6–8, 10–12] investigated experimentally only relatively thin freely suspended films (2–20 layers) and care must be taken in drawing conclusions about the bulk properties of liquid crystals from the behavior of films, as the surface layers of the film may be in a phase with higher (or lower) order than the bulk system. The surface phases cannot even exist as bulk phases. Particularly in [6, 8] instead of point-like defects predicted theoretically in [9] for the bulk phase, domain walls, i.e., two-dimensional defects in smectic layers were observed.

In our paper, we present a simple theory describing consistently experimental facts known for such low symmetrical smectic films formed by bent-shaped molecules. Our justification for adding one more paper on this topic is the fact that we alone seem to have taken into account entropic (and therefore, universal by the nature) third-order coupling between polarization, orientation, and chirality, specific just for these kind of systems. The organization of our paper is the following. In Section 2, we formulate our model and introduce (in the frame work of the Landau theory) the basic thermodynamics necessary for our discussions. In Section 3,

¶This article was submitted by the authors in English.

[1]Unlike non-polar nematic phases, where $\mathbf{n}$ and $-\mathbf{n}$ are equivalent, in ferroelectric smectics, the symmetry requires that $\mathbf{n} \longrightarrow -\mathbf{n}$ simultaneously with $\mathbf{n} \longrightarrow -\mathbf{n}_0$ give equivalent states.

we discuss different types of domain structures which may appear in smectics under consideration, and inspected the role of external influences (electric or magnetic fields and the concentration of chiral impurities). In Section 4, the analysis of how the coupling of the local order parameters near domain walls modify its structure is presented. Section 5 is devoted to a discussion and summary of our main results.

2. THEORETICAL MODEL

According to experimental data presented in the literature (see [6, 8, 10–12]), new smectic structures (labeled often in these papers as smectics B_2), are formed by polar but achiral molecules ("banana"-shaped) having the symmetry group C_{2v}, and macroscopic behavior of these structures is characterized by three spontaneous symmetry-breaking leading to the appearance of the following properties: molecular tilt, ferroelectric polarization, and chirality. The maximal macroscopic point symmetry group allowing these three types of symmetry breaking is C_2, where the second order axis should be parallel to smectic planes.

The tilt order parameter in any tilted smectic phases can be characterized by the two-component order parameter $\psi = \theta\exp(i\phi)$, where θ is the polar angle (tilt) and ϕ is the azimuthal angle of the nematic director **n**. Instead of ψ one can use the so-called **c**-director, which is the projection of the director **n** onto the layer plane. The magnitude of the tilt order parameter $|\mathbf{c}| = \sin\theta$. The ferroelectric polarization **P** is also a vectorial quantity, and it is only possible along the symmetry axis C_2.

From the general point of view, the chirality of the system is a third-order antisymmetric tensor which can be reduced for the system under study to the pseudoscalar χ. However, we have the only symmetry-breaking, namely $C_{2v} \longrightarrow C_2$ and therefore, all three order parameters should be interrelated, and the problem we now face is to find this relation. In fact, since the bend of **c** removes the $\mathbf{c} - \mathbf{n}_0$ mirror symmetry plane, it produces a local chiral symmetry breaking. This breaking of chiral symmetry can occur on two distinct length scales (microscopic or macroscopic). The distinction between microscopic and macroscopic chiral symmetry breaking is similar to the distinction between spontaneous and induced order parameters. From the macroscopic symmetry point of view to describe chiral, tilted, ferroelectric smectic films we have to introduce three order parameters (χ, **c**, **P**). Note that these three order parameters are not independent ones, and condensation of any pair amid them inevitably induces the non-zero value for the third one. Indeed, by the definition of **c**-director as the preferred direction (or projection of the director) in a smectic layer $\mathbf{c} = \mathbf{n} \times \mathbf{n}_0$, and thus if we have not only the tilt but as well non-zero ferroelectric polarization **P**, the system becomes macroscopically chiral, and as a measure of the chirality one can chose naturally

$$\chi = \mathbf{P} \cdot [\mathbf{n} \times \mathbf{n}_0] \equiv \mathbf{P} \cdot \mathbf{c}.$$

However, in this paper (unlike, e.g., [7]) we are interested in mainly microscopic causes of macroscopic symmetry breaking.

From the microscopic viewpoint, the existence of a tilt in smectic phases comes from the requirement of the molecular packing, (i.e., steric forces). These requirements fix for the polar molecules in our case (thin free standing films) the azimuthal angle ϕ and the module of the **c**-director, and therefore there are two allowed values of molecular tilt $\pm\theta$. Thus, any molecule in a smectic layer i[2] can be framed by two state systems labelled by indexes $\pm$ according to the sign of its tilt. In the same manner, the dipole moment **P** can be oriented either parallel or anti-parallel to the second order symmetry axis and it gives two more states attached to each molecular site. Therefore, each molecular site is a four state system: (+, +), (+, –), (–, +), (–, –), where the first sign corresponds to the tilt, and the second one to the dipole moment. If among the N^i molecules in a certain smectic layer i, the number of molecules in each state is $N^i(+, +)$, $N^i(+, –)$, $N^i(–, +)$, $N^i(–, –)$ then evidently

$$N^i = N^i(+, +) + N^i(+, –) + N^i(–, +) + N^i(–, –). \quad (2.1)$$

Analogously it is easy to see, that the tilt angle for the layer i can be represented as:

$$N^i\theta^i = N^i(+, +) + N^i(+, –) - N^i(–, +) - N^i(–, –), \quad (2.2)$$

and the polarization is given by

$$N^iP^i = N^i(+, +) + N^i(–, +) - N^i(+, –) - N^i(–, –). \quad (2.3)$$

It is important to note that, for each molecular site, the product of $P\theta$ represents the chirality of the given molecule, independently of site and of layer i. We follow here the idea and method developed recently for solid racemic solutions [13]. However, though for each individual molecular site $\chi \equiv P\theta$, this relation generally is not valid for the local mean values for a layer i, i.e., $\theta^iP^i \neq \langle\chi^i\rangle$, since analogously to (2.2), (2.3) one can write:

$$N^i\chi^i = N^i(+, +) + N^i(–, –) - N^i(–, +) - N^i(+, –). \quad (2.4)$$

In the spirit of the Bragg–Williams mean-field approximation, we can compute the entropy of the layer i

$$S = \ln\left[\frac{N^i!}{N^i(+, +)!\,N^i(+, –)!\,N^i(–, +)!\,N^i(–, –)!}\right]. \quad (2.5)$$

Solving the equations (2.1)–(2.4), introducing the found expressions for $N^i(\pm, \pm)$ in terms of the order

[2] For the simplicity and according to the layer structure of smectics, we suppose (unless the opposite will be said) that the order parameters are uniform within smectic layers.

parameters θ, P, χ, and expanding of (2.5) for small values of the order parameters we obtain

$$S^i = -N^i\Big[\frac{1}{2}(P^{i2} + \theta^{i2} + \chi^{i2}) + \frac{1}{2}(P^{i2}\chi^{i2} + P^{i2}\theta^{i2} + \chi^{i2}\theta^{i2}) + \frac{1}{12}(P^{i4} + \chi^{i4} + \theta^{i4}) - \theta^i\chi^iP^i\Big].$$

It is important to notice (and this is one of the main points of our investigation) the presence of the specific third order term $\theta^iP^i\chi^i$. The free energy of the system $F = U - TS$ (where U is the internal energy associated to intermolecular interactions) should have the same structure as the entropy but with non-universal (depending on detail of the interaction potentials) coefficients, namely for the layer i:

$$F^i = \frac{a_1}{2}\theta^{i2} + \frac{a_2}{2}P^{i2} + \frac{a_3}{2}\chi^{i2} + \frac{b_1}{2}P^{i2}\chi^{i2} + \frac{b_2}{2}\theta^{i2}P^{i2} + \frac{b_3}{2}\chi^{i2}\theta^{i2} + \frac{c_1}{2}\theta^{i4} + \frac{c_2}{2}P^{i4} + \frac{c_3}{2}\chi^{i4} + \gamma\chi^i\theta^iP^i. \quad (2.6)$$

The fact that the third order term necessarily figures in the free energy is related to the symmetry, since the product of the three representations to which θ, χ, and P belong includes the identical representation. The coefficients a_i, b_i, c_i, and γ can be considered as phenomenological parameters and a_i should become small near the corresponding symmetry-breaking transitions.

Taking into account the pair interactions between nearest neighboring layers, the total free energy can be written in the following general form:

$$F = \sum_i F_i + \sum_{i>j} a_{ij}\chi^i\chi^j + \sum_{i>j} b_{ij}P^iP^j + \sum_{i>j} c_{ij}\theta^i\theta^j.$$

Note that only pair interactions do not produce third order coupling found above. Therefore, in this approximation, (i.e., neglecting three-particle and further interactions) the third order coupling has the universal pure entropic origin, and $\gamma = -1$. The natural estimation for the entropic contribution is the temperature T (i.e., of the order of two-particle interactions responsible for the liquid-crystalline ordering) and therefore, this coupling can be larger than the three particle forces which we neglected.

To say more requires further knowledge of all these coefficients. Unfortunately, using only the data known from the literature, we are not able to extract values of all needed parameters. Therefore, we will not quantitatively compare our theory with available experimental data, since with too many unknown parameters the theory tends to become an exercise in curve fitting, which loses predictive credibility. Instead of this, we will discuss, in the next section, qualitative features of the model.

3. MODULATED DOMAIN STRUCTURES

Let us consider some very general consequences of the model. If we suppose to escape a conflict between experiment and theory that all three order parameters are uniform within smectic planes, this third order coupling means that the modulations of the order parameters $\chi(q_1)$, $\theta(q_2)$, and $P(q_3)$ along the normal to smectic layers should be matched to provide

$$q_1 + q_2 + q_3 = 0(\mathrm{mod}\, q_0),$$

where $q_0 = 2\pi/d$ is the wave vector of smectic density modulation (d is the interlayer distance). Thus, to satisfy the matching, there are only two possibilities: (1) one of the three wave vectors is zero and two others are anti-parallel; (2) all three wave vectors are zero.

In principle, in the continuous model, there are no restrictions on the period of the modulation, except for the requirement that it should be commensurate with q_0 (see also conclusion section of the paper). However, in the discrete model we use (i.e., each layer is a four state system), only wave vectors $\pm q_0/2$ and $\pm q_0/4$ are admissible. Even more, since independently on the model the chirality of the structure is always equal ± 1, only the wave vectors $\pm q_0/2$ lead to distinguishable macroscopically states for the smectic structures under consideration.

Let us assume that one from the three coefficients a_i is much smaller than the two others. Therefore, in the temperature region where this condition is fulfilled, we have only one soft order parameter, and we may neglect two others (hard in the region of parameters) degrees of freedom.

Experimentally (see [6, 8, 10–12]) for the smectic films which we consider, the main transition is associated with the condensation of the tilt θ. In this case, the theory is reduced to the well-known Landau theory for a scalar order parameter [14]. However, due to its importance for the present context (and for convenience), we repeat well-known results to apply them to our concrete case (free standing films). This is just the case where it is easy and more useful to derive these results for the concrete system under consideration than to try to find the suitable references, and to modify all expressions to apply them to the case.

There are two effects that are related to the existence of the surface in free standing films. The first is a purely geometrical one (finite size effects). The surfaces break the translational and rotational invariance (because the surface is a specific plane which breaks the translational invariance, and the normal to the surface is a specific direction which breaks the rotational invariance). Besides, certainly, there are physical modifications of the system due to the existence of the surface (surface effects). The surface can suppress the bulk ordering (this case is traditionally called the ordinary phase transition), the surface can enhance the bulk ordering (it is called the extraordinary phase transition), or as a third

possibility, the surface can experience its intrinsic critical behavior. There is also a so-called special phase transition which is intermediate between ordinary and extraordinary transitions.

Both effects related to the existence of the surface can be taken into consideration in the framework of the Landau expansion. In our particular case (film geometry and $a_1 \ll a_2, a_3$) it has the form

$$F = \int_0^L dz\left(\frac{1}{2}a_1\theta^2 + \frac{1}{4}c_1\theta^4 + \frac{1}{2}d_1(\nabla\theta)^2\right) + F_s, \quad (3.1)$$

where we added to (2.6) the gradient term (with the coefficient d_1) to describe the tilt profile over the film thickness L, and F_s is the surface energy which should have the same form as the bulk energy (3.1):

$$F_s = \frac{a'}{2}(\theta^2(0) + \theta^2(L)) + \frac{c'}{4}(\theta^4(0) + \theta^4(L)). \quad (3.2)$$

Usually, it is supposed that $a' \equiv d_1\lambda^{-1}$, where λ is called by extrapolation length and experimental data indicate that (at least as it concerns to the tilt) we have $\lambda < 0$ and it is traditionally called by extraordinary phase transition. From the microscopic viewpoint, the existence of the extraordinary phase transitions necessarily imply first- and second-layer interactions for the surface layer (see [15]).

In this case, the surface enhances the ordering and therefore, on the surface, one can expect the onset of ordering before (i.e., at higher temperatures) it occurs in the bulk. So one can expect, in this case, the surface transition for temperatures $T_s > T_c$ (by the definition the bulk transition temperature is determined from $a_1(T_c) = 0$). However, at T_c, due to the onset of the bulk order, the surface will experience some critical behavior as well. In the regime of $T_c < T < T_s$, the bulk correlation length ξ_b is finite and the order parameter decays from its maximum value at the surface. One can easily find the transition temperature for the surface layer:

$$\frac{T_s}{T_c} - 1 = \frac{d_1}{T_c}\lambda^{-2}.$$

To find the profile for the order parameter, we have to solve the Euler–Lagrange equation which follows from (3.1) supplemented by the boundary condition, which can be found from (3.2). We describe this rather routine procedure in the Appendix, and present only the results here. One can find that there are two types of configurations providing the minimum of the bulk functional (3.1) and simultaneously minimizing the surface energy (3.2). The first natural solution is symmetrical (we will term this solution by synclinic structure):

S-configuration: $\theta(z = 0) = \theta(z = L)$.

We imply that $a' = \alpha'(T - T_s)$, where T_s is the surface transition temperature (it can be extracted from experimental data for very thin films, e.g., for two-layer films). Determining the surface transition temperature, we can omit the third-order term in the equation for the bulk, and the transition in the film with N-layers occurs at T_N which can be found from the following equation

$$T_N = T_s - \frac{d_1}{\alpha'\xi_b(T_N)}\tanh\left(\frac{L}{2\xi_b(T_N)}\right). \quad (3.3)$$

The second solution (we will term it by anticlinic) is antisymmetrical:

A-configuration: $\theta(z = 0) = -\theta(z = L)$, and for this case

$$T_N = T_s - \frac{d_1}{\alpha'\xi_b(T_N)}\cosh\left(\frac{L}{2\xi_b(T_N)}\right). \quad (3.4)$$

The solution of both transcendental equations can be found very easily numerically and (as it should be) for small film thicknesses $L \ll \xi_b(T_N)$ synclinic configuration has always the higher transition temperature while for thick films with $L \gg \xi_b(T_N)$ the anticlinic solution can have the higher transition temperature. However, the anticlinic state can be only metastable due to the gradient energy (or in other words, due to the energy penalty which one must pay for the domain wall appearing inevitably for the anticlinic structure). However, the given above statement is valid only for the case $a_1 \ll a_2, a_3$, when we have to deal with one scalar order parameter condensation. This is not the case when we have two or three soft degrees of freedom (condensed order parameters) due to third order coupling between them.

4. ORDER PARAMETER COUPLING AND DOMAIN WALLS

We have mentioned already that for smectics under consideration, the first spontaneous symmetry breaking leads to the appearance of the tilt (though it could be only metastable pra-phase). When the temperature is decreased further on, the two remaining order parameters might be condensed as well. To study this scenario in the spirit of the mean-field approximation, one has to insert the average tilt $\langle\theta^i\rangle$ into the free energy expansion (2.6). The result will be the free energy expansion with two coupled-order parameters P and χ. Diagonalization of the free energy introduces two new order parameters η_1, η_2 which are linear combinations of P and χ. Since parameters a_2, a_3 depend on the temperature, the eigenvectors ensuring the diagonalization of the free energy are not fixed by the symmetry.

In principle, three different situations can arise according to the higher order terms:

(1) $\eta_1 \neq 0$; $\eta_2 = 0$;

(2) $\eta_1 = 0$; $\eta_2 \neq 0$;

(3) $\eta_1 \neq 0$; $\eta_2 \neq 0$.

Cases (1) and (2) correspond to different ordered phases with the same energy, while case (3) corre-

sponds to a situation of different energy. Clearly, the condition of lowest energy must prevail, and to find it, we have to know all coefficients entering the Landau expansion, which we do not have. However, as we mentioned already, the specific third-order coupling of the order parameters came mainly from the entropy of the system, i.e., from the number of states corresponding to smectic layers framed by four level systems. Therefore, in the main approximation, (i.e., neglecting non-entropic interactions) the coefficients $a_2 \simeq a_3$, and $\eta_1 \simeq \chi - P$, $\eta_2 \simeq \chi + P$. Evidently in this case, the solutions (1) or (2) should take place. Thus, below this second phase transition, eight different domains can appear in the system. They are characterized by $\pm\theta$, $\pm\eta_1$, $\pm\eta_2 = 0$ or $\pm\theta$, $\pm\eta_2$, $\pm\eta_1 = 0$.

In the case when all three order parameters are condensed, the minimization of the third order coupling energy leads to the following possible structures of smectics under consideration

(i)

$$\theta(q = 0), \quad P\left(q = \frac{q_0}{2}\right), \quad \chi\left(q = -\frac{q_0}{2}\right),$$

where q_0 is the wave vector of smectic modulation. This structure corresponds to the synclinic, antiferroelectric, and racemic phases;

(ii)

$$\theta\left(q = \frac{q_0}{2}\right), \quad P(q = 0), \quad \chi\left(q = -\frac{q_0}{2}\right),$$

i.e., anticlinic, ferroelectric, and racemic;

(iii)

$$\theta\left(q = \frac{q_0}{2}\right), \quad P\left(q = -\frac{q_0}{2}\right), \quad \chi(q = 0),$$

i.e., anticlinic, antiferroelectric, and homochiral;

(iv)

$$\theta(q = 0), \quad P(q = 0), \quad \chi(q = 0),$$

i.e., synclinic, ferroelectric, and homochiral.

It is worth noting that all four types of the predicted structures are really observed in experiments [6, 8]. Even more, it is clear that the application of the external electric field should stimulate the ferroelectric ordering of dipoles and therefore, only (ii) and (iv) structures will be stable in a strong enough field. This is also exactly what was observed in [6]. The same manner the external field conjugated to the chirality should induce (iii) and (iv) structures. As a physical realization of this field, one can have in mind the concentration of homochiral impurities. And the field conjugated to the tilt angle must induce (i) and (iv) structures only. Physically, such a field can be provided by the anchoring or magnetic field.

Note that for all structures (i)–(iv), the uniformly condensed order parameter is not necessarily uniform over the whole smectic layer. In the case when we have the condensation of two order parameters (there are three types of such pairs), one can observe a very rich behavior with many types of domain walls. For each type of wall at the variation of the parameters a_1, a_2, a_3, the wall transformations can be observed, which can be understood as an Ising–Bloch phase transition with the domain wall symmetry breaking (this type of transition was predicted theoretically in the neighborhood of the Curie point of a magnetic system in [18, 19]; here we have analogous transitions in chiral, tilted, and polar smectics under consideration).

For the illustration, we consider here one particular case of such a transformation in the structure of domain walls. As we mentioned already in the plausible scenario, the first transition is associated with the condensation of tilt θ. Let us imagine that in a certain smectic plane we have two domains with opposite signs of the tilt angle θ. Therefore, we get a domain wall in between them. After further cooling of the system at a certain temperature, the second order parameter will be condensed (let it will be the polarization P). However, owing to specific third order coupling, any variation of the order parameter χ at a given θ should lead to the corresponding variation of the polarization P. Depending on the coupling constant, different types of walls could exist. This corresponds to different trajectories between stable points in the order parameter space (on the plane χ–P for the case). When the coefficients a_2 and a_3 in the Landau expansion (2.6) are not very different (in the case when the entropic contribution into the free energy is dominant $a_2 = a_3$ ($\equiv a$)), the main contribution into the free energy (2.6) can be rewritten in the form

$$\frac{a}{2}(\eta_+^2 + \eta_-^2) - \langle\theta\rangle(\eta_+^2 - \eta_-^2), \tag{4.1}$$

where $\eta_\pm \equiv [P \pm \quad]/\sqrt{2}$. Therefore, in the domain of positive $\langle\theta\rangle$, one has the condensation of η_+, while in the domain of negative $\langle\theta\rangle$, the order parameter η_- will be condensed first. Thus, we have four stable points in the order-parameter space. All trajectories linking these points correspond to different structures of Ising-like domain walls (see Fig. 1 for the illustration).

Even between the same stable points depending on parameters, different types of domain walls (trajectories) can be realized. We schematically show three possible trajectories in Fig. 2.

Note that as it follows from (4.1), the anisotropy of the domain wall structures leading to Ising-like behavior of the walls is proportional to $\langle\theta\rangle$, and therefore, strictly speaking, the Ising–Bloch phase transition, accompanied by the symmetry breaking in domain walls, can be reached only at $\langle\theta\rangle = 0$.

As the wall thickness depends on both the uniform terms and the gradient energy, it is clear that these two components of the wall (namely χ and P) have different

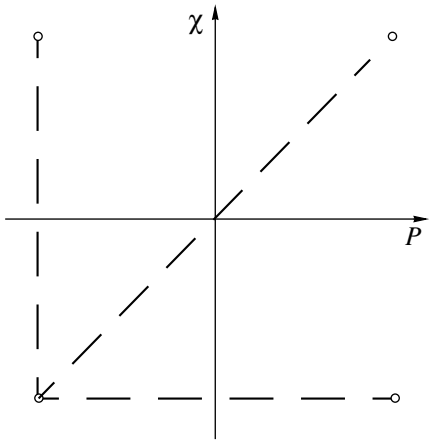


Fig. 1. Schematical trajectories on χ–P plane describing Ising domain walls.

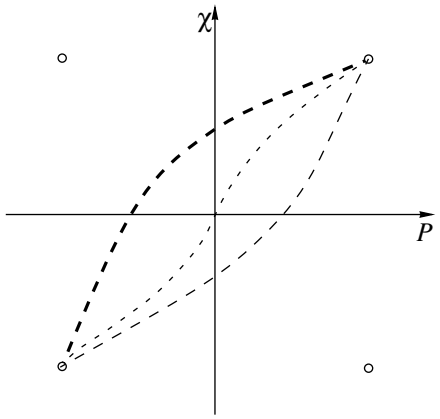


Fig. 2. Schematical trajectories on χ–P plane describing Bloch domain walls.

thicknesses. Following the arguments given in the paper [20], we can relate the trajectory to the chirality of the wall. One can call the domain wall of the type shown in Fig. 2, a double chiral one, since it can be characterized by two different types of chirality. The first one is the order parameter χ which defines the chirality of the system in a real coordinate space. It is a chiral structure in the space of the order parameters χ and P as well. The latter one can be defined via the gradient of the "angle" between two order parameters in order parameter space. As a measure for this second type of chirality, it is convenient to choose the Lifshits type of invariant, since it is not zero only for chiral domain walls:

$$\chi\left(P_x\frac{dP_y}{dz} - P_y\frac{dP_x}{dz}\right), \tag{4.2}$$

where z-axis is taken in the direction perpendicular to the wall and x, y are arbitrary axes in the plane of the wall. This contribution to the energy gives the energy of double chirality defects (χ–P domain walls) for smectics under considerations.

Unfortunately, we can find no guidance from experimental or theoretical sources for choosing all phenomenological coefficients that appear in these expressions. Thus, the primary function of this section must be to give a qualitative interpretation of our results and to demonstrate the possibility of chiral and ferroelectric ordering and domain ways in basically non-chiral systems, as opposed to proving exactly their existence.

5. CONCLUSION

We formulated a simple Landau-type model describing macroscopic behavior that recently discovered new smectic phases composed of achiral bent-shaped molecules. Films of such smectics exhibit three types of ordering related to dipole polarization, molecular tilt, and chirality. However, due to specific third order coupling of the order parameters, these three types of symmetry-breaking are not independent, and this fact leads to specific structures (i)–(iv) actually observed in experiments. This inhomogeneous ordering physically means that, over a large region of thicknesses of free standing films, they can be considered as some effective interfaces. It is typical for liquid crystals [21] that the width of the interface of experimental mesogenes is 40–100 times the length of molecules. We observed the example of how the presence of an interface may induce a type of ordering in the inhomogeneous region (for free standing films it may be the whole thickness of the system) that does not occur in the bulk phases. The analogous phenomena are also known for Langmuir monolayers where chiral symmetry can be spontaneously broken [22], and it leads to a chiral phase composed of non-chiral molecules. In fact, for a thick free-standing film, the top and bottom layers are each equivalent to Langmuir monolayers. We have seen the same type of structures for A-configuration where the tilt arrangement is anticlinic; i.e., the top and the bottom of the film are tilted in opposite directions. If it is anticlinic with the smectic periodicity, one can get conventional antiferroelectric structures, as has been recently found from ellipsometric studies in [23].[3]

Physical mechanisms providing the polarization properties of non-chiral and chiral-free standing films are very different. For the non-chiral systems the polar order is induced by the steric packing of anisotropic (but non-chiral) molecules, whereas in the ordinary (chiral) ferroelectric liquid crystalline phases, the polar order is a consequence of the molecular chirality. Owing to this fact, the value of spontaneous polarization for smectic constructed from achiral bent-shaped molecules is rather large (according to [8] it is around

[3] Note also the recent reflectivity and ellipsometric studies [24] which are fitted for thick free standing films by a model of the film consisting of surface anticlinic layers and an interior–azimuthal helix.

$300nC/cm^2$) and, therefore, the systems (unlike classical smectic C^* liquid crystals which are ferro-elastic ones, or free standing films of achiral smectics C showing piezoelectric polarization [7]), are genuine strong ferroelectrics. Note also that, as we have seen for such smectics, the polarization is parallel to the smectic layers and therefore, the depolarization field (appearing due to finite lateral size of the system) is very small.

In the presence of an external electric field E aligning the order parameter P (i.e., polarization $P = \kappa E$, where κ is the dipole polarizability of the system), the quadratic part of the free energy expansion (2.6) reduces to

$$\frac{1}{2}a_1\theta^2 + \frac{1}{2}a_3\chi^2 + \gamma\kappa E\chi\theta.$$

Excluding χ, one can find the renormalization of the main phase transition temperature (associated to the tilt angle condensation) or, what is the same, the renormalization of the coefficient a_1. The renormalized value of a_1 is

$$a_{1R} = a_1 - \frac{\gamma^2\kappa^2E^2}{2a_3}.$$

As it should be, the external field stimulates the phase transition (independently on signs of the coupling constants γ and κ). Thus, in the presence of the external electric field, the transition temperature into the smectic C field (associated to the tilt ordering) increases proportional to E^2.

Since the low-temperature phase of our system possesses the chiral and polar ordering, it should be characterized by non-zero Lifshits invariant of the type

$$\chi(\mathbf{P}\mathrm{rot}\mathbf{P}). \tag{5.1}$$

This term, in the presence of electromagnetic waves, leads to the natural optical activity, (i.e., it exhibits double circular refraction, and when a linearly polarized wave is propagated in the system possessing such a contribution into the energy, the plane of polarization is rotated and the angle of rotation per unit path length of the ray is proportional to $\omega\chi(\mathbf{P}\mathrm{rot}\mathbf{P})$, where ω is optical frequency).

Certainly, in a general case, (e.g., for thick films) we have to include into the consideration the azimuthal angle ϕ of the **c**-director, and the corresponding degree of freedom for **P** as well. Due to third order coupling, both azimuthal degrees of freedom are not independent ones. In this case, we get two order parameters **c** and **P** having the common one Goldstone degree of freedom ϕ for both of them. The contribution to the elastic energy of the smectic, analogous to the Lifshits invariant (5.1),

$$\chi(\mathbf{c}\mathrm{rot}\mathbf{c}), \tag{5.2}$$

leads to a spiral ordering of **c**-director as takes place for chiral smectics C^* or cholesterics.

ACKNOWLEDGMENTS

We are very grateful to Prof. P. Barois for sending all relevant publications on the subject. This work was supported in part by the Russian Foundation for Basic Research and INTAS grants and by the Russian State Program "Statistical Physics." E. K. thanks Prof. M. Vallade for supporting his stay at the Lab. Spectro., Joseph-Fourier University Grenoble-1 and for fruitful discussions. One the authors (E. K.) has greatly benefited from conversations with V.K. Dolganov and S.V. Malinin.

APPENDIX

LANDAU THEORY FOR FILMS

Minimization of (3.1) and (3.2) gives the following Euler–Lagrange equation

$$\frac{d^2m}{dx^2} = m + \frac{\xi_b^2}{a^2}m^3, \tag{A.1}$$

which should be supplemented by the boundary conditions

$$\pm\left(\frac{dm}{dx}\right)_{\pm l} + \alpha'(T - T_s)m(\pm l) + b'm^3(\pm l) = 0. \tag{A.2}$$

Here we introduced the following notations:

$$z = x\xi_b,\ \theta = \tilde{\theta}m,$$

where $\tilde{\theta}$ is the magnitude of the order parameter at $T \ll T_c$, $l = L/\xi_b$, and used the natural form of the Landau coefficients

$$a_1 = 2\alpha(T - T_c),\quad d_1 = 2\alpha T_c a^2,\quad c_1 = 2\alpha\frac{T_c}{\tilde{\theta}},$$

$$\theta_b = \tilde{\theta}\sqrt{1 - \frac{T}{T_c}},\quad \tau = 1 - \frac{T}{T_c}.$$

Let us consider first the case when we have the surface transition, but the bulk phase is still not tilted ($\theta_b = 0$). Neglecting into the bulk equation m^3, we find for the symmetrical case ($m = \mathrm{const}\cosh x$) expression (3.3) from the paper. The same manner for the antisymmetric case $m(x) = \mathrm{const}\sinh x$ and we get the expression (3.4) from the text. To find the magnitude of the order parameter induced by the surface phase transition, one should explicitly take into account the nonlinear terms into the equations.

The same manner near the bulk transition point we should include the bulk non-linear terms. For S-configuration from (A.1), (A.2) follow two relations:

$$\frac{L}{2a\sqrt{2}} = \int_1^u \frac{dx}{\{(x^2 - 1)[2\tau + m_0^2(x^2 + 1)]\}^{1/2}} \tag{A.3}$$

and

$$2a^2\frac{u^2}{\lambda^2} = (u^2-1)[2\tau + m_0^2(u^2+1)], \qquad (A.4)$$

where $m_0 = m(L/2)$.

Transition point for a finite thickness of the film L corresponds in the symmetrical case to $m_0 = 0$. Thus, the solution of both equations can be written in the parametric form

$$s\tanh s = \frac{L}{2\lambda}, \quad s = \frac{L\sqrt{\tau}}{2a}, \qquad (A.5)$$

and the explicit solution for $\tau_c(L)$ which is in fact the point where the surface order parameter appears (in the limit of $L \ll \lambda$):

$$\tau_c(L) = \frac{a^2}{2\lambda}\frac{1}{L}. \qquad (A.6)$$

For *A*-configuration instead of (A.3), (A.4) we get a system

$$u^2 = \left(2\tau - \frac{2a^2}{\lambda^2}\right)$$

and

$$\frac{\sqrt{\tau}}{2a} = \ln\frac{u}{\sqrt{2\tau}}$$

and again in the limit $L \ll \lambda$ we obtain

$$\tau_c(L) = \frac{a^6}{\lambda^4}\frac{1}{L^2}. \qquad (A.7)$$

Thus, expressions (3.3) and (3.4) from the text give us the modification of the surface phase transition point due to the interaction with the bulk, and (A.6) and (A.7) express the bulk transition temperature dependences on the thickness owing to the existence of the surface layers.

REFERENCES

1. S. Chandrasekhar, *Liquid Crystals* (Cambridge Univ. Press, New York, 1977).
2. R. B. Meyer, L. Liebert, L. Strzelecki, and P. Keller, J. Phys. Lett. **36**, L69 (1975).
3. A. D. L. Chandani, E. Gorecka, Y. Ouchi, *et al.*, Jpn. J. Appl. Phys. **28**, L1265 (1989).
4. P. G. de Gennes, *The Physics of Liquid Crystals* (Clarendon, Oxford, 1974).
5. L. M. Blinov, Liq. Cryst. **24**, 143 (1998).
6. D. R. Link, G. Natale, R. Shao, *et al.*, Science **278**, 1924 (1997).
7. P. O. Andreeva, V. K. Dolganov, C. Gors, *et al.*, Phys. Rev. E **59**, 4143 (1999).
8. G. Heppke, A. Jakli, S. Rauch, and H. Sawade, Phys. Rev. E **60**, 5575 (1999).
9. H. R. Brand, P. E. Cladis, and H. Pleiner, Eur. Phys. J. B **6**, 347 (1998).
10. H. T. Nguyen, J. C. Rouillon, J. P. Marcerou, and P. Barois, in *Proceedings of the International Conference on Liquid Crystals, Strasbourg, 1999*, D2-O3.
11. P. Mach, R. Pindak, A.-M. Levelut, *et al.*, Phys. Rev. Lett. **81**, 1015 (1998).
12. J. P. Bedel, H. T. Nguyen, J. C. Rouillon, *et al.*, Mol. Cryst. **328**, 177 (1999).
13. J. Lajzerowicz, J. Lajzerowicz-Bonneteau, and B. Suchod, J. Phys. I **1**, 573 (1991).
14. K. Binder, *Phase Transitions and Critical Phenomena*, Ed. by C. Domb and J. L. Leibowitz (Academic, London, 1986), Vol. 8.
15. B. Houchmanzadeh, J. Lajzerowicz, and E. Salje, J. Phys.: Condens. Matter. **4**, 9779 (1992).
16. L. M. Blinov, *Electro-Optical and Magneto-Optical Properties of Liquid Crystals* (Nauka, Moscow, 1978; Wiley, New York, 1983).
17. P. I. C. Teixeira and E. M. Terentjev, Eur. Phys. J. B **3**, 237 (1998).
18. L. N. Bulaevskiĭ and V. L. Ginzburg, Zh. Éksp. Teor. Fiz. **45**, 772 (1963) [Sov. Phys. JETP 18, 530 (1964)].
19. J. Lajzerowicz and J. J. Niez, J. Phys. Lett. **40**, 165 (1979).
20. B. Houchmanzadeh, J. Lajzerowicz, and E. Salje, J. Phys.: Condens. Matter. **3**, 5163 (1991).
21. S. Faetti and V. Palleschi, Phys. Rev. A **30**, 3241 (1984).
22. J. V. Selinger, Z. G. Wang, R. F. Bruinsma, and C. M. Knobler, Phys. Rev. Lett. **70**, 1139 (1993).
23. D. Schlauf and Ch. Bahr, Phys. Rev. E **57**, R1235 (1998).
24. P. M. Johnson, S. Pankratz, P. Mach, *et al.*, Phys. Rev. Lett. **83**, 4073 (1999).

Journal of Experimental and Theoretical Physics, Vol. 90, No. 5, 2000, pp. 834–841.
Translated from Zhurnal Éksperimental'noĭ i Teoreticheskoĭ Fiziki, Vol. 117, No. 5, 2000, pp. 960–968.

SOLIDS
Electronic Properties

Radiofrequency Magnetoresistance of Fe/Cr Superlattices

A. B. Rinkevich*, L. N. Romashev, and V. V. Ustinov
Institute of Metal Physics, Ural Division, Russian Academy of Sciences, Yekaterinburg, 620219 Russia
**e-mail: rin@imp.uran.ru*
Received July 6, 1999

Abstract—The rf magnetoresistance of Fe/Cr superlattices is studied for two orientations of the current: parallel and across the superlattice layers. A mutually single-valued correspondence is established between the relative magnetoresistance measured at dc current and the change in the transmission coefficient of electromagnetic waves in the magnetic field. When rf currents flow across the layers, the relative change in the signal amplitude is proportional to twice the change in the electrical resistance of the superlattice and is of opposite sign. It is shown that the rf losses are determined by the surface resistance which is proportional to the superlattice thickness and inversely proportional to its conductivity. An equation is derived for the rf electric field distribution in the superlattice. It is established that when the thickness of the superlattice is small compared with the skin layer depth, field and current components which penetrate through the entire superlattice exist.

1. INTRODUCTION

Studies of the magnetic and electrical properties of metal superlattices are one of the main trends in the modern physics of nanostructures. Of particular interest here is the giant magnetoresistance of metal multilayers [1]. This effect is caused by a change in the relative orientation of the magnetic moments of the ferromagnetic layers in the superlattice when an external magnetic field is applied and depends strongly on the nature of the exchange interaction between neighboring layers separated by a thin "nonmagnetic" layer. This effect is also sensitive to electron scattering processes at the interface between the magnetic and nonmagnetic layers. Of particular interest is the situation when electric current flows across the superlattice layers. In this case, the giant magnetoresistance effect is most clearly defined and has important features associated with the phenomena of spin injection and electron accumulation. It should be noted that, because of the extremely low electrical resistance of the samples, it is difficult to carry out experiments in a geometry where the current in the superlattice flows across the layers. Fairly complex methods of preparing samples for such measurements are described in the literature. In particular, in order to increase their transverse electrical resistance, the superlattices in these experiments were prepared in one case by depositing atomic layers on specially prepared grooved substrates [2] and in another case by electrochemical deposition in small-diameter apertures and the fabrication of so-called nanowires [3].

Promising methods of studying the giant magnetoresistance effect involve using microwave electromagnetic oscillations and waves. By using different types of oscillations and waves, it is fairly easy to achieve various orientations of rf electric and magnetic fields and currents relative to one another and to an external magnetic field. In addition, the rf technique can be used to answer the separate question: up to what frequencies can the giant magnetoresistance of superlattices be observed? The answer to this contains information on the rf conductivity of superlattices, which is relevant to predicting the practical application of metal superlattices in high-speed devices.

Only a few publications have dealt with the rf properties of metal superlattices [4–6]. It was shown in [4] that a direct relationship exists between the absorption of rf electromagnetic waves and the giant magnetoresistance effect. The electromagnetic properties of superlattices in the infrared were studied in [5]. One-to-one correlation between the giant magnetoresistance effect and the propagation of microwaves through the superlattice was established in [6].

In the present paper we make a detailed study of the rf magnetoresistance of superlattices. Particular attention is paid to achieving different relative orientation of the rf fields and currents in the layers. The objects studied are Fe/Cr superlattices for which the magnetic moments of the neighboring Fe layers exhibit noncollinear ordering [7]. The rf properties are studied in comparison with the giant magnetoresistance effect.

The superlattice sample is positioned either in the cross section of a waveguide or in an empty resonator. In the second case, any change in the electrical resistance of the sample in the magnetic field changes the electromagnetic losses. The complex resistance Z near the resonance at frequency ω_0 may be expressed as $Z \simeq R(1 + 2iQ\Delta\omega/\omega_0)$, where R is the resistance of the circuit, Q is the Q factor, and $\Delta\omega$ is the detuning from resonance. The oscillation power in resonance at $\Delta\omega = 0$ is

proportional to $1/R^2$. The relative change in the power T in the magnetic field H is given by

$$\frac{1/R^2(H) - 1/R^2(0)}{1/R^2(0)} = -2r\frac{1 + r/2}{(1 + r)^2},$$

where $r = [R(H) - R(0)]/R(0)$ is the relative magnetoresistance. For small r the relative change in the rf power is $T \simeq -2r$. When the negative magnetoresistance of the superlattice is high, T may appreciably exceed $|r|$. This essentially describes the enhancement of the giant magnetoresistance of superlattices: the change in the rf power is a factor of two or more greater than the relative magnetoresistance measured at dc current.

2. EXPERIMENTAL CONDITIONS

The samples were Fe/Cr superlattices grown by molecular beam epitaxy. The substrates were single-crystal magnesium oxide (MgO) wafers cut along the (100) plane. The wafers measured 30 × 30 × 0.5 mm. The substrate temperature during deposition of the Fe and Cr layers was 200°C. Details of the superlattice fabrication technology are given in [8]. In the present study we took into account several factors when selecting the samples. First, the superlattice samples must have a fairly high magnetoresistance. Second, it is desirable that the field dependence of the magnetoresistance in samples of different thickness should be different. The characteristics of the superlattices used in this study are given in the table. Sample 4 is not a superlattice but an Fe film whose thickness is of the order of the total layer thickness in the superlattices. By comparing the microwave properties of the thin Fe film and Fe/Cr superlattices, we can identify the contribution made by the multilayer structure of the superlattices.

All the measurements were made at room temperature. The relative magnetic resistance $r = [R(H) - R(0)]/R(0)$ was measured in magnetic fields H up to 32 kOe. The dc electrical resistance was measured by a standard four-contact method. In this case the magnetic field was oriented parallel to the plane of the superlattice.

The microwave measurements were made by two methods. In the first the sample was positioned in the cross section of a rectangular waveguide operating in the TE_{01} mode. The position of the sample in the waveguide is shown in Fig. 1 which gives the directions of the external magnetic field **H**, the rf electric field $\mathbf{E}_\sim$, and the wave vector **q** of the electromagnetic wave. The vectors are mutually perpendicular: $\mathbf{H} \perp \mathbf{E}_\sim$, $\mathbf{E}_\sim \perp \mathbf{q}$. In this case, the rf field $\mathbf{E}_\sim$ and the currents are parallel to the superlattice layers. An experiment was carried out using sample 1 in the frequency range between 5.4 and 11 GHz. We measured the relative change r_m in the modulus of the transmission coefficient D: $r_m = [D(H) - D(0)]/D(0)$. The transmission coefficient was introduced as the ratio of the transmitted power P_{out} to the input power P_{in}.

Table

Sample no.	Fe layer thickness t_{Fe}, Å	Cr layer thickness t_{Cr}, Å	Cr buffer layer thickness Cr, Å	Number of periods, n	Total metal thickness, Å
1	20.5	10	70	12	436
2	21	17	70	12	526
3	14.5	9	70	30	775
4	710	–	85	1	795
5	8	10	80	40	800

In the second method the sample was positioned at the maximum of the electric field of a coaxial resonator (Fig. 2). The resonant frequency was varied by varying the gap between the central rod and the bottom of the resonator. This method used frequencies between 0.5 and 2.1 GHz. Eddy currents appear in the superlattice and the Joule losses reduce the resonator Q factor. Any change in the electrical resistance in the magnetic field as a result of the giant magnetoresistance effect changes the resonator losses and the Q factor. Since the electric field $\mathbf{E}_\sim$ is oriented perpendicular to the layers, the rf current flows in the same direction so that the magnetoresistance can be measured fairly easily using the so-called current-perpendicular-to-plane experimental geometry. A method of measurement for this geometry was developed in [9].

A coaxial resonator was connected into a microwave circuit as a transfer element. As we know, the transmission coefficient of a resonant transmission element can be expressed in terms of the Q factor as follows [10]:

$$D = \frac{2Q_L^2}{Q_1 Q_2}\frac{1}{1 + Q_L^2(\Delta\omega/\omega_0)^2}, \quad (1)$$

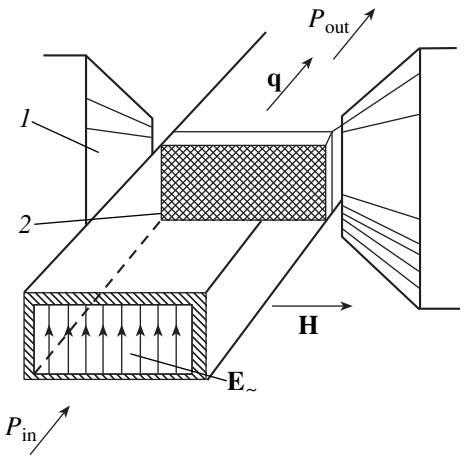


Fig. 1. Position of sample in waveguide: (*1*) electromagnet; (*2*) superlattice.

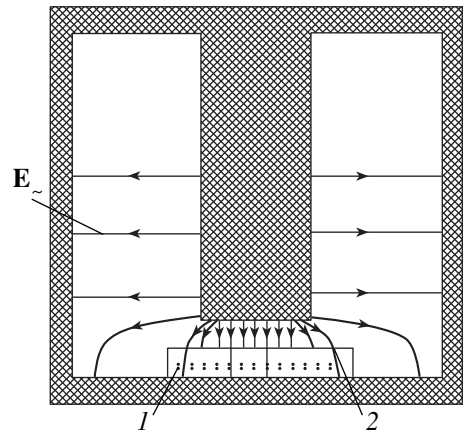


Fig. 2. Position of sample in resonator: (*1*) MgO; (*2*) superlattice.

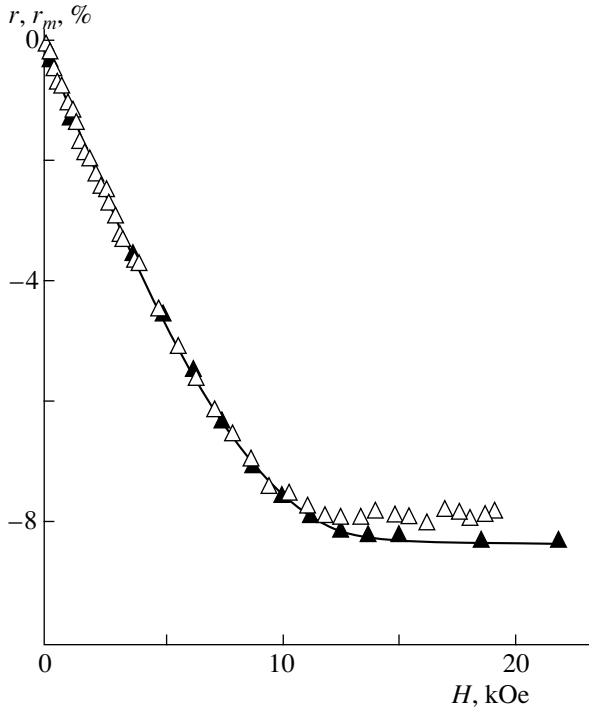


Fig. 3. Magnetoresistance r (▲) and transmission coefficient r_m (△) at frequency f = 8.4 GHz.

where Q_L is the loaded resonator Q factor, Q_1 and Q_2 are the Q factors of the input and output coupling elements, and $\Delta\omega$ is the detuning of the frequency ω from the resonant frequency ω_0. The relative change T in the transmission coefficient D in a magnetic field at the resonant frequency is given by

$$T = \frac{D(H) - D(0)}{D(0)} = \frac{Q_L^2(H)}{Q_L^2(0)} - 1. \qquad (2)$$

The stored oscillation energy of the resonator is given by

$$W = \int_V \frac{|H_\sim|^2}{2} dV,$$

where $H_\sim$ is the amplitude of the rf magnetic field and V is the resonator volume. The power loss in the sample is

$$P_{\text{loss}} = \frac{1}{2}\int_s |j^2| R_s dS,$$

where j is the surface density of the rf current at the surface S of the sample and R_s is the surface resistance. In our case $R_s = d/\sigma_{\text{eff}}$ where d is the total thickness of the superlattice metal layer and σ_{eff} is the effective conductivity. If $Q_L \ll Q_{1,2}$, then bearing in minding the definition of the loaded Q factor, we obtain the following formula for the relative change T:

$$T = -2\frac{Q_L(0)}{Q_0(0)}\frac{R_s(H) - R_s(0)}{R_s(0)}, \qquad (3)$$

where $Q_0(0)$ and $Q_L(0)$ are the intrinsic and loaded resonator Q factors without a magnetic field. Formula (3) holds if the relative change in the surface resistance is small. Thus, the value of T is directly related to the change in the resistance in the magnetic field. Since the magnetoresistance of superlattices is negative, it is predicted that the change T in the magnetic field as given by (3) will be positive. If the losses in a high-Q resonator are caused mainly by Joule losses in the superlattice, we have $Q_0(0) \approx Q_L(0)$. The changes in the coefficient T are then twice the modulus of the relative magnetoresistance measured at dc current.

3. MICROWAVE MAGNETORESISTANCE

These experiments show that a single-valued correlation exists between the transmission coefficient of the electromagnetic waves and the magnetoresistance of the superlattices measured at dc current. We shall first consider the waveguide measurement method. Figure 3 gives the results of measurements of the magnetoresistance r of sample 1 and the dependence of the rf transmission coefficient r_m on the static magnetic field. The rf measurements were made at f = 8.4 GHz. These results convincingly demonstrate the similarity between the dependences $r(H)$ and $r_m(H)$ and the numerical values contained in them. This confirms the correlation between the rf transmission coefficient and the dc magnetoresistance. Figure 4 shows dependences of the coefficient $r_m(H)$ measured at different frequencies. It can be seen that these curves are identical.

The experiments have shown that the coefficient of transmission of electromagnetic waves across the superlattice is completely determined by the giant magnetoresistance effect. The region of existence of the rf

analog of the giant magnetoresistance effect is at least tens of Gigahertz.

We shall now discuss the second method, where the sample is placed in a coaxial resonator. Figure 5 shows the transmission coefficient T measured as a function of frequency f. The symbols give the experimental data obtained for sample 5 for $H = 0$; the solid curve gives the calculations using formula (1) for the loaded Q factor $Q_L = 1089$.

The correlation between the magnetoresistance r measured for a dc current flowing along the superlattice layers and the changes in the rf losses T is not obvious *a priori* since under these experimental conditions the microwave currents flow across the layers. Nevertheless, a clear correlation can be identified. Figure 6 shows the correspondence between the field dependences of r (Fig. 6a) and T (Fig. 6b) for samples 2, 3, and 4. The microwave measurements were made at $f = 779$ MHz. As we can see, the dependences $r(H)$ and $T(H)$ are similar. Sample 2 typically has a relatively weak saturation field, around 5 kOe. It can be seen from Fig. 6 that saturation is observed for the magnetoresistance and for the microwave coefficient T. Sample 3 does not reach saturation in the magnetic fields used in the present study and for this the dependences $r(H)$ and $T(H)$ are close to linear. The iron film (sample 4) exhibits a very low magnetoresistance, considerably lower than the giant magnetoresistance of the superlattices, and the change in the microwave coefficient T is also negligible.

Thus, the results obtained by the resonator method indicate an undisputed correlation between the magnetoresistance of superlattices measured at dc current and the change in the rf losses. It will be shown in the following section that a current-perpendicular-to-plane geometry is achieved for this orientation of the superlattice plane.

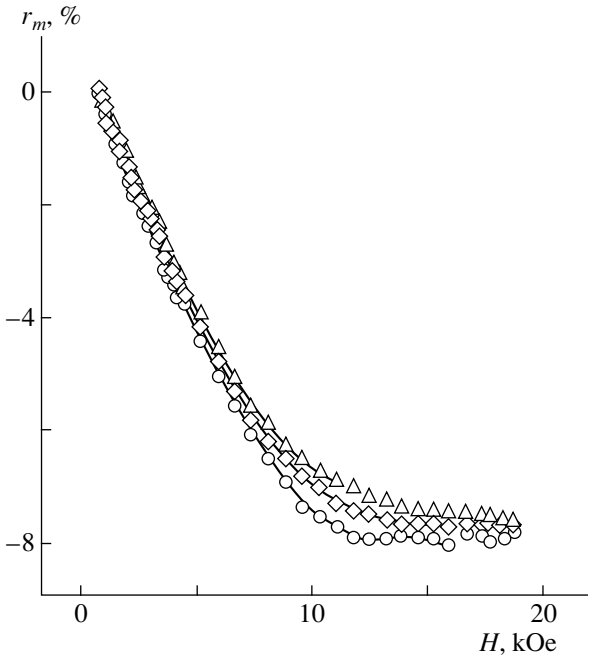


Fig. 4. Dependence of the transmission coefficient on the magnetic field at frequencies $f = 6.85$ (△), 8.15 (○), and 8.4 (◇) GHz.

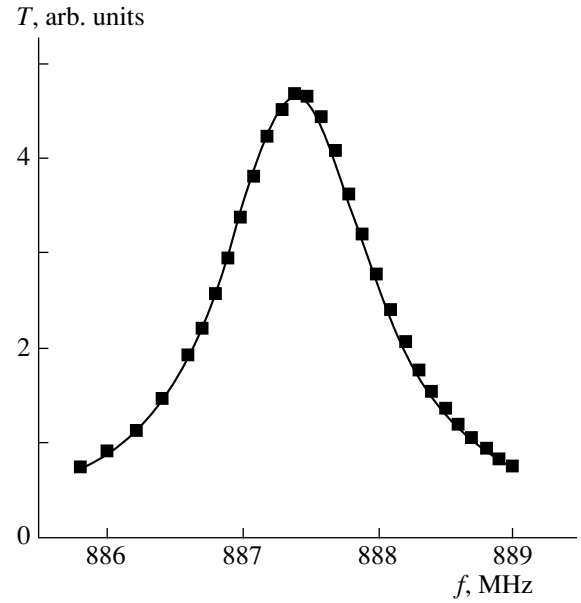


Fig. 5. Frequency dependence of the transmission coefficient of a resonator with a superlattice (sample 5), $H = 0$.

4. DISCUSSION OF RESULTS

Fairly weighty evidence indicates that a mutually single-valued correlation exists between the change in the electrical conductivity in the magnetic field and the microwave absorption and propagation through the superlattice. We shall first consider the change in the coefficient of propagation of electromagnetic waves through the superlattice. Let us assume that a plane wave is incident normally on a thin metal object. For the amplitudes of the incident E_i, reflected E_r, and transmitted E_t waves at the metal interface we can write the system of equations

$$\begin{aligned} E_i + E_r &= \frac{c}{4\pi}[Z_{00}(E_i - E_r) - Z_{01}E_t], \\ E_r &= \frac{c}{4\pi}[Z_{01}(E_i - E_r) - Z_{11}E_t]. \end{aligned} \tag{4}$$

This system includes the elements Z_{ij} of the impedance matrix. For a good conductor the inequality $\delta \ll q_0^{-1}$ is satisfied, where δ is the skin layer depth and q_0 is the wave number ($q_0 = \omega/c$). If this condition is satisfied,

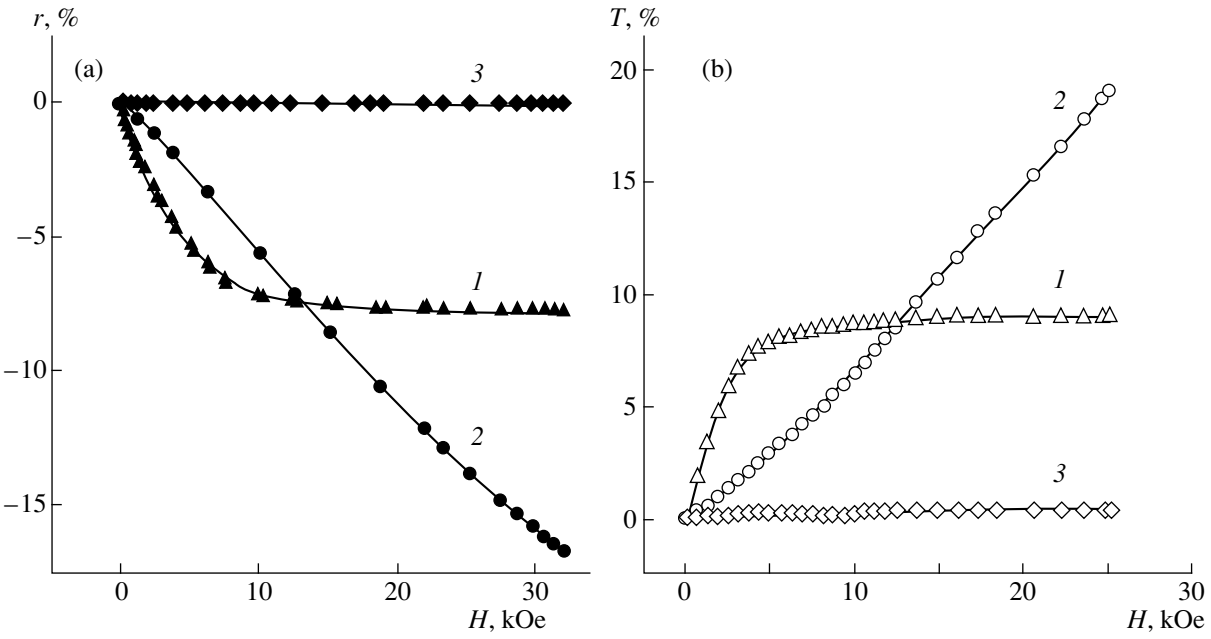


Fig. 6. Correlation between the magnetoresistance (a) and the resonator transmission coefficient (b) when rf currents flow perpendicular to the superlattice layers for various samples: (*1*) 2; (*2*) 3; (*3*) 4.

simple expressions are known [11] for the coefficients of reflection R and transmission D:

$$R \simeq -1 + \frac{c}{4\pi} Z_{00}, \quad D \simeq \frac{c}{4\pi} Z_{10}. \tag{5}$$

The reflection coefficient R is close to –1. Under our experimental conditions we are dealing with a situation where the total metal thickness d is much smaller than the skin-layer depth ($d \ll \delta$) under conditions of the normal skin effect. In this case, the transmission coefficient D is expressed simply in terms of the effective electrical conductivity σ_{eff}:

$$D \simeq \frac{q_0 \delta^2}{d} = \frac{c^2}{2\pi \sigma_{\text{eff}} d}. \tag{6}$$

On the right-hand side of equation (6) only the electrical conductivity σ_{eff} depends on the magnetic field. Expression (6) yields a mutually single-valued correlation between the relative magnetoresistance r and the relative change r_m in the transmission coefficient of the microwave electromagnetic waves. The data plotted in Fig. 3 convincingly demonstrate that such a correlation exists for Fe/Cr superlattices.

Incidentally, it should be noted that expression (6) for the transmission coefficient was obtained for conditions slightly different from the experimental ones. First, the electromagnetic wave in the waveguide exhibits velocity dispersion. This factor may prove significant if the experiments are carried out near the waveguide cutoff frequency. In the present study the frequencies at which the measurements were made are considerably higher than the cutoff frequency so that the influence of the wave velocity dispersion can be neglected. Second, the real structure of the superlattice was not taken into account in deriving equation (6). The possibility of introducing the effective conductivity σ_{eff} for a finely layered medium [12] arises because the layer thicknesses t_{Fe} and t_{Cr} are small compared with the skin layer depth $\delta \gg t_{\text{Fe}}, t_{\text{Cr}}$.

Note that under these assumptions the transmission coefficient D does not depend on the electromagnetic wave frequency. This is also demonstrated by the experimental results plotted in Fig. 4.

We shall now discuss the results obtained when the sample is placed in the electric field region of a coaxial resonator. If the sample possessed infinitely high conductivity, the boundary conditions would require the vector of the rf electric field $\mathbf{E}_\sim$ to be strictly perpendicular to the sample plane. In a real sample a tangential component E_t appears and we shall estimate this using the Leontovich condition [13]. If $H_\sim$ is the amplitude of the magnetic field at the metal surface, the tangential component E_t is given by [14]

$$E_t = \frac{1}{2}(1 + i) q_0 \delta H_\sim. \tag{7}$$

Under our conditions of a quasisteady-state electromagnetic field ($\delta \simeq 10^{-4}$–10^{-3} cm, $q_0 \simeq 0.2$ cm^{-1}) the product is $q_0\delta \ll 1$. In addition, the value of $H_\sim$ is low in the region of strong electric field in the coaxial resonator where the sample is located. Since the tangential component of the electric field is small, we can assume that the rf field enters the superlattice along the normal to its surface (along the z axis). Bearing in mind that the

metal thickness d is small compared with the skin layer depth δ, we assume that the electric field inside the superlattice only has a z-component E_z. The subscript "z" will be omitted subsequently.

As we have noted, at these frequencies the superlattice is a small-scale medium. The electrodynamics equations written for the electric fields averaged over the period of the superlattice

$$\overline{E(z)} = \frac{1}{t_{\mathrm{Fe}} + t_{\mathrm{Cr}}} \int_{z}^{z + t_{\mathrm{Fe}} + t_{\mathrm{Cr}}} E(z)dz$$

are equivalent to the equations for a homogeneous thin film having a certain effective conductivity [12]. Under the experimental conditions, a normal skin effect is achieved where the spatial dispersion can be neglected. The possibility of introducing the effective conductivity tensor of the superlattice was analyzed in [15] where it was shown that this is determined by the matrices of the electron scattering frequencies and uniquely describes the motion inside the layer, reflection from the layer boundaries, and the transition to a neighboring layer with change of spin. The expression for the effective conductivity in [15] was obtained for the case where the electric field lies in the plane of the superlattice layers. The calculations of the field distribution given below refer to the current-perpendicular-to-plane geometry and allow for screening of the field by the conductor. The aim of these qualitative calculations is to demonstrate the existence and estimate the magnitude of the current component flowing across the entire thickness of the superlattice considerably greater than the Debye length r_D.

Jointly solving the transport equation in the τ-approximation, the Maxwell equations for the divergences, and the equation of continuity yields the following equation:

$$\omega_0^2\overline{E(z)} - \omega_0^2 Z_D^2 \frac{\partial^2 \overline{E(z)}}{\partial z^2} = -(\omega^2 + i\omega\nu)E(0), \qquad (8)$$

where $\omega_0 = \sqrt{4\pi n e^2/m}$ is the plasma frequency, n and m are the electron density and mass, $E(0)$ is the electric field at the metal interface, and Z_D is the screening parameter, which will be determined subsequently. A simplified derivation of equation (8) is given in the Appendix. Equation (8) is the equation for the plasma oscillations of the carriers in the metal. The oscillations are induced by the electric field $E(0)$ at the interface. The presence of a metal boundary is reflected in (8) by the second partial derivative with respect to the coordinate z. In equation (8) the time dispersion of the carriers is taken into account by replacing the square of the wave frequency with $\omega^2 + i\omega\nu$, where ν is the collision frequency.

The screening parameter

$$Z_D = \frac{v_F}{\sqrt{\omega_0^2 - i\omega(\nu - i\omega)}}$$

is expressed simply in terms of the Debye screening length $r_D = v_F/\omega_0$, where v_F is the Fermi velocity:

$$Z_D \simeq r_D\left(1 + \frac{\omega^2}{\omega_0^2} + i\frac{\omega\nu}{\omega_0^2}\right). \qquad (9)$$

Formula (9) holds for $\omega \ll \omega_0$ and $\nu \ll \omega_0$.

The solution (8) consists of a general solution of the corresponding homogeneous equation and a particular solution of the inhomogeneous equation (8). The integration constants are determined from the conditions at the metal interface: (1) the conduction current is zero at $z = 0$; (2) the electric field is $E(0)$. The result of solving (8) is the distribution of the average electric field $\overline{E(z)}$, conduction current density $\overline{j(z)}$, and the volume charge density $\overline{\rho(z)}$:

$$\overline{E(z)} = \left(\frac{\omega_0 Z_D}{v_F}\right)^2 e^{-z/Z_D}E(0) + \left[1 - \left(\frac{\omega_0 Z_D}{v_F}\right)^2\right]E(0),$$

$$\overline{j(z)} = \frac{1}{4\pi} i\omega\left(\frac{\omega_0 Z_D}{v_F}\right)^2 (e^{-z/Z_D} - 1)E(0), \qquad (10)$$

$$\overline{\rho(z)} = -\frac{1}{4\pi}\frac{\omega_0^2 Z_D}{v_F^2} e^{-z/Z_D}E(0).$$

As was to be expected, an increased charge density forms near the superlattice surface, which decreases rapidly to a depth of the order of Z_D. This is a dynamic analog of electrostatic screening [13]. In the present study we shall not investigate the detailed characteristics of screening [16], since we are mainly interested in the electromagnetic losses inside the metal. The first two expressions in the system (10) deserve close attention. In addition to the field and current components associated with screening, penetrating components also exist. The penetrating component of the electric field is fairly small when ω, $\nu \ll \omega_0$,

$$\overline{E(z \longrightarrow \infty)} \simeq iE(0)\frac{2\omega\nu}{\omega_0^2}, \qquad (11)$$

but together with the penetrating current component, it makes a contribution to the losses which must be taken into account. It should be noted that the field and current structure in (10) does not reflect the skin effect. We recall the conditions required for the existence of the distribution (10) and the penetrating component (11): small metal thickness d compared with the skin layer depth δ, quasi-steady-conditions, and high superlattice

conductivity. Expressions (10) and (11) allow us to introduce the effective conductivity

$$\sigma_{\text{eff}} = \frac{\overline{j(z \longrightarrow \infty)}}{\overline{E(z \longrightarrow \infty)}} = \frac{ne^2}{m(\nu - i\omega)}.$$

Under the conditions of weak time dispersion ($\omega/\nu \ll 1$) achieved experimentally at room temperature, σ_{eff} is the dc conductivity. Bearing in mind that the surface resistance $R_s \simeq d\sigma_{\text{eff}}^{-1}$ is present in formula (3), we can conclude that the changes in the transmission coefficient T in the magnetic field are attributable to the magnetoresistance of the superlattice in the current-perpendicular-to-plane geometry.

We shall now calculate the electromagnetic losses and determine

$$Q_{\text{loss}} = P_{\text{loss}}/S = \text{Re}\int j^* E dV,$$

which are the Joule losses per unit sample area. Under the experimental conditions $\omega, \nu \ll \omega_0$ and

$$Q_{\text{loss}} \approx \frac{\omega^2 R_s}{8\pi^2} E^2(0). \tag{12}$$

The electromagnetic losses (12) are caused by the penetrating component of the electric field. The value of R_s on the right-hand side ensures a correlation between the magnetoresistance and the electromagnetic losses of the superlattice.

5. CONCLUSIONS

We have studied the giant magnetoresistance of Fe/Cr superlattices at microwave frequencies. We have established that there is a mutually single-valued correlation between the microwave transmission coefficient and the static magnetoresistance when current flows in the superlattice plane. This correlation is explained theoretically by solving the problem of the reflection and transmission of an electromagnetic wave through a thin metal object. It is shown that at room temperature the rf analog of the giant magnetoresistance effect is determined by the static conductivity at least up to frequencies of tens of Gigahertz.

Experimental measurements were made of the rf analog of giant magnetoresistance when an rf current flows perpendicular to the plane of the layers. It has been shown theoretically that a correlation exists between the microwave losses and the static magnetoresistance. The structure of the rf fields and currents averaged over the period of the superlattice has been calculated. It has been shown that penetrating components exist which induce microwave losses.

ACKNOWLEDGMENTS

This work was supported financially by the Russian Foundation for Basic Research (projects nos. 98-02-17553 and 98-02-17517) and was supported in part by INTAS (project no. 96-0531).

APPENDIX

We shall analyze an rf electric field $E_z(z) = E(z)$ in a thin metal film under quasi-steady-state conditions, $\omega \ll c/L_{\max}$, where $L_{\max}$ is the maximum dimension of the sample. We introduce the nonequilibrium correction $\varphi = \varphi(v, z)$ to the electron distribution function. We assume that φ depends on the coordinate z and electron velocity v. We write the transport equation in the τ approximation:

$$\frac{\partial\varphi}{\partial t} + v\frac{\partial\varphi}{\partial z} + evf_0' E = -\nu(\varphi - \bar{\varphi}), \tag{A.1}$$

where $f_0' = \partial f_0/\partial\varepsilon$, f_0 is the equilibrium distribution function, ν is the collision frequency, and $\bar{\varphi} = [\varphi(v) + \varphi(-v)]/2$. To simplify the calculations we make the model assumption that the electron velocity is constant at the Fermi surface $|v| = v_F$. Writing equation (A.1) for electrons having the velocities $v = v_F$ and $v = -v_F$ and taking into account the relations

$$\begin{aligned}\varphi(v_F) + \varphi(-v_F) &= -\frac{2v_F^2 m}{n}\frac{f_0'}{e}\rho,\\ \varphi(v_F) - \varphi(-v_F) &= -\frac{2v_F^2 m}{n}\frac{f_0'}{ev}j,\end{aligned} \tag{A.2}$$

where m and n are the electron mass and velocity, ρ is the charge density, and j is the current density.

Taking into account the current continuity equation, we can derive the following equation:

$$\frac{\partial j}{\partial t} + v_F^2\frac{\partial\rho}{\partial z} - \frac{ne^2}{m}E = -\nu j. \tag{A.3}$$

In our case it follows from $\text{div}\mathbf{E} = 4\pi\rho$ that

$$\frac{\partial E}{\partial z} = 4\pi\rho.$$

The conduction current at the metal interface should vanish, $j(0) = 0$. We then obtain

$$i\omega[E(z) - E(0)] = 4\pi j(z). \tag{A.4}$$

Substituting (A.4) into (A.3), we obtain

$$(\nu - i\omega)\frac{i\omega}{4\pi}[E(z) - E(0)] + \frac{v_F^2}{4\pi}\frac{\partial^2 E(z)}{\partial z^2} - \frac{ne^2}{m}E(z) = 0.$$

We introduce the notation: ω_0 is the plasma frequency, Z_D is the screening parameter,

$$\omega_0^2 = 4\pi ne^2/m,$$

$$Z_D = \frac{v_F}{\sqrt{\omega_0^2 - i\omega(\nu - i\omega)}} \simeq \frac{v_F}{\omega_0}\left[1 + i\frac{\omega(\nu - i\omega)}{\omega_0^2}\right]$$

assuming $\nu \ll \omega_0$, $\omega \ll \omega_0$. Finally the equation for the electric field has the form

$$\omega_0^2 E(z) - \omega_0^2 Z_D^2 \frac{\partial^2 E}{\partial z^2} = -(\omega^2 + i\omega\nu)E(0). \qquad \text{(A.5)}$$

This determines the electromagnetic oscillations under the action of the inducing field $E(0)$.

REFERENCES

1. H. N. Baibich, J. M. Broto, A. Fert, *et al.*, Phys. Rev. Lett. **61**, 2472 (1988).
2. M. A. Gijs, J. B. Giesbers, M. T. Johnson, *et al.*, J. Appl. Phys. **75**, 6709 (1994).
3. L. Piraux, S. Dubois, and A. Fert, J. Magn. Magn. Mat. **159**, L287 (1996).
4. J. J. Krebs, P. Lubitz, A. Chaiken, and J. A. Prinz, J. Appl. Phys. **69**, 4795 (1991).
5. J. C. Jackuet and T. Valet, Mater. Res. Soc. Symp. Proc. **384**, 477 (1995).
6. V. V. Ustinov, A. B. Rinkevich, L. N. Romashev, and V. I. Minin, J. Magn. Magn. Mat. **177–181**, 1205 (1998).
7. V. V. Ustinov, N. G. Bebenin, L. N. Romashev, *et al.*, Phys. Rev. B **54**, 15 958 (1996).
8. V. V. Ustinov, L. N. Romashev, V. I. Minin, *et al.*, Fiz. Met. Metalloved. **80**, 71 (1995).
9. A. B. Rinkevich and L. N. Romashev, Radiotekh. Élektron. (Moscow) **44**, 597 (1999).
10. J. L. Altman, *Microwave Circuits* (Van Nostrand, New York, 1964; Mir, Moscow, 1968).
11. E. A. Kaner and V. G. Skobov, Adv. Phys. **17**, 605 (1968).
12. S. M. Rytov, Zh. Éksp. Teor. Fiz. **29**, 605 (1955) [Sov. Phys. JETP **2**, 466 (1956)].
13. L. D. Landau and E. M. Lifshitz, *Course of Theoretical Physics, Vol. 8: Electrodynamics of Continuous Media* (Nauka, Moscow, 1982; Pergamon Press, New York, 1984).
14. A. G. Gurevich and G. A. Melkov, *Magnetic Oscillations and Waves* (Nauka, Moscow, 1994).
15. V. V. Ustinov, J. Magn. Magn. Mat. **165**, 125 (1997).
16. A. M. Gabovich, L. G. Il'chenko, *et al.*, Zh. Éksp. Teor. Fiz. **75**, 249 (1978) [Sov. Phys. JETP **48**, 124 (1978)].

Translation was provided by AIP

Journal of Experimental and Theoretical Physics, Vol. 90, No. 5, 2000, pp. 842–849.
Translated from Zhurnal Éksperimental'noĭ i Teoreticheskoĭ Fiziki, Vol. 117, No. 5, 2000, pp. 969–977.

SOLIDS
Electronic Properties

Influence of the Coulomb Blockade Effect on Heat Transfer in a One-Dimensional System of Spinless Electrons

M. V. Moskalets
e-mail: moskalets@kpi.kharkov.ua
Received September 28, 1999

Abstract—An analysis is made of some characteristics of the low-temperature thermal conductivity of a ballistic quantum dot, attributed to the influence of long-range Coulomb interaction in the geometric capacitance approximation. It is shown that at fairly low temperatures the thermal conductivity K exhibits Coulomb oscillations as a function of the electrostatic potential of the quantum dot. At the maximum of the Coulomb peak we find $K \propto T$ whereas at the minimum $K \propto T^3$. The dependence $K(T)$ is essentially nonmonotonic at temperatures corresponding to the characteristic spacing between the size-quantization levels in the quantum dot. © *2000 MAIK "Nauka/Interperiodica".*

1. INTRODUCTION

Electrostatic energy strongly influences charge transport in mesoscopic systems connected by tunneling junctions with the surroundings (supply conductors) at low temperatures [1–4]. Electron tunneling through a potential barrier is accompanied by a change in the charge of the mesoscopic sample by $1e$ and a change in the system energy by $E_c = e^2/(2C)$, where e is the electron charge and C is the electrostatic capacitance of the sample. At temperatures $T \ll E_c$ charge transport is generally strongly suppressed (Coulomb blockade effect) [5–10].[1] However, for certain values of the potential difference V_g between the sample and the surroundings the electrostatic energy of the system $E = (q - eN)^2/(2C)$ (where q is the charge of the sample; $N = CV_g/e$) is degenerate with respect to change in the charge by $1e$: $q \rightleftarrows q + e$ (this occurs for half-integer values of N) [7, 11]. In this case, the Coulomb blockade is broken and this is manifest as an appreciable increase in the conductance of the system.

A similar effect should be observed for other kinetic coefficients, in particular for the thermal conductivity K which is studied in the present paper. In a one-dimensional system the nontrivial appearance of the Coulomb blockade effect in the heat transport case consists in (neutral) electron–hole pairs making a considerable contribution to the heat transfer because of the following circumstance. It was shown in [12] that for a quantum dot (i.e., a phase-coherent mesoscopic sample connected by two quantum point contacts to supply conductors) in the strong tunneling regime, integration of the charge fluctuations of the quantum dot at low temperatures $T \ll E_c$ can reduce the problem of tunneling of spinless Fermi electrons through a double barrier to the tunneling of a Luttinger liquid [13–16] with $g = 1/2$ (where g is the Haldane parameter) through an isolated impurity [17]. Heat transfer in a Luttinger liquid with an isolated impurity was considered in [18]. It was shown that at low temperatures in addition to the electron (caused by electron tunneling) contribution to the thermal conductivity $K_e \propto T^{2/g-1}$, electron–hole pairs [plasmons, i.e., small-amplitude fluctuations of the boson (phase) field describing a Luttinger liquid] also make a significant contribution $K_p \propto T^3$. In the case $g = 1/2$ these contributions are of the same order of magnitude. It should be noted that the important role of plasmons in heat transfer in a Luttinger liquid was noted in [19–21].

Thus, heat transfer across a double potential barrier (quantum dot) under conditions when the electrostatic energy is substantial (i.e., when $T \ll E_c$) is accomplished by electrons and by neutral particles (plasmons).[2] At low temperatures both contributions are of the same order of magnitude. This factor can be used to develop a theory of heat transport based on a self-consistent harmonic approximation [22, 23]. This approximation in fact describes plasmon propagation. However, electron tunneling processes are also partly taken into account by renormalizing the potential barrier height.

It should be noted that studies of heat transport are important first from the point of view of observing the non-Fermi liquid behavior of an electron system [19–21]. Second, heating effects also influence the properties of

[1] For mesoscopic samples the capacitance may reach $C \leq 10^{-15}$ F which corresponds to $E_c \geq 1$ K.

[2] The present study only takes into account the long-range Coulomb interaction of electrons in the quantum dot described in the approximation of the geometric capacitance C. Allowance for short-range interelectron interaction ($g \neq 1$) like the spin will modify the dependence $K(T, V_g)$ and is not considered here. Note that the spinless electron model can be used in the presence of a strong magnetic field which polarizes the electron gas near the quantum dot.

mesoscopic systems which exhibit a Coulomb blockade effect (single-electron transistors and so on) [24], which requires a study of heat transfer processes in these systems.

In the present study we calculate the thermal conductivity K of a quantum dot as a function of temperature T and potential V_g. The dependence $K(V_g)$ contains peaks corresponding to destruction of the Coulomb blockade. The form of the Coulomb peak depends on temperature. The dependence $K(T)$ is essentially nonmonotonic at temperatures corresponding to the spacing between the size-quantization levels in a quantum dot.

2. FORMULATION OF THE PROBLEM AND BASIC EQUATIONS

We shall consider a one-dimensional ballistic channel containing spinless noninteracting electrons as our model. Two point potential barriers of height V_1 and V_2 positioned at points $x_1 = -d/2$ and $x_2 = d/2$ simulate the quantum dot. With respect to the rest of the channel, the quantum dot has the potential V_g which can be varied by using an additional metal electrode (gate). One-dimensional conductors corresponding to $x < x_1$ and $x > x_2$ connect the quantum dot to remote reservoirs having the temperature T and chemical potential μ. We shall neglect any inelastic processes in the system (quantum dot plus supply conductors) and we shall consider the quantum dot as a purely elastic scatterer. This holds at fairly low temperatures when the phase coherence length of the electrons $L_\phi(T)$ is greater than the distance between the electron reservoirs. Heat transport (like charge transport) is understood in the usual meaning for mesoscopic physics [25, 26] as transport between (remote) reservoirs where electron energy relaxation takes place. In calculations of the thermal conductivity in the linear approximation we assume that the electron reservoirs have the same chemical potentials $\mu_1 = \mu_2 = \mu$ and their temperatures differ by the small amount $\Delta T = |T_1 - T_2| \ll T$.

This highly simplified model can nevertheless allow for the influence of the electrostatic energy and the spatial quantization in the quantum dot [27] on the electron transport at low temperatures. Note that this model [28] corresponds to the experimental situation where a ballistic quantum dot is connected to supply conductors using single-mode quantum point contacts whose transmission coefficient may vary between zero and one.

As we know, in the one-dimensional case the electrostatic energy (in the geometric capacitance approximation) can be taken into account exactly (beyond the limits of perturbation theory) using the bosonization method [14, 15]. In order to describe the low-energy properties of the system ($\Delta\epsilon \ll \mu$) we can linearize the electron spectrum near the Fermi energy μ. In this case, the Lagrangian density of the spinless electrons has the form [17]

$$L_0 = \frac{\hbar v}{2g}\left\{\frac{1}{v^2}\left(\frac{\partial\theta(x,t)}{\partial t}\right)^2 - \left(\frac{\partial\theta(x,t)}{\partial x}\right)^2\right\}, \qquad (1)$$

where v, g are the Haldane parameters [16]. In the present case we confine our analysis to noninteracting electrons: $g = 1$; $v = v_F = \pi\hbar\rho_0/m^*$, where v_F is the Fermi velocity, ρ_0 is the average electron density, and m^* is the effective electron mass. The boson (phase) field $\theta(x, t)$ determines the deviation $\delta\rho$ of the electron density from the average density and the electron current j:

$$\delta\rho = \frac{1}{\pi^{1/2}}\frac{\partial\theta(x,t)}{\partial x}, \quad j = \frac{e}{\pi^{1/2}}\frac{\partial\theta(x,t)}{\partial t}.$$

The presence of potential barriers is taken into account by the following Lagrangian [17]:

$$\begin{aligned} L_V = &-V_1\cos(2\pi^{1/2}\theta(x,t) - k_F d)\delta(x + d/2) \\ &- V_2\cos(2\pi^{1/2}\theta(x,t) + k_F d)\delta(x - d/2), \end{aligned} \qquad (2)$$

where $k_F = \pi\rho_0$ is the Fermi wave number. We shall assume that in the energy range of interest to us $\Delta\epsilon \sim T \ll \mu$, the values of V_1 and V_2 do not depend on the electron energy. Following [12], we can assume $V_i = \mu\pi^{-1}r_i/t_i$ ($i = 1, 2$), where r_i and t_i are the moduli of the reflection coefficient and the transmission coefficient for electrons having the Fermi energy, which characterize a point potential barrier at the point x_i. Note that according to the Landauer–Büttiker approach [29], the conductance of an isolated barrier for the case of one-dimensional spinless noninteracting electrons is $G_i = G_0 t_i^2$, where $G_0 = e^2/(2\pi\hbar)$ is the conductance quantum. For interacting electrons ($g \neq 1$) this is not the case which leads to a temperature dependence of the conductance at low temperatures [17].

The electrostatic energy associated with the capacitance C of the quantum dot is described by

$$\int_{-\infty}^{\infty} L_C(x,t)dx = -E_c\left\{\int_{x_1}^{x_2}\delta\rho(x,t)dx - N\right\}^2. \qquad (3)$$

The partition function Z required to describe the properties of the system may be expressed as a functional integral:

$$Z = \int\exp\left(-\frac{S_E}{\hbar}\right)D\theta. \qquad (4)$$

The Euclidean (calculated at imaginary time $\tau = it$) action is

$$S_E = -\int_{-\infty}^{\infty} dx \int_0^{\beta} d\tau (L_0 + L_V + L_C),$$

where $\beta = \hbar/T$.

Since the nonquadratic part of the Lagrangian (L_V) with respect to θ only depends on the fields at two fixed points, we can integrate over the fluctuations of the field θ at all points apart from $x = x_1$ and $x = x_2$. As a result, we obtain the effective Euclidean action [27]:

$$\begin{aligned} S_{\text{eff}}[\theta, \phi] &= \frac{\hbar}{\beta} \sum_{n=-\infty}^{\infty} \left\{ \frac{2|\omega_n||\theta_n|^2}{1 + \exp(-|\omega_n|/\Delta\omega)} \right. \\ &\left. + \frac{|\omega_n||\phi_n|^2}{2[1 - \exp(-|\omega_n|/\Delta\omega)]} \right\} \\ &+ \int_0^{\beta} d\tau \{ V_1 \cos(\pi^{1/2}[2\theta(\tau) - \phi(\tau)] - k_F d) \\ &+ V_2 \cos(\pi^{1/2}[2\theta(\tau) + \phi(\tau)] + k_F d) \} \\ &+ \frac{E_c}{\pi} \int_0^{\beta} \{ \phi(\tau) - \pi^{1/2} N \}^2. \end{aligned} \tag{5}$$

Here we introduce the following notation: $\omega_n = 2\pi n/\beta$ is the Matsubara frequency (n is an integer); $\Delta\omega = v_F/d$; $\theta(\tau) = [\theta(x_2, \tau) + \theta(x_1, \tau)]/2$; $\phi(\tau) = \theta(x_2, \tau) - \theta(x_1, \tau)$; θ_n and ϕ_n are the coefficients of the Fourier series expansion:

$$x(\tau) = \frac{1}{\beta} \sum_n \exp(-i\omega_n \tau) x_n + \zeta_x$$

($x \equiv \theta, \phi$; ζ_x is the zeroth-order mode). Note that the field $\phi(\tau)$ determines the excess charge of the quantum dot $\delta q = e\pi^{-1/2}\phi$ and the field $\theta(\tau)$ determines the transmitted current:

$$I = \frac{ie}{\pi^{1/2}} \frac{\partial \theta(\tau)}{\partial \tau}$$

[17, 27].

We then obtain an approximate expression for S_{eff} which corresponds to small-amplitude oscillations of the boson field $\theta(\tau)$. For this we first assume that V_1 and V_2 are small (i.e., $V_1, V_2 \ll \mu$) and we integrate over charge fluctuations of the quantum dot (over fluctuations of the field ϕ).

If the condition

$$T \ll E_c \tag{6}$$

is satisfied, we obtain the effective action in the following form:

$$\begin{aligned} S'_{\text{eff}}[\theta] &= \frac{\hbar}{\beta} \sum_n \frac{2|\omega_n|}{1 + \exp(-|\omega_n|/\Delta\omega)} |\theta_n|^2 \\ &+ V \int_0^{\beta} d\tau \cos(2\pi^{1/2}\theta(\tau)). \end{aligned} \tag{7}$$

The value of V is defined as

$$\begin{aligned} V &= \left(\frac{2\gamma E_c}{\pi\mu} \right)^{1/2} \\ &\times [V_1^2 + V_2^2 + 2V_1 V_2 \cos(2\pi N + 2k_F d)]^{1/2}, \end{aligned} \tag{8}$$

where $\gamma = e^C$, $C \approx 0.5772$ is the Euler constant. We shall subsequently only allow for small fluctuations of the field θ and set $\cos(2\pi^{1/2}\theta) \approx 1 - 2\pi\theta^2$. We then need to renormalize the potential $V \longrightarrow V^*$ [22, 23] after integrating over high-frequency fluctuations whose energy $\hbar\omega_N$ exceeds the renormalized potential V^*. This is possible since these fluctuations are not sensitive to the potential V^* and may be considered as free-field fluctuations. The final expression for the effective action in the self-consistent harmonic approximation has the following form:

$$S_{\text{SCHA}}[\theta] = \frac{\hbar}{\beta} \sum_n \left\{ \frac{2|\omega_n|}{1 + \exp(-|\omega_n|/\Delta\omega)} - \omega_V \right\} |\theta_n|^2. \tag{9}$$

The value of ω_V is obtained from

$$\hbar\omega_V = 2\pi V^2/\mu, \tag{10}$$

where the Fermi energy μ plays an rf cutoff role. We shall use this action (9) to describe heat transport across a quantum dot assuming that the results will be valid for any (and not only small) value of the scattering potentials V_1 and V_2. As was noted in the Introduction, this approximation describes plasmon transport which in this particular case yields the same temperature dependence of the thermal conductivity as for electron transport ($K \propto T^3$ for $V \gg T$ and $K \propto T$ for $V = 0$).

We shall analyze heat transfer Q across a quantum dot in the linear regime characterized by the thermal conductivity $K = -Q/\Delta T$, where $\Delta T = T_2 - T_1 \ll T$ is the temperature difference between the electron reservoirs to the left ($x < x_1$) and right ($x > x_2$) of the quantum dot ($\mu_1 = \mu_2 = \mu$) [26]. It was shown in [18] that the thermal conductivity is expressed in terms of the electrical conductance $G(\omega)$:

$$K = \frac{\hbar^3}{4e^2T^2} \int_0^{\infty} d\omega \frac{\omega^2 \text{Re}G(\omega)}{\sinh^2(\hbar\omega/(2T))}, \tag{11}$$

where $\text{Re}G(\omega)$ is the real part of $G(\omega)$.

In order to calculate the dependence $G(\omega)$ in the model (9), we use the Kubo formula [30, 31]:

$$G(\omega) = \frac{-i}{\hbar\omega}\lim_{\omega_n \to -i\omega}\int_0^\beta d\tau \exp(i\omega_n\tau)\langle I(\tau)I(0)\rangle. \quad (12)$$

Substituting the formula for the current I expressed in terms of $\theta(\tau)$we obtain

$$G(\omega) = \frac{e^2T}{\pi\hbar^2}\lim_{\omega_n \to -i\omega}\omega_n\Phi_n,$$

where $\Phi_n = \langle\theta_n\theta_{-n}\rangle$. Averaging

$$\Phi_n = \frac{1}{Z}\int\theta_n\theta_{-n}\exp(-S_E/\hbar)D\theta$$

can easily be performed if the following fictitious term is added to the action

$$S_j = \hbar\sum_n j_n\theta_{-n}.$$

Then we have

$$\Phi_n = \frac{1}{Z}\left(\frac{\delta^2 Z}{\partial j_n \delta j_{-n}}\right)_{j=0}.$$

Direct calculations using the effective action (9) give

$$\Phi_n = \frac{\beta}{2}\left[\frac{2\omega_n}{1+\exp(-\omega_n/\Delta\omega)} - \omega_V\right]^{-1}.$$

Finally the rf conductivity of the quantum dot in this model is given by

$$\mathrm{Re}G(\omega) = G_0\frac{\omega^2\left(1+\cos\frac{\omega}{\Delta\omega}\right)}{\omega_V^2\left(1+\cos\frac{\omega}{\Delta\omega}\right)+2\omega_V\omega\sin\frac{\omega}{\Delta\omega}+2\omega^2}. \quad (13)$$

Expressions (13) and (11) determine the thermal conductivity of the quantum dot allowing for the Coulomb blockade effect in the self-consistent harmonic approximation. The dependence $K(T, \omega_V)$ will be analyzed in the following section.

It should be stated that the expression obtained for $G(\omega)$ describes the plasmon contribution to the conductivity of a double barrier (a similar expression for a single barrier was given in [22]). In the limit $\omega \longrightarrow 0$ this contribution vanishes since these plasmons (small-amplitude oscillations of the boson field corresponding to electron–hole pairs) are neutral particles and do not carry charge [22]. In the formalism used, topological excitations of the boson (phase) field [16] carry charge which corresponds to an appreciable change in the value of θ (transport of a single electron across the barrier corresponds to a change in θ by $\pi^{1/2}$ [17]). Note that expression (13) was obtained assuming that the characteristic energy scale (T or $\hbar\omega$) is much less than E_c (6).

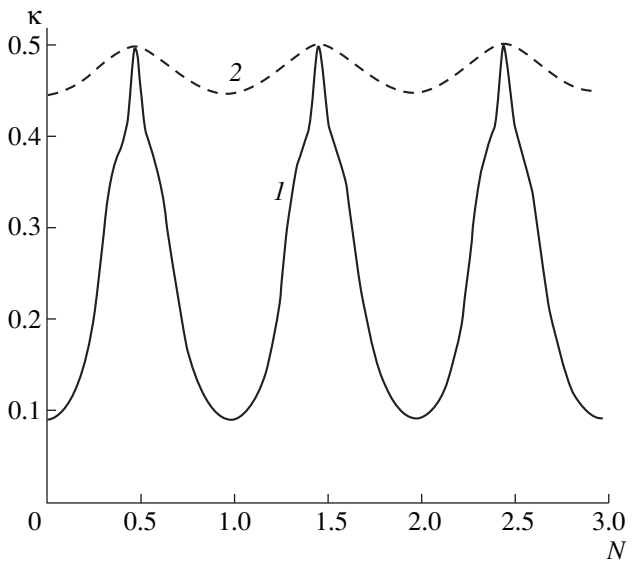


Fig. 1. Dependences of the normalized thermal conductivity $\kappa = K/K_0$ ($K_0 = \pi T/(6\hbar)$) on the quantum dot potential $N = CV_g/e$ for low ($\hbar\omega_{V0}/T = 10$, curve *1*) and high ($\hbar\omega_{V0}/T = 0.1$, curve *2*) temperatures. The curves were plotted for a symmetric quantum dot ($V_1 = V_2$) for $T = \Delta_F$.

3. THERMAL CONDUCTIVITY OF A SYMMETRIC QUANTUM DOT

In this section we consider the case where the potential barriers separating the quantum dot from the supply conductors have the same height: $V_1 = V_2$, $r_i = r$, $t_i = t$ ($i = 1, 2$). In this case we have

$$\omega_V = \omega_{V0}\cos^2(\pi N + k_F d),$$
$$\hbar\omega_{V0} = \frac{16\gamma E_c}{\pi^2}\left(\frac{r}{t}\right)^2. \quad (14)$$

It can be seen from this expression that for certain values of $N(V_g)$ the effective potential barrier disappears ($\omega_V = 0$) as a result of the energy degeneracy of the system with respect to a change in the number of particles in the quantum dot by one ($q \rightleftharpoons q+1$) [7, 11] and corresponds to destruction of the Coulomb blockade. As a result, the heat flux increases appreciably and this is observed as a series of peaks on the dependence $K(V_g)$ (Fig. 1). This effect is exactly the same as the conductance oscillations [1–4].

For the case of an asymmetric quantum dot ($V_1 \neq V_2$) the value of V (8) does not vanish for any V_g so that the oscillations on the dependence $K(V_g)$ are weak.

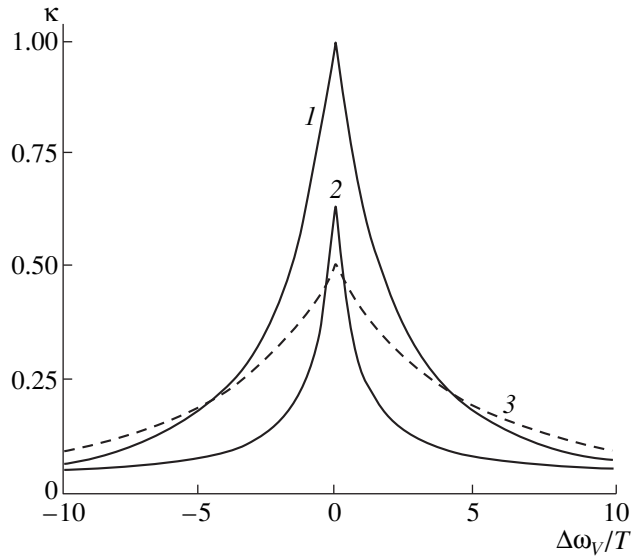


Fig. 2. Dependences of the normalized thermal conductivity κ on the potential barrier height for a symmetric quantum dot. The value of $\Delta\omega_V$ is measured from the value of ω_V corresponding to the position of the maximum. The curves correspond to $T \ll T^*$ (*1*); $T = 2T^*$ (*2*); $T \gg T^*$ (*3*) ($T^* = \Delta_F/\pi^2$).

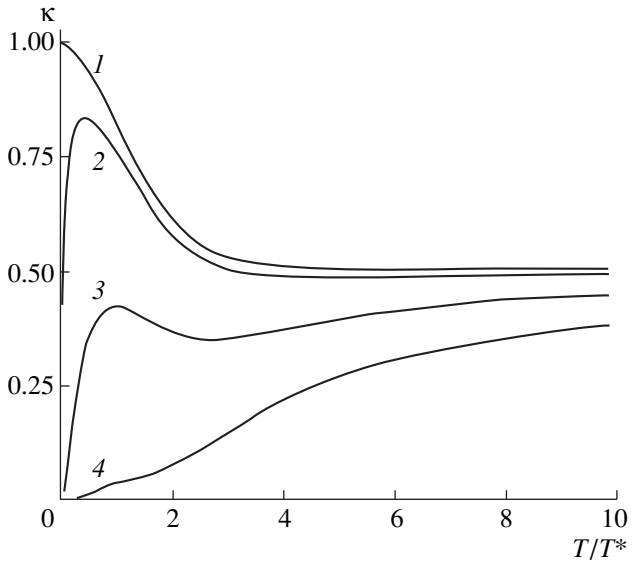


Fig. 3. Temperature dependences of the normalized thermal conductivity for a symmetric quantum dot for $\omega_V = 0$ (*1*); $\hbar\omega_V/T^* = 0.1$ (*2*), 1 (*3*), and 10 (*4*).

3.1. Shape of Coulomb Peak

The shape of the peak on the dependence $K(V_g)$ depends strongly on temperature (see Fig. 2). A crossover takes place at $T \sim T^*$ where $T^* = \Delta_F/\pi^2$; $\Delta_F = \pi\hbar v_F/d$ is the spacing between the spatial quantization levels near the Fermi energy μ in an isolated ($V_{1,2} \longrightarrow \infty$) quantum dot. Near the maximum of the peak ($\hbar\omega_V \ll T$) the thermal conductivity is given by

$$\frac{K}{K_0} = 1 - \frac{3\hbar\omega_V}{2\pi T}, \quad \hbar\omega_V \ll T \ll T^*, \tag{15}$$

$$\frac{K}{K_0} = \frac{1}{2} - \frac{3\hbar\omega_V}{2^{5/2}\pi T}, \quad T \gg T^*, \quad \hbar\omega_V. \tag{16}$$

Here $K_0 = \pi T/(6\hbar)$ is the thermal conductivity of a one-dimensional ballistic channel [32, 33]. Thus, as the temperature increases, the thermal conductivity at the peak maximum is halved. This is attributable to the influence of the electrostatic energy (capacitance C) which is responsible for the frequency dispersion of the conductance $G(\omega)$ (13).[3] At low temperatures ($T \ll T^*$) the main contribution to the thermal conductivity is made by low-frequency (long-wavelength) plasmons which do not "sense" the internal structure of the double barrier. In this case, the thermal conductivity of the system (at the peak maximum) is determined by the thermal conductivity of the one-dimensional ballistic channel. As the temperature increases, as a result of the destructive interference of plasmon contributions at different frequencies the thermal conductivity is halved. This can effectively be considered to be the result of the incoherent (at $T \gg T^*$) propagation of plasmons through two barriers. In this case, at the peak maximum ($\omega_V = 0$) we have two series-connected incoherent (classical) contacts (barriers having the transmission coefficient $t = 1$) each characterized by the thermal conductivity K_0. In this case, the thermal conductivity of the system will be $K = K_0/2$ (for similar reasoning on the electrical conductance see [12, 17]). However, it should be stressed that the halving of the thermal conductivity is caused by averaging over temperature in the phase-coherent system in the absence of real inelastic processes (which take place far from the system in the electron reservoirs). The destructive interference effect with increasing temperature is fairly general for the mesoscopic physics of ballistic structures. The characteristic energy scale $T^* \sim \Delta_F/\pi^2$ was first introduced in the persistent current problem [34]. This energy scale is also important for describing the kinetic properties of ballistic mesoscopic samples [35].

3.2. Temperature Dependence of the Thermal Conductivity

Figure 3 shows temperature dependences of the thermal conductivity for various values of the effective potential barrier. It can be seen that for a small barrier

$$\hbar\omega_V \sim T^*, \tag{17}$$

the dependence $K(T)$ is essentially nonmonotonic for $T \sim T^*$ because of the relative influence of two effects.

[3] It should be stated that for $E_c = 0$ and $\hbar\omega_V = 0$ we obtained $G(\omega) = G(0)$.

First the destructive interference of plasmon contributions at different energies leads to a reduction in K with increasing temperature. Second, an increase in the thermal conductivity is caused by an increase in the number of above-barrier (ballistic) plasmons with increasing temperature. Note that the condition (17) can only be satisfied in the immediate vicinity of the maximum of the $K(V_g)$ peak or for barriers having a low reflection coefficient ($r \longrightarrow 0$; $t \sim 1$) (strong tunneling). In this last case, the dependence $K(V_g)$ only contains weakly defined oscillations (see Fig. 1, curve *2*).

Analytic expressions for the dependence $K(T)$ can be obtained for $\omega_V = 0$ and $\hbar\omega_V \gg T$.

3.2.1. Thermal conductivity at the maximum of the coulomb peak. Assuming that $\omega_V = 0$, we obtain from (13) and (11)

$$\frac{K}{K_0} = \frac{1}{2}\left[1 + 3\frac{(T/T^*)\tanh(T/T^*) - 1}{\sinh^2(T/T^*)}\right], \qquad (18)$$

where $T^* = \Delta_F/\pi^2$. The dependence $K(T)$ is plotted in Fig. 3 (curve *1*). Note that a similar crossover for the conductance at the maximum of the Coulomb peak (at $g = 1$) was obtained in [27].

3.2.2. Thermal conductivity far from the maximum of the coulomb peak. We shall now assume that the following condition is satisfied

$$\hbar\omega_V \gg T, T^*. \qquad (19)$$

In this case the main heat transfer mechanism is plasmon tunneling. For $T \ll T^*$ expression (13) yields $\mathrm{Re}\,G(\omega) = G_0(\omega/\omega_V)^2$. Substituting into (11), we obtain

$$K(T) = K_0\frac{4\pi^2}{5}\left(\frac{T}{\hbar\omega_V}\right)^2, \quad T \ll T^*. \qquad (20)$$

A similar expression was first obtained in [18, 21] for a single potential barrier.

At higher temperatures ($T \gg T^*$) the thermal conductivity is strongly influenced by an effect involving the resonant tunneling of plasmons through the quantum dot. The importance of allowing for this effect was emphasized in [20, 21].

Resonant tunneling occurs for plasmons of frequency $\omega \simeq \omega_n^r = \omega_n[1 + 2\Delta_F/(\pi\hbar\omega_V)]$, where $\hbar\omega_n = \Delta_F(2n + 1)$, $n = 0, 1, \ldots$. Under condition (19) expression (13) may be represented as a sum of Breit–Wigner resonances and a quadratic background in terms of frequency:

$$\mathrm{Re}\,G(\omega) = G_0\left\{\sum_{n=0}^{\infty}\frac{\Gamma_n^2}{(\omega - \omega_n^r)^2 + \Gamma_n^2} + \left(\frac{\omega}{\omega_V}\right)^2\right\}, \qquad (21)$$

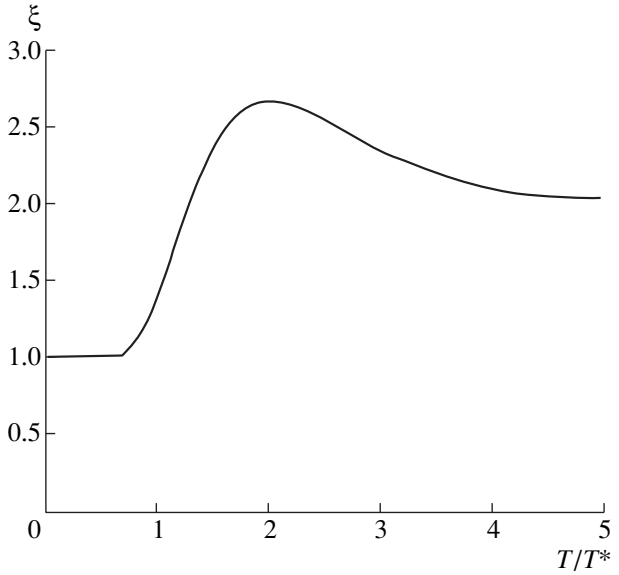


Fig. 4. Temperature dependence of the coefficient ξ in the expression $K/K_0 = (4\pi^2/5)\xi(T)(T/\hbar\omega_V)^2$ for the thermal conductivity of a symmetric quantum dot far from the maximum of the Coulomb peak.

where the resonance width is

$$\Gamma_n = \frac{2\Delta_F}{\pi}\left(\frac{\omega_n}{\omega_V}\right)^2. \qquad (22)$$

Substituting (21) into (11), we obtain

$$K = K_0\left\{\frac{4\pi^2}{5}\left(\frac{T}{\hbar\omega_V}\right)^2 + \frac{24T\Delta_F}{(\pi\hbar\omega_V)^2}\sum_{n=0}^{\infty}\frac{[\omega_n/(2T)]^4}{\sinh^2[\omega_n/(2T)]}\right\} \qquad (23)$$

(here we neglected the negligible difference between ω_n and ω_n^r). In the limit $T \gg T^*$ we can substitute

$$\sum_n \ldots \longrightarrow \int \ldots \frac{d\omega}{2\Delta_F},$$

which gives

$$K(T) = K_0\frac{8\pi^2}{5}\left(\frac{T}{\hbar\omega_V}\right)^2, \quad T \gg T^*. \qquad (24)$$

On comparing expressions (20) and (24) we can see that in the plasmon tunneling regime, $K \sim T^3$ is obtained over the entire temperature range. However, the proportionality factor is doubled for $T \sim T^*$ (see Fig. 4) as a result of the resonant tunneling effect.

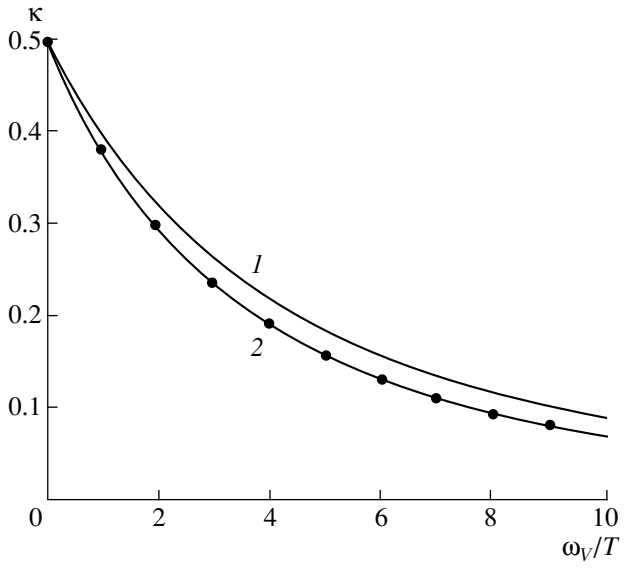


Fig. 5. Dependence of the normalized thermal conductivity κ for a symmetric quantum dot in the strong cotunneling regime ($E_c \gg T \gg \Delta_F$; $r \longrightarrow 0$) on the ratio $\hbar\omega_V/T$. The thermal conductivity was calculated using the self-consistent harmonic approximation (curve *1*), the exact expression for the conductance obtained in [12] (curve *2*), and the self-consistent harmonic approximation using a refined value of the renormalized potential $\omega_V^* = 1.2\omega_V$ (circles on curve *2*).

4. DISCUSSION AND CONCLUSIONS

In the present study we have considered the thermal conductivity of a quantum dot in the Coulomb blockade regime at low temperatures for the case of spinless electrons. The thermal conductivity was calculated using a self-consistent harmonic approximation [22, 23] which describes the plasmon heat conduction mechanism. It was shown that the dependence of the thermal conductivity on the potential of the quantum dot contains peaks caused by destruction of the Coulomb blockade.

At the maximum of the Coulomb peak, the thermal conductivity is linear with respect to temperature $K \propto T$ as a result of the ballistic heat transfer regime. However, the proportionality factor is halved for $T > T^*$ (18) because of a transition from coherent plasmon propagation through a two-barrier potential to incoherent propagation through two series-connected barriers. This is consistent with the behavior of the conductance at the maximum of the Coulomb peak (for $g = 1$) [12, 27].

In the plasmon tunneling regime ($\hbar\omega_V \gg T$) the thermal conductivity is $K \propto T^3$ (20), (24). We shall compare the plasmon contribution to the electron tunneling contribution. At low temperatures in the Coulomb blockade regime the dominant electron transport mechanism is elastic ($T \ll \Delta_F$) and inelastic ($T \gg \Delta_F$) cotunneling [8, 9]. Both in the case of a weakly reflecting potential ($r \longrightarrow 0$; $t \longrightarrow 1$) [12] and for a potential having the transmission coefficient $t \longrightarrow 0$ [8, 9] inelastic cotunneling processes lead to a quadratic dependence of the conductance on temperature $G \propto T^2$ (for $T \ll \hbar\omega_V$) which corresponds to the electron contribution to the thermal conductivity $K_e \propto T^3$. This temperature dependence was obtained in our study (24). At lower temperatures ($T \ll T^*$) when the dominant charge transport mechanism is elastic electron cotunneling, the electrical conductance does not depend on temperature $G \propto (\Delta_F/E_c)^2$. In this case the electron contribution to the thermal conductivity exceeds the plasmon contribution $K_p \propto (T/E_c)^2$ (20): $K_e/K_p \sim (\Delta_F/T)^2 \gg 1$. For $T \sim \Delta_F$ (more accurately $T \sim T^*$) these contributions are comparable and the nonmonotonicity caused by the resonant plasmon tunneling (see Fig. 4) may be observed in the total thermal conductivity ($K = K_e + K_p$).

In the strong cotunneling limit ($r \longrightarrow 0$; $E_c \gg T \gg \Delta_F$) the conductance of the quantum dot can be calculated exactly [12]. Substituting this expression into (11), we obtain the thermal conductivity in the strong cotunneling regime (SC):

$$K_{\mathrm{SC}} = \frac{\hbar^2}{16\pi T^2}\int_0^\infty d\omega \frac{\omega^2}{\sinh^2(\hbar\omega/(2T))} \times\left\{1 + \frac{1}{\hbar\omega}\int_{-\infty}^{\infty} d\epsilon \frac{\Gamma_0^2}{\epsilon^2+\Gamma_0^2}[f_0(\epsilon+\hbar\omega) - f_0(\epsilon)]\right\}, \tag{25}$$

where $f_0(x) = [\exp(\beta x) + 1]^{-1}$ is the Fermi function,

$$\Gamma_0 = \frac{2\gamma E_c}{\pi^2}[r_1^2 + r_2^2 + 2r_1 r_2 \cos(2\pi N + 2k_F d)].$$

We compare the coefficient (25) with the thermal conductivity in the self-consistent harmonic approximation (11), (13) (it should be borne in mind that $\hbar\omega_V = 2\Gamma_0$). Figure 5 gives dependences of K_{SC} (curve *1*) and K (curve *2*) on the ratio $\hbar\omega_V/T$. It can be seen that these curves show fairly good agreement. For example, for $T \gg \Gamma_0$ we have

$$K_{\mathrm{SC}} = \frac{K_0}{2}\left(1 - \frac{3\pi\Gamma_0}{16T}\right).$$

On comparing with (16) we can see that the deviation from the ballistic value is described by the plasmon approximation with a relative accuracy of around 15%. At low temperatures $\Delta_F \ll T \ll \Gamma_0$, the thermal conductivity is

$$K_{\mathrm{SC}} = K_0\frac{3\pi^2}{10}\left(\frac{T}{\Gamma_0}\right)^2,$$

and a comparison with (24) shows that the accuracy is around 30%. The agreement between the plasmon approximation and (25) can be improved by introducing the correction factor $a \approx 1.2$ in the definition of the

renormalized potential (10) $\hbar\omega_V^* = 2\pi a V^2/\mu$, which in fact implies a negligible change in the rf cutoff. The resulting dependence $K(\hbar\omega_V^*/T)$ is shown by the circles in Fig. 5.

REFERENCES

1. D. V. Averin and K. K. Likharev, in *Mesoscopic Phenomena in Solids*, Ed. by B. Altshuler, P. A. Lee, and R. A. Webb (North-Holland, Amsterdam, 1991), p. 176.
2. *Single Charge Tunneling*, Ed. by H. Grabert and M. H. Devoret (Plenum, New York, 1992).
3. M. A. Kastner, Rev. Mod. Phys. **64**, 849 (1992).
4. L. P. Kouwenhoven, C. M. Marcus, P. L. McEuen, S. Tarucha, R. M. Westervelt, and N. S. Wingreen, in *Mesoscopic Electron Transport*, Ed. by L. Sohn, L. P. Kouwenhoven, and G. Schön, Vol. 345 of NATO ASI E: Applied Sciences (Kluwer, Dordrecht, 1997).
5. R. I. Shekhter, Zh. Éksp. Teor. Fiz. **63**, 1410 (1972) [Sov. Phys. JETP **36**, 747 (1973)].
6. I. O. Kulik and R. I. Shekhter, Zh. Éksp. Teor. Fiz. **68**, 623 (1975) [Sov. Phys. JETP **41**, 308 (1975)].
7. L. I. Glazman and R. I. Shekhter, J. Phys.: Condens. Matter. **1**, 5811 (1989).
8. D. V. Averin and A. A. Odintsov, Zh. Éksp. Teor. Fiz. **96**, 1349 (1989) [Sov. Phys. JETP **69**, 766 (1989)]; D. V. Averin and A. A. Odintsov, Phys. Lett. A **140**, 251 (1989).
9. D. V. Averin and Yu. V. Nazarov, Phys. Rev. Lett. **65**, 2446 (1990).
10. L. I. Glazman and K. A. Matveev, Zh. Éksp. Teor. Fiz. **98**, 1834 (1990) [Sov. Phys. JETP **71**, 1031 (1990)].
11. H. van Houten and C. W. J. Beenakker, Phys. Rev. Lett. **63**, 1893 (1989).
12. A. Furusaki and K. A. Matveev, Phys. Rev. B **52**, 16676 (1995).
13. J. M. Luttinger, J. Math. Phys. **15**, 609 (1963).
14. Sólyom, Adv. Phys. **28**, 201 (1979).
15. V. J. Emery, in *Highly Conducting One-Dimensional Solids*, Ed. by J. T. Devreese (Plenum, New York, 1979).
16. F. D. M. Haldane, J. Phys. C **14**, 2585 (1981).
17. C. L. Kane and M. P. A. Fisher, Phys. Rev. B **46**, 15233 (1992).
18. C. L. Kane and M. P. A. Fisher, Phys. Rev. Lett. **76**, 3192 (1996).
19. R. Fazio, F. W. J. Hekking, and D. E. Khmelnitskii, Phys. Rev. Lett. **80**, 5611 (1998).
20. I. V. Krive, Fiz. Nizk. Temp. **24**, 498 (1998) [Low Temp. Phys. **24**, 377 (1998)].
21. I. V. Krive, Phys. Rev. B **59**, 12 338 (1999).
22. F. Guinea, G. Gómez Santos, M. Sassetti, and M. Ueda, Europhys. Lett. **30**, 561 (1995).
23. M. P. A. Fisher and W. Zwerger, Phys. Rev. B **32**, 6190 (1985).
24. V. A. Krupenin, S. V. Lotkhov, H. Scherer, *et al.*, Phys. Rev. B **59**, 10778 (1999).
25. M. Büttiker, Y. Imry, R. Landauer, and S. Pinhas, Phys. Rev. B **31**, 6207 (1985).
26. U. Sivan and Y. Imry, Phys. Rev. B **33**, 551 (1986).
27. A. Furusaki and N. Nagaosa, Phys. Rev. B **47**, 3827 (1993).
28. K. Flensberg, Phys. Rev. B **48**, 11156 (1993).
29. R. Landauer, Philos. Mag. **21**, 863 (1970); M. Büttiker, Phys. Rev. Lett. **57**, 1761 (1986).
30. L. D. Landau and E. M. Lifshitz, *Statistical Physics* (Nauka, Moscow, 1976; Pergamon Press, Oxford, 1980), Chap. 1.
31. E. M. Lifshitz and L. P. Pitaevskiĭ, *Physical Kinetics* (Nauka, Moscow, 1979; Pergamon Press, Oxford, 1981).
32. H.-L. Engquist and P. W. Anderson, Phys. Rev. B **24**, 1151 (1981).
33. L. W. Molenkamp, Th. Gravier, H. van Houten, *et al.*, Phys. Rev. Lett. **68**, 3765 (1992).
34. H. F. Cheung, Y. Gefen, E. K. Riedel, and W. H. Shih, Phys. Rev. B **37**, 6050 (1988).
35. M. V. Moskalets, Zh. Éksp. Teor. Fiz. **114**, 1827 (1998) [JETP **87**, 991 (1998)].

Translation was provided by AIP

Journal of Experimental and Theoretical Physics, Vol. 90, No. 5, 2000, pp. 850–860.
Translated from Zhurnal Éksperimental'noĭ i Teoreticheskoĭ Fiziki, Vol. 117, No. 5, 2000, pp. 978–989.

SOLIDS
Electronic Properties

Conductivity of a Two-Dimensional System with a Periodic Distribution of Circular Inclusions

B. Ya. Balagurov and V. A. Kashin*
Emanuel Institute of Biochemical Physics, Russian Academy of Sciences, Moscow, 117997 Russia
**e-mail: vkashin@deom.chph.ras.ru*
Received September 29, 1999

Abstract—A systematic method of solving the problem of the conductivity of a two-dimensional model with circular inclusions forming a square lattice is proposed. The complex potential outside the inclusions is expressed in terms of the Weierstrass zeta function and its derivatives. For a low concentration of inclusions an analytic expression is obtained for the effective conductivity σ_e which is a rapidly converging series in powers of the concentration. A numerical analysis of general formulas is used to determine σ_e and tabulate this in graphical form over the entire range of variation of the parameters appearing in the problem. The vicinity of the metal–insulator phase transition point is studied and the corresponding critical indices are estimated with an exact value being obtained for one of these. Four two-parameter functions contained in the expressions for the effective Hall coefficient and the effective magnetoresistance are also determined and tabulated so that it is possible to give a complete description of the galvanomagnetic properties of this system in a weak magnetic field.

1. INTRODUCTION

A theoretical study of various physical properties (in particular the electrical conductivity) of inhomogeneous media, especially disordered ones, encounters well-known mathematical difficulties. Hence, very few nontrivial analytic results are available in this field and they mainly apply to the two-dimensional case. In particular, these include the reciprocity relation [1, 2], and also the conductivity [2], galvanomagnetic [3] and thermoelectric [4, 5] characteristics of two-dimensional systems of critical composition (having the concentration $p = 1/2$). It was also shown in [6, 7] that the results of [3–5] can be applied to the case of arbitrary concentrations, for which it is sufficient to know the dimensionless effective conductivity $f(p, h)$ (where h is the conductivity ratio of the components) of the corresponding two-dimensional two-component model. However, for randomly inhomogeneous systems the function $f(p, h)$ is only known as the result of numerical analyses over the entire range of variation of the arguments p and h. A similar numerical analysis is required for each macroscopically nonequivalent disordered system.

Periodic models, for which the problem is simplified considerably, being reduced to finding the potential within a single unit cell, are more interesting. Various exact results have been obtained for these structures, i.e., two-dimensional systems with a regular distribution of similar inclusions (see, for example, [8, 9]). Various models with dielectric or ideally conducting inclusions were analyzed in [8, 9] allowing the solution to be confined to the external problem. A closed solution of a considerably more complex problem for the case where both components have finite (nonzero) conductivity is only known for a single model, comprising a two-dimensional system with a checkerboard structure [8].

An exact solution for a checkerboard-structure system can be used to find various effective characteristics (conductivity, galvanomagnetic and thermoelectric properties, and so on) of this model as a function of the parameter $h = \sigma_2/\sigma_1$ (σ_i is the conductivity of the ith component) for a fixed concentration $p = 1/2$, the same as the critical concentration. This then makes it possible to study the metal–insulator phase transition in the system merely in terms of one of two parameters (p or h). At the same time, the effective conductivity of a two-dimensional system with square inclusions (dielectric or ideally conducting) distributed in a checkerboard pattern was determined in [9] (see also [8]). This can be used to study the critical behavior of the electrical conductivity as a function of concentration for a fixed argument h ($h = 0$ or $h = \infty$). Thus, the results of [8, 9] for a checkerboard-structure model can be used to obtain some representation (albeit incomplete) of the critical conductivity behavior of this system. However, the absence of a unified approach to the analysis of the vicinity of the phase transition point, the impossibility of studying the entire range of variation of the parameters p and h, and also some characteristics associated with the angularity of the inclusions (see [9]) reduce the value of this model.

Of considerable interest for the theory of transport phenomena in inhomogeneous media is a detailed study of a two-dimensional system with periodically distributed circular inclusions. Although this model is one of the "simplest" and has already been studied by Rayleigh [10], no closed exact solution exists for it. Also no sufficiently comprehensive study of this system has been

made over the entire range of parameters p and h. However, a comprehensive study of this model is interesting for various reasons. First, it can be used to study a metal–insulator phase transition in a regular structure with smooth inclusions. Second, this model being the simplest and most studied, serves as the "touchstone" for various approximate methods. Finally, these structures are widely used in microelectronics so that their study is necessary from the practical point of view.

In the present paper we propose a systematic method of calculating the effective electrical characteristics of a two-dimensional system with a doubly periodic distribution (at the apexes of a square lattice) of circular inclusions. The method is based on expanding the solution of the Laplace equation in terms of the formally small parameter R/a, where R is the radius of the inclusion and $2a$ is the size of the unit cell. In the lowest approximation with respect to R/a, the potential outside the circle is sought as the sum of the potentials from dipoles induced by an external homogeneous electric field at each inclusion. The corresponding sum has a particularly simple form when a complex representation is used: the dipole complex potential is in fact the same as the Weierstrass zeta function $\zeta(z)$. In this representation the algorithm to allow for multipoles of any (in this case, odd) orders in successive approximations is also extremely simple since it reduces to an even number of differentiations of the function $\zeta(z)$. In view of the properties of the Weierstrass functions (see [11–13]), the potential thus obtained automatically satisfies all the necessary conditions at the boundaries of the unit cell.

The standard boundary conditions for $r = R$ yield an infinite system of equations for the unknown coefficients contained in the general expression for the potential. For $R/a \ll 1$ this system can easily be solved iteratively, which gives rapidly converging expansions in powers of the parameter R/a (or in powers of the concentrations $c = \pi R^2/(2a)^2$) for the conductivity and other effective characteristics of the model. For fairly large R/a we used numerical methods; instead of an infinite system we solved a subsystem of ten equations and near the phase transition point (i.e., for $R/a \longrightarrow 1$) this subsystem was increased to forty equations. As a result of this numerical analysis, the dimensionless effective conductivity $f(p, h)$, its derivative with respect to the argument h, and various other characteristics of this model were obtained with a high degree of accuracy over the entire range of R ($0 \le R \le a$) for various values of the parameter h. In particular, we investigated the vicinity of the metal–insulator phase transition point and estimated both critical indices. In addition, we obtained an exact value for the index t: $t = 1/2$.

In accordance with [6, 7], for two-dimensional two-component isotropic systems a knowledge of the function $f(p, h)$ over the entire range of variation of the arguments p and h is sufficient to describe the galvanomagnetic (and also thermogalvanomagnetic [14]) properties of these systems for arbitrary magnetic fields. Nevertheless, we make a separate analysis of the case of a weak magnetic field **H** when the Hall coefficient and the magnetoresistance are particularly simply related to the dimensionless conductivity $f(p, h)$ and its derivative with respect to the argument h. For this model all the functions (which depend on p and h) contained in the expressions for the effective Hall coefficient and the effective magnetoresistance were calculated and tabulated in graphical form.

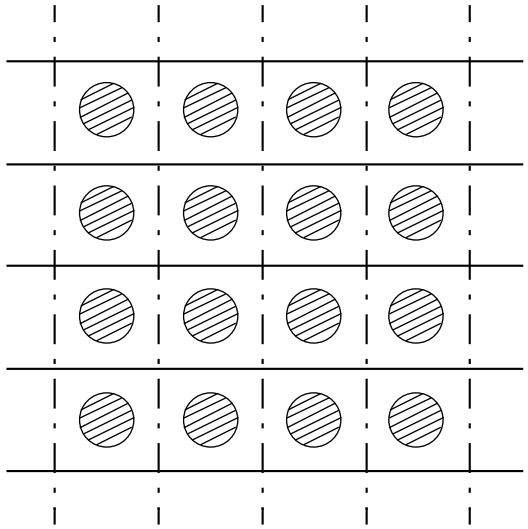


Fig. 1.

The model discussed is analyzed in terms of the conductivity problem (using the corresponding terminology). Solutions of similar problems involving thermal conductivity, permittivity, steady-state diffusion, and so on can be obtained by obvious changes in notation. We also note that the proposed method (for the case of circular inclusions) can be generalized to structures of different symmetry and systems with a higher number of components.

2. ELECTRIC FIELD IN THE MEDIUM

The model comprises a two-dimensional isotropic matrix of conductivity σ_1 with circular inclusions of radius R and conductivity σ_2. The inclusions form a regular structure with their centers located at the apexes of a square lattice having the period $2a$ (see Fig. 1). The effective conductivity of this system is isotropic so that the direction of the average electric field $\langle \mathbf{E} \rangle$ can be selected arbitrarily. We shall direct $\langle \mathbf{E} \rangle$ along the x axis. In this case, in order to determine the electric field (and current) in the entire plane, it is sufficient to analyze the problem in the unit cell shown in Fig. 2. In the selected geometry the electric field strength $\mathbf{E}(x, y)$ possesses specific symmetry:

$$\begin{aligned} E_x(-x, y) &= E_x(x, -y) = E_x(x, y), \\ E_y(-x, y) &= E_y(x, -y) = -E_y(x, y). \end{aligned} \tag{1}$$

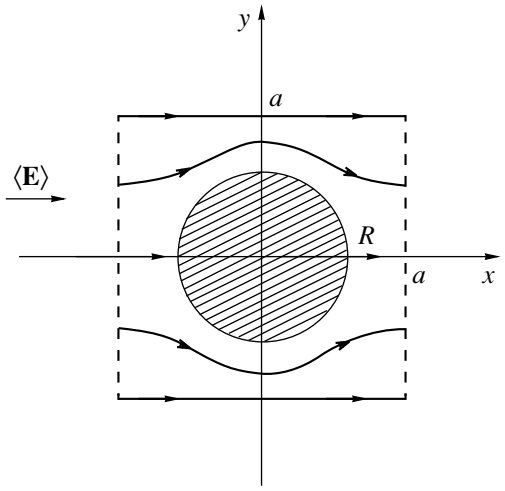


Fig. 2. Unit cell. Dashed lines are equipotential contours, solid lines with arrows are flow lines.

In particular, the vertical boundaries of the unit cell (and also the line $x = 0$) are equipotential contours on which $E_y = 0$ and the horizontal boundaries (and the line $y = 0$) are lines of flow on which $E_y = 0$ is also found.

The solutions of the two-dimensional Laplace equation in polar coordinates (r, θ) have the form $r^{\pm k}\cos k\theta$ and $r^{\pm k}\sin k\theta$, where $k = 0, 1, 2, \ldots$. Inside the circle $(r < R)$ allowing for the symmetry of the electric field (1) we have

$$\varphi^{(i)}(\mathbf{r}) = \sum_{n=0}^{\infty} A_{2n+1} r^{2n+1}\cos(2n+1)\theta. \qquad (2)$$

In the two-dimensional case, it is convenient to use the complex potential $\Phi(z)$ (where $z = x + iy$) whose derivative is expressed in terms of the strength components of $\mathbf{E}$ as follows:

$$\Phi'(z) = -E_x + iE_y, \qquad (3)$$

where $\varphi(\mathbf{r}) = \mathrm{Re}\,\Phi(z)$. For the complex potential inside the circle $(|z| < R)$ we obtain from (2) and (3)

$$\Phi^{(i)}(z) = \sum_{n=0}^{\infty} A_{2n+1} z^{2n+1} \qquad (4)$$

with the real (in the selected geometry) coefficients A_{2n+1}.

Outside the circle $(r > R)$ it is permissible to have solutions with both positive and negative exponents $k = 2n + 1$. From these solutions we need to construct an analytic function possessing the required properties (see above) at the boundaries of the unit cell. We shall attempt to construct this function assuming that the parameter R/a is formally small. In the absence of any inclusions, the complex potential contains only one (linear with respect to z) term which is responsible for the external homogeneous electric field applied to the system. In the lowest approximation with respect to R/a the existence of inclusions can be taken into account by summing the fields from the dipole moments induced at them, taking them as points. The potential created at point z by a dipole positioned at site z_{lm} is proportional to $(z - z_{lm})^{-1}$. Summation over all l and m gives the dipole component $\Phi_d(z)$ of the complete potential. However, the corresponding sum does not converge so it must be regularized. After this procedure the complex potential in the dipole approximation for $|z| > R$ has the form

$$\Phi_d(z) = \beta z + B_0\zeta(z), \qquad (5)$$

$$\zeta(z) = \frac{1}{z} + \sum_{l,m}{}'\left[\frac{1}{z - z_{lm}} + \frac{1}{z_{lm}} + \frac{z}{z_{lm}^2}\right], \qquad (6)$$

$$z_{lm} = 2la + i2ma.$$

Here $\zeta(z)$ is the Weierstrass zeta function [11–13]; summation in (6) is performed over all integer values of l and m except $l = m = 0$. The coefficients β and B_0 in (5) are real. It follows from (5) that

$$\Phi_d'(z) = \beta - B_0\mathcal{P}(z), \qquad (7)$$

where $\mathcal{P}(z) = -\zeta'(z)$ is an elliptic Weierstrass function [11–13]. Since for $z = iy$, $z = \pm a + iy$, $z = x$, and $z = x \pm ia$, we have $\mathrm{Im}\,\mathcal{P}(z) = 0$ (see [11, 13]) and $E_y = 0$ on these curves. Thus, the complex potential (5) satisfies the conditions at the unit cell boundaries (and on their axes of symmetry) imposed by the selected formulation of the problem.

In the next approximations with respect to R/a we need to allow for higher-order odd multipoles. The corresponding contributions to the potential can be obtained by differentiating the function $\zeta(z)$ from (6) an even number of times. As a result, for the complex potential in the region outside an inclusion we obtain the following expression for $|z| > R$:

$$\Phi^{(e)}(z) = \beta z + \sum_{n=0}^{\infty} B_{2n}\zeta^{(2n)}(z) \qquad (8)$$

with the real coefficients β and B_{2n} (in the selected geometry). In (8) $\zeta^{(2n)}(z)$ is a $2n$th order derivative of the zeta function (6). On account of the properties of the Weierstrass functions $\zeta(z)$ and $\mathcal{P}(z)$ (see [11–13]) the potential $\Phi^{(e)}(z)$ satisfies the same conditions at the unit cell boundaries and on its axes of symmetry as its dipole part (5).

In accordance with [11–13], the zeta function is quasi-periodic:

$$\begin{aligned}\zeta(z + 2\omega) &= \zeta(z) + 2\eta, \quad \eta = \zeta(\omega),\\ \zeta(z + 2\omega') &= \zeta(z) + 2\eta', \quad \eta' = \zeta(\omega'),\end{aligned} \qquad (9)$$

and the function $\mathcal{P}(z)$ and its derivatives are doubly periodic having the periods 2ω and $2\omega'$. For the square

lattice studied (having the half-periods $\omega = a$, $\omega' = ia$, known as the lemniscate case [11, 12]) we have

$$\eta = \frac{\pi}{4a}, \quad \eta' = -i\frac{\pi}{4a}, \qquad g_2 = \frac{1}{a^4}[K(1/\sqrt{2})]^4, \quad g_3 = 0, \tag{10}$$

where g_2 and g_3 are invariants of the Weierstrass function [11–13], $K(1/\sqrt{2}) = 1.85407\ldots$ is a complete elliptic integral of the first kind with the modulus $k = 1/\sqrt{2}$.

The electric potentials $\varphi^{(e)} = \mathrm{Re}\,\Phi^{(e)}(z)$ and $\varphi^{(i)} = \mathrm{Re}\Phi^{(i)}(z)$ should satisfy standard conditions at the inclusion boundary ($r = R$):

$$\varphi^{(e)} = \varphi^{(i)}, \qquad \frac{\partial\varphi^{(e)}}{\partial r} = h\frac{\partial\varphi^{(i)}}{\partial r}, \quad h = \frac{\sigma_2}{\sigma_1}. \tag{11}$$

From (11) we obtain a system of equations for the unknown coefficients A_{2n+1} and B_{2n} appearing in expressions (2), (4), and (8). In this procedure we use an expansion for the function $\zeta(z)$ which has the following form in the lemniscate case [12]

$$\zeta(z) = \frac{1}{z} - \sum_{k=1}^{\infty}\frac{c_{2k}}{4k-1}z^{4k-1}, \tag{12}$$

where

$$c_2 = g \equiv \frac{g_2}{20}, \quad c_4 = \frac{1}{3}g^2, \qquad c_6 = \frac{2}{3\times 13}g^3, \quad c_8 = \frac{5}{3\times 13\times 17}g^4, \ldots \tag{13}$$

with g_2 from (10). The coefficients c_{2k} satisfy the recurrence formula [12]

$$c_{2k} = \frac{3}{(4k+1)(2k-3)}\sum_{m=1}^{k-1}c_{2m}c_{2k-2m} \quad (k \geq 2), \tag{14}$$

which can be used to successively find the values of c_{2k} with increasing index k.

It follows from (6) that the coefficients c_{2k} are expressed in terms of the double sums

$$c_{2k} = 4\frac{4k-1}{(2a)^{4k}}F_k,$$

$$F_k = \frac{1}{4}\sum_{l,m}{}'\frac{1}{(l+im)^{4k}}.$$

For $k \gg 1$ the following asymptotic expansion holds for F_k

$$F_k \simeq 1 + \frac{(-1)^k}{2^{2k}} + \frac{1}{4^{4k}} + (-1)^k\frac{2\cos(2k\arccos 4/5)}{5^{2k}} + \ldots,$$

which can be used to calculate the coefficients c_{2k} with large indices.

Instead of the coefficients B_{2n} we introduce the "variables" ξ_n:

$$B_{2n} = \xi_n R^2\delta \quad (n = 0, 1, 2, \ldots); \quad \delta = \frac{1-h}{1+h}. \tag{15}$$

The finite system of equations for ξ_n derived from (11) then has the form (see Appendix A)

$$\sum_{m=0}^{\infty}(\delta_{nm} + M_{nm})\xi_m = \beta\delta_{n0}, \tag{16}$$

where

$$M_{nm} = \frac{(2n+2m)!}{(2n)!(2n+1)!}c_{n+m+1}R^{4n+2}\delta. \tag{17}$$

The coefficients c_{n+m+1} with even indices ($n + m + 1 = 2k$) are determined in (12)–(14) and those with odd indices are zero. Thus, the value of M_{nm} is only nonzero when the indices n and m have different parity. The coefficients A_{2n+1} are expressed in terms of ξ_n as follows:

$$A_{2n+1} = \frac{2}{1+h}\frac{(2n)!}{R^{4n}}\xi_n \quad (n = 0, 1, 2, \ldots). \tag{18}$$

The equations (16) and expressions (15) and (18) provide the fundamental possibility of expressing all the coefficients B_{2n} and A_{2n+1} in terms of β which in turn is related to the potential difference U [see (20)]. Thus, expressions (4), (8), and (15)–(18) together with (20) provide a complete formal solution of the formulated problem.

3. EFFECTIVE CHARACTERISTICS

For the unit cell shown in Fig. 2, the potential difference U and the total current I are expressed in terms of the complex potential $\Phi(z)$ as follows:

$$U = -\mathrm{Re}[\Phi(a+iy) - \Phi(-a+iy)], \qquad I = -\sigma_1\mathrm{Im}[\Phi(x+ia) - \Phi(x-ia)]. \tag{19}$$

Substituting into (19) the expressions (8) allowing for (9) and (10) gives

$$U = -2a\left(\beta + B_0\frac{\pi}{4a^2}\right),$$
$$I = -2\sigma_1 a\left(\beta - B_0\frac{\pi}{4a^2}\right). \tag{20}$$

For the dimensionless effective conductivity $f = \sigma_e/\sigma_1 = I/(\sigma_1 U)$ we obtain from (20) ($B_0 = \xi_0 R^2\delta$)

$$f = \left(\alpha - \frac{\pi R^2}{4a^2}\delta\right)\left(\alpha + \frac{\pi R^2}{4a^2}\delta\right)^{-1}, \quad \alpha = \frac{\beta}{\xi_0}, \tag{21}$$

so that the function f can be determined if we know ξ_0 (i.e., the coefficient B_0).

For $R \ll a$ (low concentrations) the system (16) can be solved iteratively by expanding in powers of the matrix $\hat{M}$. As a result, we find

$$\xi_n = \beta\left\{\delta_{n0} - M_{n0} + \sum_m M_{nm}M_{m0} - \sum_l\sum_m M_{nl}M_{lm}M_{m0} + \sum_k\sum_l\sum_m M_{nk}M_{kl}M_{lm}M_{m0} - \ldots\right\}, \tag{22}$$

whence it follows that for $n = 0$

$$\xi_0 = \beta\left\{1 + \sum_m M_{0m}M_{m0} + \sum_k\sum_l\sum_m M_{0k}M_{kl}M_{lm}M_{m0} + \ldots\right\}. \tag{23}$$

In (22) and (23) summation over each of the indices is performed between 0 and ∞. In (23) allowance is made for the property of the $\hat{M}$ matrix elements noted above: $M_{nm} = 0$ if n and m have the same parity. Determining ξ_0/β from (23) with the required accuracy and substituting into (21), we obtain the function f with the same accuracy.

However, it is more convenient to invert equality (23), expressing β in terms of ξ_0. For $n = 0$, in accordance with (16) we have

$$\beta = \xi_0 + \sum_{m\neq 0} M_{0m}\xi_m. \tag{24}$$

Thus, for $n \neq 0$ (16) yields the equality

$$\xi_n = -\xi_0 M_{n0} - \sum_{m\neq 0} M_{nm}\xi_m. \tag{25}$$

Solving equation (25) iteratively (expanding in powers of $\hat{M}$) and then substituting the expression obtained for ξ_n (for $n \neq 0$) into (24), we obtain

$$\alpha = \frac{\beta}{\xi_0} = 1 - {\sum_m}' M_{0m}M_{m0} - {\sum_k}'{\sum_l}'{\sum_m}' M_{0k}M_{kl}M_{lm}M_{m0} - {\sum_k}'{\sum_l}'{\sum_m}'{\sum_n}'{\sum_s}' M_{0k}M_{kl}M_{lm}M_{mn}M_{ns}M_{s0} - \ldots. \tag{26}$$

Here the prime after the summation sign indicates that summation is performed between 1 and ∞. Note that formula (26) only contains even powers of $\hat{M}$.

Using the explicit form of the matrix $\hat{M}$ [see (17)], from (26) we can obtain the value of α as an expansion in powers of R/a. Thus, to within terms $\sim(R/a)^{24}$ inclusive, we have

$$\alpha = 1 - \frac{1}{3}(gR^4)^2\delta^2 - \frac{1}{63}(gR^4)^4\delta^2 - \frac{5}{9}\left(\delta^2 + \frac{4}{5\times 11\times 13^2}\right)(gR^4)^6\delta^2 - \ldots, \tag{27}$$

where $g = g_2/20$, $g_2$was determined in (10), and δ in (15). The expansion (27) converges rapidly for $R/a \ll 1$. A comparison with the results of a numerical analysis (see below) of the system (16) shows that expressions (21) and (27) give the function f to within ~1% over the entire range of concentration (i.e., for $0 \le R \le a$) for $|\delta| \lesssim 0.7$ and for $0 \le R \le 0.95a$ for $|\delta| = 1$ (for $h = 0$ or $h = \infty$). Such a wide range of validity of this approximation is obtained because, as is deduced from (27), the parameter of the virial expansion in this case is not the inclusion concentration $c = \pi R^2/(2a)^2$ but its fourth power. For this reason even for $\alpha = 1$ formula (21) gives the function f to within ~1% in the range $0 \le R \le 0.7a$ (for $|\delta| = 1$).

Note that in accordance with (26) [see also (27)] the value of α does not depend on the sign of δ since it is a function of δ^2. Thus, making the substitution $\sigma_1 \rightleftharpoons \sigma_2$ (i.e., $h \longrightarrow 1/\hbar$, $\delta \longrightarrow -\delta$) we have $f \longrightarrow 1/f$. Consequently,

$$f(p, h)\, f(p, 1/h) = 1, \tag{28}$$

so that for this model the reciprocity relation [1, 2] (see also [4]) is satisfied automatically.

For arbitrary R/a the system (16) was solved by numerical methods (in the range $0 \le R \le a$). In this case we analyzed a finite subsystem of N equations with N unknowns $\xi_0, \xi_1, \ldots, \xi_{N-1}$ with the maximum $N = 40$ (in the range $R \simeq a$). Corresponding results for f as a func-

tion of the concentration of the first component p for three values of the argument h (0.5; 0,1; 10^{-5}) are plotted in Fig. 3. Using these data the function $f(p, h)$ for $h > 1$ can be determined from the reciprocity relation (28).

The mean-square characteristics of the electric field are directly related to the effective conductivity σ_e

$$\psi_i = \langle \mathbf{e}^2 \rangle^{(i)} = \frac{\partial \sigma_e}{\partial \sigma_i}, \qquad \mathbf{e}(\mathbf{r}) = \mathbf{E}(\mathbf{r})/|\langle \mathbf{E} \rangle|. \tag{29}$$

Here $\langle \mathbf{E} \rangle$ is the average over the volume (area in the two-dimensional case) of the sample V and $\langle \ldots \rangle^{(i)}$ is the integral over the volume of the ith component divided by V. For a two-component medium we have [15]

$$\psi_1 = \langle \mathbf{e}^2 \rangle^{(1)} = f - h f', \tag{30}$$

$$\psi_2 = \langle \mathbf{e}^2 \rangle^{(2)} = f', \quad f' \equiv \frac{\partial f(p, h)}{\partial h}. \tag{31}$$

It is easy to see that for this model $\langle E_x \rangle = U/(2a)$ with U from (20). Determining the value of $\langle \mathbf{E}^2 \rangle^{(2)}$ using (2) and (4), we obtain

$$\psi_2 = \langle \mathbf{e}^2 \rangle^{(2)} = \frac{\pi R^2}{U^2} \sum_{n=0}^{\infty} (2n+1) R^{4n} (A_{2n+1})^2$$

or allowing for (18) and (20)

$$\psi_2 = \frac{4}{(1+h)^2} \frac{1}{\left(\alpha + \dfrac{\pi R^2}{4a^2}\delta\right)^2} \frac{\pi R^2}{(2a)^2} \times \left\{ 1 + \sum_{n=1}^{\infty} \frac{(2n)!(2n+1)!}{R^{4n}} \left(\frac{\xi_n}{\xi_0} \right)^2 \right\}. \tag{32}$$

Calculating (32) to within $(R/a)^{24}$ inclusive and comparing the expression obtained for ψ_2 with the derivative of h from (21) with α from (27), we can confirm that the relationship $\psi_2 = f'$ is satisfied in this approximation. A proof of the validity of this relationship for this particular model for arbitrary R/a is given in Appendix B.

Results of a numerical analysis for ψ_1 and ψ_2 are plotted in Figs. 4 and 5. In this case ψ_2 was calculated using formula (32) and ψ_1 from the relationship $\psi_1 = f - h\psi_2$ [see (30)]. The functions ψ_1 and ψ_2 for $h > 1$ can be obtained from their values for $h < 1$ using the equalities

$$\psi_1(p, 1/h) = \frac{\psi_1(p, h)}{[f(p, h)]^2},$$

$$\psi_2(p, 1/h) = \left[\frac{h}{f(p, h)} \right]^2 \psi_2(p, h),$$

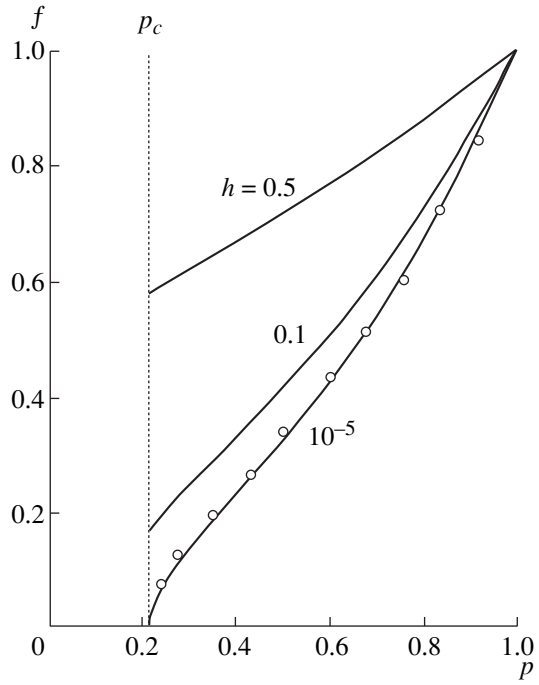


Fig. 3. Dimensionless effective conductivity f as a function of the concentration of the first component p for three values of the argument h. Circles give results of model experiment for $h = 0$ (see Fig. 39 from [8]).

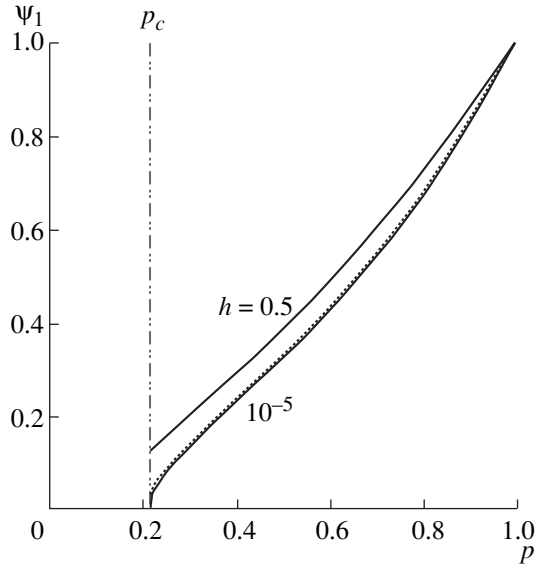


Fig. 4. Dependence of $\psi_1 = \langle \mathbf{e}^2 \rangle^{(1)} = f' - hf'$ on the concentration p for three values of the argument h. The dotted curve corresponds to the case $h = 0.1$.

which are obtained from the definitions (30) and (31) and the reciprocity relation (28).

4. CRITICAL REGION

If the conductivity of the second component (inclusions) is zero, for $R \longrightarrow a$ the effective electrical conductivity of the medium σ_e goes to zero since a metal–

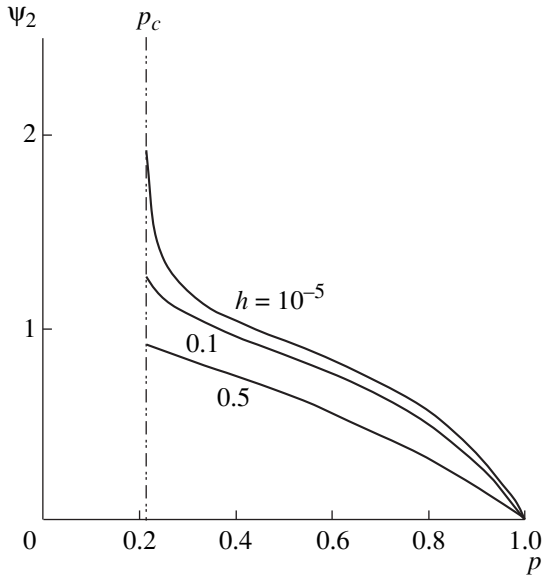


Fig. 5. Dependence of $\psi_2 = \langle \mathbf{e}^2 \rangle^{(2)}$ on the concentration p for three values of the argument h.

insulator phase transition takes place in the system. The corresponding critical concentration of the conducting (first) component is $p_c = (4 - \pi)/4 = 0.21460\ldots$. For randomly inhomogeneous media the properties of σ_e (or the function f) in the critical region $h \ll 1$ or $|\tau| \ll 1$ ($\tau = (p - p_c)/p_c$ is the similarity parameter at the transition point in terms of concentration) are described in terms of the similarity hypothesis [16]. In a certain sense this description is also suitable for periodic systems, see for example [9]. Below we confine our analysis to the region $p \geq p_c$ (i.e., $R \leq a$) since the method proposed in the present study cannot be applied for $p < p_c$ ($R > a$).

In accordance with [16] (see also, for example [15]) the function $f(p, h)$ in the critical region has the following expansions ($\tau > 0$):

$$f \simeq \tau^t \left\{ A_0 + A_1 \frac{h}{\tau^{t/s}} + A_2 \left(\frac{h}{\tau^{t/s}} \right)^2 + \ldots \right\} \quad (33)$$

$$\text{for } \Delta_0 \ll \tau \ll 1,$$

$$f \simeq h^s \left\{ a_0 + a_1 \frac{\tau}{h^{s/h}} + a_2 \left(\frac{\tau}{h^{s/t}} \right)^2 + \ldots \right\} \quad (34)$$

$$\text{for } \tau \ll \Delta_0.$$

Here

$$\Delta_0 = h^{s/t} \quad (35)$$

is the size of the smearing region [16]. Corresponding expansions for the functions ψ_1 and ψ_2 are obtained after substituting (33) and (34) into relations (30) and (31). In particular, for $\Delta_0 \ll \tau \ll 1$

$$\psi_2 \simeq \frac{1}{\tau^{q'}} \left\{ A_1 + 2A_2 \frac{h}{\tau^{t/s}} + \ldots \right\}, \quad (36)$$

where

$$q' = \frac{t}{s} - t. \quad (37)$$

For randomly inhomogeneous media the value of q' is the same as the critical index q characterizing the behavior of the function f for $\tau < 0$ [16].

For this particular model an investigation of the critical region encounters certain difficulties because of its closeness to the limit of validity of the model used. For $R \longrightarrow a$ the corresponding expansions (in powers of R/a) converge fairly slowly which necessitates allowing for a considerably larger number of terms of these expansions. For example, using expression (27) for α for $h = 0$ and $R = a$ gives $f(p_c, 0) \simeq 0.044$. Analyzing an infinite system (16) of a subsystem of nine equations we obtain $f(p_c, 0) \simeq 0.013$. Finally, by enlarging this subsystem to forty equations we obtain $f(p_c, 0) \simeq 0.0037$.

In this last case we observed critical behavior of the functions f and f' of the type (33), (34), and (36) in the concentration range $p_c < p < 0.217$. An analysis of the numerical results for f in the range $\Delta_0 \ll \tau \ll 1$ gives

$$t \simeq 0.51, \quad A_0 \simeq 0.18. \quad (38)$$

Thus from the results for $\psi_2 = f'$ in the same range of concentrations we find

$$q' \simeq 0.03, \quad A_1 \simeq 2.1. \quad (39)$$

Finally from the dependences of f and f' on h for $p = p_c$ ($R = a$) we obtain

$$s \simeq 0.95, \quad a_0 \simeq 1.88. \quad (40)$$

We note that relation (37) allowing for (38) and (40) yields $q' \simeq 0.027$. It is still possible that the "true" value is $q' = 0$ (where $s = 1$) which corresponds to a logarithmic dependence of f on τ for $\tau < 0$.

For R close to a the function f for $h = 0$ (or $h = \infty$) can be determined approximately by calculating the resistance of the appropriate "contact." This problem is solved in terms of bipolar coordinates [17]. As a result, for dielectric inclusions we obtain

$$f(p, 0) \simeq \frac{1}{\pi} \ln \left[\frac{a + \sqrt{a^2 - R^2}}{R} \right]. \quad (41)$$

For ideally conducting inclusions the effective conductivity may be determined from the reciprocity relation (28). Note that the accuracy of formula (41) increases as the ratio R/a approaches one.

For $R \longrightarrow a$ we obtain from (41) (cf. [18])

$$f \simeq \frac{1}{\pi}\sqrt{2\frac{a-R}{a}},$$

which, when converted to concentration, gives

$$f \simeq A_0\tau^{1/2},$$
$$A_0 = \frac{1}{\pi}\sqrt{\frac{4}{\pi}-1} = 0.166\ldots, \quad (42)$$

which agrees with (38). Expression (41) describes the results of numerical calculations for f in the range $0.216 \le p \le 0.218$ to within ≤1%.

This analysis shows that a scaling description can be applied to this periodic model. However, compared with "ordinary" disordered systems this case has various differences. This particularly applies to the critical index t. In accordance with the assumption put forward in [9], the value of t is less than one. This then has the result that $\partial f/\partial\tau \longrightarrow \infty$ for $\tau \longrightarrow +0$ whereas in the ordinary case $t > 1$ (for two-dimensional systems $t \simeq 1.3$) and $\partial f/\partial\tau \longrightarrow 0$ for $\tau \longrightarrow +0$. Another important difference is the smallness of the smearing region (35) since for this particular model $s/t \simeq 2$ whereas for a randomly inhomogeneous two-dimensional system $s/t \simeq 0.4$. Finally attention is drawn to the extreme smallness of the index q' (instead of the ordinary $q' = t \simeq 1.3$) which indicates that the dependence of f on τ may be logarithmic for $\tau < 0$.

5. GALVANOMAGNETIC PROPERTIES

The problem of the galvanomagnetic properties of two-dimensional two-component systems in a transverse magnetic field **H** with the conductivity tensor

$$\hat{\sigma} = \begin{pmatrix} \sigma_x & \sigma_a \\ -\sigma_a & \sigma_x \end{pmatrix} \quad (43)$$

has an exact solution for arbitrary **H** [6, 7]. In this case, the components of the effective conductivity tensor $\hat{\sigma}_e$ are expressed in terms of the galvanomagnetic characteristics of individual components and the function $f(p, \lambda)$ obtained from the dimensionless effective electrical conductivity (for $\mathbf{H} = 0$) $f(p, h)$ by substituting $h \longrightarrow \lambda$. The parameter λ is also expressed in terms of the galvanomagnetic characteristics of the components and thus depends on **H**. The relevant fairly cumbersome formulas for $\hat{\sigma}_e$ and λ are given in [6, 7]. We shall subsequently confine our analysis to the case of weak magnetic fields when the components of the tensor $\hat{\sigma}_e$ can be expressed directly in terms of the electrical conductivity for $\mathbf{H} = 0$, i.e., in terms of the function $f(p, h)$.

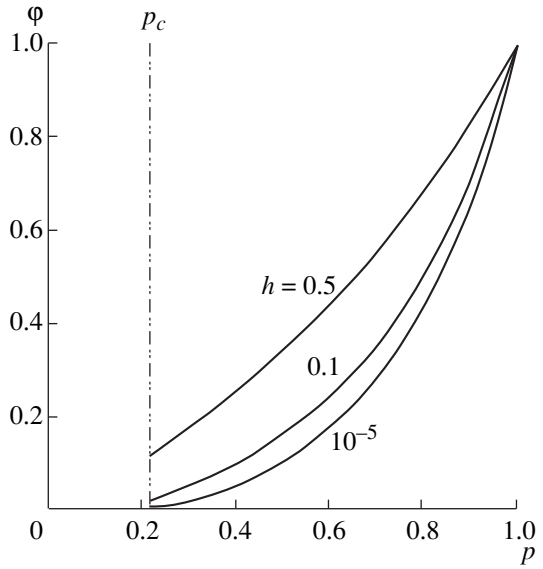


Fig. 6. The function φ determined using formula (45).

In the linear approximation with respect to **H** the Hall component of the tensor $\hat{\sigma}_e$ for a two-component system has the form [15]

$$\sigma_{ae} = \sigma_{a2} + (\sigma_{a1} - \sigma_{a2})\varphi(p, h), \quad (44)$$

where $\varphi(p, h)$ is a certain function which depends on the properties of the medium for $\mathbf{H} = 0$. In the two-dimensional case $\varphi(p, h)$ can be expressed in terms of $f(p, h)$ [15]:

$$\varphi(p, h) = \frac{f^2 - h^2}{1 - h^2}. \quad (45)$$

For the effective Hall coefficient $R_e = H^{-1}\sigma_{ae}/\sigma_e^2$ we obtain from (44) [15]

$$R_e = h^2 R_2 f^{-2} + (R_1 - h^2 R_2)\mathcal{R},$$
$$\mathcal{R}(p, h) = \frac{\varphi(p, h)}{[f(p, h)]^2} \quad (46)$$

with $\varphi(p, h)$ from (45). For the model considered in the present study the dependence of the function φ on the concentration p is plotted in Fig. 6 for three values of the argument h ($h < 1$). For $h > 1$ φ can be determined from the equality

$$\varphi(p, 1/h) = \frac{\varphi(p, h)}{[f(p, h)]^2},$$

which is an obvious consequence of the reciprocity relation (28).

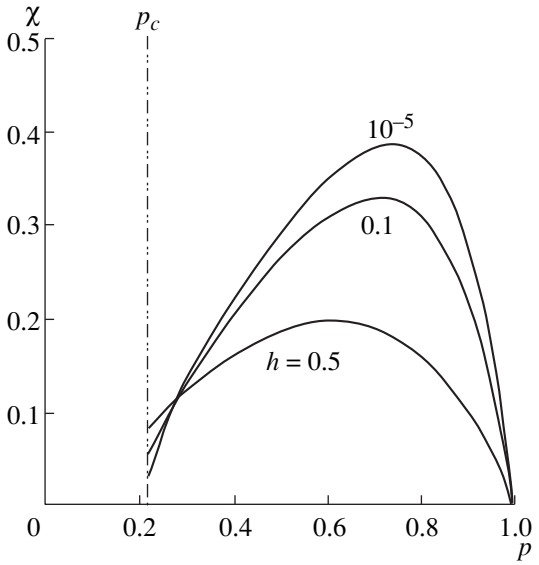


Fig. 7. The function χ determined using formula (49).

We express the diagonal components of the conductivity tensors $\hat{\sigma}_e$ and $\hat{\sigma}_i$ in a weak magnetic field in the form

$$\sigma_{xe} = \sigma_e + \gamma_e, \quad \sigma_{xi} = \sigma_i + \gamma_i, \tag{47}$$

where σ_e and σ_i are the conductivities for $\mathbf{H} = 0$, and γ_e and γ_i are quadratic with respect to $\mathbf{H}$. In accordance with [15] for γ_e we have

$$\gamma_e = \gamma_1\psi_1 + \gamma_2\psi_2 + \frac{(\sigma_{a1} - \sigma_{a2})^2}{\sigma_1}\chi. \tag{48}$$

Here the functions $\psi_1 = \psi_1(p, h)$ and $\psi_2 = \psi_2(p, h)$ are the same as in (30), (31). The function $\chi(p, h)$ in the two-dimensional case can be expressed in terms of $f(p, h)$ [15]:

$$\chi = (\psi_1 - f\varphi)/(1 - h^2) \tag{49}$$

with ψ_1 from (30) and φ from (45). For the model being studied the dependence of χ on the concentration p is plotted in Fig. 7 for three values of the argument h ($h < 1$). The following equality obtained from the definition (49) and the reciprocity relation for the functions f, ψ_1, and φ

$$\chi\left(p, \frac{1}{h}\right) = -\frac{h^2}{1 - h^2}\left\{\frac{\psi_1(p, h)}{[f(p, h)]^2} - \frac{\varphi(p, h)}{[f(p, h)]^3}\right\},$$

can be used to determine χ for $h > 1$.

For the magnetoresistance

$$\frac{\Delta\rho_e}{\rho_e} = \frac{\rho_{xe}(\mathbf{H}) - \rho_{xe}(0)}{\rho_{xe}(0)},$$

where $\rho_x = \sigma_x/(\sigma_x^2 + \sigma_a^2)$, in the quadratic approximation with respect to $\mathbf{H}$ we have

$$\frac{\Delta\rho_e}{\rho_e} = -\left(\frac{\gamma_e}{\sigma_e} + \frac{\sigma_{ae}^2}{\sigma_e^2}\right). \tag{50}$$

Here σ_e is the electrical conductivity for $\mathbf{H} = 0$, σ_{ae} is defined in (44) and (45), and γ_e in (48) and (49). Expression (50) can be used to describe the magnetoresistance of a two-dimensional two-component system over the entire concentration range and in particular in the critical range.

6. CONCLUSIONS

We shall briefly discuss the main differences between the results of the present study and those obtained (for the two-dimensional case) in [10]. An advantage of the proposed approach compared with that developed in [10] is first that the lattice structure of the model is taken into account in the lowest dipole and in each successive approximation. Thus, in all orders of the virial expansion accurate allowance is made for the symmetry of the electric field and in particular, the conditions at the unit cell boundaries. (In [10] these conditions were not discussed despite their importance.) Then the potential outside the inclusions is in fact expressed in terms of the same quantity, the Weierstrass zeta function, which has been fairly well studied. Using the known properties of the zeta function significantly simplifies the solution of the main problem of determining the potential, and simplifies the calculations of the effective model characteristics. Another important advantage is the establishment of an explicit form of M_{nm} [see formulas (16) and (17)] which means that in principle, numerical methods can be applied to find the conductivity and other effective characteristics over a wide range of parameters p and h with arbitrary accuracy. This last factor allowed us to study the vicinity of the metal–insulator phase transition point. (In [10] however only the first few terms of the corresponding virial expansion were obtained so that it is not possible to consider the critical region.) The existence of an exact formal solution also allows us to study problems of a general nature. For instance, relation (31) can be proved by direct calculations (for arbitrary p and h) which is a serious argument in support of the accuracy of this solution. Finally, a considerably wider range of physical problems was considered in the present study compared with [10].

APPENDIX A

We write the expansion (12) in the form

$$\zeta(z) = \frac{1}{z} - \sum_{l=2}^{\infty}\frac{c_l}{2l-1}z^{2l-1}, \tag{A.1}$$

where the coefficients c_l with odd indices are zero in this lemniscate case (square lattice). Differentiating (A.1) $2n$ times gives

$$\zeta^{(2n)}(z) = \frac{(2n)!}{z^{2n+1}} - \sum_{m=0}^{\infty} \frac{(2n+2m)!}{(2m+1)!} c_{n+m+1} z^{2m+1}. \quad \text{(A.2)}$$

We substitute (A.2) into (8), set $z = r\exp\{i\theta\}$, and separate the real part. As a result, for $r > R$ we obtain

$$\varphi^{(e)}(\mathbf{r}) = \beta r\cos\theta + \sum_{n=0}^{\infty} \left\{ B_{2n} \frac{(2n)!}{r^{2n+1}} - \sum_{m=0}^{\infty} B_{2m} \frac{(2n+2m)!}{(2n+1)!} c_{n+m+1} r^{2n+1} \right\} \cos(2n+1)\theta. \quad \text{(A.3)}$$

Substituting (2) and (A.3) into (11) yields the following system of equations

$$\begin{gathered} \beta\delta_{n0} + B_{2n}\frac{(2n)!}{R^{4n+2}} \\ - \sum_{m=0}^{\infty} B_{2m} \frac{(2n+2m)!}{(2n+1)!} c_{n+m+1} = A_{2n+1}, \\ \beta\delta_{n0} - B_{2n}\frac{(2n)!}{R^{4n+2}} \\ - \sum_{m=0}^{\infty} B_{2m} \frac{(2n+2m)!}{(2n+1)!} c_{n+m+1} = hA_{2n+1}, \end{gathered} \quad \text{(A.4)}$$

where δ_{n0} is the Kronecker symbol. Subtracting the second term from the first in (A.4), we find

$$A_{2n+1} = \frac{2}{1-h}\frac{(2n)!}{R^{4n+2}} B_{2n}. \quad \text{(A.5)}$$

Eliminating the coefficient A_{2n+1} from (A.4), we obtain

$$B_{2n} + \frac{1-h}{1+h} \sum_{m=0}^{\infty} B_{2m} \frac{(2n+2m)!}{(2n)!(2n+1)!} R^{4n+2} c_{n+m+1} = \frac{1-h}{1+h} \beta R^2 \delta_{n0}. \quad \text{(A.6)}$$

Substituting into (A.5) and (A.6) $B_{2n} = \xi_n R^2 \delta$ with δ from (15), we arrive at formulas (16)–(18).

APPENDIX B

We shall show that expressions (21), (26), and (32) identically satisfy relation (31).

We introduce the matrix $\hat{N}$ which is related to $\hat{M}$ from (17) as follows:

$$N_{nm} = \begin{cases} 0, & m = 0 \\ M_{nm}, & m \neq 0. \end{cases} \quad \text{(B.1)}$$

Using $\hat{N}$ the solution of equation (25) may be written in the form

$$\xi_n = -\xi_0((1+\hat{N})^{-1}\hat{M})_{n0} \quad \text{for} \quad n \neq 0, \quad \text{(B.2)}$$

so that for $\alpha = \beta/\xi_0$ from (24) allowing for (B.2) we obtain

$$\alpha = 1 - (\hat{M}(1+\hat{N})^{-1}\hat{M})_{00}. \quad \text{(B.3)}$$

Expressions (B.3) and (26) are equivalent since only even powers of $\hat{N}$ make a nonzero contribution to α

Differentiating f from (21) with respect to h gives

$$\frac{\partial f}{\partial h} = \frac{4}{(1+h)^2}\frac{\pi R^2}{(2a)^2}\left(\alpha + \frac{\pi R^2}{4a^2}\delta\right)^{-2}\left(\alpha - \frac{\partial\alpha}{\partial\delta}\delta\right). \quad \text{(B.4)}$$

The matrices $\hat{M}$ and $\hat{N}$ depend linearly on δ so that for the last factor from (B.4) allowing for (B.3) we obtain

$$\alpha - \frac{\partial\alpha}{\partial\delta}\delta = 1 + (\hat{M}(1+\hat{N})^{-1}(1+\hat{N})^{-1}\hat{M})_{00}. \quad \text{(B.5)}$$

We write ψ_2 from (32) in the form

$$\psi_2 = \langle \mathbf{e}^2 \rangle^{(2)} = \frac{4}{(1+h)^2}\frac{\pi R^2}{(2a)^2}\left(\alpha + \frac{\pi R^2}{4a^2}\delta\right)^{-2} J, \quad \text{(B.6)}$$

$$J = 1 + \sum_{n=1}^{\infty} \frac{(2n)!(2n+1)!}{R^{4n}}\left(\frac{\xi_n}{\xi_0}\right)^2. \quad \text{(B.7)}$$

Substituting (B.2) into (B.7) gives

$$J = 1 + \sum_{n=1}^{\infty} \frac{(2n)!(2n+1)!}{R^{4n}} \times ((1+\hat{N})^{-1}\hat{M})_{n0}((1+\hat{N})^{-1}\hat{M})_{n0}. \quad \text{(B.8)}$$

It is easy to see that for $n \neq 0$

$$\begin{gathered} (\hat{N}\hat{M})_{n0} = \frac{R^{4n}}{(2n)!(2n+1)!}(\hat{M}\hat{N})_{0n}, \\ (\hat{N}^2\hat{M})_{n0} = \frac{R^{4n}}{(2n)!(2n+1)!}(\hat{M}\hat{N}^2)_{0n}, \ldots \end{gathered} \quad \text{(B.9)}$$

Thus, for $n \neq 0$ we have

$$\frac{(2n)!(2n+1)!}{R^{4n}}((1+\hat{N})^{-1}\hat{M})_{n0} = (\hat{M}(1+\hat{N})^{-1})_{0n}, \quad \text{(B.10)}$$

so that for J from (B.8) we obtain

$$J = 1 + \sum_{n=1}^{\infty} (\hat{M}(1+\hat{N})^{-1})_{0n}((1+\hat{N})^{-1}\hat{M})_{n0}. \quad \text{(B.11)}$$

It follows from the definition (B.1) that

$$(\hat{M}(1+\hat{N})^{-1})_{00} = 0.$$

Thus, summation in (B.11) can be extended to all $n \geq 0$ and we then finally obtain for J:

$$J = 1 + (\hat{M}(1+\hat{N})^{-1}(1+\hat{N})^{-1}\hat{M})_{00}. \quad \text{(B.12)}$$

A comparison of (B.4) and (B.5) with (B.6) and (B.12) yields relation (31).

REFERENCES

1. J. B. Keller, J. Math. Phys. **5**, 548 (1964).
2. A. M. Dykhne, Zh. Éksp. Teor. Fiz. **59**, 110 (1970) [Sov. Phys. JETP **32**, 63 (1971)].
3. A. M. Dykhne, Zh. Éksp. Teor. Fiz. **59**, 641 (1970) [Sov. Phys. JETP **32**, 348 (1971)].
4. B. Ya. Balagurov, Zh. Éksp. Teor. Fiz. **81**, 665 (1981) [Sov. Phys. JETP **54**, 355 (1981)].
5. B. Ya. Balagurov, Fiz. Tekh. Poluprovodn. (Leningrad) **16**, 259 (1982) [Sov. Phys. Semicond. **16**, 1204 (1982)].
6. B. Ya. Balagurov, Zh. Éksp. Teor. Fiz. **82**, 1333 (1982) [Sov. Phys. JETP **55**, 774 (1982)].
7. B. Ya. Balagurov, Zh. Éksp. Teor. Fiz. **85**, 568 (1983) [Sov. Phys. JETP **58**, 331 (1983)].
8. Yu. P. Emets, *Electric Properties of Composites with Regular Structure* (Naukova Dumka, Kiev, 1986).
9. B. Ya. Balagurov, Zh. Éksp. Teor. Fiz. **79**, 1561 (1980) [Sov. Phys. JETP **52**, 787 (1980)].
10. Lord Rayleigh, Philos. Mag. **34**, 481 (1892).
11. *Heigher Transcendental Functions (Bateman Manuscript Project)*, Ed. by A. Erdelyi (McGraw-Hill, New York, 1955; Nauka, Moscow, 1967), Vol. 3.
12. *Handbook of Mathematical Functions*, Ed. by M. Abramowitz and I. A. Stegun (Dover, New York, 1965 (1971); Nauka, Moscow, 1979).
13. Yu. S. Sikorskiĭ, *Elements of Elliptic-Function Theory* (ONTI, Moscow, 1936).
14. B. Ya. Balagurov, Fiz. Tverd. Tela (Leningrad) **28**, 2068 (1986) [Sov. Phys. Solid State **28**, 1156 (1986)].
15. B. Ya. Balagurov, Zh. Éksp. Teor. Fiz. **93**, 1888 (1987) [Sov. Phys. JETP **66**, 1079 (1987)].
16. A. L. Efros and B. I. Shklovskiĭ, Phys. Status Solidi B **76**, 475 (1976).
17. P. M. Morse and H. Feshbach, *Methods of Theoretical Physics* (McGraw-Hill, New York, 1953; Inostrannaya Literatura, Moscow, 1959), Vol. 2.
18. J. B. Keller, J. Appl. Phys. **34**, 991 (1963).

Translation was provided by AIP

Journal of Experimental and Theoretical Physics, Vol. 90, No. 5, 2000, pp. 861–871.
Translated from Zhurnal Éksperimental'noĭ i Teoreticheskoĭ Fiziki, Vol. 117, No. 5, 2000, pp. 990–1002.

SOLIDS
Electronic Properties

Superfluid Transition Temperature in a Fermi Gas with Repulsion Allowing for Higher Orders of Perturbation Theory

D. V. Efremov *, **, M. S. Mar'enko*, *, M. A. Baranov****, and M. Yu. Kagan***

**Kapitza Institute of Physical Problems, Russian Academy of Sciences, Moscow, 117334 Russia*
***e-mail: efremov@kapitza.ras.ru*
****e-mail: maxim@kapitza.ras.ru*
*****Russian Research Centre Kurchatov Institute, pl. Kurchatova 1, Moscow, 123182 Russia*
Received October 13, 1999

Abstract—Higher-order perturbation-theory corrections to the superfluid transition temperature in a weakly nonideal Fermi gas with repulsion are determined. This involves calculating the contribution of third- and fourth-order diagrams in terms of the gas parameter ap_F to the effective interaction which determines the superfluid transition temperature and also allowing for effects associated with retardation and renormalization of the polar part of the Green's function. The expressions obtained provide evidence in support of attraction in the effective interaction in the second, third, and fourth orders of perturbation theory. It is shown that the critical temperature is mainly determined by second- and third-order terms of perturbation theory. Calculations are made of the superfluid transition temperature for a gas comprising neutral Fermi particles in a magnetic field. The limits of validity of the theory are analyzed and the possibility of applying the results to dilute solutions of ^{3}He in ^{4}He and neutral-particle Fermi gases in magnetic traps is discussed. © *2000 MAIK "Nauka/Interperiodica".*

1. INTRODUCTION

Nonphononic Cooper pairing mechanisms have recently started to attract greater attention. This is primarily attributable to the discovery of high-temperature superconducting (HTSC) systems, superconductivity in organic compounds and heavy-fermion compounds, and also because of the search for superfluidity in solutions of ^{3}He in ^{4}He and in atomic Fermi gases in traps. Moreover, HTSC systems and heavy-fermion systems belong to a class of strongly correlated systems whose theoretical analysis requires the development of new methods. At the same time, solutions of ^{3}He in ^{4}He and atomic Fermi gases in traps can be described using the model of a weakly nonideal Fermi gas. In this case the interparticle interaction can be either attractive or repulsive. In the attractive case normal singlet Cooper pairing takes place where the orbital momentum of the pair is $l = 0$ (according to the BCS type), for which the critical temperature was first calculated by Gor'kov and Melik-Barkhudarov [1]. In systems with repulsive interaction, the formation of $l = 0$ Cooper pairs is clearly impossible and in order to investigate the existence of superfluidity, we need to study the possibility of $l \neq 0$ Cooper pairing.

The possible existence of superfluidity in Fermi systems with repulsion was first indicated by Kohn and Luttinger in 1965. In [2] they examined the contribution of collective effects to the scattering amplitude in a particle–hole channel which lead to effective quasiparticle interaction at the Fermi surface via polarization of the Fermi background. A principal role in the formation of attractive harmonics in the effective interaction and consequently the superfluidity is played by the Kohn singularity in the effective interaction. In the three-dimensional case, this has the form

$$\tilde{\Gamma}_{eff}^{sing}(q) \sim [(2p_F)^2 - q^2]\ln\left|(2p_F)^2 - q^2\right| + \Gamma_{reg}(q^2). \quad (1)$$

In coordinate space the Kohn singularity leads to alternating oscillating RKKY interaction between quasiparticles:

$$\tilde{\Gamma}_{eff}^{sing}(r) \sim \frac{1}{r^3}\cos(2p_F r + \varphi).$$

It should be noted that the singular part of the effective interaction decreases over large distances more slowly that the seed interaction $U_0(r - r')$ and consequently makes the main contribution to the scattering amplitude in the limit of large momenta l. Integrating the singular part $\tilde{\Gamma}_{eff}(q)$ using Legendre polynomials yields the following results [2]:

$$\tilde{\Gamma}_{eff}^{(l)} \sim \frac{(-1)}{l^4}.$$

A simple extrapolation made by the authors of [2] yields extremely low estimates at superfluid transition temperatures in the limit $l \longrightarrow 2$, 10^{-16} K and 10^{-11} K

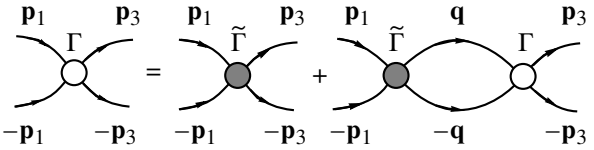


Fig. 1. Bethe–Salpeter equation for complete vertex Γ.

for ^{3}He and the electron subsystem in the metal, respectively.

It was subsequently shown in [3, 4] that effective attraction is also observed for the angular momentum $l = 1$ which gives the following expression for the critical triplet-pairing temperature in the second order of perturbation theory:

$$T_{c1} \sim \tilde{\varepsilon}\exp\left\{-\frac{5\pi^2}{4(2\ln 2-1)(ap_F)^2}\right\} \approx \tilde{\varepsilon}\exp\left\{-\frac{13.0}{\lambda^2}\right\}, \quad (2)$$

where $\lambda = 2ap_F/\pi$ is the effective gas parameter, a is the s-scattering length, p_F is the Fermi momentum, and $\tilde{\varepsilon}$ is the energy parameter, of the order of the Fermi energy, which functions as a cutoff parameter at high energies. Substituting real values for ^{3}He in which triplet pairing takes place gives good agreement with experiment: $T_{c1} \sim 10^{-3}$ K. (Obviously, the seed interaction in real ^{3}He is far more complex than that in this model).

The aim of the present paper is to determine the critical superfluid transition temperature of a weakly nonideal Fermi gas with repulsive interparticle interaction to within the preexponential factor. For this purpose we calculate the irreducible vertex in the Cooper channel in the third and fourth orders of perturbation theory with respect to the gas parameter λ. We also allow for renormalization of the polar parts of the Green's function (corrections associated with the Z factor and the effective mass) in the Bethe–Salpeter equation (3) and take into account retardation effects (the influence of the frequency and pulse momentum dependences of the irreducible vertex).

This article is constructed as follows. In Section 2 we derive and analyze an equation for the critical transition temperature in a weakly nonideal Fermi gas with repulsion. In Section 3 we calculate the irreducible vertex in the Cooper channel in the second, third, and fourth orders of perturbation theory. In Section 4 we examine the contribution of retardation effects. In Sections 5 and 6 we give the final formula for the critical temperature and discuss the contribution of the seed scattering in the p-channel. In Section 7 we note the possibility of an abrupt increase of T_{c1} in an external magnetic field. In Section 8 we discuss possible experimental applications of these results. In particular, we discuss the possibility of triplet Cooper pairing in solutions of ^{3}He in ^{4}He and in a neutral-particle Fermi gas in magnetic traps at ultralow temperatures.

2. SUPERFLUID TRANSITION IN A FERMI GAS WITH REPULSION

We shall consider a weakly nonideal Fermi gas described by the Hamiltonian

$$\hat{H} = \hat{H}_0 + \hat{H}_{int}$$
$$= \sum_{\alpha\mathbf{p}}(\varepsilon_{\mathbf{p}} - \mu)\hat{a}^\dagger_{\mathbf{p}\alpha}\hat{a}_{\mathbf{p}\alpha} + g\sum_{\alpha\beta\mathbf{p}\mathbf{p}'\mathbf{q}}\hat{a}^\dagger_{\mathbf{p}\alpha}\hat{a}^\dagger_{\mathbf{p}'\beta}\hat{a}_{\mathbf{p}'+\mathbf{q}\beta}\hat{a}_{\mathbf{p}-\mathbf{q}\alpha},$$

where the indices $\alpha, \beta = 1, 2$ number the system components which we assume to have equal masses m and concentrations $n_{1,2} = p_F^3/6\pi^2$, μ is the chemical potential, and the constant g characterizes the interparticle interaction which we shall assume to be point (here and subsequently we assume $\hbar = 1$). The specific physical content of the concept of component depends on the particular system. For example, for a solution of ^{3}He in ^{4}He it corresponds to an "upward" and "downward" projection of the spin, whereas in the case of an atomic gas in a magnetic trap, it corresponds to a hyperfine-structure component (or projection of the nuclear spin). The interparticle interaction selected by us assumes that only s-scattering takes place in the system, characterized by the scattering length a. (In the principal order of perturbation theory $a = mg/4\pi$.) The corresponding small dimensionless parameter, the gas parameter λ, is given by

$$\lambda = 2|a|p_F/\pi.$$

We subsequently show how the final result is modified in the presence of scattering in channels with nonzero orbital momenta.

As we well know, the appearance of superfluid pairing is associated with the presence of a pole in the complete two-particle vertex function Γ in the particle–particle channel (Cooper channel) for zero total momentum and frequency [5]. This vertex function Γ is a solution of the Bethe–Salpeter integral equation (Fig. 1):

$$\Gamma(p_1, -p_1;\ p_3, -p_3) = \tilde{\Gamma}(p_1, -p_1;\ p_3, -p_3)$$
$$- T\sum_{n=-\infty}^{+\infty}\int\tilde{\Gamma}(p_1, -p_1;\ q, -q)G(\omega_n, \mathbf{q}) \quad (3)$$
$$\times G(-\omega_n, -\mathbf{q})\Gamma(q, -q;\ p_3, -p_3)\frac{d^3q}{(2\pi)^3},$$

where $\tilde{\Gamma}$ is the irreducible vertex in the Cooper channel (having no singularities at zero total momentum and frequency), G is the single-particle Green's function, and the arguments of the vertex functions denote the corresponding sets of Matsubara frequencies and momenta: $q = (\omega_n, \mathbf{q})$, $p_1 = (\omega_{n1}, p_1)$, and so on. Note that in formula (3) (and in the following formulas) we do not give the explicit form of the indices distinguishing the components of the Fermi gas (for example, Γ should be taken to mean $\Gamma_{\alpha\beta\gamma\delta}$, and so on). Writing

these in explicit form does not present any difficulties. We also note that the nonsymmetrized (in terms of the component indices) irreducible vertex function $\tilde{\Gamma}$ is used in equation (3).

The vertex functions Γ and $\tilde{\Gamma}$ appearing in (3) are in fact functions of the Matsubara frequencies, the moduli of the incoming and outgoing momenta, and the angle between them. For example, we have

$$\Gamma(p_1, -p_1; p_3, -p_3) = \Gamma(\omega_1, \omega_3, |\mathbf{p}_1|, |\mathbf{p}_3|, \cos(\theta_{\mathbf{p}_1\mathbf{p}_3})).$$

Thus, expanding Γ and $\tilde{\Gamma}$ as a series in terms of Legendre polynomials

$$\begin{aligned}\tilde{\Gamma}(\ldots, \cos(\theta)) &= \sum_{l=0}^{+\infty}(2l+1)\tilde{\Gamma}_l(\ldots)P_l(\cos\theta),\\ \Gamma(\ldots, \cos(\theta)) &= \sum_{l=0}^{+\infty}(2l+1)\Gamma_l(\ldots)P_l(\cos\theta),\end{aligned} \tag{4}$$

and integrating over angles, we easily obtain from (3) the following equation for the polar part $\Gamma_l^{(s)}$ of the lth harmonic of the vertex function:

$$\begin{aligned}\Gamma_l^{(s)}(\omega_1, \omega_3, |\mathbf{p}_1|, |\mathbf{p}_3|) &= -T\sum_{n=-\infty}^{+\infty}\int\frac{d^3g}{(2\pi)^3}\\ &\times\tilde{\Gamma}_l(\omega_1, \omega_n, |\mathbf{p}_1|, |\mathbf{q}|)G(\omega_n, \mathbf{q})\\ &\times G(-\omega_n, -\mathbf{q})\Gamma_l^{(s)}(\omega_n, \omega_3, |\mathbf{q}|, |\mathbf{p}_3|).\end{aligned} \tag{5}$$

As usual, the critical temperature corresponds to the appearance of a nontrivial solution of this equation which is associated with singular (logarithmic) behavior of the Cooper loop near the Fermi surface. Thus, in order to determine the critical temperature in the principal order with respect to λ in all the vertex functions contained in (5), it is sufficient to set the frequencies to zero and the moduli of the momenta to p_F. We then have

$$\begin{aligned}-T\sum_n\int\frac{d^3k}{(2\pi)^3}G(\omega_n, q)G(-\omega_n, -q)\\ \longrightarrow \frac{m^*}{m}Z^2\nu_F\ln\frac{\tilde{\varepsilon}}{T_{c1}},\end{aligned} \tag{6}$$

where $\nu_F = mp_F/2\pi^2$ is the density of states at the Fermi surface, m^* is the effective mass, and Z is the residue in the polar part of the single-particle Green's function. In equation (6) $\tilde{\varepsilon} \sim \varepsilon_F$ is the cutoff parameter which depends on the behavior of $\tilde{\Gamma}$ at high momenta and frequencies. Now equation (5) can be rewritten in the form

$$\Gamma_l^{(s)} = \tilde{\Gamma}_l Z^2\frac{m^*}{m}\nu_F\ln\left(\frac{\tilde{\varepsilon}}{T}\right)\Gamma_l^{(s)}, \tag{7}$$

where

$$\tilde{\Gamma}_l = \tilde{\Gamma}_l(p_i = p_F, \omega_i = 0),$$

$$\Gamma_l^{(s)} = \Gamma_l^{(s)}(p_i = p_F, \omega_i = 0),$$

so that a nontrivial solution is only possible for $\tilde{\Gamma} < 0$ and occurs at temperature $T = T_{cl}$ where

$$T_{cl} = \tilde{\varepsilon}\exp\left\{-\frac{1}{\nu_F|\tilde{\Gamma}_l|}\frac{m}{m^*Z^2}\right\}. \tag{8}$$

For the case of a Fermi gas with interparticle attraction we have

$$\tilde{\Gamma}_0 \simeq 4\pi a/m < 0,$$

and the system is unstable with respect to traditional s-pairing ($l = 0$). The superconducting transition temperature in this case was obtained in [1] to within terms $O(\lambda^0)$ inclusive and is given by

$$\begin{aligned}T_{c0} &= \frac{1}{\pi}e^C\left(\frac{2}{e}\right)^{7/3}\varepsilon_F\exp\left\{-\frac{\pi}{2|a|p_F}\right\}\\ &\approx 0.28\varepsilon_F\exp\left\{-\frac{1}{\lambda}\right\},\end{aligned} \tag{9}$$

where $C = 0.58\ldots$ is the Euler constant. This expression only differs from the normal expression for T_{c1} in BCS theory in that the Debye frequency ω_D is replaced by ε_F in the expression for the preexponential function. This replacement means that in this particular case, the entire Fermi sphere and not only its vicinity of the order of the Debye frequency, is involved in the pairing. Note that in order to find the preexponential factor in [1] we need to calculate $\tilde{\Gamma}_0$ to within terms of the second order of perturbation theory inclusive.

For repulsive interaction, $a > 0$, equation (5) for $l = 0$ only has a trivial solution and s-pairing is impossible. In this case, superfluid pairing will take place in the channel having the orbital momentum l for which $\tilde{\Gamma}_l$ is negative and has the maximum absolute value. As we well know [6], the scattering amplitude of slow particles having the orbital momentum l at the short-range potential has the order of magnitude $a(ap)^{2l}$, where p is the particle momentum and a is the s-scattering length. Thus, in our particular case the corresponding contribution to $\tilde{\Gamma}_l$ from scattering at the seed interparticle potential has a maximum for $l = 1$ and has the order

$$(ap_F)^3 \sim \lambda^3.$$

At the same time, many-particle effects associated with polarization processes of the Fermi background in a Fermi gas have the order λ^2 and are consequently a determining factor for $l \geq 4$ [4]. Corresponding dia-

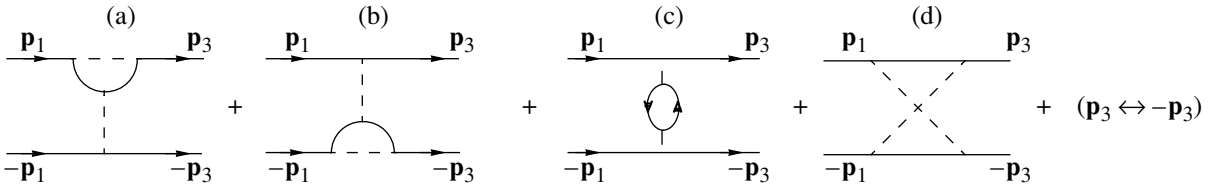


Fig. 2. Second-order diagrams in terms of gas parameter for irreducible interaction.

grams for $\tilde{\Gamma}$ in the second order with respect to λ are plotted in Fig. 2. For our particular case of point interaction the first three diagrams cancel out so that $\tilde{\Gamma}$ is completely determined by the last exchange diagram and is given by (we assumed $T = 0$)

$$\tilde{\Gamma}(\omega_1, \omega_3, \mathbf{p}_1, \mathbf{p}_3) = \left(\frac{4\pi a}{m}\right)^2 \Pi(\omega_1 + \omega_3, \mathbf{p}_1 + \mathbf{p}_3),$$

where

$$\Pi(\Omega, \mathbf{q}) = \int\frac{d\omega}{2\pi}\int\frac{d^3p}{(2\pi)^3}G(\omega, \mathbf{p})G(\Omega + \omega, \mathbf{p} + \mathbf{q}) = \int\frac{d^3p}{(2\pi)^3}\frac{n(\mathbf{p} + \mathbf{q})n(\mathbf{p})}{\Omega + \xi(\mathbf{p} + \mathbf{q}) - \xi(\mathbf{p})}. \quad (10)$$

In this expression $n(p)$ is the Fermi particle distribution function for $T = 0$, $\xi(p) = p^2/2m - \mu$, and θ is the angle between $\mathbf{p}_1$ and $\mathbf{p}_3$.

From formula (10) we can easily obtain an expression for the irreducible vertex at zero external frequencies and momenta lying on the Fermi surface. In terms of the angle θ between $\mathbf{p}_1$ and $\mathbf{p}_3$ we have

$$\nu_F\tilde{\Gamma}(0, 0, p_1 = p_F, p_3 = p_F, \cos\theta) = \frac{\pi\lambda}{2} \times\left(1 - \frac{\lambda}{2}\left[1 + \frac{\sqrt{2}(1 + \cos\theta)}{4\sqrt{1 - \cos\theta}}\ln\frac{\sqrt{2} + \sqrt{1 - \cos\theta}}{\sqrt{2} - \sqrt{1 - \cos\theta}}\right]\right).$$

As a result of integrating using Legendre polynomials we obtain

$$\nu_F\tilde{\Gamma}_1 = \frac{1}{5}\lambda^2(1 - 2\ln 2) < 0. \quad (11)$$

All the other partial components $\tilde{\Gamma}_l$ with $l > 1$ also correspond to attraction but are smaller than $\tilde{\Gamma}_1$ and their absolute value decreases rapidly with increasing l (see [4]). Thus, we conclude that a weakly nonideal Fermi gas with interparticle repulsion is unstable with respect to triplet p-pairing. The corresponding critical temperature in the principal order with respect to λ is

$$T_{c1} = \tilde{\varepsilon}\exp\left\{-\frac{1}{\nu_F|\tilde{\Gamma}_1|}\right\} = \tilde{\varepsilon}\exp\left\{-\frac{5}{(2\ln 2 - 1)}\frac{1}{\lambda^2}\right\}. \quad (12)$$

It can be seen from this formula that in order to determine the preexponential factor $\tilde{\varepsilon}$ in equation (5) we need to retain terms through order λ^4. (This follows from the fact that since $\tilde{\Gamma}_1$ begins from terms λ^2, in order to obtain terms of the order λ^0 in the argument of the exponential function, we need to know $\tilde{\Gamma}_1$ to within terms λ^4 inclusive.)

Note that the contribution of triple collisions can be neglected to within the selected accuracy since this has the order λ^5 [5, §6].

3. CONTRIBUTION OF HIGHER ORDERS OF PERTURBATION THEORY

The irreducible vertex $\tilde{\Gamma}$ in the third and fourth orders of perturbation theory is given by the diagrams shown in Figs. 3 and 4, respectively. The points on these diagrams correspond to antisymmetrized two-particle interaction. In expanded notation when the interaction is represented as a dashed line (as in Fig. 2) these corresponds to two different methods of connecting the incoming and outgoing lines.

Figures 3 and 4 only give skeletal diagrams (neglecting the self-energy components) and Fig. 4 only gives "nonoriented" diagrams. The corresponding Feynman diagrams are obtained by arranging the arrows (taking into account the particle number conservation law at the vertexes) and also the incoming and outgoing momenta. Figure 5 shows an example of such an arrangement.

Direct calculations of these diagrams using standard rules of the diagram technique yield diverging expressions obtained from integrating over large momenta in subdiagrams containing Cooper loops (loops formed from two lines in the same direction). By way of example we consider the first third-order diagram in Fig. 3 together with its corresponding diagram in which **p**. In expanded form this corresponds to the sum of the diagrams in Fig. 6 where the dashed line corresponds to interparticle interaction. It is easily established that for our point potential the first three diagrams cancel out leaving only the fourth diagram which contains a diverging subdiagram at large momenta, corresponding to a Cooper loop between two parallel dashed lines. However, it is easy to see that this subdiagram is the first correction of a ladder series to one of the dashed

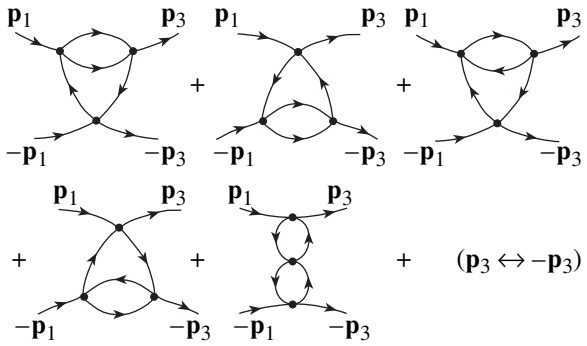


Fig. 3. Skeleton diagrams of the third order of perturbation theory for the irreducible vertex $\tilde{\Gamma}_1$.

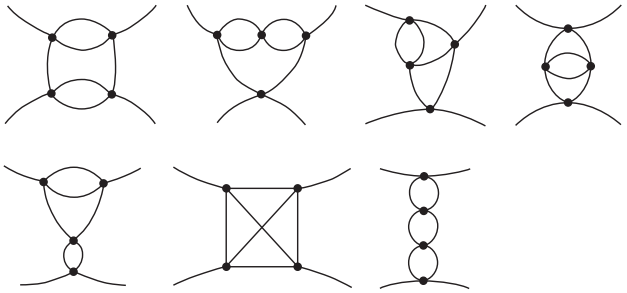


Fig. 4. Skeleton "nonoriented" diagrams of the fourth order of perturbation theory for the irreducible vertex $\tilde{\Gamma}_1$.

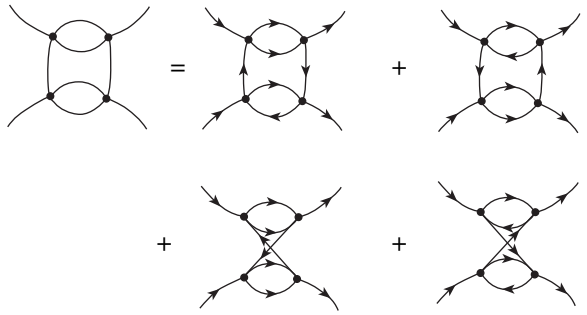


Fig. 5. Example of decoding "nonoriented" diagrams (first diagram in Fig. 4).

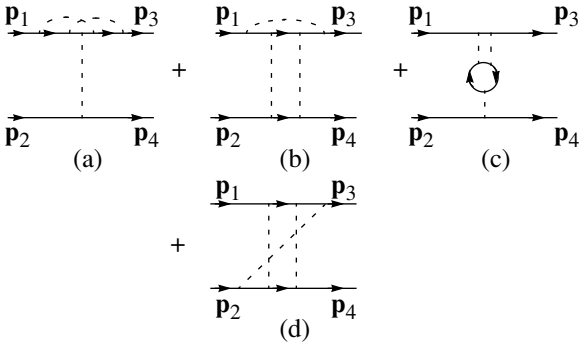


Fig. 6. First third-order diagram from Fig. 3 showing diagram corresponding to the substitution $\mathbf{p}_3 \leftrightarrow -\mathbf{p}_3$ in expanded representation.

lines on the fourth diagram in Fig. 2 which gives a contribution to $\tilde{\Gamma}$ in the second order with respect to λ. Quite clearly, the second diagram in Fig. 3 paired with its corresponding diagram where the substitution $\mathbf{p}$ is made in the sum is the correction to the dashed line in Fig. 2.

These corrections only differ from the first term of the Born series for the scattering amplitude in that they contain the single-particle Green's functions in the medium G and not in vacuum $G^{(0)}$. However, at large momenta the difference between G and $G^{(0)}$ disappears so that the divergence in the diagram in Fig. 7 can be eliminated by changing from the seed interaction g to the scattering length a (renormalization procedure). This length is determined by the scattering amplitude of two particles in vacuum in the limit where the energies of the colliding particles tend to zero and may be obtained from

$$\frac{4\pi a}{m} = g + \int\frac{d\omega}{2\pi}\int\frac{d^3p}{(2\pi)^3} g G^{(0)}(\omega, p) G^{(0)}(-\omega, -p)\frac{4\pi a}{m} = g + \int\frac{d^3p}{(2\pi)^3} g \frac{1}{2\varepsilon(p) + i0}\frac{4\pi a}{m}. \quad (13)$$

After integrating in diagram Fig. 6d over the intermediate frequency of the Cooper loop ω, we obtain the expression

$$\int\frac{d^4p}{(2\pi)^4}\int\frac{d^3p}{(2\pi)^3}\left(\frac{1-\theta(\xi_1)-\theta(\xi_2)}{\Omega-(\xi_1+\xi_2)+i\delta(\mathrm{sgn}\,\xi_1+\mathrm{sgn}\,\xi_2)}\right) \times \frac{1}{(\Omega-\xi_3+i\delta\,\mathrm{sgn}\,\xi_3)(\Omega-\xi_4+i\delta\,\mathrm{sgn}\,\xi_4)}, \quad (14)$$

where

$$\xi_1 = \xi\left(\mathbf{p}+\frac{\mathbf{q}+\mathbf{w}}{2}\right), \quad \xi_2 = \xi\left(-\mathbf{p}+\frac{\mathbf{q}+\mathbf{w}}{2}\right).$$

$$\xi_3 = \xi(\mathbf{q}-\mathbf{s}), \quad \xi_4 = \xi(\mathbf{q}+\mathbf{s}),$$

$$\mathbf{p}_1 = \mathbf{s}+\mathbf{w}, \quad \mathbf{p}_3 = \mathbf{s}-\mathbf{w}.$$

The integral over the internal momentum $\mathbf{p}$ from the expression in brackets diverges at the upper limit. As we have noted, this is because it is the same as the Born correction to the scattering amplitude in this region. This divergence can be eliminated in the second-order diagram in Fig. 2d by going over from the seed interaction constant g to the zero-energy particle scattering amplitude a in the center-of-mass system.

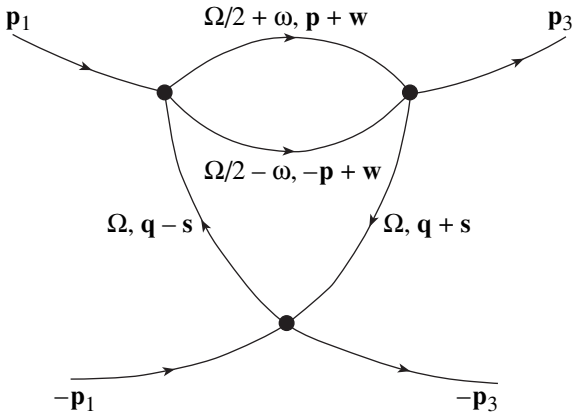


Fig. 7. Diagram of the third order of perturbation theory containing a Cooper loop as subdiagram.

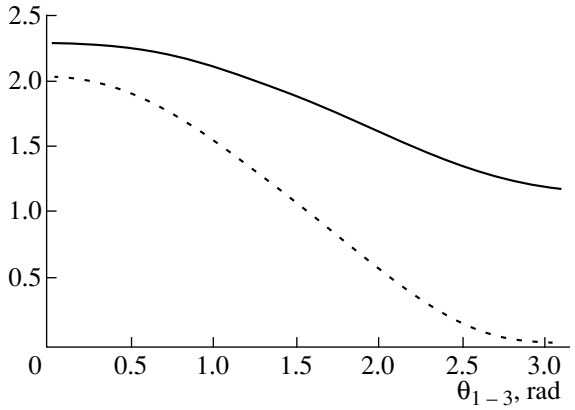


Fig. 8. Dependence of the irreducible vertex in the second and third orders on the angle between the incoming and outgoing momenta θ_{1-3}: $\Gamma^{(2)}/\lambda^2$—solid curve, $\Gamma^{(3)}/\lambda^3$—dashed curve.

To within the required accuracy the relationship between g and a is obtained from (13) and has the form

$$g = \frac{4\pi}{m}a + \left(\frac{4\pi}{m}a\right)^2 \int \frac{d^3p\,d\omega}{(2\pi)^4} G^{(0)}(\omega, \mathbf{p}) G^{(0)}(-\omega, -\mathbf{p}) = \frac{4\pi}{m}a + \left(\frac{4\pi}{m}a\right)^2 \int \frac{d^3p}{(2\pi)^3} \frac{1}{2\varepsilon(p) + i\delta}. \quad (15)$$

where $\varepsilon(p) = p^2/2m$. The renormalization procedure is shown schematically as follows:

$$\nu_F\{g^2\Pi + 2g^3(GG)\tilde{\Pi}\} \longrightarrow \lambda^2\Pi + 2\lambda^3[(GG) - (G^{(0)}G^{(0)})]\tilde{\Pi}, \quad (16)$$

where (GG) corresponds to the first cofactor in formula (14), $\tilde{\Pi}$ corresponds to the second cofactor,

$$(G^{(0)}G^{(0)}) = \int \frac{d\omega}{2\pi} G^{(0)}(\omega, -\mathbf{p}) G^{(0)}(\omega, -\mathbf{p}) = \frac{1}{2\varepsilon(p) + i\delta},$$

and the pair factor in the second term allows for the contribution of the second diagram in Fig. 3. As we can easily see, the expression for $\tilde{\Pi}$, being integrated over frequency, exactly gives the polarization operator Π which appears in the first term in formula (14). The last term in formula (14) has the explicit form

$$\Gamma^{(3a)} = -2i\left(\frac{2\pi^2\lambda}{m p_F}\right)^3 \int \frac{d^4p}{(2\pi)^4} \int \frac{d^3p}{(2\pi)^3} \times \left(\frac{1 - \theta(\xi_1) - \theta(\xi_2)}{\Omega - (\xi_1 + \xi_2) + i\delta(\mathrm{sgn}\xi_1 + \mathrm{sgn}\xi_2)} - \frac{1}{2\varepsilon(p) + i\delta}\right) \times \frac{1}{(\Omega - \xi_3 + i\delta\,\mathrm{sgn}\xi_3)(\Omega - \xi_4 + i\delta\,\mathrm{sgn}\xi_4)}. \quad (17)$$

This expression contains no divergences and can be integrated numerically. It can be established that all the order third-order diagrams contain no divergences and, as a result of numerical calculations, we obtain the final result for the third-order contribution to the p-harmonic of the irreducible vertex:

$$\nu_F\tilde{\Gamma}_1^{(3)} \simeq -0.33\lambda^3. \quad (18)$$

It should be noted that formula (18) contains no contribution from Hartree–Fock self-energy components in the second-order diagrams since this contribution corresponds to renormalization of the chemical potential. We also note that the appearance of a large numerical coefficient 0.33 (compared with the coefficient of 0.077 for the second-order contribution) is associated with the stronger angular dependence of the third-order diagrams (see Fig. 8). This dependence is mainly determined by the first two diagrams in Fig. 3 and can be attributed to the existence of subdiagrams with Cooper loops.

All divergences in the fourth-order diagrams are eliminated in exactly the same way. For this purpose, in the third-order diagrams we need to go over from g to a in accordance with formula (15) and in the second-order diagrams, renormalize in the expression for g in terms of a allowing for the term $\sim a^3$, which can easily be obtained from equation (13). (This term is required to eliminate the divergences in the second diagram in Fig. 4.) As a result, the contribution of the fourth-order diagrams in Fig. 4 is given by

$$\nu_F\tilde{\Gamma}_1^{(4)} \simeq -0.39\lambda^4. \quad (19)$$

In order to calculate $\tilde{\Gamma}_1$ to within λ^4 we also need to allow for the contribution of the self-energy components of the second order in λ in the second-order diagrams $\tilde{\Gamma}_1^{(2)}$, see Fig. 2. These contributions can no longer be reduced to renormalization of the chemical

potential. They also result in the appearance of a nontrivial Z-factor and the effective mass m^* [7]:

$$Z = 1 - \lambda^2 \ln 2, \quad \frac{m^*}{m} = 1 + \frac{2}{15}(7\ln 2 - 1)\lambda^2 \quad (20)$$

in the polar part of the single-particle Green's function which now also contains a nonpolar part proportional to λ^2. By means of direct numerical calculations of the corresponding diagrams we can establish that the contribution of the latter to $\tilde{\Gamma}_1$ is negligible. Thus, we finally obtain the following expression with the required accuracy in terms of λ for the irreducible vertex in a Cooper channel with orbital momentum $l = 1$:

$$\nu_F Z^2 \frac{m^*}{m}\tilde{\Gamma}_1 = \nu_F Z^2 \frac{m^*}{m}\left(Z^2\frac{m^*}{m}\tilde{\Gamma}_1^{(2)} + \tilde{\Gamma}_1^{(3)} + \tilde{\Gamma}_1^{(4)}\right) = -0.077\lambda^2 - 0.33\lambda^3 - 0.26\lambda^4. \quad (21)$$

4. ALLOWANCE FOR RETARDATION EFFECTS

In order to determine the critical temperature in Section 2, in equation (5) we replaced the irreducible vertex $\tilde{\Gamma}_1$, which is a function of the incoming and outgoing frequencies and moduli of momenta $\tilde{\Gamma}_1(\omega_i, \mathbf{p}_i)$, by its value at zero frequencies and momenta lying on the Fermi surface $\tilde{\Gamma}_1(\omega_i = 0, \mathbf{p}_i = p_F)$. In this section we shall show that allowance for the difference between $\tilde{\Gamma}_1(\omega_i, \mathbf{p}_i)$ and $\tilde{\Gamma}_1(\omega_i = 0, \mathbf{p}_i = p_F)$ (retardation effects) introduces a correction of the order of λ^4 to the vertex $\tilde{\Gamma}_1$. In other words, these effects influence the numerical coefficient in the preexponential factor.

Retardation effects are most conveniently taken into account using a method proposed in [8, Chapter 2]. Omitting the appropriate procedures, which are a trivial generalization of the derivation of [8] to the case of p-pairing, we arrive at the following integral equation:

$$\Phi_1(\xi) = -\int_{-\varepsilon_F}^{\infty} d\xi' \frac{\tanh(\xi'/2T)}{2\xi'} R_1(\xi, \xi')\Phi_1(\xi'), \quad (22)$$

for which the condition for existence of a trivial solution determines the critical temperature T_{c1}. The unknown function Φ_1 in (22) can be related to the spectral density of the anomalous Green's function (or more accurately its first harmonic in the expansion in terms of Legendre polynomials), and the kernel $R_1(\xi, \xi')$ is given by

$$R_1(\xi, \xi') = \frac{m}{4\pi^2 p^2(\xi)} \int_{|p(\xi) - p(\xi')|}^{p(\xi) + p(\xi')} q\,dq \times \int_0^{\infty} \frac{dE\,\sigma(E, m)}{E + |\xi| + |\xi'|} \frac{p^2(\xi) + k^2(\xi') - q^2}{2k(\xi')}, \quad (23)$$

where $\sigma(E, q)$ is related to $\tilde{\Gamma}(\omega = \omega_1 - \omega_3, \mathbf{q} = \mathbf{p}_1 - \mathbf{p}_3)$ by

$$\tilde{\Gamma}(i\omega_n, \mathbf{p}) = \int_0^{\infty} \frac{dE^2 \sigma(E, \mathbf{p})}{E^2 + \omega_n^2}, \quad (24)$$

and the factor $((p^2 + k^2 - q^2)/2pk$ is precisely the cosine of the angle between the incoming and outgoing momenta which isolates the first harmonic in the expansion (4) in terms of Legendre polynomials.

Dividing the region of integration over ξ' in equation (22) into three parts: $|\xi'| \le z\varepsilon_F$, $-\varepsilon_F < \xi' < -z\varepsilon_F$, and $\xi' > z\varepsilon_F$, where z is an arbitrary number satisfying the condition $T_{c1} \ll z\varepsilon_F \ll \varepsilon_F$, and integrating by parts (where the dependence on ξ' in $R(\xi, \xi')$ and $\Phi_1(\xi')$ can be neglected in the first region and the hyperbolic tangent in the second and third regions can be replaced by ∓ 1, respectively), equation (22) can be reduced to the form

$$\Phi_1(\xi) = -\ln\left(\frac{2\gamma\varepsilon_F}{\pi T_{c1}}\right)\Phi_1(0)R_1(\xi, 0) + \frac{1}{2}\int_{-\varepsilon_F}^{\infty} d\xi' \ln\left(\frac{|\xi'|}{\varepsilon_F}\right)\frac{d}{d\xi'}(R_1(\xi, \xi')\Phi_1(\xi')). \quad (25)$$

(As was to be expected, the arbitrary constant z was dropped from this equation.) We introduce the new variable

$$\chi(\xi) = \frac{\Phi_1(\xi)}{\Phi_1(0)\ln\dfrac{\pi T_{c1}}{2\gamma\varepsilon_F}}, \quad (26)$$

which allows us to write the expression for the critical temperature in the form

$$T_{c1} = \frac{2e^C}{\pi}\varepsilon_F \exp\left(-\frac{1}{\chi(0)}\right), \quad (27)$$

where the function $\chi(\xi)$ satisfies

$$\chi(\xi) = R_1(\xi, 0) + \frac{1}{2}\int_{-\varepsilon_F}^{\infty} d\xi' \ln\left(\frac{|\xi'|}{\varepsilon_F}\right)\frac{d}{d\xi}(R_1(\xi, \xi')\chi(\xi')). \quad (28)$$

Since the kernel R_1 contains the small parameter ($R \sim \lambda^2$), equation (22) can be solved by an iterative method. In the zeroth approximation we set:

$$\chi^{(0)}(\xi) = R_1(\xi, 0).$$

The first correction $\chi^{(1)}$ is given by the integral on the right-hand side of (28) with $\chi = \chi^{(0)}$:

$$\chi^{(1)} = \frac{1}{2}\int_{-\varepsilon_F}^{\infty} d\xi' \ln\left(\frac{|\xi'|}{\varepsilon_F}\right)\frac{d}{d\xi}(R_1(\xi, \xi')R_1(\xi', 0)) \quad (29)$$

and, as can easily be seen, begins with terms of the order λ^4. The principal term with respect to λ in χ_1 is

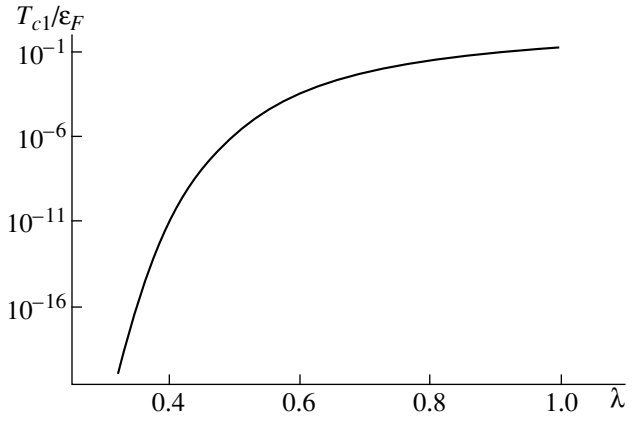


Fig. 9. Dependence of T_{c1}/ε_F on the gas parameter λ.

obtained if only the leading ($\sim\lambda^2$) terms are retained in the kernel R_1 in formula (29). In this case, the spectral function $\sigma(E, \mathbf{q})$ is the same as the imaginary part of the polarization operator

$$\sigma(E, \mathbf{q}) = -\frac{1}{\pi}\mathrm{Im}\Pi(E, \mathbf{q})$$

$$= -\frac{1}{\pi}\begin{cases} -\dfrac{mp_F}{4\pi\tilde{q}}\left[1-\left(\dfrac{\tilde{E}}{\tilde{q}}-\dfrac{\tilde{q}}{2}\right)^2\right] \text{ for } \left|\dfrac{\tilde{q}^2}{2}-\tilde{q}\right| \le \tilde{E} \le \dfrac{\tilde{q}^2}{2}+\tilde{q}, \\ -\dfrac{mp_F}{4\pi\tilde{q}}2\tilde{E} \text{ for } 0 \le \tilde{E} \le \tilde{q}-\dfrac{\tilde{q}^2}{2}, \end{cases}$$

where

$$\tilde{q} = q/p_F, \quad \tilde{E} = Em/p_F^2.$$

Direct calculations using formula (29) give $\chi^{(1)} \ge 0.004\lambda^4$ which is equivalent to adding $\Delta\tilde{\Gamma}_1$ in formula (8),

$$\nu_F\Delta\tilde{\Gamma}_1 \simeq 0.004\lambda^4. \tag{30}$$

Note that a similar estimate of the contributions of retardation effects was made in [9] where the authors used the step function approximation for the frequency dependence of the polarization operator.

5. CRITICAL TEMPERATURE T_{c1}

Collecting these results together [formulas (21) and (30)], we obtain the following expression for the critical temperature which is determined numerically to within two decimal places:

$$T_{c1} = \frac{2}{\pi}e^C\varepsilon_F\exp\{-(0.077\lambda^2+0.33\lambda^3+0.26\lambda^4)^{-1}\}$$
$$\xrightarrow{\lambda\to 0} \frac{2}{\pi}e^C\varepsilon_F\exp\left\{-\frac{13.0}{\lambda^2}+\frac{42.0}{\lambda}-190\right\}, \tag{31}$$

where the omitted terms have the order λ. This last formula assumes $\lambda < 0.23$ since for $\lambda = 0.23$ the second- and third-order terms with respect to λ in the exponential function in (31) are the same. For $0.23 \le \lambda \le 1$ the fourth-order term with respect to λ in (31) is much smaller than the first two so that (31) can accurately be rewritten in the form

$$T_{c1} \approx \frac{2}{\pi}e^C\varepsilon_F\exp\left\{-\frac{13.0}{\lambda^2(1+4.3\lambda)}+\frac{42.0}{(1+4.3\lambda)^2}\right\}. \tag{32}$$

This formula may be considered as an extrapolation of the expression for the critical temperature from $\lambda \ll 1$ [formula (31)] to the region $\lambda < 1$ [formula (32)]. A graph of the dependence $T_{c1}(\lambda)$ is plotted in Fig. 9.

6. INFLUENCE OF SEED *p*-SCATTERING

So far we have only considered *s*-scattering between particles, assuming that the interparticle potential is a point. However, as we have already noted for the finite-radius potential, the problem will always contain scattering having an arbitrary orbital momentum l whose amplitude for particles having momenta equal to the Fermi momentum p_F may be estimated as $f_l \sim a(ap_F)^{2l}$ [6]. From this it follows that with the required accuracy we can confine our analysis to *p*-scattering ($l = 1$). In this case, only two contributions will be important: a contribution of the order λ^3 from *p*-scattering at the seed interparticle potential and a contribution of the order λ^4 corresponding to the diagram in Fig. 2d where one of the dashed lines corresponds to *s*-scattering and the other to *p*-scattering. More precisely, if the amplitude of *p*-scattering of two particles having momenta p_F is written in the form

$$f_1 = \alpha_1 a\left(\frac{2ap_F}{\pi}\right)^2 = \alpha_1 a\lambda^2, \tag{33}$$

where α_1 is a numerical coefficient, the contribution of triplet scattering to the irreducible vertex $\tilde{\Gamma}_1$ has the form

$$\nu_F\tilde{\Gamma}_1 = \alpha_1\lambda^3(1+0.008\lambda).$$

We can see that the fourth-order contribution with respect to λ can be neglected and consequently the critical temperature is given by

$$T_{c1} = \frac{2}{\pi}e^C\varepsilon_F \tag{34}$$
$$\times\exp\left\{-\frac{13.0}{\lambda^2[1+(4.3+\alpha_1)\lambda]}+\frac{42.0}{[1+(4.3+\alpha_1)\lambda]^2}\right\}.$$

Nevertheless, we can specify a physical situation when the contribution of *p*-scattering can be neglected. This corresponds to the case when a shallow level having the orbital momentum $l = 0$ (resonance scattering)

exists in the potential. In this case, the *p*-scattering amplitude is estimated as

$$f_1 \sim r_0(r_0 p_F)^2,$$

where r_0 is the radius of action of the potential while the *s*-scattering amplitude (length) is given by

$$a = (1/2m|E|)^{1/2} \gg r_0,$$

where E is the discrete level energy (we assume that the condition $|E| \gg \varepsilon_F$ is satisfied so that $ap_F \ll 1$). Then for α_1 in formula (1) we obtain the estimate

$$\alpha_1 \sim \left(\frac{r_0}{a}\right)^3 \ll 1,$$

and if the condition

$$\alpha_1 \ll \lambda$$

is satisfied, the contribution of the *p*-harmonic of the seed interparticle interaction can be neglected compared with the fourth order of the effective interaction which allows only for *s*-scattering.

7. CRITICAL TEMPERATURE IN A MAGNETIC FIELD

In this section we study the influence of an external magnetic field on the irreducible vertex $\tilde{\Gamma}_1$ and consequently on the critical temperature T_{c1} to within terms of the order λ^3. As was shown in [10], in the principal approximation with respect to λ in the model being studied the critical *p*-pairing temperature may increase appreciably if a static magnetic field is applied to the system. This is because for conventional singlet pairing the role of a magnetic field is always destructive due to of the paramagnetic suppression of Cooper pairing caused by the flipping of one of the pair spins. However for triplet *p*-pairing no paramagnetic effect occurs so that the role of the magnetic field was not clear a priori.

In this approach the mechanism for variation of T_{c1} in a magnetic field is based on the magnetic field dependence of the many-particle effects which determine the effective interaction. On the one hand, as a result of a difference in the number of particles (and consequently Fermi momenta) having spins directed parallel and antiparallel to the field, the Kohn singularity increases sharply, causing an increase in $\tilde{\Gamma}_1$. On the other hand, the value $\nu_{F\downarrow}\tilde{\Gamma}_1$ decreases with increasing magnetic field because of a monotonic decrease in the number of particles with spin antiparallel to the field. (We recall that *s*-scattering can only occur between Fermi particles having different spin projections.) Competition between these two effects leads to an abruptly nonmonotonic dependence of the critical temperature on the magnetic field (more accurately, on the degree of polarization $\alpha = (n_\uparrow - n_\downarrow)/(n_\uparrow + n_\downarrow)$) with an abrupt increase in T_{c1} for small α, a maximum at intermediate α, and a decrease for $\alpha \longrightarrow 1$. (In this case of a completely polarized Fermi gas seed *p*-scattering between parallel spins can only take place between the particles.) The dependence of $\tilde{\Gamma}$ on the polarization α to within second-order terms was calculated in [10]:

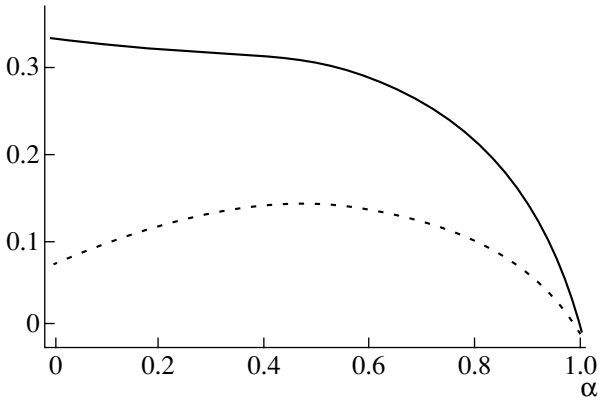


Fig. 10. Dependence of the second- and third-order contributions to the irreducible vertex on the degree of polarization α: $\Gamma^{(2)}/\lambda^2$—dashed curve, $\Gamma^{(3)}/\lambda^3$—solid curve.

$$\tilde{\Gamma}^{(2)}(\delta) = -\lambda^2 \frac{2\ln 2 - 1}{5}\frac{1}{\delta^3} \times \left(\frac{2}{1+\delta^3}\right)^{2/3}\left[1 + \frac{\delta - 1}{3(2\ln 2 - 1)}\Psi_\delta\right], \tag{35}$$

where

$$\Psi_\delta = (\delta+1)[10\ln(\delta+1) - \delta^2 - 3] + \frac{\delta-1}{2}(\delta^3 + 2\delta^2 + 8\delta + 4)\ln\frac{\delta+1}{\delta-1} + \frac{6}{\delta-1}\ln\frac{(\delta+1)}{2},$$

$$\delta = \frac{p_{F\uparrow}}{p_{F\downarrow}} = \left(\frac{1+\alpha}{1-\alpha}\right)^{1/3}.$$

In the third order with respect to λ the result can only be obtained numerically. The corresponding contribution is given by the diagrams in Fig. 3 where the spins on the outer lines are directed parallel to the field and those in the inner loops can be oriented either parallel or antiparallel to the field. The calculations (including renormalization of the diverging diagrams) are exactly the same as the case of no external magnetic field and the result is plotted in Fig. 10 (solid curve) which also gives the second-order contribution (35) for comparison (dashed curve).

It can be seen that the maximum of $\tilde{\Gamma}_1^{(2)}(\alpha)$ is obtained at $\alpha_m = 48$ whereas $\tilde{\Gamma}_1^{(3)}(\alpha)$ decreases mono-

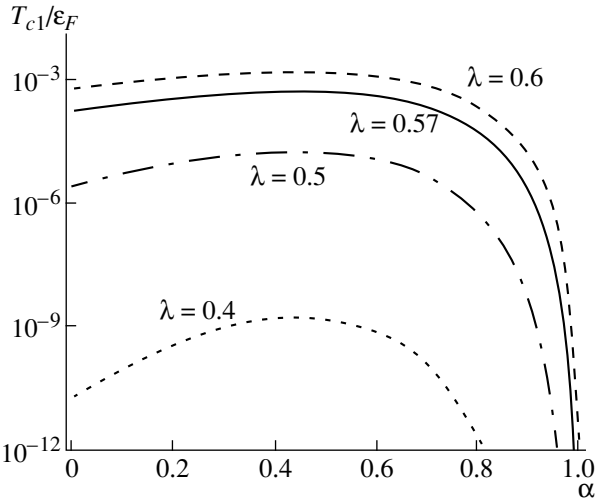


Fig. 11. Dependence of T_{c1}/ε_F on the degree of polarization α for various λ.

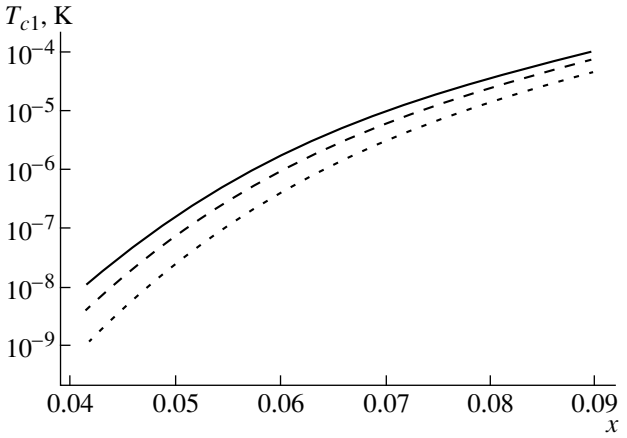


Fig. 12. Dependence of T_{c1} on the concentration x in a solution of ^{3}He in ^{4}He for various degrees of polarization: α = 0.2—solid curve, α = 0.1—dashed curve, and α = 0—dotted curve.

tonically. Thus, the maximum of T_{c1} is determined by competition between increasing $\tilde{\Gamma}_1^{(2)}(\alpha)$ and decreasing $\tilde{\Gamma}_1^{(3)}(\alpha)$. For typical λ this is in the region of 0.4. Graphs of the critical temperature as a function of the degree of polarization are plotted in Fig. 11 for typical values of λ. For λ = 0.6 the value of T_{c1} at the maximum is approximately six times the value in the absence of the field. In this case the maximum is mainly determined by the second order and is reached at λ ~ 0.45.

8. DISCUSSION OF RESULTS

The experimental search for nontrivial $l \neq 0$ pairing in isotropic Fermi systems has recently been actively pursued. Until recently the main candidate was a solution of ^{3}He in ^{4}He. So far superfluidity has not yet been observed in this system although temperatures of the order of 97 mK have been achieved experimentally [11]. In the concentration range $x < x_0 \approx 3\%$ the scattering length corresponds to attraction so that singlet *s*-pairing may be achieved. The critical temperature is given by formula (9) allowing for

$$\varepsilon_F = \varepsilon_{F0} x^{2/3}, \quad p_F = p_{F0} x^{1/3},$$

where ε_{F0} and p_{F0} are the Fermi energy and momentum of pure ^{3}He. According to estimates made in [12], we have

$$\max T_{c0} = T_{c0}(1\%) \approx 10^{-4}\ \text{K}.$$

The authors of [13] predict an even lower critical temperature:

$$\max T_{c0} = T_{c0}(2\%) \approx 4 \times 10^{-6}\text{–}10^{-5}\ \text{K}.$$

Note that the value $T_{c0} \approx 10^{-5}$ K was obtained in [13] as a fitting parameter to describe magnetostriction experiments and $T_{c0} \approx 4 \times 10^{-6}$ K was obtained in spin diffusion experiments. It should be noted that for a given concentration x the gas parameter of the theory $ap_{F0}x^{1/3}$ depends weakly on pressure. Hence the pressure cannot be considered as an instrument to obtain optimum parameters for *s*-pairing.

For high concentrations ($x > x_0$) the scattering length changes sign $a > 0$ and *s*-pairing becomes impossible. Nevertheless, in this case the subsystem of ^{3}He atoms may become superfluid but because of the instability with respect to *p*-pairing. The critical temperature is given by formula (32) with λ replaced by $\lambda x^{1/3}$ and ε_F replaced by $\varepsilon_{F0}x^{1/3}$. It has a maximum at $P = 10$ atm when the maximum ^{3}He concentration of 9.5% is achieved. Figure 12 gives the dependence of T_{c1}on the concentration calculated by using the extrapolation formula (32). At maximum concentration $x = 9.5\%$ the temperature T_{c1} is of the order of 10^{-5}K. A further increase in T_{c1} in solution may occur in strong magnetic fields. For example, at $x = 9.5\%$ the maximum of T_{c1} in a field is more than six times that in the absence of a field, leading us to experimentally measurable temperatures of 6×10^{-5} K.

Recently the properties of Bose-condensed gases of alkali elements in traps (^{23}Na, ^{7}Li, ^{87}Rb) have been studied intensively. A combination of laser and evaporative cooling in magnetic traps can reach gas-phase densities of the order of 10^{12}–10^{14} cm^{-3} and temperatures of 10^{-6}–10^{-8} K. In addition these elements have a long scattering length *a* of quasi-resonant origin. For Rb and Na the scattering lengths are positive. It is also found that the scattering length may cover a broad spectrum of values from negative to positive as a result of the Feshbach effect. This effect was observed for ^{23}Na [14].

A logical continuation of studies of Bose condensation in gases of low-density alkali elements in magnetic traps would be to obtain Fermion superfluidity in low-density Fermion systems in bounded geometry. The case of a negative scattering length makes it possible to achieve *s*-pairing with a transition temperature determined by formula (9). For ^{6}Li, for example, we have $a = -2.3 \times 10^{-3}$ Å < 0. Thus, for $n \sim 10^{14}$ cm^{-3} the critical temperature T_{c0} is of the order of magnitude of 10^{-6} K. Note that because of the Pauli principle the wave function of an $l = 0$ Cooper pair should be antisymmetric with respect to the transposition of quantum numbers characterizing the internal state of the atoms forming the pair. These numbers are indices determining the multiplet component of the hyperfine interaction for the case of zero (optical trap) or weak magnetic field, or projections of the nuclear spin when the strong external magnetic field of the trap destroys the hyperfine coupling. Thus, $l = 0$ pairing can only take place between atoms of different gas components. This imposes a very stringent constraint on the closeness of their densities from the experimental point of view:

$$|n_1 - n_2|/(n_1 + n_2) \leq T_{c0}/\varepsilon_F \ll 1.$$

In the opposite case the Cooper pair would have a velocity higher than the critical velocity $v_c \sim T_{c0}/p_F$. It may therefore prove difficult to achieve this type of pairing experimentally.

For *p*-pairing, a Cooper pair may be formed by atoms of the same component (an analog of the A2 phase in superfluid ^{3}He). Note that the superfluid transition temperature in the triplet case may be increased substantially by using the existence of several components in the trap. This increase is similar to the increase in T_{c1} in a magnetic field and is associated with the idea of channel separation: Cooper pairing is achieved between particles of one component as a result of the polarization of the other components. In this case, it is possible to obtain a superfluid *p*-pairing temperature of the order of 10^{-7}–10^{-5} K, which is quite feasible experimentally. By virtue of this fact this type of pairing may be quite promising from the experimental point of view.

ACKNOWLEDGMENTS

In conclusion the authors thank D. Rainer, Kh. Kapel, G. Frossati, R. Iokhemsen, A.F. Andreev, I.A. Fomin, I.M. Suslov, Yu. Kagan, A.S. Aleksandrov, and A.A. Golubov for many useful discussions. This work was supported by the Russian Foundation for Basic Research (project nos. 98-02-17077, 97-02-16532, and 96-15-96942) and INTAS (project no. 97-0963).

REFERENCES

1. L. P. Gor'kov and T. K. Melik-Barkhudarov, Zh. Éksp. Teor. Fiz. **40**, 1452 (1961) [Sov. Phys. JETP **13**, 1018 (1961)].
2. W. Kohn and J. H. Luttinger, Phys. Rev. Lett. **15**, 524 (1965).
3. D. Fay and A. Layzer, Phys. Rev. Lett. **20**, 187 (1968).
4. M. Yu. Kagan and A. V. Chubukov, Pis'ma Zh. Éksp. Teor. Fiz. **47**, 525 (1988) [JETP Lett. **47**, 614 (1988)].
5. E. M. Lifshitz and L. P. Pitaevskiĭ, *Course of Theoretical Physics. Vol. 5. Statistical Physics* (Nauka, Moscow, 1978; Pergamon Press, New York, 1980), Chap. 2.
6. L. D. Landau and E. M. Lifshitz, *Quantum Mechanics: Non-Relativistic Theory* (Nauka, Moscow, 1989, 4th ed.; Pergamon Press, Oxford, 1977, 3rd ed.).
7. A. A. Abrikosov, L. P. Gor'kov, and I. E. Dzyaloshinskiĭ, *Methods of Quantum Fields Theory in Statistical Physics* (Nauka, Moscow, 1962; Prentice-Hall, Englewood Cliffs, 1963).
8. *Problems in High-Temperature Superconductivity*, Ed. by V. L. Ginzburg and D. A. Kirzhnits (Nauka, Moscow, 1977).
9. A. S. Alexandrov and A. A. Golubov, Phys. Rev. B **45**, 4769 (1992).
10. M. Yu. Kagan and A. V. Chubukov, Pis'ma Zh. Éksp. Teor. Fiz. **50**, 483 (1989) [JETP Lett. **50**, 517 (1989)].
11. G.-H. Oh, Y. Ishimoto, T. Kawae, *et al.*, J. Low Temp. Phys. **95**, 525 (1994).
12. E. Ostgaard and E. P. Bashkin, Physica B (Amsterdam) **178**, 134 (1992).
13. P. G. van der Haar, G. Frossati, and K. S. Bedell, J. Low Temp. Phys. **77**, 35 (1989).
14. J. Stenger, S. Inouye, M. R. Andrews, *et al.*, Phys. Rev. Lett. **82**, 2422 (1999).
15. M. A. Baranov, M. Yu. Kagan, and Yu. Kagan, Pis'ma Zh. Éksp. Teor. Fiz. **64**, 301 (1996) [JETP Lett. **64**, 301 (1996)].

Translation was provided by AIP

Journal of Experimental and Theoretical Physics, Vol. 90, No. 5, 2000, pp. 872–880.
Translated from Zhurnal Éksperimental'noĭ i Teoreticheskoĭ Fiziki, Vol. 117, No. 5, 2000, pp. 1003–1012.

SOLIDS
Electronic Properties

Measurement of Landau Coefficients in Ferroelectrics Using Nonlinear Dielectric Spectroscopy

S. P. Palto*, G. N. Andreev, N. N. Petukhova, S. G. Yudin, and L. M. Blinov
Shubnikov Institute of Crystallography, Russian Academy of Sciences, Moscow 117333, Russia
**e-mail: lev@glasnet.ru; lbf@ns.crys.ras.ru*
Received December 9, 1999

Abstract—A method of measuring the phenomenological Landau–Ginzburg coefficients in an expansion of the free energy of ferroelectrics near the phase transition is proposed. This method is based on simultaneously recording five Fourier components of the current response to a sinusoidal voltage as a function of temperature. An analysis is made of a specific application of the method to Langmuir–Blodgett ultrathin ferroelectric films. *© 2000 MAIK "Nauka/Interperiodica".*

1. INTRODUCTION

A vast number of experimental methods are now available for studying polar structures, including ferroelectrics [1, 2]. In most cases, each method is generally based on measuring a specific physical quantity (such as the pyroelectric coefficient, polarization, dielectric susceptibility, and so on). Studies of phase transitions require temperature measurements. Since a single measurable quantity is generally insufficient to gain a complete understanding of the phase transition, investigations must be made by other methods which frequently involves not only repeated temperature cycles but also preparing new samples having suitable characteristics for a particular method. All this undoubtedly leads to additional errors caused by changes in the structure of the sample as a result of temperature cycles and by the characteristics of "new" samples prepared for a different experiment. In the present study we consider a method which, in the authors' opinion, can fairly comprehensively characterize ferroelectric phase transitions and which is based on measurements in a single temperature cycle. The method involves simultaneously recording the linear and nonlinear dielectric properties of ferroelectrics which are particularly clearly pronounced near the phase transition. It should be noted that the idea of studying the nonlinear dielectric properties in order to obtain information on the phase transition parameters such as the Landau–Ginzburg coefficients in the expansion of the free energy is not a new one. For example, in [1] (relevant literature is cited there) we can familiarize ourselves with several approaches to determining the Landau–Ginzburg coefficients based on studying the nonlinearity in the paraelectric phase. The most significant development in methods of studying nonlinear dielectric properties was clearly achieved in [3–5] in which the first three harmonic components of the response to a sinusoidal voltage were recorded. However, in the view of the present authors, the methodological approach presented in [3, 5] has some inaccuracies which may substantially influence the interpretation of the data. The authors [3–5] attribute the number of the recorded harmonic solely to the magnitude of the dielectric susceptibility of the corresponding order. For example, the third-order dielectric susceptibility is exclusively attributed to the third harmonic. This is incorrect, even in the limit of very weak fields. Since the nth-order dielectric susceptibility is by definition a coefficient in the term proportional to the nth power of the electric field in the expansion of the electric displacement as a power series, even if the third-order susceptibility is zero, a nonzero fifth-order susceptibility will lead to the appearance of the third as well as the fifth harmonic. In other words, the authors do not analyze possible intermodulation contributions caused by the presence of higher-frequency spectral components in the response. The intermodulation contributions may exceed the fundamental contribution, which means that all the harmonic components in the signal spectrum must be measured so that they can be taken into account in the interpretation of the results. There are also various other important characteristics which, in our view, have been omitted or not analyzed in the literature. For example, data obtained in [3] for ferroelectric films of 75/25 vinylidene fluoride/trifluoroethylene copolymer using nonlinear dielectric spectroscopy are at variance with other results, even with regard to the order of the phase transition, which is acknowledged by the authors but the reasons for the contradictions are not clarified. Thus, by developing a method of measuring nonlinear dielectric properties, we can give a clear idea of what is specifically being measured experimentally and how the measured quantities relate to the model parameters of the theory. In the present study we demonstrate that it is important to measure the first five harmonics to adequately determine the Landau coefficients. Simultaneous measurements of the linear and nonlinear contributions result in a new quality: not only can all the coef-

ficients of the Landau–Ginzburg model be measured in a single temperature cycle and the nature of the phase transition thus determined but also information can be obtained on various characteristics such as the spontaneous polarization and pyroelectric coefficient.

The samples to which the method is applied are Langmuir–Blodgett films of 70/30 vinylidene fluoride/trifluoroethylene (PVDF–TFE) copolymer. First, PVDF–TFE has been extensively studied, including using nonlinear dielectric spectroscopy [3, 5]. Second, Langmuir–Blodgett ferroelectric films are fundamentally new multilayer systems whose thickness can be varied to within a single monomolecular layer [6, 8]. Thus, the experimental data obtained in the present study are also of independent interest and importance.

2. EXPERIMENTAL METHOD

The apparatus is shown schematically in Fig. 1. The sample temperature is varied between –20 and +120°C in a thermostat using Peltier elements and a platinum resistance as temperature sensor. The measuring system is a set of virtual devices (generator, synchronous detector, digital oscilloscope, and so on) realized by the PhysLab program, developed for a different purposes [9].

The virtual generator synthesizes the sinusoidal voltage applied to the sample and the amplitude and phase relations of the harmonic components of the current response are recorded using the virtual synchronous detector. PhysLab can be used with a sinusoidal voltage at frequencies between 10 Hz and 20 kHz and amplitude between 1 and 2500 mV without an external amplifier. The load for the current measurements is 100 Ω. The PhysLab synchronous detector can simultaneously record not only the real and imaginary components of the current at the fundamental frequency but can also record the corresponding components at different harmonics. This possibility of simultaneously recording the harmonics forms the basis of the proposed method of studying the temperature behavior of the linear and nonlinear components of the dielectric susceptibility typical of ferroelectric phase transitions. The total coefficient of the intrinsic nonlinear distortions of the generator and the PhysLab recording system is approximately 0.005% so that the nonlinear properties of our samples can be studied up to the fifth harmonic with a large margin.

In accordance with the equivalent circuit diagram of the sample shown in Fig. 1, the current in the circuit is given by

$$I = I_R + I_C = \frac{U}{R} + \frac{d(CU)}{dt}, \tag{1}$$

where I_R and I_C are the resistive and capacitive components of the current, respectively, R is the sample resistance, and U is the sample voltage. In (1) the sample capacitance C cannot be removed from the derivative operator because in nonlinear dielectrics it is a function

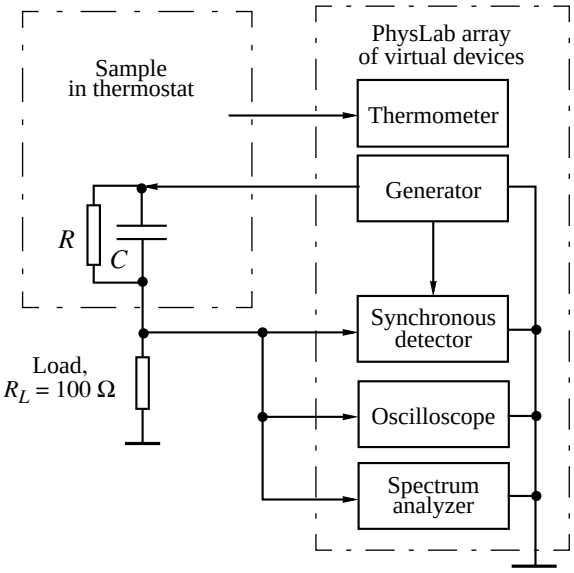


Fig. 1. Schematic of apparatus.

of voltage. Taking this into account, the capacitive current component can be expressed in the form

$$I_C = C(U)\frac{dU}{dt} + U(t)\frac{dC(U)}{dt} \simeq \left(C_0 + U(t)\frac{dC}{dU}\right)\frac{dU}{dt} + U(t)\frac{dC}{dU}\frac{dU}{dt} = C_0\frac{dU}{dt} + 2U(t)\frac{dC}{dU}\frac{dU}{dt}, \tag{2}$$

where $C_0 = C(U = 0)$. In (2) we used the weak dependence of the capacitance on the field at low sample voltages, confining ourselves to two terms in the Taylor series. An analysis shows that these constraints do not significantly influence the accuracy. Allowance for the next term in the expansion, which is proportional to the second derivative of the capacitance, yields a correction at the third harmonic which, however, does not exceed 25% of the contribution already taken into account.

If a sinusoidal voltage $U(t) = U_0\sin(\omega t)$ is applied to the sample, where U_0 is the amplitude and ω is the cyclic frequency, formula (1) allowing for (2) has the form

$$I = \frac{U_0}{R}\sin(\omega t) + \omega U_0 C_0\cos(\omega t) + \omega U_0^2\sin(2\omega t)\frac{dC}{dU}. \tag{3}$$

It is easy to see that by using phase-sensitive detection (which essentially performs a Fourier analysis of the signal) at the fundamental frequency, we can measure the first two components, which are proportional to the conductivity and capacitance of the sample. These components can easily be separated by the synchronous detector since they have a relative phase shift of 90°. The third term reflects the nonlinear contribution and is observed at higher harmonics. Here we consider the sit-

uation where the resistive component of the conductivity is much smaller than the capacitive component so that the nonlinear contribution from the conductivity can be neglected. The nonlinear contribution from the resistive component (conduction or dielectric losses), if this exists, is also recorded by the synchronous detector but with a phase shift of 90° and can be taken into account if necessary.

3. THEORETICAL PRINCIPLES OF THE METHOD AND RELATIONSHIP BETWEEN MEASURED QUANTITIES AND COEFFICIENTS OF THE LANDAU–GINZBURG MODEL

In accordance with the phenomenological Landau–Ginzburg model [10], in the vicinity of a phase transition the contribution to the free-energy density of the ferroelectric caused by the polarization P can be expressed in the form

$$F = \frac{1}{2}\alpha P^2 + \frac{1}{4}\beta P^4 + \frac{1}{6}\gamma P^6 - EP, \qquad \alpha = \alpha_0(T - T_0), \tag{4}$$

where α_0, β, and γ are the temperature-dependent Landau coefficients, T_0 is the Curie temperature, and E is the external electric field. Note that for $\beta > 0$, $\gamma \geq 0$, formula (4) describes a second-order phase transition and for $\beta < 0$, $\gamma > 0$ it describes a first-order one.

By minimizing the free energy (4) we can easily find the equation of state in the electric field:

$$E = P[\alpha_0(T - T_0) + \beta P^2 + \gamma P^4]. \tag{5}$$

In order to obtain the nonlinear contribution determined by the last term in (3), we need to find the derivative of the dielectric susceptibility χ with respect to the field which is determined from (5) by differentiating with respect to the polarization:

$$\chi^{-1} = (\alpha + 3\beta P^2 + 5\gamma P^4)\varepsilon_0, \tag{6}$$

where $\varepsilon_0 \simeq 8.85 \times 10^{-12}$ F/m is the dielectric constant of vacuum.

The derivative of interest to us is obtained by repeated differentiation of (6):

$$\frac{d\chi}{dE} = -\chi^3\varepsilon_0^2(6\beta P + 20\gamma P^3). \tag{7}$$

In (7) the polarization depends on the field in accordance with the equation of state (5) although for weak fields (below the coercive field) corresponding to the experiment, the polarization can be conveniently expressed in the approximate form:

$$P \simeq P_s + \varepsilon_0\chi E, \tag{8}$$

where P_s is the spontaneous polarization which is only nonzero in the ferroelectric phase.

Substituting (7) into (3) allowing for (8), we obtain the nonlinear contributions as far as the fifth harmonic:

$$I_\omega = S\omega\frac{U_0}{d}(1 + \chi)\varepsilon_0\cos(\omega t), \tag{9.1}$$

$$I_{2\omega} = -2\Big\{(\varepsilon_0\chi)^3[3\beta P_s + 10\gamma P_s^3] + 15(\varepsilon_0\chi)^5\Big(\frac{U_0}{d}\Big)^2\gamma P_s\Big\}S\omega\Big(\frac{U_0}{d}\Big)^2\sin(2\omega t), \tag{9.2}$$

$$I_{3\omega} = 3\Big\{(\varepsilon_0\chi)^4[\beta + 10\gamma P_s^2] + \frac{5}{2}(\varepsilon_0\chi)^6\Big(\frac{U_0}{d}\Big)^2\gamma\Big\}S\omega\Big(\frac{U_0}{d}\Big)^3\cos(3\omega t), \tag{9.3}$$

$$I_{4\omega} = 15(\varepsilon_0\chi)^5\gamma P_s S\omega\Big(\frac{U_0}{d}\Big)^4\sin(4\omega t), \tag{9.4}$$

$$I_{5\omega} = -\frac{5}{2}(\varepsilon_0\chi)^6\gamma S\omega\Big(\frac{U_0}{d}\Big)^5\cos(5\omega t), \tag{9.5}$$

where S is the overlap area of the sample electrodes and d is the film thickness.

It is easy to see that by measuring these current harmonics, we can determine all the Landau coefficients. In (9.1) we neglect the contribution of the conductivity and also the contributions of the third and fifth harmonics since these are much smaller than the fundamental contribution at fairly low sample voltages. The second terms in (9.2) and (9.3) reflect the intermodulation contributions. These must be treated with caution: they can be neglected if the fourth and fifth harmonics in the signal spectrum are negligible compared with the second and third. This condition can sometimes be satisfied experimentally by selecting a suitable sample voltage. A very useful tool here is the signal spectrum analyzer, which in PhysLab uses a fast Fourier transformation algorithm, which allows us to observe all existing harmonics of the current response and facilitates the correct choice of experimental conditions. We subsequently consider the situation when the intermodulation contributions to the first three harmonics are not significant. In this case, it is convenient to use quantities defined in terms of the ratios of the Fourier components measured by the synchronous detector as follows:

$$A_2 = \frac{-\sqrt{2}\Phi_{2x}}{2\Big(\sqrt{2}\Phi_{1y} - \omega\varepsilon_0\dfrac{U_0 S}{d}\Big)^3}\frac{U_0}{d}(\omega S)^2, \tag{10.1}$$

$$A_3 = \frac{\sqrt{2}\Phi_{3y}}{3\left(\sqrt{2}\Phi_{1y} - \omega\varepsilon_0\frac{U_0 S}{d}\right)^4}\frac{U_0}{d}(\omega S)^3, \quad (10.2)$$

$$A_4 = \frac{\sqrt{2}\Phi_{4x}}{15\left(\sqrt{2}\Phi_{1y} - \omega\varepsilon_0\frac{U_0 S}{d}\right)^5}\frac{U_0}{d}(\omega S)^4, \quad (10.3)$$

$$A_5 = \frac{-2\sqrt{2}\Phi_{5y}}{5\left(\sqrt{2}\Phi_{1y} - \omega\varepsilon_0\frac{U_0 S}{d}\right)^6}\frac{U_0}{d}(\omega S)^5, \quad (10.4)$$

where Φ_{1y}, Φ_{2x}, Φ_{3y}, Φ_{4x}, and Φ_{5y} are the effective values of the x and y Fourier components recorded by the synchronous detector from the first to the fifth harmonic, respectively. In cases when $\chi \gg 1$, which is usually the case near a phase transition, the term $\varepsilon_0\omega U_0 S/d$ in the denominator can be neglected. Note that the x component corresponds to the signal in phase with the reference signal $\propto \sin(k\omega t)$, where k is the harmonic number] and the y-component has a phase shift of 90° relative to the reference signal. Since abrupt changes in the amplitudes and phases of the harmonics may occur near the phase transition, simultaneous recording of the kth and first harmonics is a necessary condition for obtaining correct values of A_k.

In the Landau–Ginzburg model the quantities A_k introduced in accordance with expressions (9.1)–(9.5) are related to the Landau coefficients as follows:

$$A_2 = P_s(3\beta + 10\gamma P_s^2), \quad (11.1)$$

$$A_3 = \beta + 10\gamma P_s^2, \quad (11.2)$$

$$A_4 = \gamma P_s, \quad (11.3)$$

$$A_5 = \gamma. \quad (11.4)$$

It is easy to see that the measured values of A_5 and A_3 in the paraelectric phase ($P_s^2 = 0$) give us the values of the coefficients γ and β. With these values we can easily obtain the temperature dependence of the spontaneous polarization. The value of the coefficient α_0 is determined from measurements at the first harmonic which are equivalent to measurements of the capacitance [11].

The temperature behavior of A_k depends on the order of the phase transition. For a first-order transition we have $\beta < 0$, $\gamma > 0$. Thus, at certain temperatures T_2 and T_3 it is possible to have zero points [$A_2(T_2) = 0$ and $A_3(T_3) = 0$] and even a change in the sign of the signals. This situation is achieved subject to the conditions:

$$3\beta + 10\gamma(P_s(T_2))^2 = 0, \quad (12.1)$$

$$\beta + 10\gamma(P_s(T_3))^2 = 0. \quad (12.2)$$

For these points it is also easy to obtain relationships to determine the Landau coefficients:

$$\beta = -\frac{1}{2}A_3(T_2), \quad (13.1)$$

$$\gamma = \frac{1}{20}A_3(T_2)\left(\frac{A_3(T_2)}{A_2(T_3)}\right)^2 \quad (13.2)$$

and the spontaneous polarization:

$$P_s(T_3) = -\frac{A_2(T_3)}{A_3(T_2)}, \quad (14.1)$$

$$(P_s(T_2))^2 = 3(P_s(T_3))^2. \quad (14.2)$$

In the temperature range (T_2, T_3) it is easy to estimate the pyroelectric coefficient:

$$p \simeq \frac{P_s(T_2) - P_s(T_3)}{T_3 - T_2} \simeq \frac{P_s(T_3)(\sqrt{3} - 1)}{T_3 - T_2} = -\frac{A_2(T_3)}{A_3(T_2)(T_3 - T_2)}(\sqrt{3} - 1). \quad (15)$$

At this point it is convenient to note that the model proposed above refers to a homogeneous polarized ferroelectric. In practice, we need to deal with many-domain samples. The measured values are the result of averaging over the sample volume. If the sample is not polarized, even harmonics may not be observed ($A_2 = 0$) because the contributions from various sections of the sample are mutually compensated so that the average polarization is zero. Odd harmonics are also observed in an unpolarized sample since the average of the square of the polarization is always nonzero in the ferroelectric phase. In order to obtain the most comprehensive information, the sample must be polarized.

In formulas (4)–(7) the dielectric susceptibility χ corresponds to a state of thermodynamic equilibrium. However, if a fairly high-frequency alternating electric field is applied to the sample, the vector of the induced polarization may not have time to follow the direction of the electric field, for example, as a result of the presence of dielectric relaxation processes. If a sinusoidal field is applied to the sample, the frequency dispersion may be taken into account by using the following expression in (7) instead of (8):

$$P(\omega) \simeq P_s + \varepsilon_0 K(\omega)\chi E, \quad (16)$$

where the frequency dependence of the induced polarization is reflected using the complex transfer function

$$K(\omega) = |K(\omega)|e^{i\phi(\omega)}, \quad (17)$$

and $|K(\omega)|$ and $\phi(\omega)$determine the amplitude-frequency and phase characteristics of the dielectric susceptibility. Thus, an additional phase shift appears in relations (9.1)–(9.5). The existence of this phase shift may cause problems in determining the order of the phase transition from the sign of the recorded components at

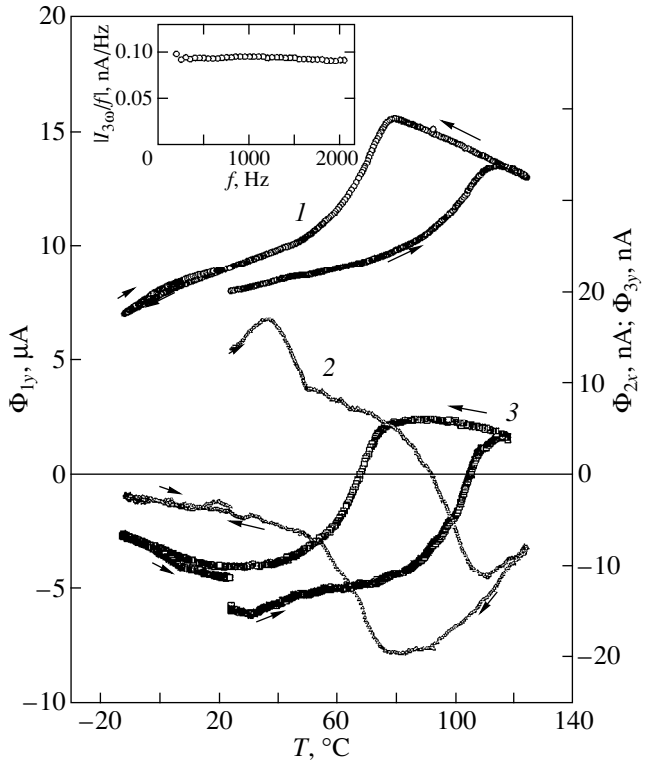


Fig. 2. Temperature dependences of the Fourier components of the current response when a voltage $U = U_0\sin(2\pi ft)$ (f = 1000 Hz, U_0 = 1 V) is applied to the sample. Curve *1* corresponds to Φ_{1y}, *2*—Φ_{2x}, and *3*—Φ_{3y}. The arrows indicate the direction of change in temperature. The inset gives the amplitude-frequency characteristic of the current response at the third harmonic $|I_{3\omega}| = (\Phi_{3x}^2 + \Phi_{3y}^2)^{1/2}$ normalized to the frequency of the applied voltage (U_0 = 2 V, T = 20°C).

harmonics higher than the second. It is easy to establish that even when the intermodulation contributions are small, the phase shift ϕ is present at the third harmonic and increases threefold at the fifth harmonic. Hence, the second harmonic may be a more reliable additional criterion for correctly determining the order of the phase transition using the sign of the corresponding Fourier component. Measurements of the frequency dependence of the corresponding harmonics may also be very useful for determining the optimum operating frequency but when the temperature varies, both the amplitudes of the Fourier components and their phases must be recorded simultaneously for subsequent use to obtain a correct interpretation of the measured values of A_k.

4. EXPERIMENTAL RESULTS AND DISCUSSION

The samples for the investigations were Langmuir–Blodgett films of 70/30 VDF (TFE) copolymer consisting of 20 monomolecular layers. The monomolecular layers were transferred from the surface of water onto glass substrates with a deposited aluminum electrode at room temperature (20–22°C) and 3 mN/m surface pressure. At room temperature this pressure corresponds to a close-packed monolayer 5 Å thick when the probability of local collapse (the formation of a bimolecular layer 10 Å thick at various points on the water surface) is low. In the present study, instead of the horizontal lift method used in [6] when the substrate plane is parallel to the water surface during transfer of the monolayer, we used the classical Langmuir–Blodgett method and the apparatus described in [12]. However, transfer was made by pulling the substrate from the water (*Z*-type transfer) so that the normal to the substrate surface was oriented approximately at an angle of 45° to the water surface, which helped to improve the homogeneity of the films. A second aluminum electrode was deposited on the polymer film from above. The total overlap area of the electrodes was 0.01 ± 0.0005 cm^2. The thickness of the polymer films calculated using the number of transfers and the thickness of a single monomolecular layer was 10 nm.

The measurements were made after polarizing the film with a single 15 V pulse of 200 μs duration. Both positive and negative pulses were used for polarization and this was reflected in the sign of the recorded values of Φ_{2x}.

Figure 2 gives temperature dependences of the first, second, and third harmonics of the current response Φ_{1y}, Φ_{2x}, and Φ_{3y} measured under conditions when the amplitudes of the fourth and fifth harmonics are negligible (corresponding to the noise level) so that the intermodulation contributions in (9.2) and (9.3) can be neglected for the analysis. The first harmonic (curve *1*) gives the well-known temperature hysteresis of the dielectric constant which can be attributed to the first order of the phase transition [11]. The first order is also confirmed by the change in the sign of the third (curve *3*) and second (curve *2*) harmonics. The sign of the second-harmonic Fourier component corresponds to the first order of the phase transition. In fact, curve *2* was obtained after polarizing the film with a negative pulse (without changing the sample switching geometry) which corresponds to negative P_s. In this case, according to (9.2) in the temperature range $T > T_2$ [see also (12.1)] the sign of the second-harmonic signal should be determined by the sign of the coefficient β. Thus, the negative second-harmonic signal at temperatures above 90°C reflects the first order of the phase transition. The temperature behavior of the phases of the second- and third-harmonic signals is an important factor, see Fig. 3. It is easy to see that the second-harmonic signal corresponds to the *x*-component. In fact, the phase shift over the entire temperature range (except for the point where the sign changes and the signal phase is not determined) is either zero or ±180°. The third harmonic signal has an *x*-component as well as a dominant *y* component. This contribution is particularly significant at temperatures below 40°C where, according to various

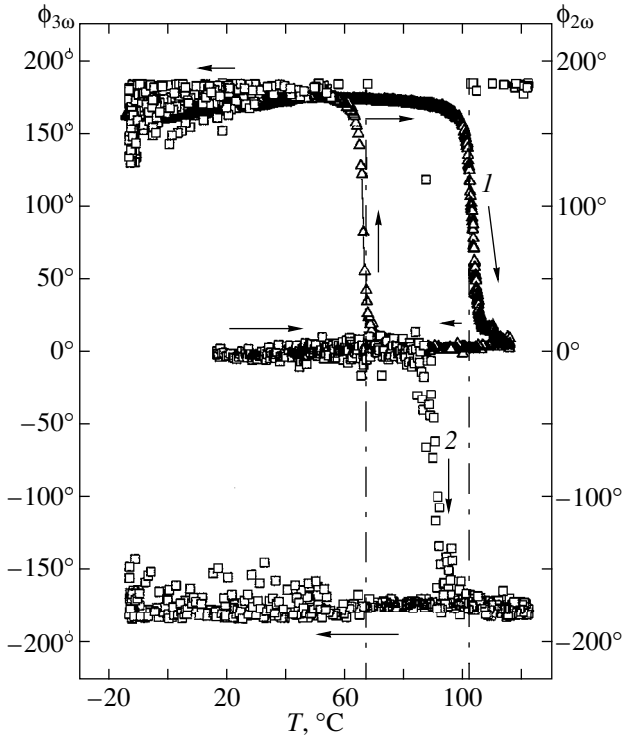


Fig. 3. Temperature dependences of the phase shifts of the current response relative to a reference signal when a voltage $U = U_0\sin(2\pi ft)$ (f = 1000 Hz, U_0 = 1 V) is applied to the sample; curve *1* corresponds to the third harmonic (reference signal $\propto\cos(6\pi ft)$); *2*—second harmonic (reference signal $\propto\sin(4\pi ft)$). The arrows indicate the direction of change in temperature.

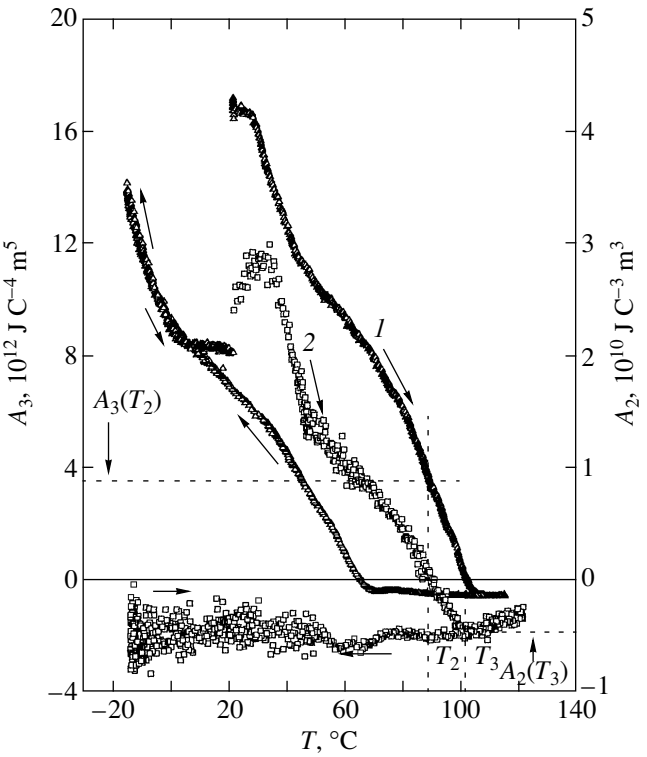


Fig. 4. Temperature dependences of A_3 (curve *1*) and A_2 (curve *2*). The dependence of A_3 allows for a sign inversion caused by a phase shift (see text).

data [13–15], a "surface" phase transition takes place. In the temperature range between +40 and +115°C the phase shift corresponds to a dominant y-component. The phase behavior of the third-harmonic signal very clearly shows a range of temperature hysteresis corresponding to a region where the ferroelectric and paraelectric phases may coexist although the sign of the signal, as in [3], does not correspond to the first order of the phase transition. Bearing in mind that the dominant contribution at the third harmonic corresponds to the y-component and the second harmonic correctly reflects the order of the phase transition, the "sign" contradiction may be attributed to a $\phi \simeq 180°$ phase shift [see formulas (16) and (17) and their discussion above]. According to the model, this phase shift should also have an influence at the fifth harmonic in the form of a sign inversion since, being trebled, it remains equivalent to 180°. Thus, we can see that the sign of Φ_{2x} is a more reliable criterion for directly determining the sign of β than the sign of Φ_{3y}.

The experimental values of A_2 and A_3 determined above using formulas (10.1) and (10.2) respectively and plotted in Fig. 4 are more informative. We stress that the dependence of A_3 allows for a sign inversion caused by the phase shift $\phi \simeq 180°$ at the third harmonic [the amplitude-frequency dependence of the third-harmonic signal (see inset to Fig. 2) is weak and was neglected.] It is easy to see that the dependences of A_2 and A_3 can have a very clear interpretation in terms of the Landau–Ginzburg model and the relationships (11.1) and (11.2) obtained by us for A_2 and A_3, respectively. The zero conditions (12.1) and (12.2) are achieved at temperatures $T_2 \simeq 89°$C and $T_3 \simeq 102°$C, respectively. In this case we have $A_2(T_3) \simeq -5 \times 10^9$ J C^{-3} m^3 and $A_3(T_2) \simeq 3.9 \times 10^{12}$ J C^{-4} m^5. In accordance with (13.1), we have $\beta \simeq -2 \times 10^{12}$ J C^{-4} m^5 which is close to the value $\beta \simeq -1.5 \times 10^{12}$ J C^{-4} m^5 obtained in [16]. However it is important to note at this point that the point T_2 belongs to the region of temperature hysteresis where the ferroelectric and paraelectric phases may coexist. The existence of a certain fraction of the paraelectric phase reduces the measured value of A_3 which in turn corresponds to reduced absolute values of β. For example, in [11] the value of β calculated using the characteristic temperature points of the transition and the determined Curie–Weiss constant was -3.9×10^{12} J C^{-4} m^5. At this point attention should be drawn to another contradiction. According to the model, in the paraelectric phase where $P_s = 0$, A_3 should be independent of temperature. In fact, during heating above the transition temperature $T \simeq 110°$C, this dependence of A_3 almost disappears

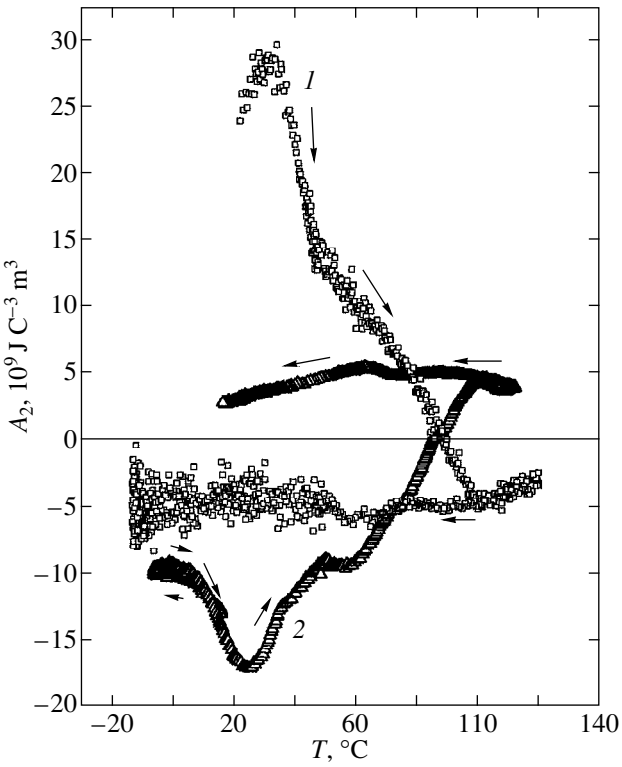


Fig. 5. Temperature dependences of A_2 obtained after polarizing the sample using pulses of different polarity. Curves *1* and *2* were obtained after polarization using single negative and positive pulses, respectively. The amplitude of the single pulses is 15 V and the duration 200 μs.

(curve *1*) becomes parallel to the abscissa). This property is also conserved during cooling to $T_0 \simeq 70$°C which indicates that a small quantity of incipient polar phase is formed. However, the value of A_3 which in this case, according to (11.2), should correspond to the coefficient β is -6×10^{12} J C^{-4} m^5 which is considerably lower than the value obtained above in terms of absolute value. All this could be explained by the fact that even above the transition temperature, 110°C, "remanent" polarization is found and this depends weakly on temperature. In the following analysis of the measured temperature dependence of A_2 we can see that even above the transition temperature $T \simeq 110$°C nonzero polarization exists.

In accordance with (15), an estimate of the pyroelectric coefficient in the temperature range (T_2, T_3) gives 0.7×10^{-4} C m^{-2} K^{-1} which is approximately 1.5 times higher than the value obtained in this temperature range by direct pyroelectric measurements [17]. Everything again looks as if the value of $A_3(T_2)$ was slightly too low.

The temperature behavior of A_2 (curve *2* in Fig. 4) is extremely interesting. According to the model, above the upper transition temperature which, as can be seen from the first-harmonic measurements, is 110°C the value of A_2 should abruptly go to zero, which would correspond to zero spontaneous and therefore average (remanent) polarization. The experiment gives a nonzero value of A_2 above the transition temperature and indicates that not only is the spontaneous polarization nonzero but the average polarization (in our geometry this corresponds to the component along the normal to the plane of the film) also does not go to zero above the upper transition temperature. The fact that we are dealing with the polarization of a film and not some parasitic effect from the surface which produces a signal at even harmonics is demonstrated by the sign inversion of the temperature dependence of A_2 after the film has been polarized by a pulse of opposite polarity, see Fig. 5. Thus, we can see that the observed phase transition is fairly complex. It follows from these data that only part of the film volume is responsible for the transition in the range $T \simeq 110$°C. The remainder of the film (to be specific we call this the *X*-state) undergoes a phase transition at a higher temperature. The existence of the *X*-state obviously distorts the picture determined by the homogeneous Landau–Ginzburg model and introduces additional error in the determination of the coefficients. If the fraction of the film volume belonging to the *X*-state were larger for some reason, it would be possible for us not to observe a change in sign of A_2 and A_3. In this case, the influence of the *X*-state makes it difficult to estimate the coefficient γ using formula (13.2) since the values of $A_3(T_2)$ and $A_2(T_3)$ are raised to some power, and even an acceptable error in determining $A_3(T_2)$ and $A_2(T_3)$ separately has an unfavorable influence on the resultant error in the calculation of γ. According to (11.4) an alternative method of determining the coefficient γ involves recording the Fourier *y*-component at the fifth harmonic. Figure 6 shows corresponding curves for the Fourier *y* and *x* components at the fifth harmonic. It can be seen that in the low-temperature range (up to +50°C) corresponding to a surface phase transition, the active losses make a very strong contribution to the *x*-component (curve *2* in Fig. 6). Consequently this temperature range cannot be used to determine the coefficient γ and in addition it is some distance from the phase transition of interest to us. Only at temperatures above +100°C is the *x*-component fairly small. Figure 7 shows the temperature dependence of A_5 (as in the case of the third harmonic, allowance is made for a sign inversion caused by a 180° phase shift). Above +90°C the value of A_5 does not depend on temperature to within measurement error and it can be assumed that in this temperature range the contribution of the surface phase transition is negligible and in accordance with (11.4) we have $A_5 = \gamma \simeq (5 \pm 3) \times 10^{14}$ J C^{-6} m^9. This last value agrees within measurement error with the values of $\gamma \simeq 8 \times 10^{14}$ J C^{-6} m^9 and $\gamma \simeq 2 \times 10^{14}$ J C^{-6} m^9 obtained for Langmuir–Blodgett

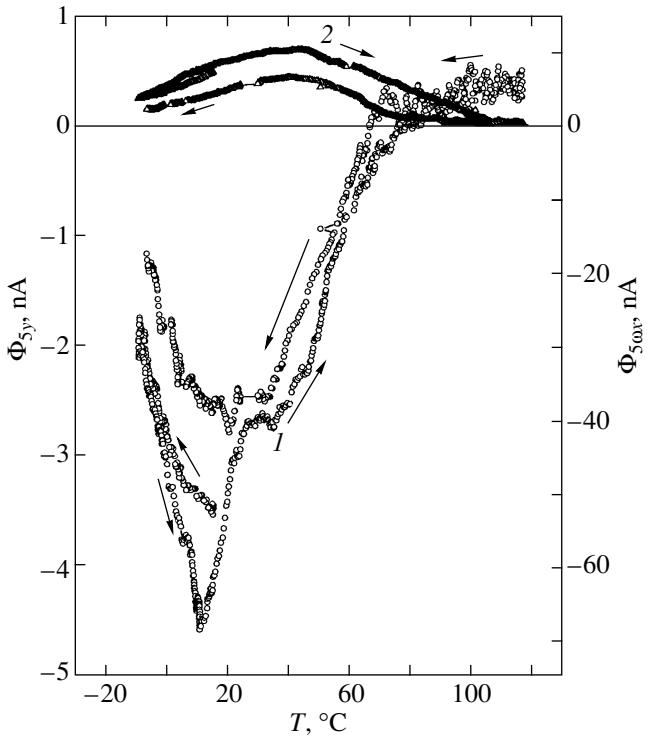


Fig. 6. Temperature dependences of the Fourier *y* and *x* components of the fifth-harmonic current response (curves *1* and *2*, respectively) when the voltage $U = U_0\sin(2\pi ft)$ ($f = 1000$ Hz, $U_0 = 2$ V) is applied to the sample. The arrows indicate the direction of change in temperature.

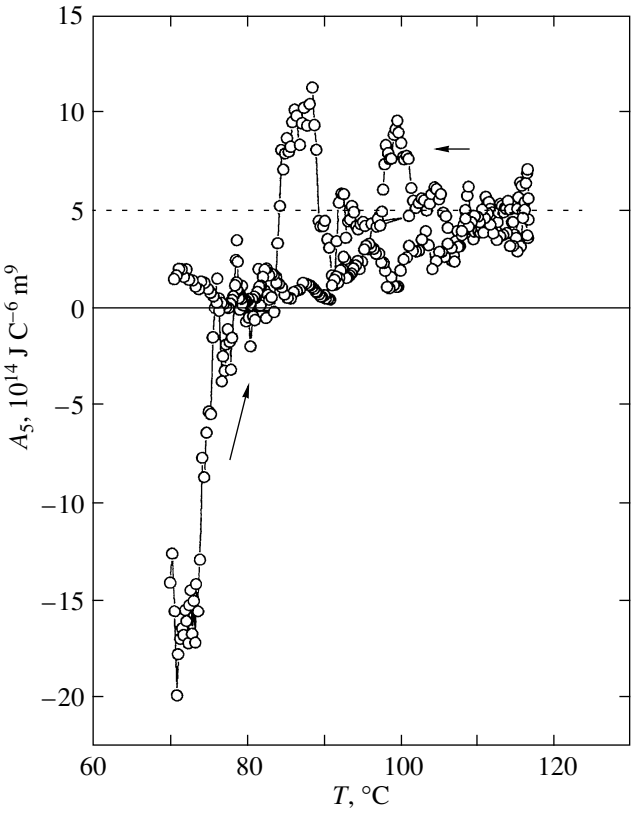


Fig. 7. Temperature dependences of A_5. The dependence of A_5 allows for a sign inversion caused by a phase shift (see text).

and thick films of a similar polymer, respectively, in [11] and [16].

The coefficient α_0 (or the Curie–Weiss constant $C' = 1/\varepsilon_0\alpha_0$) could be determined using the temperature dependence of the linear dielectric susceptibility which corresponds to the first-harmonic data (Fig. 2). However, problems arise first, because of the possible influence of the amorphous (nonferroelectric) phase and the oxide film at the electrodes and second, an additional assumption must be made on the "slow" formation of a new phase in the temperature hysteresis region. Details of an approach to solve these problems may be found in [11]. Another method of determining α_0 based on nonlinear measurements is also possible. Solving the equation of state (5) in a zero field ($E = 0$) gives a region of temperature hysteresis determined by

$$\Delta T = \beta^2/4\alpha_0\gamma. \qquad (18)$$

This region of temperature hysteresis is clearly expressed in the temperature dependence of the phase at the third harmonic (Fig. 3, curve *1*) giving $\Delta T \simeq 35$ K. Using the values of β and γ, from (18) we find $\alpha_0 = (0.9 \pm 0.5) \times 10^8$ J m C^{-2} K^{-9} which corresponds to the Curie–Weiss constant $C' = (2 \pm 1) \times 10^3$ K.

ACKNOWLEDGMENTS

The authors thank V.V. Lazarev and M.I. Barnik for many useful recommendations. This work was supported financially by the Russian Foundation for Basic Research (projects nos. 99-02-16484, 98-02-17071) and the Copernicus Program (Copernicus IC15-CT96-0744).

REFERENCES

1. G. A. Smolenskiĭ, V. A. Bokov, V. A. Isupov, N. N. Kraĭnik, R. E. Pasynkov, A. I. Sokolov, and N. K. Yushin, *The Physics of Ferroelectric Phenomena* (Nauka, Leningrad, 1985).
2. M. E. Lines and A. M. Glass, *Principles and Applications of Ferroelectrics and Related Materials* (Clarendon Press, Oxford, 1977).
3. T. Furukawa, Phase Transit. **18**, 143 (1989).
4. T. Furukawa, H. Kodama, O. Uchinokura, and Y. Takahashi, Ferroelectrics **171**, 33 (1995).
5. H. Kodama, Y. Takahashi, and T. Furukawa, Jpn. J. Appl. Phys. **38**, 3589 (1999).
6. S. Palto, L. Blinov, A. Bune, *et al.*, Ferroelectr. Lett. Sect. **19**, 65 (1995).

7. S. Palto, L. Blinov, E. Dubovik, V. Fridkin, *et al.*, Europhys. Lett. **34**, 465 (1996).
8. A. Bune, V. Fridkin, S. Ducharme, L. Blinov, *et al.*, Nature **391**, 874 (1998).
9. S. P. Palto, *Effects of Molecular Field in Langmuir–Blodgett Films: Optics and Stark Spectroscopy*, Doctoral Dissertation in Mathematical Physics (Moscow, 1998).
10. V. L. Ginzburg, Zh. Éksp. Teor. Fiz. **19**, 36 (1949).
11. S. P. Palto, A. M. Lotonov, K. A. Verkhovskaya, *et al.*, Zh. Éksp. Teor. Fiz. **117** (2000) (in press) [JETP **90** (2000) (in press)].
12. S. G. Yudin, S. P. Palto, V. A. Khavrichev, *et al.*, Thin Solid Films **210-211**, 46 (1992).
13. J. Choi, P. Dowben, S. Pebley, *et al.*, Phys. Rev. Lett. **80**, 1328 (1998).
14. J. Choi, P. A. Dowben, A. V. Bune, *et al.*, Phys. Lett. A **249**, 505 (1999).
15. J. Choi, P. A. Dowben, C. N. Borca, *et al.*, Phys. Rev. B **59**, 1819 (1999).
16. T. Furukawa, Ferroelectrics **57**, 63 (1984).
17. A. Bune, C. Zhu, S. Ducharme, *et al.*, J. Appl. Phys. **85**, 7869 (1999).

Translation was provided by AIP

Journal of Experimental and Theoretical Physics, Vol. 90, No. 5, 2000, pp. 881–885.
From Zhurnal Éksperimental'noĭ i Teoreticheskoĭ Fiziki, Vol. 117, No. 5, 2000, pp. 1013–1017.

SOLIDS
Electronic Properties

The Conductivity of the Two Component System of the Chessboard Type¶

Yu. N. Ovchinnikov*, ** and A. M. Dyugaev***

**Centre de Recherches sur les Tres Basses Temperatures, associe a UIF, CNRS, BP 166, P-38042 Grenoble Cedex 09, France*
***L. D. Landau Institute for Theoretical Physics, Russian Academy of Sciences, Chernogolovka, Moscow oblast, 142432 Russia*
e-mail: ovc@itp.ac.ru
****Grenoble High Magnetic Field Laboratory, MPI-FKF and CNRS, BP 166, F-38042 Grenoble Cedex 09, France*

Received January 12, 2000

Abstract—The effective conductivity of the sample with a chessboard structure is found. The corner points are singular for the electric field and charge density. The effective conductivity, charge density and electric potential are expressed in terms of a Weierstrass elliptic function. The results are valid for films and for cylindrical samples with the same cross-section structure. We confirm the conjecture of Dykhne about the value of the effective conductivity for the chessboard structure. © *2000 MAIK "Nauka/Interperiodica".*

The conductivity of two or more component systems is an object of intensive study [1]. Probably the first strong result of the effective conductivity of a two component system with conductivities $\sigma_{1,2}$ and for random symmetric and isotropic distribution of both components was obtained by Dykhne [1]. He found a very simple expression for the effective conductivity σ_{eff}, that is valid for the whole region of ratio σ_1/σ_2

$$\sigma_{\text{eff}} = (\sigma_1\sigma_2)^{1/2}. \tag{1}$$

Expression (1) looks so simple, that hope arises that it has an even larger region of application, that was supported by derivation. The conjecture that equation (1) can be valid also for ordered systems was made in the same paper of Dykhne [1]. If it is true, then the simple symmetry arguments can help to restore the effective conductivity of a more complicated system [6].

To prove the conjecture of Dykhne, we solved the problem of conductivity of the chessboard for arbitrary values of the conductivities $\sigma_{1,2}$ of components. The result is that the effective conductivity is given by the expression

$$\sigma_{\text{eff}} = (\sigma_1\sigma_2)^{1/2}\tilde{f}(Z^2), \quad Z = \frac{\sigma_1 - \sigma_2}{\sigma_1 + \sigma_2}, \tag{2}$$

and the function $\tilde{f}(Z^2)$ is given by the ratio of two single integrals. We prove that the ratio of these two integrals is such that $\tilde{f} = 1$.

The in-plane conductivity tensor

$$\hat{\sigma}_{\text{eff}} = \delta_{\alpha\beta}\sigma_{\text{eff}} \tag{3}$$

where $\delta_{\alpha\beta}$ is the Kronecker symbol. For this reason, we shall consider the orientation of external electric field along the diagonals of cells, as given in Fig. 1.

The chessboard possess translational symmetry. Elementary cell vectors are

$$\sqrt{2}a(1, 0), \quad \sqrt{2}a(0, 1), \tag{4}$$

where a is the size of a cell.

Translational symmetry and reflectional symmetry over the y-axis reduce the number of independent charge distribution functions to two: $\rho_{1,2}$. We will show below, that

$$\rho_1(t) = -\rho_2(t) \tag{5}$$

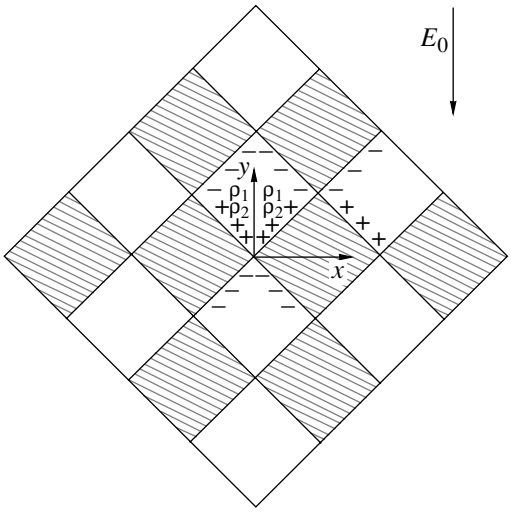


Fig. 1. The chessboard structure of the sample.

¶This article was submitted by the authors in English.

where t is the distance along the edge of a cell counted from the corner of the cell.

The scalar potential φ on the chessboard can be taken in the form

$$\varphi = E_0 y - 4\pi \int d^2 r_1 G(r - r_1)\rho(r_1), \tag{6}$$

where G is the two-dimensional Green's function

$$G(r - r_1) = \frac{1}{2\pi}\ln|r - r_1|. \tag{7}$$

On the edge of a cell we have also

$$\begin{aligned} E_n^{(1)} - E_n^{(2)} &= 4\pi\rho, \\ \sigma_1 E_n^{(1)} &= \sigma_2 E_n^{(2)} \end{aligned} \tag{8}$$

where n is normal vector to the edge of the cell, E is the electric field, and ρ is the charge density per unit of length. The equation system (6), (8) gives an integral equation for the charge distribution on the edges of cells. Simple proof shows that

$$\rho_1(t) = -\rho_2(t),$$

and hence, we obtain only one integral equation for the quantity ρ_1;

$$\begin{aligned} &\frac{Z}{2\pi}\left\{\frac{1}{\sqrt{2}}\right. \\ &\left. + 4\sum_{k,\,l=-\infty}^{\infty}\int_0^1 d\tilde{t}\frac{\tilde{t}+k+l}{(\tilde{t}+k+l)^2+(\tilde{t}'+k-l)^2}\tilde{\rho}(\tilde{t})\right\} = \tilde{\rho}(\tilde{t}'). \end{aligned} \tag{9}$$

In equation (9), we have used the new variables

$$t = a\tilde{t}, \quad \rho_1 = E_0\tilde{\rho}(\tilde{t}). \tag{10}$$

The effective conductivity σ_{eff} is connected with a charge density $\tilde{\rho}$ by a simple expression:

$$\sigma_{\text{eff}} = \frac{4\pi\sqrt{2}\sigma_1\sigma_2}{\sigma_1 - \sigma_2}\int_0^1 d\tilde{t}\tilde{\rho}(\tilde{t}). \tag{11}$$

The change $\tilde{t} \longrightarrow 1 - \tilde{t}$ shows that σ_{eff} is a symmetric function of variables $\sigma_{1,2}$ and has a form given by equation (2). For this reason, we will suppose below that $\sigma_1 > \sigma_2$.

One sum over l in equation (9) can be taken in explicit form. As a result, we have

$$\begin{aligned} &\frac{Z}{2\pi}\left\{\frac{1}{\sqrt{2}} + 2\pi\sum_{k=-\infty}^{\infty}\int_0^1 dt\right. \\ &\left.\times\frac{\sin(\pi(t+k))}{\cosh(\pi(t'+k)) - \cos(\pi(t+k))}\tilde{\rho}(t)\right\} = \tilde{\rho}(t'). \end{aligned} \tag{12}$$

From equation (12), we obtain that $\tilde{\rho}(t)$ is an analytical double periodic function of t with periods $\{1 \pm i\}$. In each cell, it has two branch points

$$t_{N,M} = \{N + M;\ i(N - M)\}.$$

$$t^1_{N,M} = \{1 + N + M;\ i(N - M)\}.$$

If $\sigma_1 > \sigma_2$, we have near the branch points

$$\begin{aligned} \tilde{\rho}(t) &\sim \left[\frac{1}{(t - t_{N,M})^2}\right]^{\kappa}, \\ \tilde{\rho}(t) &\sim [(t - t^1{}_{N,M})^2]^{\kappa}. \end{aligned} \tag{13}$$

Parameter κ is the solution of equation

$$\sin(\pi\kappa) = Z. \tag{14}$$

With the help of equations (13), (14) we obtain that near the branch points the following expansion for $\tilde{\rho}(t)$ is valid:

$$\begin{aligned} \tilde{\rho}(t_1) &= \frac{1}{t_1^{2\kappa}}\left[\sum_{n=0}^{\infty} C_n t_1^{4n}\right], \quad t_1 = t - 2k - 2il, \\ \tilde{\rho}(t_1) &= t_1^{2k}\left[\sum_{n=0}^{\infty}\tilde{C}_n t_1^{4n}\right], \quad t_1 = t - 2k - 1 - 2il. \end{aligned} \tag{15}$$

Equations (14), (5) mean that function $\tilde{\rho}(t)$ can be presented in the form

$$\tilde{\rho}(t) = B\Psi^{\kappa}(t) \tag{16}$$

where $\Psi(t)$ is the elliptic function, and B is some constant. Inside the unit cell, this function has only one pole. The order of this pole cannot be higher than two. Since the elliptic function of the first order does not exist, $\Psi(t)$ is the Weierstrass elliptic function $\wp$ [7]

$$\tilde{\rho}(t) = B\wp^{\kappa}. \tag{17}$$

From equation (15), it is trivial to prove that the solution (17) is unique. The value of the constant B will be found below. First, we give the parameters of the elliptic function $\wp(t)$

$$g_3 = 0, \quad g_2 = -4K^4(1/\sqrt{2}) = -47.26817928, \tag{18}$$

where K is a complete elliptic integral of the first kind. Note, that the existence of the corner points on the chessboard leads to the singular distribution of a electric charge, given by equation (17), because for small values of the t Weierstrass function have a pole of the second order:

$$\wp(t) = \frac{1}{t^2} + \frac{g_2}{20}t^2 + \ldots, \quad |t| \ll 1. \tag{19}$$

Near the point $t = 1$ the function $\wp(t)$ is given by equation

$$\wp(t) = -\frac{g_2}{4}(1-t)^2 + \ldots, \quad |1-t| \ll 1. \tag{20}$$

The complete Loran expansion of the function $\wp(t)$ in the circle $|t| < \sqrt{2}$ is [8]

$$\wp(t) = \frac{1}{t^2} + \frac{g_2}{20}t^2 + \frac{g_3}{28}t^4 + \sum_{k=4}^{\infty} C_k t^{2k-2}, \tag{21}$$

where for $k \geq 4$

$$C_k = \frac{3}{(2k+1)(k-3)}\sum_{m=2}^{k-2} C_m C_{k-m}.$$

In our case, the constant $g_3 = 0$, and thus, all odd coefficients C_{2k+1} are equal to zero. As the result, the Loran expansion of the function $\wp(t)$ is given by the equation

$$\wp(t) = \frac{1}{t^2} + \frac{g_2}{20}t^2 + \sum_{k=2}^{\infty} \tilde{C}_k t^{4k-2}, \tag{22}$$

$$\tilde{C}_k = \frac{3}{(4k+1)(2k-3)}\sum_{m=1}^{k-1} \tilde{C}_m \tilde{C}_{k-m}, \quad k \geq 2, \tag{23}$$

$$\tilde{C}_1 = \frac{g_2}{20} = -\frac{1}{5}K^4\left(\frac{1}{\sqrt{2}}\right) = -2.36340896.$$

The representation of Weierstrass function $\wp(t)$ with the help of Jacobi elliptic function cn is even more convenient for use than equation (23)

$$\wp(t) = K^2\frac{1 + \mathrm{cn}(2Kt, 1/\sqrt{2})}{1 - \mathrm{cn}(2Kt, 1/\sqrt{2})} \tag{24}$$

where $K = K(1/\sqrt{2}) = 1.854074677$ and

$$\mathrm{cn}\left(2Kt, \frac{1}{\sqrt{2}}\right) = \frac{2\sqrt{2}\pi}{K} \times \sum_{n=1}^{\infty} \frac{e^{-\pi(n-1/2)}}{1 + e^{-\pi(2n-1)}}\cos(\pi t(2n-1)) \tag{25}$$

is the Jacobi elliptic function.

For calculation of the quantity B, it is convenient to introduce two functions $G_{1,2}$. By the definition, we put

$$G_1 = \int_0^1 dt\left(\frac{\wp(t)}{K^2}\right)^{\kappa},$$

$$G_2 = \int_0^1 dt' \int_0^1 dt \times \sum_{n=-\infty}^{\infty} \frac{\sin(\pi(t+n))}{\cosh(\pi(t'+n)) - \cos(\pi(t+n))}\left(\frac{\wp(t)}{K^2}\right)^{\kappa}. \tag{26}$$

From equations (26), we obtain symmetry properties of functions $G_{1,2}$

$$G_1(\kappa) = G_1(-\kappa), \quad G_2(\kappa) = -G_2(-\kappa). \tag{27}$$

The expression (26) for quantity G_2 can be essentially simplified. We have

$$G_2 = \int_0^{\infty} dt' \int_0^1 dt \frac{\sin(2\pi t)}{\cosh^2(\pi t') - \cos(\pi t)}\left(\frac{\wp(t)}{K^2}\right)^{\kappa} = \int_0^1 dt(1-2t)\left(\frac{\wp(t)}{K^2}\right)^{\kappa}.$$

Integration of the equation (12) over t' leads to the following expression for B:

$$B = \frac{Z}{2\pi\sqrt{2}K^{2\kappa}}\frac{1}{G_1 - zG_2}. \tag{28}$$

Inserting this expression for the coefficient B into (11), we obtain the final expression for the effective conductivity

$$\sigma_{eff} = \frac{2\sigma_1\sigma_2}{\sigma_1 + \sigma_2}\frac{G_1}{G_1 - zG_2}. \tag{29}$$

The point $\sigma_1/\sigma_2 \longrightarrow 0$ is bifurcation point of equation (12). In the vicinity of this point, we have

$$\kappa = \frac{1}{2} - \frac{2}{\pi}\sqrt{\frac{\sigma_2}{\sigma_1}}. \tag{30}$$

This point is singular for the quantity G_1. In the main approximation, we obtain from equations (23), (26)

$$G_1 = \frac{\pi}{4K}\sqrt{\frac{\sigma_1}{\sigma_2}}. \tag{31}$$

In this point, the quantity $G_1 - zG_2$ is finite and equal to

$$G_1 - zG_2 = 2\int_0^1 dtt\frac{\mathrm{cn}(kt, 1/\sqrt{2})}{\mathrm{sn}(kt, 1/\sqrt{2})\mathrm{dn}(kt, 1/\sqrt{2})}$$

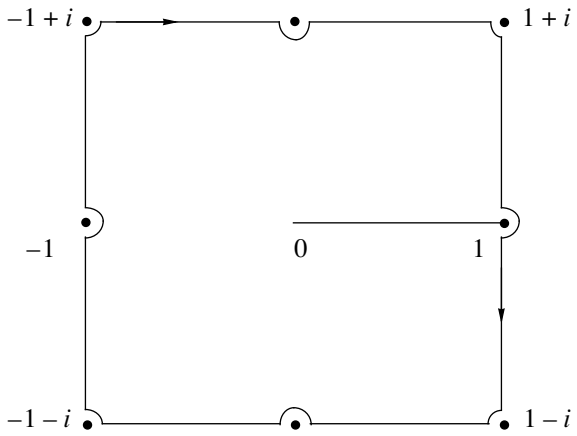


Fig. 2. The contour of integration for the quantity I.

$$= \frac{\pi}{k}\int_0^1 dtt\left[\cot\left(\frac{\pi t}{2}\right) - 4\left(\sum_{k=1}^{\infty}\frac{(-1)^n e^{-\pi n}}{1+(-1)^n e^{-\pi n}}\sin(\pi n t)\right)\right] \quad (32)$$

$$= \frac{1}{k}\left[2\ln 2 + 4\sum_{n=1}^{\infty}\frac{e^{-\pi n}}{n(1+(-1)^n e^{-\pi n})}\right],$$

sn, dn are also Jacobi elliptic function.

Inserting expressions (31), (32) into equation (29), we obtain the value of effective conductivity near the bifurcation point,

$$\sigma_{\text{eff}} = (\sigma_1\sigma_2)^{1/2}\frac{\pi}{2} \times\left\{2\ln 2 + 4\sum_{n=1}^{\infty}\frac{e^{-\pi n}}{n(1+(-1)^n e^{-\pi n})}\right\}^{-1}. \quad (33)$$

The expression in figured brackets is equal to $\pi/2$, and thus, we obtain near the bifurcation point

$$\sigma_{\text{eff}} = (\sigma_1\sigma_2)^{1/2}. \quad (34)$$

Now we will find the value of σ_{eff} for all values of the parameter z. First, we will find the value of quantity G_1. To do this, we use the substitution

$$\frac{\wp(t)}{K^2} = \sqrt{y} \quad (35)$$

and relation

$$\frac{\partial\wp(t)}{\partial t} = -2\sqrt{\wp(t)}\sqrt{\wp^2(t)+K^4} \quad (36)$$

with the help of equations (35), (36) we obtain

$$G_1 = \int_0^1 dt\left(\frac{\wp(t)}{K^2}\right)^{\kappa} = \frac{1}{4\sqrt{\pi}K}\Gamma\left(\frac{1}{4}-\frac{\kappa}{2}\right)\Gamma\left(\frac{1}{4}+\frac{\kappa}{2}\right), \quad (37)$$

where $\Gamma(x)$ is the Euler I-function. Now we calculate the integral I, defined as

$$I = \int_0^1 dtt\left(\frac{\wp(t)}{K^2}\right)^{\kappa} = \frac{1}{1-e^{-4\pi i\kappa}}\int_{\substack{0\\ \Longrightarrow\\ 1}} dtt\left(\frac{\wp(t)}{K^2}\right)^{\kappa}. \quad (38)$$

The integral over the contour between two branch points (0, 1) can be transformed into the integral over the contour, given in the Fig. 2.

Simple calculation leads to the answer

$$2I = G_1\left\{1 + \frac{1+\cos(2\pi\kappa)}{\sin(2\pi\kappa)} - \frac{2\cos(\pi\kappa)}{\sin(2\pi\kappa)}\right\} = G_1\left[1 - \tan\left(\frac{\pi\kappa}{2}\right)\right]. \quad (39)$$

Hence, we obtain

$$\frac{G_2}{G_1} = \tan\left(\frac{\pi\kappa}{2}\right). \quad (40)$$

Inserting this expression for the quantity G_2/G_1 into equation (29) we obtain

$$\sigma_{\text{eff}} = (\sigma_1\sigma_2)^{1/2} \quad (41)$$

for all values of the parameter z. And thus, we confirm the Dykhne conjecture about the value of σ_{eff} for chessboard structure.

Now we do not know whether the simple expression (11) for the effective conductivity is also valid for other regular two component systems or it is the property only of the chessboard structure.

Note, that the problem of the calculation of σ_{eff} was reduced to the calculation of two single integrals, given by equation (26). We have obtained the explicit expression for conductivity as well as for field and current distribution in samples with a chessboard structure.

It looks very plausible that effective conductivity of more complicated many component periodic structures can also be expressed in terms of elliptic functions, as it is given by equations (17), (26), and (29) for the chessboard sample.

ACKNOWLEDGMENTS

Yu.N. Ovchinnikov thanks Prof. P. Monceau for his hospitality in CRTBT (Grenoble). A.M. Dyugaev thanks Prof. P. Wyder for his hospitality in HMFL (Grenoble).

REFERENCES

1. A. M. Dykhne, Zh. Éksp. Teor. Fiz. **59**, 110 (1970) [Sov. Phys. JETP **32**, 63 (1971)].
2. A. L. Efros and B. I. Shklovskii, Phys. Status Solidi B **76**, 475 (1976).
3. D. I. Bergman and D. Stroud, Phys. Rev. B **32**, 6097 (1985).
4. A. M. Dykhne and I. M. Kaganova, Phys. Rep. **288**, 263 (1997).
5. Yu. A. Drezin and A. M. Dykhne, Zh. Éksp. Teor. Fiz. **84**, 1756 (1983) [Sov. Phys. JETP **57**, 1024 (1983)].
6. I. M. Khalatnikov and A. Kamenshik, submitted to Zh. Éksp. Teor. Fiz. [submitted to JETP].
7. M. A. Lavrent'ev and B. V. Shabat, *Methods of Theory of Function of Complex Variables* (Fizmatgiz, Moscow, 1958).
8. *Handbook of Mathematical Functions*, Ed. by M. Abramowitz and I. A. Stegun (Dover, New York, 1971; Nauka, Moscow, 1979).
9. I. S. Gradshteyn and I. M. Ryzhik, *Table of Integrals, Series, and Products* (Nauka, Moscow, 1971; Academic, New York, 1980).

Journal of Experimental and Theoretical Physics, Vol. 90, No. 5, 2000, pp. 886–896.
Translated from Zhurnal Éksperimental'noĭ i Teoreticheskoĭ Fiziki, Vol. 117, No. 5, 2000, pp. 1018–1029.

SOLIDS
Electronic Properties

Ferromagnetism of Compounds with Hexagonal Symmetry

R. O. Zaĭtsev* and N. V. Terekhina**

**Russian Research Centre Kurchatov Institute, pl. Kurchatova 1, Moscow, 123182 Russia*
e-mail: zaitsev@mbslab.kiae.ru
***Russian National Research Institute Teplopribor, Moscow, 129085 Russia*
e-mail: tnv_tnv@yahoo.com
Received January 10, 2000

Abstract—The concept of strong interaction in the same unit cell is used to establish the possible existence of ferromagnetism in hexagonal Co and manganese compounds: MnAs, MnSb, and MnBi. A phase diagram is constructed for the existence of ferromagnetic ordering and it is established that the Curie temperature depends on the occupancy of the transition-element 3*d*-shell. © *2000 MAIK "Nauka/Interperiodica".*

1. INTRODUCTION

Exchange interaction of free electrons, if considered as a perturbation, invariably results in a negative energy correction which corresponds to a tendency of the electrons toward ferromagnetism.

In studies of the excitations of localized *s*-electrons strong Hubbard repulsion at low electron density also leads to a substantial increase in the spin component of the paramagnetic susceptibility although no ferromagnetism occurs (see, for example, [1]).

In order to explain the reasons for the appearance of ferromagnetic ordering in α-iron, cobalt, and nickel, the lowest-energy 3*d* states corresponding to a particular crystal lattice must be considered as the zeroth approximation.

An analysis of the ferromagnetic states of manganese having an NiAs structure is exactly the same as an analysis of hexagonal cobalt which at $T < 723$ K has an almost ideal hcp structure.

The Hubbard energy for manganese is 15.27 eV and 17.77 eV for cobalt so that in both cases, this energy is the largest energy parameter and is subsequently considered to be infinite.

In hexagonal cobalt the fourfold degenerate $x^2 - y^2$, $2xy$-shell is filled by holes. In order to calculate the orbital magnetism it is convenient to diagonalize the atomic Hamiltonian and consider the filling of the $(x \pm iy)^2$-shell.

When studying the band-structure of the high-spin manganese states, we need to consider all possible single-particle transitions from three-electrons *xz*, *yz*, and $3z^2 - r^2$ states with spin $S = 3/2$ to the four-electron states *xz*, *yz*, $3z^2 - r^2$, $(x \pm iy)^2$, for which the sign of the projection of the orbital momentum differs and $S = 2$.

It follows from general reasoning, and will become apparent from the following, that in zero magnetic field, single-particle excitations only differ from single-hole ones in terms of sign.

2. GENERAL EQUATIONS

In compounds having a Γ12 structure the crystal field splits the atomic levels into levels for which the projection of the orbital momentum is zero and which are doubly degenerate with respect to the electron spin projection, and fourfold degenerate levels for which the momentum projection is nonzero and which are degenerate with respect to the sign of the momentum projection and with respect to the sign of the spin projection.

For 3*d* cations the levels are split into the lowest fourfold degenerate *xz*, *yz* level, the doubly degenerate $3z^2 - r^2$ level, and also the highest-energy quadruple $2xy$, $x^2 - y^2$ level.

We shall assume that hopping takes place between neighboring atoms, each having a wave function proportional to one of the components $2xy$ or $x^2 - y^2$. In order to study the magnetic properties associated with orbital splitting we need to use those linear combinations of atomic wave functions which diagonalize the atomic Hamiltonian. This condition is satisfied by two complex-conjugate wave functions $(x \pm iy)^2$having different energies in a given magnetic field ($\pm 2H$) where the magnetic field is measured in energy units.

The matrix of hopping between nearest ions $t_n^m(\mathbf{r})$ does not depend on the spin index σ but depends strongly on the orbital indices *m* and *n*.

Introducing the creation $\hat{a}^+_{(m,\sigma)}(\mathbf{r})$ and annihilation $\hat{a}_{(m,\sigma)}(\mathbf{r})$ operators of holes states in a cell having the coordinates **r**, we write the interaction Hamiltonian in

terms of the matrix of transitions between neighboring atoms:

$$\hat{H} = \sum_{m, n, \sigma, \mathbf{r}, \mathbf{r}'} t_n^m(\mathbf{r}-\mathbf{r}')\hat{a}^+_{(m,\sigma)}(\mathbf{r})\hat{a}_{(n,\sigma)}(\mathbf{r}) - \sum_{m=\pm,\sigma,\mathbf{r}} (\mu + \sigma H - 2mH)\hat{a}^+_{(m,\sigma)}(\mathbf{r})\hat{a}_{(m,\sigma)}(\mathbf{r}). \quad (1)$$

After diagonalizing the zeroth Hamiltonian corresponding to nonoverlapping atomic states, the creation and annihilation operations are expressed as an expansion in terms of all possible transitions between N and $N \pm 1$-particle states (see [2]):

$$\hat{a}^+_{(m,\sigma)}(\mathbf{r}) = \sum_\alpha g_\alpha^{m,\sigma}\hat{X}^\alpha_{\mathbf{r}}, \quad \hat{a}_{(n,\sigma)}(\mathbf{r}) = \sum_\beta g_\beta^{n,\sigma}\hat{X}^\beta_{\mathbf{r}}. \quad (2)$$

Here the indices α and β correspond to the mutually inverse transitions $s \longrightarrow p$, i.e., $\beta(p, s) = -\alpha(s, p)$. The quantities $g_\alpha^{m,\sigma}$ are called genealogical coefficients and are calculated below.

Assuming that the Hubbard energy is infinite, we write the inverse matrix of the virtual Green's function:

$$\hat{G}^{-1} = \begin{matrix} e^{2i\phi} \\ e^{-2i\phi} \end{matrix} \begin{pmatrix} e^{2i\phi} & e^{-2i\phi} \\ (i\omega - \epsilon^+_\alpha)\delta_{\alpha,\beta} - g_\alpha f^{(+)}_\alpha U_{\mathbf{p}} g_\beta & -g_\alpha f^{(+)}_\alpha B_{\mathbf{p}} g_\gamma \\ -f^{(-)}_\nu g_\nu B^*_{\mathbf{p}} g_\beta & (i\omega - \epsilon^-_\nu)\delta_{\nu,\gamma} - f^{(-)}_\nu g_\nu U_{\mathbf{p}} g_\gamma \end{pmatrix}. \quad (3)$$

Here we introduce the external magnetic field H and the chemical potential μ:

$$\epsilon^m_\alpha = -\mu + \Sigma_\alpha(H) - 2mH - \sigma H,$$
$$f^{(\sigma,\pm)}_{\alpha(k,i)} = n^{(i)}_N + n^{(k)}_{N-1},$$

and $2m = \pm 2$ is the projection of the orbital momentum. The so-called end factors $f^{(\sigma,\pm)}_{\alpha(k,i)}$ are equal to the sum of the average occupation numbers of the N- and $N-1$-particle states, where i, k is a set of indices characterizing the N- and $N-1$-particle states corresponding to the $i \longrightarrow k$ transition.

The coefficients U and B are expressed in terms of the integral of hopping to nearest neighbors (t) and the angle of rotation of the angles of the unit cell $\varphi = \pi/3$:

$$U_{\mathbf{p}} = t\left\{\cos p_x + \cos\left(\frac{1}{2}p_x + \frac{\sqrt{3}}{2}p_y\right) + \cos\left(\frac{1}{2}p_x + \frac{\sqrt{3}}{2}p_y\right)\right\},$$
$$B_{\mathbf{p}} = t\left\{\cos p_x + e^{4i\varphi}\cos\left(\frac{1}{2}p_x + \frac{\sqrt{3}}{2}p_y\right) + e^{-4i\varphi}\cos\left(\frac{1}{2}p_x - \frac{\sqrt{3}}{2}p_y\right)\right\}, \quad (4)$$

the same as in the hopping matrix for the strong coupling approximation. In this case, the matrix elements are calculated in terms of the atomic wave functions assigned to the orbital momentum $l = 2$ and differing in respect of the sign of its projection $l_z = \pm 2$.

The equations for the variations of the N-particle states of the occupation numbers δn^s_N, where $s = 1, 2, \ldots, m$ are the numbers of the lowest-energy N-particle states, can be obtained from the general equation for the averages of the T-products of the annihilation operator $\hat{a}_\nu$ and the linear combination of m conjugate X-operators with arbitrary coefficients β_s:

$$\hat{a}_\nu = \sum_{s=1}^m g_s \hat{X}^{N-1, N(s)}_{\boldsymbol{\alpha}},$$
$$\sum_s g_s\beta_s n^s_N = T\sum_{s,k} g_k\beta_s \sum_{\omega,\mathbf{p}} G^{k,s}_\omega(\mathbf{p}) f_s e^{i\omega\delta}. \quad (5)$$

Here g_k are the given genealogical coefficients, δ is an infinitely small positive quantity, and $G^{k,s}_\omega(\mathbf{p})$ are the Fourier components of the single-particle Green's function which in the zero-loop approximation is defined in terms of its inverse matrix (3).

In the particular case of transitions between high-spin states, the indices s can be conveniently replaced by the level numbers of the Zeeman multiplets S_z split under the influence of a weak magnetic field δH.

3. ZERO-LOOP APPROXIMATION

We note that the variations of the occupation numbers of the various multiplet components are interrelated by the extremely simple equation [3]:

$$\delta n^{(S_z)} \approx \delta\left(n_0 \exp\left(\frac{w}{T}S^z H\right)\right) = \frac{w}{T}S^z \delta H n^{(S_z)},$$

where

$$S_z = -S, -S+1, \ldots, S.$$

We then obtain relationships which do not depend on the gyromagnetic factor w nor on the "zero" occupation numbers n_0:

$$\delta n^{k\sigma} = k\delta n^{\sigma}, \quad \sigma = \pm 1, \quad k = 0, \pm 1, \pm 2, \ldots, \pm S \tag{6a}$$

for integer values of the total spin S;

$$\delta n^{(k+1/2)\sigma} = (2k+1)\delta n^{\sigma/2}, \quad \sigma = \pm 1, \quad k = 0, \pm 1, \pm 2, \ldots, \pm(S-1/2) \tag{6b}$$

for half-integer values of the total spin S.

We establish coupling relationships between variations of the various high-spin multiplets formed by particles whose number differs by one. We shall confine our analysis to the zero-loop approximation (Hubbard I approximation [2]) where all the self-energy components Σ are assumed to be constant and are added to the chemical potential.

We shall first assume that the coefficients β_s in equation (5) satisfy the orthogonality conditions: $\sum_s f_c \beta_s = 0$. As a result of varying the occupation numbers according to the magnitude of the external magnetic field and going to the limit $H \longrightarrow 0$, we obtain relationships which do not depend explicitly on the applied external field:

$$\sum_{k=1} g_k \beta_k \delta n_N^k = \frac{T}{g^2} \sum_{s,n} \sum_{\omega, \mathbf{p}} g_s g_n G_\omega^{s,n}(\mathbf{p}) e^{i\omega\delta} \times \sum_k g_k \beta_k \delta f_N^K = K_0 \sum_k g_k \beta_k \delta f_N^k. \tag{7}$$

Here K_0 is the average virtual Green's function (3) calculated for zero external magnetic field. In our case

$$K_0 = \frac{1}{2} \sum_{\mathbf{p}, \lambda = \pm} n_F(\xi_{\mathbf{p}}^{\lambda}), \tag{8}$$

and f_N^k are the so-called end factors which are equal to the sum of the occupation numbers corresponding to a particular transition between various multiplets. Accordingly, for $N = 2S$ and integer spin S we have [3, 4]

$$|g_k| = \sqrt{\frac{S+k}{2S}}, \quad \delta f_N^k = \delta n_N^{k\sigma} + \delta n_{N-1}^{(k-1/2)\sigma}, \tag{9}$$

where $k = 1-S, 2-S, \ldots, S$. Substituting this definition into equation (7) and using expressions (6a) and (6b) instead of the orthogonality condition, we obtain

$$(1-K_0)\delta n_{N=2S}^{\sigma} = 2K_0 \delta n_{2S-1=N-1}^{\sigma/2}. \tag{10}$$

Relations (6) and (10 can be used to express all the variations in terms of one another and substitute them into the equation of state which we obtain from the equation of state (5) under the condition $\beta_s = g_s$:

$$\sum_{s=1}^{m} g_s^2 \delta n_N^s = T \sum_{s,k} g_s g_k \sum_{\omega, \mathbf{p}} \delta\{G_\omega^{s,k}(\mathbf{p}) f_k\} e^{i\omega\delta} = (K_0 + g^2 f D_1) \sum_k g_k^2 \delta f_N^k - g^2 f \sigma D_0 \delta H. \tag{11}$$

Here all the coefficients are calculated for zero magnetic fields and are expressed in terms of integrals of the Fermi function $n_F(\epsilon)$ and its derivative $n_F'(\epsilon)$.

$$K_0 = \frac{1}{2} \sum_{\mathbf{p}, \lambda = \pm} n_F(\xi_{\mathbf{p}}^{\lambda}),$$
$$D_n = \frac{1}{2} \sum_{\mathbf{p}, \lambda = \pm} (\epsilon_{\mathbf{p}}^{\lambda})^n n_F'(\xi_{\mathbf{p}}^{\lambda}),$$
$$\xi_{\mathbf{p}}^{\lambda} = g^2 f \epsilon_{\mathbf{p}}^{\lambda} - \mu. \tag{12}$$

The chemical potential μ is determined for $H = 0$ in terms of the total number of n_s and n_d states defined in terms of the condition of electroneutrality for each transition-group element:

$$n_d = [n_d] + f R_{[n_d]+1} K_0, \quad \text{or} \quad h_d = [h_d] + f R_{[h_d]+1} K_0. \tag{13}$$

Here the brackets $[n_d]$ denote the integer part of the average number of particles (n_d) or holes ($h_d = 10 - n_d$) in the incompletely filled d-shell. The end factors f for zero field and all the coefficients are defined for each integer-value range of the variables n_d or h_d and are given in table ($z^2 \equiv 3z^2 - r^2$).

The final equation to determine the magnetic susceptibility can be obtained by substituting into equation (11) all the variations for all possible values of the spin projection from equation (6a) or (6b). Finally, the condition for positive magnetic susceptibility has the simplest form:

$$K_0(1-K_0) > g^2 f D_1(\gamma_d + K_0). \tag{14}$$

In this relationship the dimensionless quantity γ_d is expressed in terms of the squares of the genealogical coefficients g_k^2. For transitions between smaller (in terms of number N and spin S) and larger-numbered ($N = 2S$) high-spin states we have

$$\gamma_d = \frac{2S-1}{3}, \quad g^2 = S + \frac{1}{2}. \tag{15}$$

The specific values of g^2 and γ_d are given in table.

Table

Particles	f	g^2	$S_{[n_d]}$	$R_{[n_d]+1}$	γ_d	High-spin states
$1 < n_d < 2$	$(2 + n_d)/12$	3/2	1/2	3	1/3	$ze^{i\phi}$, $ze^{-i\phi}$
$2 < n_d < 3$	$(6 - n_d)/12$	2	1	4	2/3	$ze^{i\phi}$, $ze^{-i\phi}$, z^2
$3 < n_d < 4$	$(14 - 3n_d)/20$	5/2	3/2	10	1	$ze^{i\phi}$, $ze^{-i\phi}$, z^2, $e^{\pm 2i\phi}$
$4 < n_d < 5$	$(2n_d - 5)/30$	3	2	6	4/3	$ze^{i\phi}$, $ze^{-i\phi}$, z^2, $e^{2i\phi}$, $e^{-2i\phi}$
Holes	f	g^2	$S_{[n_d]}$	$R_{[n_d]+1}$	γ_d	High-spin states
$4 < h_d < 5$	$(2h_d - 5)/30$	3	2	6	4/3	$e^{2i\phi}$, $e^{-2i\phi}$, z^2, $ze^{i\phi}$, $ze^{-i\phi}$
$3 < h_d < 4$	$(14 - 3h_d)/20$	5/2	3/2	10	1	$e^{2i\phi}$, $e^{-2i\phi}$, z^2, $ze^{\pm i\phi}$
$2 < h_d < 3$	$(6 - h_d)/12$	2	1	4	2/3	$e^{2i\phi}$, $e^{-2i\phi}$, z^2
$1 < h_d < 2$	$(2 + h_d)/12$	3/2	1/2	3	1/3	$e^{2i\phi}$, $e^{-2i\phi}$

4. SINGLE-LOOP APPROXIMATION

The final expression (14) in the previous section was obtained in the simplest Hubbard I approximation where the temperature and field dependences of the end factors are taken into account. In this case, all the self-energy components are assumed to be the same so that their contribution reduces to an additive correction to the chemical potential.

In the next single-loop approximation the self-energy components Σ_k do not depend on the frequency or momentum although their dependence on the magnetic field differs substantially for different numbers (k) of single-particle transitions.

In order to find the single-loop self-energy components, it is sufficient to calculate the various loops and then sum these taking into account commutation rules which determine the nonzero vertex components of the kinematic interaction, see Fig. 1 and also [3].

The equations for the single-loop self-energy components can be expressed in terms of the product of the hopping integrals $t^{k,n}(\mathbf{p})$ and the Fourier components of the virtual Green's function:

$$A_m^\sigma = \sum_{s,\omega,a',\mathbf{p}} t_{a,a'}^{m,s}(\mathbf{p}) G_{\omega,a',a}^{s,m}(\mathbf{p}),$$
$$B_m^\sigma = \sum_{s,\omega,b',\mathbf{p}} t_{b,b'}^{m,s}(\mathbf{p}) G_{\omega,b',b}^{s,m}(\mathbf{p}), \tag{16}$$

$$\Sigma_\alpha^\sigma = K_\alpha^k A_k^\sigma + L_\alpha^k A_k^{-\sigma} + M_\alpha^k B_k^\sigma + N_\alpha^k B_k^{-\sigma}. \tag{17}$$

Here summation is performed over the recurrent indices k and the coefficients K_α^k, L_α^k, M_α^k, and N_α^k are determined using perturbation theory in accordance with the rules of the diagram technique for Hubbard operators [3]. We can see from Fig. 1 that these coefficients are ± 1.

Allowing for the symmetry conditions of the crystal lattice, we can write: $A_m^\sigma = B_m^\sigma$.

In each specific case we can show that for zero magnetic field the self-energy components do not depend on the transition number α or the spin index σ so that their influence is reduced to renormalization of the chemical potential.

After applying an infinitely weak magnetic field δH, we have the following relationships: $\delta A_m^\sigma = \delta B_m^\sigma = -\delta A_m^{-\sigma} = -\delta B_m^{-\sigma}$. Note that each of the values of A_m^σ and B_m^σ is proportional to the square of the corresponding genealogical coefficient g_m^2. Accordingly we determine the matrix $\hat{S}$ corresponding to the right-hand side of the system (17):

$$S_{\alpha,k} = \{K_\alpha^k + L_\alpha^k - M_\alpha^k - N_\alpha^k\} g_k^2. \tag{18}$$

Differentiating (16) and (17) allowing for the explicit expression for the inverse single-particle Green's function (3), we obtain the following equations for the variations:

$$\delta\Sigma_\alpha^\sigma = -\delta\Sigma_\alpha^{-\sigma} = -[F_{\alpha,n}^{(0)} - D_{\alpha,n}^{(1)}]\delta\Sigma_n^\sigma + g^2 D_{\alpha,n}^{(2)} \delta f_n^\sigma - \sigma\delta H R_\alpha D_1. \tag{19}$$

Here the matrices $\hat{D}^{(k)} = D_k \hat{U}$ differ by the temperature factor and are proportional to the same matrix

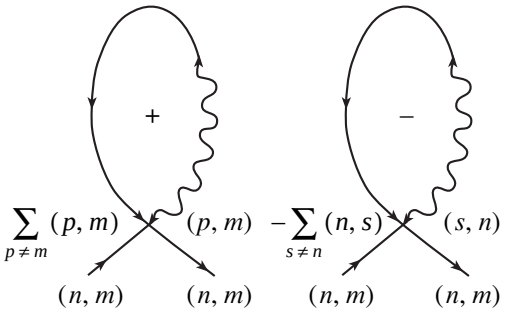


Fig. 1. Single-loop self-energy components.

$\hat{U}_{\alpha,m}$ which is expressed in terms of the numerical values of the matrix $(\hat{S})_{\alpha,m}$ constructed in accordance with the definition of the self-energy matrix (17):

$$R_\alpha = \sum_n S_{\alpha,n}, \quad U_{\alpha,n} = \frac{R_\alpha g_n^2}{g^2}, \qquad D_k = \frac{1}{2}\sum_{\mathbf{p},\lambda}(\epsilon_{\mathbf{p}}^\lambda)^k n_F(\xi_{\mathbf{p}}^\lambda). \tag{20}$$

The operator is $\hat{F}^{(0)} = Q(\mu)\hat{W}$, where

$$g^2 = \sum_k g_k^2, \qquad W_{\alpha,n} = U_{\alpha,n} - S_{\alpha,n} = \frac{g_n^2}{g^2}\sum_k S_{\alpha,k} - S_{\alpha,n}. \tag{21}$$

The matrix $\hat{W}$ has a zero sum of elements of each row:

$$Q(\mu) = \frac{1}{2}\sum_{\mathbf{p},\lambda=\pm}\frac{[n_F(\xi_{\mathbf{p}}^\lambda) - n_F(-\mu)]}{fg^2} = \frac{[K_0 - n_F(-\mu)]}{fg^2}. \tag{22}$$

The equations for the variations of the end factors also contain variations of the self-energy components:

$$\sum_{s=1}^m g_s^2 \delta n_N^s = (K_0 + g^2 f D_1)\sum_k g_k^2 \delta f_k^\sigma + f D_0 \sum_k g_k^2 \delta\Sigma_k^\sigma - g^2 f_0 \sigma D_0 \delta H. \tag{23}$$

Here the index k has all the same possible transition numbers corresponding to a given change in the projection of the momentum and spin as the index α.

Deficient equations are written using the auxiliary quantities β_k satisfying the orthogonality condition $\sum_k \beta_k g_k = 0$.

$$\sum g_k \beta_k \delta n_N^k = K_0 \sum_k g_k \beta_k \delta f_k^\delta + A(\mu)\sum_k \beta_k g_k \Sigma_k^\sigma. \tag{24}$$

$$A_m^\sigma = \frac{1}{2}\sum_{\mathbf{p},\lambda=\pm}\frac{n_F(\xi_{\mathbf{p}}^\lambda) - n_F(-\mu)}{\xi_{\mathbf{p}}^\lambda + \mu}. \tag{25}$$

Half of the variations of the N-particle states δn_N^k are expressed in terms of each other: $\delta n_N^{\sigma k} = -\delta n_N^{-\sigma k}$. All remaining independent variations of N-particle states are expressed in terms of sums of the variations of the end factors:

$$\delta n_N^{(1)} = \sum_{k=1}^m \delta f_k^\sigma, \quad \delta n_N^{(2)} = \sum_{k=2}^{m-1}\delta f_k^\sigma, \quad \delta n_N^{(3)} = \sum_{k=3}^{m-2}\delta f_k^\sigma \ldots, \tag{26}$$

so that equations (19), (23), (24), and (26) completely determine the variations of all the occupation numbers.

The condition for solvability of this system of equations for zero magnetic field $\delta H = 0$ is the condition for formation of ferromagnetic instability.

5. FERROMAGNETISM OF COBALT

Zero-loop approximation. Experiments show that the structure of hexagonal cobalt is very close to ideal close-packed: $c/a \approx 1.62$ instead of $c/a = \sqrt{8/3} \approx 1.63$ (see [5]). In accordance with the condition of electroneutrality, the difference between the number of holes (n_h) at the fourfold degenerate $2xy$, $x^2 - y^2$ shell and the number of electrons in the 4s-shell (n_s) for cobalt is 1: $n_h = n_s + 1$. According to different estimates, the observed magnetic moment is 1.6–1.75μ_B [5] so that $1 < n_h < 2$.

The end factors $f_{a,b}^{(\sigma,\pm)}$ are expressed in terms of the average occupation numbers of the two-hole $n_{\mathrm{II}}^{(0,\pm1)}$ and single-hole $n_{\mathrm{I}}^{(\sigma,\pm)}$ states

$$f_1^{(\sigma,\pm)} = n_{\mathrm{II}}^\pm + n_{\mathrm{I}}^{(\sigma,\pm)}, \quad f_2^{(\sigma,\pm)} = n_{\mathrm{II}}^0 + n_{\mathrm{I}}^{(-\sigma,\pm)}. \tag{27}$$

The matrix elements of the transition between layers in an hcp-lattice are proportional to the fourth power of the sine of the angle between the z axis and the vector joining the nearest atoms between the layers. This value is close to 1/9 so that for the wave functions used $(x \pm iy)^2$ transitions between layers can be neglected and the following analysis is made in the xy plane.

Dividing the equations into even and odd in terms of spin index which gives odd and even solutions in terms of the sign of the orbital momentum projection, we find the conditions for solving these equations in the following general form:

$$K_0(1-K_0) = g^2 f_h D_1\left(\frac{1}{3} + K_0\right), \qquad K_0(1-K_0) = Q_l(\gamma_l + K_0). \tag{28}$$

The first equation corresponds to the occurrence of spin instability and is the same as (14).

For $T = 0$ the condition for the occurrence of ferromagnetism is determined by the density of states at the Fermi surface:

$$K_0(1-K_0) = -\mu\frac{1}{2}\sum_{\lambda,\mathbf{p}}\delta(\xi_{\mathbf{p}}^{\lambda})\left(\frac{1}{3}+K_0\right) = -\bar{\mu}\frac{1}{2}\sum_{\lambda}\rho^{\lambda}(\epsilon=\bar{\mu})\left(\frac{1}{3}+K_0\right) \quad (29)$$

for $\bar{\mu} = \mu/f_h$. The condition for the occurrence of orbital instability (28) which was obtained in a study by one of the authors [6] contains the coefficient Q_l whose absolute value does not exceed one:

$$Q_l = \sum_{\mathbf{p}}\left\{\frac{(\epsilon_{\mathbf{p}}^{(+)}+\epsilon_{\mathbf{p}}^{(-)})(n_F(\xi_{\mathbf{p}}^{(+)})-n_F(\xi_{\mathbf{p}}^{(-)}))}{2(\epsilon_{\mathbf{p}}^{(+)}-\epsilon_{\mathbf{p}}^{(-)})}\right\}.$$

In our case, $\gamma_l = -1$ so that orbital instability does not occur at these temperatures.

A study of spin instability involves calculating two densities of states:

$$\rho^{(\pm)}(\epsilon) = \sum_{\mathbf{p}}\delta(\epsilon-\epsilon_{\mathbf{p}}^{(\pm)}) = \sum_{\mathbf{p}}\delta(\epsilon-U_{\mathbf{p}}\mp\sqrt{U_{\mathbf{p}}^2-3V_{\mathbf{p}}}). \quad (30)$$

For a triangular lattice the functions $U_{\mathbf{p}}$ and $V_{\mathbf{p}}$ have the following form:

$$U_{\mathbf{p}} = t\sum_{k=1}^{3}\cos p_k, \quad V_{\mathbf{p}} = t^2\sum_{k,n;\,k>n}\cos p_k\cos p_n, \quad (31)$$

where

$$p_1 = p_x, \quad p_2 = \frac{p_x}{2}+\frac{p_y\sqrt{3}}{2}, \quad p_3 = \frac{p_x}{2}-\frac{p_y\sqrt{3}}{2}.$$

For the case of hole filling, the integral of hopping to nearest neighbors is $t > 0$ so that this can be set at 1 assuming that the temperature and chemical potential are expressed in units of t. This value is known from band calculations and for cobalt is (8/9) eV $\approx 10^4$ K.

In the range adjacent to $n \approx 1$ the system has a fairly high transition temperature $\approx 0.3|t|$ which decreases with increasing hole density (see Fig. 2a).

In the range $1.062 < n_h < 1.139$ the transition temperature goes to zero and then it has a small maximum. This maximum is attributed to the logarithmic singularity in the density of states at $\epsilon = \epsilon^* = -33/16$ and may be calculated with logarithmic accuracy:

$$T_m \approx \frac{\gamma}{\pi}|t|e^{-1/\Lambda},$$

where

$$\Lambda = \frac{11}{\pi^2\sqrt{42}}\frac{1/3+K(\epsilon^*)}{K(\epsilon^*)(1-K(\epsilon^*))} \approx 0.172\frac{1/3+K(\epsilon^*)}{K(\epsilon^*)(1-K(\epsilon^*))}. \quad (32)$$

Here $\ln\gamma = C \approx 0.577$ is the Euler constant,

$$K(\epsilon^*) = \frac{1}{2}\sum_{\mathbf{p},\lambda}\Theta\left(-\frac{33}{16}-\epsilon_{\mathbf{p}}^{\lambda}\right) \approx 0.187,$$

$$\Lambda = 0.5887, \quad T_m \approx 0.567te^{-1/\Lambda} \approx 0.11t, \quad (33)$$

$$n_h = 2\frac{2+K_0}{4-K_0},$$

if $K_0 = 0.187$ then $n_h = 1.147$.

The transition temperature goes to zero for $\tilde{\mu} = -33/16 \pm \delta\mu = -33/16 \pm (\pi/2\gamma)T_m$ so that $T_m = 2\delta\mu\gamma/\pi \approx 1.13\delta\mu$.

Single-loop approximation. In order to refine the results of the previous section we calculate the single-loop self-energy components.

It may be noted that the off-diagonal self-energy components in terms of transition numbers (k, $p = 1, 2$) go to zero whereas the diagonal ones are determined in terms of integrals of the Green's functions with a given projection of the spin σ (see Fig. 1):

$$\begin{aligned}\Sigma_1^{a,\sigma} &= -A_2^{-\sigma}+B_1^{\sigma},\\ \Sigma_2^{a,\sigma} &= -A_1^{-\sigma}+B_2^{-\sigma}+A_2^{-\sigma}+B_2^{\sigma},\\ \Sigma_1^{b,\sigma} &= -B_2^{-\sigma}+A_1^{\sigma},\\ \Sigma_2^{b,\sigma} &= -B_1^{-\sigma}+A_2^{-\sigma}+B_2^{-\sigma}+A_2^{\sigma}.\end{aligned} \quad (34)$$

It can be shown that quantities on the right do not depend on the number (superscript a or b) of the atomic state and for given σ differ by a factor proportional to the square of the corresponding genealogical coefficient: $g_1^2 = 1$ and $g_2^2 = 1/2$.

$$A_k^{\sigma} = B_k^{\sigma} = T\sum_{n,\omega,\mathbf{p}}t_{\mathbf{p}}^{k,n}G_{\omega}^{n,k} \sim g_k^2C(\sigma). \quad (35)$$

In a zero magnetic field, the functions $C(\sigma)$ do not depend on the spin projection so that after substituting these into (34), we can observe that in this limit both self-energy components reduce to a constant component which gives a correction to the chemical potential and is subsequently neglected.

We obtain two equations for $\delta\Sigma_k^\sigma$ using their different definitions in terms of the Green's function:

$$\begin{aligned}\delta\Sigma_k^\sigma &= -\delta\Sigma_k^{-\sigma} = -[F_{k,n}^{(0)} - D_{k,n}^{(1)}]\delta\Sigma_n^\sigma \\ &+ g^2 D_{k,n}^{(2)}\delta f_n^\sigma - \sigma\delta H R_k D_1.\end{aligned} \tag{36}$$

Here the matrices

$$\hat{D}^{(n)} = \frac{1}{2}\sum_{\mathbf{p},\,\lambda=\pm}(\xi_{\mathbf{p}}^\lambda + \mu)^n n_F'(\xi_{\mathbf{p}})\hat{U}$$

differ by the temperature factor and are proportional to the same matrix $\hat{U}_{n,m} = R_n g_m^2/g^2$ which is expressed in terms of numerical values of the matrix $\hat{S}$ constructed in accordance with the definition of the self-energy matrix (for $A_k = B_k$):

$$\hat{S} = \begin{pmatrix} g_1^2 = 1, & g_2^2 = 1/2 \\ g_1^2 = 1, & -g_2^2 = -1/2 \end{pmatrix}, \tag{37}$$

$$R_k = \sum_n S_{k,n} = \left(\frac{3}{2}, \frac{1}{2}\right).$$

The matrix $\hat{U}$ is expressed in the form of the products

$$U_{k,n} = \frac{R_k}{g^2}g_n^2, \quad \text{so that} \quad \hat{U} = \begin{pmatrix} 1, & 1/2 \\ 1/3, & 1/6 \end{pmatrix}. \tag{38}$$

The operator is $\hat{F}^{(0)} = Q(\mu)\hat{W}$ where

$$\begin{aligned}Q(\mu) &= \frac{1}{fg^2}\sum_{\mathbf{p}}[n_F(\xi_{\mathbf{p}}) - n_F(-\mu)] \\ &= \frac{1}{fg^2}[K_0 - n_F(-\mu)].\end{aligned}$$

The matrix $\hat{W}$ has a zero sum of elements in each row:

$$\begin{aligned}\hat{W} &= \hat{U} - \hat{S} \\ &= \begin{pmatrix} U_{1,1} - g_1^2 = 0 & U_{1,2} - g_2^2 = 0 \\ U_{2,1} - g_1^2 = -2/3 & U_{2,2} + g_2^2 = 2/3 \end{pmatrix}.\end{aligned} \tag{39}$$

Thus, it is convenient to use the first equation:

$$\begin{aligned}&\sigma\Sigma_1^\sigma - D_1\sum_{k=1,2} g_k^2\delta\Sigma_k^\sigma \\ &- D_2 g^2\sum_{k=1,2} g_k^2\delta f_k^\sigma = g^2\sigma\delta H D_1.\end{aligned} \tag{40}$$

Here we introduce the squares of the genealogical coefficients: $g_1^2 = 1$, $g_2^2 = 1/2$, and $g^2 = \sum_k g_k^2 = 3/2$.

Since the influence of the external field and also the relation with the variation of the end factors are determined by the vector **R**, we can obtain general relationships which depend only on the variations $\delta\Sigma_k$. In order to achieve this aim, we multiply both sides of equations (36) by the components of the vector **N** = (–1/2, –3/2) orthogonal to the vector **R**. As a result, we obtain an equation which does not depend on the external field or the end factors f_k:

$$g_2^2(\delta\Sigma_1^{(\sigma)} + \delta\Sigma_2^{(\sigma)}) = (g_1^2 + Q(\mu))(\delta\Sigma_1^{(\sigma)} - \delta\Sigma_2^{(\sigma)}). \tag{41}$$

The occurrence of ferromagnetic instability is determined by the condition that it is impossible to solve this system of equations for $\delta H = 0$. In other words, the determinant of the following matrix must be equated to zero:

$$\begin{pmatrix} 1 - g_1^2(K_0 + fg^2D_1) & 1 - g_2^2(K_0 + fg^2D_1) & -fD_0g_1^2 & -fD_0g_2^2 \\ 1 - K_0 & 1 + K_0 & -A(\mu) & A(\mu) \\ -D_2g^2g_1^2 & -D_2g^2g_2^2 & 1 - D_1g_1^2 & -D_1g_2^2 \\ 0 & 0 & g_2^2 - g_1^2 - Q & g_1^2 + g_2^2 + Q \end{pmatrix}. \tag{42}$$

The matrix thus written can be represented as a determinant in a simple form where the only singular operator $A(\mu)$ is multiplied by the operator D_2 which completely compensates for the logarithmic singularity of $\hat{A}$ in the limit $\mu \longrightarrow 0$.

For the fcc phase of cobalt, which exists at high temperatures (723 < T < 1768 K), we have

$$U_{\mathbf{p}} = \cos p_x\cos p_y + \cos p_y\cos p_z + \cos p_z\cos p_x,$$

$$V_{\mathbf{p}} = \cos p_x\cos p_y\cos p_z\sum_k \cos p_k.$$

In the temperature range $T < 723$ K where cobalt exists in the hcp phase, the functions $U_{\mathbf{p}}$ and $V_{\mathbf{p}}$ were determined in (31).

The logarithmic divergence appearing in the coefficient $A(\mu)$ for $\mu \longrightarrow 0$ after calculating the determinant (42) is compensated by the factor D_2:

$$K_0(1-K_0) = D_1\left[fg^2\left(\frac{1}{3}+K_0\right)+g^2\frac{K_0(1-K_0)}{Q+g^2}\left(Q+\frac{7}{6}\right)\right] - \frac{D_2}{2}f\frac{A(\mu)}{Q+g^2}+fg^2\frac{D_0D_2-D_1^2}{Q+g^2}\left(\frac{1}{3}+K_0\right)\left(Q+\frac{7}{6}\right). \quad (43)$$

The first term on the right-hand side is the same as the formula in the zero-loop approximation (28). The second and third terms give a contribution of opposite sign but their difference remains positive. As a result, it is found that the transition temperature remains finite over the entire range of hole concentrations between 1 and 1.17 and the height of the logarithmic maximum increases substantially.

6. FERROMAGNETISM OF MnAs COMPOUNDS

Zero-loop approximation. The experiment shows that all manganese compounds having an NiAs structure: MnAs, MnSb, MnBi have an extremely high saturation magnetic moment higher than $3.4\mu_B$ (see, for example [7]).

Thus, the 3*d* electrons of manganese in these compounds resonate between states with spin 3/2 and a state with spin 2. We therefore need to consider the situation where the system resonates between three-particle $\{(3z^2-r^2), t(zx), t(yz)\}$ and four-particle $\{(3z^2-r^2), t(zx), t(zy), \{2xy, x^2-y^2\}\}$states, each corresponding to the lowest-energy states with the highest possible total spin.

The integrals of hopping to nearest neighbors (t) for single-particle electron transitions have a negative sign so that t may be replaced by –1 and it can be assumed that the temperature and chemical potential are expressed in units of $|t|$:

$$\xi_{\mathbf{p}}^{\pm} = fg^2\epsilon_{\mathbf{p}}^{\pm}-\mu, \quad \epsilon_{\mathbf{p}}^{\pm} = U_{\mathbf{p}} \pm \sqrt{U_{\mathbf{p}}^2-3V_{\mathbf{p}}}, \qquad f = \frac{14-3n}{20}, \quad (44)$$

$g^2 = 5/2$ is the sum of the squares of all the genealogical coefficients, $3 < n < 4$.

The excitation spectrum $\xi_{\mathbf{p}}^{(\pm)}$ only depends on the transverse quasi-momenta $\alpha = p_x/2 + p_y\sqrt{3}/2$, $\beta = p_x/2 - p_y\sqrt{3}/2$:

$$U_{\mathbf{p}} = -\cos\alpha - \cos\beta - \cos(\alpha+\beta),$$
$$V_{\mathbf{p}} = \cos\alpha\cos\beta + (\cos\alpha+\cos\beta)\cos(\alpha+\beta).$$

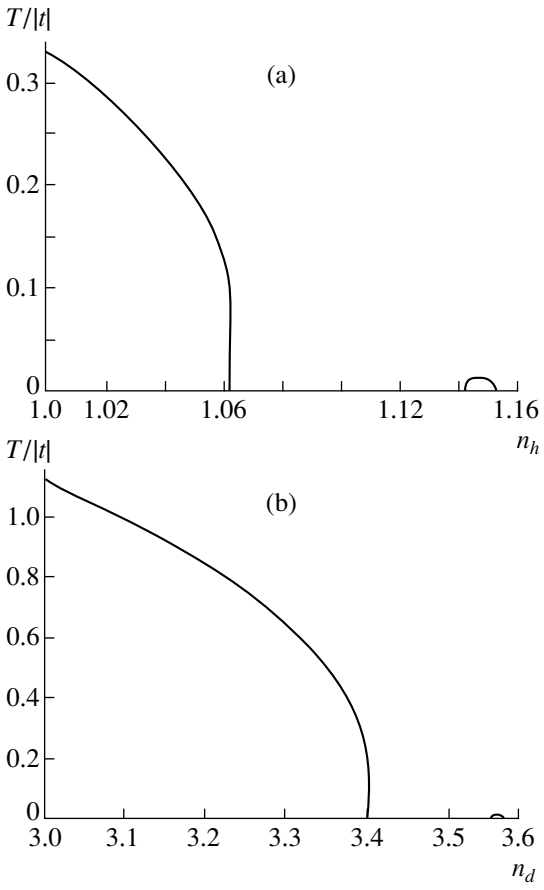


Fig. 2. (a) Dependence of the temperature of transition to the ferromagnetic state on the number of holes calculated in the zero-loop approximation for cobalt. (b) Dependence of the temperature of transition to the ferromagnetic state on the number of *d*-electrons calculated in the zero-loop approximation for MnAs.

To these equations we need to add the equation of state for $H = 0$:

$$n = 2\frac{3+7K_0}{2+3K_0}, \quad 3 < n < 4. \quad (45)$$

The condition for the occurrence of ferromagnetic instability in the zero-loop approximation has the same general form as (14) in which the amplitude $\gamma_s = 1$ is substituted.

$$K_0(1-K_0) = fg^2D_1(1+K_0), \qquad n_d = 2\frac{3+7K_0}{2+3K_0}. \quad (46)$$

In the range adjacent to $n = 3$ the system has a fairly high transition temperature $\approx 1.1|t|$ which decreases with increasing electron density (see Fig. 2b).

The transition temperature goes to zero in the range $3.41 < n_d < 3.56$. With further increasing concentration we observed a small maximum (see Fig. 2b). This max-

imum is associated with the presence of singular saddle points ($\alpha = \beta = \pi$); ($\alpha = 0$, $\beta = \pi$); ($\alpha = \pi$, $\beta = 0$) which leads to a logarithmic singularity in the density of states for $\epsilon = -1$. The maximum transition temperature is determined from the same relationship as for cobalt but with a different constant Λ:

$$T_m \approx \frac{\gamma}{\pi}|t|e^{-1/\Lambda},$$

where

$$\Lambda = \frac{3}{2\pi^2\sqrt{7}}\frac{1+K(-1)}{K(-1)(1-K(-1))} \approx 0.05745\frac{1+K(-1)}{K(-1)(1-K(-1))}. \quad (47)$$

Here $\ln\gamma = C$ is the Euler constant, $6|t|$ is the width of the conduction band,

$$K(-1) = \frac{1}{2}\sum_{\mathbf{p},\lambda}\Theta(-1-\epsilon_{\mathbf{p}}^{\lambda}) \approx 0.348$$

is the number of states within the Fermi surface $\epsilon = -1$. In this case, $n_d \approx 3.57$,

$$\Lambda \approx 0.341, \quad T_m \approx 0.567|t|e^{-1/\Lambda} \approx 0.03|t|.$$

As for cobalt the maximum transition temperature is related to the energy range $\delta\mu$ in which the ferromagnetic transition temperature has a very steep maximum. This relationship has the very simple form: $T_m = 2\gamma|\delta\mu|/\pi \approx 1.13|\delta\mu|$.

Single-loop approximation. We obtain four equations for $\delta\Sigma_m$ directly from (16) and (17) in terms of the integrals of the Green's function, the so-called single-loop approximation (see Fig. 1):

$$\begin{aligned}
\Sigma_1^{\sigma} &= B_1^{\sigma} + A_4^{-\sigma} + B_4^{-\sigma},\\
\Sigma_2^{\sigma} &= B_2^{\sigma} + A_3^{-\sigma} + B_3^{-\sigma} - A_4^{-\sigma},\\
\Sigma_3^{\sigma} &= A_2^{-\sigma} + B_2^{-\sigma} + B_3^{\sigma} - A_3^{-\sigma},\\
\Sigma_4^{\sigma} &= A_1^{-\sigma} + B_1^{-\sigma} - A_2^{-\sigma} + B_4^{\sigma}.
\end{aligned} \quad (48)$$

Taking into account the symmetry conditions, we can write:

$$A_m^{\sigma} = B_m^{\sigma},$$

so that we obtain four equations for the variation $\delta\Sigma_m$:

$$\delta\Sigma_m^{\sigma} = -\delta\Sigma_m^{-\sigma} = -[F_{m,n}^{(0)} - D_{m,n}^{(1)}]\delta\Sigma_n^{\sigma} + g^2 D_{m,n}^{(2)}\delta f_n^{\sigma} - \sigma\delta H R_m D_1. \quad (49)$$

Here the matrices $\hat{D}^k = D_k\hat{U}$ differ by the temperature factor and are proportional to the same matrix $\hat{U}_{n,m} = R_n g_m^2/g^2$ which is expressed in terms of the numerical values of the matrix $\hat{S}$ constructed in accordance with the definition of the self-energy matrix:

$$\hat{S} = \begin{pmatrix} g_1^2 = 1 & 0 & 0 & -2g_4^2 = -\frac{1}{2} \\ 0 & g_2^2 = \frac{3}{4} & -2g_3^2 = -1 & g_4^2 = \frac{1}{4} \\ 0 & -2g_2^2 = -\frac{3}{2} & 2g_3^2 = 1 & 0 \\ -2g_1^2 = -2 & g_2^2 = \frac{3}{4} & 0 & g_4^2 = \frac{1}{4} \end{pmatrix}, \quad (50)$$

$$R_k = \sum_n S_{k,n} = \frac{1}{2}(1, 0, -1, -2). \quad (51)$$

The matrix $\hat{U}$ is represented as the products $U_{k,n} = R_k g_n^2/g^2$ or

$$\hat{U} = \frac{2}{5}\begin{pmatrix} 1/2 & 3/8 & 1/4 & 1/8 \\ 0 & 0 & 0 & 0 \\ -1/2 & -3/8 & -1/4 & -1/8 \\ -1 & -3/4 & -1/2 & -1/4 \end{pmatrix}. \quad (52)$$

The operator is $\hat{F}^{(0)} = Q(\mu)\hat{W}$ where $\hat{W} = \hat{U} - \hat{S}$,

$$Q(\mu) = \frac{1}{2}\sum_{\mathbf{p},\lambda=\pm}\frac{[n_F(\xi_{\mathbf{p}}^{\lambda}) - n_F(-\mu)]}{fg^2} = \frac{[K_0 - n_F(-\mu)]}{fg^2}.$$

The matrix $\hat{W}$ has a zero sum of matrix elements in each row:

$$\hat{W} = \begin{pmatrix} -4/5 & 3/20 & 1/10 & 11/20 \\ 0 & -3/4 & 1 & -1/4 \\ -1/5 & 27/20 & -11/10 & -1/20 \\ 8/5 & -21/20 & -1/5 & -7/20 \end{pmatrix}. \quad (53)$$

The inhomogeneous term is proportional to the vector quantity $\mathbf{R} = [1/2, 0, -1/2, -1]$.

Since the influence of the external field and also the relationship with the variation of the end factors is determined by the vector $\mathbf{R}$, we can obtain general relationships containing only $\delta\Sigma_k$. To achieve this we multiply both sides of equations (49) by the components of the vectors $\mathbf{N}^{(\lambda)}$ orthogonal to the vector $\mathbf{R}$, i.e., $(\mathbf{R}\mathbf{N}^{(\lambda)}) = 0$, or

$$\mathbf{N}^{(1)} = (0, 1, 0, 0),$$

$$\mathbf{N}^{(2)} = (1, 0, 1, 0),$$
$$\mathbf{N}^{(3)} = (1, 0, -1, 1),$$
$$\delta\Sigma_k^\sigma N_k^\lambda = Q(\mu) N_k^\lambda W_n^k \delta\Sigma_n^\sigma. \quad (54)$$

Here $\lambda = 1, 2, 3$, and the third equation can be obtained by multiplying by the vector $\mathbf{M}$ whose components satisfy the condition for orthogonality of the result of action of the operator $\hat{W}$. Thus, for an arbitrary vector H_k: $M_n W_k^n H_k = 0$.

In our case we can select $\mathbf{M} = (8, 9, 8, 5)$.

This possibility arises because the determinant of the matrix $\hat{W}$ is zero whereas none of its principal minors goes to zero. As a result, we obtain

$$\delta\Sigma_k^{(\sigma)} M_k - S_n D_1 \sigma \Sigma_k^{(\sigma)} = g^2 S_n D_2 \sigma f_n^{(\sigma)} - \sigma \delta H (M_k R_k) D_1. \quad (55)$$

Here we have

$$D_k = \frac{1}{2}\sum_{\mathbf{p},\lambda} t_{\mathbf{p}}^k n_F'(\xi_{\mathbf{p}}^\lambda), \quad S_n = M_k U_n^k.$$

In our case $S_k = -2g_k^2$.

The occurrence of ferromagnetic instability is determined from the condition that this system of equations can be solved for $\delta H = 0$. In other words, we need to equate to zero the determinant of the following matrix (where $\tilde{K} = K_0 + fg^2 D_1$):

$$\begin{pmatrix}
1 - g_1^2\tilde{K} & 3/2 - g_2^2\tilde{K} & 3/2 - g_3^2\tilde{K} & 1 - g_4^2\tilde{K} & -fD_0g_1^2 & -fD_0g_2^2 & -fD_0g_3^2 & -fD_0g_4^2 \\
0 & 1 & 1 - K_0 & K_0 & 0 & 0 & -A(\mu) & A(\mu) \\
1 - K_0 & K_0 & 0 & 1 & -A(\mu) & A(\mu) & 0 & 0 \\
1 - K_0 & 3 - K_0 & 3 + K_0 & 1 + K_0 & -A(\mu) & -A(\mu) & A(\mu) & A(\mu) \\
2D_2g^2g_1^2 & 2D_2g^2g_2^2 & 2D_2g^2g_3^2 & 2D_2g^2g_4^2 & 8 + 2D_1g_1^2 & 9 + 2D_1g_2^2 & 8 + 2D_1g_3^2 & 5 + 2D_1g_4^2 \\
0 & 0 & 0 & 0 & 0; & 1 + (3/4)Q(\mu) & -Q(\mu) & (1/4)Q(\mu) \\
0 & 0 & 0 & 0 & 1 + Q(\mu) & -(3/2)Q(\mu) & 1 + Q(\mu) & -(1/2)Q(\mu) \\
0 & 0 & 0 & 0 & 1 - Q(\mu) & (9/4)Q(\mu) & -1 - Q(\mu) & 1 - (1/4)Q(\mu)
\end{pmatrix}. \quad (56)$$

Here the first row corresponds to the equation of state (23) differentiated with respect to the external field. The second, third, and fourth rows correspond to the equations (24) written in terms of the variation of the end factors δf_k. The fifth row corresponds to the linear combination of equations (55) for the three self-energy components thus summed to eliminate the action of the linear operator $\hat{F}$. The last three rows are written in accordance with equations (54).

The final equation to determine the ferromagnetic instability temperature generalizes the corresponding equation (46):

$$K_0(1 - K_0) = D_1\left[fg^2(1 + K_0) - \frac{K_0(1 - K_0)Q}{6(Q + 2/3)}\right] - \frac{5}{6} f \frac{A(\mu)D_2}{Q + 2/3} - g^2 f (D_2D_0 - D_1^2)\frac{(1 + K_0)Q}{6(Q + 2/3)}, \quad (57)$$

where $g^2 = 5/2$, $Q = (K_0 - n_F(-\mu))/fg^2$.

For a small number of excitations $n - 3 \ll 1$ we find $K_0 \approx 5Q/8 \ll 1$ so that the left-hand side of equation (57) vanishes whereas the right-hand side remains positive, which corresponds to ferromagnetic instability.

The first term on the right-hand side is the same as the right-hand side of the equation in the zero-loop approximation (46) and in the low-temperature limit is proportional to the product of the density of states and the Fermi energy having the opposite sign.

As the number of excitations increases, the left-hand side of equation (57) increases whereas the increase of the right-hand side is slowed by the increased contribution of the second and third terms, each having a negative sign at negative Fermi energy.

Thus, allowance for single-loop corrections changes the condition of occurrence of ferromagnetism compared with the zero-loop condition (46).

In the limit $T = 0$ when $D_2D_0 = D_1^2$, the concentration range of existence of ferromagnetism becomes slightly narrower and the small logarithmic region originating from the van Hove singularity completely disappears.

Another two numbers also exist, $Q_+ \approx 1.457$ and $Q_- \approx -0.457$, for which the determinant of the matrix (56) goes to zero. The first number should not be considered since it is too large and corresponds to a nonphysical value of $K_0 > 1$.

The second number $Q_- < 0$ corresponds to positive Fermi energy and for $T = 0$ corresponds to a concentration $n \approx 3.95$ close to four. It can be shown that in the region $\mu > 0$ the amplitude of the scattering of excitations with opposite spin projection is negative and thus the system cannot be transferred to a state of ferromagnetic ordering. Thus, the question of stability for $4 - n$ requires special analysis.

7. DISCUSSION OF RESULTS

For cobalt and for manganese compounds we have therefore observed a ferromagnetic region adjacent to the integer-value concentrations: $n_{Co} \geq 1$ and $n_{Mn} \geq 3$. Here the transition temperature is high and has the order of the hopping integral.

As the concentration increases, the transition temperature decreases rapidly to $0.1|t|$ when it becomes multivalued, corresponding to a first-order phase transition.

As the concentration increases further, a fairly narrow region of existence of ferromagnetism appears whose existence is attributed to the presence of a van Hove singularity in the density of states (Fig. 2).

In this range of concentrations the Curie temperature does not exceed $0.1|t|$ and a difference is observed between the ferromagnetism of cobalt and the high-spin states of manganese.

For hexagonal cobalt on going from the zero-loop to the single-loop approximation the Curie temperature increases so substantially that it remains finite over the entire range of hole concentrations between 1 and 1.174. The height of the Curie temperature maximum attributed to the van Hove singularity is at least doubled.

For high-spin manganese on going from the zero-loop to the single-loop approximation the Curie temperature decreases. The van Hove maximum disappears and the transition temperature remains finite over the range of *d*-electron concentrations between 3 and 3.6.

ACKNOWLEDGMENTS

The present paper was supported by the Russian Foundation for Basic Research (project no. 98-02-17388).

REFERENCES

1. J. Hubbard and K. P. Jain, J. Phys. C **2**, 1650 (1968).
2. J. Hubbard, Proc. R. Soc. London, Ser. A **277**, 237 (1964).
3. R. O. Zaĭtsev, Zh. Éksp. Teor. Fiz. **112**, 2223 (1997) [JETP **85**, 1218 (1997)].
4. D. T. Smirnov and Yu. F. Smirnov, *Theory of Optical Spectra of Transition-Metal Ions* (Nauka, Moscow, 1977).
5. J. B. Goodenough, *Magnetism and the Chemical Bound* (Wiley, New York, 1963; Metallurgiya, Moscow, 1988).
6. R. O. Zaĭtsev, Pis'ma Zh. Éksp. Teor. Fiz. **70**, 451 (1999) [JETP Lett. **70**, 456 (1999)].
7. *Physical Quantity* (Énergoatomizdat, Moscow, 1991).

Translation was provided by AIP

Journal of Experimental and Theoretical Physics, Vol. 90, No. 5, 2000, pp. 897–904.
From Zhurnal Éksperimental'noĭ i Teoreticheskoĭ Fiziki, Vol. 117, No. 5, 2000, pp. 1030–1038.

NONLINEAR PHYSICS

Multiple Separatrix Crossing: A Chaos Structure¶

B. V. Chirikov* and V. V. Vecheslavov**

Budker Institute of Nuclear Physics, Siberian Branch, Russian Academy of Sciences, Novosibirsk, 630090 Russia
**e-mail: chirikov@inp.nsk.su*
***e-mail: vecheslavov@inp.nsk.su*

Received January 20, 2000

Abstract—Numerical experiments on the structure of the chaotic component of motion under multiple-crossing of the separatrix of a nonlinear resonance with a time-varying amplitude are described with the emphasis on the ergodicity problem. The results clearly demonstrate nonergodicity of this motion due to the presence of a regular component of a relatively small measure with a very complicated structure. A simple 2D-map per crossing is constructed that qualitatively describes the main properties of both chaotic and regular components of the motion. An empirical relation for the correlation-affected diffusion rate is found including a close vicinity of the chaos border where evidence of the critical structure is observed. Some unsolved problems and open questions are also discussed. © *2000 MAIK "Nauka/Interperiodica".*

1. INTRODUCTION

The present work continues the studies of chaotic motion under a slow separatrix crossing. This is a particular case of adiabatic processes that is very important in physics because of the adiabatic invariance, that is, of the conservation of action variables (J) under a slow parametric perturbation (even though this is only an approximate invariance). The main problem here is the degree of accuracy or of violation of the adiabatic invariance. Separatrix crossing produces the largest chaotic component in phase space whose size does not depend on the adiabatic parameter $\epsilon \longrightarrow 0$ (which nevertheless affects the detailed structure of the motion and its time scale).

In our previous paper [1], the single separatrix crossing for a particular model was described in detail. Remarkably, a fairly simple relation, that we used for the model of [2], turned out to be surprisingly accurate within a large part of the chaotic component.

In this paper, we describe the results of numerical experiments on multiple separatrix crossing. We focus on statistical properties of the motion, including the structure and measure of the regular component disseminated into the chaotic "sea" in a rather tricky way. The existence of the regular component means nonergodicity of the motion, the question which has remained unclear for a long time up until recently. To our knowledge, the nonergodicity of motion in a similar model was first predicted theoretically and estimated numerically in [3]. We have confirmed this result by different methods and found many other characteristics of the motion structure. The present work, as well as the previous one [1], was stimulated by a very interesting study of the corresponding quantum adiabaticity [4]. We use the same classical model, which is briefly described, for the reader's convenience, in the next section (for details, see [1]).

2. THE MODEL AND TECHNIQUES

The model is determined by the Hamiltonian

$$H(x, p, t) = \frac{p^2}{2} + A_0 \sin(\Omega t)\cos x, \tag{2.1}$$

which describes a single nonlinear resonance in the pendulum approximation (see, e.g., [5, 6]) with a time-varying amplitude

$$A(t) = A_0 \sin(\Omega t). \tag{2.2}$$

The dimensionless adiabaticity parameter is defined in the usual way as the ratio of perturbation/oscillation where the tilde denotes the quantities rescaled by the frequencies,

$$\epsilon = \frac{\Omega}{\sqrt{A_0}}, \tag{2.3}$$

where $\sqrt{A_0}$ is a constant frequency of the small pendulum oscillation for the maximal amplitude.

Two branches of the instant, or "frozen", separatrix at some t = const are given by the relation

$$p_s(x'; t) = \pm 2\sqrt{|A(t)|}\sin\left(\frac{x'}{2}\right),$$
$$x' = \begin{cases} x, & A(t) > 0, \\ x - \pi, & A(t) < 0. \end{cases} \tag{2.4}$$

Following previous studies of the separatrix crossing, we restrict ourselves to this frozen approximation in

¶This article was submitted by the authors in English.

Regular component under separatrix crossing

n	ϵ	$\mu_r \times 10^2$	$T \times N_{tr}$	N_b
1	0.1	0.68 ± 0.2	$2 \times 10^3 \times 1000$	200
2	0.05	0.75 ± 0.06	$4 \times 10^5 \times 200$	500
3	0.033	0.70 ± 0.2	$4 \times 10^5 \times 200$	200
4	0.033	0.81 ± 0.08	$4 \times 10^5 \times 150$	500
5	0.02	0.60 ± 0.05	$2 \times 10^6 \times 100$	200
6	0.01	0.75 ± 0.04	$4 \times 10^6 \times 100$	200

Note: ϵ is the adiabaticity parameter; μ_r is the total relative measure of regular component; T is the number of separatrix crossings for each of the N_{tr} trajectories; N_b is the number of histogram bins in Fig. 1. n is the reference number for Fig. 1.

what follows. As shown in [1], the latter provides good accuracy of rather simple theoretical relations.

In this approximation, the action variable is defined in the standard way as

$$J = \frac{1}{2\pi}\oint p(x)dx, \tag{2.5}$$

where the integral is taken over the whole period for x rotation (off the resonance) and over a half of that for x oscillation (inside the resonance). This distinction is necessary to avoid the discontinuity of J at the separatrix where the action is given by a simple expression

$$J = J_s(t) = \frac{4}{\pi}\sqrt{|A(t)|} \le J_{\max} = \frac{4}{\pi}\sqrt{A_0}. \tag{2.6}$$

At $\Omega t = 0 \pmod{\pi}$, the action is $J = |p|$, and the conjugated phase is $\theta = x$. Note that unlike p, the action $J \ge 0$ is never negative.

In what follows, we set $A_0 = 1$, and introduce the dimensionless action by the transformation $J/J_{\max} \longrightarrow J$. The crossing region swept by the separatrix is then the unit interval, and J is simply related to the crossing time $t = t_{cr}$ by

$$|A(t_{cr})| = J^2, \quad 0 \le J \le 1, \tag{2.7}$$

while the adiabaticity parameter becomes $\epsilon = \Omega$.

Numerical integration of the equations of motion for Hamiltonian (2.1) was performed in (x, p) variables using two algorithms. In most cases, it was the so-called bilateral symplectic fourth-order Runge-Kutta algorithm as in [1]. However, in a few long runs, we applied a very simple first-order algorithm as in [2], which also is symplectic and which actually amounts to the well-known standard map [5] with the time-varying parameter

$$\bar{\tilde{p}} = \tilde{p} + \tilde{A}_0 \sin(\tilde{\Omega}\tilde{t})\sin x, \quad \bar{x} = x + \bar{\tilde{p}}, \tag{2.8}$$

where the tilde denotes the quantities rescaled by the transformation

$$\tilde{A}_0 = \frac{1}{s^2}, \quad \tilde{t} = st, \quad \tilde{\Omega} = \frac{\Omega}{s}, \quad \tilde{p} = \frac{p}{s}. \tag{2.9}$$

Here, s is the scaling parameter and we remind the reader that $A_0 = 1$. The primary goal of the rescaling was to decrease the parameter $\tilde{A}_0$ that controls the computation accuracy. Usually, it was around $\tilde{A}_0 \approx 0.1$.

As is well known, the variation of J under an adiabatic perturbation consists of two qualitatively different parts: (i) the average action, which is nearly constant between the crossings up to an exponentially small correction, and which is of primary interest in our problem, and (ii) the rapid oscillation with the motion frequency. The ratio of the two time scales is $\sim\epsilon/\sqrt{|A(t)|} \ll 1$, which allows one to efficiently suppress the second unimportant part of the J variation by simply averaging $J(t)$ over a long time interval $\sim 1/\epsilon$(see [1]).

3. ERGODICITY

The ergodicity is the weakest statistical property in dynamical systems (see, e.g., [7]). Nevertheless, it is an important characteristic of the motion, necessary in statistical theory (see, e.g., [8]).

The question of ergodicity of the motion under the separatrix crossing remained open for a long time up until recently. The upper bound for the measure (the phase - space area) of a separate domain with the regular motion (a "stability islet") was estimated in [9] as $\mu_1 \lesssim \epsilon$.

To our knowledge, the nonergodicity of motion in a similar model was first predicted theoretically and estimated numerically in [3]. The authors directly calculated the number and positions of stable trajectories for two different periods. Moreover, they were able to locate some of these trajectories in the computation, thereby measuring their area in phase space (which turned out to be surprisingly small).

Here, we use a different, statistical, approach. To this end, we first obtain, from numerical experiments, the steady-state distribution $f_s(J)$ in the action. For the ergodic motion, it must be constant. Examples of the distribution are shown in Fig. 1 with the parameters listed in the table. The striking feature of all the distributions is a clear and rather specific inhomogeneity, reminiscent of a burst of icicles hanging down from a nearly "ergodic roof". This directly demonstrates the generic nonergodic character of motion under the separatrix crossing.

The histograms are normalized such that $f_s(J) = 1$ for the ergodic motion, and the sum over all the bins is also unity for any distribution. As a result, the dips in the distribution ("icicles"), indicating the regular component, are compensated by an increase in the ergodic background. The latter is clearly seen in all the distributions, especially for small J, and is a measure of the regular component. Namely, the relative measure (share) is given by the approximate relation

$$\mu_r \approx \langle f_s(J) - 1 \rangle, \quad J < J_1, \tag{3.1}$$

where J_1 is the position of the first dip from the bottom. The approximation comes from the border effects around $J = 1$ for any finite ϵ. Typically, at this theoretical border $f_s(1) \approx 0.5$, and drops to zero within the interval $|J - 1| \sim \epsilon$. For this reason, we also used other methods for measuring μ_r. One of them was the direct calculation of the area of dips in Fig. 1. Scattering of the values provides an estimate for the accuracy of measurement of μ_r which is also given in the table.

If we are interested in statistical data only, as in Fig. 1, the computation of the J value after each crossing is not necessary, nor is the averaging of $J(t)$ done in [1]. This can be used to further speed up the computation by applying a simple relation $J = |p|$ at $A(t) = 0$, that is, at every second passage between crossings (see Section 2). It is especially important for the simple code in equation (2.8) that was used, and in particular, for the longest run $n = 6$ in Fig. 1. With the main standard code, this also was used for calculating two different distributions, after odd and even passages. Both are shown in Fig. 1 for $n = 1$ and 5. The total regular areas for both distributions are close to each other,. Yet the positions of dips are different, sometimes significantly. Another interesting peculiarity is the concentration of a regular component near $J \approx 0.9$.

Even though the total regular area is very small (~1%), its local share can be as large as 20%. In spite of stability islets, the chaotic component remains connected in the whole crossing region.

The dependence $\mu_r(\epsilon)$ is weak, if any. Apparently, the measured value already is close to the asymptotic one $\mu_r(0) \approx \langle \mu_r \rangle = 0.0072$ where the average is taken over all six cases in the table.

All these peculiarities are further discussed in Section 5.

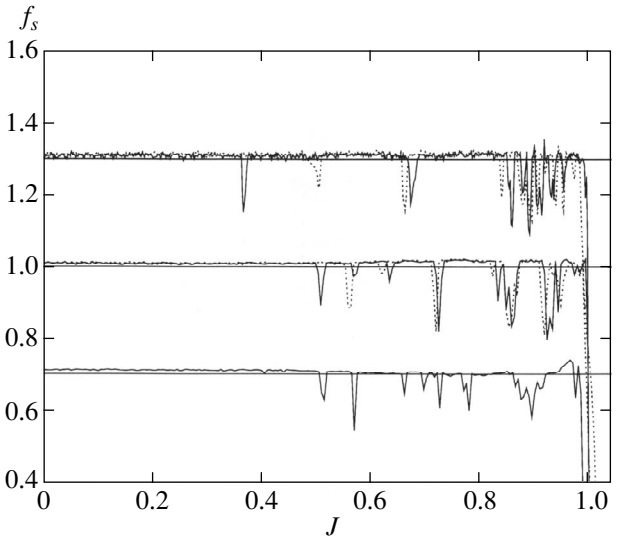


Fig. 1. Histogram of the steady-state distribution for three values of ϵ (see the table): ($n = 4$) the upper curve shifted up by 0.3; ($n = 5$) the middle curve; and ($n = 6$) the lower curve shifted down by 0.3. Solid lines correspond to J values at $|A(t)| = 1$, and the dotted ones are related to $A(t) = 0$ (see the text).

4. DIFFUSION, INSTABILITY, AND THE CRITICAL STRUCTURE

The diffusion in J was studied for a similar model in [2]. The essential difference from our mode (2.1) was the restriction of the separatrix oscillation in (2.2) by the requirement that $A(t) > 0$. In this case, the diffusive kinetics is valid in the whole crossing region. In our model, the diffusive regime is restricted to the domain $J > \epsilon^{1/3}$, while the ballistic regime takes over for $J < \epsilon^{1/3}$ with completely different kinetics (see [1] and below).

The diffusion rate in the random phase approximation (RPA) immediately follows from a simple expression for the change of J per separatrix crossing

$$\Delta J(J, \phi, \epsilon) = \mp \frac{\epsilon}{2} \frac{\sqrt{1 - J^4}}{J^2} \ln|2 \sin\phi|, \qquad (4.1)$$

where the sign coincides with that of $\dot{A}(t)$, and is given by the relation

$$D_0 = \langle (\Delta J)^2 \rangle = \frac{\epsilon^2 \pi^2}{48} \left(\frac{1}{J^4} - 1 \right), \qquad (4.2)$$

where the subscript zero indicates the RPA (see [2] and [14] therein).

The simple relation in equation (4.1) was carefully checked in [1], and proved to be surprisingly accurate in the whole diffusive region $J > \epsilon^{1/3}$. However, as was shown already in [2], the correlation-free diffusion rate (4.2) is valid for few crossings only (see also [1]). After that, the correlation in ϕ builds up, thereby decreasing the diffusion rate D by a factor of 2. We present the results of more systematic local diffusion rate measurements than in the RPA theory (4.2). To this end, we computed the correlation factor as the ratio

$$R(\langle J \rangle) = \frac{\langle D \rangle}{\langle D_0 \rangle}. \qquad (4.3)$$

This was done as follows. The number of trajectories $N_{tr} = 100$ with initial value $J = J_0$ and random x were run during $T = 800$ to 1600 separatrix crossings. The empirical diffusion rate was then calculated in the standard way, as

$$\langle D \rangle = \frac{\langle (J(T) - J_0)^2 \rangle}{T}$$

with averaging over all the trajectories, while the RPA theoretical rate $\langle D_0 \rangle$ was computed by averaging

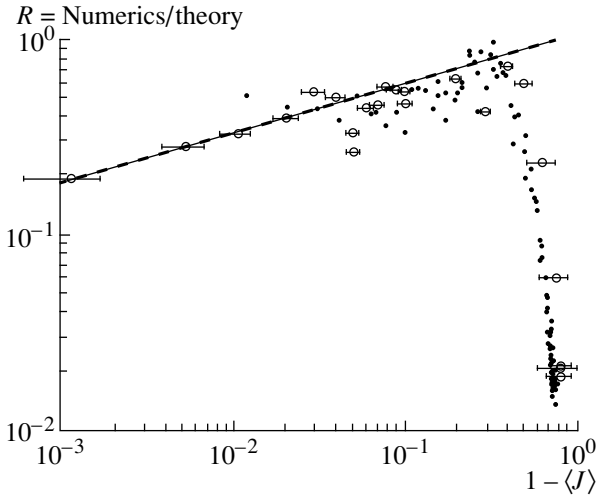


Fig. 2. The ratio of empirical to theoretical diffusion rate (the correlation factor (4.3)) vs. the mean action $\langle J\rangle$: $\epsilon = 0.001$ (circles); $\epsilon = 0.003$ (dots). Error bars show the spreading of trajectories during diffusion. The dashed straight line is fit (4.4) to four extreme left points ($\epsilon = 0.001$).

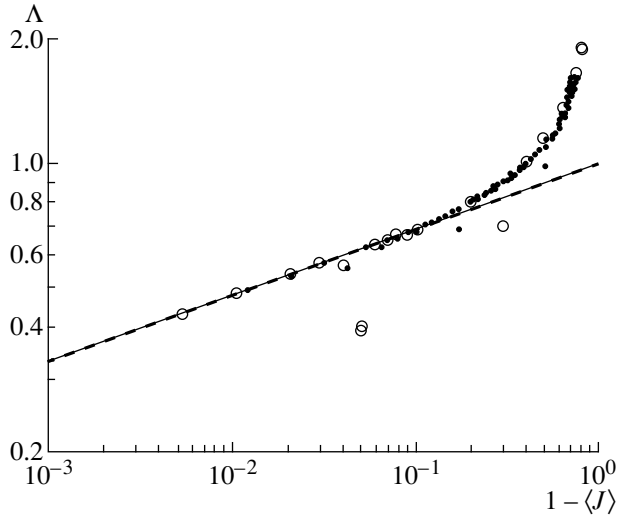


Fig. 3. The Lyapunov exponent Λ per crossing vs. mean action $\langle J\rangle$: $\epsilon = 0.001$ (circles); $\epsilon = 0.003$ (dots). The dashed straight line is fit (4.6) to ten extreme left points ($\epsilon = 0.001$).

expression (4.1) over all $N_{tr} \times T$ crossings. Altogether, 23 groups of trajectories with different initial J_0 in the whole range $0 \le J_0 < 1$ (and with random x) were run and related to the mean value $\langle J\rangle \neq J_0$ over all the crossings. Actually, all the $\langle J\rangle$ values were found to lie outside the ballistic domain because the trajectory quickly leaves the latter [1]. Nevertheless, for the initial value $J_0 > \epsilon^{1/3}$, the trajectory spent some time within this domain, and we needed a certain empirical relation for the "diffusion rate" to perform averaging $\langle D_0\rangle$. This was obtained from the results of [1] in the form

$$D_0 = 0.16\epsilon^{2/3}, \quad J < \epsilon^{1/3}.$$

It depends on ϵ but not on J.

The results of these numerical experiments are presented in Fig. 2 in the log–log scale using the quantity $1 - \langle J\rangle$ rather than $\langle J\rangle$ as the argument. The reason for this is our special interest in the asymptotic regime $J \longrightarrow 1$ at the chaos border in phase space on the edge of the crossing region. Typically, one would expect a very peculiar critical structure here (see, e.g., [8]). This interesting question is discussed later in this section.

We show in Fig. 2, the fit of the four extreme left points in the immediate vicinity of the chaos border to a power law expected in the critical structure. The result is

$$R(J) = 1.05(1 - J)^{0.25}. \qquad (4.4)$$

It is interesting that this simple relation also describes, to a reasonable accuracy, the rest of points except the five with the smallest $\langle J\rangle$ that are affected by the ballistic regime as explained in what follows. Some clear deviations from the smooth relation (4.4) reveal a certain fine structure of the diffusion of an unknown origin.

The factor R in (4.3) is always less than one, which means there is suppression of the diffusion by the correlation. The minimal suppression (maximal R) occurs at $J = J_D \approx 5\epsilon^{1/3}$, which is much larger than the crossover to the ballistic region at $J = \epsilon^{1/3}$. This is the answer to the question about the width of the ballistic-affected region put forward in the conclusion of our previous publication [1]. For $J \lesssim J_D$, the correlation strongly suppresses the diffusion down to a very low rate, which is apparently determined by fluctuations. These unusual kinetics certainly deserve further study. In any event, such a suppression explains a surprisingly long-motion time required for a good steady-state distribution in Fig. 1. The value of J_D marks the diffusion crossover from a big to a small correlation (cf. Fig. 3). In the complementary region $J \gtrsim J_D$, the correlation factor also decreases, although very slowly, see (4.4). Within fluctuations, which increase with ϵ, the factor R does not depend on ϵ (for the explanation, see Section 5).

The diffusion rate itself is given by the empirical relation

$$D(J) \approx \frac{\pi^2}{48}\epsilon^2\frac{(1 - J^4)(1 - J)^{1/4}}{J^4} \longrightarrow \frac{\pi^2}{12}\epsilon^2(1 - J)^{c_D}, \quad (4.5)$$

where the latter expression represents the asymptotics as $J \longrightarrow 1$, and $c_D \approx 5/4$ is the diffusion critical exponent.

A power law in equation (4.5) suggests the existence of a critical structure at the chaos border $J = 1$. Detailed study of this structure is hampered by some additional border effects as discussed in Section 3. Even for a rather small $\epsilon = 0.001$, we managed to follow the asymptotic behavior to $1 - J \sim 10^{-3}$ only (see Fig. 2). Also, we are not able, as yet, to calculate the critical exponent c_D from the existing resonant theory of the critical phenomena [8]. However, there is another way to test our conjecture. Namely, besides the local diffu-

sion rate, we might measure the asymptotic behavior of the Lyapunov exponent $\Lambda(J)$ In fact, we did both simultaneously in the same run.

A positive Lyapunov exponent ($\Lambda > 0$) is the main condition for the strongest statistical properties in a dynamical system, including the randomness of most trajectories [10] (see also [11, 12]). The other condition for chaos is the boundedness of motion in the phase space. The first measurement of Λ (for the same model) was reported in [13], just as a criterion for chaos. Formally, the Lyapunov exponent is defined in the ergodic theory of dynamical systems in the limit as $t \longrightarrow \infty$ [7] (as is the diffusion rate, by the way). However, for rather different time scales of motion, the local Lyapunov exponent $\Lambda(J)$ also becomes a meaningful and, moreover, a very important characteristic of the motion. Roughly, the ratio of time scales is that of error bars to the corresponding J values in Fig. 2 provided the number of crossings T per trajectory is sufficiently large for Λ to saturate.

In Fig. 3, we present the results for $\Lambda(J)$ measured, as $D(J)$, per one separatrix crossing, and for the same parameters and initial conditions as in Fig. 2. A clear crossover to asymptotic behavior is seen at $\langle J\rangle = J_\Lambda \approx 0.8$. The latter was also fitted to the power law

$$\Lambda(J) = 0.98(1 - J)^{c_\Lambda}, \tag{4.6}$$

with the critical exponent $c_\Lambda = 0.156$. In fitting, we used ten extreme left points besides the two at $\langle J\rangle = 0.95$ that represent some unknown fine structure (cf. Fig. 2). Below the crossover ($J > J_\Lambda$), the dependence is approximately linear,

$$\Lambda(J) \approx 1.9 - 1.4J. \tag{4.7}$$

The fluctuations are now much less than for $D(J)$. In both cases, the ϵ-dependence, if any, is weak. Interestingly, no effect of the ballistic region is seen for $\Lambda(J)$ (cf. Fig. 2).

The theory of critical phenomena [8] allows one to calculate the ratio of the two exponents, irrespective of other details of the critical structure. The ratio is

$$r_{th} = \frac{c_D}{c_\Lambda} = 8, \tag{4.8}$$

while the empirical value for this ratio from equations (4.5) and (4.6) is $r_{\exp} = 8.01$, a surprising agreement!

To illustrate this result, we plot, in Fig. 4, the dependence $D(\Lambda)/\epsilon^2$ together with the expected asymptotic relation

$$\frac{D}{\epsilon^2} = \Lambda^8. \tag{4.9}$$

This appealing result strongly suggests the existence of a critical structure at the chaos border $J = 1$, and further studies of this interesting problem are needed.

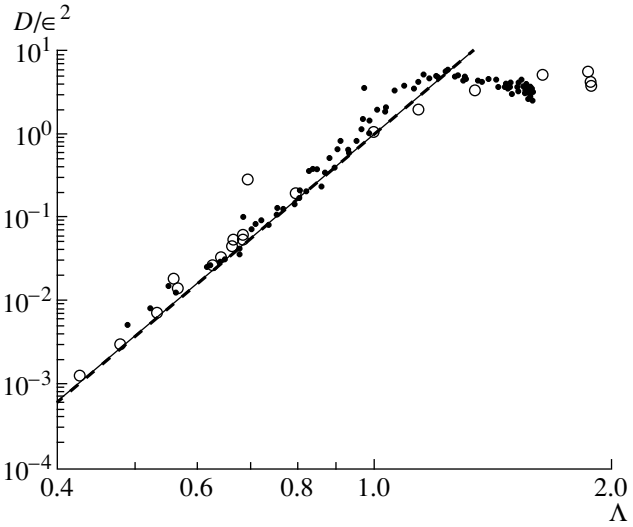


Fig. 4. Diffusion rate vs. the Lyapunov exponent: $\epsilon = 0.001$ (circles); $\epsilon = 0.003$ (dots). The dashed straight line is the theoretical prediction for the critical structure (4.9)

5. A SIMPLE MAP

Because the principal change in the adiabatic invariant J occurs at the separatrix crossing, it is natural to derive a 2D-map per crossing. These sorts of maps were considered by many authors [2, 3, 14, 15]. All these maps are rather complicated, at least for theoretical analysis. For the model under consideration here, the global map (in J) has the form

$$\bar{J} = J \mp \frac{\epsilon}{2}\frac{\sqrt{1 - J^4}}{J^2}\ln|2\sin\phi|, \tag{5.1}$$

$$\bar{\phi} = \phi + \Phi(\bar{J}),$$

where the sign coincides with that of $\dot{A}(t)$ (see equation (4.1)). The difficulty of constructing and using such a map lies in the second equation. Note that both equations are approximate and cannot be substitutes for the exact equations of motion even in the simplest form of another map (2.8).

To simplify the global map (5.1), we first transform it to a local one by the standard procedure, the linearization of the second equation (see, e.g., [5, 6]):

$$\Phi(J) \longrightarrow \pi n + \left(\frac{d\Phi}{dJ}\right)_{J = J_n}\Delta J. \tag{5.2}$$

The new parameter J_n satisfies the equation $\Phi(J_n) = \pi n$ with any integer n, and $\Delta J = J - J_n$. In our problem, this approximation is rather accurate for sufficiently small $\epsilon \longrightarrow 0$. In particular, we can consider the discrete variable J_n as a continuous one (see below).

Typically, the derivative $\Phi' = d\Phi/dJ$ is still very complicated, and we assume another principal approximation; calculating the change in ϕ between succes-

sive separatrix crossings, we use the limiting motion frequencies neglecting the change of those near the separatrix. They are

$$\omega_r = \frac{4}{\pi}J \quad \text{for phase rotation,}$$
$$\omega_o = \sqrt{A(t)} \quad \text{for phase oscillation.} \tag{5.3}$$

The rotation frequency (off the resonance) remains constant between crossings, while the oscillation slowly varies due to the separatrix motion. Now, the full period of the phase ϕ, which is equal to π, corresponds to the full period of the rotation, but only to a half of that for the oscillation. Therefore, the speed of the o variation in this approximation becomes

$$\frac{d\phi}{dt} = \begin{cases} \dfrac{\omega_r}{2} = \dfrac{2}{\pi}J, & J > \sqrt{A(t)}, \\ \omega_o = \sqrt{A(t)}, & J < \sqrt{A(t)}. \end{cases} \tag{5.4}$$

The latter inequalities determine the transition from rotation to oscillation and back, which occurs at the crossing time $t = t_{cr}$ where (see equation (2.7))

$$\epsilon t_{cr} = \arcsin(J^2). \tag{5.5}$$

For the local map in question, we need only the derivatives Φ', which are expressed in terms of elementary functions as

$$\Phi'(J) = \begin{cases} \dfrac{8}{\pi\epsilon}\left(\dfrac{1}{2}\arcsin(J^2) + \dfrac{J^2}{\sqrt{1-J^4}}\right), & J > \sqrt{A(t)}, \\ -\dfrac{4}{\epsilon}\dfrac{J^2}{\sqrt{1-J^4}}, & J < \sqrt{A(t)}. \end{cases} \tag{5.6}$$

Since the most interesting part of the motion structure is essentially concentrated near sufficiently large $J \approx 0.9$ (see Fig. 1), we can keep in the first equation (5.6) only the second term with the coefficient $4/\epsilon$ from the second equation. In fact, the difference between the two factors is less than it appears just because of the contribution of the omitted term. However, the latter correction would be certainly an excess in accuracy for our rather crude map. Finally, we assume

$$\Phi'(J) \approx \pm\frac{4}{\epsilon}\frac{J^2}{\sqrt{1-J^4}}. \tag{5.7}$$

The local map is now derived from equations (5.1), (5.2), and (5.7) in the standard way (see, e.g., [5, 6, 16]), and has the form

$$\bar{P} \approx P \mp K\ln|2\sin\phi| \mod \pi,$$
$$\bar{\phi} \approx \phi \mp \bar{P} + \frac{\pi}{4}, \tag{5.8}$$

where the signs in both equations change simultaneously at each crossing, and where

$$P = \frac{4}{\epsilon}\frac{J_n^2}{\sqrt{1-J_n^4}}\Delta J \mod \pi \tag{5.9}$$

is a new, local, momentum, and the only parameter $K \approx 2$ is simply a constant in the approximation assumed. An additional phase change by $\pi/4$ comes from the shift of the separatrix by π in x each time it crosses zero (see equation (2.4)). Literally, this change in φ is equal to $\pi/4 \pm \pi/4$, but the alternating part simply shifts P by a constant $\pi/4$ and, thus, can be omitted.

The phase space of the local map (5.8) is a 2D-torus $\pi \times \pi$. It approximately represents a narrow strip $\Delta_1 J \times \pi$ in the phase space of our main system (2.1), where

$$\Delta_1 J = \frac{\pi\epsilon}{4}\frac{\sqrt{1-J_n^4}}{J_n^2}. \tag{5.10}$$

For the local map to be applicable, the following two conditions are to be satisfied:

$$\frac{\Delta_1 J}{J_n} \approx \frac{\epsilon}{J_n^3} \lesssim 1, \tag{5.11}$$

and

$$\frac{\Delta_1 J}{1-J_n} \approx \frac{\epsilon}{\sqrt{1-J_n}} \lesssim 1. \tag{5.12}$$

The latter condition excludes a very narrow domain $1 - J_n \lesssim \epsilon^2$, which is practically impossible to observe, while the former comprises the whole ballistic region.

The density of local strips (5.10) in J_n,

$$\frac{dn}{dJ_n} \approx \frac{1}{\Delta_1 J} = \frac{4}{\pi\epsilon}\frac{J_n^2}{\sqrt{1-J_n^4}}, \tag{5.13}$$

is rapidly increasing with J_n, which explains the concentration of the regular component near the chaos border (Fig. 1). This also explains the shift δJ of the dips between two different groups in Fig. 1. The largest $\delta J \approx 0.15$ on the upper curve between the two extreme left dips is close to the full width of the corresponding local strip $\Delta_1 J \approx 0.16$.

An interesting feature of the 4-step map in equation (5.8) over a period of the adiabatic perturbation (four separatrix crossings) is a singularity at $\phi = 0 \pmod{\pi}$. The Fourier spectrum of this singularity

$$\ln|2\sin\phi| = -\sum_{n=1}^{\infty}\frac{\cos(2n\phi)}{n} \tag{5.14}$$

is similar to that of the function with a finite discontinuity. As is well known (see, e.g., [8, 17] and references therein), the chaotic component of such a motion is always connected. This means that there is no invariant

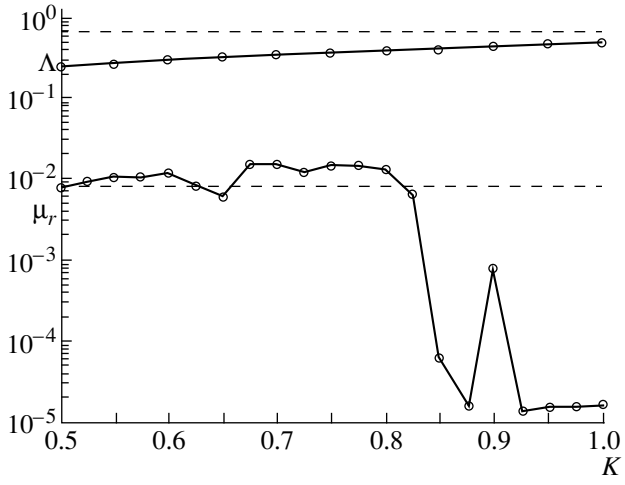


Fig. 5. Comparison of local map (5.8) (circles connected by lines to guide the eye) and the main system (2.1) (dashed lines) with respect to: the relative measure μ_r of the regular component (lower data), and the Lyapunov exponent Λ (upper data). For the main system, the dashed lines give $\mu_r = 0.007$ and $\Lambda\,(J = 0.9) = 0.67$ (see the text).

curve in the entire range $0 \le \phi \le \pi$ that would cut through and disconnect the chaotic component.

This confirms earlier conjectures on the universality of chaos under the separatrix crossings (see, e.g., [13]). The motion in such a system is typically nonergodic, that is, it contains a regular component. For a particular model under consideration, it was first found in [3], and studied in detail in the present work (Section 3). Using a simple map in equation (5.8), we are able to analyze and understand particular features of this less-known component of the motion.

To this end, we first measured the relative area μ_r of the regular component (stability islets) within the local phase-space cell ($\pi \times \pi$) as a function of the parameter K. The result is shown in Fig. 5 (lower circles). In the approximation of a constant parameter K, the relative area is the same in each cell, and thus, is approximately equal to the relative area in the whole range of J in the main system. The latter is also shown in Fig. 5 (the lower dashed line). The agreement, within a factor of 2, seems reasonable provided the local parameter is $K \lesssim 0.8$, which is about half of the estimated value. Assuming $K \approx 0.8$, we can further compare the Lyapunov exponent in the local map (upper circles in Fig. 5) with that of the main system at J = 0.9, the latter being larger by a factor of 2 (the upper dashed line).

Besides a qualitative description, therefore, a simple local map (5.8) leads to quantitative estimates within a factor of 2, which is not that bad for such a primitive map.

The local map is independent of ϵ, and so are all the dimensionless quantities of the variables and the parameters of this map. These include the relative area μ_r (cf. Fig. 1 and the table), the Lyapunov exponent Λ per separatrix crossing (or per perturbation period) (Fig. 3), and the correlation factor R (Fig. 2) except for small J close to the ballistic region, where the local map is not applicable.

6. CONCLUSION

We studied the structure and statistical properties of the chaotic motion under the separatrix crossing in numerical experiments with a typical model (2.1) used in such studies. An interesting distinction from the previous studies (except [13]) is in that we allow the full swing of the separatrix ($-1 \le A(t) \le 1$). In this case, the chaos comprises the whole range ($0 \le J \le 1$), and there is only one chaos border at $J = 1$. Usually, the perturbation amplitude $A(t) > 0$ is strictly positive (or negative) which implies two chaos borders with the chaotic component between them ($0 < J_1 \le J \le 1$), but without an interesting ballistic region.

We have qualitatively confirmed the previous results on the existence of the regular component (nonergodicity) of motion [3] and the correlation in the chaotic component suppressing the diffusion [2]; we have found many other interesting details of the motion structure (Sections 3 and 4). For a physical interpretation and understanding of our empirical results, we have constructed a very simple but meaningful local map per separatrix crossing, which leads not merely to a qualitative description of the chaos structure, but also to a reasonable quantitative estimates within a factor of 2.

In Fig. 1, most of the regular component is seen near the chaos border, at $J \approx 0.9$. We never observed any at $J = 0$, which is at variance with the prediction in [14] based on approximating the equations of motion by the Mathieu equation at small $\epsilon \longrightarrow 0$. The resolution of this apparent contradiction is that the parametric perturbation amplitude in the Mathieu equation increases as $\propto \epsilon^{-2}$ (see equation (2.9)), and therefore, stable periodic solutions are only possible in special very narrow windows of ϵ. An interesting open question is the size of the corresponding stability islets.

Another interesting problem is the expected critical structure at the chaos border $J = 1$. The standard method—statistics of the Poincare recurrences (see, e.g., [8] and references therein)—is difficult to apply here because of the confusion with many internal chaos borders around stability islets of the regular component. Instead, we measured the $J \longrightarrow 1$ asymptotic behavior of the two quantities, $\Lambda(J)$ and $R(J)$. Unfortunately, we were not able to calculate from the existing theory [8] the two critical exponents separately, because of the singularity at $J = 1$ (see equation (5.6)). However, we have found that their ratio (4.8) is independent of the singularity and agrees surprisingly well with the empirical result (Fig. 4). This is strong evidence in favor of the critical structure, and it certainly deserves further studies.

In the present work, as well as in the previous one [1], we studied the crossing of a single separatrix that is one of the two separatrix branches of a nonlinear resonance (see equation (2.1)). As is well known, there is another, related but not identical, process, the crossing of the whole resonance with both of its branches. The latter was studied even much earlier [18] (see also [19]). From the beginning, it was found that the change in the adiabatic invariant per crossing, $\Delta J \sim \epsilon \ln \epsilon$ (in dimensionless variables), differs from that for the separatrix crossing, calculated much later, by an additional factor $\ln\epsilon$, which slowly but indefinitely grows as $\epsilon \longrightarrow 0$. The importance of this factor for the regular component of the motion was understood in [3]. Namely, it was theoretically predicted that the stable trajectories of the two particular periods are destroyed, together with the surrounding islets, for sufficiently small ϵ. An interesting open question is whether the whole regular component, containing infinitely many islets [8], also vanishes, and if so, then how fast.

In terms of our local map (5.8), the additional factor would completely change all the underlying motion structure because now the map parameter $K \sim |\ln\epsilon| \longrightarrow \infty$ does depend on the adiabaticity parameter, and moreover, indefinitely grows as $\epsilon \longrightarrow 0$. This implies the ϵ-dependence of all the dimensionless characteristics of the motion, in particular, the measure of regular component. We performed some preliminary numerical experiments to estimate the dependence $\mu_r(K)$. Asymptotically, it looks like an exponential, which would imply a power law for $\mu_r(\epsilon)$.

In the very conclusion, we would like to mention that the latter particular interesting question is a part of a very important and very difficult unsolved general problem in the theory of dynamical systems, the problem of ergodicity in the case of analytic or even sufficiently smooth equations of motion.

ACKNOWLEDGMENTS

This work was partially supported by the Russian Foundation for Basic Research (grant no. 97-01-00865).

REFERENCES

1. B. V. Chirikov and V. V. Vecheslavov, Preprint No. 99-52 IYaF SO RAN (Budker Institute of Nuclear Physics, Siberian Division, Russian Academy of Sciences, Novosibirsk, 1999); Zh. Éksp. Teor. Fiz. **117**, 644 (2000) [JETP **90**, 562 (2000)].
2. D. Bruhwiler and J. Cary, Physica D (Amsterdam) **40**, 265 (1989).
3. A. I. Neishtadt, V. V. Sidorenko, and D. V. Treschev, Chaos **7**, 2 (1997).
4. B. Mirbach and G. Casati, Phys. Rev. Lett. **83**, 1327 (1999).
5. B. V. Chirikov, Phys. Rep. **52**, 263 (1979).
6. A. Lichtenberg and M. Lieberman, *Regular and Chaotic Dynamics* (Springer, New York, 1992).
7. I. Cornfeld, S. Fomin, and Ya. Sinai, *Ergodic Theory* (Springer, New York, 1982).
8. B. V. Chirikov, Chaos, Solitons and Fractals **1**, 79 (1991).
9. Y. Elskens and F. Escande, Nonlinearity **4**, 615 (1991); Physica D (Amsterdam) **62**, 66 (1993).
10. V. M. Alekseev and M. V. Yakobson, Phys. Rep. **75**, 287 (1981).
11. B. V. Chirikov, Open Systems & Information Dynamics **4**, 241 (1997).
12. B. V. Chirikov and F. Vivaldi, Physica D (Amsterdam) **129**, 223 (1999).
13. C. Menyuk, Phys. Rev. A **31**, 3282 (1985).
14. F. Escande, in *Plasma Theory and Nonlinear and Turbulent Processes in Physics*, Ed. by V. G. Baryakhtar, V. M. Chernousenko, N. S. Erokhin, A. G. Sitenko, and V. E. Zakharov (World Scientific, Singapore, 1988), p. 398.
15. J. Cary and R. Skodje, Phys. Rev. Lett. **61**, 1795 (1988); Physica D (Amsterdam) **36**, 287 (1989).
16. B. V. Chirikov, in *Reviews of Plasma Physics*, Ed. by B. B. Kadomtsev (Énergoatomizdat, Moscow, 1984; Consultants Bureau, New York, 1987), Vol. 13.
17. I. Dana, N. Murray, and I. Percival, Phys. Rev. Lett. **62**, 233 (1989).
18. B. V. Chirikov, Dokl. Akad. Nauk SSSR **125**, 1015 (1959) [Sov. Phys. Dokl. **4**, 390 (1959)]; B. V. Chirikov and D. L. Shepelyansky, Zh. Tekh. Fiz. **52**, 238 (1982) [Sov. Phys. Tech. Phys. **27**, 156 (1982)].
19. J. Cary, F. Escande, and J. Tennyson, Phys. Rev. A **34**, 4256 (1986).

Journal of Experimental and Theoretical Physics, Vol. 90, No. 5, 2000, pp. 905–907.
From Zhurnal Éksperimental'noĭ i Teoreticheskoĭ Fiziki, Vol. 117, No. 5, 2000, pp. 1039–1042.

MISCELLANEOUS

Metastability of Life¶

V. I. Marchenko

Kapitza Institute for Physical Problems, Russian Academy of Sciences, Moscow, 117334 Russia
Institut für Festkörperforschung, Forschungszentrum Jülich, D-52425, Jülich, Germany
e-mail: mar@kapitza.ras.ru

Received October 8, 1999

Abstract—The physical idea of the natural origin of diseases and deaths has been presented. The fundamental microscopical reason is the destruction of any metastable state by thermal activation of a nucleus of a irreversible change. On the basis of this idea the quantitative theory of age dependence of the death probability has been constructed. The obtained simple Death Laws are very accurately fulfilled almost for all known diseases. © *2000 MAIK "Nauka/Interperiodica".*

All of us will die, as well as all other living organisms and plants. Each and every machine or construction will break. Mountains will fall down or earthquakes will happen.

Why? Physics gives the general answer – all of these systems are not in full equilibrium, but represent metastable states. In other words: (1) they are stable against small external influences, but (2) each of them, the worst ones, as well as the best ones, has a finite probability to be spontaneously destroyed without any external influence even in the ideal environment and at perfect conditions. According to Gibbs [1], the fundamental reason of the destruction of metastable equilibrium is the thermal activation of a critical nucleus of irreversible change in the system.

Let us consider a simple example: a stretched ideal monocrystal string. If we wait a sufficiently long time, the temperature fluctuations will produce a critical Griffith's crack [2] at some place and the string will break. It is possible that the critical crack will appear earlier if there are some defects in the crystal. Such a nucleation process occurs in different ways for different cases (activation of point defects in the crystals, condensation in a super saturated solution, nucleation of a new phase in a first order phase transition) and it is well studied in condensed matter physics.

The described phenomena can also take place in any living organism, even if the latter are much more complicated. The thermal activation of a critical nucleus is the last and unremovable killer. [Last—if we exclude all other origins of diseases and deaths. Unremovable, but, one can hope, not unanalyzable.]

I want to stress here that the known qualitative and quantitative facts for the majority of diseases can be understood from the point of view of theoretical physics in terms of metastability and the activation of a critical nucleus. So, I think that the thermodynamic killer works, and that it is the main killer.

Gompertz [3] discovered that the probability $D(x)$ to die at the age x in the time interval dt exponentially increases with age

$$D \propto \exp\left(\frac{x}{a}\right). \quad (1)$$

According to modern mortality statistics, Gompertz law is valid in the age range from 30–70 years, while an even stronger increase appears in older age groups. The exponential age dependence of D, from my point of view, is the most crucial sign of the microorigin nature of diseases leading to death.

I have no answer for many questions one can ask about the details of the relationship between a given disease and the proposed idea of their natural microorigin. Nonetheless, I believe that the age dependence of the death rate can be interpreted in terms of the probability of formation of critical nuclei of irreversible change that cause the collapse of metastable equilibrium states. In the following, I present a theory that relates the probability of the arising of a critical nucleus to the age of the system in which it takes place.

Unremovable point defects on a molecular (and macromolecular) scale can arise due to the process of oxidation [4]. Thermal fluctuations can produce configurational transformations of individual molecules [5]. The same effect can be caused also by some external agents (photons, impurity atoms or molecules, elementary particles). If a concentration of those point defects is small, then the probability of the arising of new defects does not depend upon the interaction between them. It means that the concentration of point defects should be simply proportional to the age x. This linear law is known in an absolutely analogous situation, the Zeldovich stage of nucleation in I order phase transition [6]. It is quite natural to assume that, at any age, the

¶This article was submitted by the authors in English.

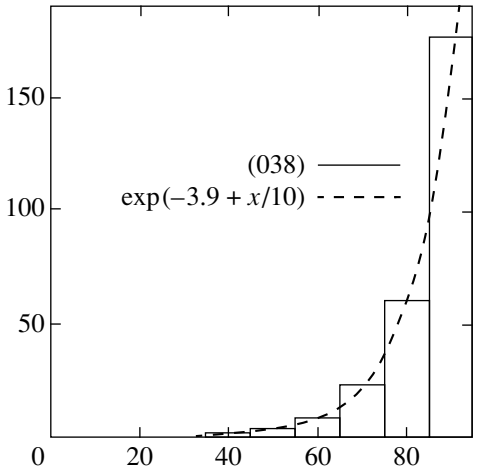


Fig. 1. Septicemia (038). Death rate.

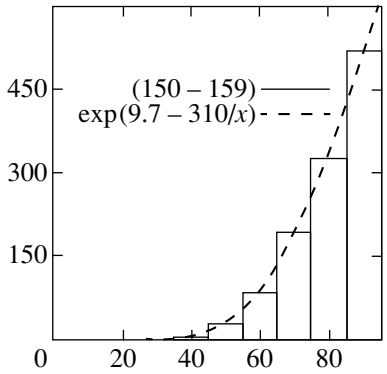


Fig. 2. Malignant neoplasms of digestive organs and peritoneum (150–159). Death rate.

dimensionless molecular concentration of the point defects remains small, so that this law is valid at any age.

A growing concentration of the point defects gives rise to small changes of the physical parameters of the body structures on a macroscopic scale (membranes, cells, as well as higher level structures). One can imagine that some functionally significant defects are thermally activated on this scale (as, e.g., the arising of a Griffith-like critical crack in a microcapillary, periodically stressed by oscillating blood pressure) or that point defects will tend to precipitate into a condensed state (as it is in supersaturated solutions), or even that some type of a structural phase transition occurs at some critical value of defect concentration. Some such types of spontaneous changing in the body can have serious functional consequences leading to diseases, and death.

The probability W of such microdamages arising is governed by the Gibbs law

$$W \propto \exp\left(-\frac{U}{T}\right), \tag{2}$$

where U is the minimum energetic barrier of the irreversible change (critical nucleus), and T is the temperature. Usually, it is possible to expand the energy of the critical nucleus in the small concentration, or equivalently in age: $U = U_0 + U'x$, and if U' is negative, the barrier diminishes with the age, we obtain the exponential law, equation (1). If U' is positive, one has the growth of the barrier, and the stability of the body increases. It is possible that the decreasing age of infant mortality is partly related to this circumstance.

The expansion of U in the concentration is impossible in the case of condensation in a supersaturated gas with a small concentration (as well as in the vicinity of I order phase transition). In a two-dimensional condensation of supersaturated gas, the energy of the critical nucleus is inversely proportional to the concentration, or in our case $U \propto x^{-1}$, corresponding to the second exponential law

$$W \propto \exp\left(-\frac{b}{x}\right). \tag{3}$$

In a three-dimensional condensation, there should be $U \propto x^{-2}$, and the third exponential law is

$$W \propto \exp\left(-\frac{c}{x^2}\right). \tag{4}$$

Let us consider the US-97 death statistics specified by selected causes [7]. If one plots $\ln(D_i)$ versus x, or versus $1/x$, and $1/x^2$, it is easy to find that almost all cases have a clearly distinguishable age behavior: 20 cases of Gompertz exponential law, equation (1); 14 cases of second exponential law (3); 4 cases with more complicated behavior, but the laws (1) or (3) are valid there in a wide age range, and some strange crossover occurs to some other behavior; 24 cases are not related with aging. Only in 3 cases statistics does not permit to make a definite conclusion on the type of age dependence. Examples of the clearly detectable exponential age behavior of the death rate are presented in Fig. 1.

The death rate here is the number of deaths per 100000 population of specified age groups 0–5, 5–14, …, 75–84, 85 years and over in 1997. There are a lot of intriguing coincidences of the parameters (a, b) for different diseases. This possibly means that a number of discussed different microorigins is substantially smaller than a number of diseases. Some of the diseases arise presumably as a combined effect of two different microorigins. This analysis is in progress.

The characteristic magnitude of function D in cases with Gompertz law (1) at $x = 0$ is exp(–13)…exp(–22) per year, or exp(–30)…exp(–39) per second. Let us compare this value with equation (2). One should introduce some preexponential value. Its simplest estimate is the characteristic frequency of oscillations of atoms in condensed matter $\omega \sim k\Theta/\hbar$, where $\Theta \sim 10^2$ K is Debye temperature; k, Boltzmann constant; $\hbar$, Planck's constant. One should introduce an additional factor, an

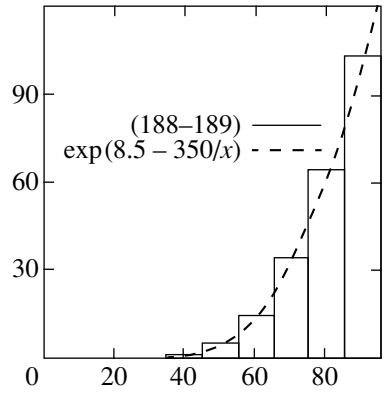


Fig. 3. Malignant neoplasms of urinary organs (188–189). Death rate.

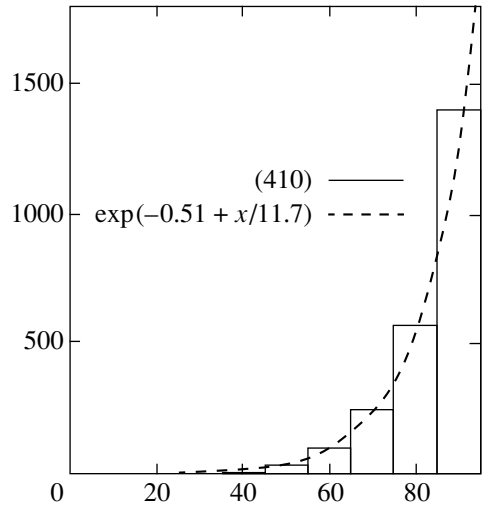


Fig. 4. Acute myocardial infarction (410). Death rate.

effective number N of possible places where the given critical nucleus can arise. The temperature of the body is $T = 273 + 36.6 \approx 310$ K. The comparison gives a reasonable estimate of barriers $U \sim (1.2–1.4) \times 10^4$ K + $T\ln N$, or $U \sim 1.1–1.3$ eV if $N \sim 1$, and only $U \sim 3$ eV even if N is equal to the total amount of the molecules in the human body (this effective number is of course unrealistic). It is noteworthy to say here that the estimated barrier values are comparable to those typically encountered in condensed matter physics for the processes mentioned in the previous pages.

In order to estimate the age change of barriers, one does not need to know the preexponential factor in the expression (2). Typical 90 years increasing factor of D_i is exp(8). It corresponds to the diminishing of barriers $\delta U \sim 8T$, this value is also reasonable $\delta U \sim 0.2$ eV $\ll U$. Two parameters $\delta U/U \ll 1$ and $U/T \gg 1$ are the main parameters of the theory.

In the framework of the presented picture, the small difference in barriers of the order of 0.02 eV for males and females corresponds to the known ratio $D_m/D_f \sim 2$, and can be directly related to the difference 1/23 in chromosome compositions. The variation of the parameters with time and specific groups of population, countries, races, etc., should be of the same order of magnitude. The situation is similar to the usual one in condensed matter physics, where the experimental data are observably dependent on the sample preparation conditions.

Note, that there is no real contradiction between the presented idea and the fact that there are a lot of diseases caused by viruses and bacteria. The age dependence of those diseases can be related to the microorigin of the destruction of the immune system.

Moreover, I think that the discussed thermal activation mechanism could play a role in the generation of congenital anomalies.

REFERENCES

1. L. D. Landau and E. M. Lifshitz, *Statistical Physics* (Nauka, Moscow, 1976; Pergamon, New York, 1980).
2. A. A. Grifith, Philos. Trans. R. Soc. London, Ser. A **221**, 163 (1920).
3. B. Gompertz, Philos. Trans. R. Soc. London, Ser. A **115**, 513 (1825).
4. T. Lindahl, Nature **362**, 709 (1993).
5. E. Shredinger, *What is Life?* (Piper, Munchen, 1989).
6. E. M. Lifshitz and L. P. Pitaevskii, *Physical Kinetics* (Nauka, Moscow, 1979; Pergamon, New York, 1981).
7. D. H. Hoyert *et al.*, Natl. Vital Stat. Rep. **47**, 19 (1999); www.cdc.gov/nchswww.

Journal of Experimental and Theoretical Physics, Vol. 90, No. 5, 2000, pp. 908–918.
From Zhurnal Éksperimental'noĭ i Teoreticheskoĭ Fiziki, Vol. 117, No. 5, 2000, pp. 1043–1053.

MISCELLANEOUS

Toward the Theory of Electron and Positron States in Dielectric Clusters¶

V. V. Pogosov*, W. V. Pogosov, and D. P. Kotlyarov
Zaporozhye State Technical University, Zaporozhye, Ukraine
**e-mail: vpogosov@zstu.zaporizhzhe.ua*
Received July 2, 1999

Abstract—Analytical expressions for the binding energy of electrons and positrons in dielectric clusters, analyzed in this work, neglect the elastic effects. Therefore, we present the density-functional theory for neutral liquid clusters that experience the spontaneous deformation. Using the 1/R-expansion, R being the cluster radius, the exact analytical expressions for the size corrections to the chemical potential, surface tension, and atomic density are derived from the condition of mechanical equilibrium. The problem of calculating these corrections is reduced to calculating the quantities for a liquid with a flat surface. The size compression and tension of density occur in the $1/R$ and $1/R^2$ orders respectively. The sizes of charged rigid and elastic critical clusters, for which the electron or positron binding energy is close to zero, are calculated for Xe_N^-, Kr_N^-, Ar_N^-, Ne_N^+, He_N^+. The calculations show significant contribution of self-compression to the binding energy of the excess electron in contrast to the positron. © *2000 MAIK "Nauka/Interperiodica".*

1. INTRODUCTION

Excess charged particles and polarization interactions are of great importance in physical chemistry and biology. The interaction of electrons with atoms, which posses large polarizabilities, exhibits an attractive character. That is why their localization is possible in clusters [1–3]. Electronic clusters (or negative cluster's ions) were discovered experimentally in a dense xenon [4, 5]. In helium, which has a very small atomic polarizability, localization of electrons happens in a void bubble [6]. Recently, the electronic bubbles were observed even in the helium microdroplets [7]. The interaction of the positrons with atoms, owing to the absence of the exchange interaction, always demonstrates their attractive character. Positron clusters were discovered in all dense gases of rare atoms [8, 9]. The temperatures of clusterization and the "optimal" sizes of clusters were estimated in [2]. Such clusters contain hundreds of atoms, and their density is close to that of a liquid cluster. On the other hand, the mass-spectrometry measurements allowed to discover the existence of xenon clusters which contain near dozen of atoms and are charged by only one electron [10]. They have a noticeable lifetime and are called "critical" clusters. The size dependence of the electron affinity and critical size of xenon solid clusters were examined by a continuum model [11], and by taking full account for the atomic structure [12]. In this work, we propose an improvement of earlier theories.

The main purpose of this paper is to discuss a true asymptotic for binding energy of quantum particle localized in a large dielectric cluster. Subsequently, we point out the importance of elastic effects in the determination of the cluster's energetics. We develop a formal density-functional theory for finite classical systems in order to account for the self-deformation of the clusters. For smallest clusters, the theory based on the continuum model retains the simplicity of the method developed for rigid clusters. Furthermore, critical sizes of single-charged elastic clusters are calculated.

2. LARGE RIGID CLUSTER

The quantum particles localized in large clusters are almost free. Their energy spectra are determined by the character of scattering on cluster atoms and depend upon the atomic density. In [13], the following expression for electron binding energy was discussed,

$$E_b = E_b^0 - \frac{\hbar^2\pi^2}{2m_{\text{eff}}R^2}, \tag{1}$$

where E_b^0 is the standard binding energy component that contains the Born correction,

$$E_b^0 = -V_0 - \frac{e^2}{2R}\frac{\varepsilon-1}{\varepsilon}, \tag{2}$$

where $V_0 < 0$ is the ground state energy of the electron in a extended dielectric (Ar, Er, Xe); $R = N^{1/3}\bar{r}$ is the cluster radius; N is the number of its atoms; and $\bar{r}$ is the

¶This article was submitted by the authors in English.

average distance between the atoms of density $\bar{\rho} = (4\pi\bar{r}^3/3)^{-1}$. The second term in (1) is the kinetic energy of the electron localized inside cluster and m_{eff} is the effective mass. The radius of critical clusters R* may be crudely estimated [11] directly from the condition $E_b^0(R) = 0$.

An alternative asymptotic expression for the binding energy of a charged particle has been derived in the effective medium approach and pseudopotential theory of scattering [15],

$$E_b = E_b^0 - \frac{\hbar^2\pi^2}{2mR^2}(1 - C\xi), \tag{3}$$

and

$$E_b^0 = T + \frac{3}{2}\frac{\alpha e^2}{\bar{r}^3\sigma}\left(1 - \frac{\sigma}{R}\right)f, \tag{4}$$

where $T = \mp\hbar^2 q_0^2/2m$. The sum of the first two terms in (4) gives $(-V_0)$ and the last term gives $-e^2(\varepsilon - 1)/(2R\varepsilon)$ The dielectric constant $\varepsilon = 1 + 3\alpha/(\bar{r}^3 - \alpha)$ was taken in the Clausius–Mossotti approximation. The second term in (4) gives the shift of the energy due to the mean polarization of infinite liquid. The minus and plus sign appearing in T correspond to $L > 0$ and $L < 0$, respectively, where $L \equiv L(\bar{r})$ is the scattering length of a quantum particle in dielectric. σ is the atomic polarizability; α is the parameter of the Lennard–Jones potential; $f = (1 + 2\alpha\bar{r}^3)^{-1}$ is the Lorentz local-field correction; $C \approx 2.86$; and $\xi = L/\bar{r}$ is the small parameter. A simple form of step function was used for the pair-correlation function for atoms,

$$g(r) = \theta(\sigma - r), \tag{5}$$

where σ corresponds to the mean closest interatomic distance in the cluster.[1] The solution of the Schrödinger equation in the Wigner–Seitz cell for the two principally different regimes of scattering [17], gives the following equation for q_0:

$$\begin{aligned} \tan[q_0\bar{r} + \delta_0(q_0)] &= q_0\bar{r}, \\ \delta_0 &= -Lq_0 + O(q_0^3), \quad \text{for } L > 0, \\ \tanh[q_0\bar{r} + \mathrm{Im}\,\delta_0(iq_0)] &= q_0\bar{r}, \\ \delta_0 &= -iLq_0 + O(-iq_0^3), \quad \text{for } L < 0. \end{aligned} \tag{6}$$

Here $\delta_0(x)$ is the phase shift of the charged particle's s-wave scattered in cellular infinite medium [18].[2]

In principle, both expressions (1) and (3) follow from the Bardeen theory [19] for the extended system. They give, however, different size-dependence of binding energy. In this section, our consideration is restricted to a special case of large clusters when both the electron mean free path in extended liquid (which is of the order of hundreds of bohrs) and the electron wave length in the cluster are close to cluster radius. In this case, the binding energy should be calculated from equation (3). Calculations using equation (1), assuming the input of effective mass, are not correct because the effective mass can be correctly calculated and entered to (1) only if the mean free path is much smaller than the cluster radius.

We describe the fluid number density ρ as of undisturbed fluid of uniform density up to spherical boundary, i.e., as for a rigid cluster with zeroth compressibility, and we put $\rho(r) = \bar{\rho}\,\theta(r - R)$. The values of V_0 and m_{eff} for electrons and positrons in considered media were measured in a wide range of densities [20–26]. We calculated these values (Table 1), taking into account the simple correlation function given by (5). The input experimental values of $L(\bar{r})$ for excess electrons are taken from [4, 25, 26]. The input calculated values of scattering lengths for positrons are used from [17].

In Fig. 1, the binding energies $E_b(N)$ calculated from (1) and (3) are shown for $Xe_N^- \equiv Xe_N + e^-$, $Kr_N^- \equiv Kr_N + e^-$, and $Ar_N^+ \equiv Ar_N + e^+$ clusters of densities corresponding to a liquid state in the triple point. As is seen from Fig. 1, the two curves for $E_b(N)$ differ considerably. The difference in the curves for Xe_N^- and Ar_N^+ originates from the effective masses m_{eff} and from the sign of scattering length L. Equation (3) predicts a smaller size "critical" cluster Xe_N^- and Kr_N^- (which correspond to the condition $E_b(N^*) = 0$). These results suggest that equation (3) is superior over (1) because it predicts smaller sizes of critical electronic clusters. In Section 4, we show that these sizes are determined by the availability of surface states. The latter effect was ignored in [14, 15].

Finally, putting aside the problem of availability of surface states, we should add that formula (3) is formally correct but exclusively for electronic clusters, $L > 0$, with $N > 100$. This is confirmed by numerical solution of eigenvalue problem for the potential well of

[1] In [16, 17] the radial distribution function $g(r)$ was used, which reflected the real structure of simple liquids in coefficients I_0, I_2, I_4 appearing in the expressions for the phase shifts scattering waves, V_0 and m_{eff}. The present version of g(r) corresponds to $I_0 = I_2 = I_4 = 1$.

[2] It should be noted that $\hbar^2 q_0^2/2m$ appearing in (4) is not the kinetic energy of the particle in the cell, as it seems to be. This term describes only scattering inside a cell. The wave number q_0 is obtained from (6) using the boundary conditions by means of the scattering length which allow to account entirely for the repulsion and partially for the attraction, i.e., the scattering at the polarization potential profile inside cell.

Table 1. The calculated input values of V_0 and m_{eff}, and used for the estimation of the binding energy $E_b(N)$. The dates are taken from [4, 17, 25, 26, 40, 41], a_0 is the Bohr radius

	T, K	$\bar{r}_0$, a_0	$L(\bar{r}_0)$, a_0	V_0, eV	$dV_0/d\rho$, eV a_0^3	m_{eff}/m	γ_0/B_0^+, a_0	ζ
Xe_N^-	161.4	4.855	0.70	–0.680	+1140	0.664	0.63	1.10
Kr_N^-	115.7	4.544	0.60	–0.454	+676	0.678	0.57	1.02
Ar_N^-	83.8	4.225	1.10	–0.201	+1122	0.711	0.49	0.97
Ar_N^+	83.8	4.225	–0.63	–0.986	–303	1.203	0.49	0.97
Ne_N^+	24.8	3.531	–0.027	–0.446	–17.8	1.099	0.46	0.93
He_N^+	4.2	4.404	–0.29	–0.259	–45.0	1.05	0.44	0.95

radius R and of depth E_b^0. The point is that the sum of exact kinetic energy and the last term that contains weight coefficient $C \equiv C(R)$ (see [15]) shows the size dependence similar to that for large clusters in equation (3).

However, in general, the cluster may be compressed under the action of surface tension and tensed by a localized quantum particle. We can neglect the pressure P_q of localized charge $q = \pm e$ in two cases: for a large and "bulk" cluster (P_q has an order R^{-4}, that is much less then the Laplace pressure) and for a critical cluster ($E_b \longrightarrow 0$, $P_q \longrightarrow 0$). In these cases, one can take into consideration, in analytical form, the effect of self-compression of cluster under the action of surface forces upon the energetics of a bound quantum particle. For intermediate sizes of clusters, the self-consistent solution of the problem of a particle localized in the liquid cluster is required.

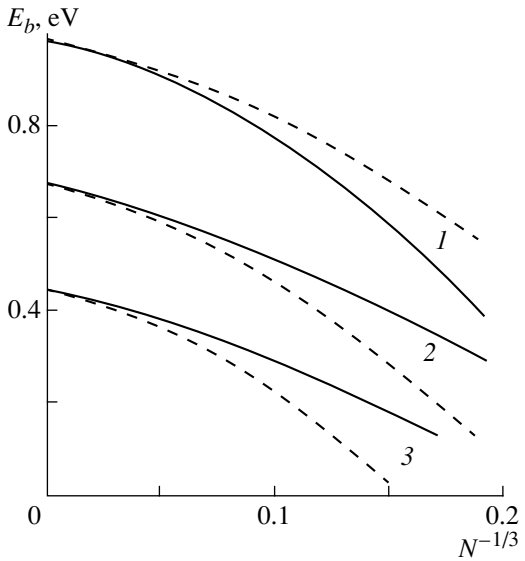


Fig. 1. The binding energy $E_b(N)$ calculated from equations (1) and (3) (dashed and solid line, respectively) for: (*1*) Ar_N^+; (*2*) Xe_N^-; (*3*) Kr_N^-.

The analytical sum-rule approach, developed for neutral metallic clusters [27, 28], describes the influence of self-compression upon ionization potential only in terms proportional to the first order in $1/R$. As will be shown, for a dielectric cluster, this approach is more progressive, and the desired corrections proportional to $1/R^2$ are obtained. In the following section, we briefly present the density-functional theory of a self-deformed cluster.

3. DENSITY-FUNCTIONAL THEORY

Consider a classical, dense vapor at temperature T, and of chemical potential μ, in a box of volume V. The free energy of a system of cluster-vapor, $F \equiv F[\rho(r, R)]$ is a functional of the inhomogeneous atomic concentration $\rho(r, R)$, R is the cluster radius. In the framework of the square-gradient theory, the free energy can be written in the form

$$F = \int d^3r\,(f + g(\nabla\rho)^2), \tag{7}$$

where $f \equiv f[\rho(\mathbf{r}, R)]$ is the energy density of the quasi-homogeneous part of the functional, $g \equiv g[\rho(\mathbf{r}, R)]$ gives the first inhomogeneity term represented by the first gradient term.

The grand free energy is found by minimizing the functional

$$\Omega_V[\rho] = F[\rho] - \mu\int d^3r\,\rho(\mathbf{r}) + \int d^3r\,\frac{\hbar^2}{2m}|\nabla\psi|^2 + \int d^3r\int d^3r'|\psi(\mathbf{r})|^2 v(\mathbf{r}-\mathbf{r}')\rho(r') \tag{8}$$

with respect to variation of $\rho(r)$ and $\psi(r)$ under the conditions

$$\int d^3r|\psi(\mathbf{r})|^2 = 1, \quad \int d^3r\rho(\mathbf{r}) = N_0.$$

Here, $V(\mathbf{r})$ is the electron/positron-atom potential, and N_0 denotes the total number of atoms in a box. By varying $\Omega_V[\rho]$ with respect to $\psi(\mathbf{r})$ while using a Lagrange multiplier, one finds the following Schrödinger equation for $\psi(\mathbf{r})$:

$$-\frac{\hbar^2}{2m}\Delta\psi(\mathbf{r}) + V(\mathbf{r})\psi(\mathbf{r}) = E\psi(\mathbf{r}), \tag{9}$$

where

$$V(\mathbf{r}) = \int d^3r' v(\mathbf{r}-\mathbf{r}')\rho(\mathbf{r}') \tag{10}$$

is the mean potential field, produced by atoms. For a given $\rho(\mathbf{r})$ we want the lowest-energy solution to equation (9). Let us denote the energy in this state by $E[\rho]$. For the equilibrium profile $\rho(\mathbf{r}, R)$, the functional $\Omega_V[\rho(\mathbf{r}, R)] = E + F - \mu N_0$ has a minimum and equals the Gibbs grand potential $\Omega = -PV$, where P is the pressure in a box.

In this paper, we use $V(\mathbf{r})$ in the form of the sum of short range (see equation (4)) and long-range (polarization) components:

$$V(\mathbf{r}) = T\delta(\mathbf{r}) + \int d^3r' V_p(\mathbf{r}-\mathbf{r}')\rho(r'), \tag{11}$$

where $\delta(\mathbf{r})$ is the Dirac δ function. For a dense cluster in the delute vapor the last term in (11) has a standard form of the interaction energy of a point charge with a dielectric sphere [16, 29]. We consider the case of weak perturbation of the atomic distribution $\rho(r)$ by the excess quantum particle (see the above discussion in Section 2). However, the effect of the correction may be estimated after the fact and such an estimate is made in the end of the Section.

3.1. Neutral Elastic Cluster

Using (7), the Euler–Lagrange equation can be written in the form

$$\mu(r, R) = \frac{\delta F[\rho]}{\delta\rho(r, R)} = \frac{\partial f}{\partial\rho} - \frac{\partial g}{\partial\rho}(\nabla\rho)^2 - 2g\left(\nabla\rho + \frac{2}{r}\nabla\rho\right). \tag{12}$$

For the equilibrium concentration profile $\rho(r, R)$ we have $\mu(r, R) \equiv \mu(R)$. By definition, the surface free energy per unit area, γ and surface tension (stress, for a solid) τ [30] are given by

$$\gamma = \frac{1}{A}[F[\rho(r, R)] - F[\rho_0^+]\theta(r-R) - F[\rho_0^-]\theta(R-r)], \tag{13}$$

$$\tau = \gamma + A\frac{d\gamma}{dA}, \tag{14}$$

where $A = 4\pi R^2$ is the area of "equimolecular surface" of cluster, which is defined by the condition

$$4\pi\int_0^\infty dr r^2(\rho(r, R) - \rho_0^+\theta(r-R) - \rho_0^-\theta(R-r)) = 0. \tag{15}$$

Here ρ_0^+ is the atomic concentration in the uniform condensed matter, ρ_0^- the density of uniform vapor beyond the surface, and $\theta(-x)$ the Heaviside step function. In the following, we employ the expansion of $Y \equiv \rho, \mu, \gamma, \tau$ quantities in powers of the inverse radius $1/R$,

$$Y = \sum_{k=0}^{\infty}\frac{Y_k}{R^k}. \tag{16}$$

The zeroth-order terms in (16) are relevant to the system with a planar boundary. Inserting this expansion into (12) and (14), and using the series

$$\frac{1}{r} = \frac{1}{R}\sum_{k=0}^{\infty}(-1)^k\left(\frac{r-R}{R}\right)^k,$$

one can compile the terms having equal powers of $1/R$, getting a set of equations for ρ_k and μ_k. The equations for $k = 0, 1, 2$ have the form

$$\mu_0 = \frac{\partial f_0}{\partial\rho_0} - \frac{\partial g_0}{\partial\rho_0}(\nabla\rho_0)^2 - 2g_0\Delta\rho_0, \tag{17}$$

$$\mu_1 = \frac{\partial^2 f_0}{\partial\rho_0^2}\rho_1 - 2\frac{\partial g_0}{\partial\rho_0}(\nabla\rho_0\nabla\rho_1 + \rho_1\Delta\rho_0) - \frac{\partial^2 g_0}{\partial\rho_0^2}\rho_1(\nabla\rho_0)^2 - 2g_0(\Delta\rho_1 + 2\nabla\rho_0), \tag{18}$$

$$\mu_2 = \frac{1}{2}\frac{\partial^3 f_0}{\partial\rho_0^3}\rho_1^2 + \frac{\partial^2 f_0}{\partial\rho_0^2}\rho_2 + \text{gradient terms}, \tag{19}$$

$$\int_{-\infty}^{\infty} dx(\rho_0 - \rho_0^+\theta(-x) - \rho_0^-\theta(x)) = 0, \tag{20}$$

$$\int_{-\infty}^{\infty} dx(\rho_1 - 2x\rho_0^+\theta(-x) - 2x\rho_0^-\theta(x)) = 0, \tag{21}$$

where we have changed the variable $x = r - R$, and we have made use of the limit $R \longrightarrow \infty$, $\rho^+ \equiv \rho(x = -\infty)$, $\rho^- \equiv \rho(x = +\infty)$. For brevity, we use the notation $\nabla = d/dx$ and $\Delta = d^2/dx^2$. The liquid under consideration occupies the half-space $x < 0$, and vapor is for $x > 0$. It

is convenient to introduce the useful definition of the "average over a planar surface"

$$\langle \mu(x) \rangle (\rho_0^+ - \rho_0^-) = -\int_{-\infty}^{\infty} dx \mu(x) \nabla \rho_0, \quad (22)$$

and "first average over spherical surface"

$$\langle\langle \mu(x) \rangle\rangle (\rho_1^+ - \rho_1^-) = -\int_{-\infty}^{\infty} dx \mu(x) \nabla \rho_1. \quad (23)$$

To transform in equation (13), we have to carry out the following procedure. Multiply equation (17) by $\nabla\rho_0(x)$ and then express the result in the form

$$\nabla(f_0 - g_0(\nabla\rho_0)^2 - \mu_0\rho_0) = 0 \quad (24)$$

which represents a microscopic analogue of the condition of mechanical equilibrium for cluster-vapor system. Next, integrate equation (24) in the limits $(-\infty, x)$ to yield

$$f_0(x) = f_0^+ + g_0(\nabla\rho_0)^2 + \mu_0(x)\rho_0(x) - \mu_0^+\rho_0^+, \quad (25)$$

where $f_0^+ \equiv f(\rho_0^+)$. This makes it possible to separate γ_0 and γ_1 in the expression (13) for $\gamma(R)$. Using equations (20)–(22), after cumbersome transformations, one gets the analogue of results obtained earlier, and in another form, in the framework of the Van der Waals theory (see [31–33], where g_0 = constant was used)

$$\gamma_0 = 2\int_{-\infty}^{\infty} dx g_0(\nabla\rho_0)^2, \quad (26)$$

$$\gamma_1 = 4\int_{-\infty}^{\infty} dx x g_0(\nabla\rho_0)^2. \quad (27)$$

A similar expression was derived earlier [31] in the two-component plasma model and stabilized jellium for a self-compressed metal cluster.

We conclude this section by deriving necessary exact sum-rules. Using equations (17), (18), (22), (23), and (26), (27) one can obtain the following expressions,

$$\mu_1^\pm = \rho_1^\pm \frac{\partial^2 f_0^\pm}{\partial \rho_0^{\pm 2}}, \quad (28)$$

$$\langle \mu_1 \rangle (\rho_0^+ - \rho_0^-) = 2\gamma_0, \quad (29)$$

and

$$\mu_2^\pm = \frac{1}{2}\rho_1^{\pm 2}\frac{\partial^3 f_0^\pm}{\partial \rho_0^{\pm 3}} + \rho_2^\pm \frac{\partial^2 f_0^\pm}{\partial \rho_0^{\pm 2}}, \quad (30)$$

$$\langle \mu_2 \rangle (\rho_0^+ - \rho_0^-) + \frac{1}{2}\langle\langle \mu_1 \rangle\rangle \left(\frac{\rho_0^{+2}\mu_1^+}{B_0^+} - \frac{\rho_0^{-2}\mu_1^-}{B_0^-} \right) = 2\gamma_1, \quad (31)$$

where $\mu^+ \equiv \mu(x = -\infty)$, $\mu^- \equiv \mu(x = +\infty)$, and $B_0^\pm = \rho_0^{\pm 2}\partial^2 f_0^\pm / \partial \rho_0^{\pm 2} \equiv \rho_0^{\pm 2} f_0^{\pm\prime\prime}$ is the bulk modulus (inverse compressibility) of liquid and vapor, respectively. In particular, equation (29) defines the size correction to the "atomic work function" or cohesive energy $\varepsilon_{coh}(R) = \varepsilon_{coh0} + \varepsilon_{coh1}/R$, where $\varepsilon_{coh1} = -2\gamma_0/(\rho_0^+ - \rho_0^-)$.

The equilibrium conditions, $\mu_{1,2}^+ = \langle \mu_{1,2} \rangle = \mu_{1,2}^- = \langle\langle \mu_{1,2} \rangle\rangle$, lead to cancellation of the second term in (31) and, after trivial algebra, we derive the desired equalities

$$\rho_1^+ = 2\gamma_0 \frac{\rho_0^{+2}}{B_0^+(\rho_0^+ - \rho_0^-)}, \quad (32)$$

$$\rho_2^+ = \rho_1^+(\delta - \chi), \quad (33)$$

which will be used in further calculations. The "size" coefficient $\delta = \gamma_1/\gamma_0$ is defined by the dependence $\gamma(R) = \gamma_0(1 + \delta(R))$, and $\chi = \rho_1^+ f_0^{+\prime\prime}/2 f_0^{+\prime\prime}$. The quantities ρ_1, ρ_2 appearing in (32) and (33) can be calculated by solving the problem for a flat surface. It should be noted that for liquid rare gases the value γ_0/B_0^+ is close to one half of the Bohr radius $a_0 = \hbar^2/me^2$, thus giving some "fundamental" length by analogy with the liquid metals [34].

Expression (32) means that atomic concentration in the bulk of the cluster increases by ρ_1^+/R compared to the ρ_0^+ case where $R \longrightarrow \infty$. Thus, self-compression is a result of surface curvature that creates extra pressure, $2\gamma_0/R$, in comparison to the planar case. It will be demonstrated below that the second correction, ρ_2^+/R^2, has a negative sign. This points to the size self-tension that appears in the term of order $1/R^2$.

The sign of coefficient δ in (32) may be derived intuitively in the following way. The response of the cluster to the decreasing of its size corresponds to the well-known Le Chatelier principle. Taking into account the size dependence of surface energy, the extra pressure inside the cluster is $2\gamma(R)/R$, where $\gamma(R) < \gamma_0$. Consequently, the decreasing of $\gamma(R)$in comparison to γ_0, counteracts the increase in capillary pressure, caused by the decreasing of cluster size. In order to make a connection to [10], we will restrict our consideration to

the cluster-vacuum system. It means that we need to set $\rho_0^- = 0$ in (32) and to make a change $\gamma(R) \longrightarrow \tau(R)$ [35]. Then, by definition of surface tension (14), we have

$$\tau(R) = \tau_0\left(1 + \frac{\delta}{2R}\right). \tag{34}$$

Here, for simplicity, we assume that $\tau_0 = \gamma_0$ (see discussion in [36–38]). Following equation (34), the correction ρ_2^+/R^2, defined by (33), decreases by a factor of 2.

Let us discuss the influence of localized quantum particle upon the atomic density in a cluster. In general, the corresponding component of pressure is defined by the two last terms of equation (8). In the considered system, the intrinsic pressure has a form

$$P = \frac{2\tau(R)}{R} + P_q[\psi(r), \rho(r)]. \tag{35}$$

The pressure P_q stipulated by the excess particle is defined by derivative dE/dV_{cl} over the volume of cluster V_{cl}, $E = -E_b$. For a large cluster,

$$\int_0^R dr 4\pi r^2 |\psi(r)|^2 \longrightarrow 1,$$

E_b corresponds to equations (2), (3), and this component of pressure can be written as follows

$$P_q \longrightarrow \left\{-\frac{e^2}{8\pi R^4}\frac{\varepsilon - 1}{\varepsilon} - \frac{\hbar^2\pi}{4mR^5}(1 - C\xi)\right\}. \tag{36}$$

Thus, we obtain an analogue of the Tompson equation [39]. The "surplus" pressure P_q of the quantum particle introduces the additional correction to atomic density $\Delta\rho_q = \rho_0 P_q/B_0^+$ (see Section 1). Simple calculations demonstrate the weak effect of the tension of Xe clusters, induced by a particle in the range for $N > 100$. With decreasing N, its effect becomes somewhat noticeable. However, for the smallest, i.e., the near-critical clusters, this physical picture become simpler because the occupation probability for the electron (or positron) is close to zero, and the pressure term P_q disappears.

3.2. Small Clusters

Consider the ground state of the particle localized in a small cluster. Using (8), let us write the wave equation for the radial wave function

$$\frac{d^2u(r)}{dr^2} - \frac{2m}{\hbar^2}[E_b + V(r)]u(r) = 0, \tag{37}$$

where $u(r) = r\psi(r)$, $\psi(r)$ is the particle wave function and the potential $V(r) \equiv V(R, r)$. The ground-state wave function is symmetrical about the center of the cluster, so that the boundary conditions $u(0) = 0$, and $u(\infty) = 0$, have to be satisfied.

With cluster size decreasing, the near-surface region occupies the considerable part of its volume and the electron mainly can be found to be outside the formal cluster boundary at the polarization tail of the potential $V(r > R)$. It is stipulated by the electrostatic component of $V(r)$, which can be calculated exactly as the interaction energy of a point charge $e^{\pm}$ with the dielectric sphere of radius R. The behavior of the electrostatic component of $V(r)$ at the boundary has a nonphysical singularity [29]. Therefore, the singularity at $r = R$ is removed by a usual cut-off procedure and replaced by a constant potential. The discontinuities of $V(r)$ are an artifact of this model and have only a small influence on $V(r)$ and the binding energy [11]. On the other hand, the short-range component of $V(r)$ can be calculated only when $r \le R$ [15] (see equation (11)). Thus, we assume that the one-particle "pseudopotential" in equation (37) has a form similar to the Heine-Abarenkov electron-ion pseudopotential for a metal, i.e., it can be written as follows

$$V(r) = \begin{cases} -E_b^0, & r < R, \\ V_p(R + \bar{r}/2), & R < r < R + \bar{r}/2, \\ V_p(r), & r > R + \bar{r}/2, \end{cases} \tag{38}$$

where for the polarization tail $V_p(r)$, the cut-off at $r = R + \bar{r}/2$ is used, and

$$V_p(r) = -\frac{e^2}{2}\frac{(\varepsilon - 1)R}{(\varepsilon + 1)r^2}\left[\frac{R^2}{r^2 - R^2} - \frac{1}{\varepsilon + 1}\left[\ln\left(\frac{r^2}{r^2 - R^2}\right) - \sum_{k=1}^{\infty}\frac{1}{k(k\varepsilon + k + 1)}\left(\frac{R}{r}\right)^{2k}\right]\right], \tag{39}$$

$$r > R.$$

The pseudopotential (38) has the right asymptotics: $V(r) \longrightarrow V_0$ for $\sigma/R \longrightarrow 0$, and $V(r) \longrightarrow -N\alpha e^2/2r^4$ for $r/R \longrightarrow \infty$. The binding energy E_b results from a competition of kinetic and polarization energies, and for a critical cluster is close to zero. Thus, solving equation (37), we find $V(R^*, r)$, and consequently $N^* = (R^*/\bar{r})^3$.

The potential in the center of a large cluster can be assumed as the nearest to the bottom V_0 of the conduction-band in the infinite liquids. For solid state, V_0 is close to zero (especially for Ar) [20], and by taking into account the Born size correction and self-compression, it becomes even positive (more incapable to retain an electron). On the other hand, the polarization tail $V_p(r)$, in the region $r > R$, depends rather weakly upon the cluster state (liquid or solid). Therefore it is clear, that when the first bound state appears, the electron will probably be localized outside the cluster, in a near-surface state.

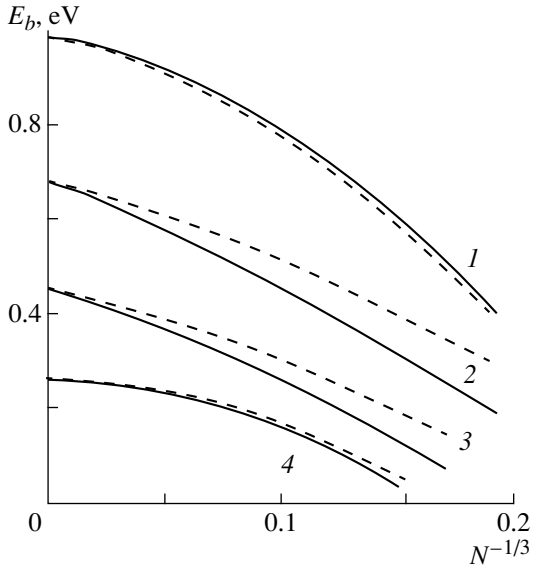


Fig. 2. The binding energy $E_b(N)$ for electron and positron in rigid and elastic clusters (dashed and solid line, respectively), calculated from equation (3): (*1*) Ar_N^+; (*2*) Xe_N^-; (*3*) Kr_N^-; (*4*) He_N^+. For Xe_N^- the experiment gives $N^* = 5–8$ [10].

For the positron, in contrast to the electron, it is more probable that it will be situated inside the cluster. In a large cluster Ar_N, the value V_0 is about –1 eV, that is, in the center of the cluster, the positron feels a deep potential well. The positron localizes on much smaller clusters of Ar than the electron. This is conditioned by the comparative prevalence of attraction over repulsion in the positron-atom interaction.

4. CALCULATIONS

4.1. Large Clusters

First, let us define $\delta = c_1 + c_2$ for the calculation of $\rho_2^+ = (1/2)\rho_1^+(\delta - \chi)$ (see equation (33) and comment below equation (31)). From the semi-empirical rule [38], derived from the vacancy formation energy and the cohesive energy results $c_1 = +0.5\bar{r}_0$. The re-definition of "equimolecular surface" for an icosahedral cluster [36] gives $c_2 = -1.32\bar{r}_0$ and thus $\delta = -0.82\bar{r}_0$. The calculation of the third derivative of free energy with respect to density is a difficult problem. On the other hand, the third derivative in (33) can be expressed by the first derivative of the (B_0^+/ρ_0^{+2}) with respect to ρ_0^+. Let us use the well-known sum-rule for compressibility

$$S_{k=0} = \frac{\rho_0^+ k_B T}{B_0^-}, \qquad (40)$$

where $S_{k=0}$ is the structure factor of a liquid for zeroth wave vector and for the constant temperature, T. For bulk properties of liquids, the hard-sphere model gives good results so we employed the Percus–Yevick fluid structure factor, $S_{HS} = (1 - \eta)^4/(1 + 2\eta)^2$. Here, $\eta = \pi d^3\rho_0^+/6$ is the packing fraction and d is the hard-sphere diameter. Then, we have

$$\chi = \frac{\gamma_0}{B_0^+}\frac{\rho_0^+}{y}\frac{\partial y}{\partial \rho_0^+},$$

where

$$y = \frac{k_B T}{S_{HS}\rho_0^+} = \frac{B_0^+}{\rho_0^{+2}}.$$

Using the experimental magnitudes of $S_{k=0}$ in the triple point [40] we determine d and then χ. This allows rewriting expression (33) with more reasonable accuracy in the following form

$$\rho_2^+ = -\zeta\bar{r}_0\rho_1^+, \qquad (41)$$

where ζ is the constant (see Table 1). Comparing the values of ρ_1^+ and ρ_2^+ one can see that size tension is a noticeable effect on the atomic density corresponding to smallest clusters.

It should be noted that the compression of the cluster leads to the rise/drop of the potential bottom V_0 for the electron/positron and to the growth of its kinetic energy owing to the decrease of the radius. The position of the bottom of the band shows strong dependence on the density of atoms (see [17] and Table 1). In the following, for simplicity, in the calculation of the $E_b^0(N)$ component in (3), for self-deformed clusters, we employed linear approximation

$$V_0(\rho) = V_0(\rho_0) + \frac{dV_0}{d\rho}(\rho - \rho_0),$$

and

$$\rho - \rho_0 = \frac{\rho_1^+}{R_0} + \frac{\rho_2^+}{R_0^2},$$

where, $R_0 = N^{1/3}\bar{r}_0$. To illustrate, Fig. 2 compares the electron binding energy, calculated from equation (3), for elastic and for rigid clusters. The difference is much greater than the energy k_BT of thermal excitation. One can see that the shrinkage of a Xe or Kr cluster leads to a strong positive shift of the electron discrete energy level. This effect was not revealed by the previous calculations for critical solid clusters [11–13]. For the positron in the Ar_N clusters, the self-compression leads to negative shift in energy. The positron in the Ar_N the E_b^0 term grows faster then the kinetic energy, therefore,

Table 2. The number of atoms N^* constituting the electron and positron critical cluster for different rare gases and two different binding potentials. The N^*_{JPC} values were determined from rule (42), which corresponds $E_b = 0$, using square potentials barrier ($I^* = \pi^2/8$). The given values of E_b, N^* and N_T are determined quantum mechanically with the potential given by (38). The values of N_T correspond to $E_b = k_BT$

	Elastic				Rigid			
	N_T	E_b, meV	N^*	N^*_{JPC}	N_T	E_b, meV	N^*	N^*_{JPC}
Xe^-_N	9	0.37	5	6–7	7	0.008	4	4–5
Kr^-_N	17	0.019	9	13–14	14	0.13	8	9–10
Ar^-_N	52	0.00002	24	78–79	32	0.27	19	28–29
Ar^+_N	6	4.33	5	4	6	3.12	5	4
Ne^+_N	23	0.07	20	18	23	0.02	19	17
He^+_N	22	0.0005	20	19	22	0.02	20	18–19

E_b^0 is bigger for the self-compressed cluster than for rigid one. For He_N, this correlation breaks down. This is also reflected in the results for the critical positron clusters presented in Table 2.

4.2. Critical Clusters

The critical size of cluster, corresponding to the number of atoms N^*, may be semi-quantitatively calculated from the Jost-Pais-Calogero (JPC) rule, i.e., from the condition for the appearance of the first bound state in the potential $V(r) < 0$, which is given by

$$\frac{m}{\hbar^2}\left|\int_0^\infty dr\,V(r)\right| = I^*. \qquad (42)$$

For different potentials, usually employed in the nuclear physics, the value of I* changes from $\pi^2/8$ to 1.6 [42]. Solving equation (42) with respect to R^*, we can calculate $N^* = (R^*/\bar{r})^3$. It should be remembered that condition (42) was formulated for zeroth binding energy of the captured particle.

As a rough estimation, we have calculated R^* and N^* by (42) using the square potential well. The results for rigid as well as for elastic clusters are presented in Table 2. According to this simple estimation, the electronic stabilization must be observed for $N^*_{JPC} > 4$, 9 and 28 atoms for rigid and for $N^*_{JPC} > 6$, 13 and 78 atoms for elastic clusters of Xe, Kr, and Ar, respectively.

As mentioned above, the absence of the exchange (repulsive) interaction creates more favorable conditions for the localization of the positron. Positron in critical Ar_N clusters feels a deep potential well. Our estimations demonstrate the stabilization of positively charged clusters for $N^*_{JPC} > 4$, 18 and 19 for Ar, Ne and He, respectively. In these materials, a small value of the derivative, $dV_0/d\rho$, causes the self-compression not to significantly influence the binding energy of the antiparticle. It is interesting to note the different influence of this effect on Ar as compared to Ne and He. Self-compressing leads to a positive shift of the positron energy level in Ne_N and to a negative one for Ne_N and He_N. This is determined by the competition between the size dependences of the bottom of the potential well, the polarization tail and the kinetic energy of antiparticle. In critical Ne_N and He_N clusters, the polarizability tail is very small and the positron encounters a nearly square potential well. Thus, our estimation of N^* based on equation (42) and the square potential well is close to realistic values (Table 2).

In a second step, we determine the sizes of critical clusters N* by the numerical solution of equation (37). To simplify the calculations, we assume that $V_p(r) = 0$ for $r > R + a$ [11]. Putting $a = 7R$, which is a good approximation because $|V_p(R + a)| < 0.1$ meV, and owing to the fact that wave function in the region $r > R + a$ has a purely exponential form, we can replace the boundary condition outside the cluster from $r = \infty$ to $r = R + a$. The new boundary condition put at $r = R + a$ is

$$\frac{d}{dr}\ln u(r) = -\sqrt{\frac{2mE_b}{\hbar^2}}.$$

We determine the critical N^* by calculating the least positive value of the binding energy. The results of calculations for N^* and $E_b(N^*)$are presented in Table 2. The actual forms of the pseudopotential (38) and the density $r^2|\psi(r)|^2$ for electron in Xe and Ar and positron in Ar critical clusters are plotted in Figs. 3.

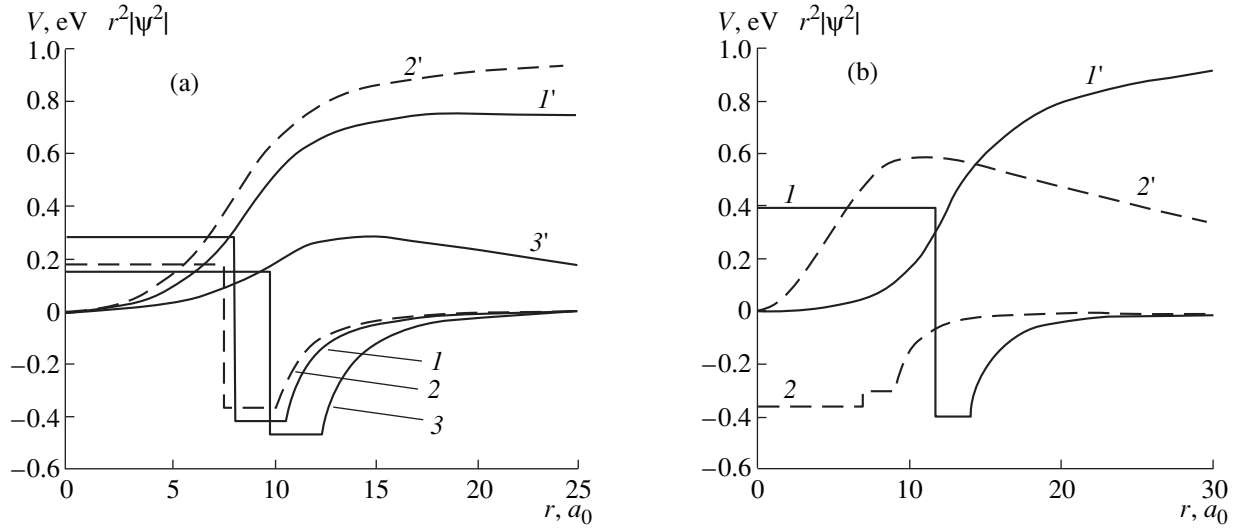


Fig. 3. Pseudopotential $V(r)$ (curves *1*, *2*, *3*) and radial density distribution $r^2|\psi(r)|^2$ (arbitrary units, curves *1*', *2*', *3*') for: (a) critical elastic Xe_5^- (solid lines *1*, *1*'), critical rigid Xe_4^- (dashed lines *2*, *2*') and "thermal" cluster Xe_9^- (solid lines *3*, *3*'); (b) critical elastic Ar_{24}^- (solid lines *1*, *1*') and critical elastic Ar_5^+ (dashed lines *2*, *2*').

It is interesting to compare the obtained values of N^* with the ones calculated from (42). As one can surmise, for electronic clusters, $N^* < N^*_{JPC}$. It is stipulated by the fact that condition (42) was derived for $V(r) < 0$. However, in the electronic clusters $V(r) > 0$ for $r < R$ and $V(r) < 0$ for $r > R$; i.e., in the interior of the cluster, an electron encounters not potential well but a barrier (see Figs. 3). Therefore, the attractive and repulsive part of $V(r)$ compensate each other in equation (42). To fulfill the equality in equation (42) the "negative" region of $V(r)$ has to be increased, which is equivalent to a fictitious increase of N^*. The results of the numerical solution of equation (37) confirm the role of self-deformation. The magnitudes of N^* for the rigid and elastic clusters differ on 30%. This is caused by the significant magnitude of the derivative $dV_0/d\rho$. In spite of that, in the smallest clusters, the size self-compression is neutralized by the size self-tension.

For positron clusters, the values of N^* are higher then N^*_{JPC}. This result is clear too: despite that $V(r) < 0$ for arbitrary r (see Fig. 3), the using $I^* = \pi^2/8$ in (42) corresponds to the use of a square potential well. This is because positron states the account for the size dependence of the polarization tail leads to $N^* > N^*_{JPC}$.

Thus, for the fixed magnitudes of the density ρ_0 and temperature T, the "optimal" clusters are: $Xe^-_{N>5}$, $Kr^-_{N>9}$, $Ar^-_{N>24}$, $Ar^+_{N>24}$, $Ar^+_{N>20}$, $He^+_{N>20}$. Note that these values of N* are underestimated, because they do not take into account the possibility of thermal excitations, i.e., $E_b > k_BT$. In a final step, we calculate the sizes of clusters N_T corresponding to the condition $E_b = k_BT$ (see Table 2).

Analyzing the results for elastic electronic clusters, it is seen that the calculated critical number for Xe^-_N agrees well with the experimental result giving $N^* = 5$–8 [10]. On the other hand, the agreement with another theoretical result for solid Xe and Kr clusters ($N^* = 8$ and 14) is quite good, but not for Ar ($N^* = 46$) [13]. Our results point on the noticeable influence of self-compression, which has not been taken into account before. Self- deformation leads to an increase of N by 30–50%. In view of the latter fact, the accurate prediction of critical number N^* by the authors of [11–13] must be rather considered as fortuitous.

5. CONCLUSIONS

The estimation presented in this work demonstrate that analytical equation (3) points on smaller sizes of the electron critical cluster and thus gives better agreement with measured values. The theory underlying this formula does not use adjustable parameters and is based on the information about electron/positron scattering length, and the Lennard-Jones potential. We have developed a formal density-functional theory of a finite classical system which allows to account for the effect of self-compression, originating from the curvature of the cluster surface, and the effect of self-tension due to the reduction of the cluster's size. The critical sizes of clusters were determined quantum-mechanically by

solving the Schrödinger equation and from the Jost-Pais-Calogero criterion. The effects of self-compression and tension give a significant contribution to the critical sizes of clusters charged by the electron and should be taken into consideration in any comparison of critical cluster's sizes with the measured ones. For positron charged clusters the elastic effects are negligible.

Our model based on the continuum approximation may not be used to describe the localization of electron/positron at a single atom having a large polarizability. The appropriate methods for the solution of this problem have been developed before [43, 44] (see also [45]). However, for the direct application of these methods to the single "ion" $Xe_1 + e^+$ one needs to be aware of the radius of the short-range core of potential.

The behavior of the one-particle potential $V(r)$ of electronic clusters qualitatively resemble that for a positron in a metal with a negative positron work function (Al, Mo, Fe, Ni) [46]. It suggests a possibility of the application of our method to large metallic clusters charged by a positron. The results of this investigation may find application in positron diagnostics in ultradispersed media and possibly in rare-gas atom nanotechnology.

ACKNOWLEDGMENTS

The authors wish to thank A. Kiejna for reading the manuscript. This work was supported in part by International Soros Education Program through the grant APU 072082.

REFERENCES

1. V. B. Shikin, Usp. Fiz. Nauk **121**, 457 (1977) [Sov. Phys. Usp. **20**, 226 (1977)].
2. I. T. Iakubov and A. G. Khrapak, Rep. Prog. Phys. **45**, 697 (1982).
3. J. P. Hernández, Rev. Mod. Phys. **63**, 675 (1991).
4. S. S. Huang and G. R. Freeman, J. Chem. Phys. **68**, 1355 (1978).
5. V. V. Dmitrenko, A. S. Romanyuk, S. I. Sushkov, and Z. M. Uteshev, Zh. Tekh. Fiz. **53**, 2343 (1983) [Sov. Phys. Tech. Phys. **28**, 1440 (1983)].
6. A. Ya. Parshin and S. V. Pereverzev, Zh. Éksp. Teor. Fiz. **101**, 126 (1992) [Sov. Phys. JETP **74**, 68 (1992)].
7. M. Rosenblit and J. Jortner, Phys. Rev. Lett. **75**, 4079 (1995).
8. K. F. Couter and L. O. Roeling, Phys. Rev. Lett. **25**, 328 (1970).
9. M. Nuomisaari, K. Rytsola, and P. Hautojarvi, Phys. Lett. A **112**, 279 (1985).
10. H. Haberland, T. Kolar, and T. Reiners, Phys. Rev. Lett. **63**, 1219 (1989).
11. P. Stampfli and K. H. Bennemann, Phys. Rev. A **38**, 4431 (1988).
12. G. J. Martyna and B. J. Berne, J. Chem. Phys. **90**, 3744 (1989).
13. P. Stampfli, Phys. Rep. **255**, 1 (1995).
14. V. D. Lakhno, Izv. Akad. Nauk., Ser. Fiz. **60**, 69 (1996).
15. I. T. Iakubov and V. V. Pogosov, J. Chem. Phys. **106**, 2306 (1997).
16. V. M. Atrazhev and I. T. Iakubov, J. Chem. Phys. **103**, 9030 (1995).
17. I. T. Iakubov and V. V. Pogosov, Phys. Rev. B **51**, 14941 (1995); **53**, 13362 (1996).
18. V. M. Atrazhev, I. T. Iakubov, and V. V. Pogosov, Phys. Lett. A **204**, 393 (1995).
19. J. Bardeen, J. Chem. Phys. **6**, 367 (1938).
20. A. O. Allen and W. S. Shmidt, Z. Naturforsch. A **37**, 316 (1982).
21. R. Reininger, U. Asaf, I. T. Steinberger, and S. Basak, Phys. Rev. B **28**, 4426 (1983).
22. R. Reininger, U. Asaf, and I. T. Steinberger, Chem. Phys. Lett. **90**, 287 (1982).
23. E. M. Gullikson, A. P. Mills, Jr., and E. E. McRae, Phys. Rev. B **37**, 588 (1988).
24. G. Perluzzo, G. Bader, L.G. Caron, and L. Sanche, Phys. Rev. Lett. **55**, 545 (1985).
25. J. A. Jahnke, L. Meyer, and S. A. Rice, Phys. Rev. A **3**, 734 (1971).
26. F. M. Jacobsen, N. Gee, and G. R. Freeman, Phys. Rev. A **34**, 2329 (1986).
27. A. Kiejna and V. V. Pogosov, J. Phys.: Condens. Matter **8**, 4245 (1996).
28. M. Seidl, J. P. Perdew, M. Brajczewska, and C. Fiolhais, Phys. Rev. B **55**, 13288 (1997).
29. V. V. Batygin and I. N. Toptygin, *Problem in Electrodynamics* (Academic, New York, 1964).
30. R. Shuttleworth, Proc. Phys. Soc. London, Sect. A **63**, 444 (1950).
31. P. A. Fisher and M. Wortis, Phys. Rev. B **29**, 6252 (1984).
32. F. P. Buff, J. Chem. Phys. **23**, 419 (1955).
33. R. Evans, Adv. Phys. **28**, 143 (1979).
34. V. V. Pogosov, Chem. Phys. Lett. **193**, 473 (1992); Vacuum **46**, 455 (1995).
35. E. L. Nagaev, Usp. Fiz. Nauk **162** (9), 49 (1992) [Sov. Phys. Usp. **35**, 747 (1992)].
36. W. Vogelsberger, H. G. Fritsche, and E. Muller, Phys. Status Solidi B **148**, 155 (1988).

37. A. Kiejna and P. Ziesche, Solid State Commun. **88**, 143 (1993).

38. V. V. Pogosov, Solid State Commun. **89**, 1017 (1994).

39. R. Kubo, *Thermodynamics, An Advanced Course with Problems and Solutions* (Horth-Holland, Amsterdam, 1968; Mir, Moscow, 1970).

40. A. V. Chadwick and H. R. Glyde, in *Rare Gas Solids*, Ed. by M. L. Klein and L. A. Venables (Academic, New York, 1977).

41. T. S. Stein and W. E. Kauppila, Adv. At. Mol. Phys. **18**, 53 (1982).

42. 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).

43. B. M. Smirnov, *Negative Ions* (Atomizdat, Moscow, 1978).

44. A. G. Khrapak, Teplofiz. Vys. Temp. **13**, 858 (1975).

45. K. F. Volykhin, A. G. Khrapak, and W. F. Schmidt, Zh. Éksp. Teor. Fiz. **108**, 1642 (1995) [JETP **81**, 901 (1995)].

46. V. V. Pogosov and I. T. Yakubov, Fiz. Tverd. Tela (St. Petersburg) **36**, 2343 (1994) [Phys. Solid State **36**, 1274 (1994)].