Journal of Experimental and Theoretical Physics, Vol. 96, No. 4, 2003, pp. 581–586.
From Zhurnal Éksperimental'noĭ i Teoreticheskoĭ Fiziki, Vol. 123, No. 4, 2003, pp. 659–664.

GRAVITATION, ASTROPHYSICS

Gyroscope Deviation from Geodesic Motion: Quasiresonant Oscillations on a Circular Orbit¶

O. B. Karpov
Moscow State Mining University, Moscow, 119991 Russia
e-mail: oleg_karpov@mail.ru; Karpov@msmu.ru
Received May 31, 2002

Abstract—The general relativistic spin–orbit interaction gives rise to a quasiresonant oscillation of the center of mass of a gyroscope along the orbital normal. The oscillation amplitude appears to be measurable by present-day instruments. The influence of oblateness of the field source is investigated. © *2003 MAIK "Nauka/Interperiodica".*

1. INTRODUCTION

In general relativity, the motion of a spinning test body (gyroscope) is affected by the spin–orbit interaction in two aspects: (1) the influence of the orbital motion on the orientation of the gyroscope's rotation axes, and (2) the influence of the gyroscope's intrinsic momentum (spin) on its orbit. The first is comparatively simple when parallel spin transport is assumed. This is admissible if the deviation from geodesic motion is small. The Fermi–Walker transport along an appointed world line is also not complicated. In a spherically symmetric field, parallel transport along a geodesic leads to a precession of the gyroscope's axes known as the geodetic or de Sitter precession [1]. In the field of a rotating mass, the gyroscope's axes undergo the Schiff precession [2], to be verified in the Gravity Probe B experiment (see [3] for details).

In this work, the second aspect of the spin–orbit interaction is considered. The orbital motion of the gyroscope is a sophisticated problem that has not been fully resolved until now even in the post-Newtonian approximation. There exist several different approaches with different results in the leading approximation (see, e.g., [4–10]). The only covariant general relativistic equations of motion of the spinning test particles are the well-known Papapetrou equations [5]. This set of equations is incomplete and requires supplementary conditions. It is generally accepted that these conditions single out a representative point as the gyroscope's center of mass, but there exist diverse other opinions [9–12]. In addition, the Papapetrou equations or alternative ones are very complicated. Their investigation is usually limited to a general analysis; and examination of the effects is typically restricted by the motion of the gyroscope with a vertical spin, i.e., with the gyroscope's axes orthogonal to the orbital plane [13]. For example, it is known that such a gyroscope moves along a circular orbit with the velocity differing from the one of a body without spin [14]. In [14], the conclusion was drawn that a gyroscope with a horizontal spin leaves the geodesic plane, but an erroneous estimation of this effect was given. The effect is much larger because of the quasiresonant character of the spin–orbit interaction, as was first revealed in [15, 16].

In the present work, the motion of a gyroscope with horizontal spin is investigated and the general relativistic effect of a quasiresonant beating is proposed. Because of a small denominator, the speed of light is cancelled in the oscillation amplitude, and the effect therefore becomes quite sizeable. The obvious physical interpretation of the effect is given. This effect is independent of supplementary conditions and is the same in different approaches [4–10]. Description is significantly simplified by expanding the equations of motion up to linear terms in the displacement from a geodesic. Instead of studying an intricate gyroscopic orbit, a small oscillation is investigated. This oscillation gives sufficient information about the gyroscopic orbit. It is shown that a Newtonian nonsphericity of the field source causes a specific effacing of the quasiresonant beating, leaving the oscillation amplitude measurable.

In what follows, orthonormal bases are used in calculations, Greek indices run from 0 to 3, and Latin indices run from 1 to 3. The signature is (– + + +).

2. THE ESSENCE OF THE EFFECT

The general relativistic spin–orbit acceleration a that causes the center of mass of a gyroscope to deviate from a geodesic is on the order of

$$a \sim \epsilon \frac{S}{\lambda} g, \tag{1}$$

¶ This article was submitted by the author in English.

where

$$\epsilon = \frac{GM}{c^2 r}$$

is the relativistic small parameter, $g = GM/r^2$ is the Newtonian acceleration due to gravity, S is the spin of the gyroscope, λ is its orbital moment, c is the speed of light, M is the source mass, and G is the gravitational constant. The motion of the center of mass of the rotating body significantly depends (in the leading approximation (1)) on the reference frame in which it is obtained. The general expression for the spin–orbit acceleration in the leading post-Newtonian approximation (1) is [7, 17]

$$\mathbf{a} = 3\frac{GM}{mc^2 r^3}[\mathbf{S}\times\mathbf{v} + (2-\sigma)\hat{\mathbf{r}}\cdot(\mathbf{S}\cdot(\hat{\mathbf{r}}\times\mathbf{v})) - (1+\sigma)(\mathbf{v}\cdot\hat{\mathbf{r}})(\mathbf{S}\times\hat{\mathbf{r}})]. \quad (2)$$

The parameter σ numbers the different mass centers: $\sigma = 0$ corresponds to the Dixon [6] and Pirani [18] conditions (the intrinsic mass center), $\sigma = 1$ corresponds to the Corinaldesi–Papapetrou conditions [19] (the center of mass defined in the "rest" frame in which the gyroscope moves with the velocity $\mathbf{v}$), and $\sigma = 1/2$ leads to the results of Fock [4] and of [9, 10]. For a circular orbit of the gyroscope ($\mathbf{v}\cdot\mathbf{r} = 0$) with the gyroscope's axis lying in the orbital plane ($\mathbf{S}\cdot(\mathbf{r}\times\mathbf{v}) = 0$), spin–orbit acceleration (2) is independent of the parameter σ,

$$\mathbf{a} = 3\frac{GM}{mc^2 r^3}\mathbf{S}\times\mathbf{v}. \quad (3)$$

Parallel transport of the spin vector $\mathbf{S}$ means that in the process of revolution, acceleration (3) is directed along the orbital normal $\mathbf{e}_3$ and is periodic in time τ,

$$\mathbf{a} = \mathbf{e}_3\epsilon\frac{S}{\lambda}g\cos(\omega_s\tau+\beta).$$

The frequency ω_s differs from the orbital frequency ω because of the geodetic precession Ω^G,

$$\Delta\omega = \omega - \omega_s = \Omega^G = \frac{3}{2}\epsilon\omega. \quad (4)$$

On the other hand, the frequency of the free tidal oscillation along the orbital normal is equal to the orbital frequency. This leads to an almost resonant beating with modulation frequency (4) and the maximum amplitude

$$A = \frac{a}{2\omega\Delta\omega} = \frac{S}{\lambda}r. \quad (5)$$

We note the cancellation of the speed of light c in amplitude (5) by the small relativistic denominator $\Delta\omega$ given by Eq. (4). During the time $\tau \ll (\Omega^G)^{-1}$, the quasiresonant oscillation is enhanced linearly at a rate of

$$A\Delta\omega = \frac{3}{2}\epsilon\frac{S}{\lambda}v \quad (6)$$

and reaches values measurable with present-day instruments. For example, in the case of a gyroscope with a dimension of 10^{-1} m and an intrinsic rotation period of 10^{-1} s in a near-Earth orbit of $r \approx 7\times10^3$ km, we obtain the values

$$\epsilon \sim 10^{-10}, \quad \frac{S}{\lambda} \sim 10^{-9}, \quad A\Delta\omega \sim 10^{-9}\ \text{cm/day}. \quad (7)$$

Parasitic effects of a nonrelativistic origin are mutually cancelled in the symmetric relative oscillations of two gyroscopes with antiparallel spins.

3. CALCULATION OF THE NET EFFECT

In the post-Newtonian approximation, the static, spherically symmetric gravitational field is described by the tetrad

$$\overset{\circ}{\mathbf{e}}{}^{\mu} = \{(1-\epsilon)cdt, (1+\epsilon)dr, r\sin\theta d\phi, -rd\theta\} \quad (8)$$

which represents the rest observers in the Schwarzschild metric. In this frame, the "electric" part E and the "magnetic" part B of the Riemann tensor $\mathcal{R}$ (see, e.g., [8, 17]),

$$E_{ij} = \mathcal{R}_{i0j0}, \quad 2B_{ij} = \mathcal{R}_{iomn}\varepsilon^{mn}{}_j,$$

are given by

$$\overset{\circ}{E}_{ij} = n^2\,\mathrm{diag}\{-2, 1, 1\}, \quad \overset{\circ}{B}_{ij} = 0, \quad n^2 = GM/r^3. \quad (9)$$

Transition to the orbital frame $\mathbf{e}^{\nu}$ is performed by the boost

$$\overset{\circ}{\mathbf{e}}{}^{\mu} = L^{\mu}_{\nu}\mathbf{e}^{\nu}$$

in the $\overset{\circ}{\mathbf{e}}_2$ direction. The Lorentz matrix L has the standard form. Namely, the components of the 4-velocity of the fiducial orbital motion $\phi = nt = \omega\tau$ are

$$u^{\mu} = L^{\mu}_0 = \gamma\{1, 0, \beta, 0\}, \quad (10)$$

where

$$\gamma = (1-\beta^2)^{-1/2}, \quad \beta = v/c, \quad v = (1+\epsilon)nr, \quad \omega = \gamma v/r = n(1+3\epsilon/2),$$

and τ is the proper time. The $\mathbf{e}_1$ axis is directed along the current radius vector, the $\mathbf{e}_2$ axis is directed along

the orbital motion velocity, and $\mathbf{e}_3$ is orthogonal to the orbital plane,

$$L^1_1 = 1, \quad L^2_2 = \gamma, \quad L^3_3 = 1.$$

The angular velocity vector $\mathbf{\Omega}$ of rotation of the orbital triad

$$\nabla_u \mathbf{e}^i = \Omega^i_k \mathbf{e}^k$$

has only the component

$$\Omega_3 = \Omega_{12} \stackrel{\text{def}}{=} \omega_s = n. \tag{11}$$

The transformation of the magnetic matrix [17]

$$\begin{aligned} B_{ij} = {} & 4\overset{\circ}{B}_{kl} L^{[k}{}_i u^{0]} L^{[l}{}_j u^{0]} - \overset{\circ}{B}_{pq} \varepsilon^p{}_{km} \varepsilon^q{}_{ln} L^k{}_i u^m L^l{}_j u^n \\ & - 4\overset{\circ}{E}_{km} \varepsilon^m{}_{ln} L^{[k}{}_{(i} u^{0]} L^l{}_{j)} u^n \end{aligned} \tag{12}$$

leads to the appearance of the component

$$B_{31} = \beta(\overset{\circ}{E}_{33} - \overset{\circ}{E}_{11}) = 3n^2\beta \tag{13}$$

in the orbital frame. The transformation of the electric matrix is analogous to (12) with the substitution

$$B \longrightarrow E, \quad E \longrightarrow -B$$

(see [17]). The result is

$$\begin{aligned} E_{11} = -2n^2(1 + \epsilon/2), \quad E_{22} = n^2, \\ E_{33} = n^2(1 + 3\epsilon) = \omega^2. \end{aligned} \tag{14}$$

We note that component E_{22} parallel to the boost is invariant and the equality $E_{33} = \omega^2$ is exact.

The equation of motion of the gyroscope's center of mass in the orbital frame is the equation of geodesic deviation with spin–orbit acceleration (2) in its right-hand side,[1]

$$\nabla_u \nabla_u \xi^i + E^i_k \xi^k = a^i, \tag{15}$$

where

$$\nabla_u \nabla_u \boldsymbol{\xi} = \ddot{\boldsymbol{\xi}} + 2\mathbf{\Omega} \times \dot{\boldsymbol{\xi}} + \dot{\mathbf{\Omega}} \times \boldsymbol{\xi} + \mathbf{\Omega} \times (\mathbf{\Omega} \times \boldsymbol{\xi}).$$

The dot denotes the derivative with respect to the proper time τ. In the post-Newtonian approximation, the spin–orbit force applied to the intrinsic center of mass of the rotating body is

$$ma^i = -c^{-1} B^i_k S^k. \tag{16}$$

This formula can be obtained, for example, by the matched asymptotic expansions method [8] or directly from the Papapetrou equations with the supplementary conditions of Pirani or Dixon (see [17]; distinctions between the exact conditions of Pirani and Dixon are also discussed there).

In Eq. (15), E_{ik} is measured on the fiducial geodesic u, but B_{ik} in (16) must be calculated in the frame comoving with the gyroscope center of mass. This "mixing" is admissible in the approximation linear in S (Eq. (1)) and linear in ξ (Eq. (15)) if displacement ξ is induced by the spin–orbit interaction,

$$\xi \sim S, \quad \xi S \sim S^2 \sim \xi^2 = 0.$$

On the same basis, we transport the spin vector along the fiducial geodesic according to Fermi and Walker,

$$\begin{aligned} \nabla_u \mathbf{S} = \dot{\mathbf{S}} + \mathbf{\Omega} \times \mathbf{S} = 0, \\ \dot{S}_1 = \omega_s S_2, \quad \dot{S}_2 = -\omega_s S_1, \quad \dot{S}_3 = 0. \end{aligned} \tag{17}$$

Parallel transport equation (17) describes the known geodetic precession (4),

$$S_1 = S\cos(\omega_s \tau + \beta), \quad S_2 = -S\sin(\omega_s \tau + \beta). \tag{18}$$

For the spin in the fiducial plane ($S_3 = 0$), equations (15) and (16) of the center-of-mass motion become

$$\left.\begin{aligned} \ddot{\xi}_1 - 2\omega_s \dot{\xi}_2 + (E_{11} - \omega_s^2)\xi_1 = 0 \\ \ddot{\xi}_2 + 2\omega_s \dot{\xi}_1 = 0 \end{aligned}\right\}, \tag{19}$$

$$\ddot{\xi}_3 + E_{33}\xi_3 = 3g\epsilon \frac{S_1}{\lambda}. \tag{20}$$

Equations (19) describe the free oscillation with the frequency

$$\omega' = \sqrt{E_{11} - 3\omega_s^2} = n(1 - 3\epsilon/2)$$

induced by the initial perturbation in the fiducial plane. The difference between ω' and the orbital frequency ω is caused by the general relativistic pericenter drift of the perturbed quasielliptic orbit,

$$\omega - \omega' = 3\epsilon n.$$

If the initial perturbation in the fiducial plane is zero, the trajectory of the gyroscope's projection onto the plane coincides with the circular geodesic.

[1] Equation (15) can be obtained by expanding the Papapetrou equations up to linear terms in displacement ξ in the leading approximation (1) of the spin–orbit interaction. At $\mathbf{S} = 0 \longrightarrow \mathbf{a} = 0$, Eq. (15) is reduced to the geodesic deviation equation.

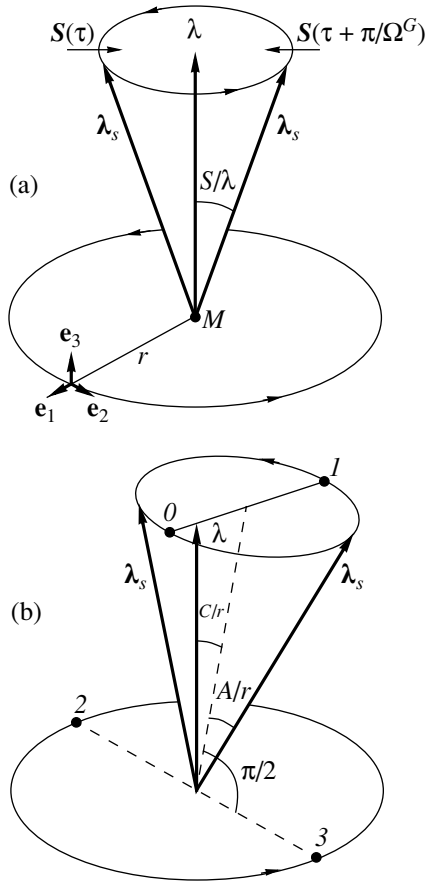


Fig. 1. Orbit of the gyroscope. Orbital moments of the fiducial geodesic and the gyroscope are λ and λ_s, respectively. (a) Precession of the gyroscope orbit at $C = 0$. (b) Variable inclination of the gyroscope orbit, the constant C is arbitrary. The orbital moment λ_s points at the positions marked *0* and *1* when $\sin((\eta - \zeta)/2)$ equals 0 and 1, respectively. At points *2* and *3* it turns out that $\cos\eta = 0$.

The equation of forced oscillations (20) along the orbital normal,

$$\ddot{\xi}_3 + \omega^2\xi_3 = 3\epsilon\frac{S}{\lambda}g\cos(\omega_s + \beta), \tag{21}$$

proves to be quasiresonant due to proximity of the frequencies of the natural tidal oscillation $\sqrt{E_{33}} = \omega$ and of the compelling force ω_s. The difference of the frequencies $\Delta\omega$ in Eq. (4), which prevents the oscillation from becoming resonance, is equal to the geodetic precession Ω^G. The general solution of Eq. (21),

$$\begin{gathered}\xi_3 = A\cos\zeta - C\cos\eta,\\ \zeta = \omega_s\tau + \beta, \quad \eta = \omega\tau + \alpha,\end{gathered} \tag{22}$$

contains the amplitude A given by (5) and two integration constants, C and α. If $C = 0$, oscillation (22) describes the precession of the gyroscope's orbit tilted by the angle

$$A/r = S/\lambda$$

relative to the fiducial plane, with the angular velocity of the geodetic precession given by (4) (Fig. 1a). The evolution of the gyroscope's orbital moment with arbitrary C is presented in the Fig. 1b. If $C = A$, pure beating occurs,

$$\xi_3 = 2A\sin\frac{\eta - \zeta}{2}\sin\frac{\eta + \zeta}{2}. \tag{23}$$

The center of mass oscillates along the orbital normal with a variable amplitude modulated by geodetic precession (4). The initial condition

$$\xi_3(\tau = 0) = 0$$

is provided by choosing the constant $\alpha = -\beta$,

$$\xi_3 = 2A\sin\frac{\Delta\omega}{2}\tau\sin\left(\frac{\omega + \omega_s}{2}\tau + \beta\right). \tag{24}$$

Within a time $\tau \ll (\Delta\omega)^{-1}$, the oscillation amplitude grows at the rate $A\Delta\omega$ given by (6) and (7). The condition

$$\dot{\xi}_3(\tau = 0) = 0$$

fixes the initial spin orientation $\sin\beta = 0$ along the radial direction (see (18)).

The problem of measuring oscillation (24) is complicated by the circumstance that initial perturbations lead to the natural tidal oscillation with the orbital frequency $\omega^2 = E_{33}$ (see (21)). Therefore, gyroscopes with antiparallel spins must be manufactured to be coaxial. In order that the Newtonian harmonic oscillation due to instrumental error be smaller than the relativistic oscillation induced by the spin–orbit interaction, strong restrictions on the initial perturbations $\xi_3(0)$ and $\dot{\xi}_3(0)$ are required,

$$\xi_3(0) \ll \xi \sim A\Delta\omega\tau_f, \quad \dot{\xi}_3(0) \ll \omega\xi, \tag{25}$$

where τ_f is the formation time of the amplitude measured.

4. THE EFFECT OF FIELD OBLATENESS

The Newtonian oblateness of the source does not lead to forced gyroscope oscillations. The oblateness affects the natural tidal oscillation frequency $(\tilde{E}_{33})^{1/2}$, the orbital frequency $\tilde{\omega}$, and consequently, the angular velocity $\tilde{\omega}_s$ of the spin rotation relative to the orbital

triad. The two frequencies, $(\tilde{E}_{33})^{1/2}$ and $\tilde{\omega}_s$, enter the equation of motion of the gyroscope's center of mass,

$$\ddot{\tilde{\xi}}_3 + \tilde{E}_{33}\tilde{\xi}_3 = 3\epsilon\frac{S}{\lambda}g\cos(\tilde{\omega}_s\tau + \beta). \quad (26)$$

Considering only the quadrupole moment J_2 (which is given by $J_2 \approx 1 \times 10^{-3}$ for the Earth), we find for an equatorial orbit that

$$\sqrt{\tilde{E}_{33}} = \omega\left(1 + \frac{9}{4}J_2\frac{R^2}{r^2}\right), \quad (27)$$

$$\tilde{\omega} = \omega\left(1 + \frac{3}{4}J_2\frac{R^2}{r^2}\right), \quad (28)$$

$$\tilde{\omega}_s = \omega_s\left(1 + \frac{3}{4}J_2\frac{R^2}{r^2}\right), \quad (29)$$

where R is the equatorial radius of the source. The frequency $(\tilde{E}_{33})^{1/2}$ differs from the orbital frequency $\tilde{\omega}$ because of the Newtonian quadrupole precession Ω^J of the orbital plane,

$$\tilde{\omega} - \sqrt{\tilde{E}_{33}} = -\frac{3}{2}\omega J_2\frac{R^2}{r^2} = \Omega^J. \quad (30)$$

The gyroscope's axis does not undergo the additional Newtonian precession,

$$\tilde{\omega} - \tilde{\omega}_s = \Omega^G.$$

As a result of the difference in Eq. (30), small denominator (4) is changed as

$$\Delta\tilde{\omega} = \sqrt{\tilde{E}_{33}} - \tilde{\omega}_s = \Omega^G - \Omega^J \approx -\Omega^J \approx \Delta\omega\frac{J_2}{\epsilon}\frac{R^2}{r^2}. \quad (31)$$

The oscillation modulation period is then given by

$$\tilde{T} = \frac{2\pi}{\Delta\tilde{\omega}}$$

and amplitude (5) becomes

$$\tilde{A} = -A\frac{\Omega^G}{\Omega^J} = \frac{S}{\lambda}\frac{\epsilon}{J_2}\frac{r^2}{R^2}. \quad (32)$$

The gyroscope's orbital moment vector describes a conic surface with the apex angle $2\tilde{A}/r$ and the time period $\tilde{T}$. The quadrupole precession period $\tilde{T}$ of a near-Earth orbit is two months. For the pure beating

$$\tilde{\xi}_3 = 2\tilde{A}\sin\frac{\Delta\tilde{\omega}}{2}\tau\sin\left[\left(\tilde{\omega}_s + \frac{\Delta\tilde{\omega}}{2}\right)\tau + \beta\right] \quad (33)$$

within a timescale of $t \ll \tilde{T}$, the oscillation increases precisely as in the case of a spherically symmetric field (see Eq. (6)),

$$\tilde{A}\Delta\tilde{\omega} = A\Delta\omega = \frac{3}{2}\epsilon\frac{S}{\lambda}v. \quad (34)$$

The maximum amplitude formed in time $\tilde{T}/2$ in a near-Earth orbit for a gyroscope with $S/\lambda \sim 10^{-9}$ (see Eq. (7)) is

$$\tilde{A} \sim 10^{-7}\ \text{cm} \quad (35)$$

and is several orders as good as the present-day limit of measuring small oscillations.

5. CONCLUSIONS

The general relativistic quasiresonant spin–orbit interaction leads to oscillation of the gyroscope's center of mass relative to the fiducial geodesic along the orbital normal. The beating amplitude does not include the speed of light and equals the ratio of the intrinsic moment of the gyroscope to its orbital moment. The modulation frequency equals the angular velocity of the geodetic precession. The oscillation represents the precession of the gyroscope orbital moment. Within an acceptable time, the oscillation amplitude reaches values that are amenable to experimental analysis.

Taking the source oblateness into account decreases the beating amplitude and increases the modulation frequency by a factor that is equal to the ratio of the quadrupole precession velocity to the geodetic precession velocity. The period of the quadrupole precession turns out to be quite sufficient time to form a measurable amplitude of the oscillation. The tidal acceleration, providing the quasiresonant character of the oscillation, leads to strong restrictions that must be imposed on the initial perturbations in order to distinguish the relativistic spin–orbit oscillation in the background of the Newtonian tidal oscillation.

REFERENCES

1. W. de Sitter, Mon. Not. R. Astron. Soc. **77**, 155 (1916).
2. L. I. Schiff, Phys. Rev. Lett. **4**, 215 (1960).
3. R. J. Adler and A. S. Silbergeit, Int. J. Theor. Phys. **39**, 1291 (2000).

4. V. A. Fock, *The Theory of Space, Time and Gravitation* (Fizmatgiz, Moscow, 1961; MacMillan, New York, 1964), Sect. 77.
5. A. Papapetrou, Proc. R. Soc. London, Ser. A **209**, 248 (1951).
6. W. G. Dixon, Nuovo Cimento **34**, 317 (1964).
7. B. M. Barker and R. F. O'Connel, J. Math. Phys. **28**, 661 (1987).
8. K. S. Thorne and J. B. Hartle, Phys. Rev. D **31**, 1815 (1985).
9. A. A. Pomeranskiĭ and I. B. Khriplovich, Zh. Éksp. Teor. Fiz. **113**, 1537 (1998) [JETP **86**, 839 (1998)].
10. A. A. Pomeransky, R. A. Sen'kov, and I. B. Khriplovich, Usp. Fiz. Nauk **170**, 1129 (2000) [Phys. Usp. **43**, 1055 (2000)].
11. R. I. Khrapko, Zh. Éksp. Teor. Fiz. **90**, 401 (1986) [Sov. Phys. JETP **63**, 231 (1986)].
12. K. Pyragas *et al.*, Astrophys. Space Sci. **149**, 39 (1988).
13. O. S. Ivanitskaya, *Lorentzian Basis and Gravitational Effects in Einstein's Gravitation Theory* (Nauka i Tekhnika, Minsk, 1979).
14. E. N. Epikhin and N. V. Mitskievich, Izv. Vyssh. Uchebn. Zaved., Fiz., No. 5, 115 (1979).
15. O. B. Karpov, Izv. Vyssh. Uchebn. Zaved., Fiz., No. 11, 68 (1982).
16. O. B. Karpov, in *Abstracts of the 10th International Conference on Gen. Relat. Grav., Padova* (1983), Vol. 2, p. 965.
17. O. B. Karpov, Preprint No. 012-90, MIFI (Moscow Inst. of Engineering Physics, 1990).
18. F. Pirani, Acta Phys. Pol. **15**, 389 (1956).
19. E. Corinaldesi and A. Papapetrou, Proc. R. Soc. London, Ser. A **209**, 259 (1959).

Journal of Experimental and Theoretical Physics, Vol. 96, No. 4, 2003, pp. 587–593.
Translated from Zhurnal Éksperimental'noĭ i Teoreticheskoĭ Fiziki, Vol. 123, No. 4, 2003, pp. 665–671.

GRAVITATION, ASTROPHYSICS

Macroscopic Einstein Equations for a System of Gravitating Particles with Unequal Masses

A. V. Zakharov* and R. K. Mukharlyamov

Kazan State Energy University, ul. Lenina 18, Kazan, Tatarstan, 420008 Russia
**e-mail: Alexei.Zakharov@ksu.ru*

Received November 10, 2002

Abstract—We derive macroscopic Einstein equations, to within terms of the second order of smallness in interaction, for a system of gravitationally interacting particles with unequal masses. We generalize the results of [1, 2], which are applicable only to a system of gravitationally interacting particles with equal masses.

1. INTRODUCTION

This work is a generalization of our works [1, 2], which are devoted to the derivation of macroscopic Einstein equations, to within terms of the second order of smallness in interaction, for a system of gravitationally interacting particles. Our previous results are generalized to systems of gravitationally interacting particles with unequal masses.

We reduced the macroscopic Einstein equations to the form

$$G_{ij} + \varphi^{k}_{ij;\,k} + \mu_{ij} = \chi T_{ij},$$

where G_{ij} is the Einstein tensor, T_{ij} is the energy–momentum tensor, and χ is the Einstein gravitational constant. The semicolon denotes a covariant derivative. The new terms on the left-hand sides of the derived equations are attributable to the interaction between particles. They are traceless tensors with a zero divergence. We give an explicit covariant expression for these terms via the integrals of expressions dependent on the single-particle distribution functions of interacting particles over momentum space. These expressions are proportional to the Einstein constant cubed and to the particle density squared. The latter implies that interaction effects can manifest themselves only in high-density systems (the Universe at early evolutionary stages, dense objects close to gravitational collapse, and others) or in macrosystems consisting of massive objects (clusters of galaxies).

2. MACROSCOPIC EINSTEIN EQUATIONS

The method of deriving the macroscopic Einstein equations is described in [1]; the notation used here is the same as that in [1]. The macroscopic Einstein equations are derived as follows. First, we write the system of microscopic Einstein equations

$$\tilde{G}^{ij} = \chi \tilde{T}^{ij}, \tag{1}$$

where, $\tilde{G}^{ij}$ is the Einstein tensor of the Riemannian space with metric $\tilde{g}_{ij}$; $\tilde{T}^{ij}$ is the microscopic energy–momentum tensor,

$$\tilde{T}^{ij} = \sum_a c \int \frac{d^4\tilde{p}_a}{\sqrt{-\tilde{g}}} \tilde{p}^i_a \tilde{u}^j_a \tilde{N}_a(q^i, \tilde{p}_j); \tag{2}$$

$\tilde{g}$ is the determinant of the metric $\tilde{g}_{ij}$; $\tilde{p}^i_a$ is the momentum of the particles of type a; $\tilde{u}^i_a = \tilde{p}^i_a / \sqrt{\tilde{g}_{kj}\tilde{p}^k_a \tilde{p}^j_a}$, $\chi = 8\pi k/c^4$ is the Einstein constant; k is the gravitational constant; and $\tilde{N}_a(q^i, \tilde{p}_j)$ is the Klimontovich random function [3]:

$$\tilde{N}_a(\tilde{q}^i, \tilde{p}_j) = \sum_{i=1}^{n_a} \int d\tilde{s}\, \delta^4(q^i - q^i_{(l)})\delta^4(\tilde{p}_j - \tilde{p}^{(l)}_j(s)). \tag{3}$$

Here, n_a is the number of particles of type a; $\tilde{s}$ is the canonical parameter along the particle trajectory: $d\tilde{s} = \sqrt{\tilde{g}_{ij} dq^i dq^j}$; and $q^i_{(l)}$ and $\tilde{p}^{(l)}_j$ are, respectively, the coordinates and momentum of the lth particle of type a, which can be determined from the equations of motion (geodesic equations). The function $\tilde{N}_a$ satisfies the Liouville equation

$$\tilde{p}^i \frac{\partial \tilde{N}_a}{\partial q^i} + \tilde{\Gamma}_{j,\,ik}\tilde{p}^j \tilde{p}^k \frac{\partial \tilde{N}_a}{\partial \tilde{p}_i} = 0. \tag{4}$$

Next, the metric $\tilde{g}_{ij}$ is represented as

$$\tilde{g}_{ij} = g_{ij} + h_{ij}, \tag{5}$$

where $g_{ij} = \langle \tilde{g}_{ij} \rangle$ is the ensemble-averaged metric [1, 2]. In system (1), (2), and (4), we change from the quantities measured in the metric $\tilde{g}_{ij}$ to the quantities measured in the metric g_{ij}. All of these quantities are denoted without an upper tilde. As a result, system (1), (2), and (4) takes the form

$$\begin{aligned} &R_{ij} + \nabla_m \Omega^m_{ij} - \nabla_j \Omega^m_{im} + \Omega^m_{mn}\Omega^n_{ij} - \Omega^m_{jn}\Omega^n_{im} \\ &= \chi \sum_a \int \frac{d^4p}{\sqrt{-g}} \alpha \sqrt{\frac{g}{\tilde{g}}} \left[\tilde{g}_{ik}\tilde{g}_{jm} - \frac{1}{2}\tilde{g}_{ij}\tilde{g}_{km} \right] p^k p^m N_a, \end{aligned} \tag{6}$$

$$p^i \frac{\partial N_a}{\partial q^i} + \Gamma_{j,ik} p^k p^j \frac{\partial N_a}{\partial p_i} = \frac{\partial}{\partial p_i}(\Omega^m_{jk}\Delta_{mi} p^j p^k N_a). \tag{7}$$

Here, p^i is the momentum measured in the metric g_{ij} and N_a is the random function in the space with metric g_{ij}. An expression for this function can be obtained from (3) by omitting the tilde in all quantities; $\Omega^m_{kj} = \tilde{\Gamma}^m_{kj} - \Gamma^m_{kj}$ is the difference between the Christoffel symbols of the second kind for the metrics $\tilde{g}_{ij}$ and g_{ij}, $\Delta_{ij} = g_{ij} - u_i u_j$, and $u_i = p_i/\sqrt{p_l p^l}$.

The next step involves expanding Eqs. (6) to within terms of the second order of smallness in h_{ij} and averaging the derived equations over the ensemble. When we restrict ourselves to deriving the averaged equations to within terms of the second order of smallness in interaction, we can obtain a closed system of equations for the single-particle distribution function $f_a(q, p) = \langle N_a \rangle / n_a$ and for the averaged metric g_{ij}. The equation for f_a was previously derived in [1, 4]:

$$\begin{aligned} &u^i \frac{\partial f_a}{\partial q^i} + u^j p^k \Gamma_{j,ik} \frac{\partial f_a}{\partial p_i} \\ &= \sum_b \int \frac{d^4p}{\sqrt{-g}} E_{ij}(p, p') \left(\frac{\partial f_a}{\partial p_j} f'_b - \frac{\partial f'_b}{\partial p'_j} f_a \right), \end{aligned} \tag{8}$$

where

$$\begin{aligned} E_{ij}(p, p') &= \frac{2\pi k^2 L n_b}{c^6 [(u, u')^2 - 1]^{3/2}} \\ &\times [2(u, u')(p, p') - (u, p)(u', p')]^2 \{ -g_{ij}[(u, u')^2 - 1] \\ &- u_i u_j - u'_i u'_j + (u, u')(u_i u'_j + u'_i u_j) \}. \end{aligned} \tag{9}$$

Here, $(u, u') = u'_i u^i$, $(u, u) = u_i u^i$, etc. The quantities with and without a prime refer to the particles of types b and a, respectively. The quantity L is an analog of the Coulomb logarithm [4, 5]

$$L = \int_{k_{\min}}^{k_\infty} \frac{dk}{k}. \tag{10}$$

For the averaged metric g_{ij}, the equations were brought in [2] to the form

$$G_{ij} + \varphi^k_{ij;\,k} + \mu_{ij} = \chi T_{ij}, \tag{11}$$

where the semicolon denotes a covariant derivative in the space with metric g_{ij}, G_{ij} is the Einstein tensor of this space, and T_{ij} is the macroscopic energy–momentum tensor.

The tensors φ^k_{ij} and μ_{ij} can be expressed in terms of the single-particle distribution function f_b specified in eight-dimensional phase space in which all four components of the four-dimensional momentum are assumed to be independent:

$$\varphi^k_{ij} = -(1/2)(\delta^k_n \delta^s_j - \delta^k_j \delta^s_n) P^n_{is},$$

where

$$\begin{aligned} P^n_{is} &= \sum_{bc} \frac{\chi^3 m_b m_c n_b n_c c^6}{8(2\pi)^3} \\ &\times \int d^4p' \int d^4p''\, f_b(x')\, f_c(x'')[1 - 10(u'u'')^2] \\ &\times \left(u'^n u'_l - \frac{1}{2}\delta^n_l \right) \left[\left(u''_i u''_s - \frac{1}{2} g_{is} \right) g^{lf} \right. \\ &- \left(u''^l u''_i - \frac{1}{2}\delta^l_i \right)\delta^f_s - \left(u''^l u''_s - \frac{1}{2}\delta^l_s \right)\delta^f_i \Big] \\ &\times (m_c u'^m K^{(1)}_{fm}(u', u'') + m_b u''^m K^{(2)}_{fm}(u', u'')) \\ &- \sum_{bc} \frac{\chi^3 m_b^2 m_c^2 n_b n_c c^7}{4(2\pi)^3} \int d^4p' \int d^4p'' \left(u'^n u'_l - \frac{1}{2}\delta^n_l \right) \\ &\times \left[\left(u''_i u''_s - \frac{1}{2} g_{is} \right) g^{lf} - \left(u''^l u''_i - \frac{1}{2}\delta^l_i \right)\delta^f_s \right. \\ &- \left(u''^l u''_s - \frac{1}{2}\delta^l_s \right)\delta^f_i \Big] \Big\{ \Big[(u'u'')^2 (\delta^m_j + u'_j u'^m) \\ &- \frac{1}{2}(\delta^m_j - u'_j u'^m) - 2(u'u'')u'^m u''_j \Big] \\ &\times f_c(x'') \frac{\partial f_b(x')}{\partial p'_j} K^{(1)}_{fm}(u', u'') \\ &+ \Big[(u', u'')^2 (\delta^m_j + u''_j u''^m) - \frac{1}{2}(\delta^m_j - u''_j u''^m) \end{aligned} \tag{12}$$

$$-2(u'u'')u''^m u'_j\Big]f_b(x')\frac{\partial f_c(x'')}{\partial p''_j}K^{(2)}_{fm}(u', u'')\Big\},$$

$$\mu_{ij} = \sum_{bc}\frac{\chi^3 m_b m_c n_b n_c c^6}{16(2\pi)^3}$$

$$\times\int\frac{d^4p'}{\sqrt{-g}}\int\frac{d^4p''}{\sqrt{-g}}f_b(x')f_c(x'')(1-10(u'u'')^2)$$

$$\times\Big\{u'^q u'_j\delta^r_i - u''^r u''_j\delta^q_i + g^{qr}\Big[\Big((u'u'')^2+\frac{1}{2}\Big)u''_i u''_j$$

$$+\frac{1}{2}\Big((u'u'')^2-\frac{1}{2}\Big)g_{ij} - 2(u'u'')u'_i u''_j\Big]$$

$$-\Big((u'u'')^2-\frac{1}{2}\Big)\delta^q_i\delta^r_j\Big\}(m_c u'^m J^{(1)}_{rqm}(u', u'')$$

$$+ m_b u''^m J^{(2)}_{rqm}(u', u'')) - \sum_{bc}\frac{\chi^3 m_b^2 m_c^2 n_b n_c c^7}{8(2\pi)^3}$$

$$\times\int\frac{d^4p'}{\sqrt{-g}}\int\frac{d^4p''}{\sqrt{-g}}\Big\{u'^q u'_j\delta^r_i - u''^r u''_j\delta^q_i \qquad (13)$$

$$+ g^{qr}\Big[\Big((u'u'')^2+\frac{1}{2}\Big)u''_i u''_j + \frac{1}{2}\Big((u'u'')^2-\frac{1}{2}\Big)g_{ij}$$

$$-2(u'u'')u'_i u''_j\Big] - \Big((u'u'')^2-\frac{1}{2}\Big)\delta^q_i\delta^r_j\Big\}$$

$$\times\Big\{\Big[\Big((u'u'')^2-\frac{1}{2}\Big)\delta^m_f + \Big((u'u'')^2+\frac{1}{2}\Big)u'_f u'^m$$

$$-2(u'u'')u''_f u'^m\Big]J^{(1)}_{rqm}(u', u'')f_c(x'')\frac{\partial f_b(x')}{\partial p'_f}$$

$$+\Big[\Big((u'u'')^2-\frac{1}{2}\Big)\delta^m_f + \Big((u'u'')^2+\frac{1}{2}\Big)u''_f u''^m$$

$$-2(u'u'')u'_f u''^m\Big]J^{(2)}_{rqm}(u', u'')f_b(x')\frac{\partial f_c(x'')}{\partial p''_f}\Big\}.$$

In (12), we use the notation $K^{(1)}_{fm}(u', u'')$ and $K^{(2)}_{fm}(u', u'')$ for the tensors that have the following form in a local Lorentz frame of reference where $g_{ij} = \eta_{ij}$ is the Minkowski tensor:

$$K^{(1)}_{fm}(u', u'') = \frac{1}{u'^0 u''^0}\int\frac{d^3k}{k^3}\int_{-\infty}^{\eta}d\eta'\int_{-\infty}^{\eta}d\eta''\int_{-\infty}^{\eta'}d\tau'$$

$$\times\int_{-\infty}^{\tau'}d\tau''(e^{ik(\eta'-\eta)} - e^{-ik(\eta'-\eta)})$$

$$\times(k^+_f e^{ik(\eta''-\eta)} - k^-_f e^{-ik(\eta''-\eta)})$$

$$\times(k^+_m e^{-ik(\tau''-\tau')} - k^-_m e^{ik(\tau''-\tau')})$$

$$\times\exp\Big[\frac{i}{c}(\mathbf{k}\cdot\mathbf{v}'')(\eta''-\tau'') + \frac{i}{c}(\mathbf{k}\cdot\mathbf{v}')(\tau'-\eta')\Big],$$

$$K^{(2)}_{fm}(u', u'') = \frac{1}{u'^0 u''^0}\int\frac{d^3k}{k^3}\int_{-\infty}^{\eta}d\eta'\int_{-\infty}^{\eta}d\eta''\int_{-\infty}^{\eta''}d\tau''$$

$$\times\int_{-\infty}^{\tau''}d\tau'(e^{ik(\eta'-\eta)} - e^{-ik(\eta'-\eta)})$$

$$\times(k^+_f e^{ik(\eta''-\eta)} - k^-_f e^{-ik(\eta''-\eta)})$$

$$\times(k^+_m e^{-ik(\tau''-\tau')} - k^-_m e^{ik(\tau''-\tau')})$$

$$\times\exp\Big[\frac{i}{c}(\mathbf{k}\cdot\mathbf{v}'')(\eta''-\tau'') + \frac{i}{c}(\mathbf{k}\cdot\mathbf{v}')(\tau'-\eta')\Big].$$

After calculating the integrals over τ', τ'', η', and η'', these expressions take the form

$$K^{(1)}_{fm}(u', u'') = \frac{2\pi c^5}{u'^0 u''^0}\int\frac{d^3k}{k^2}\delta(\mathbf{k}\cdot\mathbf{v}'' - \mathbf{k}\cdot\mathbf{v}')$$

$$\times\Big\{\frac{k^+_f k^+_m}{(kc-\mathbf{k}\cdot\mathbf{v}'')(kc+\mathbf{k}\cdot\mathbf{v}'')^3} \qquad (14)$$

$$+\frac{k^+_f k'^+_m + k^-_f k^+_m}{(kc-\mathbf{k}\cdot\mathbf{v}'')^2(kc+\mathbf{k}\cdot\mathbf{v}'')^2}$$

$$+\frac{k^-_f k^-_m}{(kc-\mathbf{k}\cdot\mathbf{v}'')^3(kc+\mathbf{k}\cdot\mathbf{v}'')}\Big\} = K_{fm}(u', u''),$$

$$K^{(2)}_{fm}(u', u'') = -K^{(1)}_{fm}(u', u'') = -K_{fm}(u', u''). \qquad (15)$$

These equalities are valid only in the local Lorentz frame of reference. To derive covariant expressions for $K^{(1)}_{fm}(u', u'')$ and $K^{(2)}_{fm}(u', u'')$, we take into account the fact that these two tensors must be calculated in the same frame of reference. It is convenient to choose a

center-of-mass frame in which $\mathbf{p}'' = -\mathbf{p}'$ as this frame. In this frame of reference,

$$K_{00} = K_{0\alpha} = 0,$$
$$K_{\alpha\beta} = \frac{4\pi^2 c}{v' u_0' u_0'' k_{\min}^2 (1 + m_b u_0'/m_c u_0'')}\left(\delta_{\alpha\beta} - \frac{v_\alpha' v_\beta'}{v'^2}\right). \quad (16)$$

Here, $v' = \sqrt{v_1'^2 + v_2'^2 + v_3'^2}$ and $v_\alpha' = v'^\alpha = u'^\alpha/u'^0$ are the spatial components of the vector $\mathbf{v}'$.

The covariant generalization of (16) is

$$K_{ij}(u', u'') = \frac{4\pi^2}{k_{\min}^2[(u'u'')^2 - 1]^{3/2}}\{-[(u'u'')^2 - 1]g_{ij} - u_i' u_j' - u_i'' u_j'' + (u'u'')(u_i' u_j'' + u_i'' u_j')\}. \quad (17)$$

The expressions for $K_{fm}^{(1)}(u', u'')$ and $K_{fm}^{(2)}(u', u'')$ diverge when $k \longrightarrow 0$, i.e., at large impact parameters. This is because we integrate over an infinite range, while actually we must restrict ourselves to integrating only over the correlation region where the metric was assumed to change only slightly. As in the case of deriving the kinetic equation, this difficulty can be obviated by introducing a truncation in the divergent integral

$$\int_0^\infty \frac{dk}{k^3}.$$

We set the lower integration limit equal not to zero but to $k_{\min} = 1/r_{\max}$, where $r_{\max}$ is the correlation length. The above integral then takes on the value of $1/2k_{\min}^2 = (1/2)r_{\max}^2$. Experience in deriving the relativistic kinetic equation (see [5, 6]) shows that in the case of a more careful analysis, the integrals converge when $r \longrightarrow \infty$, with the contribution from $r > r_{\max}$ to the integrals being negligible. Estimates for $r_{\max}$ when the averaged metric g_{ij} is the metric of an isotropic cosmological model are given in [5, 6].

Tensor (17) has the following properties:

$$K_{ij}(u', u'') = K_{ij}(u'', u');$$
$$K_{ij}u'^i = K_{ij}u''^i = 0; \quad K_{ij} = K_{ji}. \quad (18)$$

As a result, the expression for P_{is}^n is significantly simplified. The tensor $\varphi_{ij}^k = -(1/2)(\delta_n^k\delta_j^s - \delta_j^k\delta_n^s)P_{is}^n$ rather than P_{is}^n appears in the macroscopic Einstein equations. The expression for this tensor reduces to

$$\varphi_{ij}^k = -\sum_{bc}\frac{\chi^3 m_b^2 m_c^2 n_b n_c c^7}{8(2\pi)^3}\int\frac{d^4p'}{\sqrt{-g}}\int\frac{d^4p''}{\sqrt{-g}} \times\left[\frac{1}{2}g^{fk}u_i''u_j'' + u'^k(u'u'')(\delta_j^f u_i'' + \delta_i^f u_j'')\right] \times\left((u'u'')^2 - \frac{1}{2}\right)K_{fr}(u', u'') \times\left(f_c(x'')\frac{\partial f_b(x')}{\partial p_r'} - f_b(x')\frac{\partial f_c(x'')}{\partial p_r''}\right). \quad (19)$$

Note that

$$g^{ij}\varphi_{ij}^k = 0, \quad \varphi_{ij}^i = 0, \quad \varphi_{ij}^k = \varphi_{ji}^k. \quad (20)$$

In (13), we use the notation $J_{rqm}^1(u', u'')$ and $J_{rqm}^2(u', u'')$ for the tensors that have the following form in the local Lorentz frame of reference:

$$J_{lmn}^{(1)}(u', u'') = \frac{1}{u'^0 u''^0}\int\frac{d^3k}{k^3}\int_{-\infty}^{\eta}d\eta'\int_{-\infty}^{\eta}d\eta''\int_{-\infty}^{\eta'}d\tau' \times\int_{-\infty}^{\tau'}d\tau''(k_l^+ e^{-ik(\eta'-\eta)} - k_l^- e^{ik(\eta'-\eta)}) \times(k_m^+ e^{ik(\eta''-\eta)} - k_m^- e^{-ik(\eta''-\eta)}) \times(k_n^- e^{ik(\tau''-\tau')} - k_n^+ e^{-ik(\tau''-\tau')}) \times\exp\left[\frac{i}{c}(\mathbf{k}\cdot\mathbf{v}'')(\eta''-\tau'') + \frac{i}{c}(\mathbf{k}\cdot\mathbf{v}')(\tau'-\eta')\right],$$

$$J_{lmn}^{(2)}(u', u'') = \frac{1}{u'^0 u''^0}\int\frac{d^3k}{k^3}\int_{-\infty}^{\eta}d\eta'\int_{-\infty}^{\eta}d\eta''\int_{-\infty}^{\eta''}d\tau'' \times\int_{-\infty}^{\tau''}d\tau'(k_l^+ e^{-ik(\eta'-\eta)} - k_l^- e^{ik(\eta'-\eta)}) \times(k_m^+ e^{ik(\eta''-\eta)} - k_m^- e^{-ik(\eta''-\eta)}) \times(k_n^- e^{ik(\tau''-\tau')} - k_n^+ e^{-ik(\tau''-\tau')}) \times\exp\left[\frac{i}{c}(\mathbf{k}\cdot\mathbf{v}'')(\eta''-\tau'') + \frac{i}{c}(\mathbf{k}\cdot\mathbf{v}')(\tau'-\eta')\right].$$

Calculating the integrals over η', η'', τ', and τ'' yields

$$J_{lmn}^{(1)}(u', u'') = \frac{c^4}{u'^0 u''^0}\int\frac{d^3k}{k^3}\frac{\text{V.p.}}{(\mathbf{k}\cdot\mathbf{v}'' - \mathbf{k}\cdot\mathbf{v}')}$$

$$\times\left\{\frac{k_l^+k_m^+k_n^+}{(kc+\mathbf{k}\cdot\mathbf{v}'')^3}+\frac{k_l^+k_m^+k_n^-+k_l^+k_m^-k_n^++k_l^-k_m^+k_n^+}{(kc+\mathbf{k}\cdot\mathbf{v}'')^2(kc-\mathbf{k}\cdot\mathbf{v}'')}\right. \quad (21)$$

$$\left.+\frac{k_l^+k_m^-k_n^-+k_l^-k_m^+k_n^-+k_l^-k_m^-k_n^+}{(kc+\mathbf{k}\cdot\mathbf{v}'')(kc-\mathbf{k}\cdot\mathbf{v}'')^2}+\frac{k_l^-k_m^-k_n^-}{(kc-(\mathbf{k}\cdot\mathbf{v}''))^3}\right\}.$$

The expression for $J^{(2)}_{lmn}(u', u'')$ can be obtained from the expression for $J^{(1)}_{lmn}(u', u'')$ by the substitution $\mathbf{v}' \longrightarrow \mathbf{v}''$ and vice versa. The symbol V.p. means that a principal-value integral is calculated.

Just as in the previous case, we specify (21) in the center-of-mass frame in which $\mathbf{p}'' = -\mathbf{p}'$. In this frame of reference, the components $J^{(1)}_{lmn}(u', u'')$ are (the spatial indices in the three-dimensional velocity vector v'^{α} are omitted using the three-dimensional Kronecker symbol $\delta_{\alpha\beta}$)

$$J_{000} = -\frac{2}{(1+m_bu_0'/m_cu_0'')}\frac{1}{u_0'u_0''}\times\left(\frac{m_bu_0'}{m_cu_0''}\right)^3\alpha\left(\frac{m_bu_0'}{m_cu_0''}v'\right)\frac{v'^2}{c^2}, \quad (22)$$

$$J_{00\alpha} = -\frac{2}{(1+m_bu_0'/m_cu_0'')}\frac{1}{u_0'u_0''}\times\left(\frac{m_bu_0'}{m_cu_0''}\right)^2\alpha\left(\frac{m_bu_0'}{m_cu_0''}v'\right)\frac{v'_\alpha}{c^2}, \quad (23)$$

$$J_{0\alpha\beta} = -\frac{2}{(1+m_bu_0'/m_cu_0'')}\frac{1}{u_0'u_0''}\left(\frac{m_bu_0'}{m_cu_0''}\right)\times\alpha\left(\frac{m_bu_0'}{m_cu_0''}v'\right)\delta_{\alpha\beta}+\frac{2}{(1+m_bu_0'/m_cu_0'')}\frac{1}{u_0'u_0''}\times\left(\frac{m_bu_0'}{m_cu_0''}\right)\beta\left(\frac{m_bu_0'}{m_cu_0''}v'\right)\left(\delta_{\alpha\beta}-\frac{v'_\alpha v'_\beta}{v'^2}\right), \quad (24)$$

$$J_{\alpha\beta\gamma} = -\frac{2}{(1+m_bu_0'/m_cu_0'')}\frac{1}{u_0'u_0''}\alpha\left(\frac{m_bu_0'}{m_cu_0''}v'\right)\times\frac{c^2}{v'^2}\left[\delta_{\alpha\beta}\frac{v'_\gamma}{c}+\delta_{\alpha\gamma}\frac{v'_\beta}{c}+\delta_{\beta\gamma}\frac{v'_\alpha}{c}-2\frac{v'_\alpha v'_\beta v'_\gamma}{cv'^2}\right]+\frac{2}{(1+m_bu_0'/m_cu_0'')}\frac{1}{u_0'u_0''}\beta\left(\frac{m_bu_0'}{m_cu_0''}v'\right)\times\frac{c^2}{v'^2}\left[\left(\delta_{\alpha\beta}-\frac{v'_\alpha v'_\beta}{v'^2}\right)\frac{v'_\gamma}{c}+\left(\delta_{\alpha\gamma}-\frac{v'_\alpha v'_\gamma}{v'^2}\right)\frac{v'_\beta}{c}+\left(\delta_{\beta\gamma}-\frac{v'_\beta v'_\gamma}{v'^2}\right)\frac{v'_\alpha}{c}\right]. \quad (25)$$

The functions α and β in (22)–(25) depend only on the argument $w = (m_bu_0'/m_cu_0'')v'$. In explicit form, they are

$$\alpha = \frac{\pi c^3}{w^3k_{\min}}\left[\frac{2\frac{w}{c}\left(1+\frac{w^2}{c^2}\right)}{\left(1-\frac{w^2}{c^2}\right)^2}+\ln\left(\frac{1-\frac{w}{c}}{1+\frac{w}{c}}\right)\right], \quad (26)$$

$$\beta = \frac{\pi c^3}{2w^3k_{\min}}\left[\frac{2\frac{w}{c}\left(3-2\frac{w^2}{c^2}+3\frac{w^4}{c^4}\right)}{\left(1-\frac{w^2}{c^2}\right)^2}+3\left(1+\frac{w^2}{c^2}\right)\ln\left(\frac{1-\frac{w}{c}}{1+\frac{w}{c}}\right)\right]. \quad (27)$$

Here, we use the notation for the integral

$$\frac{1}{k_{\min}} = \int_{k_{\min}}^{\infty}\frac{dk}{k^2}.$$

For the reasons given above, we again set the lower limit equal to $k_{\min} = 1/r_{\max}$.

The covariant generalization of these results obtained in the local Lorentz center-of-mass frame of reference to arbitrary frames of reference is

$$J^{(1)}_{ijk}(u', u'') = J^{(2)}_{ijk}(u'', u') = J_{ijk}(u', u''), \quad (28)$$

$$\begin{aligned}J_{ijk}(u', u'') &= A(g_{ij}u'_k+g_{ik}u'_j+g_{jk}u'_i)\\ &\quad -z(g_{ij}u''_k+g_{ik}u''_j+g_{jk}u''_i)\\ &\quad -(u'_iu''_ju''_k+u''_iu'_ju''_k+u''_iu''_ju'_k)+3zu''_iu''_ju''_k]\\ &\quad +C[u'_iu'_ju'_k-z(u'_iu'_ju''_k+u'_iu''_ju'_k+u''_iu'_ju'_k)\\ &\quad +z^2(u'_iu''_ju''_k+u''_iu'_ju''_k+u''_iu''_ju'_k)-z^3u''_iu''_ju''_k],\end{aligned} \quad (29)$$

where $z = (u'u'') = (u'^iu''_i)$,

$$A = \frac{\pi}{k_{\min}}\frac{(\mu^2+2\mu z+1)^{1/2}(1+\mu z)^2}{\mu^3(z^2-1)^{5/2}}\times\left[\frac{2\mu\sqrt{z^2-1}(1+3\mu^2+2\mu z-2\mu^2z^2)}{(1+\mu z)(\mu^2+2\mu z+1)}+\frac{1-3\mu^2+2\mu z+4\mu^2z^2}{(1+\mu z)^2}\ln\left(\frac{1+\mu z-\mu\sqrt{z^2-1}}{1+\mu z+\mu\sqrt{z^2-1}}\right)\right], \quad (30)$$

$$C = \frac{\pi}{k_{\min}}\frac{(\mu z+1)}{(1+2\mu z+\mu^2)^{1/2}\mu^3(z^2-1)^{7/2}}$$
$$\times\Bigg[2\mu\sqrt{z^2-1}(5+7\mu^2+10\mu z-2\mu^2 z^2)$$
$$+\frac{(1+2\mu z+\mu^2)}{(1+\mu z)}(5-7\mu^2+10\mu z+12\mu^2 z^2)$$
$$\times\ln\left(\frac{1+\mu z-\mu\sqrt{z^2-1}}{1+\mu z+\mu\sqrt{z^2-1}}\right)\Bigg]. \quad (31)$$

Here, $\mu = m_b/m_c$.

For $\mu = 1$, these results are identical to those obtained previously in [2].

Thus, we generalized the results of [2] to a multicomponent system of gravitationally interacting particles with unequal masses.

The tensor $J_{ijk}(u', u'')$ satisfies the identity

$$J_{ijk}(u', u'')u''^k = 0. \quad (32)$$

Because of the properties (28) and (32), we can write expression (13) for μ_{ij} in covariant form:

$$\mu_{ij} = -\sum_{bc}\frac{\chi^3 m_b^2 m_c^2 n_b n_c c^7}{8(2\pi)^3}\int\frac{d^4p'}{\sqrt{-g}}\int\frac{d^4p''}{\sqrt{-g}}$$
$$\times\Bigg\{\Bigg[\left(z^2+\frac{1}{2}\right)(u_i''u_j''+u_i'u_j')$$
$$+\left(z^2-\frac{1}{2}\right)g_{ij}-2z(u_i'u_j''+u_i''u_j')\Bigg]g^{qr}-2\left(z^2-\frac{1}{2}\right)\delta_i^q\delta_j^r\Bigg\},$$
$$f_c(x'')\frac{\partial}{\partial p_f'}\Bigg\{f_b(x')\Bigg[\left(z^2-\frac{1}{2}\right)\delta_f^m$$
$$+\left(z^2+\frac{1}{2}\right)u_f'u'^m-2zu_f''u'^m\Bigg]\Bigg\}J_{rqm}(u', u''). \quad (33)$$

Here, we use the identity

$$\frac{\partial}{\partial p_f'}\Bigg[\left(z^2-\frac{1}{2}\right)\delta_f^m+\left(z^2+\frac{1}{2}\right)u_f'u'^m-2zu_f''u'^m\Bigg]$$
$$=\left(5z^2-\frac{1}{2}\right)\frac{u'^m}{m_b c}.$$

Note that μ_{ij} is a traceless tensor:

$$g^{ij}\mu_{ij} = 0. \quad (34)$$

The tensors φ_{ij}^k and μ_{ij} can be expressed using formulas (19) and (33) in terms of the single-particle distribution functions f_b specified in eight-dimensional phase space in which all four components of the four-dimensional momentum are assumed to be independent. The transformation to the seven-dimensional distribution function F_b is made as follows:

$$n_b f_b(q^i, p_j) = F_b(q^i, p_\alpha)\delta(\sqrt{g^{lm}p_l p_m}-m_b c).$$

Here, the function F_b depends only on the spatial momentum components (the spatial components are denoted by Greek indices).

Integrating (19) and (33) over p_0' and p_0'', we reduce the tensors φ_{ij}^k and μ_{ij} to

$$\varphi_{ij}^k = -\sum_{bc}\frac{\chi^3 m_b^3 m_c^3 c^9}{8(2\pi)^3}\int\frac{d^3p'}{p'^0\sqrt{-g}}\int\frac{d^3p''}{p''^0\sqrt{-g}}$$
$$\times\Bigg[\frac{1}{2}g^{fk}u_i''u_j''+u'^k(u'u'')(\delta_j^f u_i''+\delta_i^f u_j'')\Bigg]$$
$$\times\left((u'u'')^2-\frac{1}{2}\right)K_{f\alpha}(u', u'')$$
$$\times\left(F_c(x'')\frac{\partial F_b(x')}{\partial p_\alpha'}-F_b(x')\frac{\partial F_c(x'')}{\partial p_\alpha''}\right), \quad (35)$$

$$\mu_{ij} = -\sum_{bc}\frac{\chi^3 m_b^3 m_c^3 c^9}{8(2\pi)^3}\int\frac{d^3p'}{p'^0\sqrt{-g}}\int\frac{d^3p''}{p''^0\sqrt{-g}}$$
$$\times\Bigg\{\Bigg[\left(z^2+\frac{1}{2}\right)(u_i''u_j''+u_i'u_j')$$
$$+\left(z^2-\frac{1}{2}\right)g_{ij}-2z(u_i'u_j''+u_i''u_j')\Bigg]g^{qr}$$
$$-2\left(z^2-\frac{1}{2}\right)\delta_i^q\delta_j^r\Bigg\}F_c(x'')\frac{\partial}{\partial p_f'}\Bigg\{F_b(x')\Bigg[\left(z^2-\frac{1}{2}\right)\delta_f^m$$
$$+\left(z^2+\frac{1}{2}\right)u_f'u'^m-2zu_f''u'^m\Bigg]\Bigg\}J_{rqm}(u', u''). \quad (36)$$

Here,

$$\frac{d^2p'}{p'^0\sqrt{-g}}, \quad \frac{d^3p''}{p''^0\sqrt{-g}}$$

are invariant volume elements in the three-dimensional momentum space of the particles of types b and c, respectively. The Greek index α in (35) take on only 1, 2, 3 (the spatial index). The derivative with respect to p_f' in (36) should be calculated as if all four momentum components are independent. The dependence of p_0' on p_α'

is taken into account after differentiation with respect to p'_f.

The tensors φ^k_{ij} and μ_{ij} should satisfy the additional condition

$$g_{lj}(\varphi^k_{ij;\,k} + \mu_{ij})_{;l} = 0, \tag{37}$$

because the divergences of the tensors G_{ij} and T_{ij} become zero.

Equation (37) imposes constraints on the dependence of $k_{\min}$ on particle coordinates and relative velocity (the latter can be expressed in terms of z).

The macroscopic energy–momentum tensor appears on the right-hand sides of the macroscopic Einstein equations. This tensor can be expressed in terms of the single-particle distribution functions F_b:

$$T_{ij} = \sum_b \int \frac{d^3p}{p^0\sqrt{-g}} p_i p_j F_b(p). \tag{38}$$

The kinetic equation for F_b, which can be obtained from (8) by the integration over p_0 and has the form (53) from [1], should be added to system of equations (11)–(38).

3. POSSIBLE APPLICATIONS OF THE THEORY

The derived gravitational field equations for continuous media differ from the classical Einstein equations by the additional terms

$$\varphi^k_{ij;\,k} + \mu_{ij}$$

on the left-hand side.

These terms are proportional to the Einstein constant cubed, but they are proportional to the particle density squared. Consequently, these additional terms can play a role only in high-density continuous media. Such densities are possible at early evolutionary stages of the Universe and inside objects close to gravitational collapse. Therefore, the first applications of the derived equations should naturally be sought in the theory of early evolutionary stages of the Universe and in the theory of gravitational collapse.

REFERENCES

1. A. V. Zakharov, Zh. Éksp. Teor. Fiz. **110**, 3 (1996) [JETP **83**, 1 (1996)].
2. A. V. Zakharov, Zh. Éksp. Teor. Fiz. **112**, 1153 (1997) [JETP **85**, 627 (1997)].
3. Yu. L. Klimontovich, Zh. Éksp. Teor. Fiz. **37**, 735 (1959) [Sov. Phys. JETP **10**, 535 (1959)].
4. A. V. Zakharov, Zh. Éksp. Teor. Fiz. **99**, 769 (1989) [Sov. Phys. JETP **69**, 437 (1989)].
5. G. S. Bisnovatyĭ-Kogan and I. G. Shukhman, Zh. Éksp. Teor. Fiz. **82**, 3 (1982) [Sov. Phys. JETP **55**, 1 (1982)].
6. A. V. Zakharov, Astron. Zh. **66**, 1208 (1989) [Sov. Astron. **33**, 624 (1989)].

Translated by V. Astakhov

Journal of Experimental and Theoretical Physics, Vol. 96, No. 4, 2003, pp. 594–609.
From Zhurnal Éksperimental'noĭ i Teoreticheskoĭ Fiziki, Vol. 123, No. 4, 2003, pp. 672–687.

GRAVITATION, ASTROPHYSICS

Radiation from Cosmic Chiral String Loops¶

E. Babichev and V. Dokuchaev

Institute for Nuclear Research, Russian Academy of Sciences, Moscow, 119312 Russia
e-mail: babichev@inr.npd.ac.ru, dokuchaev@inr.npd.ac.ru

Received November 13, 2002

Abstract—The gravitational and electromagnetic radiation from chiral superconducting cosmic string loops is calculated. The formulas for energy, momentum, and angular momentum losses due to gravitational and electromagnetic radiation from chiral loops of an arbitrary configuration are derived. After summation over all modes, expressions for the corresponding radiation rates averaged over the loop oscillation period have the form of four-dimensional integrals. These formulas are reduced to sums over the kinks for loops composed of piecewise linear strings. For three examples of string loops, the total radiation rates are calculated numerically depending on the chiral current along the string. In the limit of a nearly maximum current, which corresponds to a stationary loop (vorton) configuration, we determine the upper bounds on the gravitational and electromagnetic radiation. We also estimate the oscillation damping time of a nearly stationary loop. © *2003 MAIK "Nauka/Interperiodica".*

1. INTRODUCTION

We investigate the properties of the gravitational and electromagnetic radiation of energy, momentum, and angular momentum from superconducting closed cosmic strings with a chiral current. Formation of cosmic strings in early universe phase transitions is predicted by many particle-physics models (see, e.g., reviews in [1, 2]). In 1985, Witten showed that cosmic strings can carry a superconducting electromagnetic current [3]. Exact solutions of the equations of motion for current-carrying cosmic strings were found by Carter and Peter [4], Davis *et al.* [5], and Blanco-Pillado *et al.* [6] in the case of a chiral (or null) current $J^a J_a = 0$, which does not couple to any gauge field.

Ordinary cosmic strings (without a current) radiate energy [7–12], momentum [7, 13, 14], and angular momentum [14] in the form of gravitational waves. If cosmic strings carry the electromagnetic current, cosmic string loops radiate both gravitational and electromagnetic waves. For a small current, the most intense radiation is generated by a cusp on the loop. The radiation from a single cusp of the chiral string loop with a small current was studied by Blanco-Pillado and Olum [15]. The radiation of loops in the opposite case of a nearly maximum current was considered in [16]. In this paper, we study the gravitational and electromagnetic radiation from closed chiral string loops in the entire range of the string current. The rates of energy $\dot{E}$, momentum $\dot{P}$, and angular momentum $\dot{L}$ losses (averaged over the oscillation period) to the gravitational and electromagnetic waves can be expressed in the general form as

$$\begin{aligned} \dot{E}^{gr} &= \Gamma_E^{gr} G\mu^2, \quad \dot{P}^{gr} = \Gamma_P^{gr} G\mu^2, \\ \dot{L}^{gr} &= \Gamma_L^{gr} G\mu^2, \quad \dot{E}^{em} = \Gamma_E^{em} \mu q^2 \\ \dot{P}^{em} &= \Gamma_P^{em} \mu q^2, \quad \dot{L}^{em} = \Gamma_L^{em} \mathcal{L} \mu q^2, \end{aligned} \tag{1}$$

where the coefficients Γ_E^{gr}, Γ_P^{gr}, Γ_L^{gr}, Γ_E^{em}, Γ_P^{em}, and Γ_L^{em} depend on the particular string configuration and the current on the string, $\mathcal{L}$ is the string invariant length, μ is string mass per unit length, and q is the electromagnetic charge; we use the units $\hbar = c = 1$. In what follows, we calculate the coefficients Γ_E^{gr}, Γ_P^{gr}, Γ_L^{gr}, Γ_E^{em}, Γ_P^{em}, and Γ_L^{em} as functions of the current on the string. It is known that for ordinary loops (without a current), the corresponding coefficients for the gravitational radiation are of the respective orders $\Gamma_E^{gr} \sim 100$, $\Gamma_P^{gr} \sim 10$, and $\Gamma_L^{gr} \sim 10$. We found that for loops with a chiral current, the same coefficients Γ_E^{gr}, Γ_P^{gr}, and Γ_L^{gr} behave as follows: they rapidly decrease with the current at small current values and slowly decrease at large current values. In general, the gravitational radiation rates are decreasing functions of the current on the string. For the electromagnetic radiation, the situation is quite different: the loss rates of the energy, momentum, and angular momentum to electromagnetic waves for all examples considered have a maximum near some rather small value of the current.

The total rates of the energy, momentum, and angular momentum per unit time (averaged over the period)

¶ The article was submitted by the authors in English.

are usually calculated by summing the losses in different Fourier modes. As noted by Allen *et al.* [11], such calculations may not be accurate in practice because of a slow convergence of the corresponding sums over mode numbers. In this paper, we perform the summation over all radiation modes analytically and derive formulas for the energy, momentum, and angular momentum loss rates to the gravitational and electromagnetic radiation from the chiral string loops of a general configuration. As a result, the corresponding rates of radiation into the unit solid angle averaged over the loop oscillation period are reduced to four-dimensional integrals. In general, these integrals can be calculated only numerically. For chiral loops composed of piecewise linear strings, these formulas lead to analytic expressions for the energy, momentum, and angular momentum radiation into the unit solid angle. For large currents (close to the maximum value), we determine the upper bounds of the gravitational and electromagnetic radiation. For weak radial oscillations of a chiral ring, we find the temporal behavior of the loop energy and current analytically. For some other less symmetric loop examples, we estimate the damping time of small-amplitude loop oscillations.

This paper is organized as follows. In Section 2, we review some general properties of chiral cosmic strings. In Section 3, we derive new expressions for the energy, momentum, and angular momentum gravitational radiation rates by chiral loops of a general configuration into the unit solid angle. These expressions are reduced to four-dimensional integrals where summation over all radiation modes is performed analytically. In Section 4, we derive similar formulas for the electromagnetic radiation rates. In Section 5, the radiation and oscillation damping to the vorton state of nearly stationary loops are described. In Section 6, we present numerical calculations of the electromagnetic and gravitational radiation rates for some illustrative examples of chiral loops and study the dependence of the chiral string radiation on the current. In Section 7, we describe the results obtained and discuss some qualitative features of gravitational and electromagnetic radiation from chiral loops.

2. MOTION OF A CHIRAL STRING IN FLAT SPACE–TIME

In this section, for pedagogical reasons, we describe some general properties of chiral cosmic strings, i.e., strings with a chiral current of $J^a J_a = 0$. The general solution of the equations of motion of the chiral string can be written as [4–6]

$$x^0 = t, \quad \mathbf{x}(t, \sigma) = \frac{\mathscr{L}}{4\pi}[\mathbf{a}(\xi) + \mathbf{b}(\eta)], \tag{2}$$

where t is the Minkowski time, σ parameterizes the string total energy as

$$E = \mu \int d\sigma, \tag{3}$$

$\mathscr{L}$ is the invariant length of the string, and $\mathbf{a}(\xi)$ and $\mathbf{b}(\eta)$ are arbitrary vector functions of $\xi = (2\pi/\mathscr{L})(\sigma - t)$ and $\eta = (2\pi/\mathscr{L})(\sigma + t)$ satisfying the conditions

$$\mathbf{a}'^2 = 1, \quad \mathbf{b}'^2 = k^2(\eta) \le 1. \tag{4}$$

For closed chiral strings (loops), the vector functions $\mathbf{a}(\xi)$ and $\mathbf{b}(\eta)$ form closed loops, called *a*- and *b*-loops. The function $k(\eta)$ in (4) can be expressed as [6]

$$k^2(\eta) = 1 - \frac{4F'^2(\eta)}{\mu}, \tag{5}$$

where the function $F(\eta)$ defines the auxiliary scalar field

$$\phi(\sigma, t) = \frac{\mathscr{L}}{2\pi} F(\eta). \tag{6}$$

According to (6), the scalar field $\phi(\sigma, t)$ is an arbitrary function of the single parameter η. The four-dimensional current on the string is expressed through this scalar field $\phi(\sigma, t)$ as [21]

$$j^\mu(\mathbf{x}, t) = q \int d\sigma \phi'(\sigma, t)(x'^\mu - \dot{x}^\mu)\delta^{(3)}(\mathbf{x} - \mathbf{x}(\sigma, t)), \tag{7}$$

where x' denotes $\partial x/\partial\sigma$ and $\dot{x}$ denotes $\partial x/\partial t$. The energy-momentum tensor of the string in this gauge is

$$T^{\mu\nu} = \mu \int d\sigma (\dot{x}^\mu \dot{x}^\nu - x'^\mu x'^\nu)\delta^{(3)}(\mathbf{x} - \mathbf{x}(\sigma, t)). \tag{8}$$

Correspondingly, the total momentum and angular momentum of the string are given by

$$\mathbf{P} = \mu \int d\sigma \dot{\mathbf{x}}(\sigma, t), \tag{9}$$

$$\mathbf{L} = \mu \int d\sigma [\mathbf{x}(\sigma, t) \times \dot{\mathbf{x}}(\sigma, t)]. \tag{10}$$

3. GRAVITATIONAL RADIATION FROM CHIRAL LOOPS

We consider a periodic system with the period T. In this system, the Fourier transform of the energy-momentum tensor $T^{\mu\nu}(\mathbf{x}, t)$ is given by [14]

$$\hat{T}^{\mu\nu}(\omega_l, \mathbf{n}) = \frac{1}{T}\int_0^T dt \int d^3x T^{\mu\nu}(\mathbf{x}, t) \exp\{i\omega_l(t - \mathbf{n}\cdot\mathbf{x})\}, \tag{11}$$

where $\omega_l = 2\pi l/T$ and $\mathbf{n}$ is an arbitrary unit vector. It is

useful to also define the Fourier transform of the first moment,

$$\hat{T}^{\mu\nu p}(\omega_l, \mathbf{n}) = \frac{1}{T}\int dt \int d^3x T^{\mu\nu}(\mathbf{x}, t) x^p \exp\{i\omega_l(t - \mathbf{n}\cdot\mathbf{x})\}. \quad (12)$$

For convenience, we define the four-dimensional symbol $n^\mu \equiv (1, \mathbf{n})$. For any periodic system, the corresponding gravitational energy, momentum, and angular momentum radiation rates per solid angle $d\Omega$ (averaged over the period T) are given by the series

$$\frac{d\dot{P}^\mu}{d\Omega} = \sum_{n=1}^{\infty} \frac{d\dot{P}^\mu(\omega_n)}{d\Omega}, \quad \frac{d\dot{\mathbf{L}}}{d\Omega} = \sum_{n=1}^{\infty} \frac{d\dot{\mathbf{L}}(\omega_n)}{d\Omega}, \quad (13)$$

where [17]

$$\frac{d\dot{P}^\mu(\omega)}{d\Omega} = -n^\mu \frac{G\omega^2}{\pi} P_{ij} P_{lm} \left[\hat{T}^*_{il}\hat{T}_{jm} - \frac{1}{2}\hat{T}^*_{ij}\hat{T}_{lm}\right] \quad (14)$$

and [14]

$$\frac{d\dot{L}_i(\omega)}{d\Omega} = -\frac{G}{2\pi}\epsilon^{ijk} n^j \left[i\omega n^l P^{pq}\left(3\hat{T}^*_{kl}\hat{T}_{qp} + 6\hat{T}^*_{kp}\hat{T}_{ql}\right) + \omega^2 P^{lm}P^{pq}\left(2\hat{T}^*_{kmq}\hat{T}_{lp} - 2\hat{T}^*_{km}\hat{T}_{lpq} - \hat{T}^*_{lpk}\hat{T}_{mq} + \frac{1}{2}\hat{T}^*_{lmk}\hat{T}_{pq}\right) + \text{c.c}\right]. \quad (15)$$

Here, $P_{ij} = \delta_{ij} - n_i n_j$ is the projection operator to the plane perpendicular to the unit vector $\mathbf{n}$. It is possible to simplify (14) and (15) further by rewriting them in the corotating basis $(\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3) \equiv (\mathbf{n}, \mathbf{v}, \mathbf{w})$, where $\mathbf{v}$ and $\mathbf{w}$ are arbitrary unit vectors perpendicular to each other and to $\mathbf{n}$. In this corotating basis, Eqs. (14) and (15) become [14]

$$\frac{d\dot{P}^\mu(\omega)}{d\Omega} = n^\mu \frac{G\omega^2}{\pi}\left[\tau^*_{pq}\tau_{pq} - \frac{1}{2}\tau^*_{qq}\tau_{pp}\right], \quad (16)$$

$$\frac{d\dot{\mathbf{L}}(\omega)}{d\Omega} = \frac{d\dot{L}_2}{d\Omega}\mathbf{v} + \frac{d\dot{L}_3}{d\Omega}\mathbf{w}, \quad (17)$$

where

$$\frac{d\dot{L}_2}{d\Omega} = \frac{G}{2\pi}\left[-i\omega(3\tau^*_{13}\tau_{pp} + 6\tau^*_{3p}\tau_{p1}) - \omega^2\left(2\tau^*_{3pq}\tau_{pq} - 2\tau^*_{3p}\tau_{pqq} - \tau^*_{pq3}\tau_{pq} + \frac{1}{2}\tau^*_{qq3}\tau_{pp}\right) + \text{c.c}\right], \quad (18)$$

$$\frac{d\dot{L}_3}{d\Omega} = \frac{G}{2\pi}\left[i\omega(3\tau^*_{12}\tau_{pp} + 6\tau^*_{2p}\tau_{p1}) + \omega^2\left(2\tau^*_{2pq}\tau_{pq} - 2\tau^*_{2p}\tau_{pqq} - \tau^*_{pq2}\tau_{pq} + \frac{1}{2}\tau^*_{qq2}\tau_{pp}\right) + \text{c.c}\right].$$

Here, τ_{pq} and τ_{pqr} are the respective Fourier transforms of the energy-momentum tensor and its first moment in the new corotating basis. We note that only the subscripts p and q with the values 2 and 3 appear in Eqs. (16) and (18). For chiral loops, the Fourier transforms τ_{pq} can be expressed as

$$\tau_{pq}(\omega_l, \mathbf{n}) = -\frac{\mathcal{L}\mu}{2}[I_p(l)Y_q(l) + Y_p(l)I_q(l)], \quad (19)$$

where the functions $I_p(l)$ and $Y_q(l)$ are expressed through the "fundamental integrals,"

$$I_i(l) \equiv \frac{1}{2\pi}\int_0^{2\pi} d\xi \exp\{-il(\xi + \mathbf{n}\cdot\mathbf{a})\}\mathbf{a}'\cdot\mathbf{e}_i,$$
$$Y_j(l) \equiv \frac{1}{2\pi}\int_0^{2\pi} d\eta \exp\{il(\eta - \mathbf{n}\cdot\mathbf{b})\}\mathbf{b}'\cdot\mathbf{e}_j. \quad (20)$$

For the first moment (12), we can similarly find that

$$\tau_{ijk}(\omega_l, \mathbf{n}) = -\frac{\mathcal{L}\mu}{8\pi}[I_i(l)N_{jk}(l) + I_j(l)N_{ik}(l) + Y_i(l)M_{jk}(l) + Y_j(l)M_{ik}(l)], \quad (21)$$

where

$$M_{ij}(l) \equiv \frac{1}{2\pi}\int_0^{2\pi} d\xi \exp\{-il(\xi + \mathbf{n}\cdot\mathbf{a})\}(\mathbf{a}'\cdot\mathbf{e}_i)(\mathbf{a}\cdot\mathbf{e}_j),$$
$$N_{ij}(l) \equiv \frac{1}{2\pi}\int_0^{2\pi} d\eta \exp\{il(\eta - \mathbf{n}\cdot\mathbf{b})\}(\mathbf{b}'\cdot\mathbf{e}_i)(\mathbf{b}\cdot\mathbf{e}_j). \quad (22)$$

The crucial point of the calculation to follow is the summation over all mode numbers l in expressions (13) for the requested rates of the radiated gravitational energy, momentum, and angular momentum. For this summation, we first integrate expressions (20) and (22) by parts to obtain an additional l in the denominator. For example, the function I_i becomes

$$I_i(l) = \frac{1}{2\pi}\int_0^{2\pi} d\xi [\exp\{-il(\xi + \mathbf{n}\cdot\mathbf{a})\}(1 + \mathbf{n}\cdot\mathbf{a}')] \times \frac{\mathbf{a}'\cdot\mathbf{e}_i}{1 + \mathbf{n}\cdot\mathbf{a}'} = -\frac{1}{2\pi i l} \times \frac{\mathbf{a}'\cdot\mathbf{e}_i}{1 + \mathbf{n}\cdot\mathbf{a}'}\exp\{-il(\xi + \mathbf{n}\cdot\mathbf{a})\}\Big|_0^{2\pi} + \frac{1}{2\pi i l}\int_0^{2\pi} d\xi \left[\frac{\mathbf{a}'\cdot\mathbf{e}_j}{1 + \mathbf{n}\cdot\mathbf{a}'}\right]' \exp\{-il(\xi + \mathbf{n}\cdot\mathbf{a})\}, \quad (23)$$

where the first term is equal to zero because of the periodicity of a- and b-loops. Expressions for the functions

Y_j, M_{ij}, and N_{ij} can be integrated by parts similarly. We finally obtain

$$I_i = \frac{1}{2\pi i l}\int_0^{2\pi} d\xi \mathcal{I}_i \exp\{-il(\xi + \mathbf{n}\cdot\mathbf{a})\},$$

$$Y_j = -\frac{1}{2\pi i l}\int_0^{2\pi} d\eta \mathcal{Y}_j \exp\{il(\eta - \mathbf{n}\cdot\mathbf{b})\},$$

$$M_{ij} = \int_0^{2\pi} d\xi\left(\frac{1}{2\pi i l}\mathcal{M}_{ij} - \frac{1}{2\pi l^2}\tilde{\mathcal{M}}_{ij}\right)\exp\{-il(\xi + \mathbf{n}\cdot\mathbf{a})\}, \tag{24}$$

$$N_{ij} = -\int_0^{2\pi} d\eta\left(\frac{1}{2\pi i l}\mathcal{N}_{ij} + \frac{1}{2\pi l^2}\tilde{\mathcal{N}}_{ij}\right)\exp\{il(\eta - \mathbf{n}\cdot\mathbf{b})\},$$

where

$$\mathcal{I}_i = \left[\frac{\mathbf{a}'\cdot\mathbf{e}_i}{1+\mathbf{n}\cdot\mathbf{a}'}\right]', \quad \mathcal{Y}_j = \left[\frac{\mathbf{b}'\cdot\mathbf{e}_j}{1-\mathbf{n}\cdot\mathbf{b}'}\right]',$$

$$\mathcal{M}_{ij} = \left[\frac{\mathbf{a}'\cdot\mathbf{e}_i}{1+\mathbf{n}\cdot\mathbf{a}'}\right]'(\mathbf{a}\cdot\mathbf{e}_j),$$

$$\tilde{\mathcal{M}}_{ij} = \left[\frac{(\mathbf{a}'\cdot\mathbf{e}_i)(\mathbf{a}'\cdot\mathbf{e}_j)}{(1+\mathbf{n}\cdot\mathbf{a}')^2}\right]', \tag{25}$$

$$\mathcal{N}_{ij} = \left[\frac{\mathbf{b}'\cdot\mathbf{e}_i}{1-\mathbf{n}\cdot\mathbf{b}'}\right]'(\mathbf{b}\cdot\mathbf{e}_j),$$

$$\tilde{\mathcal{N}}_{ij} = \left[\frac{(\mathbf{b}'\cdot\mathbf{e}_i)(\mathbf{b}'\cdot\mathbf{e}_j)}{(1-\mathbf{n}\cdot\mathbf{b}')^2}\right]'.$$

Substituting (24) in (19) and (21), we find

$$\tau_{ij} = -\frac{\mathcal{L}\mu}{8\pi^2 l^2}\int_0^{2\pi}\int_0^{2\pi} d\xi d\eta \mathcal{T}_{ij} \times \exp\{-il[\xi - \eta + \mathbf{n}\cdot(\mathbf{a}+\mathbf{b})]\},$$

$$\tau_{ijk} = -\frac{\mathcal{L}^2\mu}{32\pi^3 l^2}\int_0^{2\pi}\int_0^{2\pi} d\xi d\eta\left(\mathcal{T}_{ijk} + \frac{1}{il}\tilde{\mathcal{T}}_{ijk}\right) \times \exp\{-il[\xi - \eta + \mathbf{n}\cdot(\mathbf{a}+\mathbf{b})]\}, \tag{26}$$

where

$$\mathcal{T}_{ij} = \mathcal{I}_i\mathcal{Y}_j + \mathcal{I}_j\mathcal{Y}_i,$$
$$\mathcal{T}_{ijk} = \mathcal{I}_i\mathcal{N}_{jk} + \mathcal{I}_j\mathcal{N}_{ik} + \mathcal{Y}_i\mathcal{M}_{jk} + \mathcal{Y}_j\mathcal{M}_{ik}, \tag{27}$$
$$\tilde{\mathcal{T}}_{ijk} = -\mathcal{I}_i\tilde{\mathcal{N}}_{jk} - \mathcal{I}_j\tilde{\mathcal{N}}_{ik} + \mathcal{Y}_i\tilde{\mathcal{M}}_{jk} + \mathcal{Y}_j\mathcal{M}_{ik}.$$

Next, substituting (26) in (16) and (18), we find the radiation rates of E, $\mathbf{P}$, and $\mathbf{L}$ on the particular eigenfrequency $\omega_l = 2\pi l/T$,

$$\frac{d\dot{P}^\mu(\omega)}{d\Omega} = n^\mu\frac{G\mu^2}{4\pi^3 l^2}\int d^4\xi \mathcal{P}\cos(l\Delta x), \tag{28}$$

$$\frac{d\dot{L}_v}{d\Omega} = -\frac{G\mathcal{L}\mu^2}{16\pi^4} \times\int d^4\xi\left[\frac{\sin(l\Delta x)}{l^3}(3\lambda_2 + \tilde{\Lambda}_2) + \frac{\cos(l\Delta x)}{l^2}\Lambda_2\right],$$

$$\frac{d\dot{L}_w}{d\Omega} = \frac{G\mathcal{L}\mu^2}{16\pi^4} \times\int d^4\xi\left[\frac{\sin(l\Delta x)}{l^3}(3\lambda_3 + \tilde{\Lambda}_3) + \frac{\cos(l\Delta x)}{l^2}\Lambda_3\right], \tag{29}$$

where we use the notation

$$\Delta x = \xi - \xi' - (\eta - \eta') + \mathbf{n}[\mathbf{a}(\xi) - \mathbf{a}(\xi') + \mathbf{b}(\eta) - \mathbf{b}(\eta')],$$

$$\mathcal{P} = \mathcal{T}'_{pq}\mathcal{T}_{pq} - \frac{1}{2}\mathcal{T}'_{qq}\mathcal{T}_{pp},$$

$$\lambda_2 = \mathcal{T}'_{13}\mathcal{T}_{pp} + 2\mathcal{T}'_{3p}\mathcal{T}_{p1},$$
$$\lambda_3 = \mathcal{T}'_{12}\mathcal{T}_{pp} + 2\mathcal{T}'_{2p}\mathcal{T}_{p1}, \tag{30}$$

$$\Lambda_2 = 2\mathcal{T}'_{3pq}\mathcal{T}_{pq} - 2\mathcal{T}'_{3p}\mathcal{T}_{pqq} - \mathcal{T}'_{pq3}\mathcal{T}_{pq} + \frac{1}{2}\mathcal{T}'_{qq3}\mathcal{T}_{pp},$$

$$\tilde{\Lambda}_2 = 2\tilde{\mathcal{T}}'_{3pq}\mathcal{T}_{pq} + 2\mathcal{T}'_{3p}\tilde{\mathcal{T}}_{pqq} - \tilde{\mathcal{T}}'_{pq3}\mathcal{T}_{pq} + \frac{1}{2}\tilde{\mathcal{T}}'_{qq3}\mathcal{T}_{pp},$$

$$\Lambda_3 = 2\mathcal{T}'_{2pq}\mathcal{T}_{pq} - 2\mathcal{T}'_{2p}\mathcal{T}_{pqq} - \mathcal{T}'_{pq2}\mathcal{T}_{pq} + \frac{1}{2}\mathcal{T}'_{qq2}\mathcal{T}_{pp},$$

$$\tilde{\Lambda}_3 = 2\tilde{\mathcal{T}}'_{2pq}\mathcal{T}_{pq} + 2\mathcal{T}'_{2p}\tilde{\mathcal{T}}_{pqq} - \tilde{\mathcal{T}}\mathrm{J}_{pq2}\mathcal{T}_{pq} + \frac{1}{2}\tilde{\mathcal{T}}'_{qq2}\mathcal{T}_{pp}.$$

It is assumed that integration in (28) and (29) is over the four-dimensional cube with the side $(0, 2\pi)$; we also use the notation $d^4\xi = d\xi d\xi' d\eta d\eta'$.

We now find the form of expressions (28) and (29) suitable for summing over the modes l. Using the known values for infinite series [18]

$$\sum_{l=1}^{\infty}\frac{\cos(lx)}{l^2} = \frac{1}{4}(x-\pi)^2 - \frac{\pi^2}{12}, \quad 0 \le x \le 2\pi,$$

$$\sum_{l=1}^{\infty}\frac{\sin(lx)}{l^3} = \frac{1}{12}[(x-\pi)^3 - \pi^2 x] + \frac{\pi^3}{12}, \tag{31}$$

$$0 \le x \le 2\pi,$$

we obtain the final expressions for the gravitational radiation of energy, momentum, and angular momen-

tum rates [19] from (28) and (29) as

$$\frac{d\dot{P}^\mu}{d\Omega} = n^\mu \frac{G\mu^2}{16\pi^3}\int d^4\xi \mathcal{P}(\Delta x \bmod 2\pi - \pi)^2, \quad (32)$$

$$\frac{d\dot{L}^\nu}{d\Omega} = -\frac{G\mathcal{L}\mu^2}{64\pi^4}\int d^4\xi \left\{ [(\Delta x \bmod 2\pi - \pi)^3 - \pi^2 \Delta x \bmod 2\pi]\left(\lambda_2 + \frac{1}{3}\tilde{\Lambda}_2\right) + (\Delta x \bmod 2\pi - \pi)^2 \Lambda_2 \right\}, \quad (33)$$

$$\frac{d\dot{L}_w}{d\Omega} = \frac{G\mathcal{L}\mu^2}{64\pi^4}\int d^4\xi \left\{ [(\Delta x \bmod 2\pi - \pi)^3 - \pi^2 \Delta x \bmod 2\pi]\left(\lambda_3 + \frac{1}{3}\tilde{\Lambda}_3\right) + (\Delta x \bmod 2\pi - \pi)^2 \Lambda_3 \right\}.$$

We note that the integrals in (32) and (33) do not contain the terms $\pi^2/12$ and $\pi^3/12$ originating in (31) because the corresponding contributions vanish in the integrals. The advantage of formulas (32) and (33) over the corresponding formulas (16) and (17) is that there are no summations over modes. However, because of the presence of the function $\Delta x \bmod 2\pi$, the four-dimensional integrals in (32) and (33) cannot be reduced to products of lower dimensional integrals, and therefore numerical calculations of the four-dimensional integrals become more complicated.

4. ELECTROMAGNETIC RADIATION FROM CHIRAL LOOPS

We now consider the electromagnetic radiation from an arbitrary relativistic periodic system in a similar way. We calculate the electromagnetic radiation by analogy with Durrer's calculations of the gravitational radiation [14]. In the Lorentz gauge, a retarded solution for the electromagnetic potential A_μ in such a system is given by

$$A_\mu(\mathbf{x}, t) = -\int \frac{j_\mu(\mathbf{x}', t_{\rm ret})}{|\mathbf{x} - \mathbf{x}'|} d\mathbf{x}', \quad (34)$$

where j_μ is the four-dimensional current, and we set $t_{\rm ret} = t - |\mathbf{x} - \mathbf{x}'|$ We consider formula (34) in the limit $r = |\mathbf{x}| \gg |\mathbf{x}'|$. Expanding (34) in a series in $1/r$ and taking the first two terms into account, we obtain

$$A_\mu(\mathbf{x}, t) = \frac{1}{r}\int j_\mu(\mathbf{x}', t_{\rm ret}) d\mathbf{x}' - \frac{1}{r^2}\int j_\mu(\mathbf{x}', t_{\rm ret}) x'^i d\mathbf{x}' + O(r^{-3}), \quad (35)$$

where $\mathbf{n} = \mathbf{x}/r$. Expanding $t_{\rm ret}$ in a series in $|\mathbf{x}'|/r$, we then find

$$t_{\rm ret} = t - r + \mathbf{n}\cdot\mathbf{x}' - \frac{1}{2r}P_{ij}x'^i x'^j + O(|\mathbf{x}'|^2/r^2)|\mathbf{x}'|. \quad (36)$$

Equation (36) implies the useful relation

$$A_{\mu,j} = -A_{\mu,0}n_j + O(A_\mu/r). \quad (37)$$

Similarly to the case of the gravitational field ($T^{\mu\nu} \longleftrightarrow j^\mu$, $h^{\mu\nu} \longleftrightarrow j^\mu$, etc.), we have the Fourier transforms of the current $\tilde{j}^\mu$ and its first and second moments $\tilde{j}^{\mu p}$ and $\tilde{j}^{\mu pq}$,

$$\tilde{j}^\mu(\omega_l, \mathbf{n}) = \frac{1}{T}\int_0^T dt \int d^3x j^\mu(\omega_l, \mathbf{x}) \exp\{i\omega_l(t - \mathbf{n}\cdot\mathbf{x})\},$$

$$\tilde{j}^{\mu p}(\omega_l, \mathbf{n}) = \frac{1}{T}\int_0^T dt \int d^3x j^\mu(\omega_l, \mathbf{x}) x^p \times \exp\{i\omega_l(t - \mathbf{n}\cdot\mathbf{x})\}, \quad (38)$$

$$\tilde{j}^{\mu pq}(\omega_l, \mathbf{n}) = \frac{1}{T}\int_0^T dt \int d^3x j^\mu(\omega_l, \mathbf{x}) x^p x^q \times \exp\{i\omega_l(t - \mathbf{n}\cdot\mathbf{x})\}.$$

These quantities satisfy the conditions

$$\tilde{j}^0 - n^k \tilde{j}^k = 0,$$
$$-i\omega\tilde{j}^{0p} - \tilde{j}^p + i\omega n_k \tilde{j}^{kp} = 0, \quad (39)$$
$$i\omega P_{mn}(\tilde{j}^{0mn} - n_p \tilde{j}^{pmn}) + 2P_{pq}\tilde{j}^{pq} = 0,$$

which follow from the relations

$$j^\mu_{,\mu} = 0,$$
$$\int j^\mu(t, \mathbf{x}')_{,\mu}[x'^p \exp\{i\omega(t - \mathbf{n}\cdot\mathbf{x}')\}] dt d^3x = 0,$$
$$\int j^\mu(t, \mathbf{x}')_{,\mu}\{[x'^2 - (\mathbf{n}\cdot\mathbf{x}')] \times \exp\{i\omega(t - \mathbf{n}\cdot\mathbf{x}')\} dt d^3x = 0. \quad (40)$$

Using (38) and (40), we obtain from (35)

$$A_\mu(\mathbf{x}, t) = \frac{1}{r}\sum_{l=1}^{\infty} e^{-i\omega_l(t-r)} \times \left[\tilde{j}_\mu(\omega_l, \mathbf{n}) + \frac{n^p}{r}\tilde{j}_{\mu p}(\omega_l, \mathbf{n}) + \frac{i\omega_l}{2r}P^{pq}\tilde{j}_{\mu pq}(\omega_l, \mathbf{n})\right] + \text{c.c.} + O(r^{-3}). \quad (41)$$

To calculate the energy and momentum radiation losses, we keep only terms on the order of $1/r$ in (41).

The radiation of energy from the system is determined by the Poynting vector as [20]

$$\frac{d\dot{E}^{em}}{d\Omega} = \frac{|\mathbf{E}\times\mathbf{H}|}{4\pi}, \tag{42}$$

where **E** and **H** are the electric and magnetic fields. Using (41), we obtain from (42)

$$\frac{d\dot{P}^{\mu}_{em}}{d\Omega} = \sum_{n=1}^{\infty}\frac{d\dot{P}^{\mu}(\omega_n)}{d\Omega}, \tag{43}$$

where

$$\frac{d\dot{P}^{\mu}_{em}(\omega)}{d\Omega} = n^{\mu}\frac{\omega^2}{2\pi}P^{pq}\tilde{j}^{*}_{p}\tilde{j}_{q}. \tag{44}$$

We now calculate electromagnetic radiation of the angular momentum. The angular momentum rate per unit solid angle is given by [20]

$$\frac{d\dot{\mathbf{L}}^{em}}{d\Omega} = \frac{r^3}{4\pi}[[\mathbf{n}\times\mathbf{E}](\mathbf{n}\cdot\mathbf{E}) + [\mathbf{n}\times\mathbf{H}](\mathbf{n}\cdot\mathbf{H})]. \tag{45}$$

In calculating $[\mathbf{n}\times\mathbf{E}]$ and $[\mathbf{n}\times\mathbf{H}]$, it suffices to keep only terms on the order of $1/r$, but the longitudinal components $\mathbf{n}\cdot\mathbf{E}$ and $\mathbf{n}\cdot\mathbf{H}$ arise from terms on the order of $1/r^2$. As a result, the term r^3 is canceled in (45). This implies that the distance from the system r does not enter the final formula, as should be the case. Using (41) and (40), we obtain

$$[\mathbf{n}\times\mathbf{E}]^i = -\epsilon^{ijk}n_j A_{k,0}$$

$$= -\sum_{l=1}^{\infty}\frac{i\omega_l}{r}\exp\{-i\omega_l(t-r)\}\epsilon^{ijk}n_j\tilde{j}_k + \text{c.c.},$$

$$[\mathbf{n}\times\mathbf{H}]^i = \sum_{l=1}^{\infty}\frac{i\omega_l}{r}\exp\{-i\omega_l(t-r)\}P^{ij}\tilde{j}_k + \text{c.c.}, \tag{46}$$

$$\mathbf{n}\cdot\mathbf{E} = -\sum_{l=1}^{\infty}\frac{i\omega_l}{r^2}\exp\{-i\omega_l(t-r)\}P^{pq}\tilde{j}_{pq} + \text{c.c.},$$

$$\mathbf{n}\cdot\mathbf{H} = \sum_{l=1}^{\infty}\frac{i\omega_l}{r^2}\exp\{-i\omega_l(t-r)\}\epsilon^{pqr}n_p\tilde{j}_{rq} + \text{c.c.}$$

Substituting (46) in (45), we obtain

$$\frac{d\dot{\mathbf{L}}^{em}}{d\Omega} = \sum_{n=1}^{\infty}\frac{d\dot{\mathbf{L}}^{em}(\omega_n)}{d\Omega}, \tag{47}$$

where

$$\frac{d\dot{L}^{em}_{i}(\omega)}{d\Omega} = \frac{\omega^2}{4\pi}[(\epsilon^{ijk}P_{pq} - P_{ik}\epsilon_{ipq})n_j\tilde{j}^{*}_{k}\tilde{j}_{pq} + \text{c.c.}]. \tag{48}$$

As for the gravitational field, we rewrite (44) and (48) in the corotating basis $(\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3) = (\mathbf{n}, \mathbf{v}, \mathbf{w})$,

$$\frac{d\dot{P}^{\mu}_{em}(\omega)}{d\Omega} = n^{\mu}\frac{\omega^2}{2\pi}\tilde{\iota}^{*}_{p}\tilde{\iota}_{p}, \tag{49}$$

$$\frac{d\dot{\mathbf{L}}^{em}(\omega)}{d\Omega} = \frac{d\dot{L}^{em}_{2}}{d\Omega}\mathbf{v} + \frac{d\dot{L}^{em}_{3}}{d\Omega}\mathbf{w}, \tag{50}$$

where

$$\frac{d\dot{L}^{em}_{2}}{d\Omega} = -\frac{\omega^2}{4\pi}[\iota^{*}_{3}\iota_{pp} + \iota^{*}_{2}(\iota_{23} - \iota_{32}) + \text{c.c.}],$$
$$\frac{d\dot{L}^{em}_{3}}{d\Omega} = \frac{\omega^2}{4\pi}[\iota^{*}_{2}\iota_{pp} - \iota^{*}_{3}(\iota_{23} - \iota_{32}) + \text{c.c.}], \tag{51}$$

and ι^{p} and ι^{pq} are components of j^{μ} and $j^{\mu p}$ in this corotating basis.

For superconducting chiral strings, we obtain from expression (7) for the current that

$$\iota_i(\omega_l, \mathbf{n}) = \frac{\mathscr{L}q\sqrt{\mu}}{2}[I_i(l)X(l)], \tag{52}$$

where the function $I_i(l)$ is given by (20) and $X(l)$ is

$$X(l) \equiv \frac{1}{2\pi}\int_0^{2\pi} d\eta e^{il(\eta - \mathbf{n}\cdot\mathbf{b})}\sqrt{1 - |\mathbf{b}'|^2}. \tag{53}$$

Similarly, for the first moment ι_{pq}, we obtain

$$\iota_{pq}(\omega_l, \mathbf{n}) = \frac{\mathscr{L}^2 q\sqrt{\mu}}{8\pi}[I_p(l)Z_q(l) + X(l)M_{pq}(l)], \tag{54}$$

where M_{pq} is given by (22) and Z_q is

$$Z_i(l) \equiv \frac{1}{2\pi}\int_0^{2\pi} d\eta e^{il(\eta - \mathbf{n}\cdot\mathbf{b})}\sqrt{1 - |\mathbf{b}'|^2}(\mathbf{b}\cdot\mathbf{e}_i). \tag{55}$$

We now integrate expressions (53) and (55) by parts to

obtain an additional l in the denominator,

$$X = -\frac{1}{2\pi i l}\int_0^{2\pi} d\eta \mathcal{X}\exp\{il(\eta-\mathbf{n}\cdot\mathbf{b})\},$$
$$Z_j = -\int_0^{2\pi} d\eta\left(\frac{1}{2\pi i l}\mathcal{Z}_j+\frac{1}{2\pi l^2}\tilde{\mathcal{Z}}_j\right)\times\exp\{il(\eta-\mathbf{n}\cdot\mathbf{b})\}, \quad (56)$$

where

$$\mathcal{X} = \left[\frac{\sqrt{1-|\mathbf{b}'|^2}}{1-\mathbf{n}\cdot\mathbf{b}'}\right]',$$
$$\mathcal{Z}_j = \left[\frac{\sqrt{1-|\mathbf{b}'|^2}}{1-\mathbf{n}\cdot\mathbf{b}'}\right]'(\mathbf{b}\mathbf{e}_j), \quad (57)$$
$$\tilde{\mathcal{Z}}_j = \left[\frac{\sqrt{1-|\mathbf{b}'|^2}(\mathbf{b}'\cdot\mathbf{e}_j)}{1-\mathbf{n}\cdot\mathbf{b}'}\right]'.$$

Substituting (20), (22), and (56) in (52) and (54), we obtain

$$\iota_i = \frac{\mathcal{L}q\sqrt{\mu}}{8\pi^2 l^2}\int_0^{2\pi}\int_0^{2\pi} d\xi d\eta \mathcal{J}_i\exp\{-il[\xi-\eta+\mathbf{n}\cdot(\mathbf{a}+\mathbf{b})]\}$$
$$\iota_{ij} = \frac{\mathcal{L}^2 q\sqrt{\mu}}{32\pi^3 l^2}\int_0^{2\pi}\int_0^{2\pi} d\xi d\eta\left(\mathcal{J}_{ij}+\frac{1}{il}\tilde{\mathcal{J}}_{ij}\right)\times\exp\{-il[\xi-\eta+\mathbf{n}\cdot(\mathbf{a}+\mathbf{b})]\}, \quad (58)$$

where

$$\mathcal{J}_i = \mathcal{I}_i\mathcal{X},\quad \mathcal{J}_{ij} = \mathcal{I}_i\mathcal{Z}_j+\mathcal{X}\mathcal{M}_{ij},$$
$$\tilde{\mathcal{J}}_{ij} = -\mathcal{I}_i\tilde{\mathcal{Z}}_j+\mathcal{X}\tilde{\mathcal{M}}_{ij}. \quad (59)$$

Finally, substituting (58) in (49) and (51), we find the expressions for electromagnetic radiation rates of the energy, momentum, and angular momentum in a unit solid angle at the frequency ω_l,

$$\frac{d\dot{P}^{\mu}_{em}(\omega_l)}{d\Omega} = n^{\mu}\frac{q^2\mu}{8\pi^3 l^2}\int d^4\xi\mathcal{P}^{em}\cos(l\Delta x), \quad (60)$$

$$\frac{d\dot{L}^{em}_{v}}{d\Omega} = \frac{\mathcal{L}q^2\mu}{32\pi^4}\times\int d^4\xi\left[\frac{\sin(l\Delta x)}{l^3}\tilde{\Lambda}_2^{em}-\frac{\cos(l\Delta x)}{l^2}\Lambda_2^{em}\right],$$
$$\frac{d\dot{L}^{em}_{w}}{d\Omega} = -\frac{\mathcal{L}q^2\mu}{32\pi^4}\times\int d^4\xi\left[\frac{\sin(l\Delta x)}{l^3}\tilde{\Lambda}_3^{em}-\frac{\cos(l\Delta x)}{l^2}\Lambda_3^{em}\right], \quad (61)$$

where

$$\mathcal{P}^{em} = \mathcal{J}'_p\mathcal{J}_p,$$
$$\Lambda_2^{em} = \mathcal{J}'_3\mathcal{J}_{pp}+\mathcal{J}'_2(\mathcal{J}_{23}-\mathcal{J}_{32}),$$
$$\tilde{\Lambda}_2^{em} = \mathcal{J}'_3\tilde{\mathcal{J}}_{pp}+\mathcal{J}'_2(\tilde{\mathcal{J}}_{23}-\tilde{\mathcal{J}}_{32}), \quad (62)$$
$$\Lambda_3^{em} = \mathcal{J}'_2\mathcal{J}_{pp}-\mathcal{J}'_3(\mathcal{J}_{23}-\mathcal{J}_{32}),$$
$$\tilde{\Lambda}_3^{em} = \mathcal{J}'_2\tilde{\mathcal{J}}_{pp}-\mathcal{J}'_3(\tilde{\mathcal{J}}_{23}-\tilde{\mathcal{J}}_{32}).$$

As for the gravitational radiation, we use the values for infinite series (31) to obtain the total electromagnetic radiation rates of the energy, momentum, and angular momentum [19],

$$\frac{d\dot{P}^{\mu}_{em}}{d\Omega} = n^{\mu}\frac{q^2\mu}{32\pi^3}\int d^4\xi\mathcal{P}^{em}(\Delta x\bmod 2\pi-\pi)^2, \quad (63)$$

$$\frac{d\dot{L}^{em}_{v}}{d\Omega} = \frac{\mathcal{L}q^2\mu}{128\pi^4}\int d^4\xi\times\left[\frac{1}{3}((\Delta x\bmod 2\pi-\pi)^3-\pi^2\Delta x\bmod 2\pi)\tilde{\Lambda}_2^{em}-(\Delta x\bmod 2\pi-\pi)^2\Lambda_2^{em}\right],$$
$$\frac{d\dot{L}^{em}_{w}}{d\Omega} = -\frac{\mathcal{L}q^2\mu}{128\pi^4}\int d^4\xi\times\left[\frac{1}{3}((\Delta x\bmod 2\pi-\pi)^3-\pi^2\Delta x\bmod 2\pi)\tilde{\Lambda}_3^{em}-(\Delta x\bmod 2\pi-\pi)^2\Lambda_3^{em}\right]. \quad (64)$$

As a result we found expressions for the electromagnetically radiated energy, momentum, and angular momentum from chiral string loops in which the summation over modes l is carried out.

5. RADIATION OF NEARLY STATIONARY LOOPS

We can now consider small-amplitude oscillations of the chiral string loop (i.e., a string that is close to its vorton state) in more detail. An arbitrary function $b(\eta)$ in the solution for string motion (2) is then such that $b'(\eta) = k(\eta) \ll 1$. If three-dimensional coordinates are chosen such that the b-loop is near the origin of the coordinate system (e.g., exactly intersects the origin of the coordinate system), then $b(\eta) \ll 1$. We now return to expressions (13). The expressions for the functions $Y_j(l)$ and $N_{ij}(l)$ can be integrated by parts twice to increase the power of l in the denominator, while the expressions for $I_j(l)$ and $M_{ij}(l)$ are left unchanged. We

assume that $b(\eta)$ is twice continuously differentiable and $b'''(\eta)$ is piecewise continuous. Integrating in (20) and (22) by parts twice and using the smallness of $b'(\eta)$, we obtain

$$Y_i(l) = -\frac{1}{2\pi l^2}\int_0^{2\pi} d\eta e^{il\eta}\mathbf{b}''' \cdot \mathbf{e}_i,$$
$$N_{ij}(l) = -\frac{1}{2\pi l^2}\int_0^{2\pi} d\eta e^{il\eta}[(\mathbf{b}' \cdot \mathbf{e}_i)(\mathbf{b} \cdot \mathbf{e}_j)]''. \quad (65)$$

As has been noted, we are free to add any numerical coefficient to $(\Delta x \bmod 2\pi - \pi)^2$ in (32) without changing the value of the integral. Using this, we add $-\pi^2/2$, and then Eq. (32) implies

$$\left|\frac{d\dot{P}^\mu}{d\Omega}\right| = \left|n^\mu \frac{G\mu^2}{16\pi^3}\int d^4\xi \mathcal{P}[(\Delta x \bmod 2\pi - \pi)^2 - \pi^2/2]\right| \quad (66)$$
$$\leq \frac{G\mu^2\pi^3}{2}|\mathcal{P}|.$$

It only remains to estimate the function $\mathcal{P}$ in (66). Using (20), (27), (30), and (65) we easily find

$$|\mathcal{P}| \leq 12 b_3^2, \quad (67)$$

where b_3 is the maximum value of $|b'''(\eta)|$ on the segment $\eta \in (0, 2\pi)$. From (66) and (67), we then estimate the energy losses as

$$\left|\dot{E}^{gr}\right| \leq 24 G\mu^2\pi^4 b_3^2, \quad \left|\dot{\mathbf{P}}^{gr}\right| \leq 24 G\mu^2\pi^4 b_3^2. \quad (68)$$

We next estimate the upper bounds of the radiated angular momentum. Similarly to the case of energy and momentum radiation, we use Eqs. (33), (30), (27), and (65) to find the upper bounds of losses of the angular momentum to gravitational waves,

$$\left|\dot{L}^{gr}\right| \leq 12\sqrt{2}\pi^4\left(1 + \frac{4}{3\sqrt{3}}\right) G\mathcal{L}\mu^2 b_3^2. \quad (69)$$

We now consider the electromagnetic radiation in the case of a large current. To find the first-order expansion with respect to k in (53), we must take not only zero, but also the first term in the expansion of $\exp(-il\mathbf{n} \cdot \mathbf{b})$, into account. Subsequent integration of the resulting expression by parts gives

$$X(l) \simeq -\frac{1}{2\pi l^2}\int_0^{2\pi} d\eta e^{il\eta}\mathbf{n} \cdot \mathbf{b}'''. \quad (70)$$

For the function Z_i, we have

$$Z_i(l) \approx -\frac{1}{2\pi l^2}\int_0^{2\pi} d\eta e^{il\eta}(\mathbf{b}'' \cdot \mathbf{e}_i). \quad (71)$$

Similarly to the gravitational case, we can find the bounds of the electromagnetic radiation for a large current. Using (70), (71), (64), (63), (62), and (59), we obtain

$$\left|\dot{E}^{em}\right| \leq \pi^4 q^2 \mu b_3^2, \quad \left|\dot{P}^{em}\right| \leq \pi^4 q^2 \mu b_3^2,$$
$$\left|\dot{L}^{em}\right| \leq \sqrt{2}\pi^3\left(1 + \frac{4\pi}{9\sqrt{3}}\right)\mathcal{L}q^2\mu b_3^2. \quad (72)$$

The presence of the third derivative $\mathbf{b}'''(\eta)$ in (68), (69), and (65) is not surprising and resembles the quadruple gravitational radiation formula (see, e.g., [20])

$$\dot{E} = \frac{G}{45}\dddot{D}_{ij}^2 \quad (73)$$

involving the third time derivative of the quadruple moment D_{ij}. Electromagnetic radiation involves $\ddot{\mathbf{d}}$ in the dipole approach ($\mathbf{d}$ is the dipole moment). Arguing similarly, we can conclude that in this case, the second derivative of $\mathbf{b}(\eta)$, not the third, must be restricted. However, in the first order of the expansion in k, the dipole radiation is equal to zero, the first nonzero term is quadruple, and we therefore again obtain the dependence on $\mathbf{b}'''$.

We note that it is not necessary to restrict the third derivative $\mathbf{b}'''$ in general. For example, if the string has kinks (see below), the first derivative $\mathbf{b}'$ is discontinuous (and consequently, $Y_p(l) \propto 1/l$, $M(l) \propto 1/l$). Convergence of series (13), (43), and (47) is then ensured by the behavior of fundamental integrals $I_p(l) \propto 1/l$ at $l \gg 1$.

It is possible to derive rather simple expressions for the total energy, momentum, and angular momentum radiated by chiral loops in the limit, as the loops are very close to their stationary states, i.e., $k \ll 1$ in (4). Additionally, it is supposed that k is independent of η and the current j^μ is therefore constant along the string. Using expansions of (32), (33), (63), and (64) in powers of k, we can write the corresponding gravitational and electromagnetic rates as

$$\dot{E}^{gr} = K_E^{gr} G\mu^2 k^2, \quad \left|\dot{\mathbf{P}}^{gr}\right| = K_P^{gr} G\mu^2 k^2,$$
$$\left|\dot{\mathbf{L}}^{gr}\right| = K_L^{gr}\mathcal{L}G\mu^2 k^2, \quad \dot{E}^{em} = K_E^{em} q^2\mu k^2, \quad (74)$$
$$\left|\dot{\mathbf{P}}^{em}\right| = K_P^{em} q^2\mu k^2, \quad \left|\dot{\mathbf{L}}^{em}\right| = K_L^{em}\mathcal{L}q^2\mu k^2,$$

where K^{em} and K^{gr} are numerical coefficients depending only on the loop geometry. We see that radiation rates of nearly stationary chiral loops are proportional to k^2. The geometrical numerical factors K in Eq. (74)

are in turn related to the corresponding coefficients Γ in Eq. (1) as

$$\Gamma = Kk^2. \tag{75}$$

We now evaluate the damping time of small-amplitude oscillations of nearly stationary chiral strings corresponding to the limit $k \ll 1$. For simplicity, we again assume that k is independent of η in the considered limit (this assumption is valid in the solvable examples considered above). The total loop charge conservation in (7) then gives

$$\frac{q\sqrt{\mu}}{2}L\sqrt{1-k^2} = \text{const}. \tag{76}$$

From this equation, we find the relation between the energy E and the parameter k of the chiral string with small-amplitude oscillations,

$$E \simeq E_v\left(1+\frac{k^2}{2}\right), \tag{77}$$

where $E_v = L\mu$ is the energy of the stationary (vorton) chiral loop configuration at $k = 0$. Comparing (77) with (74), we estimate the damping time of string oscillations [16] as

$$\tau \sim \frac{E_v}{2(K^{gr}G\mu^2 + K^{em}q^2\mu)}. \tag{78}$$

We next express (78) through the vorton length. We have $E_v = L\mu$, where L is the invariant length and the physical length of a stationary string is equal to half the invariant length $L_{ph} = L/2$ [22]. We find

$$\tau \sim \frac{L_{ph}}{K^{gr}G\mu + K^{em}q^2}. \tag{79}$$

Also assuming for simplicity that k depends only on time and using Eqs. (74) and (77), we find the oscillation damping law

$$k^2 \sim k_0^2\exp\left\{-t\left(\frac{1}{\tau_c^{gr}}+\frac{1}{\tau_c^{em}}\right)\right\}, \tag{80}$$

where $k_0 = k(t = 0)$; therefore, the damping time due to gravitational radiation is

$$\tau_c^{gr} \sim \frac{E_v}{2K^{gr}G\mu^2} = \frac{L_{ph}}{K^{gr}G\mu} \tag{81}$$

and the time due to electromagnetic radiation is

$$\tau_c^{em} \sim \frac{E_v}{2K^{em}q^2\mu} = \frac{L_{ph}}{K^{em}q^2}. \tag{82}$$

Substituting (80) in (77), we obtain

$$E \sim E_v\left[1+\frac{k_0^2}{2}\exp\left\{-t\left(\frac{1}{\tau_c^{gr}}+\frac{1}{\tau_c^{em}}\right)\right\}\right]. \tag{83}$$

The effective number of oscillations during the damping time (oscillator quality) is

$$Q = \frac{\tau}{T} \sim \frac{2}{L}\frac{\tau^{gr}\tau^{em}}{\tau^{gr}+\tau^{em}}. \tag{84}$$

To restore the standard CGS units, we replace $G\mu^2 \longrightarrow G\mu^2c$, $q^2\mu \longrightarrow q^2\mu c^2/\hbar$ and choose the standard normalization for the string mass per unit length $G\mu/c^2 = 10^{-6}\mu_6$ and $q_e = q/e$ for the dimensionless charge carrier on the string, where the elementary electric charge is $e = 4.8\times10^{-10}$. As a result, the damping times are expressed as

$$\tau^{gr} \sim \frac{L_{ph}c}{K^{gr}G\mu}, \quad \tau^{em} \sim \frac{L_{ph}\hbar}{K^{em}q^2}. \tag{85}$$

Oscillator quality (84) for the gravitational and electromagnetic radiation is given by the respective formulas

$$Q^{gr} \sim \frac{1}{K^{gr}}\frac{c^2}{G\mu}, \quad Q^{em} \sim \frac{1}{K^{em}}\frac{1}{\alpha_{em}q_e^2}, \tag{86}$$

with $\alpha_{em} = e^2/c\hbar$. The ratio of the damping times is

$$\frac{\tau^{gr}}{\tau^{em}} \sim \frac{q^2}{G\mu\hbar}\left(\frac{K^{em}}{K^{gr}}\right) \approx 1.4\times10^{-4}\frac{q_e^2}{\mu_6}\frac{K^{em}}{K^{gr}}. \tag{87}$$

If $q_e^2/\mu_6 \gtrsim 1.4\times10^{-3}$, the electromagnetic radiation prevails in the chiral loop evolution (this is valid for the standard values $\mu_6 \sim 1$ and $q_e \sim 1$). If, on the contrary, $q_e^2/\mu_6 \lesssim 1.4\times10^{-3}$ (for example, if the current is neutral and there is no electromagnetic radiation at all), then the pure gravitational radiation determines the evolution.

6. NUMERICAL EXAMPLES OF RADIATING LOOPS

In this section, we apply analytical formulas (32), (33), (63), and (64) derived above for gravitational and electromagnetic radiation to some particular examples of chiral loops. At the final steps, numerical calculations of four-dimensional integrals are used to find the energy, momentum, and angular momentum radiation rates as functions of the current on the string.

We first consider the class of piecewise linear kinky loops. Let $\mathbf{a}(\xi)$ and $\mathbf{b}(\eta)$ be piecewise linear functions; that is, vector functions $\mathbf{a}(\xi)$ and $\mathbf{b}(\eta)$ are closed loops

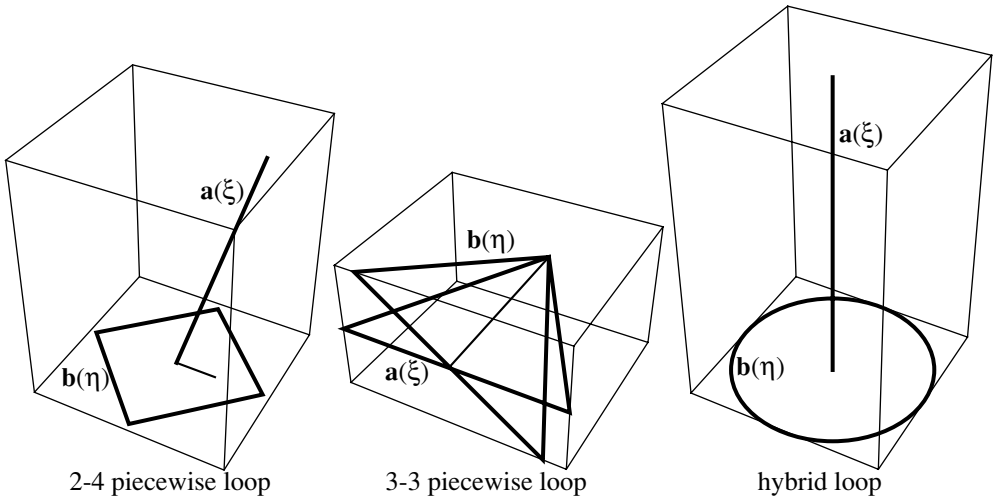


Fig. 1. Schematic view of the vector functions **a**(ξ) and **b**(η) for radiating loop examples considered in Sections 6.1, 6.2, and 6.3.

consisting of connected straight parts. The join points of segments of *a*- and *b*-loops, where **a**'(ξ) and **b**'(η) are discontinuous, are called "kinks". We take the *a*-loop consisting of N_a and the *b*-loop consisting of N_b segments (parts). Kinks are labeled by $i = 0, 1, \ldots, N_a - 1$, and the value of ξ on the kink labeled by *i* is denoted as ξ^i. In what follows, we use superscripts for the segment labels and subscripts for tensor components. Because we use only spatial tensor components, there should be no confusion. Without loss of generality, we can set $\xi^0 = 0$. We note that $\xi^{i+N_a} = \xi^i + 2\pi$ because of periodicity. Using the notation $\Delta\xi^i = \xi^{i+1} - \xi^i$, $\mathbf{A}^i = a(\xi^i)$, and $\mathbf{a}^i = (\mathbf{A}^{i+1} - \mathbf{A}^i)/\Delta\xi^i$, and similarly for the *b*-loop, we find

$$\begin{aligned} \mathbf{a}(\xi) &= \mathbf{A}^i + (\xi - \xi^i)\mathbf{a}^i, \quad \xi \in [\xi^i, \xi^{i+1}], \\ \mathbf{b}(\eta) &= \mathbf{B}^j + (\eta - \eta^j)\mathbf{b}^j, \quad \eta \in [\eta^j, \eta^{j+1}]. \end{aligned} \tag{88}$$

For piecewise linear loops, the functions $\mathcal{I}_p$, $\mathcal{Y}_p$, $\mathcal{M}_{pq}$, $\mathcal{N}_{pq}$, $\mathcal{X}$, and $\mathcal{Z}_p$ in (25) and (57) become the sums of delta functions because of the discontinuity of **a**' and **b**' at the kinks. For example, the function $\mathcal{I}_p$ in (25) is given by

$$\mathcal{I}_p = \sum_i \left(\frac{\mathbf{a}^i \cdot \mathbf{e}_p}{1 + \mathbf{a}^i \cdot \mathbf{n}} - \frac{\mathbf{a}^{i-1} \cdot \mathbf{e}_p}{1 + \mathbf{a}^{i-1} \cdot \mathbf{n}} \right) \delta(\xi - \xi^i). \tag{89}$$

Similar expressions can be obtained for the other functions. Due to the presence of delta functions in $\mathcal{I}_p$, $\mathcal{Y}_p$, $\mathcal{M}_{pq}$, $\mathcal{N}_{pq}$, $\mathcal{X}$, and $\mathcal{Z}_p$, the integrations in (32), (33), (63), and (64) can be carried out easily. To obtain the expressions for the gravitational and electromagnetic radiation from the general formulas, we must replace integrations in (32), (33), (63), and (64) by summations over the kinks and make the substitutions

$$\begin{aligned} \Delta x \longrightarrow x^{ijkl} &= \xi^i - \xi^k - (\eta^j - \eta^l) \\ &\quad + \mathbf{n} \cdot (\mathbf{a}^i - \mathbf{a}^k + \mathbf{b}^j - \mathbf{b}^l), \\ \mathcal{I}_p \longrightarrow \mathcal{I}_p^i &= \frac{\mathbf{a}^i \cdot \mathbf{e}_p}{1 + \mathbf{a}^i \cdot \mathbf{n}} - \frac{\mathbf{a}^{i-1} \cdot \mathbf{e}_p}{1 + \mathbf{a}^{i-1} \cdot \mathbf{n}}, \\ \mathcal{Y}_p \longrightarrow \mathcal{Y}_p^j &= \frac{\mathbf{b}^j \cdot \mathbf{e}_p}{1 - \mathbf{b}^j \cdot \mathbf{n}} - \frac{\mathbf{b}^{j-1} \cdot \mathbf{e}_p}{1 - \mathbf{b}^{j-1} \cdot \mathbf{n}}. \end{aligned} \tag{90}$$

Similar substitutions must be performed for the functions M_{pq}, $\mathcal{N}_{pq}$, *X*, and Z_p.

6.1. 2-4 Piecewise Loop

As the first example, we consider the chiral string loop shown in Fig. 1. In this example, the *a*-loop consists of two segments and lies along the *z* axis. One kink of the *a*-loop is positioned at the origin (ξ = 0) and the other kink (ξ = π) has the coordinates π(cos α, 0, sin α). The positions of the *b*-loop kinks are as follows: the first kink at η = 0 has the coordinates $(\pi k/2\sqrt{2})(1, 0, 0)$, the second kink at η = π/2 has the coordinates $(\pi k/2\sqrt{2})(0, 1, 0)$, the third kink at η = π has the coordinates $(\pi k/2\sqrt{2})(-1, 0, 0)$, and the position of the fourth kink at η = 3π/2 is given by $(\pi k/2\sqrt{2})(0, -1, 0)$. We call this loop the 2-4 piecewise loop. The dependence of the radiated gravitational and electromagnetic energy on the mode number *l* is shown in Fig. 2 for the 2-4 piecewise loop with α = π/2. The decrease of the radiated energy with the mode number *l* is more pronounced for the larger current, as it should be physically, because the maximal speed of the string

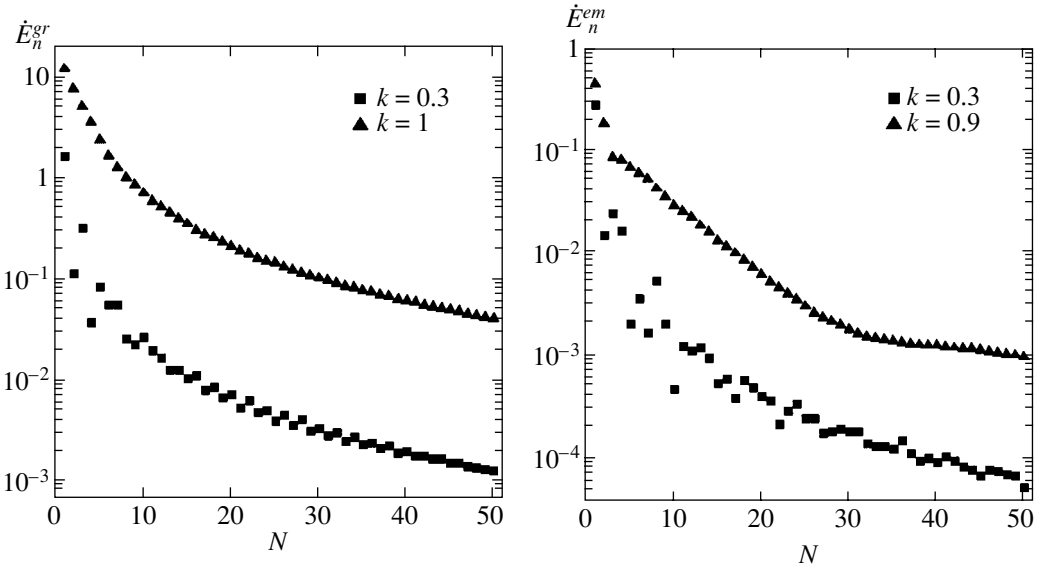


Fig. 2. Radiated gravitational energy rate (left graph) $\dot{E}_n^{gr}$ in the units $G\mu^2$ and the electromagnetic energy rate (right graph) $\dot{E}_n^{em}$ in the units $q^2\mu$. For the 2-4 kinky loop, the energy radiation is drawn as a logarithmic function of the mode number N for different values of the parameter k.

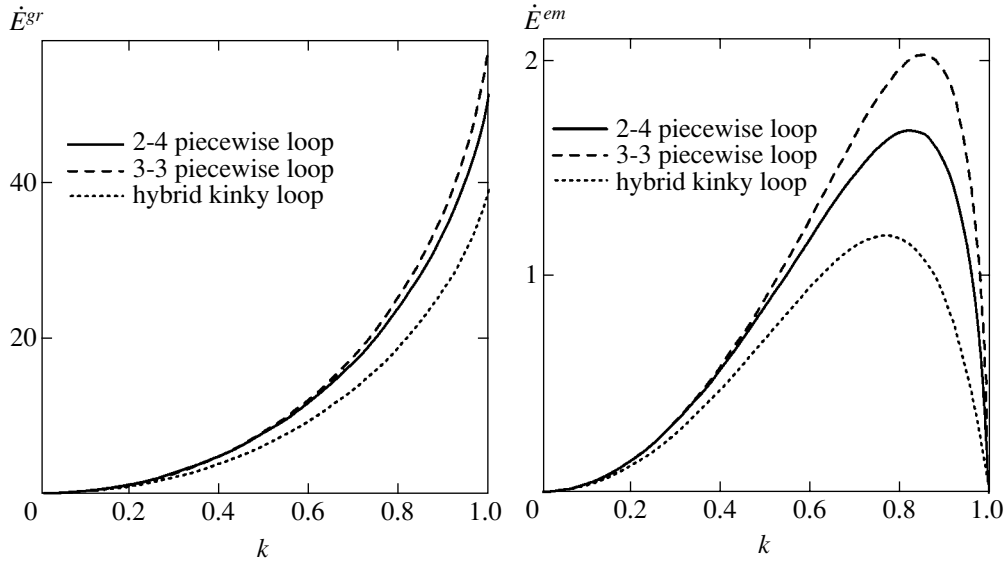


Fig. 3. The total radiated gravitational and electromagnetic energy rates $\dot{E}^{gr}$ in the units $G\mu^2$ (left graph) and $\dot{E}^{em}$ in the units $q^2\mu$ (right graph) correspondingly for the 2-4, 3-3 piecewise and hybrid kinky loops as a function of the parameter k. The following parameters are chosen: $\alpha = \pi/2$, $\beta = \pi/2$, $\gamma = 0$.

decreases as the current increases. In Fig. 3, the dependence of the total radiated energy on the parameter k is shown for $\alpha = \pi/2$. We can see a monotonic increase of the gravitational energy radiation with k (i.e., with a decrease in the string current). At the same time, the electromagnetic energy radiated by the string has a maximum near $k \sim 0.9$. The value of $k = k_{\max}^{gr}$ at which the maximum for the gravitational radiation rate is reached, is exactly 1, and for the electromagnetic radiation rate, $k_{\max}^{em} \sim 0.9$.

The corresponding rates for the angular momentum as a function of the mode number are shown in Fig. 4. For the electromagnetic radiation, we can also see weak oscillations of the angular momentum rate in addition to the overall decrease in the radiated angular momentum with the mode number.

The total angular momentum radiation to electromagnetic and gravitational waves is shown in Fig. 5. The graphs for the angular momentum rates look very similar to the graphs for the energy radiation. The cor-

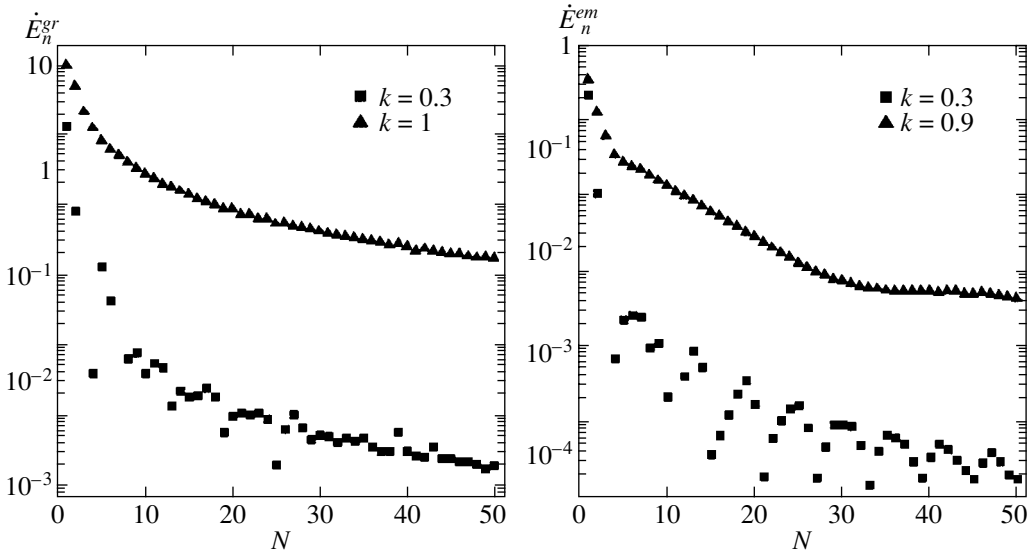


Fig. 4. The angular momentum $\dot{E}_n^{gr}$ radiated to gravitational waves in the units $G\mu^2$ (left graph) and the angular momentum $\dot{E}_n^{em}$ radiated to electromagnetic waves in the units $q^2\mu$ (right graph). For the 2-4 kinky loop, the energy radiation is drawn as a logarithmic function of the mode number N for different values of parameter k.

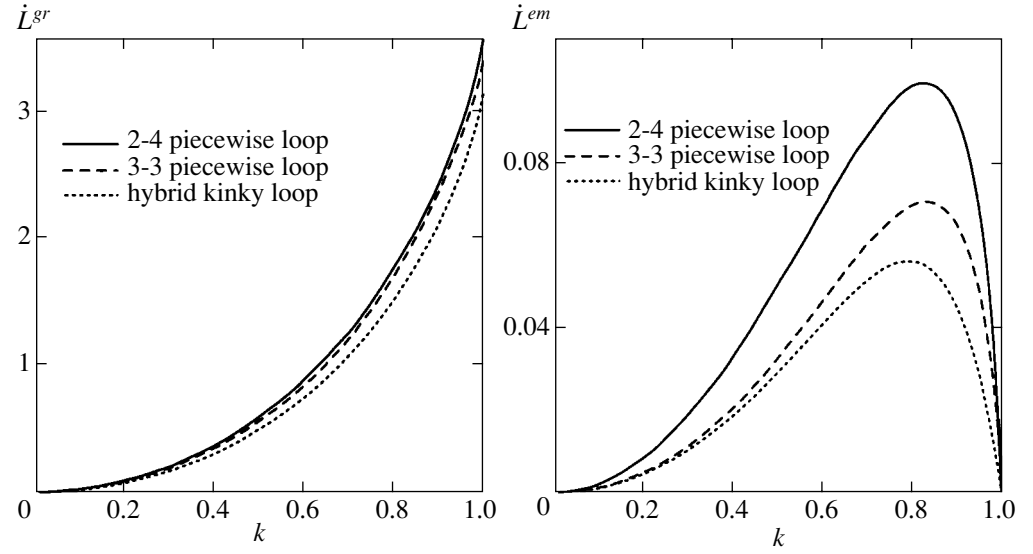


Fig. 5. The total angular momentum radiated to gravitational and electromagnetic waves, $\dot{L}^{gr}$ in the units $G\mu^2$ and $\dot{L}^{em}$ in the units $q^2\mu$, respectively, for the 2-4, 3-3 piecewise and hybrid kinky loops as a function of the parameter k. For the 2-4 loop, $\alpha = \pi/2$, for 3-3 loop, $\beta = \pi/2$, and for the hybrid loop, $\gamma = 0$.

responding gravitational radiation rates increase monotonically with k, and the electromagnetic radiation of momentum has maxima near $k = 0.9$. Using the general expressions for gravitational and electromagnetic radiation in Eqs. (32), (33), (63), and (64), we can easily calculate the coefficients K in the case of large currents. For $\alpha = \pi/2$, we find that $K_E^{gr} = 28.36$, $K_L^{gr} = 1.41$, $K_E^{gr} = 4$, and $K_L^{gr} = 0.25$. The radiated gravitational $\dot{E}^{gr}$ and electromagnetic $\dot{E}^{em}$ powers are approximately equal to Kk^2 in accordance with Eq. (75).

Durrer [14] found that for some particular class of ordinary cosmic string loops, the radiated angular momentum $\dot{\mathbf{L}}^{gr}$ is antiparallel to the stationary angular momentum $\mathbf{L}_{st}$ of the loop. This implies that the angular momentum of the loops always decreases with time due to gravitational radiation. Our results for the angular momentum radiation to electromagnetic and gravitational waves for string loops with the chiral current agree with the results of Durrer in general. The chiral loops considered in this paper also lose angular momentum with time. However, in contrast to the

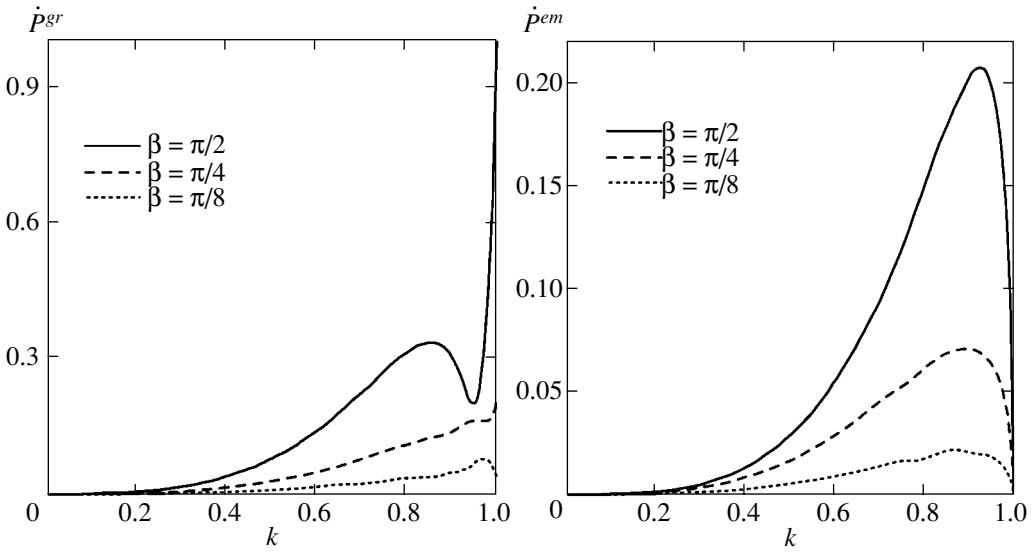


Fig. 6. The total momentum radiated to gravitational and electromagnetic waves, respectively, $\dot{P}^{gr}$ in the units $G\mu^2$ and $\dot{P}^{em}$ in the units $q^2\mu$, for the 3-3 piecewise loop with different parameters β_2 and with $\beta_1 = 0$ as a function of the parameter k. The three cases are considered with $\beta_2 = \pi/2$, $\pi/4$, and $\pi/8$.

examples considered by Durrer, we found that for some configurations of chiral loops, $\dot{\mathbf{L}}^{gr}$ and $\dot{\mathbf{L}}^{em}$ are not exactly antiparallel to the total angular momentum of the loop $\mathbf{L}_{st}$, but deviate by a small angle. In Table 1, the values $\varepsilon^{gr} = (\dot{\mathbf{L}}^{gr} \cdot \mathbf{L}_{st})/|\dot{\mathbf{L}}^{gr}||\mathbf{L}_{st}|$ and $\varepsilon^{em} = (\dot{\mathbf{L}}^{em} \cdot \mathbf{L}_{st})/|\dot{\mathbf{L}}^{em}||\mathbf{L}_{st}|$ determining the angle between $\dot{\mathbf{L}}$ and $\mathbf{L}_{st}$ are presented for the 2-4 piecewise loop with $\alpha = \pi/4$. We note that for symmetric configurations with $\alpha = \pi/2$, the angular momentum radiation $\dot{\mathbf{L}}^{gr}$, $\dot{\mathbf{L}}^{em}$ is exactly antiparallel to $\mathbf{L}_{st}$ at any k.

6.2. 3-3 Piecewise Loop

As the second example, we consider the two-parameter piecewise linear loop with *a* and *b*-loops consisting of three segments (Fig. 1). Positions of the *a*-loop kinks are given by the following coordinates: the first kink at $\eta = 0$ is at the origin, the second kink at $\eta = 2\pi/3$ has the coordinates $-(\pi/3)(\cos\beta_1, \sqrt{3}, \sin\beta_1)$, and the third kink at $\eta = 4\pi/3$ has the coordinates $(\pi/3)(\cos\beta_1, -\sqrt{3}, \sin\beta_1)$. The *b*-loop is given by almost the same conditions, except for the angle β_1 replaced by β_2. We call this loop the 3-3 piecewise loop. The total radiated energy rates to the gravitational and electromagnetic waves for $\beta_1 = 0$ and $\beta_2 = \pi/2$ are shown in Fig. 3. This loop also radiates momentum and angular momentum. The total angular momentum radiation rates are shown in Fig. 5. In Fig. 6, the total momentum radiation rates to electromagnetic and gravitational waves are shown for different values of the parameters β_1 and β_2. For the momentum radiation, we can see a different situation from that in the case of the energy and angular momentum radiation: for each value of k, the momentum rate has a local maximum on the interval $k \in (0, 2\pi)$.

6.3. Hybrid Kinky Loop

As the third example, we consider the loop with the configuration

$$\mathbf{a} = \mathbf{A}\begin{cases}\xi, & 0 \le \xi \le \pi,\\ \pi - \xi, & \pi \le \xi \le 2\pi,\end{cases} \qquad (91)$$

$$\mathbf{b} = k(\sin\eta, -\cos\eta, 0).$$

The *b*-loop in this example is a circle in the *x*, *y* plane and $\mathbf{A} = (\cos\gamma; 0; \sin\gamma)$ (Fig. 1). For $\gamma = \pi/2$, the gravitational and electromagnetic radiated energy rates and

The cosine of the angle between $\dot{\mathbf{L}}^{gr}$ and $\mathbf{L}_{st}$ and between $\dot{\mathbf{L}}^{em}$ and $\mathbf{L}_{st}$ for the 2-4 piecewise loop

	k	0.2	0.4	0.6	0.8	1.0
2–4 loop, $\alpha = \pi/4$	ε^{gr}	−0.94	−0.95	−0.95	−0.96	−0.97
	ε^{em}	−0.97	−0.99	−0.99	−0.99	–

angular momentum rates are shown in Figs. 3 and 5. The total gravitational energy radiation for $k = 1$ coincides with the result of Allen *et al.* [12] ($\dot{E}^{gr} \approx 39.0G\mu^2$).

6.4. Weakly Oscillating Ring Loop

As the final example, we consider the radially oscillating loop,

$$\begin{aligned} \mathbf{a} &= (\cos\xi, -\sin\xi, 0), \\ \mathbf{b} &= k(\cos\eta, -\sin\eta, 0). \end{aligned} \tag{92}$$

Unfortunately, because the calculation of integrals (32) and (63) would take an enormous amount of computer time, we cannot present the results for radiation rates of oscillating rings for the entire range of currents (we note that the radiated power diverges as the current tends to zero). However, in the large-current limit, the radiated power rate is easy to calculate. For loop (92), the first nonzero term in the expansion of the radiated power in k is proportional to k^2 in agreement with (74). It suffices to take only the first term in (13) and (43); the other terms are of higher orders in k. Substituting (92) in (20) and (53) and keeping the leading nonzero term at $k \ll 1$, we obtain

$$\begin{aligned} I_2(1) &= \frac{1}{2}e^{-i\phi}[J_2(-\sin\theta) + J_0(-\sin\theta)]\cos\theta, \\ I_3(1) &= \frac{i}{2}e^{-i\phi}[J_2(-\sin\theta) - J_0(-\sin\theta)], \\ Y_2(1) &= \frac{k}{2}e^{i\phi}\cos\theta, \quad Y_3(1) = i\frac{k}{2}e^{i\phi}, \\ X(1) &= \frac{k}{2}e^{i\phi}\sin\theta. \end{aligned} \tag{93}$$

Using (93), (19), (52), (14), and (44) and integrating over the unit sphere, we next obtain the coefficients K_E^{gr} and K_E^{em},

$$\begin{aligned} K^{gr} &= \pi^2\int_0^\pi d\theta\sin\theta\{[J_2(\sin\theta) - J_0(\sin\theta)]^2 \\ &\quad + 2[3J_2^2(\sin\theta) - J_0^2(\sin\theta)]\cos^2\theta \\ &\quad + [J_2(\sin\theta) + J_0(\sin\theta)]^2\cos^4\theta\}, \end{aligned} \tag{94}$$

$$\begin{aligned} K^{em} &= \int_0^\pi d\theta\sin^3\theta\{[J_2(\sin\theta) - J_0(\sin\theta)]^2 \\ &\quad + [J_2(\sin\theta) + J_0(\sin\theta)]^2\cos^2\theta\}, \end{aligned}$$

which are numerically given by $K^{gr} = 4.73$ and $K^{em} = 2.28$.

Because of the symmetry of the oscillating ring, the estimations of damping time (79), coefficient k (80), and total string energy (83) become exact in the large-current limit.

7. DISCUSSION

Electromagnetic and gravitational radiation plays an important role in the evolution of the cosmic string network. This network could be produced in early universe phase transitions and would generate large-scale structures later. Previously, the properties of cosmic string radiation were mainly studied for strings with a small current or without any current. Here, we described the radiation properties of chiral cosmic loops for the entire possible range of currents. We succeeded in analytically summing the infinite mode series of radiation rates for periodically oscillating string loops. The expressions derived for the energy, momentum, and angular momentum rates contain four-dimensional integrals depending on loop geometry. Such an integral representation is especially convenient for numerical calculations of radiation from relativistically moving loops as compared to the method of summation of a weakly convergent mode series. To find the total rates of the radiated energy, momentum, and angular momentum, the expressions obtained were integrated over the unit sphere. Applying the derived formulas to some particular examples of chiral string loop configurations, we numerically calculated the coefficients Γ in Eq. (1) as functions of k. The corresponding calculations of the radiated energy, momentum, and angular momentum rates were performed for the following examples (see Fig. 1): (i) a piecewise linear kinky loop with the a-loop consisting of two straight parts and the b-loop consisting of four straight parts (2-4 piecewise loop); (ii) a piecewise linear loop such that the a- and b-loops consist of three segments each (3-3 piecewise loop); and (iii) the hybrid loop in which the a-loop consists of two straight parts and the b-loop is a circle (hybrid kinky loop). For the first and second examples, the four-dimensional integrals in our expressions for radiated energy, momentum, and angular momentum become multiple sums over the kinks. These sums can be calculated analytically using symbolic computer manipulations (e.g., the Mathematical program package). To find the radiation in the third example (hybrid loop), we calculated two-dimensional integrals (originating from a smooth a-loop) and summed over the kinks of the b-loop. Unfortunately, we could not perform calculations for strings with the a- and b-loops being arbitrary smooth curves, because calculations of the four-dimensional integrals would take an enormous amount of computer time.

The total gravitational radiation energy, momentum, and angular momentum rates behave similarly: they slowly increase with k when k is sufficiently small (and, respectively, the current is large) and rapidly increase at large k (i.e., at small current). Overall, the gravitational radiation rates are increasing functions of k. For the electromagnetic radiation, the situation is quite different: the energy, momentum, and angular momentum losses to electromagnetic waves for all examples considered have a maximum near $k \sim 0.9$, i.e., when the current is rather small. For the examples considered, the maximum values of the coefficients Γ in (1) are approximately equal to

$$\begin{aligned} &\Gamma_E^{gr} \approx 50, \quad \Gamma_P^{gr} \approx 1, \quad \Gamma_L^{gr} \approx 3, \\ &\Gamma_E^{em} \approx 2, \quad \Gamma_P^{em} \approx 0.1, \quad \Gamma_L^{em} \approx 0.1. \end{aligned} \tag{95}$$

We also found that for some nonsymmetric examples of chiral loops, the angular momentum $\dot{\mathbf{L}}$ that radiated to electromagnetic and gravitational waves is not exactly opposite to the angular momentum of the loop $\mathbf{L}_{st}$, but slightly differs from it (even when there is no current on the string), unlike in the loop examples considered by Durrer [14].

The asymptotic fading of chiral cosmic string loops into vortons was derived. It was found that the upper bounds on the gravitational and electromagnetic radiation rates of nearly stationary loops is proportional to the squared third derivative of the oscillation amplitude (see Eqs. (68), (69), and (72)). We showed that if the oscillation amplitude is small ($k \ll 1$) and the current j^μ is constant along the string, the energy, momentum, and angular momentum rates of radiation to gravitational and electromagnetic waves are proportional to k^2 and the proportionality coefficient depends only on the form of the loop (see Eq. (74)). In some examples of chiral loops, we calculated the total radiated power in the limit of the small-amplitude oscillations. For the chiral ring with small-amplitude radial oscillations, the radiated power per solid angle $d\Omega$ for the electromagnetic and gravitational radiation is found analytically, Eq. (94). We also estimated the damping time of chiral loops (78) with small-amplitude oscillations. In the case of the gravitational radiation prevalence over the electromagnetic one, this time is $\tau^{gr} \sim L_{ph}/K^{gr}G\mu$, where K^{gr} is a numerical coefficient depending on the string geometry. The damping time due to the gravitational radiation of the chiral loops considered is by an order of magnitude longer than the lifetime of ordinary cosmic strings. On the contrary, if the electromagnetic radiation prevails, the decay time is $\tau^{em} \sim L_{ph}/K^{em}q^2$. For a radially oscillating chiral ring with a large current, expressions (80) and (83) for the temporal evolution of the total energy and the amplitude parameter k become asymptotically exact.

We can find the characteristic size of the string L_ν with oscillation damping time (85) equal to the universe lifetime $t_0 \simeq 10^{18}$ s. In the case of the gravitational radiation predominance, we find

$$L_\nu^{gr} \sim \frac{G\mu K^{gr} t_0}{c} \approx 10^2 \mu_6 \text{ kpc} \tag{96}$$

for $K^{gr} \sim 1$. Chiral strings with a length of $L < L_\nu^{gr}$ (i.e., with the size of a typical galactic halo or less) therefore have enough time to fade into vortons. On the other hand, if the electromagnetic radiation prevails, we have

$$L_\nu^{em} \sim \frac{q^2 K^{em} t_0}{\hbar} \approx 70 q_e^2 \text{ Mpc} \tag{97}$$

for $K^{em} \sim 1$ and the electromagnetically radiated chiral loops with length shorter than the size of galactic clusters have therefore transformed into vortons. We can see that only sufficiently long superconducting cosmic strings oscillate up to the present time. On the contrary, small-scale chiral loops are transformed into stationary vortons due to the oscillation damping.

It is interesting to estimate the current parameter k at which the electromagnetic and gravitational radiation rates become equal. From [15], we know that at small currents, the electromagnetic radiation is given by $\dot{E}^{em} \sim q^2\mu\sqrt{1-k}$. For the gravitational radiation of small currents, we have $\dot{E}^{gr} \sim 10^2 G\mu^2$. Comparing these two expressions, we can easily find the string current value at which these radiation rates are equal, $k \sim 1 - (10^2 G\mu/q^2)^2 \sim 1\text{–}10^{-4}$.

ACKNOWLEDGMENTS

This work was supported in part by the Russian Foundation for Basic Research (grant nos. 00-15-96632 and 00-15-96697) and by INTAS (grant no. 99-1065).

REFERENCES

1. E. P. S. Shellard and A. Vilenkin, *Cosmic Strings and Other Topological Defects* (Cambridge Univ. Press, Cambridge, 1994).
2. M. B. Hindmarsh and T. W. B. Kibble, Rep. Prog. Phys. **58**, 477 (1995).
3. E. Witten, Nucl. Phys. B **249**, 557 (1985).
4. B. Carter and P. Peter, Phys. Lett. B **466**, 41 (1999).
5. A. C. Davis, T. W. B. Kibble, M. Pickles, and D. A. Steer, Phys. Rev. D **62**, 083516 (2000).
6. J. J. Blanco-Pillado, Ken D. Olum, and A. Vilenkin, Phys. Rev. D **63**, 103513 (2001).
7. T. Vachaspati and A. Vilenkin, Phys. Rev. D **31**, 3052 (1985).
8. C. Burden, Phys. Lett. B **164B**, 277 (1985).

9. D. Garfinkle and T. Vachaspati, Phys. Rev. D **36**, 2229 (1987).
10. T. Damour and A. Vilenkin, Phys. Rev. Lett. **85**, 3761 (2000).
11. B. Allen and P. Casper, Phys. Rev. D **50**, 2496 (1994).
12. B. Allen, P. Casper, and A. Ottewill, Phys. Rev. D **50**, 3703 (1994).
13. B. Allen, P. Casper, and A. Ottewill, Phys. Rev. D **51**, 1546 (1995).
14. R. Durrer, Nucl. Phys. B **328**, 238 (1989).
15. J. J. Blanco-Pillado and Ken D. Olum, Nucl. Phys. B **599**, 435 (2001).
16. E. Babichev and V. Dokuchaev, Phys. Rev. D **66**, 025007 (2002).
17. S. Weinberg, *Gravitation and Cosmology: Principles and Applications of the General Theory of Relativity* (Wiley, New York, 1974; Mir, Moscow, 1975).
18. I. S. Gradshteyn and I. M. Ryzhik, *Tables of Integrals, Series and Products*, 5th ed. (Nauka, Moscow, 1971; Academic, New York, 1980).
19. E. Babichev and V. Dokuchaev, Nucl. Phys. B **645**, 134 (2002).
20. L. D. Landau and E. M. Lifshitz, *The Classical Theory of Fields*, 6th ed. (Nauka, Moscow, 1973; Pergamon, New York, 1975).
21. A. Vilenkin and T. Vachaspati, Phys. Rev. Lett. **58**, 1041 (1987).
22. D. A. Steer, Phys. Rev. D **63**, 083517 (2001).

Journal of Experimental and Theoretical Physics, Vol. 96, No. 4, 2003, pp. 610–616.
Translated from Zhurnal Éksperimental'noĭ i Teoreticheskoĭ Fiziki, Vol. 123, No. 4, 2003, pp. 688–695.

NUCLEI, PARTICLES, AND THEIR INTERACTION

Quantum Mechanical Equations of Motion of Particles and Spin in Polarized Media

A. Ya. Silenko

Research Institute of Nuclear Problems, Belarussian State University, Minsk, 220050 Belarus
e-mail: silenko@inp.minsk.by

Received September 11, 2002

Abstract—Quantum mechanical equations of motion are obtained for particles and spin in media with polarized electrons in the presence of external fields. The motion of electrons and their spins is governed by the exchange interaction, while the motion of positrons and their spin is governed by the annihilation interaction. For particles with spin $S \geq 1$, second-order terms in spin are taken into account. The equations obtained can be applied to describe the motion of particles and spin both in magnetic and nonmagnetic media. © *2003 MAIK "Nauka/Interperiodica".*

1. INTRODUCTION

One of the most important problems in the study of interaction of particles with matter is the quantum mechanical description of the motion of particles and spin. The classical theory of motion of particles and spin has been developed in great detail (see [1, 2]). A quantum mechanical equation of motion of relativistic particles in an electromagnetic field was derived by Derbenev and Kondratenko [3]. The equation of motion for spin was found still earlier. It is well-known that the motion of the spin of relativistic particles in an electromagnetic field is described by the Bargmann–Michel–Telegdi (BMT) equation [4]. The quantum mechanical version of this equation for spin-1/2 particles was consistently derived in [5] and, for particles with arbitrary spin, in [6, 7]. Derbenev and Kondratenko investigated the effect of radiative effects on the spin motion [3]. A detailed analysis of the dynamics of the polarization of high-energy particles with regard to radiative effects is given in [8]. A consistent quantum mechanical description of the interaction between relativistic particles with arbitrary spin and an electromagnetic field with regard to terms quadratic in spin was carried out in [7, 9]. The equation of spin motion derived by the Lagrangian obtained in [7, 9] is presented in [10].

The study of the motion of particles and the dynamics of their polarization in media with polarized electrons presents an important problem. The characteristic features of the interaction between polarized particles and polarized matter were analyzed in [11]. In the present paper, we apply the results of [11] and the expression for the Hamiltonian in the Foldy–Wouthuysen (FW) [13] representation derived in [12] to find quantum mechanical equations of motion of particles and spin for relativistic particles with arbitrary spin that move in media with polarized electrons in the presence of external fields.

Here, we do not consider strong interaction, which may also change the polarization of particles in matter [11, 14–16]. Also, we do not take into consideration the effect of radiative phenomena on the polarization of a beam. In principle, these phenomena may lead to an additional rotation of the polarization vector and to the radiative self-polarization (see [3, 17]). However, they become appreciable only at high energies of particles (for electrons and positrons, at energies greater than 10 GeV).

In this paper, we use the relativistic system of units $\hbar = c = 1$.

2. HAMILTONIAN FOR PARTICLES IN POLARIZED MEDIA

The FW representation [13] is the most convenient means for describing the interaction of relativistic particles with matter and external fields by means of equations of motion of particles and spin. In this representation, the operators have the simplest form, which corresponds to the form of operators in nonrelativistic quantum mechanics. The FW representation is especially convenient for describing polarization phenomena because, in other representations (for example, in the Dirac representation for spin-1/2 particles), the polarization operator is given by cumbersome expressions (see [18]).

A formula for the Hamiltonian in the FW representation that describes the interaction of spin-1/2 particles

with an electromagnetic field was obtained in [12]. Up to terms linear in field, it is expressed as[1]

$$\mathcal{H} = \beta\epsilon' + e\Phi + \frac{1}{4}\left\{\left(\frac{\mu_0 m}{\epsilon' + m} + \mu'\right)\frac{1}{\epsilon'},\right.$$

$$\left.(\boldsymbol{\Sigma}\cdot[\boldsymbol{\pi}\times\mathbf{E}] - \boldsymbol{\Sigma}\cdot[\mathbf{E}\times\boldsymbol{\pi}] - \nabla\cdot\mathbf{E})\right\}_+$$

$$+\frac{\mu_0 m}{16}\left\{\frac{2\epsilon'^2 + 2\epsilon' m + m^2}{\epsilon'^4(\epsilon' + m)^2}, \boldsymbol{\pi}\cdot\nabla(\boldsymbol{\pi}\cdot\mathbf{E} + \mathbf{E}\cdot\boldsymbol{\pi})\right\}_+$$

$$-\frac{1}{2}\left\{\left(\frac{\mu_0 m}{\epsilon'} + \mu'\right), \boldsymbol{\Pi}\cdot\mathbf{H}\right\}_+ \quad (1)$$

$$+\frac{\mu'}{4}\left\{\frac{1}{\epsilon'(\epsilon' + m)}, [(\mathbf{H}\cdot\pi)(\boldsymbol{\Pi}\cdot\boldsymbol{\pi})\right.$$

$$\left.+ (\boldsymbol{\Pi}\cdot\boldsymbol{\pi})(\boldsymbol{\pi}\cdot\mathbf{H}) + 2\pi(\boldsymbol{\pi}\cdot\mathbf{j} + \mathbf{j}\cdot\pi)]\right\}_+,$$

where

$$\epsilon' = \sqrt{m^2 + \boldsymbol{\pi}^2}, \quad \beta = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix},$$

$$\boldsymbol{\Sigma} = \begin{pmatrix} \boldsymbol{\sigma} & 0 \\ 0 & \boldsymbol{\sigma} \end{pmatrix}, \quad \boldsymbol{\Pi} = \begin{pmatrix} \boldsymbol{\sigma} & 0 \\ 0 & -\boldsymbol{\sigma} \end{pmatrix};$$

$\mu_0 = e/2m$, $\mu' = \mu - \mu_0$, and μ are the Dirac, anomalous, and total magnetic moments, respectively; e and m are the particle charge and mass, respectively; $\pi = -i\nabla - e\mathbf{A}$ is the operator of kinetic momentum; Φ, $\mathbf{A}$ and $\mathbf{E}$, $\mathbf{H}$ are the potentials and the strengths of the electromagnetic field; $\boldsymbol{\sigma}$ is the Pauli matrix; and $\{\ldots, \ldots\}_+$ denotes the anticommutator.

Transition to the quasiclassical description consists in averaging the operators over the wave functions of stationary states. For free particles, the lower spinor in the FW representation vanishes [19]. For particles in an external field, the ratio of the lower spinor to the upper is of the same order of magnitude as $|W_{\text{int}}|/E$, where W_{int} is the energy of interaction between particles and the field and E is the total energy of a particle. Thus,

$$\frac{\chi^\dagger\chi}{\phi^\dagger\phi} \sim (W_{\text{int}}/E)^2.$$

Therefore, the contribution of the lower spinor is negligible, and the upper spinor is sufficient for the quasiclassical description. The motion of particles in external electric and magnetic fields admits a quasiclassical description; this allows one to speak of the trajectory of particles. One can also neglect the commutators of the operators of dynamic variables $\mathbf{r}$ and $\mathbf{p}$ and their functions; this allows one to arbitrarily change the order of these operators in quantum mechanical expressions (see [17–20]). In [20], the present author considered the accuracy of these assumptions when describing the motion of particles in a magnetic field. However, the spin of a particle remains an essentially quantum quantity in any case.

For particles with arbitrary spin, the Hamiltonian can be derived with the use of the interaction Lagrangian $\mathcal{L}$ obtained in [7, 9]:

$$\mathcal{L} = \mathcal{L}_1 + \mathcal{L}_2,$$

$$\mathcal{L}_1 = \frac{e}{2m}\left\{\left(g - 2 + \frac{2}{\gamma}\right)(\mathbf{S}\cdot\mathbf{B})\right.$$

$$-(g-2)\frac{\gamma}{\gamma+1}(\mathbf{S}\cdot\mathbf{v})(\mathbf{v}\cdot\mathbf{B})$$

$$\left.+\left(g - 2 + \frac{2}{\gamma+1}\right)(\mathbf{S}\cdot[\mathbf{E}\times\mathbf{v}])\right\},$$

$$\mathcal{L}_2 = \frac{Q}{2s(2s-1)}\left[(\mathbf{S}\cdot\nabla) - \frac{\gamma}{\gamma+1}(\mathbf{S}\cdot\mathbf{v})(\mathbf{v}\cdot\nabla)\right] \quad (2)$$

$$\times\left[(\mathbf{S}\cdot\mathbf{E}) - \frac{\gamma}{\gamma+1}(\mathbf{S}\cdot\mathbf{v})(\mathbf{v}\cdot\mathbf{E}) + (\mathbf{S}\cdot[\mathbf{v}\times\mathbf{B}])\right]$$

$$+\frac{e}{2m^2}\frac{\gamma}{\gamma+1}(\mathbf{S}\cdot[\mathbf{v}\times\nabla])\left[\left(g - 1 + \frac{1}{\gamma}\right)(\mathbf{S}\cdot\mathbf{B})\right.$$

$$\left.-(g-1)\frac{\gamma}{\gamma+1}(\mathbf{S}\cdot\mathbf{v})(\mathbf{v}\cdot\mathbf{B}) + \left(g - \frac{\gamma}{\gamma+1}\right)(\mathbf{S}\cdot[\mathbf{E}\times\mathbf{v}])\right],$$

$$\gamma = \frac{1}{\sqrt{1-\mathbf{v}^2}} = \frac{\sqrt{m^2 + \boldsymbol{\pi}^2}}{m},$$

where $g = 2\mu m/(eS)$, $\mathbf{v}$ is the velocity operator, γ is the Lorentz factor, Q is a quadrupole moment, $\mathbf{S}$ is the spin operator, and $\mathbf{B}$ is the magnetic induction. In Lagrangian (2), $\mathcal{L}_1$ contains terms that are linear in spin, while $\mathcal{L}_2$ contains quadratic terms. The Hermitian form of relation (2) is obtained by the substitution

$$\mathcal{L} \longrightarrow \frac{\mathcal{L} + \mathcal{L}^\dagger}{2}.$$

The spin-dependent part of the Hamiltonian is equal to the interaction Lagrangian taken with the

[1] A similar formula for spin-one particles was obtained in [10].

opposite sign:

$$\mathcal{H}^{(s)} = -\mathcal{L}.$$

The spin-independent part of the Hamiltonian that is significant for deriving the equations of motion of particles can easily be determined by the well-known procedure of "extension" of a derivative. For free particles in the FW representation, the relationship between the Hamiltonian and the momentum operator **p** for the upper spinor is given by

$$\mathcal{H} = \sqrt{m^2 + \mathbf{p}^2}.$$

In the presence of an external field, the extension procedure leads to the following substitutions:

$$\mathcal{H} \longrightarrow \mathcal{H}^{(0)} = \mathcal{H} - e\Phi, \quad \mathbf{p} \longrightarrow \boldsymbol{\pi} = \mathbf{p} - e\mathbf{A}.$$

As a result, the total Hamiltonian is given by

$$\mathcal{H} = \mathcal{H}^{(0)} + \mathcal{H}^{(s)},$$

$$\mathcal{H}^{(0)} = \sqrt{m^2 + \boldsymbol{\pi}^2} + e\Phi = \gamma m + e\Phi,$$

$$\mathcal{H}^{(s)} = -\mathcal{L},$$

or

$$\mathcal{H} = \sqrt{m^2 + \boldsymbol{\pi}^2} + e\Phi - (\mathcal{L}_1 + \mathcal{L}_2). \tag{3}$$

We stress that, in Eq. (3), only the upper spinor is retained and the commutators of the operators of dynamic variables and their functions are neglected. An approximate relation between the operators of velocity and kinetic momentum is given by

$$\mathbf{v} \approx \boldsymbol{\pi}/\epsilon' = \boldsymbol{\pi}/\gamma m.$$

Therefore, formulas (1) and (3) completely agree with each other.

The polarization of the electrons of the medium does not change the form of Hamiltonian (3) provided that a beam contains neither electrons nor positrons. This is attributed to the fact that the average field acting on particles in the medium is characterized by the electric field strength **E** and the magnetic induction **B**.

However, the form of the Hamiltonian is changed if the beam consists of electrons or positrons. There is an exchange interaction between electrons, which is very strong.[2] The main contribution to the exchange interaction is made by the Coulomb exchange interaction or the Coulomb scattering. The exchange magnetic scattering yields a much smaller contribution to the Hamiltonian; however, in magnetic crystals, it is comparable in order of magnitude to the contribution of an ordinary magnetic field (see [11]). The well-known formulas for the scattering amplitudes from which one can obtain an expression for the effective interaction Hamiltonian (see [11]) were derived for the nonrelativistic case $v \ll c$. In this case, the magnetic field **B** for electrons should be replaced in (3) by the effective quasimagnetic field [11]

$$\mathbf{B} \longrightarrow \mathbf{G}_e = \mathbf{B} + \mathbf{H}^{\mathrm{c}}_{\mathrm{eff}} + \mathbf{H}^{\mathrm{m}}_{\mathrm{eff}},$$

$$\mathbf{H}^{\mathrm{c}}_{\mathrm{eff}} = -\frac{4\pi|eP|N}{mv^2}\mathbf{P}, \tag{4}$$

$$\mathbf{H}^{\mathrm{m}}_{\mathrm{eff}} = \frac{2\pi|e|N}{m}(\mathbf{P}\cdot\mathbf{n})\mathbf{n},$$

where N and $\mathbf{P} = \langle\boldsymbol{\sigma}'\rangle$ are the polarization density and vector (average spin), respectively, of polarized electrons of matter and $\mathbf{n} = \mathbf{v}/v$ is the propagation direction of the beam. Usually, nonrelativistic electrons, for which Eqs. (4) hold, satisfy the following inequalities:

$$\left|\mathbf{H}^{\mathrm{c}}_{\mathrm{eff}}\right| \gg |\mathbf{B}|, \quad \left|\mathbf{H}^{\mathrm{c}}_{\mathrm{eff}}\right| \gg \left|\mathbf{H}^{\mathrm{m}}_{\mathrm{eff}}\right|$$

(see [11]).

In addition to the interaction described by Hamiltonian (3), there exists an annihilation interaction between positrons and the electrons of matter; in the nonrelativistic case, this interaction is determined by the operator [19]

$$W_{\mathrm{a}} = \frac{\pi e^2}{2m^2}[3 + (\boldsymbol{\sigma}\cdot\boldsymbol{\sigma}')]\delta(\mathbf{r}),$$

where $\boldsymbol{\sigma}$ and $\boldsymbol{\sigma}'$ are the Pauli matrices for positrons and electrons, respectively. After averaging, the expression for the spin-dependent part of the operator of annihilation interaction takes the form ($\mathbf{P} = \langle\boldsymbol{\sigma}'\rangle$)

$$W^{(s)}_{\mathrm{a}} = \frac{\pi e^2 N}{2m^2}(\boldsymbol{\sigma}\cdot\mathbf{P}). \tag{5}$$

The effective field $\mathbf{H}^{\mathrm{a}}_{\mathrm{eff}}$ corresponding to the annihilation interaction is determined by

$$W^{(s)}_{\mathrm{a}} = -\boldsymbol{\mu}\cdot\mathbf{H}^{\mathrm{a}}_{\mathrm{eff}},$$

where

$$\boldsymbol{\mu} = \frac{e}{2m}\boldsymbol{\sigma}$$

is the operator of magnetic moment of a positron. Hence, one can easily derive an expression for the effective field that acts on nonrelativistic positrons in polarized media:

$$\mathbf{G}_{\mathrm{p}} = \mathbf{B} + \mathbf{H}^{\mathrm{a}}_{\mathrm{eff}} = \mathbf{B} - \frac{\pi|e|N}{m}\mathbf{P}. \tag{6}$$

It is convenient to represent formulas (4) and (6) in a more compact form by introducing the magnetization vector (magnetic moment of a unit volume) **M**:

$$\mathbf{M} = -\frac{|e|N}{2m}\mathbf{P}.$$

[2] Recall that the exchange interaction is responsible for the ferromagnetism.

This formula takes into consideration dia-, para-, and ferromagnetic phenomena, and formulas (4) and (6) are rewritten as

$$\mathbf{G}_e = \mathbf{B} + \frac{8\pi}{v^2}\mathbf{M} - 4\pi(\mathbf{M}\cdot\mathbf{n})\mathbf{n}, \quad \mathbf{G}_\mathrm{p} = \mathbf{B} + 2\pi\mathbf{M}. \tag{7}$$

For isotropic magnetic materials, one can introduce a magnetic permeability μ_m. Then,[3]

$$\mathbf{G}_e = \mathbf{B} + \frac{2(\mu_\mathrm{m}-1)}{\mu_\mathrm{m} v^2}\mathbf{B} - \frac{\mu_\mathrm{m}-1}{\mu_\mathrm{m}}(\mathbf{B}\cdot\mathbf{n})\mathbf{n}, \qquad \mathbf{G}_\mathrm{p} = \frac{3\mu_\mathrm{m}-1}{2\mu_\mathrm{m}}\mathbf{B}. \tag{8}$$

Formulas (8) for electrons and positrons in a polarized medium have been derived in a nonrelativistic approximation that admits the consideration of basic relativistic corrections. In this approximation, an appropriate expression for the Hamiltonian, obtained from (3) by replacing **B** by **G**, is expressed as

$$\mathcal{H} = \sqrt{m^2 + \boldsymbol{\pi}^2} + e\Phi + \frac{e}{2m}\{g(\mathbf{S}\cdot\mathbf{G}) + (g-1)(\mathbf{S}\cdot[\mathbf{E}\times\mathbf{v}])\}, \tag{9}$$

where $\mathbf{G} = \mathbf{G}_e, \mathbf{G}_\mathrm{p}$.

Formulas (4) and (7)–(9) for electrons, which contain terms proportional to $1/v^2$, are valid only when the velocities of particles are not too small. The analysis of the limiting case of small velocities is of interest.[4] This case is described by the ordinary theory of exchange interaction. As is known, this interaction is characterized by the operator [21]

$$V = -\frac{1}{2}J[1 + (\boldsymbol{\sigma}\cdot\boldsymbol{\sigma}')],$$

where J is the exchange integral and $\boldsymbol{\sigma}$ and $\boldsymbol{\sigma}'$ are the Pauli matrices for electrons in the beam and the medium, respectively. After averaging over the coordinates and summing up over the electrons of the medium, the operator V is rewritten as

$$\langle V\rangle = -\frac{1}{2}\langle J\rangle[1 + (\boldsymbol{\sigma}\cdot\mathbf{P})]. \tag{10}$$

As above, it is convenient to introduce a quasimagnetic field $\mathbf{H}^\mathrm{c}_\mathrm{eff}$ defined by

$$\langle V\rangle = -\boldsymbol{\mu}\cdot\mathbf{H}^\mathrm{c}_\mathrm{eff} = \frac{|e|}{2m}(\boldsymbol{\sigma}\cdot\mathbf{H}^\mathrm{c}_\mathrm{eff}).$$

Then,

$$\mathbf{H}^\mathrm{c}_\mathrm{eff} = -\frac{m}{|e|}\langle J\rangle\mathbf{P}, \tag{11}$$

and the total quasimagnetic field acting on the electrons of the beam is

$$\mathbf{G} = \mathbf{B} + \mathbf{H}^\mathrm{c}_\mathrm{eff}.$$

Introducing the magnetization vector **M** and the magnetic permeability μ_m, we rewrite this formula as

$$\mathbf{G} = \mathbf{B} + \frac{2m^2}{e^2 N}\langle J\rangle\mathbf{M} = \mathbf{B} + \frac{m^2}{2\pi e^2 N}\frac{\mu_\mathrm{m}-1}{\mu_\mathrm{m}}\langle J\rangle\mathbf{B}. \tag{12}$$

Let us estimate the magnitude of the quasimagnetic field defined by formulas (11) and (12) for particles in condensed media (paramagnetic and ferromagnetic materials). The exchange integral J is determined by averaging the interaction operator of electrons e^2/r. Therefore, for our estimates, we can take

$$J \sim e^2/a_0, \quad N \sim 1/a_0^3,$$

where a_0 is the Bohr radius of atom. Hence,

$$\frac{m^2}{e^2 N}\langle J\rangle \sim m^2 a_0^2 \sim \frac{1}{\alpha^2},$$

where $\alpha \approx 1/137$ is the fine structure constant. According to this estimate,

$$\mathbf{H}^\mathrm{c}_\mathrm{eff} \sim \frac{\mu_\mathrm{m}-1}{\alpha^2\mu_\mathrm{m}}\mathbf{B} \sim 10^4\frac{\mu_\mathrm{m}-1}{\mu_\mathrm{m}}\mathbf{B}.$$

The actual value of $\mathbf{H}^\mathrm{c}_\mathrm{eff}$ may be several times or an order of magnitude smaller due to the incomplete overlapping of the wave functions in the exchange integral (see [21]). The estimate obtained for the quasimagnetic field $\mathbf{H}^\mathrm{c}_\mathrm{eff}$ completely agrees with the theory of ferromagnetism.

3. EQUATIONS OF MOTION OF PARTICLES AND SPIN

In the FW representation, we obtain operator equations of motion of particles and spin by calculating the commutators of the Hamiltonian $\mathcal{H}$ (defined by formulas (3) and (9)) with the operators $\boldsymbol{\pi}$ and **S**:

$$\frac{d\boldsymbol{\pi}}{dt} = i[\mathcal{H}, \boldsymbol{\pi}] - e\frac{\partial\mathbf{A}}{\partial t}, \quad \frac{d\mathbf{S}}{dt} = i[\mathcal{H}, \mathbf{S}]. \tag{13}$$

The general solution to the problem of the motion of a classical particle with spin in electromagnetic and gravitational fields is given in [1, 2]. The quantum mechanical equation of motion of relativistic particles

[3] One should take into account that the magnetic permeability depends on the magnetic field H.

[4] Under the Coulomb scattering, one can assume that electrons are "fast" ("slow") if their velocity is large (small) as compared to $v_0 = e^2/\hbar \approx c/137$ [21]. The quantity v_0 corresponds to the kinetic energy of electrons of $m v_0^2/2 = 13.5$ eV.

with arbitrary spin in an electromagnetic field, obtained by Derbenev and Kondratenko in [3], is given by[5]

$$\frac{d\boldsymbol{\pi}}{dt} = e\mathbf{E} + e[\mathbf{v}\times\mathbf{B}] - \frac{e}{2m}\nabla\left\{\left(g-2+\frac{2}{\gamma}\right)(\mathbf{S}\cdot\mathbf{B}) - (g-2)\frac{\gamma}{\gamma+1}(\mathbf{S}\cdot\mathbf{v})(\mathbf{v}\cdot\mathbf{B}) + \left(g-2+\frac{2}{\gamma+1}\right)(\mathbf{S}\cdot[\mathbf{E}\times\mathbf{v}])\right\}. \quad (14)$$

The form of the equations of motion of particles in polarized media does not differ from (14) if the expression for the Hamiltonian remains unchanged. In this case, the expression for the Hamiltonian is changed only for electrons and positrons.

Naturally, the Hamiltonians for particles with spins $S = 1/2$ and $S \geq 1$ are slightly different because Hamiltonian (3) contains terms that are quadratic in spin. However, as a rule, these terms are very small and can be neglected when analyzing the motion of particles. Within the accuracy adopted in this study, the equations of motion for particles with spins $S = 1/2$ and $S \geq 1$ coincide.

The equation of motion (14) of a particle yields a quantum mechanical expression for the force exerted on the particle (14) by the external field and matter. In addition to the Lorentz force, the particle is subject to a force due to the interaction between its magnetic moment and a nonuniform field (the Stern–Gerlach force [22]). This force is exerted by both magnetic and electric fields. Depending on the orientation of the magnetic moment (and, consequently, the spin), the Stern–Gerlach force either pulls a particle into a region of stronger field or expels it from this region. As a result, a beam of particles is split according to the polarizations of particles.

For electrons and positrons in polarized matter, the form of equations of motion is substantially changed. According to (9) and (13), the equation of motion for these particles is given by

$$\frac{d\boldsymbol{\pi}}{dt} = e\mathbf{E} + e[\mathbf{v}\times\mathbf{B}] - \frac{e}{2m}\nabla\{g(\mathbf{S}\cdot\mathbf{G}) + (g-1)(\mathbf{S}\cdot[\mathbf{E}\times\mathbf{v}])\}. \quad (15)$$

Analysis of (15) shows that, in magnetic crystals, where $\mu_m - 1 \gtrsim 1$, the Stern–Gerlach force exerted on electrons is substantially greater. For positrons, this effect is less pronounced.

As we have mentioned above, for particles with arbitrary spin, the motion of spin in a uniform electromagnetic field is described by the BMT equation [4]. When describing the effects associated with the nonuniformity of the field, one should take into consideration the terms that are quadratic in spin. These terms are contained in the Hamiltonian only for particles with spin $S \geq 1$. The equation of motion of spin with regard to these terms, which is given in [10], has the form

$$\frac{d\mathbf{S}}{dt} = \left(\frac{d\mathbf{S}}{dt}\right)_{\mathrm{BMT}} + \left(\frac{d\mathbf{S}}{dt}\right)_{\mathrm{q}},$$

$$\left(\frac{d\mathbf{S}}{dt}\right)_{\mathrm{BMT}} = \frac{e}{2m}\left\{\left(g-2+\frac{2}{\gamma}\right)[\mathbf{S}\times\mathbf{B}] - (g-2)\frac{\gamma}{\gamma+1}[\mathbf{S}\times\mathbf{v}](\mathbf{v}\cdot\mathbf{B}) + \left(g-2+\frac{2}{\gamma+1}\right)[\mathbf{S}\times[\mathbf{E}\times\mathbf{v}]]\right\},$$

$$\left(\frac{d\mathbf{S}}{dt}\right)_{q} = \frac{Q}{4s(2s-1)} \times\left(\left\{\left([\mathbf{S}\times\nabla] - \frac{\gamma}{\gamma+1}[\mathbf{S}\times\mathbf{v}](\mathbf{v}\cdot\nabla)\right),\right.\right.$$

$$\left.\left((\mathbf{S}\cdot\mathbf{E}) - \frac{\gamma}{\gamma+1}(\mathbf{S}\cdot\mathbf{v})(\mathbf{v}\cdot\mathbf{E}) + (\mathbf{S}\cdot[\mathbf{v}\times\mathbf{B}])\right)\right\}$$

$$+\left\{\left((\mathbf{S}\cdot\nabla) - \frac{\gamma}{\gamma+1}(\mathbf{S}\cdot\mathbf{v})(\mathbf{v}\cdot\nabla)\right), \quad (16)\right.$$

$$\left.\left([\mathbf{S}\times\mathbf{E}] - \frac{\gamma}{\gamma+1}[\mathbf{S}\times\mathbf{v}](\mathbf{v}\cdot\mathbf{E}) + [\mathbf{S}\times[\mathbf{v}\times\mathbf{B}]]\right)\right\}\right)$$

$$+\frac{e}{4m^2}\frac{\gamma}{\gamma+1}\left(\left\{[\mathbf{S}\times[\mathbf{v}\times\nabla]], \left[\left(g-1+\frac{1}{\gamma}\right)(\mathbf{S}\cdot\mathbf{B}) - (g-1)\frac{\gamma}{\gamma+1}(\mathbf{S}\cdot\mathbf{v})(\mathbf{v}\cdot\mathbf{B}) + \left(g-\frac{\gamma}{\gamma+1}\right)(\mathbf{S}\cdot[\mathbf{E}\times\mathbf{v}])\right]\right\} + \left\{(\mathbf{S}\cdot[\mathbf{v}\times\nabla]),\right.\right.$$

$$\left.\left.\left[\left(g-1+\frac{1}{\gamma}\right)[\mathbf{S}\times\mathbf{B}] - (g-1)\frac{\gamma}{\gamma+1}[\mathbf{S}\times\mathbf{v}](\mathbf{v}\cdot\mathbf{B}) + \left(g-\frac{\gamma}{\gamma+1}\right)[\mathbf{S}\times[\mathbf{E}\times\mathbf{v}]]\right]\right\}\right).$$

[5] In the present paper, we do not take into consideration the motion of particles due to the radiation reaction. This factor is significant only for large energies. In this case, the force of radiation reaction is mainly responsible for the deceleration of a particle and gives rise to a comparatively small component of the velocity in the direction perpendicular to the direction of its original motion.

The quantities $(d\mathbf{S}/dt)_{\mathrm{BMT}}$ and $(d\mathbf{S}/dt)_{\mathrm{q}}$ describe the motion of spin determined by the terms that are linear and quadratic in spin, respectively. This equation can be obtained by formulas (2), (3), and (13). Note that the operator $\epsilon' = \sqrt{m^2 + \boldsymbol{\pi}^2}$ has no influence on the motion of spin.

The essential feature of these equations for spin-1/2 particles is the vanishing of the terms that are quadratic in spin. For particles with $S \geq 1$, the terms quadratic in spin in Eq. (16) lead not only to the rotation but also to the oscillations of the spins of particles in a nonuniform field [23–25].

In polarized media, the form of the equations of motion of spin is not changed for all particles except for electrons and positrons. For nonrelativistic electrons and positrons, the equation of motion of spin obtained by Eqs. (9) and (13) is expressed as

$$\frac{d\mathbf{S}}{dt} = \frac{e}{2m}\{g[\mathbf{S}\times\mathbf{G}] + (g-1)[\mathbf{S}\times[\mathbf{E}\times\mathbf{v}]]\}. \quad (17)$$

The comparison of Eqs. (15) and (17) shows that the effect of the exchange interaction, which determines the quasimagnetic field **G**, on the motion of spin is much stronger than that on the motion of particles.

4. DISCUSSION OF THE RESULTS AND CONCLUSIONS

The investigation carried out shows that the motion of particles and spin in media with polarized electrons exhibits a number of features. Although the equations of motion of particles and spin in polarized media and in vacuum coincide (except for electrons and positrons), the magnitudes of the magnetic filed in these cases may substantially differ. The relation between the tangential components of the vector of magnetic induction in matter (B_τ) and in vacuum ($B_{0\tau}$) is as follows:[6]

$$B_\tau = \mu_{\mathrm{m}} B_{0\tau}.$$

For an appropriately chosen geometry of a crystal,[7] the magnetic induction in a crystal is several times greater than that in vacuum. For instance, for $\mu_{\mathrm{m}} = 10^3$, we have $B_0 = B_{0\tau} = 10^{-3}$ T and $B = B_\tau = 1$ T. These values of magnetic induction can be achieved due to the residual magnetization even in the absence of external field. Therefore, magnetic crystals can be effectively used for the rotation of the beam polarization vector.

Even for neutrons, whose magnetic moment is relatively small, for $B \sim 1$ T, the angle of rotation of the polarization vector per unit length is of the order of

$$\Delta\Phi/\Delta l \sim (c/v)\times 10^{-3}\ \text{rad/cm}.$$

For fast neutrons with $v/c \sim 10^{-2}$, the rotation of the polarization vector through an angle of about 1 rad is attained at a crystal length of $l \sim 10$ cm. For slow neutrons, this effect is still greater.

For example, positrons generated by the β^+ decay are usually relativistic. Therefore, we estimate the rotation angle of spin for these neutrons by the general formula (16). For $B = 1$ T and $\gamma = 5$, the rotation angle of the polarization vector per unit length is of the order of

$$\Delta\Phi/\Delta l \sim 1\ \text{rad/cm}.$$

If a beam of positrons passes through a crystal in a planar channeling mode, the vector of magnetic induction must lie in the plane containing the propagation direction of the beam (z axis) and the normal to the system of crystallographic planes (x axis). If the vector **B** is collinear to the y axis, then, due to the shift of the beam trajectory under a magnetic field, the arising electric component of the Lorentz field compensates for the magnetic component. As a result, the mean value of the Lorentz force is equal to zero. In this case, $\langle\mathbf{E}\rangle = -[\mathbf{v}\times\mathbf{B}]$ and the angle of spin rotation decreases, approximately by a factor of γ.

The rotation of the polarization vector in magnetic crystals reaches especially large values for nonrelativistic electrons. It follows from (8) and (9) that the angular velocity of the spin precession for nonrelativistic electrons is increased by a factor of $(c/v)^2$ due to the exchange interaction. For $B \sim 1$ T, we have

$$\Delta\Phi/\Delta l \sim (c/v)^3 \times 1\ \text{rad/cm}$$

in order of magnitude. In particular, when $v/c \sim 0.1$, we have

$$\Delta\Phi/\Delta l \sim 10^3\ \text{rad/cm}.$$

The use of magnetic crystals may also be sufficiently effective for the rotation of the polarization vector of relativistic electrons.

The polarization of a medium also significantly affects the motion of particles. In this case, the Stern–Gerlach force, which splits a beam according to the polarization of particles, is considerably increased. A large value of the magnetic-field gradient can be achieved both by placing a ferromagnetic sample in a strongly nonuniform external field and by appropriately shaping a sample (for example, making it in the form of a triangular prism) so that it guarantees a nonuniform magnetic field. However, the use of polarized media for splitting beams according to the polarization of particles is seriously hampered by the small value of the energy of interaction between the spin of electrons and a quasimagnetic field $\mathcal{H}^{(s)}$ (of the order of 1 eV or less) and a multiple scattering that increases the transverse energy of electrons of the beam. If the transverse

[6] One should take into account that the permeability μ_{m} of ferromagnetic materials depends on the strength of the magnetic field.

[7] It is desirable that the length of a crystal in the direction of the magnetic field should be much greater than its dimensions in two other directions. This fact guarantees that the demagnetizing factor is small.

energy of electrons in the beam is greater than $|\mathcal{H}^{(s)}|$, the splitting of the beam according to the polarization of particles becomes impossible.

We also mention other phenomena that are associated with the transmission of polarized beams of particles through polarized matter. The multiple scattering by the electrons of matter leads to the depolarization of a beam [26, 27]. The dependence of the cross section of scattering by polarized electrons on the polarization of the beam leads to the dichroism. As a result, one of polarization directions (either parallel or antiparallel to the field) in the beam leaving the matter becomes dominant. For an arbitrarily polarized beam, the dichroism leads to an additional rotation of the polarization vector [11, 15]. One should take into consideration the depolarization and dichroism in exact calculations. Note also that the transverse energy levels of electrons are split during channeling through magnetic crystals [28].

Thus, in this paper, we have found quantum mechanical equations of motion of particles and spin in media with polarized electrons. When deriving these equations, we took into account the exchange interaction for electron beams and the annihilation interaction for positron beams. For particles with $S \geq 1$, we took into account second-order terms in spin. The analysis of the equations obtained has shown that, in many cases, the exchange interaction several times increases the angular velocity of the rotation of the spin of electrons in polarized matter.

All the results obtained in this work are also valid for a beams of polarized nuclei.

ACKNOWLEDGMENTS

I am grateful to professor V.G. Baryshevsky for valuable remarks and discussion of the results.

REFERENCES

1. I. M. Ternov and V. A. Bordovitsyn, Usp. Fiz. Nauk **132**, 345 (1980) [Sov. Phys. Usp. **23**, 679 (1980)].
2. K. Yee and M. Bander, Phys. Rev. D **48**, 2797 (1993).
3. Ya. S. Derbenev and A. M. Kondratenko, Zh. Éksp. Teor. Fiz. **64**, 1918 (1973) [Sov. Phys. JETP **37**, 968 (1973)].
4. V. Bargmann, L. Michel, and V. L. Telegdi, Phys. Rev. Lett. **2**, 435 (1959).
5. D. M. Fradkin and R. H. Good, Rev. Mod. Phys. **33**, 343 (1961).
6. D. Zwanziger, Phys. Rev. **139**, B1318 (1965).
7. A. A. Pomeranskiĭ and I. B. Khriplovich, Zh. Éksp. Teor. Fiz. **113**, 1537 (1998) [JETP **86**, 839 (1998)].
8. D. P. Barber, Preprint DESY-98-096A (1998); E-print archives/physics/9901038.
9. A. A. Pomeransky and R. A. Sen'kov, Phys. Lett. B **468**, 251 (1999).
10. A. Ya. Silenko, Yad. Fiz. **64**, 1054 (2001) [Phys. At. Nucl. **64**, 983 (2001)].
11. V. G. Baryshevskiĭ, *Nuclear Optics of Polarized Media* (Énergoatomizdat, Moscow, 1995).
12. A. Ya. Silenko, Teor. Mat. Fiz. **105**, 46 (1995).
13. L. L. Foldy and S. A. Wouthuysen, Phys. Rev. **78**, 29 (1950).
14. V. G. Baryshevskiĭ, Dokl. Akad. Nauk BSSR **35**, 416 (1991).
15. V. G. Baryshevsky, Phys. Lett. A **171**, 431 (1992).
16. V. G. Baryshevsky, J. Phys. G: Nucl. Part. Phys. **19**, 273 (1993).
17. V. N. Baĭer, V. M. Katkov, and V. S. Fadin, *Emission of Relativistic Electrons* (Atomizdat, Moscow, 1973).
18. I. M. Ternov, V. R. Khalilov, and V. N. Rodionov, *Interaction of Charged Particles with Strong Electromagnetic Field* (Mosk. Gos. Univ., Moscow, 1982).
19. V. B. Berestetskiĭ, E. M. Lifshitz, and L. P. Pitaevskiĭ, *Quantum Electrodynamics*, 3rd ed. (Nauka, Moscow, 1989; Pergamon Press, Oxford, 1982).
20. A. Ya. Silenko, Zh. Éksp. Teor. Fiz. **114**, 1153 (1998) [JETP **87**, 629 (1998)].
21. L. D. Landau and E. M. Lifshitz, *Course of Theoretical Physics*, Vol. 3: *Quantum Mechanics: Non-Relativistic Theory*, 4th ed. (Nauka, Moscow, 1989; Pergamon, New York, 1977).
22. K. Heinemann, Preprint No. 96-11001; E-print archives/physics/9611001.
23. V. G. Baryshevskiĭ and A. A. Sokol'skiĭ, Pis'ma Zh. Tekh. Fiz. **6**, 1419 (1980) [Sov. Tech. Phys. Lett. **6**, 612 (1980)].
24. V. G. Baryshevskiĭ, *Channeling, Radiation, and Reactions in Crystals at High Energies* (Belorus. Gos. Univ., Minsk, 1982).
25. V. G. Baryshevsky and A. J. Shechtman, Nucl. Instrum. Methods Phys. Res. B **83**, 250 (1993).
26. V. L. Lyuboshits, Yad. Fiz. **32**, 702 (1980) [Sov. J. Nucl. Phys. **32**, 362 (1980)].
27. V. G. Baryshevsky, Nucl. Instrum. Methods Phys. Res. B **44**, 266 (1990).
28. V. I. Vysotskiĭ, R. N. Kuz'min, and N. V. Maksyuta, Poverkhnost, No. 12, 137 (1998).

Translated by I. Nikitin

Journal of Experimental and Theoretical Physics, Vol. 96, No. 4, 2003, pp. 617–628.
Translated from Zhurnal Éksperimental'noĭ i Teoreticheskoĭ Fiziki, Vol. 123, No. 4, 2003, pp. 696–709.

ATOMS, SPECTRA, RADIATION

Room-Temperature Fluctuations in the Fluorescence of a Single Polymer Molecule

I. S. Osad'ko

Lebedev Institute of Physics, Russian Academy of Sciences, Moscow, 119991 Russia

Received October 18, 2002

Abstract—A theoretical model of light absorption and emission by a polymer molecule has been developed using recent experimental data on the room-temperature fluctuations in the fluorescence intensity of single molecules of a PPV–PPyV copolymer containing several tens of chromophores. An analysis of the experimental data based on the proposed model shows evidence of a change in the conformation of a polymer molecule in the triplet state. By applying the theory to the PPV–PPyV copolymer, it is possible to determine the rate constants of the conformation variation, the rates of the transition from the singlet to the triplet state, and the lifetime of the triplet state of the molecule. The theory also predicts some new effects which can be verified by experiment. *© 2003 MAIK "Nauka/Interperiodica".*

1. INTRODUCTION

In past years, there has been successful expansion in the field of applications of single molecule spectroscopy (SMS). Five to ten years ago, the main objects of investigation for SMS were relatively simple molecules of the aromatic series, dissolved in a polymer cooled below 4.2 K [1, 2]. In recent years, the number of objects studied by SMS has rapidly increased. SMS techniques have proved to be effective for studying even very complicated molecular complexes, including those playing a significant role in biological systems.

It was established that SMS can also be effectively employed at room temperature. SMS provides important information about the dynamics of complex quantum systems at such temperatures, because spectral investigations traditional for the molecular ensembles can be supplemented by measurements of fluorescence fluctuations. This is a very important circumstance, since many objects, such as proteins and light-collecting systems of centers involved in photosynthesis, function in living organisms at room temperature.

A question always naturally arises as to what is implied by a single molecule. Indeed, why do a dozen anthracene molecules (chromophores) dissolved in a polymer constitute a molecular ensemble, the room-temperature fluorescence of which contains little information, whereas a polymer chain containing about a hundred such chromophores with a total molecular weight exceeding 20000 amu can be considered as a single molecule and the corresponding fluorescence spectrum is treated as providing for much more valuable information than the spectrum of a molecular ensemble?

In answering this question, it has to be recalled that a single molecule in SMS is probed by the light of a continuous laser and the information on the molecular dynamics is obtained by counting the photons emitted from the laser-excited molecule. In the early investigations considered in [1, 2], SMS was used to study molecules incapable of absorbing the second photon until the photon absorbed previously is emitted. Such a molecule can be called a single photon absorber and a single photon emitter. Unlike this, polymer molecules and light-collecting antennas of photosynthesis centers are characterized by so-called exciton-type absorption. This implies that many chromophores are involved in the absorption process, so that two or more photons can be simultaneously absorbed and the molecules can no longer be referred to as single absorbers. However, at a sufficiently small intensity of the exciting light, the exciton band may contain only one exciton and, hence, these molecules still remain single photon emitters and exhibit fluorescence fluctuations. Note that many complicated systems, such as semiconductors with quantum dots, can behave as single photon emitters. Some quantum dots exhibit intermittency in fluorescence [3–5]. One quantum dot in a semiconductor with such fluorescence may contain up to a million atoms, but the emission will obey the same laws as in any other single photon emitter.

The fluctuations in fluorescence have been observed in single quantum dots of various semiconductors [3–5], in single light-collecting antennas (LH2) of photosynthesis centers [6, 7], and in single molecules of polymers [8–11], dendrimers [12–14], and proteins [15–17]. Unfortunately, no theoretical models capable of explaining the observed fluorescence fluctuations have been proposed in the papers cited.

This paper is aimed at demonstrating, using a particular example, how the existing theory can be applied to the interpretation of experimental results on the fluctuating fluorescence. The experimental material represents data on the fluorescence of single molecules of a PPV–PPyV copolymer reported by Barbara and

1063-7761/03/9604-0617$24.00

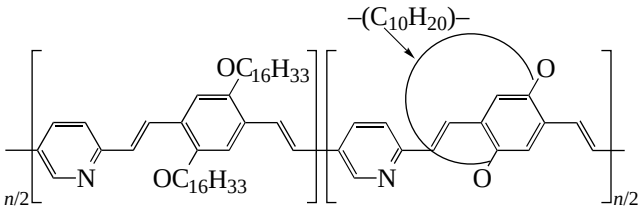


Fig. 1. Molecular structure of the PPV–PPyV copolymer studied in [8].

coworkers [8]. Their results will be treated within the framework of a theory developed in monograph [2]. Using this theory, it is possible to formulate a model of the electron excitation energy transfer in the polymer and to establish how changes in conformation of the polymer molecule can influence the fluorescence. Then, the observed fluorescence fluctuations will reveal the role of singlet and triplet levels of a molecule in changing the polymer conformation.

2. PRINCIPAL EXPERIMENTAL FACTS

The results of spectroscopic investigation of bulk polymer samples show that the light is absorbed by numerous chromophores, while the emission takes place after migration of the electron excitation energy toward local minima of the Franck–Condon surface of a polymer molecule [18]. For example, in a polymer with anthracene molecules attached as side chains to the backbone, the emission took place from the anthracene molecules occurring at the ends of the backbone [19]. Unfortunately, in bulk materials it is difficult to distinguish between the intermolecular and intramolecular energy transfer. For this reason, Barbara's team studied the energy transfer in single molecules of a poly(*para*-phenylene vinylene)–poly(*para*-pyridine vinylene) (PPV–PPyV) copolymer in which the electron energy transfer is known to be the intramolecular process. This process was studied by measuring the fluorescence of individual copolymer molecules. The structural formula of such a molecule is presented in Fig. 1. The PPV–PPyV copolymer has a molecular weight of about 20000 amu and contains 80–100 chromophores capable of absorbing light. Figure 2 shows the absorption and emission bands of this copolymer.

Since the optical absorption and emission bands of the polymer molecule are structureless, no information concerning the molecular dynamics can be extracted from these spectra. The principal difference between fluorescence from a single polymer molecule and that from an ensemble of such molecules is manifested in the kinetics of fluorescence, rather than in the spectral characteristics (as can be seen, the two emission bands virtually coincide). Indeed, the intensity of fluorescence from an ensemble of polymer molecules is time-independent within tens of seconds, whereas the intensity of emission from a single molecule fluctuates as depicted in Fig. 3. It was demonstrated [8] that the character of the quantum intensity transient (QIT) of the fluorescence from a single molecule excited by the light with $\lambda = 514$ nm remains the same. These very fluctuations in the intensity of the broad fluorescence band make it possible to study the dynamics of energy transfer in a single polymer molecule at room temperature.

Figures 4 and 5 show the results of statistical data processing for several dozens of QITs measured from single polymer molecules [8]. According to these data, the fluorescence from each individual polymer molecule exhibits intermittency, whereby the molecule passes from the emissive ("on") to nonemissive ("off") state. In the on state, the fluorescence intensity can reach one of the two levels, conditionally referred to as intermediate (I_1) and high (I_2). Therefore, the fluorescence is a twofold process with respect to intensity. All the experimental facts summarized in Section 2 underlie the theoretical model formulated below.

3. DOUBLE FLUORESCENCE MODEL

The process of fluorescence with two intensity levels is known not only for polymer molecules. Such a pattern was also observed for cw-laser-excited single

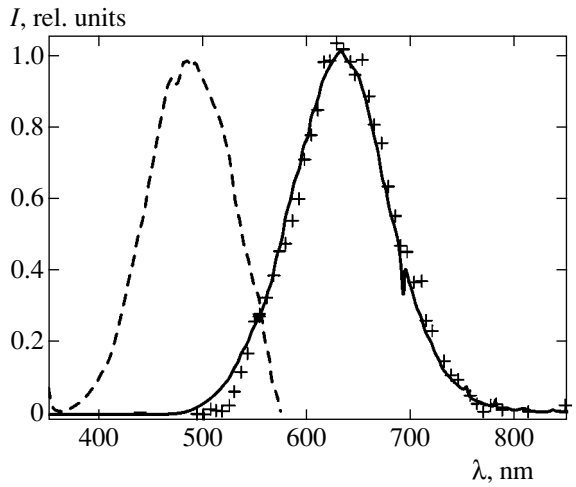


Fig. 2. The absorption (dashed curve) and emission (solid curve) bands of a concentrated sample of the PPV–PPyV copolymer. Crosses show the fluorescence band of a single PPV–PPyV molecule [8].

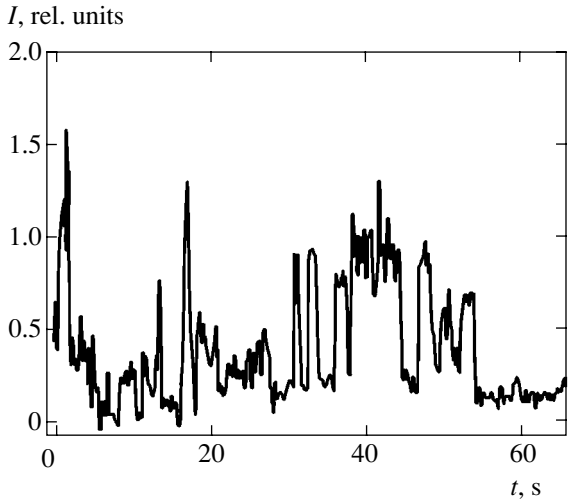


Fig. 3. Fluorescence kinetics (quantum intensity transient) of a single PPV–PPyV molecule continuously excited at λ = 457 nm (sampling window, 100 ms) [8].

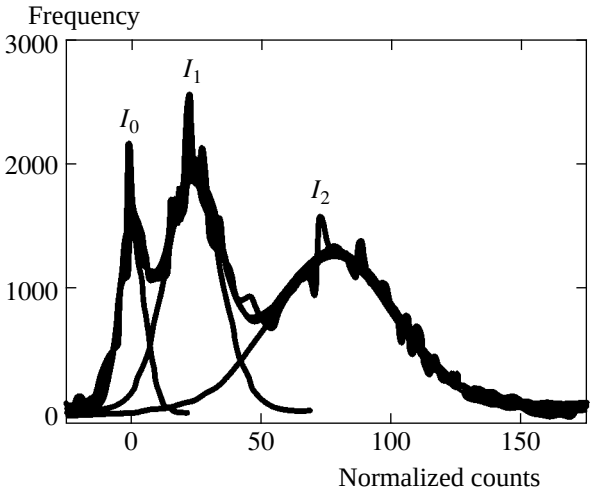


Fig. 4. Three fluorescence intensity levels of single PPV–PPyV molecule: zero (I_0), intermediate (I_1), and high (I_2) with the normalized counts (per 100-ms sampling window) 0, 25, and 75, respectively [8].

molecules of terrylene dissolved in polyethylene [20]. The phenomenon of double fluorescence can be described using the theory developed in [2, 21]. Figure 6 shows the experimental results obtained in [20] and presents an energy band diagram explaining these data in terms of the theoretical model.

The proposed model takes into account the interaction of the chromophore of a terrylene molecule with

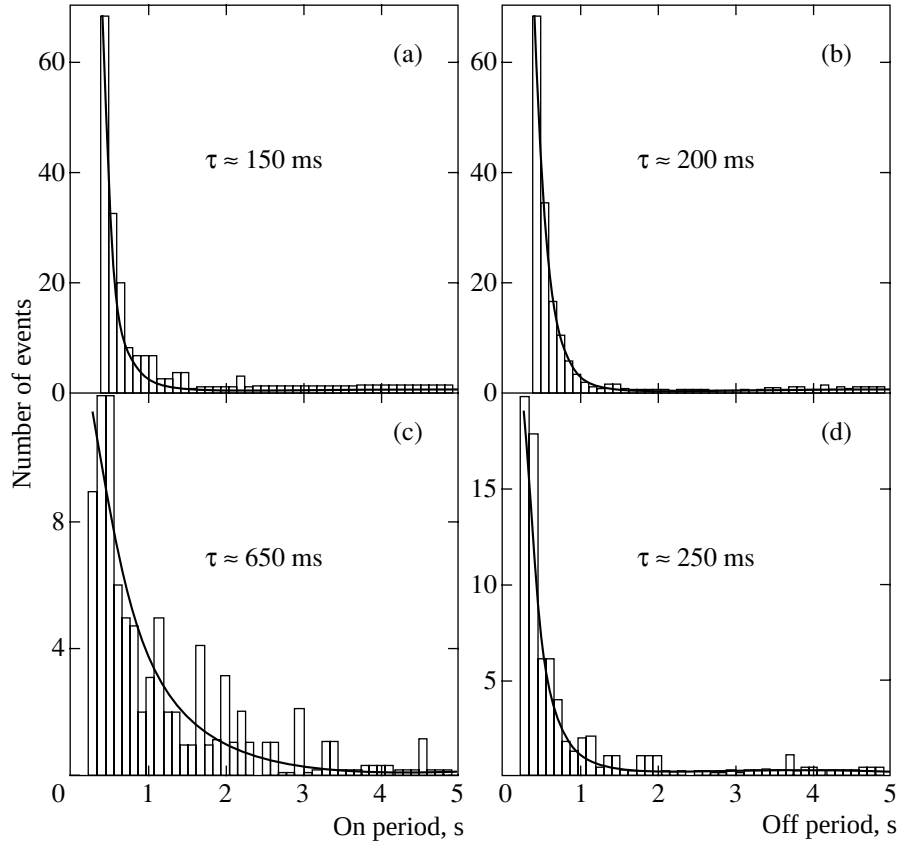


Fig. 5. Distributions of the on and off periods (with the average duration τ) in QITs observed for an excitation power of P = 1500 (a, b) and 300 W/cm^2 (c, d) [8].

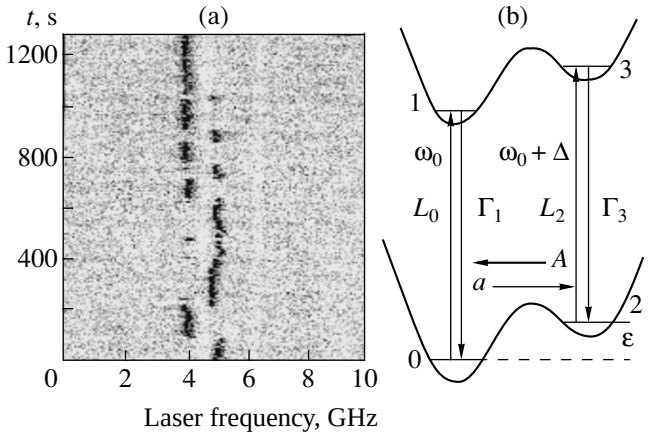


Fig. 6. (a) Quantum jumps in the absorption band of single terrylene molecules in polyethylene at 1.8 K [20] and (b) an energy band diagram explaining this behavior (*A* and *a* are the probabilities of tunneling transitions) [2, 21]. The experimental data were obtained by multiply repeated scanning over the exciting laser frequency during a single laser scan time of $t_{sc} \approx 1$s.

the so-called two-level system of the polymer. This two-level system models a change in conformation of the impurity complex comprising a terrylene molecule and the nearest environment of polymer molecules. The rates (probabilities) of the direct and reverse conformation transitions are *A* and *a*, respectively (Fig. 6b). According to the theory [2, 21], the coefficient of absorption of such a system is described by a simple expression,

$$k(\omega, t) = [1 - \rho(t)]L_0(\omega) + \rho(t)L_2(\omega), \qquad (1)$$

where

$$L_0(\omega) = 2\chi^2 \frac{1/T_2}{(\omega - \omega_0)^2 + 1/T_2^2},$$
$$L_2(\omega) = 2\chi^2 \frac{1/T_2}{(\omega - \omega_0 - \Delta)^2 + 1/T_2^2} \qquad (2)$$

are the envelopes of the optical lines corresponding to the photoinduced transitions 1–0 and 3–2 depicted in Fig. 6, $\chi = \mathbf{d} \cdot \mathbf{E}/\hbar$ is the Rabi frequency, T_2 is the optical dephasing time, and

$$\rho(t_{sc} + t_0) = f(T) + [\rho(t_0) - f(T)]\exp[-(A + a)t_{sc}] \qquad (3)$$

is the probability of finding the system in quantum state 2 at the time instant $t = t_{sc} + t_0$. Here, t_0 is the initial time moment, t_{sc} is the time of counting photons, and $f(T) = [\exp(-\varepsilon/kT) + 1]^{-1}$ is the same probability under the conditions of a thermal equilibrium established in the system at $t_{sc}(A + a) = t_{sc}R \gg 1$.

The vertical bands observed in Fig. 6a are called the spectral trails. The degree of blackening in such a trail is proportional to the number of emitted photons which, in turn, is proportional to the probability of absorption of a laser radiation quantum. According to the experimental data presented in Fig. 6a, a single molecule of terrylene randomly switches the resonance frequency (i.e., jumps from state 0 to state 2 and back). The initial condition in formula (3) is selected as follows (see Fig. 6b). If the molecule jumps from state 2 to state 0 at $t_0 = 0$, we use formulas (1) and (3) with $\rho(t_0) = \rho(0) = 0$, and when it jumps back (at a random time instant $t_0 = t_1$) from state 0 to state 2, we must take $\rho(t_0) = \rho(t_1) = 1$. Accordingly, we obtain two possible absorption coefficients,

$$k = k_0 = \{1 - f[1 - \exp(-Rt_{sc})]\}L_0 + f[1 - \exp(-Rt_{sc})]L_2 \qquad (4)$$

and

$$k = k_2 = [1 - f - (1 - f)\exp(-Rt_{sc})]L_0 + [f + (1 - f)\exp(-Rt_{sc})]L_2, \qquad (5)$$

which implies that the absorption coefficient fluctuates. These fluctuations in the absorption coefficient calculated by formulas (4) and (5) are depicted in Fig. 7. Since the absorption at the frequencies of transitions 1–0 and 3–2 is different, the intensity of fluorescence varies as well and we obtain the pattern of double fluorescence. The ratio of the integral fluorescence intensities for a large scan time equals the ratio of peak intensities in Fig. 7c. Since the fluorescence bands corresponding to the excitation of transitions 1–0 and 3–2 virtually coincide, no fluctuations in the fluorescence intensity are observed for $Rt_{sc} = 10^2$. On the contrary, maximum fluctuations will be observed at so small a scan time as $Rt_{sc} = 10^{-3}$.

4. FLUCTUATING FLUORESCENCE MODEL

The above model cannot explain the disappearance of fluorescence (the appearance of dark levels) at certain time instants, that is, the presence of off periods. This is possible within the framework of the following model.

Since a polymer molecule is composed of many chromophores, it is expedient to first consider a single chromophore molecule. Figure 8 shows an energy level diagram which is typical of organic molecules. Here, vertical arrows indicate the quantum transitions taking place in the laser-excited molecule. The emission from such a molecule under continuous irradiation was studied both experimentally [22] and theoretically [2, 22, 23]. The pattern of fluorescence in this system consists of on periods separated by off periods. This phenomenon, whereby photons tend to group in time, is called photon bunching. The on period duration is determined by the lifetime of a molecule in the singlet states 0 and 1, whereby the quantum jumps are performed between the ground and excited electron states with absorption of the laser radiation quanta and emission of the fluorescence photons. Then the molecule jumps into a long-

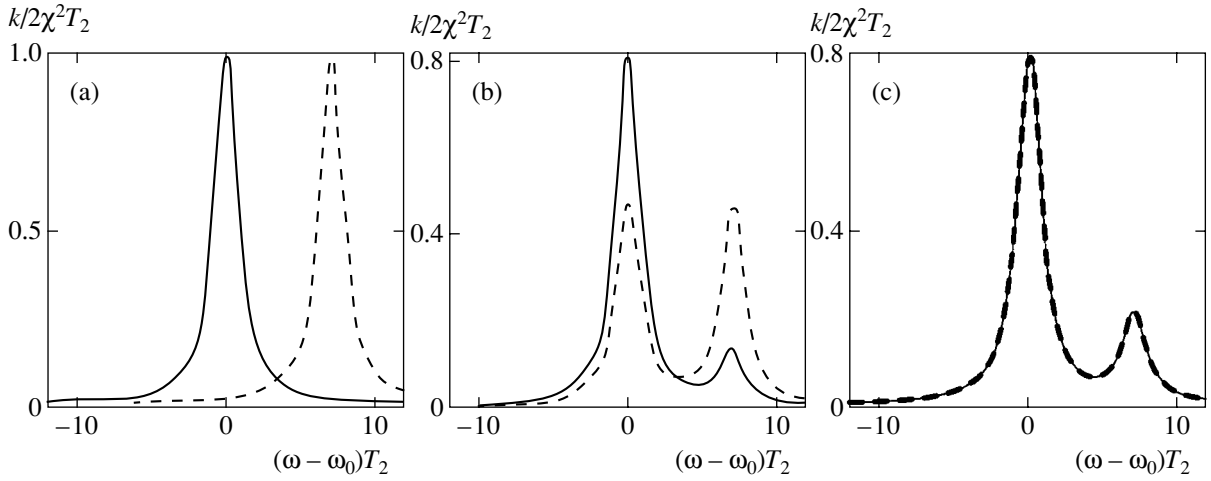


Fig. 7. Fluctuations in the absorption coefficients for various times of scanning over the laser frequency: $Rt_{sc} = 10^{-3}$ (a), 1 (b), and 10^2 (c). Solid and dashed curves show optical envelopes of the two types corresponding to the experimental lines averaged at a given photon counting time t_{sc}.

lived triplet state 2, in which it occurs for a long time without photon emission. Thus, the lifetime in the triplet state corresponds to the off interval (nonemissive state). Therefore, this model can explain alternation of the on and off periods, but does not account for the double fluorescence. Evidently, an adequate model of fluorescence from a polymer molecule must combine the properties of both partial models considered above.

5. THEORETICAL MODEL DESCRIBING THE DYNAMICS OF FLUORESCENCE OF A SINGLE POLYMER MOLECULE

An adequate theoretical model was selected based on the QIT data. The key elements determining a physical model explaining the observed QIT peculiarities are as follows:

(i) there must be three fluorescence intensity levels—zero, intermediate, and high (Fig. 4);

(ii) jumps from the intermediate (I_1) to high (I_2) intensity level and back usually follow the intensity drops to zero (Fig. 3);

(iii) the inverse value of the on period duration is proportional to the excitation power, while the off period duration is independent of this power (Fig. 5).

(iv) the duration of both on (emissive) and off (nonemissive) periods is on the order of seconds (Figs. 3 and 5).

A polymer chain in which N regularly arranged light-absorbing chromophores are arranged represents a one-dimensional crystal model. Apparently, each chromophore can be described by a triad of levels depicted in Fig. 8. Owing to the interaction between chromophores, N excited singlet levels form a band of singlet excitons responsible for the optical absorption, while $3N$ triplet levels account for the band of triplet excitons. Occurring in the latter state, the polymer produces no fluorescence.

The interaction of solvent molecules with the polymer chromophores leads to the appearance of defect levels below the bottom of the singlet and triplet exciton bands, to which the singlet and triplet electron excitations will decay, respectively. These defect levels play the role of traps for the electron excitations. The shape of the fluorescence bands for the shallow traps, depending on the temperature and the depth of these traps, was studied in [24].

Strictly speaking, several defect levels can occur below the bottom of the singlet exciton band. However, the existence of clearly pronounced fluctuations in the fluorescence intensity followed by drops to zero level (Fig. 3) indicates that the polymer molecule emits photons via a single channel, probably corresponding to the lowest defect level. The model with a single defect level seems to be quite adequate. Indeed, if there were two independent photon emitters, the resulting pattern would correspond to the sum of two QITs of the type depicted in Fig. 3, randomly shifted in the time scale. Then, the number of levels in the fluorescence intensity in the resulting QIT would be greater than that observed in Figs. 3 and 4.

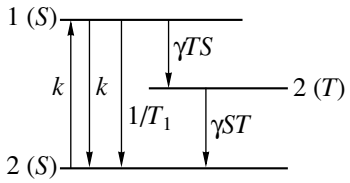


Fig. 8. Energy diagram for a molecule with two singlet (0 and 1) and one triplet (2) levels corresponding to three spin states of the molecule (T_1 is the lifetime of level 1).

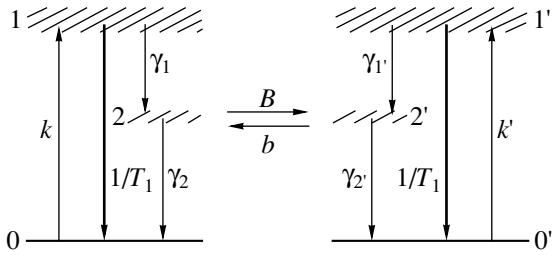


Fig. 9. Energy diagram of a PPV–PPyV molecule subject to conformational changes.

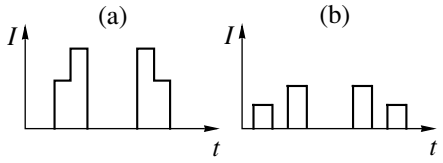


Fig. 10. Two types of the jumps between intermediate and high fluorescence levels corresponding to conformational changes taking place in the (a) ground and (b) triplet state of the polymer molecule.

The existence of an exciton band and defect levels, which play an important role in the luminescence kinetics, is practically not manifested in the shapes of the room temperature absorption and emission bands. The Gaussian shape of the bands presented in Fig. 2 is evidence that a strong electron–phonon interaction "masks" the structure of electron levels related to the exciton band and defect levels. The shape of the bands in Fig. 2 can be described using the following relations valid for the impurity centers with strong electron–phonon interaction [2]:

$$I^{a,f}(\omega) = \frac{1}{\sqrt{2\pi D}}\exp\left[-\frac{(\omega-\omega_0\mp C)^2}{2D}\right]. \qquad (6)$$

Here, $2C$ is the Stokes shift and $\Delta\omega^{1/2} = \sqrt{2D\ln 2}$ is the band halfwidth determined by the formulas

$$C = \sum_{q=1}^{N_0}\frac{a_q^2}{2}\nu_q, \quad D = \sum_{q=1}^{N_0}\frac{a_q^2}{2}\nu_q^2\coth\frac{\hbar\nu_q}{2kT}, \qquad (7)$$

where ν_q are the frequencies of normal vibrations of the polymer molecule, a_q are the shifts of their equilibrium positions, and N_0 is the number of vibrational modes.

Let us estimate the degree of correspondence between formulas (7) and the experimental values C = 2604 cm^{-1} and $\Delta\omega_{1/2}$ = 2500 cm^{-1} determined from the curves in Fig. 2. As is known, the extensions of both the absorption spectrum and the emission band in the aromatic molecules are determined by vibronic phototransitions with a frequency of intramolecular vibrations of 1400–1700 cm^{-1}. Indeed, these characteristic vibrations are most actively manifested in the optical spectra of aromatic substances. Such frequencies at room temperature obey the condition $\hbar\nu_q/kT \ll 1$ and, hence, formulas (7) yield the following relation between the Stokes shift and the band halfwidth:

$$D = \sum_{q=1}^{N_0}\frac{a_q^2}{2}\nu_q^2 = C\nu_e, \qquad (8)$$

where ν_e is the effective frequency of intramolecular vibrations determining the extension of the optical bands. Substituting the experimental values of C and D into this relation, we obtain ν_e = 1750 cm^{-1}, which agrees with the frequency most effectively manifested in the optical spectrum of aromatic substances. This estimate confirms our assumption that the singlet exciton band width is significantly narrower than the width of the optical bands, the latter being determined by the electron–phonon interaction.

According to the above considerations, the energy level diagram of the polymer molecule studied can be represented as in Fig. 9, where the system has two conformations corresponding to the sets of levels 0, 1, 2 and 0', 1', 2'. Levels 1 and 1' correspond to the singlet exciton and singlet trap in two conformations. Evidently, the transitions from states 1 and 1' to the ground state are manifested by fluorescence with different levels of intensity. Levels 2 and 2' correspond to the triplet exciton and triplet trap in these conformations. Only jumps of the polymer molecule between conformations can explain the phenomenon of double fluorescence, while the triplet states in these conformations are necessary for explaining the existence of the off periods.

The different probabilities of transitions in the scheme of Fig. 9 are indicated by different thicknesses of the corresponding arrows: greater thickness corresponds to higher transition probability. Therefore, we adopt the following hierarchy of the relaxation constants:

$$1/T_1 \gg k, k' \gg \gamma_2, \gamma_{2'}, \gamma_1, \gamma_{1'}, B, b. \qquad (9)$$

The absence of induced fluorescence in Fig. 9 is related to the fact that the absorption is mediated by excitons whose energy is thermalized before the emission event takes place. Under these conditions, no resonance fluorescence takes place.

Apparently, conformational changes in the time scale of seconds can proceed either in the ground or in the long-lived triplet state of the polymer molecule. If this change in conformation were to take place in the ground state, the transition from intermediate to high fluorescence level would proceed in a jumplike manner so that the system would not occur in the off state, as depicted in Fig. 10a. This contradicts the fact that the fluorescence intensity changes from intermediate to

high level via drops to the zero level, that is, by the scheme depicted in Fig. 10b. For this reason, Fig. 9 indicates that a change in the conformation takes place when the polymer molecule occurs in the triplet state of the trap.

Balance equations for the probability of finding a polymer molecule in one of the six quantum states according to Fig. 9 are as follows:

$$\begin{aligned}
\dot\rho_1 &= -(1/T_1 + \gamma_1)\rho_1 + k\rho_0,\\
\dot\rho_0 &= \rho_1/T_1 - k\rho_0 + \gamma_2\rho_2,\\
\dot\rho_2 &= \gamma_1\rho_1 - (\gamma_2 + B)\rho_2 + b\rho_{2'},\\
\dot\rho_{1'} &= -(1/T_1 + \gamma_{1'})\rho_{1'} + k'\rho_{0'},\\
\dot\rho_{0'} &= \rho_{1'}/T_1 - k'\rho_{0'} + \gamma_{2'}\rho_{2'},\\
\dot\rho_{2'} &= B\rho_2 + \gamma_{1'}\rho_{1'} - (\gamma_{2'} + b)\gamma_{2'}.
\end{aligned} \tag{10}$$

This set of equations describes both the fast relaxation in the time scale of T_1 and the slower process in the scale of singlet–triplet transitions and the transitions between conformations. In the experiment under consideration, the information on fast relaxation was lost because of insufficiently high temporal resolution of the experimental setup, which was capable of measuring only slow relaxation. If we are interested only in slow relaxation in the polymer molecule, the system of equations (10) can be simplified by taking $\dot\rho_1 = \dot\rho_{1'} = 0$. Then, the first and fourth equations yield

$$\rho_1 = \frac{kT_1}{1 + \gamma_1 T_1}\rho_0, \quad \rho_{1'} = \frac{k'T_1}{1 + \gamma_{1'}T_1}\rho_{0'}. \tag{11}$$

Substituting these expressions into the remaining four equations, we arrive at the following system:

$$\begin{aligned}
\dot\rho_0 &= -k_e\rho_0 + \gamma_2\rho_2,\\
\dot\rho_2 &= k_e\rho_0 - (\gamma_2 + B)\rho_2 + b\rho_{2'},\\
\dot\rho_{0'} &= -k_e'\rho_{0'} + \gamma_{2'}\rho_{2'},\\
\dot\rho_{2'} &= k_e'\rho_{0'} + B\rho_2 - (\gamma_{2'} + b)\rho_{2'},
\end{aligned} \tag{12}$$

where the coefficients

$$k_e = \frac{\gamma_1 T_1}{1 + \gamma_1 T_1}k, \quad k_e' = \frac{\gamma_{1'}T_1}{1 + \gamma_{1'}T_1}k' \tag{13}$$

describe the rate of effective pumping of the triplet states via the excited singlet states in both conformations. Equations (13) describe slow relaxation in the system after fast population of levels 1 and 1'. Using this system, it is possible to calculate distribution of the on and off periods with respect to duration and determine their average values. Since the polymer molecule can occur in two fluorescent conformations, we must observe both on and off states of two types.

6. DISTRIBUTION OF THE ON AND OFF STATES

On states. The probability of finding a polymer molecule in each of the two possible on states is determined by the formulas

$$\rho_{\text{on}}^{(1)} = \rho_0 + \rho_1, \quad \rho_{\text{on}}^{(1)} = \rho_{0'} + \rho_{1'}. \tag{14}$$

The fluorescence intensity is proportional to the probability of finding the molecule in one of the two on states and vice versa; each probability corresponds to one of the two emission intensities. In the set of Eqs. (10), there are terms responsible for the decay and pumping of the on states. The decay of the on states is determined by the set of Eqs. (10) with excluded terms describing population of these states:

$$\dot\rho_{\text{on}}^{(1)} = -\gamma_1\rho_1, \quad \dot\rho_{\text{on}}^{(2)} = -\gamma_{1'}\rho_{1'}. \tag{15}$$

Taking into account expressions (11) and (14), we obtain the formulas

$$\begin{aligned}
\rho_0 &= \frac{1 + \gamma_1 T_1}{1 + \gamma_1 T_1 + kT_1}\rho_{\text{on}}^{(1)},\\
\rho_{0'} &= \frac{1 + \gamma_{1'}T_1}{1 + \gamma_{1'}T_1 + k'T_1}\rho_{\text{on}}^{(2)}.
\end{aligned} \tag{16}$$

Substituting these formulas into expressions (15), we arrive at the set of equations determining decay of the on states:

$$\dot\rho_{\text{on}}^{(1)} = -\frac{\rho_{\text{on}}^{(1)}}{\tau_{\text{on}}^{(1)}}, \quad \dot\rho_{\text{on}}^{(2)} = -\frac{\rho_{\text{on}}^{(2)}}{\tau_{\text{on}}^{(2)}}, \tag{17}$$

where

$$\begin{aligned}
\frac{1}{\tau_{\text{on}}^{(1)}} &= \frac{\gamma_1 T_1}{1 + \gamma_1 T_1 + kT_1}k,\\
\frac{1}{\tau_{\text{on}}^{(2)}} &= \frac{\gamma_{1'}T_1}{1 + \gamma_{1'}T_1 + k'T_1}k'.
\end{aligned} \tag{18}$$

Functions describing the distribution of the on periods with respect to their duration can be determined by solving Eqs. (17):

$$\begin{aligned}
w_{\text{on}}^{(1)}(t) &= \frac{1}{\tau_{\text{on}}^{(1)}}\exp\left(-\frac{t}{\tau_{\text{on}}^{(1)}}\right),\\
w_{\text{on}}^{(2)}(t) &= \frac{1}{\tau_{\text{on}}^{(2)}}\exp\left(-\frac{t}{\tau_{\text{on}}^{(2)}}\right).
\end{aligned} \tag{19}$$

According to these results, the distributions of on periods corresponding to the fluorescence intensities I_1 and I_2 can be described by the exponents with different indices. If an experiment could provide separately measured distributions of the on periods corresponding to different fluorescence intensity levels, then it would be possible to find directly the experimental values of $\tau_{on}^{(1)}$ and $\tau_{on}^{(2)}$ having a clear physical meaning of the average on period durations for the medium and high fluorescence levels. However, the experiments performed in [8] gave a total distribution of the on period durations of

$$w_{on}^{(1)}(t) + w_{on}^{(2)}(t), \tag{20}$$

presented in Figs. 5a and 5c, from which it follows that the inverse average on period duration is a linear function of the excitation intensity. This conclusion fully agrees with formulas (18) if we take into account that $kT_1 \ll 1$ and $k'T_1 \ll 1$. This dependence of the on period on the pumping level additionally confirms that conformational changes take place in the triplet state. If it were in the ground state, the on period duration would be independent of the pumping intensity.

Using the simplified formula

$$\frac{1}{\tau_{on}} \approx \gamma_1 T_1 k = \frac{1}{150\ \text{ms}}, \tag{21}$$

it is possible to estimate the rate γ_1 of intercombination transitions for the given polymer. Measuring the emission from a single polymer molecule, the electron multiplier in [8] detected about 4000 photons per second. For a quantum efficiency of the photon detection 10^{-3} and the quantum yield of fluorescence on the order of 0.1, we obtain $k = 4000 \times 10^3 \times 10 = 4 \times 10^7\ \text{s}^{-1}$. Substituting this value into the approximate formula (21), we obtain an estimate of the intercombination transitions probability in the given polymer molecule (at $T_1 = 10^{-9}$): $\gamma_1 = 1.6 \times 10^2\ \text{s}^{-1}$. Thus, the singlet–singlet transition is a million times more probable than the singlet–triplet transition.

Off states. Now it will be shown that the theory predicts the existence of two types of off state, namely, those following the on periods corresponding to the medium and high fluorescence intensities. Evidently, the probability of finding a molecule in the "dark" off state is

$$\rho_{off} = \rho_2 + \rho_{2'}. \tag{22}$$

In order to write equations determining decay of the off state, it is necessary to reject all terms in Eqs. (12) which increase the population of levels 2 and 2'. This yields the system of equations

$$\begin{aligned} \dot{\rho}_2 &= -(\gamma_2 + B)\rho_2 + b\rho_{2'}, \\ \dot{\rho}_{2'} &= B\rho_2 - (\gamma_2' + b)\rho_{2'}, \end{aligned} \tag{23}$$

upon solution of which we obtain the expressions

$$\begin{aligned} \rho_2(t) &= \frac{(\lambda_1 - b - \gamma_{2'})\rho_2(0) - b\rho_{2'}(0)}{\lambda_1 - \lambda_2} e^{-\lambda_1 t} \\ &\quad - \frac{(\lambda_2 - b - \gamma_{2'})\rho_2(0) - b\rho_{2'}(0)}{\lambda_1 - \lambda_2} e^{-\lambda_2 t}, \end{aligned} \tag{24}$$

$$\begin{aligned} \rho_{2'}(t) &= \frac{(\lambda_1 - B - \gamma_2)\rho_{2'}(0) - B\rho_2(0)}{\lambda_1 - \lambda_2} e^{-\lambda_1 t} \\ &\quad - \frac{(\lambda_2 - B - \gamma_2)\rho_{2'}(0) - B\rho_2(0)}{\lambda_1 - \lambda_2} e^{-\lambda_2 t}, \end{aligned} \tag{25}$$

where

$$\begin{aligned} \lambda_{1,2} &= \frac{B + b + \gamma_2 + \gamma_{2'}}{2} \\ &\quad \pm \sqrt{\left(\frac{B + \gamma_2 - b - \gamma_{2'}}{2}\right)^2 + Bb} \end{aligned} \tag{26}$$

are the roots of the determinant of Eqs. (23). Substituting this solution into formula (22), we arrive at an expression for the probability of finding a polymer molecule in the off state:

$$\begin{aligned} \rho_{off}(t) &= \left[\frac{(\lambda_1 - b - \gamma_{2'})e^{-\lambda_1 t} - (\lambda_2 - b - \gamma_{2'})e^{-\lambda_2 t}}{\lambda_1 - \lambda_2} \right. \\ &\quad \left. + \frac{B(e^{-\lambda_2 t} - e^{-\lambda_1 t})}{\lambda_1 - \lambda_2}\right]\rho_2(0) \\ &\quad + \left[\frac{(\lambda_1 - B - \gamma_2)e^{-\lambda_1 t} - (\lambda_2 - B - \gamma_2)e^{-\lambda_2 t}}{\lambda_1 - \lambda_2} \right. \\ &\quad \left. + \frac{b(e^{-\lambda_2 t} - e^{-\lambda_1 t})}{\lambda_1 - \lambda_2}\right]\rho_{2'}(0). \end{aligned} \tag{27}$$

If the molecule jumped into the off state from the on state with $\rho_{on}^{(1)}$, we have to use the above formula with the initial condition

$$\rho_2(0) = 1, \quad \rho_{2'}(0) = 0. \tag{28}$$

Let us denote the corresponding solution by $\rho_{off}^{(1)}$. If the molecule jumped into the off state from the on state with $\rho_{on}^{(2)}$, we have to use another initial condition,

$$\rho_2(0) = 0, \quad \rho_{2'}(0) = 1, \tag{29}$$

which yields the corresponding solution denoted by $\rho_{off}^{(2)}$. The resulting formulas for the off-state probabili-

ties $\rho_{off}^{(1)}$ and $\rho_{off}^{(2)}$ can be written, taking into account the relation $\lambda_1 + \lambda_2 = b + B + \gamma_2 + \gamma_{2'}$, in the following simple form:

$$\rho_{off}^{(1)}(t) = \frac{\gamma_2 - \lambda_2}{\lambda_1 - \lambda_2}\exp(-\lambda_1 t) - \frac{\gamma_2 - \lambda_1}{\lambda_1 - \lambda_2}\exp(-\lambda_2 t), \quad (30)$$

$$\rho_{off}^{(2)}(t) = \frac{\gamma_{2'} - \lambda_2}{\lambda_1 - \lambda_2}\exp(-\lambda_1 t) - \frac{\gamma_{2'} - \lambda_1}{\lambda_1 - \lambda_2}\exp(-\lambda_2 t). \quad (31)$$

According to these expressions, the distribution of the off periods following the on periods corresponding to the fluorescence intensity levels I_1 and I_2 is biexponential, provided that the roots λ_1 and λ_2 are significantly different.

7. FLUORESCENCE INTENSITY

The intensity of fluorescence is proportional to the effective absorption coefficient, reflecting both the exciton absorption and the efficiency of energy transfer to the luminescent trap of a polymer molecule. Since the polymer molecule performs quantum jumps between conformations, the effective absorption coefficient will fluctuate. It is these fluctuations that appear as quantum intermittency in the fluorescence intensity. As demonstrated in [21], the fluctuating absorption coefficient of a single molecule is simply related to a two-photon correlator, theoretically described in [2], or to an experimentally measured autocorrelation function [1, 25, 26]. The latter function is determined as [1]

$$g^{(2)}(\tau) = \lim_{t\to\infty}\frac{\langle I(t)I(t+\tau)\rangle}{\langle I(t)I(t+\infty)\rangle} = \lim_{t\to\infty}\frac{\langle I(t)I(t+\tau)\rangle}{\langle I(t)\rangle^2}, \quad (32)$$

where $I(t)$ is the QIT of the fluorescence. According to this formula, the autocorrelation function is measured by shifting the QIT by τ in the time scale and counting the number of photon pairs (proportional to the QIT length and the degree of coincidence of the on periods in the two QITs).

Here, a question arises as to how the autocorrelation function can be calculated theoretically. It was demonstrated [2, 23] that this function is related to a two-photon correlator $p(\tau)$:

$$\frac{p(\tau)}{p(\infty)} = g^{(2)}(\tau). \quad (33)$$

The total two-photon correlator is determined as the counting rate of the photon pairs separated by the time interval τ. This value can be calculated for a given microscopic model of the emitting system. As was demonstrated in [2], the total two-photon correlator in the system with a single fluorescent state is expressed as

$$p(\tau) = \frac{\rho_1(\tau)}{T_1}, \quad (34)$$

where $\rho(\tau)$ is the probability of finding the system in the fluorescent state upon elapse of time τ after photon emission. For a molecule with the system of levels depicted in Fig. 9, it is easy to establish that, for $kT_1 \ll 1$, the above probability is $\rho_1(\tau) \approx kT_1\rho_0(\tau)$, so that

$$p(\tau) \approx k\rho_0(\tau). \quad (35)$$

This implies that the two-photon correlator is proportional to the absorption coefficient. Therefore, the intensity of fluorescence is proportional to this correlator.

In the case under consideration, there are two fluorescent states and, hence, two photoemission processes. Then, we have to answer the question as to which probabilities will enter the expression for a two-photon correlator.

For definiteness, let us consider the case when the fluorescence with intensities I_1 and I_2 proceeds from states 1 and 1', respectively. Assume that the fluorescence component with intensity I_2 is removed from the experimental QIT to leave only the component with intensity I_1. Then, $I = I_1$ and the case reduces to that considered in [2], with

$$g_1^{(2)}(\tau) = \lim_{t\to\infty}\frac{\langle I_1(t)I_1(t+\tau)\rangle}{\langle I_1(t)\rangle^2} = \frac{p(\tau)}{p(\infty)} \quad (36)$$

and a two-photon correlator determined by formula (34). By the same token, when the fluorescence component with the intensity I_1 is removed and only the component with the intensity I_2 is left, we obtain the formula

$$g_2^{(2)}(\tau) = \lim_{t\to\infty}\frac{\langle I_2(t)I_2(t+\tau)\rangle}{\langle I_2(t)\rangle^2} = \frac{p'(\tau)}{p'(\infty)}, \quad (37)$$

where

$$p'(\tau) = \frac{\rho_1'(\tau)}{T_1}. \quad (38)$$

Calculating the probabilities $\rho_1(\tau)$ and $\rho_1'(\tau)$ using formulas (11), we obtain the following expressions:

$$\begin{aligned} p(\tau) &= \frac{k}{1+\gamma_1 T_1}\rho_0(\tau) \approx k\rho_0(\tau), \\ p'(\tau) &= \frac{k'}{1+\gamma_{1'} T_1}\rho_{0'}(\tau) \approx k'\rho_{0'}(\tau). \end{aligned} \quad (39)$$

Therefore, there are two-photon correlators of the two types and, hence, two absorption coefficients and two fluorescence processes, such that

$$p(\tau) \propto I_1(\tau), \quad p'(\tau) \propto I_2(\tau). \quad (40)$$

The probabilities $\rho_0(\tau)$ and $\rho_{0'}(\tau)$ in formulas (39) are determined from system of equations (12) with the ini-

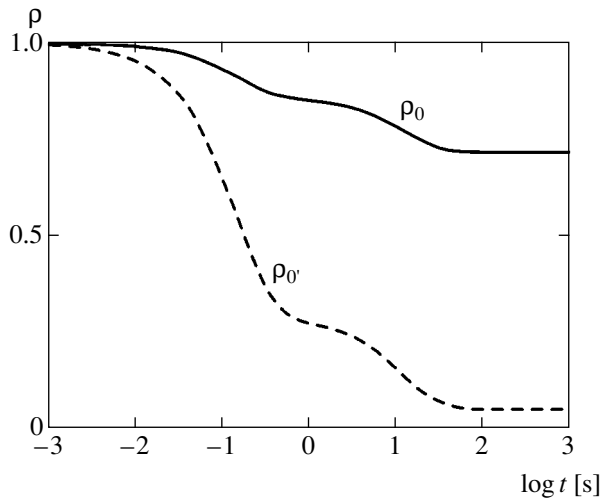


Fig. 11. Population of the ground electron state in two conformations of a polymer molecule calculated by Eqs. (39)–(44) with $B = b = 0.1\ s^{-1}$, $k_e = 1\ s^{-1}$, $\gamma_2 = 6\ s^{-1}$, $k'_e = 5\ s^{-1}$, and $\gamma_{2'} = 2\ s^{-1}$.

tial conditions $\rho_0(0) = 1$ and $\rho_{0'}(0) = 1$, respectively. The final expressions are as follows:

$$\rho_0(\tau) = -\frac{M_0(0)}{z_1z_2z_3} + \frac{M_0(z_1)e^{-z_1t}}{z_1(z_1-z_2)(z_1-z_3)} + \frac{M_0(z_2)e^{-z_2t}}{z_2(z_2-z_1)(z_2-z_3)} + \frac{M_0(z_3)e^{-z_3t}}{z_3(z_3-z_1)(z_3-z_2)}, \quad (41)$$

$$\rho_{0'}(\tau) = -\frac{M_{0'}(0)}{z_1z_2z_3} + \frac{M_{0'}(z_1)e^{-z_1t}}{z_1(z_1-z_2)(z_1-z_3)} + \frac{M_{0'}(z_2)e^{-z_2t}}{z_2(z_2-z_1)(z_2-z_3)} + \frac{M_{0'}(z_3)e^{-z_3t}}{z_3(z_3-z_1)(z_3-z_2)}, \quad (42)$$

where

$$M_0(z) = (z-B-\gamma_2)(z-b-\gamma_{2'})(z-k'_e) - \gamma_{2'}k'_e(z-B-\gamma_2) - bB(z-k'_e), \quad (43)$$

$$M_{0'}(z) = (z-B-\gamma_2)(z-b-\gamma_{2'})(z-k_e) - \gamma_2k_e(z-b-\gamma_{2'}) - bB(z-k_e), \quad (44)$$

and z_1, z_2, and z_3 are the roots of the equation

$$(z-B-b)(z-k_e-\gamma_2)(z-k'_e-\gamma_{2'}) - B\gamma_2(z-k'_e-\gamma_{2'}) - b\gamma_{2'}(z-k_e-\gamma_2) = 0, \quad (45)$$

which is obtained by equating the determinant of Eqs. (12) to zero. The fourth root of this equation ($z_0 = 0$), corresponding to the conservation of probabilities in Eqs. (12), is already taken into account in formulas (41) and (42).

In order to avoid the need of solving a cubic equation, let us consider the particular case of $k_e + \gamma_2 = k'_e + \gamma_{2'} = K$. Then, the above roots are as follows,

$$z_0 = 0, \quad z_2 = K,$$
$$z_{1,3} = \frac{K+B+b}{2} \pm \sqrt{\left(\frac{K-B-b}{2}\right)^2 + B\gamma_2 + b\gamma_{2'}}, \quad (46)$$

and their values increase in the sequence $z_0 < z_3 < z_2 < z_1$.

8. DISCUSSION OF RESULTS AND COMPARISON WITH EXPERIMENT

According to the energy diagram in Fig. 9, a single polymer molecule in Fig. 9 has two fluorescent conformations, the transition between which takes place in the triplet state. The model has been justified in Section 3, assuming that fluorescence from a single polymer molecule can measured in experiment. Experimental evidence for this was reported in [8]. The single-quantum character of the fluorescent system is also confirmed by the following considerations. If the experimentally measured fluorescence were from two polymer molecules, each in its own time-independent conformation and fluorescence fluctuating between zero and the corresponding intensity levels I_1 and I_2, respectively, the QIT of the system would also represent fluctuating fluorescence. However, the intensity of fluorescence in this system would jump between the four levels: 0, I_1, I_2, and $I_1 + I_2$, which contradicts the data presented in Fig. 4.

The proposed theory predicts the presence of the on and off periods of the two types. Therefore, the experimental distributions obtained for the periods of all the four types would allow the parameters k_e, k'_e, γ_2, and $\gamma_{2'}$ to be determined. Using the presently available experimental data (Fig. 5), it is possible only to estimate these quantities by order of magnitude.

The two fluorescence intensities are determined by formulas (39). The time fluorescence intensity kinetics is determined by the time variation of the probability of finding a molecule in the ground state of a given conformation. Using formulas (39)–(44) with the parameters determined from QITs in Fig. 3, we obtain the curves presented in Fig. 11. For these parameters, the roots calculated by formulas (46) are as follows:

$$z_0 = 0, \quad z_3 = 0.084\ s^{-1}, \quad z_2 = 7\ s^{-1},$$
$$z_1 = 7.116\ s^{-1}.$$

The corresponding relaxation times are

$$t_1 = 1/z_1 = 0.141\ s, \quad t_2 = 1/z_2 = 0.143\ s,$$
$$t_3 = 1/z_3 = 11.905\ s.$$

Let us consider the results presented in Fig. 11 in more detail. In a logarithmic time scale, a smooth step

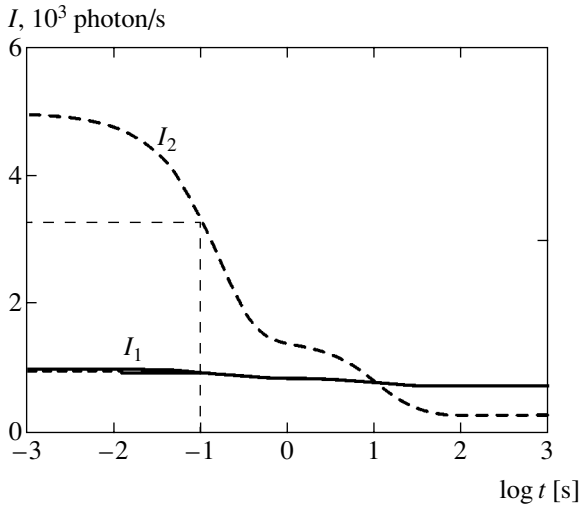


Fig. 12. Plots of the intermediate and high fluorescence intensities versus photon counting time calculated for the same parameters as in Fig. 11.

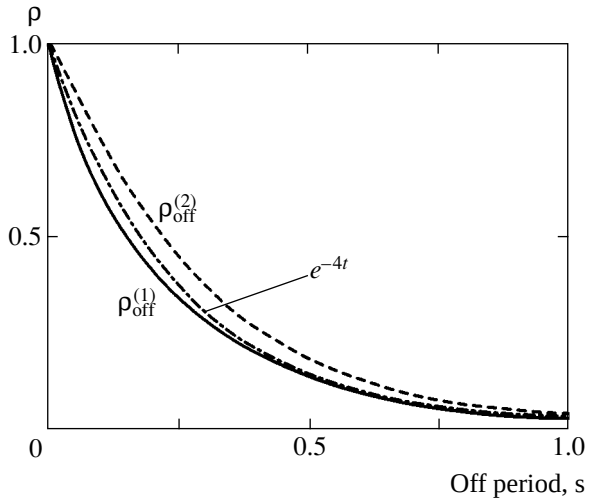


Fig. 13. Distributions of the off period duration for the off states of two types (calculated for the same parameters as in Fig. 11) in comparison to the exponential decay curve with a characteristic time of 250 ms.

(accounting for approximately one order of magnitude), describes exponential decay. There are two such steps in Fig.11, one of which (on the left) corresponds to an exponential probability for the molecule to jump to the triplet state. This transition corresponds to very close characteristic times t_1 and t_2. The right-hand step describes the exponential probability of the transition between conformations, with the corresponding time t_3.

The number of photons counted over time t is given by the formulas

$$N_1(t) = \int_0^t p(\tau)d\tau, \quad N_2(t) = \int_0^t p'(\tau)d\tau, \qquad (47)$$

while the fluorescence intensity measured in experiment is

$$I_1 = N_1(t)/t, \quad I_2 = N_2(t)/t. \qquad (48)$$

Figure 12 shows these intensities calculated via relations (48) as functions of the counting time.

For a sampling window of 100 ms used in [8], the fluorescence signal intensity fluctuated between two levels (I_1 and I_2) indicated in Fig. 12. The ratio of these intensities is about one third, which corresponds to 25/75 for the two nonzero peaks in Fig. 4.

The same parameters as in Figs. 11 and 12 can be used to calculate the distribution functions for the off period duration according to formulas (26), (30), and (31). The result of this calculation is presented in Fig. 13. As can be seen, the two distributions virtually coincide and are close to the exponential decay curve with a characteristic time of 250 ms.

9. CONCLUSION

Experimental data on fluctuating fluorescence reported in [8], processed in terms of the theory of light absorption by single molecules developed in [2, 21], allowed a theoretical model to be constructed that describes the electron excitation energy transfer in a single polymer molecule to an emission center of this fluorescent system and to determine the relaxation constants for this process.

According to the proposed model, the electron excitation energy in the polymer molecule is transferred from a singlet to triplet level, after which the molecule exhibits a conformational transition. Each conformation of the polymer molecule is characterized by a fluorescence intensity of its own, I_1 and I_2. This model does not allow determining which atoms in the polymer molecule change their equilibrium positions in the course of the conformational transition. Probably, these changes take place only in the fluorescent chromophore, rather than involving the entire molecule.

Since all quantum transitions possess a jumplike character, fluorescence of a single polymer molecule exhibits intermittency (fluctuates). These fluctuations in the fluorescence intensity studied as the function of time are represented by a quantum intensity transient.

The proposed theory explains the following experimental results [8]:

(i) the existence of intermediate and high fluorescence intensity levels in a time scale of seconds;

(ii) the alteration of dark (off) periods and two types of bright (on) fluorescence periods in this time scale;

(iii) the exponential distribution of the durations of on (τ_{on}) and off (τ_{off}) periods;

(iv) the independence of the average τ_{off} value of the exciting light intensity and the linear variation of the average τ_{on} value with the inverse intensity of excitation.

Quantitative agreement of the experimental data and calculated values is obtained for the following values of the relaxation constants of the polymer molecule studied: conformation variation rate, 0.1 s^{-1}; triplet state lifetime, 0.1–0.5 s; probability of the intercombination transition from excited singlet to triplet state, 10^2 s^{-1}.

The theory predicts the following new effects:

(i) existence of the bright (on) periods of two types with different levels of the fluorescence intensity—intermediate (I_1) and high (I_2)—and two types of the dark (off) periods, following the corresponding on periods with the intermediate (I_1) and high (I_2) intensities;

(ii) dependence of the I_1/I_2 ratio of the intermediate to bright fluorescence levels on the signal accumulation time (sampling window), which was equal to 100 ms in the experiments reported in [8].

Experimental verification of these predictions and confirmation of the important role of the triplet state in conformational transitions would provide for significant progress in our understanding of the process of energy transfer in polymer chains.

ACKNOWLEDGMENTS

This study was supported by the Russian Foundation for Basic Research, project no. 01-02-16580.

REFERENCES

1. Th. Basche, W. E. Moerner, M. Orrit, and U. P. Wild, *Single Molecule Optical Detection, Imaging, and Spectroscopy* (Wiley–VCH, Munich, 1997).
2. I. S. Osad'ko, *Selective Spectroscopy of Single Molecules* (Fizmatlit, Moscow, 2000; Springer, Berlin, 2002).
3. M. Nirmal, B. O. Dabbousi, M. G. Bawendi, *et al.*, Nature **383**, 802 (1996).
4. M. E. Pistol, P. Castrillo, D. Hessman, *et al.*, Phys. Rev. B **59**, 10725 (1999).
5. M. Sugisaki, Hong-Wen Ren, I. S. Osad'ko, *et al.*, Phys. Status Solidi B **224**, 67 (2001).
6. A. M. van Oijen, M. Ketelaars, J. Kohler, *et al.*, J. Phys. Chem. B **102**, 9363 (1998).
7. A. M. van Oijen, M. Ketelaars, J. Kohler, *et al.*, Science **285**, 400 (1999).
8. D. van der Bout, Wai-Tak Yip, D. Hu, *et al.*, Science **277**, 1074 (1996).
9. J. Yu, D. Hu, and P. Barbara, J. Am. Chem. Soc. **121**, 6936 (1999).
10. J. Yu, D. Hu, and P. Barbara, Science **289**, 1327 (2000).
11. J. White, J. Hsu, Shu-Chun Yang, *et al.*, J. Chem. Phys. **114**, 3848 (2001).
12. T. Gensch, J. Hofkens, A. Heirmann, *et al.*, Angew. Chem. Int. Ed. Engl. **38**, 3752 (1999).
13. J. Hofkens, M. Maus, Th. Gensch, *et al.*, J. Am. Chem. Soc. **122**, 9278 (2000).
14. T. Vosch, J. Hofkens, M. Cotlet, *et al.*, Angew. Chem. Int. Ed. Engl. **40**, 4643 (2001).
15. R. Vale, T. Funatsu, D. Plerce, *et al.*, Nature **380**, 451 (1996).
16. R. Dickson, A. Cubitt, R. Tsien, and W. E. Moerner, Nature **388**, 355 (1997).
17. H. Lu, L. Xun, and X. Xie, Science **282**, 1877 (1998).
18. N. Harrison, D. Baigent, I. Samuel, and R. Fried, Phys. Rev. B **53**, 15815 (1996).
19. T. Swager, C. J. Gil, and M. Wrighton, J. Phys. Chem. B **99**, 4886 (1995).
20. A. M. Boiron, Ph. Tamarat, B. Lounis, *et al.*, Chem. Phys. **247**, 119 (1999).
21. I. S. Osad'ko and E. V. Khots, Single Mol. **3**, 236 (2002).
22. J. Bernard, L. Fleury, H. Talon, and M. Orrit, J. Chem. Phys. **98**, 850 (1993).
23. I. S. Osad'ko, Zh. Éksp. Teor. Fiz. **113**, 1606 (1998) [JETP **86**, 875 (1998)].
24. I. S. Osad'ko, in *Spectroscopy and Excitation Dynamics of Condensed Molecular Systems*, Ed. by V. Agranovich and R. Hochstrasser (North-Holland, Amsterdam, 1983), p. 437.
25. I. S. Osad'ko and L. B. Yershova, J. Lumin. **86**, 211 (2000).
26. I. S. Osad'ko and L. B. Yershova, J. Lumin. **87–89**, 184 (2000).

Translated by P. Pozdeev

Journal of Experimental and Theoretical Physics, Vol. 96, No. 4, 2003, pp. 629–642.
Translated from Zhurnal Éksperimental'noĭ i Teoreticheskoĭ Fiziki, Vol. 123, No. 4, 2003, pp. 710–725.

ATOMS, SPECTRA, RADIATION

Spectroscopy of Coherent Dark Resonances in Multilevel Atoms for the Example of Samarium Vapor

Yu. V. Vladimirova[a], B. A. Grishanin[a,*], V. N. Zadkov[a,], N. N. Kolachevskiĭ[b], A. V. Akimov[b], N. A. Kisilev[b], and S. I. Kanorskiĭ[b]**

[a]*Department of Physics and the International Laser Center, Moscow State University, Moscow, 119899 Russia*
[b]*Lebedev Physical Institute, Russian Academy of Sciences, Moscow, 117924 Russia*
**e-mail: grishan@comsim1.phys.msu.su*
***e-mail: zadkov@comsim1.phys.msu.su*
Received August 30, 2002

Abstract—A universal theory for calculating coherent population trapping resonances in multilevel atoms is suggested. The theory allows arbitrary schemes of multilevel atoms and their excitations to be calculated taking into account the influence of relaxation effects in atoms, applied magnetic field, and the Doppler effect. The experimental data obtained by high-precision diode spectroscopy of coherent dark resonances in samarium vapor are systematically analyzed using the suggested theory. In the absence of a magnetic field, the model of samarium is based on consideration of a degenerate Λ system of the $4f^6 6s^2(^7F_0) \longleftrightarrow 4f^6(^7F)6s6p(^3P^0)^9F_1^0 \longleftrightarrow 4f^6 6s^2(^7F_1)$ active transitions. If the fourth $4f^6 6s^2(^7F_2)$ level is taken into account, this Λ system becomes open. Numerical simulation of coherent population trapping resonances shows that the open character of the system decreases the contrast of resonance curves in absorption spectra without changing resonance widths. The system under applied external longitudinal and transverse magnetic fields is correctly described by 7- and 12-level models of atomic transitions, respectively. © *2003 MAIK "Nauka/Interperiodica".*

1. INTRODUCTION

The interaction of an electromagnetic field with an atom is one of the most fundamental problems of quantum optics. Multilevel atoms are known to exhibit a broader spectrum of effects under these conditions than two-level atoms because of field-induced coherence between atomic states and quantum interference. Three-level systems in the Λ, Ξ, and V configurations play an important role in studying these effects, being intermediate in complexity between two-level and multilevel atoms. Quite a number of new effects are observed in three-level atoms, of which coherent population trapping is one of the most intriguing phenomena. This phenomenon has been extensively studied both theoretically and experimentally (see review [1] and the references therein). The coherent population trapping effect most strikingly manifests itself in three-level systems with two closely spaced long-lived levels and the third level distant from them (Λ or V systems), which are excited by two continuous laser fields in such a way that the distant level is optically "coupled" with two others. Tuning the exciting fields in resonance with dipole transitions results in system population trapping in the coherent superposition of two closely spaced levels. This effect manifests itself in Raman absorption spectra as a very narrow dip against the absorption line background and in resonance fluorescence spectra as the absence of emission, whence the term "dark (or coherent population trapping) resonance."

The coherent population trapping phenomenon is currently extensively used in various applications, such as magnetometry, metrology, etc. [2–6]. Since the first observation of a coherent population trapping resonance in sodium vapor [2], the majority of experimental studies of coherent population trapping resonances have been performed with alkali metal atoms [1, 7], whose hyperfine ground state components with characteristic splittings of several GHz were used as the lower Λ system levels. The long lifetimes of the coherent superposition of the lower alkali metal atom states allow high-contrast and high-Q coherent population trapping resonances to be recorded thanks to the availability of stable high-precision laser systems tunable in resonance transition regions and comparatively simple phase locking of exciting light fields. For instance, resonances about 10 kHz wide were recorded for pure cesium vapor [7]. A further decrease in the width of resonances can be achieved by introducing an inert buffer gas (Ne, He, or Ar) into the cell for measurements at pressures of several kPa. The residence time of atoms in light beams then increases without disturbing coherence of the superposition state of the lower levels, which are weakly dephased by collisions with buffer gas atoms. In particular, the narrowest resonance about

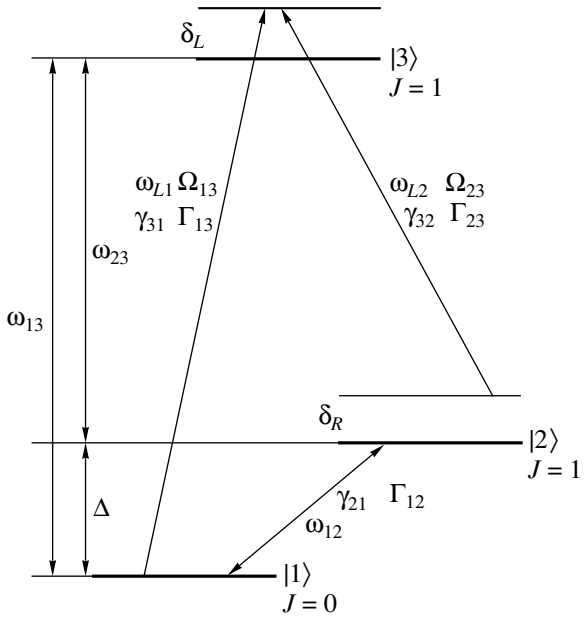


Fig. 1. Scheme of a three-level atom in the Λ configuration excited by two laser fields at frequencies ω_{L1} and ω_{L2}. Ω_{13} and Ω_{23} are the Rabi frequencies corresponding to pumping fields; δ_L is the resonance detuning at the $|1\rangle \longleftrightarrow |3\rangle$ transition; δ_R is the Raman detuning; γ_{31} and γ_{32} are the rates of radiative decay of excited states to the $|1\rangle$ and $|2\rangle$ levels, respectively; γ_{21} and w_{12} are the rates of decay and thermal pumping of level $|1\rangle$ through $|2\rangle$; and Γ_{13}, Γ_{23}, and Γ_{12} are the rates of dephasing of transitions $|1\rangle \longleftrightarrow |3\rangle$, $|2\rangle \longleftrightarrow |3\rangle$, and $|1\rangle \longleftrightarrow |2\rangle$, respectively.

50 Hz wide was obtained for cesium–neon combinations [7].

Coherent population trapping in rare-earth metal atoms has certain special features, because the characteristic distance between fine structure components used as the lower Λ system levels is substantially larger than the hyperfine splittings of alkali metal ground states and amounts to 10–100 THz. The characteristic spontaneous decay time of these levels is determined by magnetic dipole transitions and equals several seconds, which does not prevent the observation of supernarrow resonances. These levels are also weakly sensitive to atomic collisions, because they are well shielded by the outer closed shell. For this reason, rare-earth metal atoms also offer promise for use in metrological applications, for instance, for creating a secondary frequency standard (e.g., see [8]). The samarium atom is one of the most promising objects for metrological applications. Its scheme of levels is much simpler than that of cesium, especially in applied magnetic fields. Precisely for this reason, we selected it as a "touchstone" for testing the general theory of coherent population trapping in multilevel atoms developed by us.

Theoretically, the coherent population trapping phenomenon was studied in detail for the three-level model [1], which allows calculations to be performed analytically. The model, however, becomes much more complex for multilevel systems, and its analytic study turns impossible in the majority of cases. In this work, we suggest a general theoretical model for numerically analyzing the coherent population trapping spectra of atoms with an arbitrary number of levels and compare the results obtained for this model with the experimental data on samarium [9].

In Section 2, we describe the coherent population trapping effect in terms of the simplest three-level Λ system model. Section 3 contains a description of a general mathematical technique for calculating stationary states of active atoms and the corresponding level populations, absorption coefficients, and the dispersion of applied fields from the point of view of the spectroscopy of dark resonances. A method for taking into account the Doppler effect in calculating medium absorption is considered in Section 4. An experimental study of coherent population trapping in samarium vapor is described in Section 5. For samarium atoms, completely taking into account the Zeeman structure of lines involved in the formation of coherent population trapping resonances requires the use of a 12-level model. However, even a substantially simpler 4-level model gives close qualitative agreement with experiment. Section 6 contains a general description of this model and parameters necessary for performing calculations and comparing them with experimental results. The results obtained in calculating absorption in the absence of magnetic fields and under longitudinal and transverse applied fields are given in Section 7, where these results are compared with experimental data. The most important conclusions are formulated in Section 8. The special features of coherent population trapping resonances against the background of a line broadened by the Doppler effect in longitudinal and transverse magnetic fields are considered in the Appendix.

2. COHERENT POPULATION TRAPPING IN A Λ SYSTEM

In the simplest three-level system of atomic transitions in the Λ configuration, two lower long-lived levels $|1\rangle$ and $|2\rangle$ with frequency splitting Δ are coupled with the upper excited energy level $|3\rangle$ by two light fields (Fig. 1). If the $|1\rangle \longleftrightarrow |2\rangle$ transition is forbidden in the dipole approximation and two fields $E_1\exp(-i\omega_{L1}t - i\varphi_1)$ and $E_2\exp(-i\omega_{L2}t - i\varphi_2)$ are in resonance with the corresponding $|1\rangle \longleftrightarrow |3\rangle$ and $|2\rangle \longleftrightarrow |3\rangle$ transitions, a narrow coherent population trapping resonance is formed as a result of quantum interference. It manifests itself in absorption spectra by the appearance of a sharp maximum when one of the acting fields, for instance, ω_{L1}, is scanned and the Raman detuning $\delta_R = \omega_{L1} - \omega_{L2} - \Delta$ passes zero, which corresponds to the exact resonance.

To describe the nature of this physical process more visually, different basis sets are used to consider the

atomic system. In particular, the $|1\rangle$ and $|2\rangle$ ground states can conveniently be replaced by their symmetrical and antisymmetric combinations $|\pm\rangle$,

$$|+\rangle = \frac{\Omega_{R1}^*|1\rangle + \Omega_{R2}^*|2\rangle}{\Omega_{\text{eff}}}, \quad |-\rangle = \frac{\Omega_{R2}|1\rangle - \Omega_{R1}|2\rangle}{\Omega_{\text{eff}}},$$

where $\Omega_{Rk} = -d_{3k}E_k/\hbar$ ($k = 1, 2$) are the Rabi frequencies determined from the corresponding dipole moments $d_{3k} = -e\langle 3|r|k\rangle$,

$$\Omega_{\text{eff}} = \sqrt{|\Omega_{R1}|^2 + |\Omega_{R2}|^2}$$

is the effective Rabi frequency, and the phases of states $|1\rangle$ and $|2\rangle$ coincide with those of laser fields.

The matrix element of the electric dipole operator between the ground and excited states vanishes for the $|-\rangle$ state at zero Raman detuning,

$$\langle 3|V_{\text{dip}}|-\rangle \propto (1 - e^{i\delta_R t}) \xrightarrow{\delta_R \to 0} 0.$$

By far the larger part of atomic population is concentrated in this state named dark because of radiative decay. As a result, fluorescence is almost fully suppressed. This process of optical pumping into the coherent dark state is known as coherent population trapping. The coherent nature of population trapping manifests itself by a dependence of the dark state on laser field phases. It follows that acting field phase fluctuations can decrease or even destroy coherent population trapping, and it is necessary to stabilize the relative phase of laser fields. Other decoherence processes and Doppler broadening can also contribute to the destruction of coherent population trapping.

The experimentally observed line width is determined by the stability of detuning δ_R and phase difference $\Delta\varphi$ and also by Doppler broadening, time-of-flight broadening, Stark broadening (broadening caused by light and external fields), broadening in nonuniform magnetic fields, and impact broadening. The $\Delta\varphi$ phase difference can very accurately be stabilized in experiments with alkali metal atoms, for instance, by modulating lasers at a frequency corresponding to Δ. When two independent diode lasers are used in the free generation mode, we can expect that coherent population trapping resonances several MHz wide will be observable.

3. A MATHEMATICAL TECHNIQUE FOR CALCULATING COHERENT POPULATION TRAPPING IN MULTILEVEL SYSTEMS

A description of the dynamics of quantum systems in which relaxation processes occur requires modifying dynamic equations in comparison with their usual form given in traditional textbooks on quantum mechanics and only applicable to closed systems without relaxation. While the dynamics of closed systems is determined by an energy operator acting on wave functions, the dynamics of systems with relaxation can only be described by transformations of density matrix operators or of dynamic variables, that is, by superoperator transformations. The simplest transformations of this type also arise in systems without relaxation if these systems are described in terms of density matrices, in particular, by the Liouville quantum equation

$$\frac{\partial\hat{\rho}}{\partial t} = \mathcal{L}_0\hat{\rho} = -\frac{i}{\hbar}[\hat{H}, \hat{\rho}].$$

The role of a superoperator transformation is here played by the Liouvillian $\mathcal{L}_0$. To the $-i/\hbar$ imaginary factor, the Liouvillian is described by the commutator with Hamiltonian $\hat{H}$ applied to the density matrix $\hat{\rho}$.

In order to introduce the corresponding superoperators irrespective of the operators to be transformed, it suffices to introduce the substitution symbol $\odot$ for specifying the position into which the operator in question should be substituted. In the Schrödinger representation, this is the density matrix. Further, we can use the rules for handling symbolic expressions that follow from the general definitions of the algebra of linear operators [10], which are quite obvious. For instance,

$$[\hat{A}_2, [\hat{A}_1, \odot]] = \hat{A}_2(\hat{A}_1\odot - \odot\hat{A}_1) - (\hat{A}_1\odot - \odot\hat{A}_1)A_2$$
$$= \hat{A}_2\hat{A}_1\odot - \hat{A}_1 \odot \hat{A}_2 - \hat{A}_2 \odot \hat{A}_1 + \odot\hat{A}_1\hat{A}_2.$$

In the symbolic representation, the Liouvillian of a closed system has the form

$$\mathcal{L}_0 = -\frac{i}{\hbar}[\hat{H}, \odot]. \quad (1)$$

Like all linear operators, superoperators can be written in the form of the corresponding matrices after the introduction of a linear basis in the linear space of quantum operators. The use of this technique for symbolically representing superoperators is effective in calculating systems of arbitrary dimensions, especially in calculating multilevel systems. In particular, because of large problem dimensions, even merely writing down the matrices that describe the evolution superoperators becomes a technically complex task. However, if the symbolic representation of superoperators is used, these matrices can first be written in the symbolic form thanks to its physical transparency, and matrix elements can then be calculated either analytically or numerically (for large-dimensionality matrices) on a computer. Technical difficulties of reproducing them are then fully transferred to automatic computer calculations, and the results of such calculations can easily be used in numerical calculations of applied problems under consideration with programs written in the most suitable programming language. We used a combination of the MATHEMATICA computer algebra package (for analytically setting superoperators) and the Fortran language (for subsequent numerical calculations of spectra with the use of the calculated dynamic superoperator matrices).

3.1. Calculations of the Liouvillian of an N-Level Atom in the Symbolic Representation

As for the two-level system, the Liouvillian of an *N*-level atom in the rotating field approximation can be written as the sum of contributions

$$\mathcal{L}_t = \mathcal{L}_r + \mathcal{L}_e + \mathcal{L}_\delta + \mathcal{L}_i, \tag{2}$$

where $\mathcal{L}_r$ is the radiative damping superoperator, $\mathcal{L}_e$ is the elastic dephasing superoperator, $\mathcal{L}_i$ is the superoperator of the interaction with the laser field, and $\mathcal{L}_\delta$ is the laser detuning superoperator, which augments the selected unperturbed evolution operator to the superoperator of free atomic dynamics in zero laser field. It includes the corresponding detunings of all acting laser fields and takes into account that free precession at the frequencies of these fields is included into the unperturbed dynamics superoperator.

Radiative damping is described by the Liouvillian that combines the population transfer superoperator given by the $\hat{P}_{lk} \odot \hat{P}_{kl}$ projector and the polarization damping superoperator given by the $[\hat{P}_{kk}, \odot]_+$ anticommutator,

$$\mathcal{L}_r = \sum_{kl} \gamma_{kl} \left(\hat{P}_{lk} \odot \hat{P}_{kl} - \frac{1}{2} [\hat{P}_{kk}, \odot]_+ \right), \tag{3}$$

where the two-dimensional array γ_{kl} describes the rates of spontaneous decay (for $k > l$) and pumping (for $k < l$).

Elastic dephasing is introduced by the $\mathcal{L}_e$ superoperator written in terms of the squares of commutators and determined by the particular model of dephasing. In order to specify it, consider two types of dephasing. First, we can only take into account internal dephasing in the system of two electronic states k and $l > k$. In conformity with the microscopic nature of elastic dephasing caused by weak collisions [11, 12] (random transition frequency fluctuations), this dephasing is described by the corresponding random frequency shift superoperator $-(i/2)\xi(t)[(\hat{P}_{kk} - \hat{P}_{ll}), \odot]$, where $\xi(t)$ is the fluctuation transition frequency shift. The resulting relaxation superoperator averaged over random phase fluctuations has the form

$$\mathcal{L} = -\frac{\Gamma_{in}^{kl}}{4} [\hat{n}_{kl}, \odot]^2,$$

where $\hat{n}_{kl} = \hat{P}_{ll} - \hat{P}_{kk}$ is the population inversion operator for the kl subsystem and Γ_{in}^{kl} is the corresponding dephasing rate. This type of pure dephasing is not only related to the dephasing of the kl transition itself but also contributes to the dephasing of all the transitions adjacent to it. It is nevertheless convenient to consider the dephasing only of the kl transition with the use of the representation

$$[\hat{n}_{kl}, \odot]^2 = 2(\hat{P}_{kk} \odot \hat{P}_{ll} + \hat{P}_{ll} \odot \hat{P}_{kk}) + [\hat{I}_{kl}, \odot]^2,$$

where $\hat{I}_{kl} = \hat{P}_{ll} + \hat{P}_{kk}$ is the operator of the total population of the kl subsystem. The first term,

$$\mathcal{L}_{in}^{kl} = -\Gamma_{in}^{kl} (\hat{P}_{kk} \odot \hat{P}_{ll} + \hat{P}_{ll} \odot \hat{P}_{kk}), \tag{4}$$

describes purely internal dephasing and does not influence adjacent transitions. Using all Γ_{in}^{kl} independent parameters, we can write the dephasing of all transitions by (4) alone. For simplicity of describing the physical nature of dephasing, it is, however, convenient to introduce another contribution. This is equal dephasings of arbitrary other levels through the kth and lth levels in the absence of action on the kl transition itself, that is, "external dephasing,"

$$\mathcal{L}_{ex}^{kl} = -\Gamma_{ex}^{kl} [\hat{I}_{kl}, \odot]^2, \tag{5}$$

where Γ_{ex}^{kl} is the corresponding dephasing rate. Accordingly, the complete elastic dephasing superoperator is given by the sum

$$\mathcal{L}_e = \sum_{k<l} (\mathcal{L}_{in}^{kl} + \mathcal{L}_{ex}^{kl}). \tag{6}$$

The laser detuning superoperator depends on the type of resonance under consideration and can usually be written in the form of an antisymmetric superoperator given by the commutator with population operators,

$$\mathcal{L}_\delta = i \sum_k \delta_k (\hat{P}_{kk} \odot - \odot \hat{P}_{kk}), \tag{7}$$

where δ_k is the array of frequency detunings.

The interaction with the laser field can be described by the antisymmetric commutator with the polarization operators

$$\mathcal{L}_i = -\frac{i}{2} \sum_{k<l} \Omega_{kl} [(\hat{P}_{kl} + \hat{P}_{lk}), \odot], \tag{8}$$

where Ω_{kl} is the two-dimensional array of the Rabi frequencies of the kl transitions.

After the introduction of the symbolic representation of the complete evolution operator [Eq. (2)] and its components (3) and (6)–(8), we can calculate the $N^2 \times N^2$ matrix representations of the L_t, L_r, L_e, L_δ, and L_i values by the formula

$$L_{mn} = (\hat{e}_m, \mathcal{L}\hat{e}_n).$$

Here, $\{\hat{e}_k\}$ is the orthonormalized basis and parentheses denote the scalar multiplication of two operators of

the form $\mathrm{Tr}(\hat{A}^{+}\hat{B})$, which is antilinear in the first and linear in the second multiplier.

The $\{\hat{e}_k\}$ basis can conveniently be selected as Hermitian and expressed via the $\hat{P}_{kl}$ transition operators represented by $N \times N$ matrices, each with a single nonzero kl element $P_{kl}(k, l) = 1$. It is also convenient to assume that the levels are numbered in order of increasing energy, $E_1 \le E_2 \le \ldots \le E_N$. The corresponding basis is then constructed as follows:

$$\hat{e}_{j(k,l)} = \begin{cases} \hat{P}_{kk}, & k = l, \\ \dfrac{\hat{P}_{kl} + \hat{P}_{lk}}{\sqrt{2}}, & k < l, \\ -i\dfrac{\hat{P}_{kl} - \hat{P}_{lk}}{\sqrt{2}}, & k > l, \end{cases} \tag{9}$$

where $j(k, l)$ is the numbering index, that is, a one-to-one mapping of the two-dimensional set of numbers kl ($k, l = 1, N$) onto the one-dimensional set $j = 1, N^2$. This index can, in particular, be specified in the following way universal for any N:

$$j = \begin{cases} k, & k = l, \\ (2k-1)N - (k+1)^2 + k + 2l, & k < l, \\ (2l-1)N - (l+1)^2 + l + 2k + 1, & k > l. \end{cases}$$

For $N = 2$, 3, and 4, this corresponds to the following $j_N = (j_N(k, l))$ matrices:

$$j_2 = \begin{pmatrix} 1 & 3 \\ 4 & 2 \end{pmatrix}, \quad j_3 = \begin{pmatrix} 1 & 4 & 6 \\ 5 & 2 & 8 \\ 7 & 9 & 3 \end{pmatrix},$$

$$j_4 = \begin{pmatrix} 1 & 5 & 7 & 9 \\ 6 & 2 & 11 & 13 \\ 8 & 12 & 3 & 15 \\ 10 & 14 & 16 & 4 \end{pmatrix}.$$

Basis (9) is Hermitian and orthonormalized with respect to the $(\hat{A}, \hat{B})$ scalar product described above, and $(\hat{e}_m, \hat{e}_n) = \delta_{mn}$ for all $m, n = 1, N^2$.

3.2. Calculations of Coherent Population Trapping in an N-Level Atom

The technique described above can effectively be used in analytic calculations for solving the stationary state problem and the complete spectral problem for the $\mathcal{L}_t$ evolution superoperator [13, 14].

The most important properties of coherent population trapping are determined by the absorption of applied field. For the Λ resonance, absorption is described by the equation

$$\begin{aligned} W_L &= \hbar \mathrm{Im}(\omega_L g \langle \hat{\sigma}_{13}^{+} \rangle + \omega_L' g' \langle \hat{\sigma}_{23}^{+} \rangle) \\ &\approx \hbar(\omega_L \gamma + \omega_L' \gamma') n_3. \end{aligned} \tag{10}$$

Here, $\langle \hat{\sigma}_{13}^{+} \rangle$ and $\langle \hat{\sigma}_{23}^{+} \rangle$ are averaged positive-frequency operators of the complex amplitudes of the 1–3 and 2–3 transitions, respectively; ω_L and ω_L' are the frequencies of biharmonic pumping fields; g and g' are the corresponding Rabi frequencies; and γ and γ' are the corresponding radiative damping rates. The determination of the $\langle \hat{\sigma}_{13}^{+} \rangle$, $\langle \hat{\sigma}_{23}^{+} \rangle$, or n_3 stationary mean values requires calculating the corresponding vector representation $\langle 0|$ of the $\hat{\rho}_{st}$ stationary density matrix by solving the $\langle 0| L_t = 0$ equation.

In the basis under consideration with $\hat{e}_1 = \hat{P}_{11}$, $\hat{e}_2 = \hat{P}_{22}$, and $\hat{e}_3 = \hat{P}_{33}$, the first three vector $\langle 0|$ elements describe populations and should be normalized accordingly. The $\langle 0|$ bra vector of the stationary density matrix should therefore be used in the normalized form

$$\langle 0| \longrightarrow \frac{\langle 0|}{\langle 0|_1 + \langle 0|_2 + \langle 0|_3},$$

which automatically gives correct signs of the values to be calculated. The mean populations then coincide with the corresponding components, $\langle \hat{n}_k \rangle = \langle 0|_k$, $k = 1, 3$, and the complex transition amplitudes are expressed through the components with $k > 3$,

$$\langle \hat{\sigma}_{13}^{+} \rangle = \frac{\langle 0|_6 + i\langle 0|_7}{\sqrt{2}}, \quad \langle \hat{\sigma}_{23}^{+} \rangle = \frac{\langle 0|_8 + i\langle 0|_9}{\sqrt{2}},$$

$$\langle \hat{\sigma}_{12}^{+} \rangle = \frac{\langle 0|_4 + i\langle 0|_5}{\sqrt{2}}.$$

These equations allow absorption to be written in an analytic form convenient for both numerical calculations and a qualitative analysis.

We wrote a universal Fortran program for calculating level populations, absorption coefficients, and dispersion in an arbitrary N-level system. The program can be used at large $N > 10$ values. Its important feature is the use of a minimum necessary number of input parameters, which is substantially smaller than the $N^2 \times N^2$ number of Liouvillian matrix elements in the Liouvillian generalization under consideration [Eq. (1)], because it is not necessary to write down all L_t dynamic matrix elements by hand. According to (2)–(8), this matrix in reality contains a huge number of zero contributions.

4. TAKING INTO ACCOUNT THE DOPPLER EFFECT

The technique of calculations described in Section 3 can be used to obtain the dependences of the laser radiation absorption coefficient for an atom at rest on the first field δ_L and Raman δ_R detunings. In experiments, moving atoms interact with fields. For this reason, the Doppler effect influences the absorption coefficient of the medium. In the absence of simplifications, this requires performing calculations for a continuum of detunings simultaneously. In this work, the Doppler effect was taken into account by the following simplified method, which qualitatively corresponds to the approach taken in [15], but without approximately replacing the Maxwell velocity distribution by the Lorentz distribution.

The frequency of the laser field that interacts with an atom that moves in an arbitrary direction is given by the formula $\omega_{1,2} = \omega_{Lj} + \delta_{Lj}$, which takes into account the first-order correction for the Doppler effect. Here, $\delta_{Lj} = \omega_{Lj} v_n/c$, $j = 1, 2$, are the detunings of the biharmonic laser field components and v_n is the projection of the velocity of the moving atom onto the vector **n** of laser beam propagation.

The number of gas particles that move at velocity v_k at temperature T is determined by the Maxwell velocity distribution [11]

$$\frac{dN}{d\delta_L} = \frac{N}{\sqrt{\pi}}\exp\left[-\left(\frac{\delta_L}{\Delta\omega_D}\right)^2\right]\frac{d\delta_L}{\Delta\omega_D}, \quad \Delta\omega_D = \frac{\omega_{L1}v_0}{c},$$

where

$$\delta_L = \omega - \omega_{L1}, \quad v_0 = \sqrt{\frac{2kT}{m}}.$$

The $\delta_R = \omega_{L2} - \omega_{L1} - \Delta$ Raman detuning for laser beams propagating in one direction is then considered approximately constant for particles moving at different velocities. The dependence of the absorption coefficient for field ω_{L1} on the Raman detuning when field ω_{L2} is scanned therefore has the form

$$K_D(\delta_R) = \int K(\delta_L, \delta_R)\frac{dN}{d\delta_L}d\delta_L, \tag{11}$$

where $K(\delta_L, \delta_R)$ is the absorption coefficient of the atom at rest.

5. EXPERIMENTAL OBSERVATION OF COHERENT POPULATION TRAPPING RESONANCES IN SAMARIUM VAPOR

As mentioned in Section 1, samarium is a promising candidate for studying the feasibility of using coherent population trapping resonances in rare-earth metal vapors for metrological application purposes by high-resolution nonlinear spectroscopy methods. A diagram of the energy levels of the samarium atom is shown in Fig. 2. We experimentally studied samarium vapor absorption in the region of the

$$4f^6 6s^2({}^7F_0) \longleftrightarrow 4f^6({}^7F)6s6p({}^3P^0)\,{}^9F_1^0 \longleftrightarrow 4f^6 6s^2({}^7F_1) \tag{12}$$

transition lines, which formed a Λ system.

The experimental unit is schematically drawn in Fig. 3. The radiation sources were two semiconductor lasers (*1*, *2*) with external cavity resonators tuned to resonance wavelengths 672 and 686 nm. The lasers were assembled according to the Littrov scheme with a collimating aspherical objective and a holographic grating of 1800 lines/mm. The free detuning range of lasers was around 5 GHz. The lasers, with radiation wavelengths 672 and 686 nm, radiated 2.5 and 12 mW, respectively, in the single-frequency mode.

The spectra of the transitions in samarium that we were interested in were studied in detail in [16] by sub-

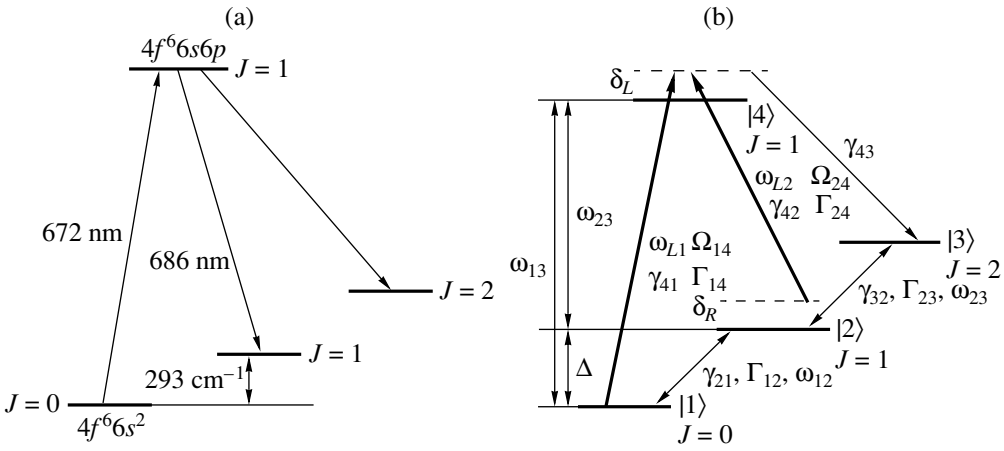


Fig. 2. (a) Diagram of energy levels of the samarium atom and (b) parameters used in calculations. The notation is the same as in Fig. 1 (for the corresponding level numbers).

Doppler absorption saturation spectroscopy. In [16], the relative isotopic shifts and hyperfine splittings of the levels were determined accurate to 1–2 MHz. The ^{154}Sm isotope (abundance 22.75%) had lines shifted with respect to the spectral lines of the other isotopes [^{144}Sm (3.07%), ^{147}Sm (14.99%), ^{148}Sm (11.24%), ^{149}Sm (13.82%), ^{150}Sm (7.38%), and ^{152}Sm (26.75%)] by 1 GHz to the red, which allowed reliable frequency locking to the transitions in this isotope to be performed. The presence of the other isotopes, however, slightly changed the wings of the working transition line.

Samarium vapor was generated in stainless steel cell *7*, which was 50 cm long and had glass windows at the ends. The cell was connected to a vacuum line and a system for buffer gas puffing. The cell was heated in its central part (15 cm long) with coaxial direct-current heater *8* (~500 W). The remanent magnetic field in the cell was of fractions of an oersted unit. The cell was placed within two pairs of Helmholtz rings *9* 30 cm in diameter, which could be used to create longitudinal and transverse magnetic fields up to 40 Oe in the central part of the cell. To obtain noticeable absorption, the cell was heated to about 1000 K [9]. Vapor concentration at this temperature was about 10^{11}–10^{12} cm^{-3}.

The 672 nm laser was tuned to the center of the $4f^6 6s^2(^7F_0) \longleftrightarrow 4f^6 6s6p(^9F_1^0)$ transition in ^{154}Sm and locked to the transmission peak of stabilized one-meter confocal interferometer *12* with a high long-term stability (about 5 MHz/h). The width of the laser generation spectrum was less than 0.5 MHz. The 686 nm laser was slowly retuned in the region of the $4f^6 6s^2(^7F_1) \longleftrightarrow 4f^6 6s6p(^9F_1^0)$ transition in such a way that its frequency passed the $\delta_R = 0$ point. Laser generation frequency variations were controlled by 0.5-m confocal interferometer *13* with a *Q*-factor of about 20 and a free dispersion region of 149.8 + 0.1 MHz. The mode composition of laser radiation was controlled using spectrum analyzer *14* with a *Q*-factor of 50 and a free dispersion region of 8 GHz. All interferometers were optically isolated from lasers to prevent the arising of feedback. Linearly polarized laser radiation converged into one beam (accurate to 10^{-3} rad) on polarization cube *6* and was launched into the cell with samarium vapor. The polarization planes of the beams were mutually orthogonal. At the entrance to the cell, radiation power density was 0.1 mW/mm^2 for the 672 nm laser and 0.2 mW/mm^2 for the 686 nm laser. After exit from the cell, the beams were divided using holographic diffraction grating *15* (2400 lines/mm) and directed to a system for recording (components *16*, *20*, and *21*).

As coherent population trapping is related to atomic system interactions with a dichromatic light field, we only recorded changes in the adsorption of laser radiation at 672 nm caused by the presence of the second light field. For this purpose, the beam from the 686 nm laser was modulated by liquid crystalline modulator *4* at a frequency of f_m = 600 Hz before it entered the cell, and the signal with the same modulation frequency was recorded in the channel of the 672 nm laser. The presence of broad excess absorption wings was caused by collisions with buffer gas atoms [17]. The spectra of induced absorption were recorded at both zero magnetic field and under applied longitudinal and transverse magnetic fields. In experiments with transverse applied fields, their direction was aligned with the direction of 672 nm laser polarization.

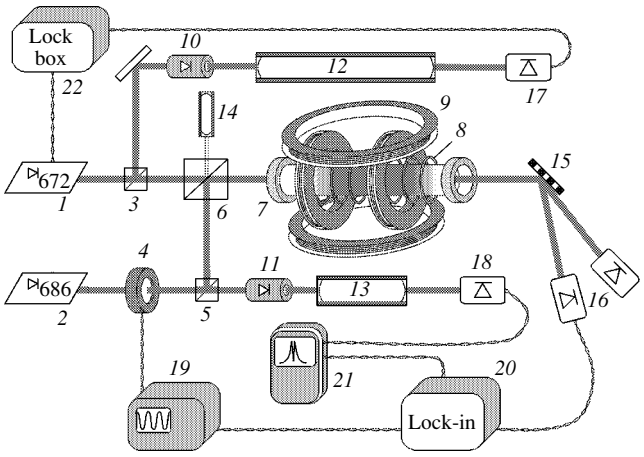


Fig. 3. Experimental unit for measuring dark resonance spectra in samarium vapor; *1* and *2* are semiconducting lasers with 672 and 686 nm wavelengths, respectively; *3* and *5* are light-splitting cubes; *4* is a modulator (600 Hz frequency); *6* is a polarizing cube; *7* is a cell with samarium vapor; *8* is a coaxial heater; rings *9* are Helmholtz rings; *10* and *11* are optical insulators; *12* is a confocal interferometer with the region of free dispersion of 74.35 ± 0.01 MHz; *13* is a confocal interferometer with the region of free dispersion of 149.8 ± 0.01 MHz; *14* is a spectroanalyzer; *15* is a 2400 lines/mm diffraction grating; *16–18* are photodiodes; *19* is a generator; *20* is a synchronous detector; *21* is an oscilloscope with memory; and *22* is the electronic block for locking the 672 nm laser frequency to the transmission peak of interferometer *12*.

6. COHERENT POPULATION TRAPPING IN THE FOUR-LEVEL MODEL

In the experiment under consideration, apart from the active samarium vapor levels that formed the Λ system, the formation of coherent population trapping resonances involved the $4f^6 6s^2$ ($J = 2$) level. Although this level did not directly participate in the excitation of the upper level, it absorbed part of the population as a consequence of radiative decay (Fig. 2a). In addition, the $J = 2$ level was populated by incoherent pumping from the lower levels that formed the Λ system. It follows that the $J = 2$ level played the role of a reservoir for coherent population trapping in the Λ system under consideration and its presence made the Λ system an open system. In the absence of magnetic field, this four-

Table 1. Wavelengths and oscillator strengths of active transitions

Transition	Wavelength λ, nm	Oscillator strength *gf*
$6s^2$ ($J = 0$) ⟶ $6s6p$ ($J = 1$)	672.5875	8.5×10^{-3}
$6s^2$ ($J = 1$) ⟶ $6s6p$ ($J = 1$)	686.0927	9.5×10^{-3}

level model takes into account the main mechanisms that determine coherent population trapping effects.

The characteristics of the samarium atom and the parameters of experiments necessary for a comparison with experimental data are summarized in Tables 1, 2, and 3. Table 1 contains the oscillator strengths of the transitions of interest, and Table 2, the energies and *g*-factors of the lower metastable levels with $J = 0$, 1, and 2 and the upper level of the Λ system. The relative populations of the metastable levels at $T = 600$°C are also given (see [16]).

Dipole moment *d*, Rabi frequencies Ω, and decay rates γ are calculated by the equations

$$|d_{JJ'}|^2 = \frac{3\hbar e^2}{2m}\frac{(2J+1)|f_{JJ'}|}{\omega_{JJ'}},$$

$$\Omega_{JJ'} = \frac{d_{JJ'}E}{\hbar}, \quad \gamma_{JJ'} = \frac{4d_{JJ'}^2\omega^3}{3\hbar c^3},$$

where *m* and *e* are the mass and the charge of the electron, respectively; *c* is the velocity of light; $\omega_{JJ'}$ is the transition frequency; and $|f_{JJ'}|$ is the oscillator strength of the $J \longrightarrow J'$ transition. The electric fields are calculated by the formula $E = \sqrt{2W/c\varepsilon_0}$; they take on the values $E_{L1} \approx 270$ V/m and $E_{L2} \approx 390$ V/m at laser radiation power densities at the entrance to the cell of $W_{L1} = 0.1$ mW/mm^2 and $W_{L2} = 0.2$ mW/mm^2, respectively [9]. The calculation results are summarized in Table 3.

The longitudinal and transverse magnetic field intensities were 15 and 29 Oe, respectively. The Zeeman splittings in the longitudinal magnetic field calculated as $\Delta = egH/2mc$ were $\Delta' = 1.98 \times 10^8$ s^{-1} for the $6s6p$ level and $\Delta'' = 4.09 \times 10^8$ s^{-1} for the $6s^2$ level. The transverse magnetic field splittings were $\Delta' = 2.50 \times 10^8$ s^{-1} for the $6s6p$ level and $\Delta'' = 5.17 \times 10^8$ s^{-1} for the $6s^2$ level.

7. THE RESULTS OF SIMULATING THE COHERENT POPULATION TRAPPING SPECTRA IN SAMARIUM VAPOR

Calculations based on the technique described in Section 3 gave the absorption coefficients of a samarium atom at rest for the three- and four-level models (Figs. 4a and 4b, respectively). An analysis of the dependences plotted in these figures shows that the introduction of the fourth $J = 2$ level into the three-level model has virtually no effect on the width of the resonance, whereas the complete absorption value for the four-level system is much smaller than for the three-level one. The reason for this is population trapping at the $J = 2$ level through the corresponding radiative decay channel.

7.1. The Modification of the Spectra in a Magnetic Field

Applying magnetic field transforms the three-level system of the samarium atom considered above into a seven-level one because of the splitting of the $J = 1$ levels. The $|3\rangle$ level splits into three components (Fig. 5), which results in the existence of three transitions to the $|1\rangle$ level allowed by the selection rules for radiative transitions. The probability of each of these transitions equals one-third of the total probability of the $|3\rangle \longleftrightarrow |1\rangle$ transition. Similarly, the $|2\rangle$ level also splits into three components, and, according to the selection rules, the $|3\rangle \longleftrightarrow |2\rangle$ transition transforms into six transitions, the probability of each of them being one-sixth of the total probability of the $|3\rangle \longleftrightarrow |2\rangle$ transition.

An additional decay channel in multilevel systems compared with three-level ones is collisional depolarization [18]. The depolarization of an atom caused by a collision with another atom is related to transitions between states with different magnetic moment projections onto the selected direction. When a magnetic field is applied, collisions cause transitions between Zeeman sublevels with different magnetic momentum projections for each multiplet, $|m\rangle \rightleftharpoons |m \pm 1\rangle$. Collisions

Table 2. Samarium energy levels determining absorption spectrum

Even levels $4f^6 6s^2(^7F)$				Odd level $4f^6(^7F)6s6p(^3P^0)^9F_1^0$		
J	Energy, cm^{-1}	*g*	Relative population at $T = 600$°C	*J*	Energy, cm^{-1}	*g*
0	0	–	1.0			
1	292.58	1.50	0.6	1	14863.85	3.10
2	811.92	1.50	0.24			

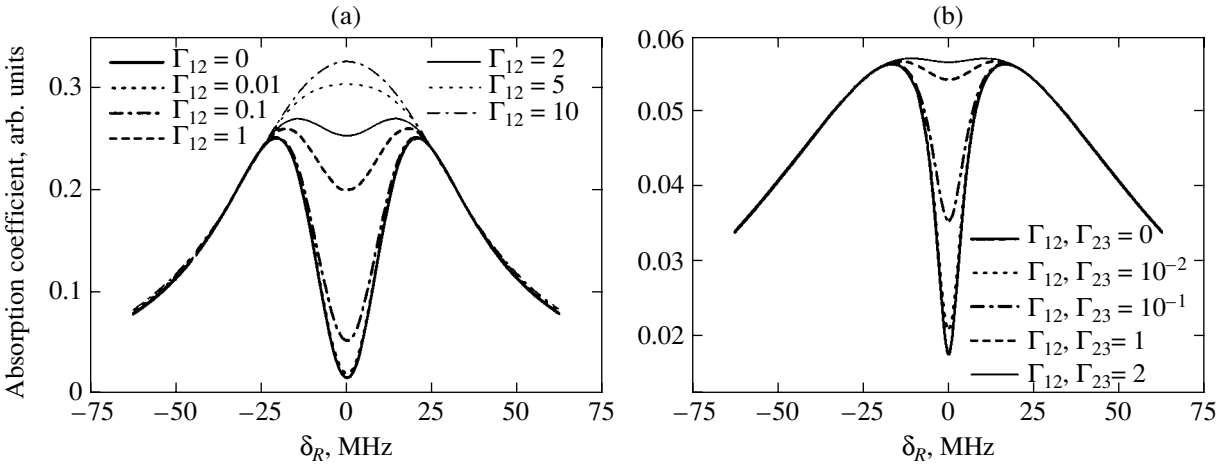


Fig. 4. Probing field absorption coefficients in (a) three-level and (b) four-level systems as functions of Raman detuning δ_R at $\delta_L = 0$ and different dephasing rates (s^{-1}). The corresponding schemes of levels are given in Figs. 1 and 2.

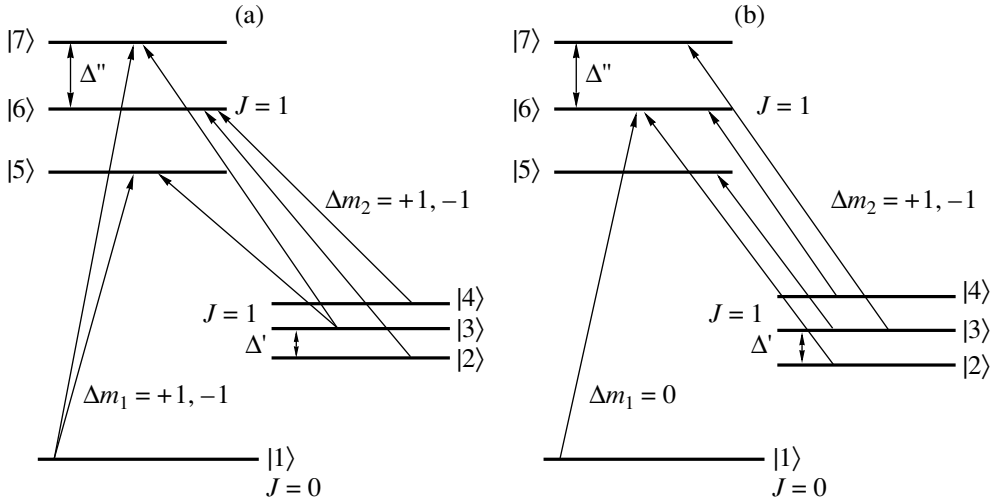


Fig. 5. Scheme of the Λ system of a samarium atom under applied (a) longitudinal magnetic field for linear orthogonal laser beam polarizations and (b) transverse magnetic field; the selection rules for the first field are $\omega_{L1} - \Delta m_1 = \pm 1$ (a) and 0 (b); for the second field, $\omega_{L2} - \Delta m_2 = \pm 1$; Δ' and Δ" are the Zeeman splittings of the lower and upper levels with $J \neq 0$, respectively.

with changes in momentum projections destroy the coherence of the lower Λ system levels, which influences the coherent population trapping resonance value. We used the numerical data given in Section 6 to perform theoretical calculations. In the calculations, this process was taken into account by introducing depolarization constant G between the levels of each multiplet. This constant was used as an adjustable parameter and was varied in the range $G = 0$–$80\gamma_{41}$.

Below, we will distinguish between two applied magnetic field configurations, namely, longitudinal and transverse.

7.2. Longitudinal Magnetic Field

A scheme of levels for the longitudinal field configuration is shown in Fig. 5a. According to the selection rules, six transitions are allowed for linearly polarized fields in the system under consideration, because $E_1 \perp H$ ($\Delta m_1 = \pm 1$) and $E_2 \perp H$ ($\Delta m_2 = \pm 1$). The $|1\rangle \longleftrightarrow |5\rangle$, $|3\rangle \longleftrightarrow |5\rangle$ and $|1\rangle \longleftrightarrow |7\rangle$, $|3\rangle \longleftrightarrow |7\rangle$ transitions form two Λ systems, whereas the $|2\rangle \longleftrightarrow |6\rangle$ and $|4\rangle \longleftrightarrow |6\rangle$ transitions, which are also allowed by the selection

Table 3. Parameters determining Λ system excitation

Rabi frequency, s^{-1}	Radiative decay rate, s^{-1}	Dephasing rate, s^{-1}
$\Omega_{14} = 0.58 \times 10^7$	$\gamma_{41} = 0.42 \times 10^6$	$\Gamma_{12} = 2.4 \times 10^4$
$\Omega_{24} = 0.83 \times 10^7$	$\gamma_{42} = 0.45 \times 10^6$	$\Gamma_{23} = 1.6 \times 10^4$
–	$\gamma_{43} = 0.42 \times 10^6$	–

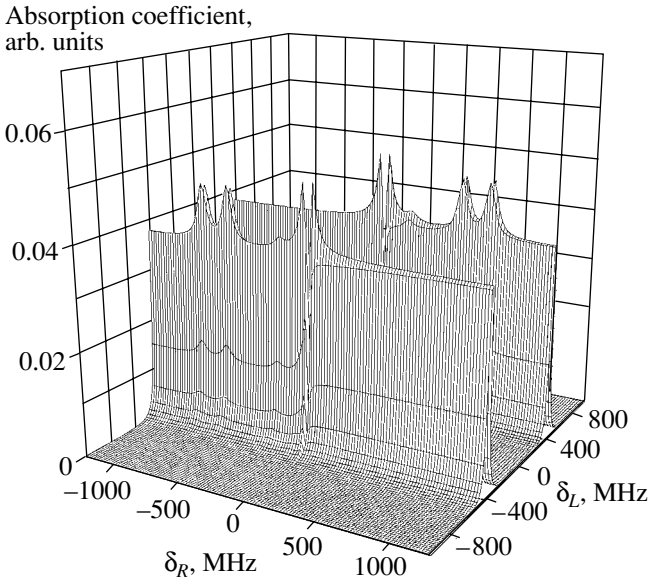


Fig. 6. Dependence of the absorption coefficient of the samarium atom at rest on δ_R and δ_L in applied longitudinal magnetic field.

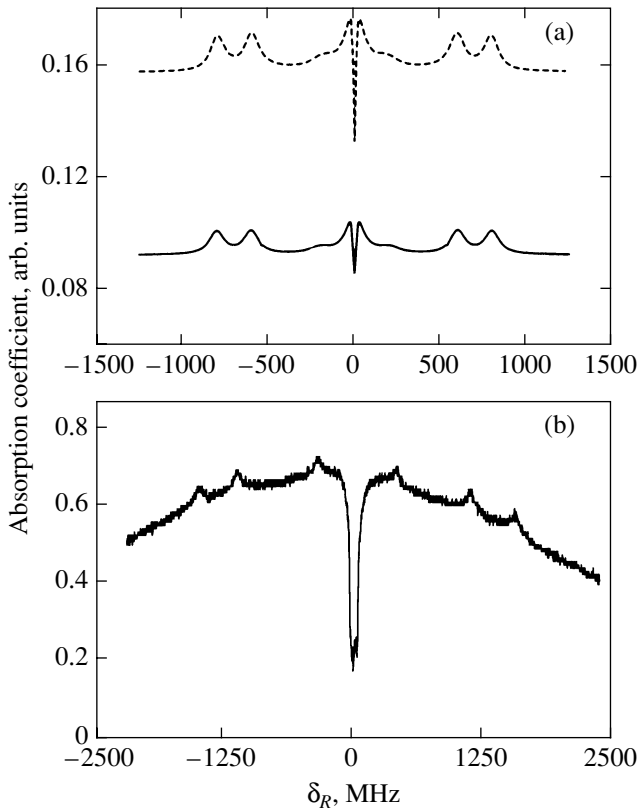


Fig. 7. (a) Theoretical dependences of the absorption coefficient of a seven-level system on Raman detuning δ_R in a longitudinal magnetic field calculated taking into account the Doppler effect at two temperatures T = 10 K (dashed line) and T = 873 K (solid line) and (b) experimental dependence of absorption coefficient in a 29 Oe longitudinal magnetic field at a 0.2 Torr buffer gas (Ar) pressure.

rules, do not participate in Λ system formation but are responsible for the formation of additional absorption peaks (see Appendix).

The dependence of the absorption coefficient of the samarium atom at rest on the Raman δ_R and laser δ_L detunings for the configuration under consideration is shown in Fig. 6; this dependence does not take into account depolarization. Including depolarization results, first, in an increase in induced absorption and, second, in a monotonic decrease in the contrast of coherent population trapping resonances. We observe hardly any depolarization effects on the width of coherent population trapping resonances.

We used (11) and the calculated absorption of the atom at rest to determine the absorption coefficient of the medium. The dependences of the absorption coefficient of the medium at ω_{L1} = const and $\delta_L = 0$ on the Raman detuning δ_R are shown in Fig. 7a for two temperatures, T_1 = 873 K (experimental temperature) and T_2 = 10 K.

Temperature variations change the absorption coefficient magnitude but have virtually no effect on its form. The reason for these changes is a temperature-induced increase in the contribution of atoms that interact with the field at large laser detunings, which decreases absorption $K(\delta_L, \delta_R)$.

For comparison, the experimental absorption spectrum of the probing laser field (672 nm, $\delta_L = 0$) obtained by scanning the frequency of the second controlling field is shown in Fig. 7b. According to Fig. 7, the typical width of experimentally observed coherent population trapping resonances is 5–6 MHz, which is in agreement with theoretical estimates. Because field absorption is measured at a fixed ω_{L1} frequency tuned in resonance with the atomic transition, the total width of the absorption contour as a function of the δ_R Raman detuning is unbounded.

The estimates given in the Appendix show that applying longitudinal magnetic field should split the coherent population trapping resonance by $2\Delta'\omega_{12}/\omega_{13}$. The positions of coherent population trapping resonances observed experimentally are in close agreement with these results, and the splitting amounts to about 3 MHz.

The experimentally observed broad absorption contour wings at large Raman detunings δ_R are explained by the influence of collisions [17], namely, by the possibility of the transfer of atoms from different speed groups to the group resonant to the light field [19]. This mechanism was not taken into account in our calculations.

7.3. Transverse Magnetic Field

The energy level diagram for the samarium atom in a transverse magnetic field is shown in Fig. 5b. Linearly polarized laser radiation with frequency ω_{L1} can only

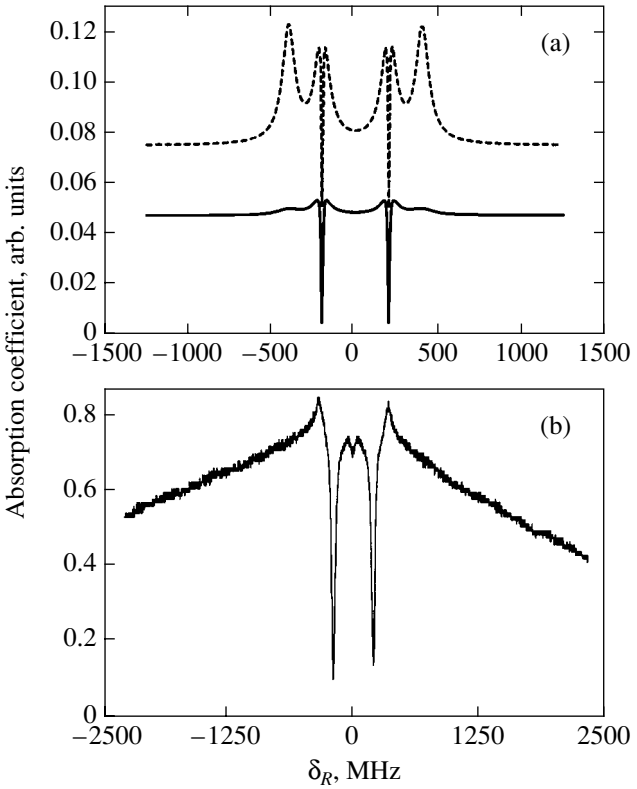


Fig. 8. (a) Theoretical dependences of the absorption coefficient of a seven-level system on Raman detuning δ_R in a transverse magnetic field calculated taking into account the Doppler effect at $T = 873$ K for two magnetic sublevel depolarization values, $G = 0$ (solid line) and $G = 0.5$ (dashed line) and (b) experimental dependence of the absorption coefficient in a 29 Oe transverse magnetic field at a 0.2 Torr buffer gas (Ar) pressure.

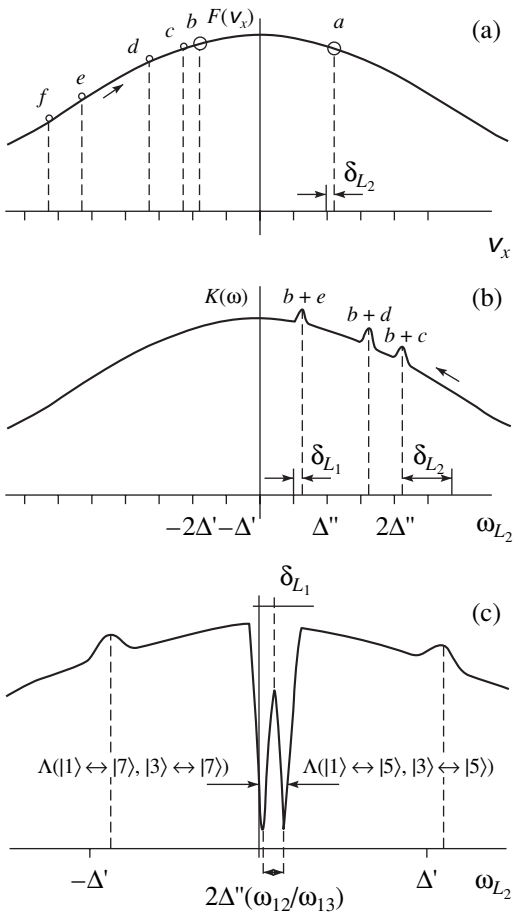


Fig. 9. Observation of coherent population trapping resonances against the background of a Doppler-broadened line; point *a* corresponds to the $|1\rangle \longleftrightarrow |7\rangle$; *b*, to $|1\rangle \longleftrightarrow |5\rangle$; *c*, to $|3\rangle \longleftrightarrow |7\rangle$; *d*, to $|2\rangle \longleftrightarrow |6\rangle$; *e*, to $|4\rangle \longleftrightarrow |6\rangle$; and *f*, to the $|3\rangle \longleftrightarrow |5\rangle$ transition. (a) $F(v_x)$ is the Maxwell velocity distribution function, points *a* and *b* correspond to particles with a definite velocity projection onto the beam propagation direction, namely, the projection at which the Doppler shift balances the detuning of field ω_{L1} frequency from the $|1\rangle \longleftrightarrow |7\rangle$ and $|1\rangle \longleftrightarrow |5\rangle$ quantum transition frequencies, respectively; points *c*, *d*, *e*, and *f* correspond to the velocity groups of particles for which the Doppler shift balances detuning δ_{L_2} of frequency ω_{L2} from the $|3\rangle \longleftrightarrow |7\rangle$, $|2\rangle \longleftrightarrow |6\rangle$, $|4\rangle \longleftrightarrow |6\rangle$, and $|3\rangle \longleftrightarrow |5\rangle$ quantum transition frequencies, respectively; (b) illustration of the formation of absorption peaks $K(\omega)$; and (c) illustration of the formation of coherent population trapping resonances against the background of a Doppler-broadened line.

cause transitions with $\Delta m = 0$ (π components) in transverse magnetic field $H_\perp$ (the $H_\perp$ vector lies in the polarization plane). At the same time, laser radiation with frequency ω_{L2} and the polarization plane orthogonal to $H_\perp$ causes transitions with $\Delta m = \pm 1$ (σ components). Two Λ systems (the transitions $|1\rangle \longleftrightarrow |6\rangle$, $|2\rangle \longleftrightarrow |6\rangle$ and $|1\rangle \longleftrightarrow |6\rangle$, $|4\rangle \longleftrightarrow |6\rangle$) are then formed. The $|3\rangle \longleftrightarrow |5\rangle$ and $|3\rangle \longleftrightarrow |7\rangle$ transitions do not participate in Λ system formation.

The absorption coefficients of the atom at rest, the depolarization effects, and the absorption coefficients of the medium taking into account the Doppler effect were calculated for the system in a transverse magnetic field using the approach described in Section 4.

The special feature of absorption spectra in a transverse magnetic field is the splitting of the coherent population trapping resonance line. The splitting of the resonance coincides with the Zeeman splitting value for level $J = 1$ sublevels $|2\rangle$ and $|4\rangle$, $\Delta\omega = 2\Delta'$ (see Appendix).

The depolarization of magnetic sublevels manifests itself similarly to the case of a longitudinal magnetic field (see Section 7.2). The resonance contrast is maximum at $G = 0$; it decreases as G grows, but its width remains virtually constant.

The experimental data are compared with the theoretically calculated absorption coefficients of the medium in a transverse magnetic field in Fig. 8. As with a longitudinal magnetic field, the positions and widths of coherent population trapping resonances observed experimentally coincide with those found in theoretical calculations (see Appendix).

8. CONCLUSIONS

In this work, we suggested a theoretical model for describing coherent population trapping in multilevel systems that allows calculations to be performed using a minimal set of input parameters. An analysis of the spectroscopic characteristics of coherent population trapping in samarium vapor in terms of this model showed that coherent population trapping resonances in the absence of an external magnetic field could well be approximated by a simple four-level model.

When a longitudinal or transverse magnetic field was applied, the spectroscopic characteristics of samarium atoms were well described by a seven-level model. The complication of the energy structure of samarium atom levels increased the number of coherent population trapping resonances and caused the appearance of additional peaks in the spectra, because the system under consideration decomposed into a set of three-level Λ systems, each being responsible for the formation of a resonance of its own. The transitions between the levels that did not directly participate in the formation of Λ systems contributed to the formation of induced absorption peaks.

In the presence of magnetic field, the depolarization of magnetic sublevels substantially influenced the shape of the absorption line and the contrast of coherent population trapping resonances, namely, the contrast of coherent population trapping resonances monotonically decreased as the depolarization constant increased.

The absorption coefficients of vapor were calculated taking into account the Maxwell velocity distribution of atoms and compared with the experimental data. It was shown that temperature variations caused changes in the absorption coefficient magnitude but had almost no effect on its form.

The results of numerical calculations accurately reproduced the experimental data on coherent population trapping resonance positions and widths and on the shape of the spectra obtained in the presence of a transverse magnetic field. Qualitatively, the theoretical estimates were also in agreement with the splitting of the coherent population trapping resonance by a small value of the order of 3 MHz observed experimentally in a longitudinal magnetic field. The reproduction of such a splitting in numerical calculations would, however, require going beyond the approximations [15] usually applied to describe the Doppler broadening effect on the formation of coherent population trapping spectra.

ACKNOWLEDGMENTS

This work was financially supported by the Russian Foundation for Basic Research (projects nos. 01-02-16311, 01-02-174-42, 01-02-174-39, 00-15-96-586), State Scientific-Technical Program "Fundamental Metrology," INTAS (INFO 00-479), and Volkswagen-Stiftung (I/73647).

APPENDIX

FEATURES OF COHERENT POPULATION TRAPPING RESONANCES AGAINST THE BACKGROUND OF A DOPPLER-BROADENED LINE IN MAGNETIC FIELDS

Longitudinal Magnetic Field

In experiments, the frequency of the first laser is constant and equal to $\omega_{L1} = \omega_{13} + \delta_{L_1}$, where δ_{L_1} is a small laser detuning. Only the particles that have a certain velocity projection along the light beam direction can be in resonance with field ω_{L1}. This projection is determined by the condition that the Doppler shift should balance the detuning of field ω_{L1} from the frequencies of the $|1\rangle \longleftrightarrow |7\rangle$ and $|1\rangle \longleftrightarrow |5\rangle$ quantum transitions (points *a* and *b* in Fig. 9a).

The absorption of wave ω_{L1} is observed at the frequencies $\omega_{13} + \Delta''$ (the $|1\rangle \longleftrightarrow |7\rangle$ transition) and $\omega_{13} - \Delta''$ (the $|1\rangle \longleftrightarrow |5\rangle$ transition). According to the definition of the Doppler effect, the corresponding group velocities are

$$\omega_{L1} = \frac{\omega_{13} + \Delta''}{1 - v_{x1}/c}$$

$$\longrightarrow \left(1 - \frac{v_{x1}}{c}\right) = \frac{\omega_{13} + \Delta''}{\omega_{L1}} \quad \text{(point } a\text{)},$$

$$\omega_{L1} = \frac{\omega_{13} - \Delta''}{1 - v_{x2}/c}$$

$$\longrightarrow \left(1 - \frac{v_{x2}}{c}\right) = \frac{\omega_{13} - \Delta''}{\omega_{L1}} \quad \text{(point } b\text{)}.$$

The interaction of particles in each velocity group with field ω_{L1} depletes level $|1\rangle$, whereas the populations of levels $|2\rangle$, $|3\rangle$, and $|4\rangle$ increase because of upper level decays. The second frequency $\omega_{L2} = \omega_{23} + \delta_{L_2}$, where δ_{L_2} is the detuning of the second field, is scanned in a wide frequency range. The *c*, *d*, *e*, and *f* points in Fig. 9a correspond to the velocity groups of particles for which the Doppler shift balances the δ_{L_2} detuning of frequency ω_{L2} from the $|3\rangle \longleftrightarrow |7\rangle$, $|2\rangle \longleftrightarrow |6\rangle$, $|4\rangle \longleftrightarrow |6\rangle$, and $|3\rangle \longleftrightarrow |5\rangle$ quantum transition frequencies. The absorption of wave ω_{L2} is observed at four frequencies, namely, $\omega_{23} \pm \Delta''$ (the $|3\rangle \longleftrightarrow |7\rangle$ and $|3\rangle \longleftrightarrow |5\rangle$ transitions) and $\omega_{23} \pm \Delta'$ ($|2\rangle \longleftrightarrow |6\rangle$ and $|4\rangle \longleftrightarrow |6\rangle$ transitions); that is, the equations for the corresponding velocity groups have the form

$$\omega_{L2} = \frac{\omega_{23} + \Delta''}{1 - v_{x3}/c}$$

$$\longrightarrow \left(1 - \frac{v_{x3}}{c}\right) = \frac{\omega_{23} + \Delta''}{\omega_{L2}} \quad \text{(point } c\text{)},$$

$$\omega_{L2} = \frac{\omega_{23} + \Delta'}{1 - v_{x4}/c}$$

$$\longrightarrow \left(1 - \frac{v_{x4}}{c}\right) = \frac{\omega_{23} + \Delta'}{\omega_{L2}} \quad \text{(point } d\text{)},$$

$$\omega_{L2} = \frac{\omega_{23} - \Delta'}{1 - v_{x5}/c}$$

$$\longrightarrow \left(1 - \frac{v_{x5}}{c}\right) = \frac{\omega_{23} - \Delta'}{\omega_{L2}} \quad \text{(point } e\text{)}.$$

$$\omega_{L2} = \frac{\omega_{23} - \Delta''}{1 - v_{x6}/c}$$

$$\longrightarrow \left(1 - \frac{v_{x6}}{c}\right) = \frac{\omega_{23} - \Delta''}{\omega_{L2}} \quad \text{(point } f\text{)}.$$

If the detuning δ_{L_2} is larger than $\omega_{23} + \Delta''$, points *c*–*f* are situated to the left of point *b*. In all these velocity groups (*a*–*f*), atoms are excited by radiation ω_{L2} or ω_{L1}. Equilibrium population distribution is disturbed. The *c*, *d*, *e*, and *f* plot points (Fig. 9a) move to the right as detuning δ_{L_2} decreases. When points *c* and *b* coincide, the atoms in these velocity groups simultaneously interact with both fields. Two transitions that do not form a Λ system are excited, but, because the distribution of particles over levels is nonequilibrium and level $|3\rangle$ is populated excessively, the absorption of the second wave increases, which corresponds to an absorption peak (Fig. 9b). The frequency of this absorption peak can be determined from the condition of the coincidence of velocity groups v_{x2} (point *b*) and v_{x3} (point *c*); that is, from the condition

$$\frac{\omega_{13} - \Delta''}{\omega_{L1}} = \frac{\omega_{23} + \Delta''}{\omega_{L2}}.$$

Using this equation, we easily find the frequency at which the first absorption peak is observed, $\omega_{L2} = \omega_{23} + 2\Delta'' + \delta_{L_1}$. Decreasing δ_{L_2} further, we observe absorption peaks corresponding to the coincidence of the v_{x2} and v_{x4} (points *b* and *d*) and v_{x2} and v_{x5} (points *b* and *e*) velocity groups. The corresponding frequencies are $\omega_{L2} = \omega_{23} + \Delta'' + \Delta' + \delta_{L_1}$ and $\omega_{L2} = \omega_{23} + \Delta' + \delta_{L_1}$. Decreasing δ_{L_2} further brings points *c* and *a* in coincidence, which corresponds to simultaneous excitation of the $|1\rangle \longleftrightarrow |7\rangle$ and $|3\rangle \longleftrightarrow |7\rangle$ transitions, that is, to a coherent population trapping resonance in the Λ system formed by the $|1\rangle$, $|3\rangle$, and $|7\rangle$ levels. Let us calculate the frequency of the coherent population trapping resonance peak. The condition of the coincidence of velocity groups v_{x1} and v_{x3} is

$$\left(1 - \frac{v_{x1}}{c}\right) = \frac{\omega_{13} + \Delta''}{\omega_{L1}}, \quad \left(1 - \frac{v_{x3}}{c}\right) = \frac{\omega_{23} + \Delta''}{\omega_{L2}}.$$

It follows that the first coherent population trapping resonance is observed at the frequency

$$\omega_{L2} = \omega_{23} + \delta_{L_1} + \frac{\omega_{12}\Delta''}{\omega_{13}}.$$

The second coherent population trapping resonance arises when the v_{x6} (point *f*) and v_{x2} (point *b*) velocity groups coincide. Accordingly, we obtain the frequency of the second coherent population trapping resonance peak in the form

$$\omega_{L2} = \omega_{23} + \delta_{L_1} - \frac{\omega_{12}\Delta''}{\omega_{13}}.$$

The distance between the two coherent population trapping resonances is

$$\omega_{L2}^{ac} - \omega_{L2}^{bf} = 2\Delta'\frac{\omega_{12}}{\omega_{13}}.$$

A further decrease in δ_{L_2} results in the appearance of absorption peaks in the left part of the plot (Fig. 9c). Note that the laser detuning of the second field being nonzero results in that the distribution of peaks and resonances is symmetrical with respect to the frequency

$$\omega_{L2} = \omega_{23} + \delta_{L_1}. \tag{A.1}$$

To summarize, we observe six absorption peaks and two coherent population trapping resonances symmetrically distributed with respect to the (A.1) frequency on a Doppler broadened contour, which differs from the picture characteristic of an atom at rest.

Transverse Magnetic Field

The reasoning is similar to that with a longitudinal magnetic field. The absorption of wave ω_{L1} is observed at the ω_{13} frequency (the $|1\rangle \longleftrightarrow |6\rangle$ transition). The second frequency $\omega_{L2} = \omega_{23} + \delta_{L_2}$, where δ_{L_2} is the detuning of the second field, is scanned in a wide frequency range. The absorption of wave ω_{L2} is observed at two frequencies, namely, $\omega_{23} \pm \Delta'$ (the $|2\rangle \longleftrightarrow |6\rangle$ and $|4\rangle \longleftrightarrow |6\rangle$ transitions). Simultaneous excitation of the $|1\rangle \longleftrightarrow |6\rangle$ and $|2\rangle \longleftrightarrow |6\rangle$ transitions and of the $|1\rangle \longleftrightarrow |6\rangle$ and $|4\rangle \longleftrightarrow |6\rangle$ transitions corresponds to coherent population trapping resonances in the Λ systems formed by the $|1\rangle$, $|6\rangle$, $|2\rangle$ and $|1\rangle$, $|6\rangle$, $|4\rangle$ levels, respectively. It follows that two coherent population trapping resonances are observed at the $\omega_{L2} = \omega_{23} \pm \Delta'$ frequencies. The distance between the resonances equals twice the Zeeman splitting of the lower level, that is, $\Delta\omega_{\perp} = 2\Delta'$. Note that the ratio between the splittings of coherent population trapping resonances for

samarium in transverse and longitudinal magnetic fields is

$$\frac{\Delta\omega_\perp}{\Delta\omega_\parallel} = \frac{2\Delta'_\perp}{2\Delta'_\parallel\omega_{12}/\omega_{13}} = 25\frac{H_\perp}{H_\parallel}.$$

REFERENCES

1. E. Arimondo, in *Progress in Optics*, Ed. by E. Wolf (Elsevier, Amsterdam, 1996), Vol. 35, p. 257.
2. G. Alzetta, A. Gozzini, L. Moi, and G. Orriols, Nuovo Cimento B **36**, 5 (1976).
3. H. R. Gray, R. M. Whitly, and C. R. Stroud, Jr., Opt. Lett. **3**, 218 (1978).
4. G. Alzetta, L. Moi, and G. Orriols, Nuovo Cimento B **52**, 209 (1979); Opt. Commun. **42**, 335 (1982).
5. A. Aspect, E. Arimondo, R. Kaiser, *et al.*, Phys. Rev. Lett. **61**, 826 (1988).
6. A. Kasapi, Phys. Rev. Lett. **77**, 1035 (1996).
7. R. Wynands and A. Nagel, Appl. Phys. B **68**, 1 (1999).
8. R. Holtzwarth, Th. Udem, and T. W. Haensh, Phys. Rev. Lett. **85**, 2264 (2000).
9. N. N. Kolachevskiĭ, A. V. Akimov, N. A. Kiselev, *et al.*, Kvantovaya Élektron. (Moscow) **31** (1), 61 (2001).
10. B. A. Grishanin, http://comsim1.phys.msu.su/people/grishanin/teaching/qsp/.
11. L. A. Vaĭnshteĭn, I. I. Sobel'man, and E. A. Yukov, *Excitation of Atoms and Broadening of Spectral Lines* (Nauka, Moscow, 1978; Springer, Berlin, 1981).
12. W. Happer, Rev. Mod. Phys. **44**, 169 (1972).
13. B. A. Grishanin, V. N. Zadkov, and D. Meschede, Phys. Rev. A **58**, 4235 (1998).
14. I. V. Bargatin, B. A. Grishanin, and V. N. Zadkov, Proc. SPIE **3736**, 246 (1998).
15. E. Kuznetsova, O. Kocharovskaya, and M. O. Scully, Proc. SPIE **4750**, 117 (2002).
16. N. N. Kolachevskiĭ, A. V. Akimov, N. A. Kiselev, *et al.*, Opt. Spektrosk. **90** (2), 201 (2001) [Opt. Spectrosc. **90**, 164 (2001)].
17. A. V. Akimov, N. N. Kolachevsky, V. N. Sorokin, and S. I. Kanorsky, in *Abstracts of International Quantum Electronics Conference, IQEC-2002, Technical Digest* (Moscow, 2002), p. 93.
18. B. M. Smirnov, *Excited Atoms* (Énergoizdat, Moscow, 1982).
19. P. F. Liao, J. E. Bjorkholm, and P. R. Berman, Phys. Rev. A **21** (6), 1927 (1980).

Translated by V. Sipachev

Journal of Experimental and Theoretical Physics, Vol. 96, No. 4, 2003, pp. 643–652.
Translated from Zhurnal Éksperimental'noĭ i Teoreticheskoĭ Fiziki, Vol. 123, No. 4, 2003, pp. 726–737.

ATOMS, SPECTRA, RADIATION

The Structural Features of Wave Collapse in Medium with Normal Group Velocity Dispersion

N. A. Zharova*, A. G. Litvak, and V. A. Mironov

Institute of Applied Physics, Russian Academy of Sciences, ul. Ul'yanova 46, Nizhni Novgorod, 603600 Russia
**e-mail: zhani@appl.sci-nnov.ru*

Received October 8, 2002

Abstract—The dynamic characteristics of self-action in three-dimensional wave packets described by the nonlinear Schrödinger equation with a hyperbolic space operator were studied analytically and numerically. The class of the initial wave field distributions for which self-focusing effects predominated over dispersion spreading and caused the arising of wave collapses was considered. The collapse of tubular wave packets was shown to be accompanied by packet shape changes during its contraction to the axis of the system. The nonlinear stabilization of collapses resulted in wave field fragmentation in the longitudinal direction followed by the expansion of the bunches thus formed along the axis. The dynamics of collapses was numerically studied taking into account medium nonlinearity saturation and nonlinear dissipation. *© 2003 MAIK "Nauka/Interperiodica".*

1. INTRODUCTION

The self-action of a wave field is one of the main processes that determine the dynamics of interaction between high-power radiation and matter. The amplitude of the envelope of three-dimensional wave packets [$E = E_{NL}\psi(x, y, z, \tau)\exp(-i\omega t + ikz)$] that propagate along axis z in a medium with characteristic field of local cubic nonlinearity E_{NL} is described by the nonlinear evolution equation [1]

$$2ik\frac{\partial\psi}{\partial z} + \Delta_\perp\psi + \frac{\partial v_{gr}}{\partial\omega}\frac{\omega^2}{kc^4}\frac{\partial^2\psi}{\partial\tau^2} + k^2|\psi|^2\psi = 0, \quad (1)$$

where $v_{gr} = (dk/d\omega)^{-1}$ is the group velocity of the pulse, $\tau = (t - z/v_{gr})$, and $\Delta_\perp = \partial^2/\partial x^2 + \partial^2/\partial y^2$. The $\partial v_{gr}/\partial\omega$ dispersion of group velocity in condensed media in the optical frequency range is as a rule normal; that is, the $\partial v_{gr}/\partial\omega$ derivative is negative [2–8]. The physical (electron–positron) vacuum is also a medium with normal dispersion [9]. In such media, Eq. (1) has a hyperbolic-type space operator rather than an elliptical operator characteristic of the standard nonlinear Schrödinger equation. In terms of dimensionless variables

$$\tilde{t} = \frac{kz}{2}, \quad (\tilde{x}, \tilde{y}) = k(x, y), \quad \tilde{z} = \frac{\tau k^{3/2}c^2}{\omega|\partial v_{gr}/\partial\omega|^{1/2}}$$

equation (1) can be written as

$$i\frac{\partial\psi}{\partial t} + \Delta_\perp\psi - \frac{\partial^2\psi}{\partial z^2} + |\psi|^2\psi = 0 \quad (2)$$

(here and below, tilde is omitted). Equation (2) can be also used to describe the self-action of waves (upper hybrid, cyclotron, etc.) in a magnetized plasma [2, 10, 11].

The special features of the self-action of wave fields described by (2) are related to the competition of wave packet contraction in transverse directions (x and y) caused by self-focusing and longitudinal dispersion spreading. Studies of the dynamics of self-action of wave packets of a Gaussian shape showed that nonuniform transverse compression of a wave packet caused its splitting in the longitudinal direction [2, 3] followed by cascade repetition of the process (multiple fragmentation) [2]. Such a scenario of the development of self-action noticeably differs from the well-known self-focusing and wave field collapse processes. The special features of cascade fragmentation and its threshold characteristics were studied in [4–7]. The nontrivial dynamics of this process is invoked to interpret the anomalous broadening of laser pulse spectra in condensed media [8, 12]. We should specially mention the works in which the coincidence of numerical simulation results and the data of special experiments in which the fragmentation of a Gaussian pulse was observed was established [4, 5].

It was shown in [13] for the two-dimensional case ($\Delta_\perp = \partial^2/\partial x^2$) that self-compression-induced effects of wave field energy concentration were most manifest if the wave packets initially had hyperbolic intensity level surfaces; this corresponds with the topology of the space operator in (2). Clearly, the role played by self-focusing effects should increase in three-dimensional systems [14]. In this work, we undertook a more detailed analytic and numerical study of the nonlinear dynamics for the class of three-dimensional wave packets that experienced collapse.

1063-7761/03/9604-0643$24.00 © 2003 MAIK "Nauka/Interperiodica"

A qualitative analysis of the dynamics of axially symmetrical three-dimensional horseshoe-shaped distributions is given in Section 2. Section 3 contains the results of a numerical study of the evolution of a broader class of tubular initial distributions.

2. AN ANALYTIC STUDY OF SELF-ACTION CHARACTERISTICS

First, consider the structural features of the dynamics of axially symmetrical wave field distributions localized near hyperbolic surfaces.

2.1. A Qualitative Study of the Structure of the Nonlinear Schrödinger Equation with a Hyperbolic Space Operator

Note that Eq. (2) for a function depending on the self-similar variable $\eta = \sqrt{r^2 - z^2}$ ($r^2 = x^2 + y^2$) is transformed to the form

$$i\frac{\partial\psi}{\partial t} + \frac{1}{\eta^2}\frac{\partial}{\partial\eta}\eta^2\frac{\partial\psi}{\partial\eta} + |\psi|^2\psi = 0. \quad (3)$$

Equation (3) describes a "spherically symmetrical" collapse, which can be classified with distributed collapses [15, 16]. It differs from the more familiar self-focusing of axially symmetrical beams by the formation of a singularity, extended in time, behind the focal region. Importantly, the power flowing (reflectionlessly) into the singularity can noticeably exceed the critical value. It is, however, difficult to directly use the results obtained in [15, 16]. In the problem under consideration, the field is not localized along the $\eta_0^2 = r^2 - z^2$ hyperbolas. It follows that the "point" of the arising of the singularity is the surface of the cone $|z| = r$. Within this region ($|z| > r$), the dynamics of the corresponding self-similar field distribution $\psi = \psi(t, \sqrt{z^2 - r^2})$ obeys the equation

$$i\frac{\partial\psi}{\partial t} - \frac{1}{\eta^2}\frac{\partial}{\partial\eta}\eta^2\frac{\partial\psi}{\partial\eta} + |\psi|^2\psi = 0. \quad (4)$$

Equation (4) differs from (3) virtually only in the sign of the "diffraction" term and, therefore, describes the self-defocusing of the wave field; that is, we observe a change in the regime of self-action. Singular solutions to (3) with a finite energy flux into the singularity were found in [15, 16]. They have singularities of converging spherical waves modified by the nonlinearity $|\psi| \propto 1/\eta(\ln\eta)^{1/2}$. Clearly, these solutions only describe a possible trend of the behavior of a wave field localized in the focusing region ($r > |z|$). In addition, they do not reflect the special features of the problem that are related to a change in the self-action regime on the $r = |z|$ cone.

2.2. Collapse of Wave Packets Localized Close to Hyperbolas

The above consideration shows that a new scenario of self-action is possible. This raises the problem of finding the initial wave field distributions for which the self-similar collapse regime is an attractor.

Our search will be oriented to distributions localized near hyperbolic surfaces. It is convenient to pass to the new variables

$$\zeta = (r^2 - z^2)/4, \quad z = z. \quad (5)$$

Such a transformation allows us to uniformly study the evolution of a wave field distribution, even with respect to z, both in the region with focusing nonlinearity (outside the cone, $|z| < r$, which corresponds to $\zeta > 0$) and in the defocusing region (inside the cone, $|z| > r$, or $\zeta < 0$). For definiteness, we will consider the dynamics of self-action in the space region $0 < z < \infty$, $0 < r < \infty$ (or, in the new variables, $0 < z < \infty$, $-\infty < \zeta < \infty$). The transformation of the equation and some simplifications can conveniently be performed using the variational approach on the basis of the corresponding Lagrangian. For (2), the action has the form

$$S = \int\Big(\frac{i}{2}(\psi\psi_t^* - \psi^*\psi_t) + |\nabla_\perp\psi|^2 - |\psi_z|^2 - \frac{|\psi|^4}{2}\Big)r\,dr\,dz. \quad (6)$$

In new variables (5), the action is written as

$$S = \int\Big(\frac{i}{2}(\psi\psi_t^* - \psi^*\psi_t) + \zeta\Big|\frac{\partial\psi}{\partial\zeta}\Big|^2 - \Big|\frac{\partial\psi}{\partial z}\Big|^2 + \frac{z}{2}(\psi_\zeta^*\psi_z + \psi_\zeta\psi_z^*) - \frac{|\psi|^4}{2}\Big)d\zeta dz. \quad (7)$$

The processes that occur in the system will be examined using the technique traditionally applied in the theory of wave self-focusing, namely, it will be assumed that the structure of the wave field along z does not change during the evolution and has a fixed, for instance, Gaussian, shape,

$$\psi = \frac{u(\zeta, t)}{\sqrt{b}}\exp\Big(-\frac{z^2}{2b^2}\Big). \quad (8)$$

This gives the contracted action in the form

$$S_c = \int\Big(\frac{i}{2}(uu_t^* - u^*u_t) + \zeta\Big|\frac{\partial u}{\partial\zeta}\Big|^2 - \frac{|u|^4}{2b\sqrt{2}}\Big)d\zeta \quad (9)$$

and allows us to find the Euler equation that describes

the evolution of the wave field along "hyperbolas,"

$$i\frac{\partial u}{\partial t} + \frac{\partial}{\partial \zeta}\zeta\frac{\partial u}{\partial \zeta} + \frac{|u|^2 u}{2b\sqrt{2}} = 0. \quad (10)$$

This equation is a variational generalization of (3) and (4) to field distributions bounded along hyperbolas (5). The above approach allows us to explicitly single out the main distinguishing feature of the problem under consideration related to a change in the self-action regime. The diffraction term in (10) is positive at $\zeta > 0$, and wave field self-focusing is therefore possible. This term is negative at $\zeta < 0$, and (10) then describes self-defocusing. At the $\zeta = 0$ boundary, where the diffraction of the wave field is balanced by the normal dispersion of the medium, the dispersion term vanishes. Note that solutions to (10) with a nonzero energy flux into the singularity are physically meaningful if $\zeta \approx 0$.

First, consider some special features of self-action by analyzing the exact integral relations that follow from (10). Note that, like the initial equation [Eq. (2)], (10) has integrals, namely, the number of quanta and the Hamiltonian,

$$I = \int_{-\infty}^{\infty} |u|^2 d\zeta, \quad H = \int_{-\infty}^{\infty} \left(\zeta\left|\frac{\partial u}{\partial \zeta}\right|^2 - \frac{u^4}{2b\sqrt{2}}\right)d\zeta. \quad (11)$$

These integrals can be used to show that the characteristic "radius" of distribution (8) localized near the one-sheeted hyperboloid ($r^2 - z^2 = \rho$),

$$\rho(t) = \int \frac{\zeta |u|^2 d\zeta}{I}, \quad (12)$$

experiences uniformly accelerated changes,

$$\frac{d^2\rho}{dt^2} = \frac{2H}{I}. \quad (13)$$

In the most interesting situation with a negative Hamiltonian ($H < 0$), the characteristic radius of the wave packet concentrated in the focusing region ($\rho \geq 0$) is "accelerated" to the boundary between the regions, crosses the boundary, and passes into the defocusing region ($\rho < 0$), where it continues to move along the z axis at the same acceleration rate. Negative ρ values correspond to field distributions localized close to the two-sheeted hyperboloid. Note that the first-order moment ρ of distribution (8) is related to second-order moments in the cylindrical system of coordinates as

$$\rho(t) = \int (r^2 - z^2)\left|\psi(r^2 - z^2, z, t^2)\right|^2 r dr dz. \quad (14)$$

This moment acquires an independent meaning for distributions localized near hyperbolas. It then determines the central hyperbola close to which the field distribution is concentrated. In the approximation that we use, Eq. (13), which describes changes in this moment, is an obvious corollary to the equations for second-order moments [2].

Changes in the characteristic scale of a wave packet that moves with acceleration (13) will be determined using the variational procedure. Let us minimize contracted action (9) on the class of trial Gaussian functions

$$u = \frac{u_0}{\sqrt{a}}\exp\left(-\frac{(\zeta - \rho)^2}{2a^2} + i\phi(\zeta - \rho)^2\right) \quad (15)$$

localized close to central hyperbola (12). Substituting (15) into (9) and integrating with respect to ζ, we obtain the corresponding contracted action and the (Euler) equation of motion

$$\rho a_{tt} - a_t \rho_t = \frac{\rho^3}{a^3} - \frac{\rho^2 u_0^2}{4a^2 b}, \quad (16)$$

which describes changes in the characteristic size of the region of wave field localization $a(t)$ during the evolution of the system.

In the problem with a negative Hamiltonian that we are interested in ($H < 0$), an approximate solution to (13), (16) can be found. Close to the boundary at which the dispersion changes sign ($\zeta \approx 0$), the characteristic radius decreases at a constant rate by the law

$$\rho \approx 2(\rho_0 |H|/I)^{1/2}(t_0 - t), \quad (17)$$

where ρ_0 is the initial position of the central hyperbola and $t_0 = (\rho_0 I/|H|)^{1/2}$ is the instant at which the central hyperbola intersects the $\zeta = 0$ boundary. In these conditions, (16) has a solution linear in $t_0 - t$,

$$a \approx A_0(t_0 - t), \quad A_0 = \frac{1}{8}\left(\frac{u_0^2}{b}\right)^{1/2}\left(\frac{\rho_0}{a_0}\right)^{1/6}. \quad (18)$$

This means that, as the central hyperbola approaches the boundary at which the self-action regime changes (from self-focusing to self-defocusing), the amplitude of the field tends to infinity by the law

$$\psi(\zeta \longrightarrow 0, t \longrightarrow t_0) \propto (t_0 - t)^{-1/2}, \quad (19)$$

that is, the wave packet experiences collapse.

2.3. Self-Similar Structures

Note that, in the vicinity of the singularity [that is, under the conditions of a decrease in the characteristic radius of distribution (15) by law (17)], there also exists a self-similar solution of the equation under consideration [Eq. (10)]. This solution has the form

$$u = \frac{v(\zeta/(t - t_0))}{\sqrt{2b\sqrt{2}(t - t_0)}} \times \exp\left\{i\left[\phi_0(t - t_0) + \frac{\zeta}{2(t - t_0)}\right]\right\}. \quad (20)$$

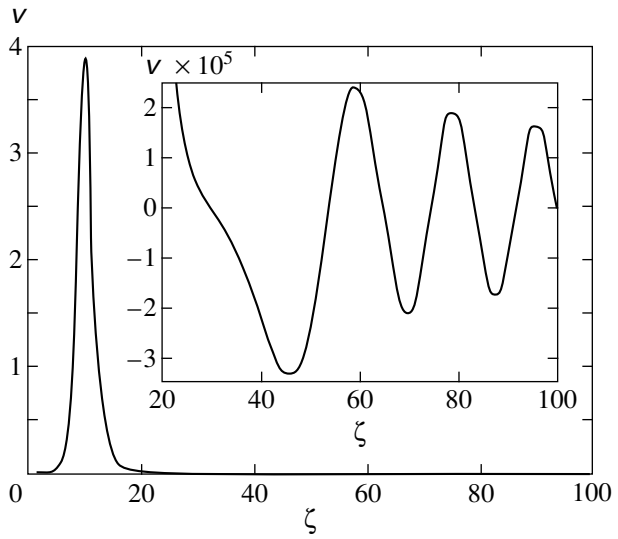


Fig. 1. Structure of a localized self-similar mode with $h = 10$ and $v(\zeta = 0) = v_0 = 3.15941666 \times 10^{-7}$. The asymptotic behavior of the solution at large ζ is shown in the inset.

The structure of the self-similar solution is determined by the equation

$$\eta \frac{d^2 v}{d\eta^2} + \frac{dv}{d\eta} + \frac{\eta}{4} v - hv + v^3 = 0, \qquad (21)$$

where $\eta = \zeta/t$. Parameter h is related to ϕ_0 by the equation $\partial\phi_0/\partial t = h/t$.

The self-similar mode under consideration, which is localized near the uniformly moving central hyperbola, is not strictly localized when $\eta \longrightarrow \infty$. The amplitude of the wave that runs away can be substantially decreased by adjusting conditions at $\eta = 0$. At $\eta \longrightarrow 0$ and $v(\zeta \longrightarrow 0) \longrightarrow 0$, (21) easily yields

$$\frac{dv}{d\eta}(\eta = 0) = hv(\eta = 0). \qquad (22)$$

It follows that the problem of finding the self-similar mode is determined by one parameter $v(\eta = 0) = v_0$. Equation (22) with a fixed h value can be used to minimize the amplitude of the wave that runs away by changing v_0. The structure of the self-similar mode at $h = 10$ is shown in Fig. 1. The amplitude of the wave running away is 2×10^{-5} at $v_0 = 3.15941666 \times 10^{-7}$ (see inset in Fig. 1). "Enery" I [see (11)] in the self-similar mode grows as h increases. A self-similar mode with two maxima can be found by increasing v_0 (at a fixed h value). The corresponding I value then increases.

To conclude this section, let us estimate the rate at which the wave field spreads over z under collapse conditions (19). We can use the initial equation (2) for the effective longitudinal size of a localized distribution, $b^2 = \int z^2 |\psi|^2 d\mathbf{r}$, to find the relation

$$\frac{d^2 b^2}{dt^2} = 8 \int \left(|\psi_z|^2 + \frac{|\psi|^4}{2} \right) d\mathbf{r}.$$

Substituting (12) and (15) into this relation yields the equation

$$\frac{d^2 b^2}{dt^2} \approx \frac{u_0^2}{4ab}, \qquad (23)$$

which describes changes in the characteristic longitudinal size of the field. For a collimated wave packet [$db/dt(t = 0) = 0$], integrating (23) with a determined by (18) easily yields

$$b^2 \approx b_0^2 + \frac{u_0^2}{4A_0 b_0} (t_0 - t) \ln|t - t_0|. \qquad (24)$$

This equation shows that the characteristic size of the wave field in the longitudinal direction hardly changes as the wave field "collapses" to the axis. What is more, the obtained $b(t)$ variations do not change the character of the behavior of the field in the vicinity of the singularity ($\zeta \longrightarrow 0$). This justifies the above approximation of a fixed field structure along hyperbolas (8). A collapse occurs if this structure remains unchanged.

It follows that the dynamics of self-action of a wave packet localized close to the $\rho_0 = r^2 - z^2 > 0$ hyperboloid in a medium with a normal dispersion of group velocity can be described as follows. A wave packet with a linear power above the critical value [which corresponds to a negative Hamiltonian (11)], virtually does not spread along hyperbolas, contracts in the transverse direction, and collapses to the axis of the system. As a result of the "fall of the wave packet onto the center," there occurs an unlimited field strengthening near the axis of the system, that is, wave collapse.

3. A NUMERICAL STUDY OF THE DYNAMICS OF SELF-ACTION

In this section, we use numerical methods to continue our study of the dynamics of the system. Consider the characteristics of long-term evolution of tubular wave packets. Numerical simulation of the dynamics of axially symmetrical localized distributions was performed with the help of a code that uses the fast Fourier transform along z and the net method along r.

Detailed numerical studies allowed us to characterize a fairly broad class of initial distributions that, during their evolution, first localized near hyperboloids and then shifted toward the axis of the system. As a result, the field close to the origin rapidly increased and the solution, whose accuracy was checked against the

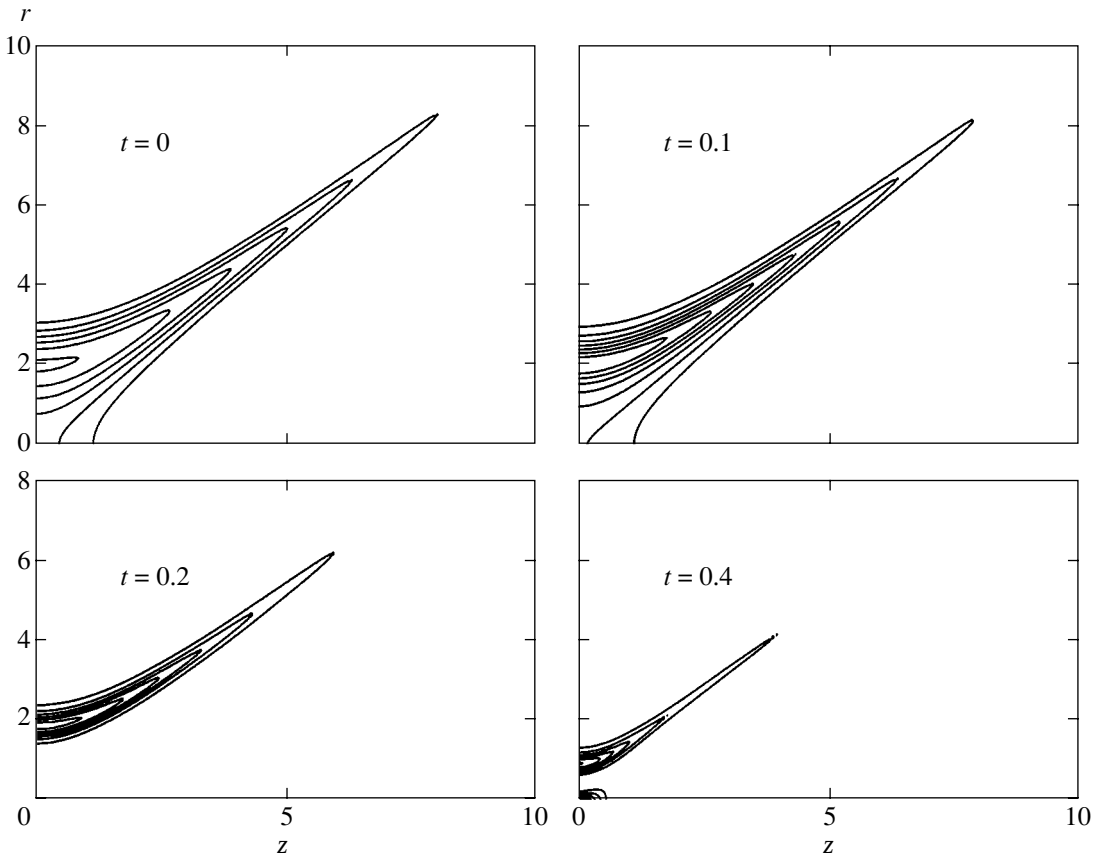


Fig. 2. Evolution of an axially symmetrical horseshoe-shaped wave packet during its "fall onto the center;" $|\psi(r, z, t = t_n)|^2$ level lines are shown.

integrals of (2), became meaningless. In order to stabilize the collapse and study the further evolution of the system, we used the already commonly accepted model procedures to regularize (2). These procedures took into account the saturation of nonlinearity or nonlinear radiation absorption in strong fields. The evolution of the field in conservative systems was studied based on the following generalization of (2):

$$i\frac{\partial\psi}{\partial t} + \frac{1}{r}\frac{\partial}{\partial r}r\frac{\partial\psi}{\partial r} - \frac{\partial^2\psi}{\partial z^2} + \frac{|\psi|^2\psi}{1 + |\psi|^2/\psi_s^2} = 0, \qquad (25)$$

where ψ_s is the field of nonlinearity saturation.

Another procedure for collapse stabilization related to taking into account nonlinear dissipation was used in numerical simulations to describe reflectionless absorption of the wave field in the singularity. The characteristics of the collapse in the problem under consideration under a change in the self-action regime were studied based on the equation

$$i\frac{\partial\psi}{\partial t} + \frac{1}{r}\frac{\partial}{\partial r}r\frac{\partial\psi}{\partial r} - \frac{\partial^2\psi}{\partial z^2} + (|\psi|^2 + i\delta|\psi|^{2n})\psi = 0, \quad (26)$$

where δ is the nonlinear dissipation parameter. The calculations were performed for $n = 3$ and 4.

First, we will study the dynamics of self-action for special initial wave field distributions localized close to hyperbolic surfaces. The principal stages of the corresponding evolution of the system were analyzed above analytically. To compare the analytic results with the numerical simulation data, we studied the behavior of the moments of wave field distributions. The evolution of a more realistic wave field distribution of the type of a tubular wave packet is analyzed at the end of this section.

3.1. A Numerical Study of the Collapse of Wave Packets Localized near Hyperbolas

The self-action picture described above finds support in numerical studies of the dynamics of initially Gaussian distributions,

$$\psi = \frac{\psi_0}{a_r}\exp\left[-\frac{(r^2 - z^2 - a_r^2)^2}{2a_r^4} - \frac{z^2}{2a_z^2}\right], \qquad (27)$$

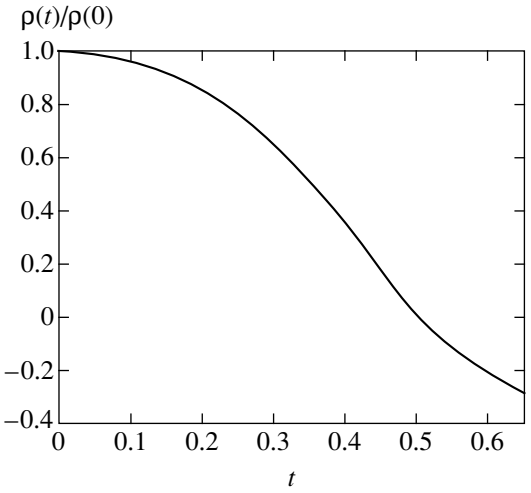


Fig. 3. Time dependence of the coordinate of the center of mass of a wave packet normalized with respect to the initial coordinate, $\rho(t)/\rho(0)$.

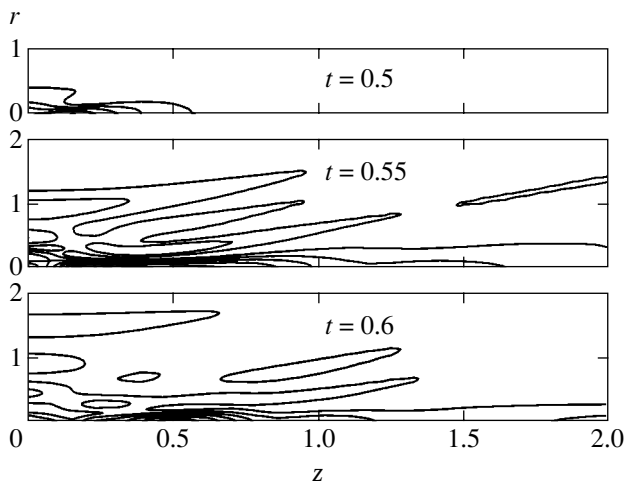


Fig. 4. Further evolution (see Fig. 2) of the wave packet in the focal region determined taking into account nonlinearity saturation; $|\psi|$ field level lines are shown.

which are localized in the vicinity of the $r^2 - z^2 = a_r^2$ hyperbolas. The level lines of the $|\psi(r, z, t)|^2$ function at successive instants are shown in Fig. 2 for the parameters $\psi_0 = 7$, $a_r = 2$, and $a_z = 6$. The corresponding Hamiltonian (11) value is negative. At the initial stage, the evolution of the wave packet predominantly occurs in the transverse direction. As a result, the field distribution contracts to the $r^2 - z^2 = 4$ "central" hyperbola. At $t > 0.2$, the displacement of the wave field toward the straight line (conic surface) that separates regions with different self-action regimes becomes noticeable. This and the absence of wave packet fragmentation along the "base" hyperbola distinguishes the evolution of the system under consideration from the picture observed in the two-dimensional case [13]. It appears that the displacement of the wave field toward the axis of the system prevents the development of longitudinal fragmentation instability. This conclusion can be drawn from an analysis of the fine details of structural changes in $|\psi|$ (rather than $|\psi|^2$) level lines corresponding to the initial conditions that decelerate the collapse to the center. One of such variants is given below in a consideration of a tubular wave packet. The specified processes (contraction and displacement) cause a rapid field increase at the origin; at $t \approx 0.45$, solving (2) by numerical methods becomes meaningless.

First, consider the structure of the wave field in the focal region at $t > 0.4$ in the conservative case with the use of (25) (at the nonlinearity saturation field value $\psi_s = 10$).

Note that the characteristic wave packet radius (Fig. 3) moves to the axis of the system ($z = 0$) with an acceleration during the collapse. To emphasize the fine details of the evolution of the wave field distribution, $|\psi|$ (rather than $|\psi|^2$) level lines are shown in Fig. 4. The contraction of the wave packet to the center of the system ($r = 0$, $z = 0$) is accompanied by the formation of a strongly nonuniform (aberrational) structure in the focal region. A comparison of the structures of the $|\psi|$ and $|\psi|^2$ level lines shows that a two-scale field distribution localized close to the center of the system ($r \approx 0$, $z \approx 0$) is formed. The field amplitude at a maximum reaches a value that is an order of magnitude larger than the nonlinearity saturation field value (Fig. 5), which is evidence of a high collapse rate.

Two processes are clearly seen in the dynamics of the collapse (see Fig. 4). Note that some part of the field is reflected and moves away from the axis of the system. The behavior of the remaining part is qualitatively similar to that observed for Gaussian initial distributions [2–5] but develops under the conditions of strong nonlinearity saturation. Contraction, which is nonuni-

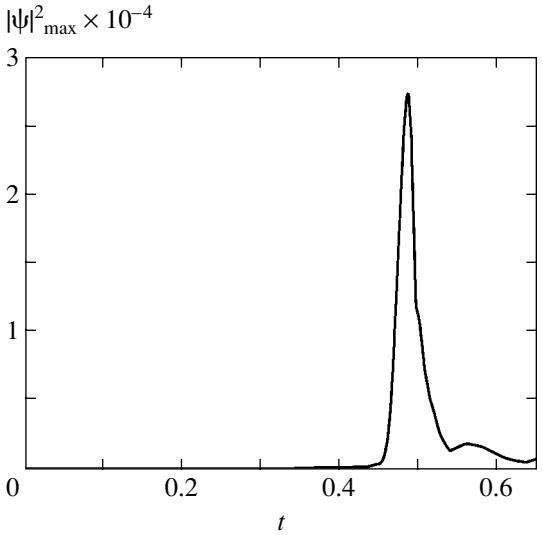


Fig. 5. Time dependence of the square of field amplitude at a maximum during a conservative collapse.

form along z (the central distribution part, where the amplitude is larger, experiences stronger radial contraction), causes the formation of a constriction in the center (see Fig. 4). The constriction grows tighter, and the wave packet therefore divides in two fragments, which move along the z axis in opposite directions. The packet expands along z because the field amplitude in it exceeds the saturating amplitude ($|\psi| > \psi_s$) and fragments without hardly changing its transverse dimension.

Next, consider the dissipative case. First, note that the position of the characteristic field distribution radius depends on time virtually in the same way as when nonlinearity saturation is taken into account (see Fig. 3); that is, its motion toward the axis of the system is uniformly accelerated. The field maximum amplitude during the collapse increases approximately (more than) tenfold. This corresponds to the selected parameters, but the amplitude is noticeably smaller than in the case of nonlinearity saturation. The energy of the wave packet during collapses decreases by 30%.

The most important structural differences manifest themselves in the high field region. The $|\psi|$ level lines are shown in Fig. 6. A comparison with the corresponding picture for the conservative case (see Fig. 4) reveals two differences. First, note the absence of "reflection" from the axis of the system. Secondly, it is as though the wave packet near the conic surface ($r = z$ straight lines in Fig. 6), which separates regions with different self-action regimes, is delayed for a time comparable with duration of the collapse. In this sense, the situation under consideration is more akin to that with a spherically symmetrical collapse. However, no complete wave field absorption occurs in the singularity region ($r = |z|$), and a weakened wave packet passes into the defocusing region.

In the immediate vicinity of the center ($r \approx z \approx 0$), the structure of the field is substantially more nonuniform than in the conservative case.

3.2. A Study of the Collapse of Tubular Wave Packets

Next, consider the evolution of the more realistic wave packet distribution

$$\psi = \frac{\psi_0 r^4}{a_r^5} \exp\left[-\frac{r^2}{2a_r^2} - \frac{z^2}{2a_z^2}\right], \qquad (28)$$

which, for instance, simulates the structure of the field of a hollow (near the axis of the system) spatially limited wave packet. It is well known that, at a power exceeding the critical value, tubular wave beams, uniform along z, experience contraction in the transverse direction and are self-focused toward the z axis. It can therefore be expected that, in a distribution of type (28),

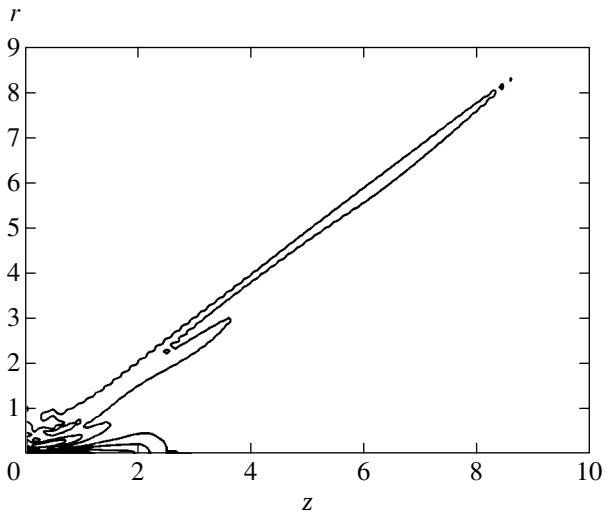


Fig. 6. Field structure during dissipative collapse stabilization (dissipation parameter $\delta = 2 \times 10^{-8}$) at the instant corresponding to the passage of the center of mass of the wave packet (see Fig. 3) across the boundary at which the regime of self-action changes [$\rho(t = 0.5) = 0$]; $|\psi(r, z)|$ field level lines are shown for the same initial conditions as in Fig. 2.

which is nonuniform along z, the central part ($z \approx 0$) will be focused toward the axis ($r = 0$) at a higher rate than the peripheral part ($z \sim a_z$). As a result, a horseshoe-shaped wave packet distribution is formed. The dynamics of such a distribution has been described above.

Numerical calculations confirm this prediction. For definiteness, let us first consider the conservative case with nonlinearity saturation. Equation (25) with the same saturating field value as previously ($\psi_s = 10$) will be used. Next, we will give additional results obtained for dissipative collapse stabilization. The following parameters were selected: $\psi_0 = 0.2$, $a_r = 2$, and $a_z = 12$. The field maximum was then situated at a distance from the axis of a system two times larger than in the variants considered above. Under these conditions, the system passed into the regime already studied.

The calculation results are shown in Fig. 7, according to which wave beam self-focusing takes a fairly long time. The self-focusing occurs without changes in the characteristic tubular field distribution radius. At the next stage, a horseshoe structure is formed. This is accompanied by the development of fragmentation instability. In the problem under consideration, as distinct from the two-dimensional problem [13], fragmentation does not play a noticeable role and is stabilized by the displacement of the horseshoe-shaped distribution to the axis of the system. During the collapse to the center, the pulse is divided in two. Plotting $|\psi|$ level lines on a larger scale (Fig. 8) gives an idea of the fine structure of the wave field in the focal plane. A comparison of this picture with the structure of $|\psi|^2$ level lines shows that the distribution of the wave field is two-

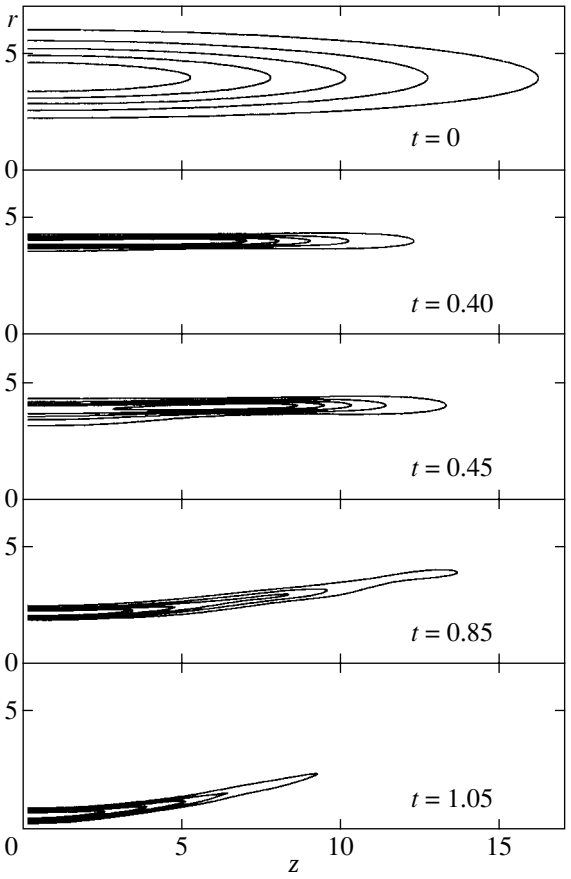


Fig. 7. Tubular wave packet evolution during collapse in the conservative case; $|\psi(r, z, t = t_n)|^2$ level lines are shown.

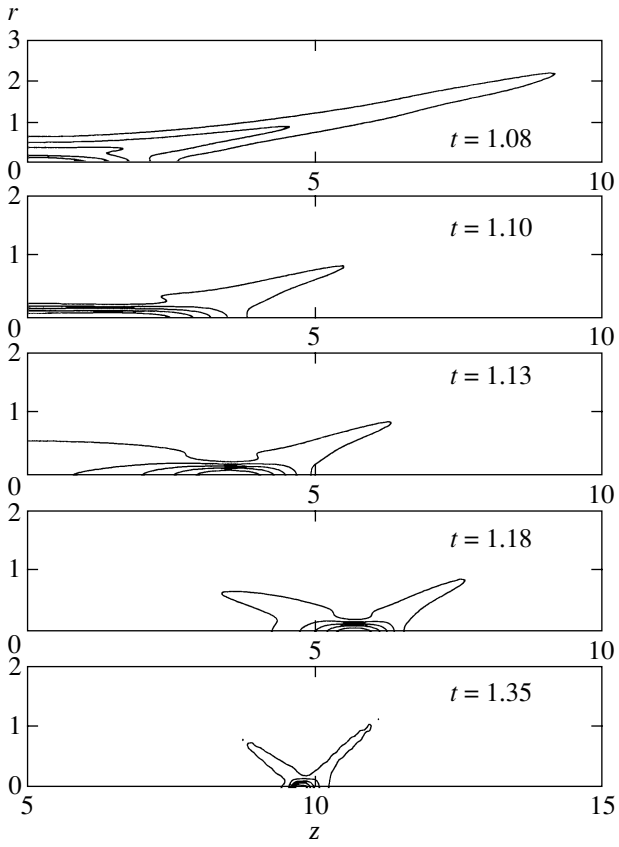


Fig. 8. Further evolution (see Fig. 7) of the tubular wave packet in the conservative case close to the axis of the system; $|\psi(r, z, t = t_n)|$ level lines are shown.

scale. The central part of the pulse has a somewhat smaller amplitude than that obtained in the corresponding variant of the initially horseshoe-shaped distribution localized near hyperbolas. The peripheral region of the wave packet consists of a reflected wave (behind the pulse) and a wave field spread along the "horseshoe;" this field continues to move toward the axis of the system. Because the problem has axial symmetry, this wave field region is strengthened near the axis of the system (*z* axis). It appears that, thanks to this strengthening, the maximum wave packet amplitude (the level lines of this packet are shown in Fig. 8) hardly changes over a fairly prolonged evolution.

The dynamics of self-action in the dissipative case is similar in many respects. The only difference is the absence of a reflected wave. The high field region created by the peripheral part of the horseshoe structure near the axis of the system (*z* axis) becomes the region of strong absorption. Complete nonlinear absorption of an electromagnetic pulse along the propagation path is about 50% in this variant of calculations. The special features of the dynamics of self-action were described in detail in [14].

Let us perform a more detailed analysis of the evolution of tubular wave packet (28) with an amplitude 1.5 times larger ($\psi_0 = 0.3$). In the conservative case, such an increase in the initial amplitude considerably complicates calculations and requires using shorter computation steps, larger arrays, a lower saturating field value (which is fairly small as it is), etc. For this reason, the results will be given for the dissipative mechanism of collapse stabilization. Figure 9 shows that, as in Fig. 7, the self-focusing of a tubular wave beam is initially accompanied by its fragmentation in the longitudinal direction. At the next stage, a horseshoe-shaped structure is formed and displaced to the axis of the system. A noticeable wave packet fraction then remains localized in the region of the initial tubular beam location. Note that, in this variant too, the development of fragmentation instability does not have any catastrophic consequences either for the horseshoe-shaped structure or for the remaining wave packet part. It can be suggested that fragmentation is weakened by the development of self-focusing instability along the radius of the beam. After the formation of the horseshoe structure, the second ring layer is formed and experiences similar evolution. Fragmentation in the center

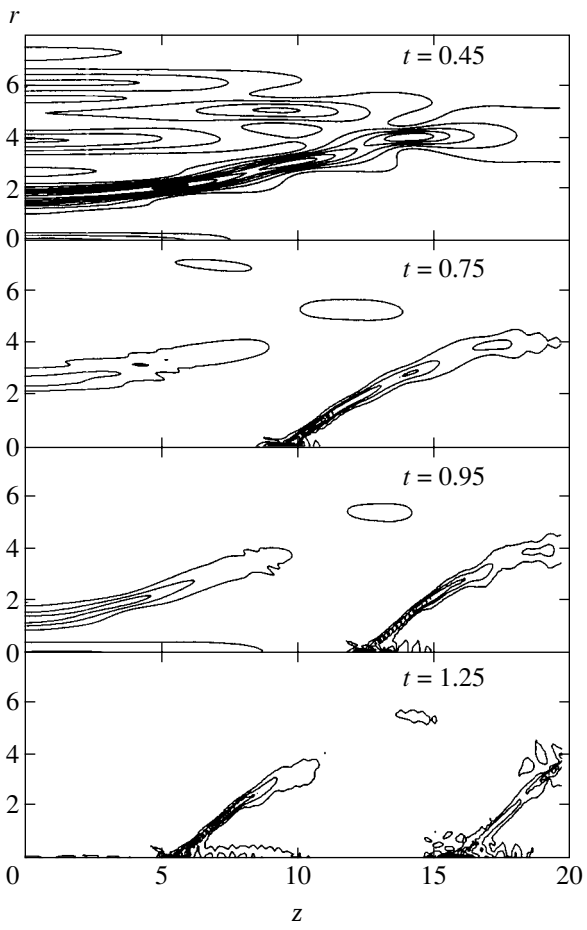


Fig. 9. Evolution of axially symmetrical initially tubular wave packet with amplitude $\psi_0 = 0.3$ under dissipative collapse stabilization conditions, $\delta = 1.8 \times 10^{-5}$; $|\psi(r, z, t = t_n)|$ level lines are shown.

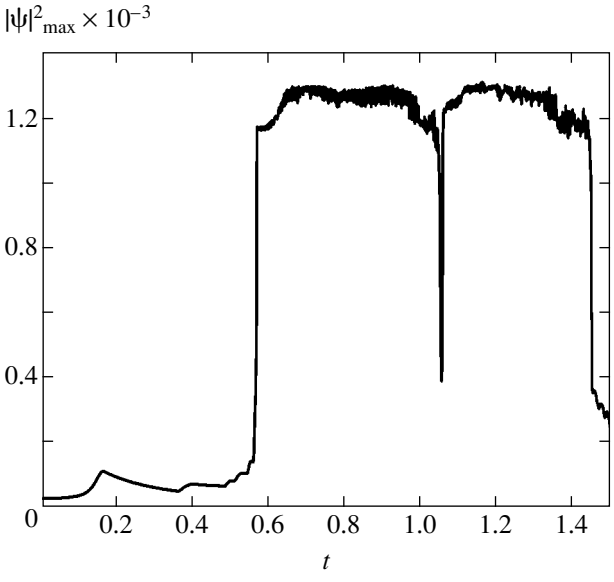


Fig. 10. Time dependence of the $|\psi|^2_{max}$ maximum field amplitude.

($r_\perp = 0$, $z = 0$) causes the formation of two pairs of strong cone-shaped field regions, which move along the axis of the system in opposite directions. The maximum field amplitude in the focal region remains nearly constant (Fig. 10). A sharp minimum of the $|\psi|^2_{max}(t)$ dependence at $t \approx 1.1$ corresponds to the displacement of the strong absorption region from the first pulse to the second. Complete nonlinear electromagnetic field absorption along the propagation path increases to 70%.

It can therefore be stated that the above numerical study of the dynamics of self-action of tubular initial field distributions substantiates the analytic results obtained in Section 2. This primarily refers to the formation of horseshoe-shaped self-similar field structures followed by their accelerated motion to the axis of the system. Also note the possibility of the existence of a self-similar mode of a higher order (with two or more maxima) predicted by (21). This possibility manifests itself by the excitation of two sequential horseshoe-shaped structures.

4. CONCLUSION

We analyzed a new scenario of the dynamics of self-action of an axially symmetrical wave packet in a system described by the nonlinear Schrödinger equation with a hyperbolic space operator. The class of initial wave field distributions whose evolution allowed the special features of the equation under study [Eq. (2)] to be illustrated was determined. In the competition between self-focusing compression and wave field dispersion spreading along orthogonal directions, self-focusing plays the predominant role under these initial conditions. It causes narrowing of tubular wave packets in the transverse direction, the formation of horseshoe-shaped structures, and the displacement of the wave packet toward the axis of the system, which eventually (by virtue of cylindrical symmetry of the problem) creates a collapse regime. The spread of a horseshoe-shaped distribution in the longitudinal direction has virtually no effect on the dynamics of singularity formation. However clearly, in the regime of falling onto the center, the peripheral part of the wave packet reaches the axis of the system at later times and contributes to maintaining the field at a high level during the whole interaction time. In this sense, the type of collapse under consideration should be considered distributed. It is in many respects similar to a spherically symmetrical collapse described by the nonlinear Schrödinger equation with an elliptical space operator. However, the absence of the symmetrization of wave packet distributions in the problem under consideration and the predominance of self-focusing over dispersion make this collapse also similar to the axially symmetrical collapse of the nonlinear Schrödinger equation. Because of the hyperbolic character of initial equation (2), wave packet fragmentation along the axis (close to $r = 0$ and $z = 0$) is accompanied by the formation of several pairs

of high field regions ("hot" points), which move in different directions. Our numerical study of this process under the conditions of nonlinear collapse stabilization shows that these hot points travel through distances longer than the longitudinal pulse length without noticeable spatial structure changes.

As concerns applications to nonlinear optics, the self-action regime under consideration should be subjected to a detailed experimental examination. First, collapse stability with respect to the loss of axial symmetry and deviations from the approximation of slow envelope changes and the paraxial approximation [17] should be studied. These effects may play an important role in the collapse of wave packets whose power is much higher than the critical self-focusing power. It is, we believe, fairly clear that the use of such a self-action regime offers promise, for instance, for artificial photon crystals and creating channels in condensed media with the purpose of using such structures for the radiation frequency conversion and acceleration of charged particles.

ACKNOWLEDGMENTS

This work was financially supported by the Russian Foundation for Basic Research (project nos. 02-02-17277, 01-02-17388). The authors thank A.A. Balakin for the determination of the structure of the self-similar mode (see Fig. 1).

REFERENCES

1. A. G. Litvak and V. I. Talanov, Izv. Vyssh. Uchebn. Zaved., Radiofiz. **10**, 539 (1967).
2. N. A. Zharova, A. G. Litvak, T. A. Petrova, *et al.*, Pis'ma Zh. Éksp. Teor. Fiz. **44**, 12 (1986) [JETP Lett. **44**, 13 (1986)]; Izv. Vyssh. Uchebn. Zaved., Radiofiz. **29**, 1137 (1986).
3. J. E. Rothenberg, Opt. Lett. **17**, 583 (1992); P. Chernev and V. Petrov, Opt. Lett. **17**, 172 (1992); Opt. Commun. **87**, 28 (1992).
4. J. E. Ranka, R. W. Schrimer, and A. L. Gaeta, Phys. Rev. Lett. **77**, 3783 (1996).
5. A. A. Zozulya, S. A. Diddams, A. G. Van Engen, and T. S. Clement, Phys. Rev. Lett. **82**, 1430 (1999); A. A. Zozulya, S. A. Diddams, and T. S. Clement, Phys. Rev. A **58**, 3303 (1998).
6. G. G. Luther, A. C. Newell, and J. V. Moloney, Physica A (Amsterdam) **49**, 4085 (1994); G. G. Luther, A. C. Newell, J. V. Moloney, *et al.*, Opt. Lett. **19**, 789 (1994); Opt. Lett. **19**, 862 (1994).
7. L. Berge and J. J. Rasmussen, Phys. Plasmas **3**, 824 (1996); L. Berge, J. J. Rasmussen, E. A. Kuznetsov, *et al.*, J. Opt. Soc. Am. B **13**, 1879 (1996); L. Berge, E. A. Kuznetsov, and J. J. Rasmussen, Phys. Rev. E **53**, R1340 (1996); G. Fibich, V. M. Malkin, and G. C. Papanicolaou, Phys. Rev. A **52**, 4218 (1995); K. Germaschevski, B. Grauer, L. Berge, *et al.*, Physica D (Amsterdam) **151**, 175 (2001).
8. A. Brouder and S. I. Chin, Phys. Rev. Lett. **80**, 4406 (1998).
9. N. N. Rozanov, Zh. Éksp. Teor. Fiz. **113**, 513 (1998) [JETP **86**, 284 (1998)].
10. A. G. Litvak, T. A. Petrova, A. M. Sergeev, and A. D. Yunakovskiĭ, Fiz. Plazmy **9**, 495 (1983) [Sov. J. Plasma Phys. **9**, 287 (1983)]; A. G. Litvak, A. M. Sergeev, and N. A. Shakhova, Pis'ma Zh. Tekh. Fiz. **5**, 862 (1979) [Sov. Tech. Phys. Lett. **5**, 33 (1979)].
11. J. R. Myra and C. S. Liu, Phys. Fluids **23**, 2258 (1980).
12. I. G. Koprinkov, A. Suda, P. Wang, and K. Midorikawa, Phys. Rev. Lett. **84**, 3847 (2000).
13. A. G. Litvak, V. A. Mironov, and É. M. Sher, Zh. Éksp. Teor. Fiz. **118**, 1463 (2000) [JETP **91**, 1268 (2000)]; A. G. Litvak, V. A. Mironov, and E. V. Sher, Phys. Rev. E **61**, 891 (2000).
14. N. A. Zharova, A. G. Litvak, and V. A. Mironov, Pis'ma Zh. Éksp. Teor. Fiz. **75**, 655 (2002) [JETP Lett. **75**, 539 (2002)].
15. S. N. Vlasov, L. V. Piskunova, and V. I. Talanov, Zh. Éksp. Teor. Fiz. **95**, 1945 (1989) [Sov. Phys. JETP **68**, 1125 (1989)]; S. N. Vlasov, Izv. Vyssh. Uchebn. Zaved., Radiofiz. **42**, 468 (1999).
16. V. E. Zakharov, N. E. Kosmatov, and V. F. Shvets, Pis'ma Zh. Éksp. Teor. Fiz. **49**, 431 (1989) [JETP Lett. **49**, 492 (1989)]; V. M. Malkin, Pis'ma Zh. Éksp. Teor. Fiz. **48**, 603 (1988) [JETP Lett. **48**, 653 (1988)].
17. S. N. Vlasov and V. I. Talanov, *Wave Self-Focusing* (Inst. Prikl. Fiz. Ross. Akad. Nauk, Nizhni Novgorod, 1997); N. N. Rozanov, N. V. Vysotina, and A. G. Vladimirov, Zh. Éksp. Teor. Fiz. **118**, 1307 (2000) [JETP **91**, 1130 (2000)].

Translated by V. Sipachev

Journal of Experimental and Theoretical Physics, Vol. 96, No. 4, 2003, pp. 653–657.
Translated from Zhurnal Éksperimental'noĭ i Teoreticheskoĭ Fiziki, Vol. 123, No. 4, 2003, pp. 738–743.

ATOMS, SPECTRA, RADIATION

Photon Production in a Dense Medium

A. V. Koshelkin

Moscow Institute of Engineering Physics, Kashirskoe sh. 31, Moscow, 115409 Russia

Received October 24, 2002

Abstract—On the basis of diagram formalism for two-particle Green functions, the bremsstrahlung of charged particles interacting in a dense medium is investigated. In the case of production of soft photons, exact expressions for two-particle Green functions corresponding to the process of particle bremsstrahlung in the substance are obtained. The Green functions found are fully determined by the set of closed irreducible diagrams. It is shown that, in the case of radiation in a sufficiently dense medium in a far long-wave region of the spectrum, the coherent multiple particle scattering results in an additional (as compared that reported earlier in [1–13]) suppression of the bremsstrahlung photon yield. © *2003 MAIK "Nauka/Interperiodica".*

1. INTRODUCTION

For the first time, the influence of the scattering medium on the radiation of fast charged particles was considered in [1, 2]. In these papers, the suppression of the bremsstrahlung in the long-wave region of the spectrum due to the multiple elastic scattering of such particles in the substance was pointed out (the Landau–Pomeranchuk effect). In [3, 4], Migdal constructed a quantitative theory of this effect. The methods proposed in [3, 4] for calculating the spectrum of the bremsstrahlung in a substance were further developed in [5–8] while investigating how the frequency distribution of the bremsstrahlung is affected by the dispersion properties of the scattering medium [5, 6], its boundary [7, 8], and inelastic processes occurring in the substance [6]. Recently, the influence of the medium on bremsstrahlung was investigated in [9–13], where the application of the continual integration method for calculating the spectrum of bremsstrahlung in the substance was developed [9, 10], and the Landau–Pomeranchuk–Migdal effect in quantum chromodynamics [10, 11] and for the Coulomb interaction of particles in a scattering medium [11] was considered.[1]

In all the above-mentioned papers, multiple scattering in the medium was treated as a set of successive acts of pair interactions of particles, and the influence of the simultaneous collision of several (more than two) particles on forming the bremsstrahlung spectrum was neglected. It is clear that effects of this kind must play an important role in photon production in sufficiently dense scattering media, where the gas parameter is greater than or approximately equal to unity, and in forming the limiting soft bremsstrahlung when particles pass through the substance.

[1] A detailed analysis of publications devoted to the Landau–Pomeranchuk–Migdal effect can be found in review [12].

In this paper, a diagram formalism is developed for calculating the spectrum of soft bremsstrahlung. The formalism is based on determining two-particle Green functions in a nonequilibrium medium. Exact relations for these Green functions corresponding to the process of formation of the spectrum of soft bremsstrahlung in a substance are obtained. The two-particle Green functions obtained are determined by a set of topologically different irreducible closed diagrams. It is shown that, in the case of a sufficiently dense medium, the coherent multiple particle scattering results is an additional (as compared with that demonstrated earlier in [1–13]) suppression of the bremsstrahlung intensity in the medium in a far long-wave region of the spectrum.

2. TWO-PARTICLE GREEN FUNCTIONS AND PARTICLE BREMSSTRAHLUNG IN A MEDIUM

The probability of photon production by the current j^ν is given by the expression [13]

$$d^4w = 4\pi e_\mu e_\nu^*(1+n_\gamma)\delta(\omega^2-\omega_k^2) \times \int d^4x_1 d^4x_2 e^{-ik(x_1-x_2)}\langle j^{\mu+}(x_1)j^\nu(x_2)\rangle \frac{d^4k}{(2\pi)^3}, \quad (1)$$

where $k = (\omega, \mathbf{k})$ and e_α are the 4-momentum and the polarization vector of the photon, n_γ is the density of the photon states, and $j^\nu(x)$ is the current of the particles producing photons. The angle brackets denote averaging over a certain state of particles of the medium (in general, this state can be nonequilibrium), and x are 4-coordinates.

1063-7761/03/9604-0653$24.00 © 2003 MAIK "Nauka/Interperiodica"

The bilinear combination of currents in Eq. (1) can be represented in the following form:

$$\langle j^{\mu+}(x_1)j^{\nu}(x_2)\rangle = \langle i|(\hat{O}^{\mu})_{\alpha,\beta}((\hat{O}^{\dagger})^{\nu})_{\gamma,\delta}|j\rangle \times \langle \Psi^{\dagger}_{\delta}(x_1)\Psi_{\gamma}(x_1)\Psi_{\beta}(x_2)\Psi^{\dagger}_{\alpha}(x_2)\rangle, \quad (2)$$

where $\langle i|(\hat{O}^{\mu})_{\alpha,\beta}((\hat{O}^{\dagger})^{\nu})_{\gamma,\delta}|j\rangle$ is the matrix element of an operator independent of 4-coordinates; $\Psi_{\alpha}(x)$ are psi operators in the Heisenberg representation; and α, β, γ, and δ are spin variables.

Thus, the problem of determining the probability of photon production in a medium is reduced to determining the correlator consisting of four Ψ functions. In turn, this correlator is proportional to the two-particle Green function.

Suppose that the influence of the particle scattering in the medium on the change in spin states of the particles is negligible. Then, expanding the correlator $\langle \Psi^{\dagger}_{\delta}(x_1)\Psi_{\gamma}(x_1)\Psi_{\beta}(x_2)\Psi^{\dagger}_{\alpha}(x_2)\rangle$ in a complete set of plane waves, we obtain from formulas (1) and (2) the following expression for the probability of photon production:

$$d^4W = \overline{d^4w} = \frac{8\pi}{(2s+1)}\frac{d^4k}{(2\pi)^3}\{(1+n_{\gamma})\delta(\omega^2-\omega_k^2) \times \int d^4p_1 d^4p_2 d^4p_3 d^4p_4 \delta^4(p_3-p_1-k)\delta^4(p_2-p_4-k) \times \mathrm{Tr}[e_{\mu}e^*_{\nu}\langle i|(\hat{O}^{\mu})_{\alpha,\beta}((\hat{O}^{\dagger})^{\nu})_{\gamma,\delta}|j\rangle \times \bar{u}^{\alpha}(p_1)\bar{u}^{\beta}(p_2)u^{\gamma}(p_3)u^{\delta}(p_4)] \times K(p_4(+);\ p_2(-)|p_3(-);\ p_1(+))\}, \quad (3)$$

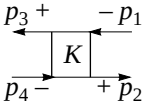
where $p_i = (p_i^0, \mathbf{p}_i)$ is the 4-momentum of the radiating particle, s is the particle spin, $u^{\alpha}(p)$ are the Dirac spinors, T is the observation time or the time of the particle motion in the medium, and the bar over d^4w denotes the averaging and summing over the corresponding spin states of the particles. The function $K(p_4(+);\ p_2(-)|p_3(-);\ p_1(+))$ is the two-particle Green function $K(4(+), 2(-), 3(-), 1(+))$ in the momentum representation,

$$K(p_4(+);\ p_2(-)|p_3(-);\ p_1(+)) = \int dX_1 dX_2 dX_3 dX_4 \exp(-ip_1X_1 - ip_2X_2 + ip_3X_3 + ip_4X_4)\bar{u}(p_3)\bar{u}(p_4)u(p_1)u(p_2) \times \langle \Psi^{\dagger}_{\delta}(x_1)\Psi_{\gamma}(x_1)\Psi_{\beta}(x_2)\Psi^{\dagger}_{\alpha}(x_2)\rangle. \quad (4)$$

Thus, the problem of determining the spectrum of the particle bremsstrahlung in a medium in the Keldysh formalism is reduced to determining the so-called nonchronologized two-particle Green function $K(p_4(+);\ p_2(-)|p_3(-);\ p_1(+))$ [14] in the momentum representation. Note that the function $K(p_4(+);\ p_2(-)|p_3(-);\ p_1(+))$ is indirectly present in formula (3), namely, in the form of integrals with respect to appropriate momenta. Therefore, to find the bremsstrahlung spectrum, we use, instead of the function $K(p_4(+);\ p_2(-)|p_3(-);\ p_1(+))$, its various convolutions with δ functions taking into account the conservation laws for photon emission with respect to the corresponding arguments (with respect to the momenta p_1; p_2; p_3; and p_4).

Introduce the following notation:

$$K(p_4(+);\ p_2(-)|p_3(-);\ p_1(+)) \equiv$$



, (5)

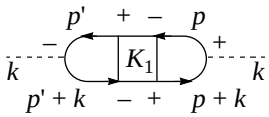
where K is the two-particle Green function [14] in the Keldysh diagram formalism [15].

Then, formula (3) can be represented in the following form:

$$d^4W = \frac{8\pi}{(2s+1)}\frac{d^4k}{(2\pi)^3} \times \{(1+n_{\gamma})\delta(\omega^2-\omega_k^2)[K_1(k)+K_2(k)]\}, \quad (6)$$

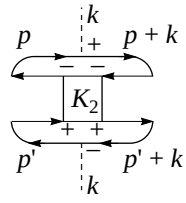
where

$K_1(k) =$



, (7)

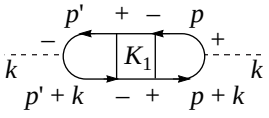
$K_2(k) =$



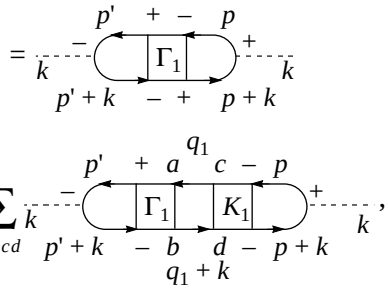
. (8)

The solid lines in relations (7) and (8) denote exact one-particle Green functions in the medium and the dotted lines denote the emitted photon.

The function $K_1(k)$ satisfies [14] the following diagram equation:[2]



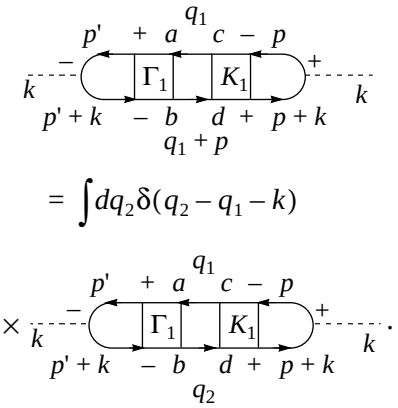
[2] Obviously, the equation for the function $K_2(k)$ has the same form.



(9)

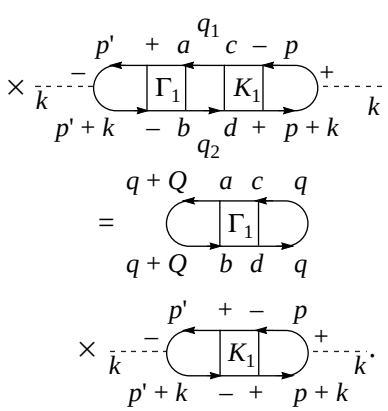
where Γ_1 is the exact irreducible vertex function consisting of all topologically different diagrams that cannot be cut into two diagrams by a vertical line crossing only two solid or dotted lines (denoting, as is usual in the Keldysh diagram formalism [15], the particle interaction). The letters a, b, c, and d can take the values + or – [14, 15].

Obviously, the following equality holds:



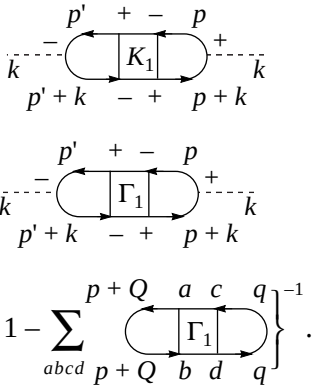
(10)

Suppose that the energy ω of produced photons is low as compared with the energy E of the radiating particle and with the energy ΔE transmitted to other particles due to a single interaction with emission of a photon: $\omega \ll \min\{E, \Delta E\}$. Then, the arguments of the vertex functions and the two-particle Green function in formula (10) can be expanded into a series in the differences $p - q_i$ and $p' - q_i$. As a result, relation (10) implies that, up to terms on the order of $\omega/\min\{E, \Delta E\}$, inclusively, we have



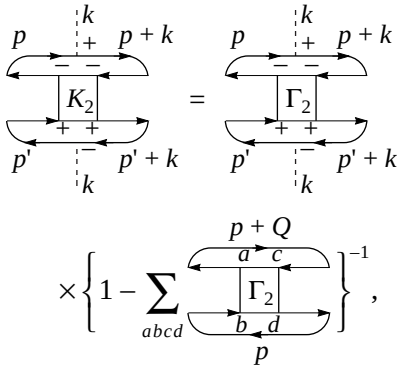
(11)

Substituting the expression obtained into Eq. (9), we find



(12)

Similarly, with the same accuracy, for the function $K_2(k)$, we obtain



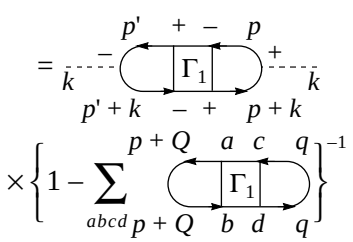
(13)

where Γ_2 is the exact vertex function consisting of diagrams with more than two solid lines in the two-particle channel; Q is the momentum transmitted as a result of a single interaction between particles with emission of a photon; in the case $\omega \ll \min\{E, \Delta E\}$, it is equal to [16]

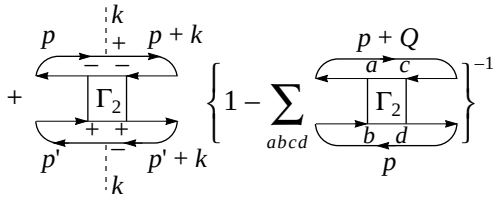
$$Q = \left(k, \frac{\partial}{\partial p}\right)p, \tag{14}$$

where the parentheses denote the usual scalar product and p and k are the 4-momenta of the particle and the photon, respectively.

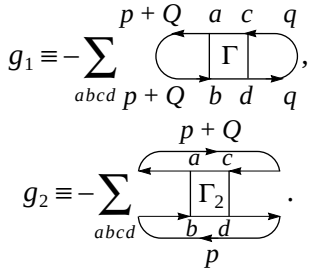
Formulas (12) and (13) imply the following expression for the function $K(k) = K_1(k) + K_2(k)$ appearing in formula (6) for the probability of photon production:



(15)



Expressions (6) and (15) completely determine the probability of soft photon production in the medium of interacting particles. The term containing the function $K_1(k)$ corresponds, as is usual in electrodynamics, to the photon production by a particle in the initial or final state. As for the function $K_2(k)$, it corresponds to photon emission by a particle outside the mass surface in the intermediate state [13, 17] (from the point of view of the Feynman diagrams, this means *photon emission* from the internal line of the corresponding diagram). The factors in brackets in formula (15) correspond to the additional (as compared with the one pointed out in [1–13, 17]) suppression of bremsstrahlung due to the coherent multiple interaction of particles in the medium. These factor are



(16)

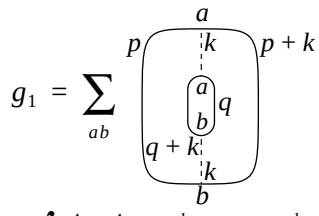
The factors g_1 and g_2 have quite a clear physical sense. They (up to a coefficient) represent the gas parameter [18] determining the significance of the coherent multiparticle interaction in the medium in the problem under consideration.

Note that the diagram summation performed above for the two-particle Green functions does not depend on the nature of the produced particles and can be made in any other situation under the following restrictions: it is necessary that the energy of the produced particles is sufficiently low, $\omega \ll \min\{E, \Delta E\}$, and that perturbation theory for the vertex corresponding to the production of appropriate particles is valid. In particular, this concerns neutrino pair production in strongly interacting media [17]. To apply the formulas obtained to such a problem, it is necessary to replace the photon polarization vector with the weak current of the neutrino pair and to sum over the corresponding states of the produced particles.

It should also be noted that the first nonvanishing approximation in formula (15) is a two-loop one. This means that, under an appropriate choice of the exact one-particle Green functions, one can always automatically ensure conservation of the currents of the corresponding particles.

3. BREMSSTRAHLUNG SPECTRUM RENORMALIZATION IN A DENSE MEDIUM IN THE TWO-LOOP APPROXIMATION

We estimate the factors g_1 and g_2 specified by formulas (16) in the first nonvanishing two-loop approximation. As is shown in [17], the contribution of functions of type K_2 in the soft bremsstrahlung spectrum is equal to zero due to conservation of the vector current. Then, in the two-loop approximation for g_1, we have



$$= -4\int d^4p\, d^4q\, G^{a,b}(p+k)G^{b,a}(p) \times G^{a,b}(q)G^{b,a}(q+k)|V(k)|^2, \quad (17)$$

where a and b take the values + or –, $V(k)$ determines the pair particle interaction, and $G^{a,b}(p)$ are the exact one-particle Green functions.

Let the particle interaction in the medium be such that the collisional breadths of the particles are small as compared with their energies. In this case, for nonrelativistic particles, the one-particle Green functions $G^{+,-}(p)$ and $G^{-,+}(p)$ are proportional [13, 17] to the following expressions:

$$G^{a,b} \propto \frac{\gamma}{[p^0 - \varepsilon(\mathbf{p}) + \mu]^2 + \gamma^2}, \quad (18)$$

where γ is the collisional breadth and μ is the chemical potential.

Then, for the production of extremely long-wave photons, $\omega \ll \gamma$, in an equilibrium medium consisting of nonrelativistic fermions, whose temperature is equal to T, we obtain

$$-g_1 \sim \frac{\sigma_B(k)T^6}{v_F^6 m^2 \gamma^2}, \quad T \ll E_F, \quad (19)$$

$$-g_1 \sim \frac{\sigma_B(k)mT^3}{\gamma^2}, \quad T \gg E_F, \quad (20)$$

where $\sigma_B(k)$ is the Born scattering cross section of particle, m is the mass of a particle of the medium, ω is the photon energy, and E_F and v_F are the energy and the velocity of a particle on the Fermi surface.

When considering the bremsstrahlung of an individual nonrelativistic particle in a scattering medium, the internal loop in relation (17) should be replaced with the factor n/E, where n is the density of scatterers in the

medium and E is the particle energy. As a result, for the factor g_1, we obtain

$$-g_1 \sim \frac{\sigma_B(k) n v_0}{\gamma} \sim \frac{\nu}{\gamma}, \quad (21)$$

where ν is the collision frequency for the particle in the scattering medium and v_0 is the particle velocity. In this case, the parameter $|g_1|$ can be interpreted as the number of scattering centers of the medium simultaneously interacting with the particle when the latter emits a photon. If this quantity is sufficiently large, $\nu \gg \gamma \gg \omega$, the coherent multicenter scattering results in an additional suppression of the bremsstrahlung in the far longwave region of the spectrum.

4. CONCLUSIONS

On the basis of diagram formalism for two-particle Green functions in a nonequilibrium medium, the bremsstrahlung in a substance consisting of interacting particles is investigated. It is shown that the probability of photon production is completely determined by the set of these Green functions. In the case of soft photon bremsstrahlung, the summation of the diagram series for the corresponding two-particle Green functions is performed. It is shown that the probability of photon production is completely determined by a set of irreducible diagrams, starting from two-loop ones, which automatically provides conservation of the current of corresponding particles. These diagrams correspond to the process of photon production, both from the final and initial states of the particles and from intermediate states outside the mass surface. The influence of coherent multiparticle interaction of nonrelativistic particles on the formation of the bremsstrahlung spectrum in a medium is estimated. It is shown that, in the case of a sufficiently dense medium, the coherent effects result in additional suppression (as compared with that indicated earlier in [1–13]) of bremsstrahlung in the longwave region of the spectrum. The possibilities of applying the developed formalism to the investigation of the processes of nonelectromagnetic particle production are discussed.

REFERENCES

1. L. D. Landau and I. Ya. Pomeranchuk, Dokl. Akad. Nauk SSSR **92**, 535 (1953).
2. L. D. Landau and I. Ya. Pomeranchuk, Dokl. Akad. Nauk SSSR **92**, 735 (1953).
3. A. B. Migdal, Dokl. Akad. Nauk SSSR **96**, 49 (1954).
4. A. B. Migdal, Phys. Rev. **103**, 1811 (1956).
5. M. L. Ter-Mikaelyan, Dokl. Akad. Nauk SSSR **94**, 1033 (1954).
6. V. M. Galitskiĭ and V. V. Yakimets, Zh. Éksp. Teor. Fiz. **46**, 1066 (1964) [Sov. Phys. JETP **19**, 723 (1964)].
7. I. I. Gol'dman, Zh. Éksp. Teor. Fiz. **38**, 1866 (1960) [Sov. Phys. JETP **11**, 1341 (1960)].
8. V. E. Pafomov, Zh. Éksp. Teor. Fiz. **49**, 1222 (1965) [Sov. Phys. JETP **22**, 848 (1965)].
9. V. G. Zakharov, Pis'ma Zh. Éksp. Teor. Fiz. **63**, 906 (1996) [JETP Lett. **63**, 952 (1996)].
10. V. G. Zakharov, Pis'ma Zh. Éksp. Teor. Fiz. **64**, 737 (1996) [JETP Lett. **64**, 781 (1996)].
11. R. Baier, Yu. L. Dokshitzer, A. H. Muller, and D. Schiff, Nucl. Phys. B **478**, 577 (1996).
12. S. Klein, Rev. Mod. Phys. **71**, 1501 (1999).
13. J. Knoll and D. N. Voskresensky, Ann. Phys. (N.Y.) **249**, 532 (1996).
14. A. V. Koshelkin, Phys. Lett. B **471**, 202 (1999).
15. L. V. Keldysh, Zh. Éksp. Teor. Fiz. **47**, 1515 (1964) [Sov. Phys. JETP **20**, 1018 (1964)].
16. V. B. Berestetskiĭ, E. M. Lifshitz, and L. P. Pitaevskiĭ, *Quantum Electrodynamics*, 2nd ed. (Nauka, Moscow, 1980; Pergamon Press, Oxford, 1982).
17. A. Sedrakian and A. E. L. Dieperink, Phys. Rev. D **62**, 083002 (2000).
18. E. M. Lifshitz and L. P. Pitaevskiĭ, *Physical Kinetics* (Nauka, Moscow, 1979; Pergamon Press, Oxford, 1981).

Translated by E. Pankratiev

Journal of Experimental and Theoretical Physics, Vol. 96, No. 4, 2003, pp. 658–668.
Translated from Zhurnal Éksperimental'noĭ i Teoreticheskoĭ Fiziki, Vol. 123, No. 4, 2003, pp. 744–756.

ATOMS, SPECTRA, RADIATION

Method for Calculating Photoionization Cross Sections of Fullerenes in the Local Density and Random Phase Approximations

V. K. Ivanov[a], G. Yu. Kashenok[a], R. G. Polozkov[a], and A. V. Solov'yov[b,*]

[a] *St. Petersburg State Technical University, ul. Politekhnicheskaya 29, St. Petersburg, 195251 Russia*
[b] *Ioffe Physicotechnical Institute, Russian Academy of Sciences, ul. Politekhnicheskaya 26, St. Petersburg, 194021 Russia*
**e-mail: solovyov@rpro.ioffe.ru*

Received September 24, 2002

Abstract—A method for calculating the photoionization cross sections of fullerenes taking into account many-electron correlations on the basis of the local density and random phase approximations is proposed and implemented. Calculations are made specifically for fullerenes C_{60} and C_{20}. It is shown that the photoionization spectrum of C_{60} acquires a plasmon resonance whose position and magnitude are in good agreement with experimental results [I.V. Hertel, H. Steger, J. de Vries *et al.*, Phys. Rev. Lett. **68**, 784 (1992)] and with the results of other calculations [M.J. Puska and R.N. Nieminen, Phys. Rev. **A47**, 1181 (1993)]. The emergence of a giant resonance is predicted in the photoionization spectrum of fullerene C_{20} with the center at a photon energy on the order of 27 eV, which corresponds to the frequency of resonant surface plasmon oscillations in a conducting sphere. © *2003 MAIK "Nauka/Interperiodica".*

1. INTRODUCTION

Atomic clusters occupy an intermediate position between individual atoms and a bulk material in respect to the number of particles constituting them. For this reason, they possess the properties typical of individual atoms as well as the bulk material, which renders them extremely interesting objects of investigation. In particular, resonant plasmon oscillations emerging during the excitation of atomic clusters are characteristic primarily of the electron density of conduction electrons in metals. The dipole mode of surface plasmon oscillations is known to be responsible for the formation of a giant resonance in the course of photoexcitation of metallic clusters [1–8]. The excitation of plasmon oscillations during scattering of charged particles from metallic clusters also plays an important role [9–13].

Surface plasmon oscillations in fullerenes have been intensely studied both experimentally [14–16] and theoretically [17–22] during the last decade. These studies were initiated by the experimental work [15] devoted to analysis of electron loss spectra on thin C_{60} films, in which a giant resonance was observed with the center at an excitation energy on the order of 28 eV. These studies stimulated theoretical investigations of collective electron oscillations in isolated fullerenes C_{60} [18]. In the framework of the random phase and tight binding approximations, the emergence of a giant resonance in the optical response function of C_{60} was predicted at an energy on the order of 20 eV. This prediction was confirmed experimentally in [14], where the relative photoionization cross section for isolated fullerenes C_{60} and C_{70} was measured. The measured cross sections exhibited a clearly manifested resonant nature with the resonance center at an energy on the order of 20 eV. In the subsequent series of publications, the optical response of fullerenes C_{60} and C_{70} was studied in detail using various methods based, as a rule, on the basis of the density functional theory (DFT), the time-dependent local density approximation (TDLDA), and the random phase approximation (RPA) [15, 17, 19–22].

In the past several years, considerable attention has been paid to the C_{20} cluster [23–31]. It is known that this cluster has three isomers with the minimal total energy, including an isomer whose structure is close to that of a dodecahedron. It is this isomer that can be regarded as a fullerene with the smallest size. A comparison of numerical results concerning the total energy of different C_{20} isomers can be found in [23–27]. Attempts to study the C_{20} structure experimentally were fruitless for a long time [28, 29]. Nevertheless, promising results concerning the existence and synthesis of fullerene C_{20} have been obtained in [30, 31]. These results served as an impetus to further theoretical investigations of this object. Dynamic parameters of fullerene C_{20} were studied in [31, 32].

In this study, we develop a new method and calculate the electronic structure and the photoionization spectra of fullerenes C_{20} and C_{60} taking many-electron

correlations into consideration. Calculations for C_{60} were made to compare the results with those obtained by other methods. The photoionization cross section for fullerene C_{20} has been calculated for the first time. We also study the role of many-electron correlations in the formation of plasmon resonances in the photoionization spectra of fullerenes.

We use the atomic system of units, where $|e| = \hbar = m_e = 1$. The energy is expressed in rydbergs (Ry).

2. METHODS AND MODELS

2.1. Local Density Approximation

The self-consistent field method, which was proposed by Kohn and Sham [33] in the framework of the density functional theory, has been used more and more often for calculating the electronic structure of multielectron systems. In this method, a system of (generally, integrodifferential) equations of the form

$$\left(\frac{\hat{\mathbf{p}}}{2} + V_{\text{eff}}(\mathbf{r})\right)\phi_i(\mathbf{r}) = \epsilon_i\phi_i(\mathbf{r}) \quad (1)$$

is solved, where $\phi_i(\mathbf{r})$ and ϵ_i are one-electron wave functions and energies.

The total electron density in this case has the form

$$n(\mathbf{r}) = \sum_{i=1}^{N_e} |\phi_i(\mathbf{r})|^2, \quad (2)$$

where N_e is the total number of electrons in the system.

The effective interaction operator in Eqs. (1) has three components:

$$V_{\text{eff}}(\mathbf{r}) = V(\mathbf{r}) + \int d\mathbf{r}' \frac{n(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|} + V_{xc}(\mathbf{r}). \quad (3)$$

Here, V is the operator describing the interaction of electrons with an external field, the second term describes the direct Coulomb interaction, and the third term, describes the exchange-correlation interaction. In an approximate solution of Eqs. (1), we replace exact interaction V_{xc} by an appropriate approximate interaction. The local density approximation (LDA), in which the electron subsystem of atomic clusters in question behaves as a homogeneous electron gas, turned out to be quite useful for practical calculations. We will employ here the widely used Gunnarsson–Lundquist parameterization [34], in which the exchange-correlation potential is represented as

$$V_{xc}^{LDA}(r) = -\left(\frac{9}{4\pi^2}\right)^{1/3}\frac{1}{r_s(r)} - 0.0333\ln\left(1 + \frac{11.4}{r_s(r)}\right), \quad (4)$$

where

$$r_s(r) = \left(\frac{3}{4\pi n(r)}\right)^{1/3}$$

is the Wigner–Seitz radius.

Thus, the exchange-correlation potential is a function of the local value of density at a given point $\mathbf{r}$, which considerably simplifies the numerical procedure of solving Eqs. (1). At first glance, the LDA is applicable only in the case of a spatially homogeneous system. Many years of experience in numerical calculations show, however, that the LDA can be used successfully for calculating strongly inhomogeneous systems such as an atom or a cluster. A considerable disadvantage of the LDA is the presence of self-action in the electron–electron interaction operator $V_{ee}(\mathbf{r})$. This circumstance must be taken into account in the calculation of the electronic structure since it leads to an incorrect asymptotic form of the self-consistent potential.

2.2. Jellium Model for Fullerene

It is generally accepted that fullerenes are formed by fragments of planar graphite lattices [35]. Carbon atoms in graphite are at the vertices of hexagons forming a plane. Different planes are coupled through weak π bonds. Atoms in a plane are connected through σ bonds. These are covalent bonds; i.e., collectivized electrons are concentrated on orbitals spatially fixed relative to the atoms being coupled. Bonds are formed with the participation of 2*s* and 2*p* electrons of carbon atoms (four electrons from each atom). This gives 240 delocalized electrons for fullerene C_{60} and 80 such electrons for fullerene C_{20}. During the formation of bonds, atomic orbitals are hybridized; i.e., the resultant molecular orbital is a combination of atomic *s* and *p* orbitals in various proportions. For example, quantum-chemical calculations [35] show that, in the case of fullerene C_{60}, the σ bond in this molecule is formed via hybridization of atomic sp^2 orbitals. The remaining *p* electron participates in the formation of a π bond. Thus, the ratio of the number of σ and π orbitals in fullerene C_{60} is 3 : 1. Molecular σ orbitals are localized in the radial direction on the fullerene radius, while molecular π orbitals, on the contrary, are directed at right angles to the surface of the imaginary fullerene sphere.

The wave functions of the remaining two 1*s* electrons of carbon atoms are deformed insignificantly and are strongly localized, as before, in the region of location of a given atom. Consequently, it appears natural to divide all the electrons of the system into the valence electrons and the core electrons and consider the self-consistent motion of the valence electrons only.

The potential of the fullerene core has two parts: the potential of interaction between the valence and core electrons and the potential of interaction between the valence electrons and the nuclei of carbon atoms:

$$V_{\text{core}}(\mathbf{r}) = \sum_{\alpha=1}^{N_{\text{at}}}\left\{-\frac{6}{|\mathbf{r}-\mathbf{R}_\alpha|} + 2\int\frac{|\psi_{1s}(\mathbf{r}'-\mathbf{R}_\alpha)|^2}{|\mathbf{r}-\mathbf{r}'+\mathbf{R}_\alpha|}d\mathbf{r}'\right\}. \quad (5)$$

Here, N_{at} is the number of atoms in the fullerene, $\psi_{1s}(\mathbf{r}'-\mathbf{R}_\alpha)$ is the wave function of the 1*s* electron at the

C^{4+} ion, and $\{\mathbf{R}_\alpha\}$ is the set of the coordinates of the atoms. Self-consistent analysis of the wave functions of valence electrons taking into account the actual position of all N_{at} carbon ions, which corresponds to the point symmetry group of the molecule, involves considerable difficulties. For this reason, a simpler but nevertheless adequate model is required to study effectively the dynamic response of fullerenes. For example, here we are using the jellium model of a spherical layer [20] for calculating the electronic structure of fullerenes. The main idea of this approximation is the replacement of the actual core potential (5) by a spherically symmetric potential averaged over positions $\{\mathbf{R}_\alpha\}$. Such a replacement considerably simplifies the numerical procedure of calculating the electronic structure of a fullerene. The accuracy of this approximation obviously increases as the shape of a cluster approaches a spherical shape. The correctness of the spherically symmetric approximation for the core potential of fullerene C_{60} in problems on calculating the dynamic response was demonstrated earlier [20].

The positive nuclear charge of all atoms of a cluster in our model is averaged over a sphere of radius R (which will henceforth be referred to as the fullerene radius). As a result, the potential created by the nuclei is transformed into

$$V_{n-e}(r) = -\frac{6N_{at}}{r_>}, \tag{6}$$

where $r_>$ is larger from r and R.

Earlier [20], the core potential of fullerene C_{60} was replaced by the potential created by the charge distribution of C^{4+} ions averaged over the sphere; i.e., the effect of electrons from inner shells was reduced to simple compensation of the corresponding part of the nuclear charge. Here, apart from screening, we also take into account the spatial extension of the density distribution of 1*s* electrons. Consequently, the contribution of the charge of the 1*s* electron cloud to the core potential is taken into account more accurately.

In accordance with the basic principle of the jellium model (indistinguishability of atoms of the same species in the cluster), we assume that 1*s* electrons of each C^{4+} ion make the same contribution to the total electron density of core electrons. Then we can average the electron density of one of the ions over the sphere and multiply the result by the total number N_{at} of ions:

$$\langle \rho_{1s^2} \rangle(r) = \frac{N_{at}}{4\pi}\int 2|\psi_{1s}(|\mathbf{r}-\mathbf{R}|)|^2 d\Omega. \tag{7}$$

Here, $d\Omega = \sin\theta d\theta d\varphi$ is an element of the solid angle and $\mathbf{R} \equiv (R, 0, 0)$ is a vector in the spherical system of coordinates with the origin at the center of the fullerene. It should be noted that we used in specific calculations the wave functions of 1*s* electrons obtained in the Hartree–Fock approximation for the carbon atom.

The potential created by the averaged electron density,

$$V_{1s^2-e}(r) = \int \frac{\langle \rho_{1s^2} \rangle(r')}{|\mathbf{r}-\mathbf{r}'|} d\mathbf{r}', \tag{8}$$

can be evaluated easily using the expansion in spherical harmonics:

$$V_{1s^2-e}(r) = \int \langle \rho_{1s^2} \rangle(r') \frac{r'^2}{r_>} dr'. \tag{9}$$

Integration with respect to r' is carried out numerically (by the trapezoid method). The total core potential is equal to the sum of two components:

$$V_{core}(r) = V_{n-e}(r) + V_{1s^2-e}(r). \tag{10}$$

In the framework of our model, we use the following standard expansion for the wave functions of valence electrons in fullerene:

$$\phi_i \equiv \phi_{n_i, l_i, m_i}(\mathbf{r}) = \frac{1}{r} P_{n_i l_i}(r) Y_{l_i m_i}(\theta, \varphi). \tag{11}$$

Here, n_i, l_i, and m_i are the principal, orbital, and magnetic quantum numbers of the ith shell, respectively; $Y_{lm}(\theta, \varphi)$ is the spherical harmonic determining the angular dependence of the wave function; and $P_{n_i l_i}(r)$ determines the radial dependence of the wave function and is determined by the specific form of potential $V_{eff}(\mathbf{r})$; then $(n - l - 1)$ determines the number of nodes in the radial wave function. In Eqs. (1), the angular components of the electron wave functions can be separated, which leads to a system of self-consistent differential equations for the radial wave functions, which can be solved numerically.

The wave functions $\phi_i^{N+1}(\mathbf{r})$ of an excited state of an electron in the frozen core approximation are solutions of Eq. (1), i.e., a linear nonself-consistent equation. In this approach, we take into account the fact that an excited electron moves in the field created by all N_e electrons and the self-consistent field at large distance behaves as $(-(Z - N_e)/r)$. Thus, for an electrically neutral system, the $(N + 1)$th electron is in a field whose range is on the order of the size of the system, i.e., the fullerene radius. Among other things, this leads to the absence of discrete excited states for neutral atoms in this approximation. The wave functions $\phi_i^{N+1}(\mathbf{r})$ for $i > F$ and $\phi_i(\mathbf{r})$ for $i \le F$ are solutions of the same system of equations; consequently, they are mutually orthogonal and form a complete set in the sense that

$$\sum_{i \le F} \phi_i^*(\mathbf{r})\phi_i(\mathbf{r}') + \sum_{i > F} \phi_i^{N+1*}(\mathbf{r})\phi_i^{N+1}(\mathbf{r}') = \delta(\mathbf{r}-\mathbf{r}'). \tag{12}$$

The numerical procedure of solving Eqs. (1) for the states of the discrete spectrum is analogous to the procedure of solving Eq. (1) for the wave functions of the

ground state. The solution of Eq. (1) for the states of the continuous spectrum differs in that the electron energy $\epsilon_{s+1} = \kappa^2 > 0$ is given beforehand. Consequently, the problem is no longer an eigenvalue problem.

2.3. Method for Calculating the Photoionization Cross Sections for Fullerenes

The total photoionization cross section is defined as [37]

$$\sigma(\omega) = \sum_{\nu_1\nu_2} \sigma_{\nu_1\nu_2}(\omega), \tag{13}$$

where $\sigma_{\nu_1\nu_2}(\omega)$ is the partial cross section describing the removal of an electron from the $\nu_1 \equiv n_1 l_1$ shell due to its transition to state $\nu_2 \equiv \epsilon l_2$ belonging to the continuum. The partial photoionization cross section is defined as

$$\sigma_{\nu_1\nu_2}(\omega) = \frac{4\pi^2\alpha}{3}\frac{N_{\nu_1}\omega}{2l_{\nu_1}+1}\left|d_{\nu_1\nu_2}\right|^2, \tag{14}$$

where ω is the photon energy and $\epsilon = \epsilon_{nl} + \omega$.

Photoexcitation to a discrete state is described by the oscillator strength

$$f_{\nu_1\nu_2} = \frac{N_{\nu_1}\omega_{\nu_1\nu_2}}{3(2l_{\nu_1}+1)}\left|d_{\nu_1\nu_2}\right|^2, \tag{15}$$

where $d_{\nu_1\nu_2}$ is the reduced dipole matrix element in the LDA approximation or in the LDA-based RPA, $\alpha = e^2/\hbar c = 1/137$ is the fine structure constant, $\omega_{\nu_1\nu_2} = \epsilon - \epsilon_{nl}$ is the photon energy, and N_{ν_1} is the number of electrons in the ν_1 orbital.

The photoionization cross sections and oscillator strengths in the LDA and in the LDA-based RPA are calculated in the form of length and velocity. In the second case, the photoionization cross sections in the form of length and velocity must coincide if the effect of all the shells in a cluster is taken into account. For this reason, the difference in these representations can be used for estimating the total numerical error in the obtained results as well as the effect of transitions from other shells except that being considered.

Both in the one-particle approximation and in the random phase approximation, cross sections and oscillator strength must obey the sum rule:

$$N = \sum_{nl}\sum_{n',l'} f_{nl,n'l'} + \frac{1}{2\pi^2\alpha}\int_I^\infty \sigma_{nl}(\omega)d\omega, \tag{16}$$

where N is the total number of valence electrons and $\{nl\}$ is the set of quantum numbers of the corresponding orbital.

Here, we use the local density approximation as the one-particle approximation; for this reason, dipole matrix elements are calculated using one-particle wave functions determined from Eqs. (1) for the ground state and in the frozen core approximation for excited states.

In order to calculate the photoionization amplitude taking into account many-electron correlations $\hat{D}(\omega)$, we used the random phase approximation. The RPA equation can be represented symbolically in the following form:

$$\hat{D}(\omega) = \hat{d} + \hat{D}(\omega)\hat{\chi}(\omega)\hat{V}. \tag{17}$$

In specific calculations, Eq. (17) is solved numerically. For this purpose, use should be made of the matrix form of Eq. (17), expressing the matrix elements of the photon absorption amplitude $\langle k_2|\hat{D}(\omega)|k_1\rangle$ in terms of their values $\langle k_2|\hat{d}|k_1\rangle$ in the one-particle approximation, the matrix elements of the operator $\hat{\chi}(\omega)$ describing the propagation of an electron-hole excitation in time, and the matrix elements $\langle k_2 k_1'|\hat{V}|k_1 k_2'\rangle$ of the electron–electron interaction [40]. It should be noted that, since we are using the local density approximation in the framework of the density functional theory as the one-particle approximation, the kernel of the electron–electron interaction V must be taken [38, 39] in the form

$$V^{DFT}(\mathbf{r},\mathbf{r}') = \frac{1}{|\mathbf{r}-\mathbf{r}'|} + \left.\frac{\delta^2 E_{xc}[n]}{\delta n(\mathbf{r})\delta n(\mathbf{r}')}\right|_{n=n_0(\mathbf{r})}, \tag{18}$$

where $E_{xc}[n]$ is the exchange-correlation functional. In the LDA [34], we have

$$\left.\frac{\delta^2 E_{xc}[n]}{\delta n(\mathbf{r})\delta n(\mathbf{r}')}\right|_{n=n_0(\mathbf{r})} = \left.\frac{\partial^2 \epsilon_{xc}^{\text{hom}}[n]}{\partial n^2}\right|_{n=n_0(\mathbf{r})}\delta(\mathbf{r}-\mathbf{r}'). \tag{19}$$

Here, $\epsilon_{xc}^{\text{hom}}[n]$ is the exchange-correlation energy of a homogeneous electron gas of density n. In our calculations, we used for $\epsilon_{xc}^{\text{hom}}[n]$ the widely used Gunnarsson–Lundquist approximation [34].

The numerical procedure of solving Eq. (17) used here is completely analogous to the procedure described in monograph [40]. However, a peculiar feature of our approach is that the matrix elements in Eq. (17) are calculated with one-particle wave functions determined from Eqs. (1) for the ground state and in the frozen core approximation for excited states, and not on the basis of the Hartree–Fock approximation. For this reason, it is appropriate to refer to the computation method as the LDA-based RPA. In the framework of the LDA-based RPA, the photoionization cross section is calculated taking into account the intra- and intershell correlations [37]. A transition is characterized by the definition of the $\{n_1 l_1\}$ vacancy and the

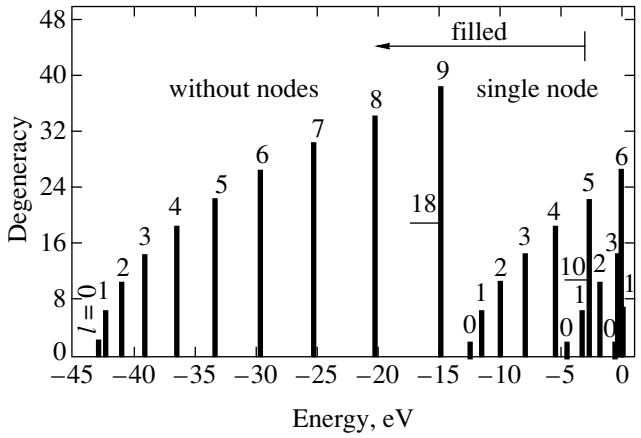


Fig. 1. Schematic diagram of the arrangement and values of one-electron energies of valence electrons in C_{60}. The figures on top of vertical segments indicate the values of the angular momentum l of the state. Numbers 18 (for the outer zero-node state) and 10 (for the outer single-node state) indicate the occupational numbers for incompletely filled states. The remaining states are filled completely.

excited $\{n_2(\epsilon_2)l_2\}$ state (n_2 for the discrete spectrum and ϵ_2 for the continuum).

3. RESULTS OF CALCULATION OF PHOTOIONIZATION CROSS SECTIONS FOR C_{60} AND C_{20}

3.1. Electronic Structure of C_{60} and C_{20}

Martins *et al.* [36] proposed a correlation that should be established between wave functions with a certain number of nodes in the radial part of the wave function and σ and π orbitals of planar graphite lattices. The correlation is quite simple: functions without nodes, which are strongly localized in the radial direction on the fullerene radius, correspond to σ orbitals, while functions with a single node correspond to π orbitals. Here, we choose the electronic configurations of the ground state of fullerenes using this correlation. Namely, since the ratio of the number of σ and π orbitals for the ground state of C_{60} is 3 : 1 [35], the ratio of the number of electrons located in node-free and single-node shells $\{n, l\}$ is taken in the same proportion. This gives 180 electrons in zero-node shells and 60 electrons in shells with a single node in the radial wave function:

$$\underbrace{1s^2 2p^6 3d^{10} 4f^{14} 5g^{18} 6h^{22} 7i^{26} 8k^{30} 9l^{34} 10m^{18}}_{\text{zero nodes, } 180e^-}$$

$$\underbrace{2s^2 3p^6 4d^{10} 5f^{14} 6g^{18} 7h^{10}}_{\text{single node, } 60e^-}$$

(here and below, the atomic scheme nl_e^N is used for the electron configuration notation [41]). For fullerene C_{20}, the number of orbitals that can be characterized unambiguously as σ and π orbitals is also known: $n_\sigma = 30$ and $n_\pi = 8$. Since each of these orbital is doubly degenerate in spin, this give 76 electrons in all. The remaining two molecular orbitals cannot be classified unambiguously as σ or π. Analogously to C_{60}, σ and π orbitals are put in correspondence with zero- and single-node wave functions. The remaining electrons are arranged in accordance with the principle of minimum total electron energy. Thus, the electronic configuration of the ground state of fullerene C_{20} has the form

$$\underbrace{1s^2 2p^6 3d^{10} 4f^{14} 5g^{18} 6h^{14\,=\,10\,+\,4}}_{\text{zero nodes, } (60\,+\,4)e^-}\ \underbrace{2s^2 3p^6 4d^8}_{\text{single node, } 16e^-}$$

It should be noted that, in the above electronic configurations, outer zero-node and single-node shells are filled only partially. Since it is impossible to construct the entire system of terms for a given electronic configuration for such a large number of electrons, we used the averaged term approximation. In this approximation, the electron charge is smeared uniformly over an orbital with certain $\{n, l\}$. In other words, the maximal occupancy $2(2l + 1)$ of the $\{n, l\}$ orbital is in fact replaced by the actual occupancy N_{nl} for the given orbital. In this case, the electron density of valence electrons can be represented in the form

$$n(r) = \sum_{nl} \frac{N_{nl}}{4\pi}\left(\frac{1}{r}P_{nl}(r)\right)^2. \qquad (20)$$

The arrangement and the values of one-particle energies of the ground state of the valence electrons in C_{60} are given in Fig. 1. Figure 2 shows the electron density and the self-consistent potential of C_{60}, obtained in the local density approximation.

The self-consistent potential decreases over large distances at a higher rate than the Coulomb potential. This is due to an irregular asymptotic form of the self-consistent potential in the LDA since each electron moves in the field of all 240 electrons, i.e., in the field of a neutral system. In such cases, the wave functions are said to be calculated in the $(N + 1)$ basis. As a result, the energy levels have values smaller than for a potential with a regular (Coulomb) asymptotic form. Indeed, a decrease in the potential well width for the same depth leads to the upward expulsion of one-particle levels. In particular, this leads to a decreased value of the ionization potential $I_p = 2.88$ eV, considerably lower than the experimentally obtained value $I_p^{\text{exp}} = 7.6$ eV [14]. This is correct from the physical point of view: each electron actually moves in the field of the neutral system and, hence, is bound less tightly. In order to verify the applicability of the scheme with the $(N + 1)$ basis, we calculated the electronic structure of fullerene C_{60} in

the LDA with partly eliminated self-action. For this purpose, we represented Eqs. (1) in the form

$$\left(\frac{\hat{\mathbf{p}}}{2} + V_{\text{core}}(r) + \sum_{j=1}^{s}\int\frac{N_j|\phi_j(\mathbf{r}')|^2}{|\mathbf{r}-\mathbf{r}'|}d\mathbf{r}' + V_{xc}(n(r))\right) \times \phi_i(\mathbf{r}) = \epsilon_i\phi_i(\mathbf{r}). \quad (21)$$

Here, as before, we have

$$i \equiv \{nl\}, \quad n(r) = \sum_{i=1}^{s} N_i|\phi_i(\mathbf{r})|^2.$$

In order to eliminate self-action, the occupation number for the shell being calculated is reduced by unity ($N_i \longrightarrow N_i - 1$):

$$\left(\frac{\hat{\mathbf{p}}}{2} + V_{\text{core}}(r) + \sum_{j\neq i}^{s}\int\frac{N_j|\phi_j(\mathbf{r}')|^2}{|\mathbf{r}-\mathbf{r}'|}d\mathbf{r}' + V_{xc}(n^*(r))\right)\phi_i(\mathbf{r}) + \int\frac{(N_i-1)|\phi_i(\mathbf{r}')|^2}{|\mathbf{r}-\mathbf{r}'|}d\mathbf{r}'\phi_i(\mathbf{r}) = \epsilon_i\phi_i(\mathbf{r}). \quad (22)$$

Here,

$$n^*(r) = \sum_{j\neq i}^{s} N_j|\phi_j(\mathbf{r})|^2 + (N_i-1)|\phi_i(\mathbf{r})|^2. \quad (23)$$

Thus, each electron is in the field of 239 electrons or, in other words, wave functions are calculated in the N basis. It can easily be seen that each ith shell will have its own potential. This means, first, that the obtained orbitals are not orthogonal and, hence, do not form a complete set. Consequently, they cannot be used for calculating the dynamic response of the system without additional orthogonalization. Second, the calculations made by using this scheme showed that the level structure remained unchanged although the value of the ionization potential (I_p^N = 5.52 eV) becomes closer to the correct value of I_p^{exp} [14]. The level structure includes their mutual arrangement and the spacing between the levels. This means that, in fact, the entire electronic structure has shifted downwards approximately by the difference $I_p^{\text{exp}} - I_p$. From the physical point of view, this means that, since the number of electrons in fullerene shells is quite large, the removal of one electron from the total electron density does not change its value significantly in the surface region of fullerene. For this reason, the form of the potential varies insignificantly over distances on the order of the fullerene radius. On the basis of the above analysis, we decided to retain a more convenient and consistent (N + 1) scheme for calculating wave functions, using in subsequent calculations the values of one-electron energies shifted by $I_p^{\text{exp}} - I_p$ on the energy scale towards higher values.

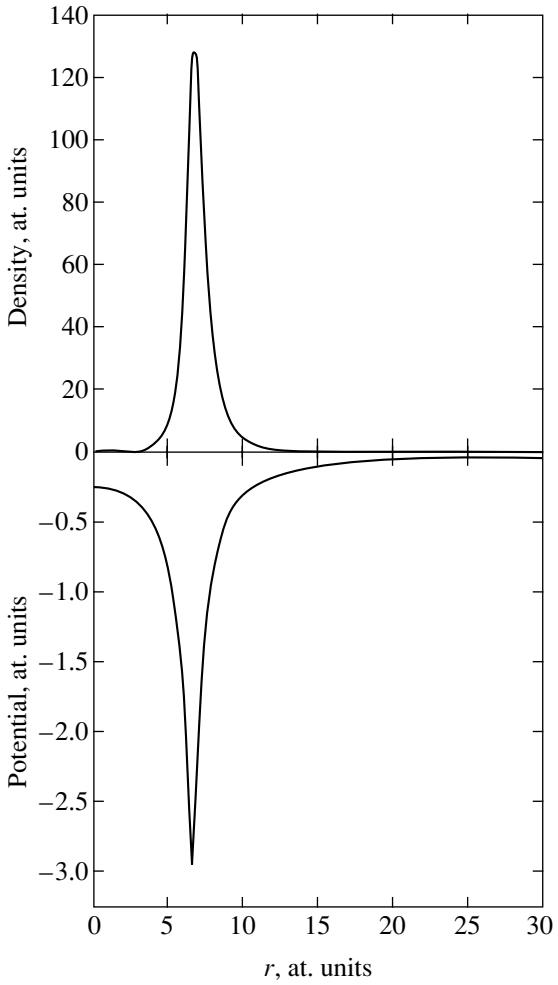


Fig. 2. Self-consistent potential of C_{60} and the total electron density of valence electrons, calculated in the local density approximation.

Since the results of calculations of the electronic structure of C_{60} in the framework of our model are in satisfactory agreement with the results obtained by other authors [17, 20, 21], we decided to use this approximation for calculating the electronic structure of a new object, viz., fullerene C_{20}.

The calculated energy spectrum is shown in Fig. 3. Zero-node orbitals are filled up to l = 5, while single-node orbitals are filled up to l = 2. The self-consistent potential and the electron density are depicted in Fig. 4. As in the case of C_{60}, the electron density is strongly localized and has a peak at the fullerene radius. Accordingly, the self-consistent potential has a sharp minimum at the surface of the imaginary sphere into which fullerene is inscribed. The values of one-particle energies are found to be slightly exaggerated. In particular,

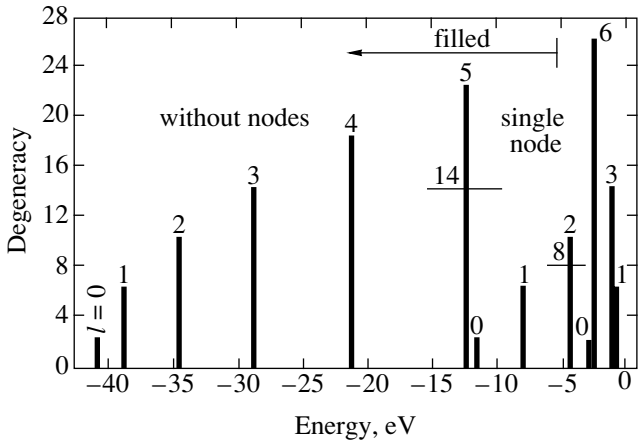


Fig. 3. Schematic diagram of the arrangement and values of one-electron energies of valence electrons in C_{20}. The figures on top of vertical segments indicate the values of the angular momentum l of the state. Numbers 14 (for the outer zero-node state) and 8 (for the outer single-node state) indicate occupational numbers for incompletely filled states. The remaining states are filled completely.

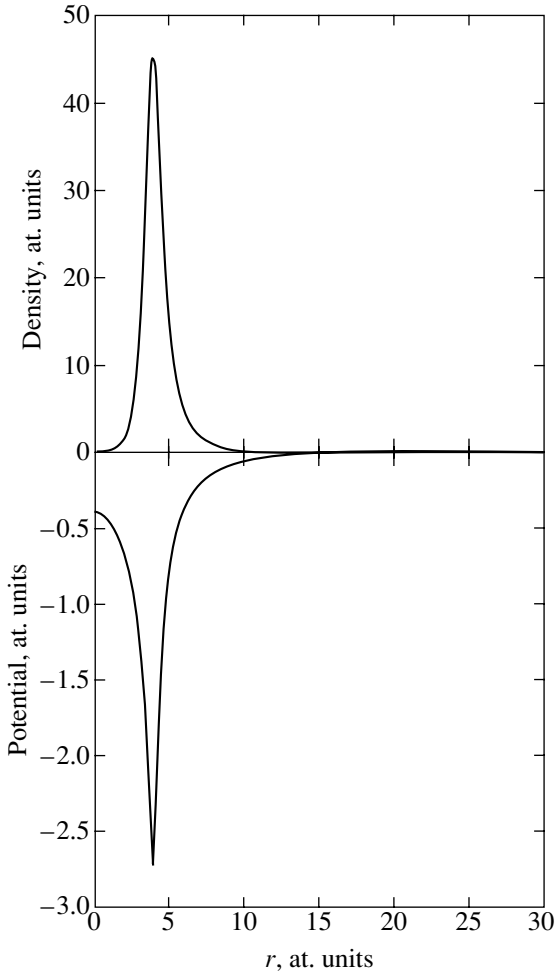


Fig. 4. Self-consistent potential of C_{20} and the total electron density of valence electrons, calculated in the local density approximation.

the value of ionization potential $I_p = 4.362$ eV is slightly lower than the value $I_p^{MO} = 6.939$ eV obtained from quantum-chemical calculations [32].

3.2. Photoionization Spectra of Fullerenes C_{60} and C_{20}

Let us consider the results of calculations of the photoionization cross section for fullerene C_{60}. The complete electron spectrum of a C_{60} molecule below and above the Fermi level is represented in Fig. 1. Excited electron states were chosen in accordance with the selection rules for dipole transitions $l \longrightarrow l \pm 1$. The wave functions of excited states were determined in the frozen core approximation. It was mentioned above that an excited electron in this approximation is in a short-range potential, in which only a limited number of discrete excited state can exist. The following discrete levels were determined: 3*s*, 4*p*, 5*d*, 4*s*, 6*f*, 8*i*, and 5*p*. It can be seen from the figure that two discrete excited states 3*s* and 4*p* "fall" below the Fermi level. This can be attributed primarily to the limited validity of the assumption concerning the spherical symmetry of the C_{60} core. It was shown by Yabana and Bertsch [20] that allowance for exact (icosahedral) symmetry of the core shifts these levels to a region above the Fermi level. These authors predicted that taking into account the exact symmetry of the core leads to a small correction to the pseudopotential describing the interaction of valence electrons with the core, which can be included in the framework of perturbation theory.

We analyzed possible transitions starting from the outermost 7*h* shell down to the 7*i* shell, which amounts to 19 transitions in all. The transitions map is as follows:

$$7h \longrightarrow (8i, Eg); \quad 6g \longrightarrow (7h, 6f); \quad 5f \longrightarrow (Eg, 5d);$$

$$4d \longrightarrow (6f, 4p(5p)); \quad 3p \longrightarrow (5d, 3s(4s));$$

$$2s \longrightarrow (4p(5p)); \; 10m \longrightarrow (En, El); \; 9l \longrightarrow (10m, Ek);$$

$$8k \longrightarrow (El, 8i); \; 7i \longrightarrow (Ek, 7h).$$

It should be noted that, since the outer zero-node 10*m* and single-node 7*h* shells are not filled completely, the transitions in these shells are allowed. Henceforth, we will refer to transition from σ and π orbitals as the σ and the π transitions, respectively.

Partial cross sections in the LDA-based RPA have a clearly manifested resonance form. The main contribution to the formation of a giant resonance comes from transitions from outer π orbitals. By way of example, Fig. 5 shows partial photoionization cross sections for the 6*g* shell of fullerene C_{60}. The shapes of the partial cross sections in the one-particle and random-phase approximations differ significantly. Moreover, a considerable redistribution of oscillator strengths takes place. In the one-particle approximation, the main contribution to the sum rule comes from oscillator

strengths (the contribution from the continuous spectrum is 22% and that from the oscillator strengths is 78%), while the main contribution in the LDA-based RPA (75%) comes from the region of the continuum. This indicates a significant role of many-electron correlations in the formation of the spectrum.

Figure 6 shows the total photoionization cross section of fullerene C_{60} taking into account many-electron correlations, which was obtained by the LDA-based RPA method, as well as the experimental cross section obtained in [14] and normalized in accordance with [16]. The cross section in the LDA-based RPA is shifted by $I_p^{\text{exp}} - I_p^{\text{our}}$ on the energy scale towards higher photon energies. The main features of the photoionization cross section are two resonances: a near-threshold resonance and a giant resonance centered approximately at 20 eV with a width of about 10 eV. A comparison of the theoretically predicted and experimental cross sections shows that the position and height of the giant resonance are reproduced correctly in our model. Figure 6 shows, however, that the theoretical and experimental cross sections in the near-threshold region differ qualitatively: the theory predicts a narrow and high resonance near the threshold. The conditions for its emergence and origin will be discussed below.

Similar calculations were made for fullerene C_{20}. The complete electron spectrum of a C_{20} molecule below and above the Fermi level is shown in Fig. 3. As in the case of C_{60}, only a few discrete levels were determined: 3*s*, 7*i*, 5*f*, and 4*p*. We analyzed all possible transitions starting from the outermost 4*d* shell down to the 3*d* shell. The transition map has the form

$$4d \longrightarrow (5f, 4p); \quad 3p \longrightarrow (4d, 3s); \quad 2s \longrightarrow (4p);$$
$$6h \longrightarrow (7i, 6g); \quad 5g \longrightarrow (6h, 5f); \quad 4f \longrightarrow (6g, 4d);$$
$$3d \longrightarrow (5f, 4p).$$

Since the outer zero-node 6*h* and single-node 4*d* shells are filled incompletely, transitions in these shells are allowed.

Figure 7 shows the total photoionization cross section for fullerene C_{20} taking into account many-electron correlations, which was obtained by the LDA-based RPA method. It can be seen that the photoionization cross section of a C_{20} molecule, calculated taking into account many-electron correlations, has a clearly manifested resonance shape. As in the case of C_{60}, the main features of the photoionization cross section are two resonances, one of which is near the threshold and the other is centered approximately at 27 eV and has a width of about 7 eV. The conditions for their emergence and origin will be discussed below.

3.3. Resonances in the Spectra of C_{60} and C_{20}

The photoionization spectra of C_{60} and C_{20} fullerene molecules have a clearly manifested resonance shape.

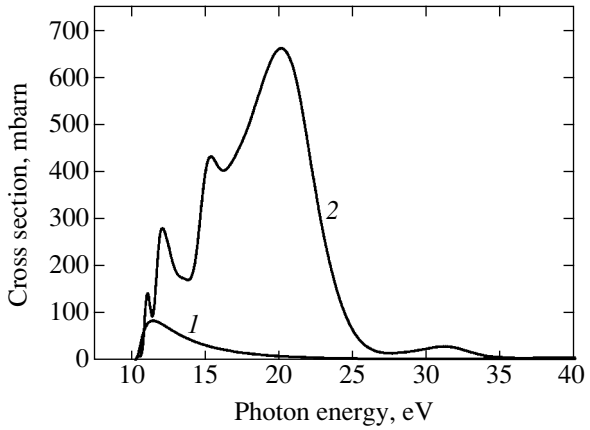


Fig. 5. Partial photoionization cross sections of the 6*g* shell of fullerene C_{60} calculated in the one-particle approximation (LDA) (*1*) and taking into account many-electron correlations (RPA and LDA).

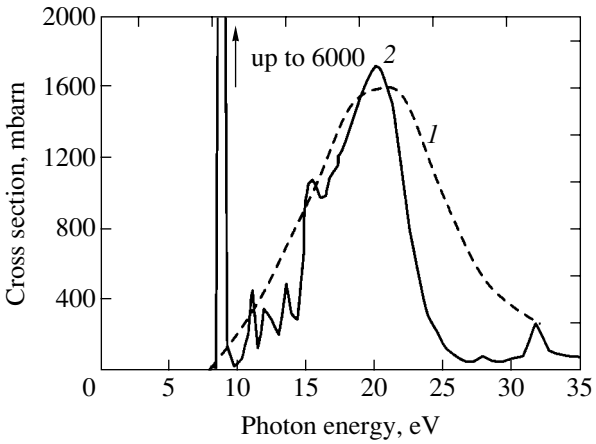


Fig. 6. Total photoionization cross section of C_{60}: experimental data [14] (*1*) and calculation in the random phase and local density approximations (*2*); the theoretical cross section is shifted towards higher values of photon energy by $\Delta = I_p^{\text{exp}} - I_p^{\text{our}}$ [14] (explanation is given in the text).

The conditions for the emergence of resonances in the photoionization spectra of two molecules and their physical origins are identical.

The near-threshold resonances in the photoionization spectra are completely determined by transitions from outer π orbitals. The oscillator strengths of π–π transitions to discrete excited states are not high in the one-particle approximation, while narrow near-threshold resonances appear in the continuous spectrum calculated in the LDA-based RPA. Thus, the oscillator strengths are redistributed when many-electron correlations are taken into account. The small width of near-threshold resonances indicates, from the physical point of view, that a remote electron with a low energy remains in the vicinity of fullerene for a long time. Since the excited wave functions were calculated in the

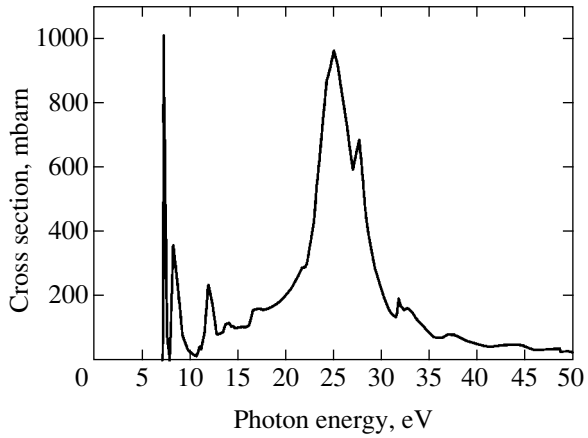


Fig. 7. Total photoionization cross section of C_{20} calculated in the random phase and local density approximations; the theoretical cross section is shifted towards higher values of photon energy by $\Delta = I_p^{\text{quant chem}} - I_p^{\text{our}}$ [32].

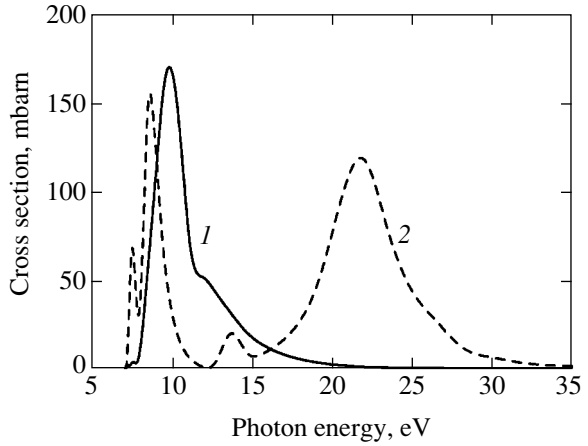


Fig. 8. Partial cross section of transition $4d \longrightarrow \epsilon p$ in C_{20} calculated in the RPA and LDA with correlations between transitions from π orbitals (*1*) and with correlations between transitions from π orbitals and the transition from the outer σ orbital (*2*).

(N + 1) basis, the remote electron actually corresponds to the added (N + 1)th electron. Therefore, we can draw an analogy between the existence of metastable states of negative fullerene ions and discrete excited states in our approximation. It is known [42] that a large number of metastable states exist for C_{60}^-. The existence of metastable states for C_{20}^-, which are manifested during inelastic scattering of slow electrons in fullerene C_{20}, was also predicted [32]. The positions of near-threshold resonances in the spectra of C_{60} and C_{20}, which were obtained in our analysis, are in good agreement with the energies of these metastable states. It should also be noted that the cross sections near the threshold are exaggerated considerably in the RPA due to limitations of the method. Indeed, a photoelectron in this case possesses a relatively low energy, and the core relaxation effects, which are beyond the RPA, must be taken into account.

The second resonance in the photoionization spectra of C_{20} and C_{60} can be classified as a collective, or giant resonance. It emerges in partial photoionization cross sections in the random phase approximation taking the electron–electron correlations into account and is absent in the one-particle spectrum. Consequently, we can state that the emergence of a giant resonance is a manifestation of collective modes of cluster excitations. Numerical calculations proved that it emerges due to the interaction between π and σ transitions. This is illustrated in Fig. 8: the interaction between π transitions does not lead to a giant resonance in the partial cross section, and only the actuation of the channel with a σ transition leads to its emergence. This result can be explained by analyzing the process of collective excitations in terms of oscillations of the electron density corresponding to the spatial configuration. Indeed, σ orbitals are presented by zero-node wave functions in the radial direction, which have a peak at the fullerene surface and are strongly localized near this surface. Consequently, the excitation of the electron density of σ orbitals has the form of electron density oscillations in a thin surface region, i.e., surface plasmon vibrations.

The formula for the resonance frequency of electron density oscillations in a conducting sphere with permittivity $\epsilon(\omega) = 1 - \omega_p^2/\omega^2$ (ω_p is the frequency of plasma oscillations in a bulk material) was derived in a number of publications in the framework of classical electrodynamics [43] and in the theory of linear response of multielectron systems [11]:

$$\omega_l = \sqrt{\frac{l(l+1)N_e}{(2l+1)R_{ful}^3}} \qquad (24)$$

(ω_l is the frequency of surface plasmon vibrations with angular momentum l and N_e is the number of delocalized electrons). For fullerene C_{60}, we have R = 6.6624 at. units [35] and N_e = 240; for l = 1 (which corresponds to the excitation of dipole plasmon vibrations), we have ω_1 = 20 eV. It can be seen from Fig. 7 that the position of the giant resonance in our calculations corresponds to the position of the dipole surface plasmon resonance. Similarly, for fullerene C_{20}, we have R = 3.86 at. units [32] and N_e = 80; for l = 1 (which corresponds to excitation of dipole plasmon vibrations), we have ω_1 = 27 eV. It can be seen from Fig. 7 that the position of the giant resonance in our calculations corresponds to the position of the dipole surface plasmon resonance.

4. CONCLUSIONS

The above analysis proved that many-electron correlations play a decisive role in the formation of the photoionization spectra of fullerenes C_{20} and C_{60}. The

photoionization cross sections of C_{20} and C_{60}, which are calculated by using a method combining the random phase and local density approximations, display giant resonances of the same physical origin. These resonances appear as a result of collective excitations of the electron density, viz., surface plasmon vibrations. The method for calculating the electronic structure of fullerenes based on the local density approximation in the framework of the spherical jellium model of the layer makes it possible to obtain self-consistent potentials of C_{20} and C_{60}, displaying a sharp minimum in the vicinity of the fullerene surface.

The calculations made for fullerene C_{60} were of a test nature since the number of theoretical and experimental publications concerning the optical response of this object is sufficiently large [14, 16–22]. The criterion for verification was the comparison of the theoretically predicted cross section with the experimental cross section measured in [14]. The position and height of the giant resonance in the photoionization spectrum of C_{60} obtained by us is in good agreement with the result obtained in the experimental work [14] and normalized in accordance with [16]. This enables us to state that the proposed method makes it possible to calculate the photoionization cross section to a fairly high degree of accuracy (at least, in the resonance region). For this reason, this method of calculation was applied to a new and scarcely investigated object, viz., fullerene C_{20}.

This method was used for the first time to calculate the photoionization cross section for fullerene C_{20} in a wide photon energy range taking into account many-electron correlations. The cross section was obtained with allowance for correlations in all transitions. The absolute values of photoionization cross sections obtained in the framework of this method are of considerable importance for experimental studies.

It was found that the photoionization spectrum of fullerene C_{20} displays a giant resonance. The conditions for the emergence of the resonance were investigated and the mechanism responsible for its emergence was identified with the excitation of surface plasmon vibrations. The decisive role of collective excitations in the formation of the photoionization spectrum of C_{20} was established.

It should be noted that our method of calculations gives too high values of photoionization cross sections in the region of the cluster ionization threshold in view of limitations of the random phase approximation. However, despite this fact, the model developed by us can be applied for studying a wide class of collision processes involving fullerenes and spherical metallic clusters.

ACKNOWLEDGMENTS

This study was supported financially by the Russian Foundation for Basic Research (grant no. 03-02-26026), the Ministry of Higher Education (grant no. E00-3.4-549), INTAS (grant no. 97-603) and the Sixth Competition of Scientific Projects by Young Scientists, Russian Academy of Sciences (1999) (grant no. 44).

REFERENCES

1. W. A. De Heer, Rev. Mod. Phys. **65**, 611 (1993).
2. M. Brack, Rev. Mod. Phys. **65**, 677 (1993).
3. C. Brechignac and J. P. Connerade, J. Phys. B: At. Mol. Opt. Phys. **27**, 3795 (1994).
4. *Clusters of Atoms and Molecules, Theory, Experiment and Clusters of Atoms*, Ed. by H. Haberland (Springer, Berlin, 1994), Springer Ser. Chem. Phys. **52**.
5. F. Alasia, R. A. Broglia, H. E. Roman, *et al.*, J. Phys. B: At. Mol. Opt. Phys. **27**, L663 (1994).
6. M. Madjet, C. Guet, and W. R. Johnson, Phys. Rev. A **51**, 1327 (1995).
7. U. Kreibig and M. Vollmer, *Optical Properties of Metal Clusters* (Springer, Berlin, 1995), p. 328.
8. A. V. Korol and A. V. Solov'yov, J. Phys. B: At. Mol. Opt. Phys. **30**, 1105 (1997).
9. W. Ekardt, Phys. Rev. Lett. **52**, 1925 (1984).
10. W. Ekardt, Phys. Rev. B **33**, 3702 (1986).
11. L. G. Gerchikov, A. V. Solov'yov, J. P. Connerade, and W. Greiner, J. Phys. B: At. Mol. Opt. Phys. **30**, 4133 (1997).
12. L. G. Gerchikov, A. N. Ipatov, and A. V. Solov'yov, J. Phys. B: At. Mol. Opt. Phys. **30**, 5939 (1997).
13. L. G. Gerchikov, A. N. Ipatov, A. V. Solov'yov, and W. Greiner, J. Phys. B: At. Mol. Opt. Phys. **31**, 3065 (1998).
14. I. V. Hertel, H. Steger, J. de Vries, *et al.*, Phys. Rev. Lett. **68**, 784 (1992).
15. J. H. Weaver, J. L. Martins, T. Komeda, *et al.*, Phys. Rev. Lett. **66**, 1741 (1991).
16. T. Liebsch, O. Plotzke, F. Heiser, *et al.*, Phys. Rev. A **52**, 457 (1995).
17. M. J. Puska and R. M. Nieminen, Phys. Rev. A **47**, 1181 (1993).
18. G. F. Bertsch, A. Bulgac, D. Tomanek, *et al.*, Phys. Rev. Lett. **67**, 2690 (1991).
19. N. Ju, A. Bulgac, and J. W. Keller, Phys. Rev. B **48**, 9071 (1993).
20. K. Yabana and G. F. Bertsch, Phys. Scr. **48**, 633 (1993).
21. G. Wendin and B. Wastberg, Phys. Rev. B **48**, 14764 (1993).
22. F. Alasia, R. A. Broglia, H. E. Roman, *et al.*, J. Phys. B: At. Mol. Opt. Phys. **27**, L643 (1994).
23. V. Parasuk and G. Almlöf, Chem. Phys. Lett. **184**, 187 (1991).
24. M. C. Domene, P. W. Fowler, P. A. Madden, *et al.*, Chem. Phys. Lett. **314**, 158 (1999).
25. R. O. Jones, J. Chem. Phys. **110**, 5189 (1999).
26. S. Sokolova, A. Luchow, and J. B. Anderson, Chem. Phys. Lett. **323**, 229 (2000).
27. J. M. L. Martin, J. El-Yazal, and J.-P. Francois, Chem. Phys. Lett. **248**, 345 (1996).
28. G. von Helden, M. T. Hsu, N. G. Gotts, *et al.*, Chem. Phys. Lett. **204**, 15 (1993).

29. C. J. Brabec, E. B. Anderson, B. N. Davidson, *et al.*, Phys. Rev. B **46**, 7326 (1992).
30. K. Hata, M. Ariff, K. Tohij, *et al.*, Chem. Phys. Lett. **308**, 343 (1999).
31. H. Prizbach, A. Weller, P. Landenberger, *et al.*, Nature **407**, 60 (2000).
32. F. A. Gianturco, G. Yu. Kashenock, R. R. Lucchese, *et al.*, J. Chem. Phys. **116**, 2811 (2002).
33. W. Kohn and L. J. Sham, Phys. Rev. **140**, A1133 (1965).
34. O. Gunnarsson and B. I. Lundqvist, Phys. Rev. B **13**, 4274 (1976).
35. R. C. Haddon, L. E. Brus, and K. Raghavachari, Chem. Phys. Lett. **125**, 459 (1986).
36. J. L. Martins, N. Troullier, and J. H. Weaver, Chem. Phys. Lett. **180**, 457 (1991).
37. M. Ya. Amus'ya and L. V. Chernysheva, *Automatized System for Study of Atoms* (Nauka, Leningrad, 1983).
38. M. J. Stott and E. Zaremba, Phys. Rev. A **21**, 12 (1980).
39. A. Zangwill and P. Soven, Phys. Rev. A **21**, 1561 (1980).
40. M. Ya. Amus'ya, N. A. Cherepkov, and L. V. Chernysheva, Zh. Éksp. Teor. Fiz. **60**, 160 (1971) [Sov. Phys. JETP **33**, 90 (1971)].
41. L. D. Landau and E. M. Lifshitz, *Course of Theoretical Physics*, Vol. 3: *Quantum Mechanics: Non-Relativistic Theory*, 4th ed. (Nauka, Moscow, 1989; Pergamon, New York, 1977).
42. O. Elhamidi, J. Pommier, and R. Abouaf, J. Phys. B: At. Mol. Opt. Phys. **30**, 4633 (1997).
43. D. A. Gorokhov, R. A. Suris, and V. V. Cheainov, Phys. Lett. A **223**, 116 (1996).

Translated by N. Wadhwa

Journal of Experimental and Theoretical Physics, Vol. 96, No. 4, 2003, pp. 669–673.
Translated from Zhurnal Éksperimental'noĭ i Teoreticheskoĭ Fiziki, Vol. 123, No. 4, 2003, pp. 757–762.

ATOMS, SPECTRA, RADIATION

Effect of Electric Field on Photoproduction of Electron–Positron Pairs

A. E. Lobanov[a,*] and A. R. Muratov[b]

[a]*Moscow State University, Vorob'evy gory, Moscow, 119992 Russia*
[b]*Institute of Oil and Gas Problems, Russian Academy of Sciences, Moscow, 117701 Russia*
**e-mail: th180@phys.msu.su*

Received October 18, 2002

Abstract—The cross section of electron–positron pair production by two circularly polarized photons in a static uniform electric field is calculated. The dependence of the cross section on the photon energy is determined. It is shown that the corrections to the cross section of the process involving photons of various polarizations are suppressed as compared to the situation with crossed fields. *© 2003 MAIK "Nauka/Interperiodica".*

The study of the production of electron–positron pairs by photons started in the classical work by Breit and Wheeler [1]. This phenomenon, which is of considerable importance for constructing astrophysical models, is treated in many theoretical works in spite of the fact that it was inaccessible to experimental investigation under laboratory conditions for a long time. For this reason, the main regularities of the process in question have been established only theoretically. In particular, expressions for the multiphoton process cross section [2–11], as well as the formulas reflecting the effect of external fields on two-photon [12–17] and multiphoton [18–22] pair production, have been obtained.[1] Naturally, after the observation of multiphoton pair production [24] using high-energy photons generated in the inverse Compton effect [25],[2] theoretical investigations were continued in [27–31].

In this paper, we analyze the effect of purely electric field on photoproduction of pairs. Namely, we consider the production of electron–positron pairs by two circularly polarized photons propagating toward each other along a static uniform electric field **E**. The choice of the problem geometry is dictated by the following considerations. If the energies of the photons differ significantly (which is the most interesting case), the field is close to that of a crossed field for photons propagating at an angle to **E** in the center-of-mass system of the pair being formed; such a configuration was studied, for example, in [16]. All the results obtained here are based on completely relativistic calculations, which enables us to take into account the effect of the photon spin orientation relative to the external field on the cross sections and to find the energy dependence of the cross sections for any (and not only at threshold) values. Consequently, this paper, together with [16] and publication [15], in which a pure magnetic field is considered, gives a complete qualitative pattern of the effect of static uniform fields on the pair production by hard polarized γ quanta at laser beam photons.

The integral representation for the pair production cross section σ in the lowest order of perturbation theory in quantized field can be determined using the results obtained in [32], where the polarization operator of a photon propagating in an external field of a complex configuration was calculated. This configuration includes constant and parallel electric and magnetic fields as well as an arbitrary plane-wave field with a wave vector oriented along the common direction of **E** and **H**. We assume that the magnetic field strength in the polarization operator is zero and choose a monochromatic plane wave with circular polarization for a plane-wave field. If we expand the obtained expression into a power series in the field strength of the external wave, the expansion term on the mass surface, which is proportional to the squared wave field strength, is equal to the forward scattering amplitude for a photon with frequency ω' from a photon of the external wave of frequency ω.

In order to determine the pair production cross section, we apply the optical theorem. Since we are primarily interested in a relatively weak field whose strength E satisfies the condition[3] $eE/m^2 \equiv \mu \ll 1$, we will disregard modifications of the unitary relation typical of pair-producing fields [33, 34]. This leads to an error on the order of $\exp(-\pi/\mu)$ associated with the contribution from spontaneous pair production.

Taking this assumption into account, we define the process cross section (σ^+ corresponds to coinciding

[1] The state of the art by 1997 is reflected in [23].

[2] The scheme for hard photon production in the inverse Compton effect was implemented for the first time back in 1964 on the synchrotron of the Institute of Physics, Academy of Sciences of the Soviet Union [26].

[3] Here, we are using the system of units in which $c = \hbar = 1$.

polarizations of photons and σ^- to the opposite case) by the formula

$$\sigma^{\pm} = \frac{ir_0^2 t}{2}\int_{-\infty}^{\infty}\frac{d\rho}{\sinh(\rho - i0)}\int_{-1}^{1}d\beta\exp\left(-i\frac{\rho}{\mu}\right)$$
$$\times\left\{\cosh\rho\left[1 + B_+^2 + B_-^2 - 2B_+\cos z\right.\right.$$
$$\left.\left. + B_-\frac{\beta\sinh\rho}{\sinh(\rho\beta_-)\sinh(\rho\beta_+)}\sin z\right]\right.$$
$$\left. + \frac{\cosh\rho\cosh(\rho\beta) - 1}{\sinh(\rho\beta_-)\sinh(\rho\beta_+)}\sin z[\sin z \pm i(B_+ - \cos z)]\right\}. \quad (1)$$

Here,

$$B_{\pm} = \frac{1}{2\rho}\int_{-1}^{1}dx\left[\frac{1}{x + \coth(\rho\beta_+)} \pm \frac{1}{x + \coth(\rho\beta_-)}\right]$$
$$\times\begin{pmatrix}\cos(xz)\\ \sin(xz)\end{pmatrix}, \quad (2)$$
$$z = \frac{2}{\mu t}\frac{\sinh(\rho\beta_-)\sinh(\rho\beta_+)}{\sinh\rho},$$

where m and e are the mass and charge of an electron, $r_0 = e^2/4\pi m$ is its classical radius, and $\beta_\pm = (1 \pm \beta)/2$. Parameter $t = m^2/\omega\omega'$ characterizes the photon energy. In zero external field, pair production is possible only for $t < 1$.

In order to find the asymptotic expansion of the integrals for $\mu \ll 1$, we use the two-dimensional method of stationary phase [35]. The application of this method for calculating the pair photoproduction cross section was described in detail in [16]; consequently, we will consider only some features of computations.

It can easily be proved that the boundaries of the domain of integration with respect of ρ and β make an exponentially small contribution to the integral. In order to determine the boundary contribution from integration with respect to variable x in functions $B_\pm$, we will use the following approach. Functions $B_\pm$ can be expressed in terms of integral sine and cosine. Using the well-known representation for the latter functions (see formulas (5.2.8) and (5.2.9) in [36]), we obtain

$$B_+ = \frac{1}{2\rho}\{\sin z[f(z_1^+) + f(z_1^-) + f(z_2^+) + f(z_2^-)]$$
$$- \cos z[g(z_1^+) - g(z_1^-) + g(z_2^+) - g(z_2^-)]\},$$
$$B_- = \frac{1}{2\rho}\{\sin z[g(z_1^+) + g(z_1^-) - g(z_2^+) - g(z_2^-)]$$
$$+ \cos z[f(z_1^+) - f(z_1^-) - f(z_2^+) + f(z_2^-)]\}, \quad (3)$$

where

$$z_1^{\pm} = z(\coth(\rho\beta_-) \pm 1), \quad z_2^{\pm} = z(\coth(\rho\beta_+) \pm 1), \quad (4)$$

$$f(y) = \int_0^{\infty}dy\frac{\exp(-xy)}{x^2 + 1},$$
$$g(y) = \int_0^{\infty}dx\frac{x\exp(-xy)}{x^2 + 1}. \quad (5)$$

The auxiliary functions $f(y)$ and $g(y)$ have the following asymptotic expansions for $|y| \longrightarrow \infty$ ($|\arg y| < \pi$):

$$f(y) \sim \frac{1}{y}\sum_{n=0}^{\infty}(-1)^n\frac{(2n)!}{y^{2n}},$$
$$g(y) \sim \frac{1}{y^2}\sum_{n=0}^{\infty}(-1)^n\frac{(2n+1)!}{y^{2n}}. \quad (6)$$

Representation (3) for $B_\pm$ transforms expression (1) to the canonical form for the application of the stationary phase approximation; there exist only two nontrivial exponents in the integrals both for σ^+ and for σ^-:

$$\frac{i}{\mu}S_{\mp}(\rho, \beta) = \frac{i}{\mu}\left[\mp\frac{4}{t}\frac{\sinh(\rho\beta_-)\sinh(\rho\beta_+)}{\sinh\rho} - \rho\right]. \quad (7)$$

Since $S_-(\rho, \beta)$ has no stationary points, in calculating the asymptotic form of the cross sections we must take into account only the stationary points $S_+(\rho, \beta)$ apart from the line of singularities $\rho = 0$:

$$A(\rho = 0, \beta = v), \quad A'(\rho = 0, \beta = -v),$$
$$B\left(\rho = -\ln\frac{1+v}{1-v}, \beta = 0\right),$$
$$C\left(\rho = \ln\frac{1+v}{1-v}, \beta = 0\right). \quad (8)$$

Parameter $v = \sqrt{1-t}$ for the process in zero fields is the velocity of formed particle in the center-of-mass system.

Let us consider the results of calculations. As in the case of pair production in a crossed field [16], the process cross section can be represented as the sum of the monotonic and oscillating parts,

$$\sigma^{\pm} = \sigma^{\pm}_{\text{mon}} + \sigma^{\pm}_{\text{osc}}, \quad (9)$$

the contribution to the monotonic part coming from the line $\rho = 0$ and from the stationary points A and A', while the oscillating part receives a contribution from stationary points B and C.

The expansion of the monotonic part of the cross section into a power series in μ^2 coincides with the

series in perturbation theory. Its first two terms are given by

$$\sigma^-_{\rm mon} = \pi r_0^2 t\left\{\left(2+t-\frac{t^2}{2}\right)\ln\frac{1+v}{1-v}-(4+t)v + \mu^2 t\left[(3t^2+2t-4)\frac{t}{4}\ln\frac{1+v}{1-v} + \frac{16-19t-4t^2-12t^3+18t^4}{12v^3}\right]\right\},$$
$$\sigma^+_{\rm mon} = \pi r_0^2 t\left\{t\left(1-\frac{t}{2}\right)\ln\frac{1+v}{1-v}+(2-t)v + \mu^2 t\left[(3t^2-2t+4)\ln\frac{1+v}{1-v} - \frac{32-56t-58t^2+223t^3-216t^4+72t^5}{48v^5}\right]\right\}. \quad (10)$$

It is interesting to note that the correction to the Breit–Wheeler cross section due to the presence of the external field reverses its sign upon a change in the energy of interacting photons. For example, the coefficient of μ^2 in the expression for $\sigma^+_{\rm mon}$ is negative in the region of high photon energies. It becomes positive for $t > 0.37$. For $\sigma^-_{\rm mon}$, the inverse dependence is observed. The coefficient of μ^2 is positive in the high-energy range. It becomes negative for $t > 0.74$.

Let us consider the oscillating part. It can be expanded into a power series in μ. The principal terms of the series have the form

$$\sigma^-_{\rm osc} = -\pi r_0^2 t\frac{2\mu^2}{v}\left(\ln\frac{1+v}{1-v}\right)^{-2} \times \sin\left[\frac{1}{\mu}\left(\frac{2v}{t}-\ln\frac{1+v}{1-v}\right)\right],$$
$$\sigma^+_{\rm osc} = -\pi r_0^2 t^2\frac{\mu}{v}\left(\ln\frac{1+v}{1-v}\right)^{-1} \times \cos\left[\frac{1}{\mu}\left(\frac{2v}{t}-\ln\frac{1+v}{1-v}\right)\right]. \quad (11)$$

As the photon energy decreases, the absolute value of corrections increases and we have

$$\sigma^- = \frac{8}{3}\pi r_0^2 v^3\left(1-\frac{\mu^2}{32v^6}-\frac{3\mu^2}{16v^6}\sin\frac{4v^3}{3\mu}\right),$$
$$\sigma^+ = 2\pi r_0^2 v\left(1+\frac{\mu^2}{32v^6}-\frac{\mu}{4v^3}\cos\frac{4v^3}{3\mu}\right) \quad (12)$$

for $v \ll 1$. Consequently, the effective expansion parameter in this energy range is the quantity μ/v^3.

The obtained formulas are inapplicable for $|v| < \mu^{1/3}$, i.e., in the immediate vicinity of the value $t = 1$ corresponding to the pair production threshold in zero external field. In order to estimate cross sections in this region, we substitute $t = 1$ into formula (1), which readily gives

$$\sigma^- = \pi r_0^2\mu\frac{2}{3\sqrt{3}}, \quad \sigma^+ = \pi r_0^2\mu^{1/3}\frac{\Gamma(5/6)}{(12)^{1/6}\sqrt{\pi}}. \quad (13)$$

The steepest descent method can be used for calculating the asymptotic cross sections in the prethreshold region, i.e., for $t > 1$:

$$\sigma^-_{\rm osc} = \pi r_0^2 t\frac{\mu^2}{4\tilde{v}\arctan^2\tilde{v}}\exp\left[\frac{2}{\mu}\left(\frac{\tilde{v}}{t}-\arctan\tilde{v}\right)\right],$$
$$\sigma^+_{\rm osc} = \pi r_0^2 t^2\frac{\mu}{4\tilde{v}\arctan\tilde{v}}\exp\left[\frac{2}{\mu}\left(\frac{\tilde{v}}{t}-\arctan\tilde{v}\right)\right], \quad (14)$$

where $\tilde{v} = (t-1)^{1/2}$. For $\tilde{v} \ll 1$, we have

$$\sigma^- = \pi r_0^2\frac{\mu^2}{4\tilde{v}^3}\exp\left(-\frac{4\tilde{v}^3}{3\mu}\right),$$
$$\sigma^+ = \pi r_0^2\frac{\mu}{4\tilde{v}^2}\exp\left(-\frac{4\tilde{v}^3}{3\mu}\right). \quad (15)$$

Let us return to formulas (14). For $t \gg 1$, these formulas fail to describe the two-photon productions. Indeed, in this case, $\sigma^\pm \sim \exp(-\pi/\mu)$; i.e., the dependence on the external field strength, typical of the probability of spontaneous production of electron–positron pairs by an electric field, takes place.

The formulas describing pair generation in the nonrelativistic approximation are similar to the well-known result of the problem on optical transitions in semiconductors near the absorption band in the presence of an electric field (the Franz–Keldysh effect). The probability of this process for an isotropic dispersion relation can be obtained from formulas (12) and (15) by replacing parameter μ/v^3 by $eE/\sqrt{2m^*}(\varepsilon-\varepsilon_0)^{3/2}$ (where ε is the total energy of photons, ε_0 is the forbidden gap width, and m^* is the reduced mass of an electron–hole pair) and assuming that the overall coefficient is equal to the absorption probability in zero field. In this case, the quantities σ^+ and σ^- determine the probabilities of the allowed [37] and forbidden transitions, respectively. The observed analogy is quite clear since the presence of oscillating terms in the expressions for the probabilities of both processes is associated with reflection from a potential barrier. Consequently, the meaning of parameter μ/v^3 is also clear: a field may noticeably change the total probability of the process if the work done by it over the de Broglie wavelength of the particle formed is comparable to its kinetic energy. The role of

this parameter in the particle physics problem was discussed in connection with the analysis of the effect of an external field on the beta decay probability in the nonrelativistic approximation (see [38–41]).

It is interesting to note that, when pair production is studied in the nonrelativistic approximation in an arbitrary static field or in the field of an electromagnetic low-frequency wave [14, 27, 28], the principal terms in the power expansion of the external field strength for the cross sections coincide with formulas (12), (13), and (15). This can be explained qualitatively by the fact that it is the electric field component that mainly affects the cross section near the reaction threshold, where the strongest effect is observed. The value of the Lorentz force, i.e., the effect of the magnetic field, is weaker by a factor of v^{-1}.[4] Consequently, the formulas derived in the nonrelativistic approximation coincide irrespective of the type of the field. In the version of nonrelativistic approximation used in the above-mentioned publications, spin effects were taken into account as follows: wave functions used in calculations described pairs with zero angular momentum [14] or with zero as well as nonzero orbital angular momenta [27, 28]. Consequently, partial cross sections averaged over the polarizations of the initial photons were calculated in fact. Obviously, the partial cross section for production of pairs with zero orbital angular momentum can also be treated as the cross section of a process involving photons with identical polarizations and propagating towards each other. However, the situation for partial cross section with nonzero angular momenta is nontrivial. Our calculations show that in the case when photons propagate along a purely electric field, the treatment of the relevant expression as the cross section of pair production by photons with opposite polarizations is quite admissible.

Let us compare formulas (12), (13), and (15) with the low-energy limit of the process cross section in a crossed field [16]. The above formulas for σ^+ coincide with the result for the crossed field if we carry out the substitution

$$\mu \longrightarrow \kappa = \frac{e\tilde{E}}{m^2}\sqrt{\frac{\omega'}{\omega}}, \qquad (16)$$

i.e., replace E by the strength of the crossed field in the center-of-mass system of the pair being formed ($\tilde{E}$ is the strength of the crossed field in the laboratory frame of reference). However, the formulas for σ^- with substitution (16) differ from the cross section in the crossed field, which has the following form near the threshold:

$$\sigma^- = \frac{8}{3}\pi r_0^2 v^3\left(1 - \frac{1}{8}\frac{\kappa^2}{v^6} + \frac{3}{8}\frac{\kappa}{v^3}\cos\frac{4v^3}{3\kappa}\right). \qquad (17)$$

First of all, the dependence on the field strength in the expression for σ^-_{osc} is obviously different. This can be explained as follows. Since the oscillating part of the cross section is formed due to reflection from a potential barrier, pairs with momenta directed along the electric field make the largest contribution to this component. In the case of interaction of photons with different polarizations, the production of a pair with zero orbital angular momentum is ruled out. In this case, the probability of emission of a pair along the field is strongly suppressed.

Considerations associated with the analyticity in energy suggest that oscillating terms appear only in the cross sections of reactions with an energy threshold. For example, no oscillating terms appear in the total cross section of the Compton effect in an external crossed electric field. In the above analogy with the problems in solid-state physics (see also [42]), the boundary of the physical region in the momentum space plays the role of the Fermi surface.

Let us now consider the possibility of experimental observation of the effect of electric field on pair production. This problem was analyzed in detail for a crossed field in [16], and we will compare our formulas with the results of that publication. For the second harmonic of a neodymium laser used in the experiment described in [24], the photon energy is $\omega \approx 2.35$ eV. The threshold value is $\omega' \approx 111$ GeV. Thus, the strengths of the crossed and purely electric fields in the center-of-mass system of a pair being formed differ by a factor of $\chi = (\omega'/\omega)^{1/2} \approx 0.22 \times 10^6$ if their values in the laboratory system are equal. Consequently (see formulas (13)), at the reaction threshold, where this effect is the strongest, the ratio of cross sections for photons with identical polarizations is $\chi^{1/3} \approx 0.6 \times 10^2$. For this reason, the conditions for observing the effect in a purely electric field are noticeably worse than for a crossed field for which the cross section is on the order of 10% of the maximal Breit–Wheeler cross section even for field strengths of a few tesla. However, the cross sections for photons with different polarizations differ by more than a factor of χ (in our example, by a factor of 1.5×10^6).

It should be noted that the results obtained for $v \lesssim \alpha$, where α is the fine structure constant, are generally only qualitative since we have not considered the interaction of the pair components in the final state. However, the inclusion of the Coulomb interaction cannot radically affect the main conclusions of this study.

ACKNOWLEDGMENTS

The authors are grateful to A.V. Borisov, V.Ch. Zhukovskiĭ, and V.N. Rodionov for fruitful discussions.

REFERENCES

1. G. Breit and J. A. Wheeler, Phys. Rev. **46**, 1087 (1934).
2. H. R. Reiss, J. Math. Phys. **3**, 59 (1962).

[4] Analysis of pair production in a constant magnetic field revealed [15] that the effective expansion parameter in this case is μ/v^2.

3. A. I. Nikishov and V. I. Ritus, Zh. Éksp. Teor. Fiz. **46**, 776 (1964) [Sov. Phys. JETP **19**, 529 (1964)].
4. A. I. Nikishov and V. I. Ritus, Zh. Éksp. Teor. Fiz. **46**, 1768 (1964) [Sov. Phys. JETP **19**, 1191 (1964)].
5. N. B. Narozhnyĭ, A. I. Nikishov, and V. I. Ritus, Zh. Éksp. Teor. Fiz. **47**, 930 (1964) [Sov. Phys. JETP **20**, 622 (1964)].
6. A. I. Nikishov and V. I. Ritus, Zh. Éksp. Teor. Fiz. **47**, 1130 (1964) [Sov. Phys. JETP **20**, 757 (1964)].
7. A. I. Nikishov and V. I. Ritus, Zh. Éksp. Teor. Fiz. **52**, 1707 (1967) [Sov. Phys. JETP **25**, 1135 (1967)].
8. I. M. Ternov, V. G. Bagrov, Yu. I. Klimenko, and B. V. Kholomaĭ, Izv. Vyssh. Uchebn. Zaved., Fiz. **11** (8), 71 (1968).
9. A. V. Borisov, O. G. Goryaga, and V. Ch. Zhukovskiĭ, Izv. Vyssh. Uchebn. Zaved., Fiz. **20** (2), 46 (1977).
10. V. A. Lyul'ka, Zh. Éksp. Teor. Fiz. **67**, 1638 (1974) [Sov. Phys. JETP **40**, 815 (1974)].
11. V. A. Lyul'ka, Zh. Éksp. Teor. Fiz. **72**, 865 (1977) [Sov. Phys. JETP **45**, 452 (1977)].
12. B. A. Lysov, O. F. Dorofeev, and O. S. Pavlova, Izv. Vyssh. Uchebn. Zaved., Fiz. **11** (2), 46 (1968).
13. B. A. Lysov, O. S. Pavlova, and A. F. Zhuravlev, Vestn. Mosk. Univ., Ser. 3: Fiz., Astron. **12** (5), 557 (1971).
14. A. I. Nikishov and V. I. Ritus, Zh. Éksp. Teor. Fiz. **85**, 1544 (1983) [Sov. Phys. JETP **58**, 898 (1983)].
15. A. E. Lobanov and A. R. Muratov, Zh. Éksp. Teor. Fiz. **87**, 1140 (1984) [Sov. Phys. JETP **60**, 651 (1984)].
16. A. E. Lobanov and A. R. Muratov, Zh. Éksp. Teor. Fiz. **90**, 409 (1986) [Sov. Phys. JETP **63**, 235 (1986)].
17. A. A. Kozlenkov and I. G. Mitrofanov, Zh. Éksp. Teor. Fiz. **91**, 1978 (1986) [Sov. Phys. JETP **64**, 1173 (1986)].
18. V. Ch. Zhukovskiĭ and I. Kherrman, Yad. Fiz. **14**, 1014 (1971) [Sov. J. Nucl. Phys. **14**, 569 (1971)].
19. V. Ch. Zhukovskiĭ and N. S. Nikitina, Yad. Fiz. **19**, 148 (1974) [Sov. J. Nucl. Phys. **19**, 77 (1974)].
20. V. N. Rodionov, I. M. Ternov, and V. R. Khalilov, Zh. Éksp. Teor. Fiz. **69**, 1148 (1975) [Sov. Phys. JETP **42**, 585 (1975)].
21. V. N. Rodionov, Zh. Éksp. Teor. Fiz. **78**, 105 (1980) [Sov. Phys. JETP **51**, 52 (1980)].
22. A. E. Lobanov, V. N. Rodionov, and V. R. Khalilov, Yad. Fiz. **32**, 174 (1980) [Sov. J. Nucl. Phys. **32**, 90 (1980)].
23. A. V. Borisov, A. S. Vshivtsev, V. Ch. Zhukovskiĭ, and P. A. Éminov, Usp. Fiz. Nauk **167**, 241 (1997) [Phys. Usp. **40**, 229 (1997)].
24. D. L. Burke, R. C. Field, G. Horton-Smith, *et al.*, Phys. Rev. Lett. **79**, 1626 (1997).
25. C. Bula, K. T. McDonald, E. J. Prebys, *et al.*, Phys. Rev. Lett. **76**, 3116 (1996).
26. O. F. Kulikov, Yu. Ya. Tel'nov, E. I. Filippov, and M. N. Yakimenko, Zh. Éksp. Teor. Fiz. **47**, 1591 (1964) [Sov. Phys. JETP **20**, 1069 (1964)].
27. V. N. Rodionov, Zh. Éksp. Teor. Fiz. **113**, 21 (1998) [JETP **86**, 11 (1998)].
28. V. G. Kadyshevskiĭ, G. A. Kravtsova, and V. N. Rodionov, Teor. Mat. Fiz. **130**, 275 (2002).
29. N. B. Narozhny and M. S. Fofanov, Phys. Rev. E **60**, 3443 (1999).
30. N. B. Narozhnyĭ and M. S. Fofanov, Zh. Éksp. Teor. Fiz. **117**, 476 (2000) [JETP **90**, 415 (2000)].
31. V. S. Popov, Zh. Éksp. Teor. Fiz. **121**, 1235 (2002) [JETP **94**, 1057 (2002)].
32. A. E. Lobanov and V. R. Khalilov, Zh. Éksp. Teor. Fiz. **77**, 548 (1979) [Sov. Phys. JETP **50**, 278 (1979)].
33. A. I. Nikishov, Tr. Fiz. Inst. im. P.N. Lebedeva, Akad. Nauk SSSR **168**, 156 (1986).
34. D. M. Gitman, E. S. Fradkin, and Sh. M. Shvartsman, *Quantum Electrodynamics with Unstable Vacuum* (Nauka, Moscow, 1991).
35. M. V. Fedoryuk, *The Steepest Descent Method* (Nauka, Moscow, 1977).
36. *Handbook of Mathematical Functions*, 2nd ed., Ed. by M. Abramowitz and I. A. Stegun (Dover, New York, 1971; Nauka, Moscow, 1979).
37. L. V. Keldysh, Zh. Éksp. Teor. Fiz. **34**, 1138 (1958) [Sov. Phys. JETP **7**, 788 (1958)].
38. A. I. Nikishov and V. I. Ritus, Zh. Éksp. Teor. Fiz. **85**, 24 (1983) [Sov. Phys. JETP **58**, 14 (1983)].
39. I. M. Ternov, V. G. Zhulego, V. N. Rodionov, *et al.*, Vestn. Mosk. Univ., Ser. 3: Fiz., Astron. **24** (4), 79 (1983).
40. M. B. Voloshin, Yad. Fiz. **38**, 814 (1983) [Sov. J. Nucl. Phys. **38**, 484 (1986)].
41. E. Kh. Akhmedov, Zh. Éksp. Teor. Fiz. **85**, 1521 (1983) [Sov. Phys. JETP **58**, 883 (1983)].
42. A. E. Lobanov, Pis'ma Zh. Éksp. Teor. Fiz. **50**, 161 (1989) [JETP Lett. **50**, 179 (1989)].

Translated by N. Wadhwa

Journal of Experimental and Theoretical Physics, Vol. 96, No. 4, 2003, pp. 674–683.
Translated from Zhurnal Éksperimental'noĭ i Teoreticheskoĭ Fiziki, Vol. 123, No. 4, 2003, pp. 763–774.

ATOMS, SPECTRA, RADIATION

Method of Transfer Relations in the Theory of Multiple Resonant Scattering of Waves as Applied to Diffraction Gratings and Photonic Crystals

Yu. N. Barabanenkov[a,*] and M. Yu. Barabanenkov[b]

[a] *Eldis Research Center For Electronic Diagnostic Systems, Russian Academy of Sciences, Moscow, 101000 Russia*

[b] *Institute of Microelectronic Technology and Ultra-High-Purity Materials, Russian Academy of Sciences, Chernogolovka, Moscow oblast, 142432 Russia*

**e-mail: Yu.Barab@g23.relcom.ru*

Received November 6, 2002

Abstract—An approach is proposed in the theory of multiple scattering of wave fields in two-dimensional inhomogeneous media, which provides a universal description for wave scattering from one-dimensional periodic interfaces between two dielectric media (optical gratings) and from two-dimensional periodic dielectric structures (photonic crystals). The approach is based on the transfer matrix methodology, which involves subdividing the scattering medium into elementary layers with gaps; however, in contrast to the transfer matrix method, it leads to invariant imbedding equations for the matrix coefficients of reflection and transmission of an inhomogeneous medium. The developed approach is applied in a quantitative analysis of two optical effects: resonant decrease in the light reflection coefficient from the grating, associated with the profile depth effect, and exponential-power decay of optical radiation in the forbidden band of a 2D photonic crystal upon an increase in the number of its layers starting from one layer. The frequency spectrum for the electromagnetic radiation power transmission through a 2D photonic crystal formed by parallel layers of infinitely long cylinders is interpreted taking into account the spectral dependence of the total cross section of scattering from a single cylinder. Such an interpretation of the frequency spectrum with two forbidden gaps combined with the analysis of layer-by-layer dynamics of its formation makes it possible to reveal the role of microscopic resonant scattering of waves from a single cylinder and of macroscopic Bragg-type resonant scattering from a periodic system of cylinders during the formation of the spectrum of radiation transmission through a photonic crystal. A physical explanation is given for the transparency peaks in one of the forbidden gaps in the spectra of radiation transmission through a perfectly ordered system of cylinders in terms of multipole resonances of scattering from a single cylinder. © *2003 MAIK "Nauka/Interperiodica".*

1. INTRODUCTION

For historical reasons, the methods for describing scattering from inhomogeneous (including periodic) surfaces and inhomogeneous (periodic) volume dielectric media in the theory of multiple scattering of wave fields were developed independently. The scattering from the surfaces, which is often referred to as surface scattering [1], is usually described in terms of the boundary value problem using perturbation theory [2, 3], or it is considered on the basis of the Fredholm integral equations written for a boundary field (which are subsequently solved numerically [4]), or it is based on the scattering (S) matrix technique [5]. In the theory of multiple scattering in volume media, the following three approaches are commonly used: the Watson method of composition of scattering operators (T matrix technique) [6], the invariant imbedding method [7, 8], and the transfer matrix method, which has become popular recently [9].

Intense experimental studies of electromagnetic wave scattering from structures with a periodic dielectric "potential" (optical gratings and photonic crystals) and from rough surfaces have stimulated the development of several new theoretical approaches. The description of some optical resonance effects in gratings, such as Wood's "perpendicular" anomalies [10] associated with excitation of surface plasmons [11] in the grating by an incident wave with a nonzero electric vector component along the normal to the unperturbed surface; suppression of specular reflection of light [12]; giant Raman scattering [13], as well as anomalous absorption of light by a rough surface [14] and coherent enhancement of back scattering of light from a rough interface in a multimode waveguide structure [15], were no longer confined to the solution of a traditional diffraction problem formulated in terms of the Fredholm equations [16, 17]. Using approximate boundary conditions in the form of surface impedance [18, 19], the waves excited on the surface were calculated exactly [20] or approximately [21], or the problem of the wave field in the transition region of a perturbed surface [20] was formulated analogously to the case of

wave propagation in a periodic layered volume structure [22].

The studies of photonic crystals initiated by Yablonovitch and John [23] led to new methods for calculating the coefficients of reflection or transmission of waves for periodic volume structures, for calculating local fields in such structures, and for analyzing the band structure of their frequency spectrum emerging due to geometrical resonance during Bragg scattering of waves. We can mention in this connection the method of plane waves [24–26], the layer doubling algorithm [27, 28] inherited from the theory of low-energy electron diffraction [29], and the finite element method [30, 31] based on discretization of the Maxwell equations, carried out in real space, as well as the above-mentioned transfer matrix technique [9].

The method of plane waves and the layer doubling algorithm can be reduced to the Fredholm matrix equation, which is typical of surface scattering problems in their traditional formulation [4]. At the same time, the effective potential method for surface scattering (although it was applied [15] to analysis of coherent amplification of backward scattering in the framework of the Rayleigh hypothesis) together with the *S* matrix technique [5] and the finite element method [30, 31] close to the invariant imbedding method [7, 8] can be regarded as a step to a uniform description of surface and volume multiple scattering of electromagnetic waves.

From the practical point of view, the development of a unified approach to the problem of multiple scattering of optical radiation is stimulated by miniaturization of optical and optoelectronic structures, which rules out an unambiguous division between their surface and volume as, for example, in the case of a planar microscopic optical fiber whose surface is formed by a layer of a colloidal hexagonal crystal made of spheres [32].

Here, we develop a unified exact approach to the theoretical description of multiple resonant scattering of waves from one-dimensional periodic interface between two dielectric medial (optical gratings) and from two-dimensional (2D) periodic dielectric structures (photonic crystals). The approach is a further development of the transfer matrix methodology [9] involving a virtual subdivision of a scattering medium into elementary nonintersecting layers. The development of the method carried out in conformity with the method of transfer relations [33] consists in the inclusion of the splitting of a transition inhomogeneous region of a periodic surface or individual scatterers of a periodic volume structure. From the physical point of view, the applicability of transfer matrices both to gratings and to photonic crystals is based on the possibility of representing photonic crystals in the form of a stack of corresponding gratings; this was demonstrated while deriving the transfer relations [33].

We will write the invariant imbedding relations for matrix wave coefficients of reflection and transmission of a periodic surface (optical grating). It will be shown that the same equations with appropriately modified coefficients can be used in the case of photonic crystals. The effectiveness of the proposed approach will be demonstrated as applied to an analysis of the following two optical effects: (1) "parallel" anomalies on gratings, observed about a century ago by Wood [34] and studied in detail by Palmer [35]; physically, these anomalies are not associated with the excitation of surface plasmons (the electric vector of an incident wave is parallel to the generator of the grating relief) in contrast to the above-mentioned "perpendicular" anomalies [10]; and (2) the dependence of the depth of penetration of optical radiation into a photon 2D structure upon an increase in the number of its layers, provided that the radiation frequency lies in the forbidden gap of the transmission frequency spectrum of the structure. The power transmission frequency spectrum for an ordered system of infinitely long dielectric cylinders is interpreted with the help of the spectral dependence of the total cross section of scattering of incident electromagnetic field from a single cylinder. A comparison of the layer-by-layer dynamics of formation of the forbidden gap in the transmission spectrum of a 2D photonic crystal with the wave field attenuation depth in a layer of a random strongly scattering dielectric 2D medium, which is calculated on the basis of the self-consistent theory of localization [36], enabled us to analyze the John hypothesis [23] on the essentially feasible realization of strong localization of photons in an ordered medium with a moderate disorder suppressing coherent backward scattering of photons in some of Bragg's directions.

2. RICCATI EQUATION FOR LIGHT SCATTERING FROM A PERIODIC 1D SURFACE AND FROM A 2D VOLUME STRUCTURE

Let us consider the scattering of a plane monochromatic electromagnetic wave from a periodic one-dimensional (1D) interface $z = f(x)$ between two dielectric media occupying the upper and lower half-spaces in the Cartesian system of coordinates x, y, z. Let us suppose that the periodic surface has a triangular profile in the xz plane, which is extended along the y axis (Fig. 1a). The plane linearly polarized wave with wavelength λ is incident from the upper half-space ($z > f(x)$), whose permittivity is ε^{bac}. The lower half-space ($z < f(x)$) consists of a transition region $0 < z < f(x)$ and a substrate $-L_o < z < 0$, which have the same permittivity ε, and a semi-infinite medium $z < -L_o$, similar to the medium of the upper half-space. We consider the case of the *TE* (or *s*) polarization, which enables us to reduce the vector Maxwell equations to the scalar Helmholtz wave equation.

In accordance with the transfer relation method [33], the transition region can be divided into a system of elementary layers perpendicular to the z axis, which

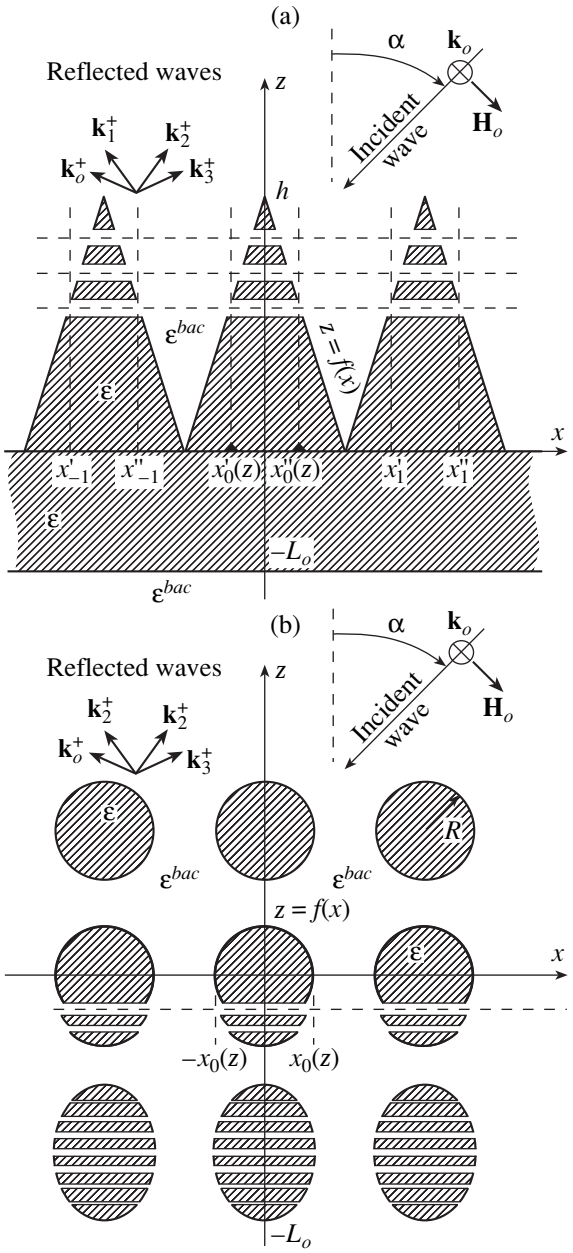


Fig. 1. Schematic diagram of splitting of the transition region between two dielectric media into auxiliary layers with gaps. The transition region has the form of (a) a periodic surface with a triangular profile on a substrate of thickness L_o; (b) two-dimensional volume structure, e.g., a system of parallel layers (starting from one layer) of infinitely long cylinders of radius R extended along the y axis, which forms a square grating (starting from two layers) with period Λ. The electric field vector $\mathbf{E}_o$ of a plane wave incident at angle α is parallel to the y axis (*TE* polarization). The wave vector $\mathbf{k}_o$ and the magnetic vector $\mathbf{H}_o$ of the wave lie in the xz plane of incidence.

are separated by vanishingly small gaps. The Watson rule [6] for a composition of scattering operators (T matrices) for the system of layers leads to a mixed system of exact matrix equations (transfer relations) for the matrix wave coefficients of reflection and transmission through the system of elementary layers and the matrix wave amplitudes of waves in the gaps between the layers (local fields). The transfer relations make it possible to write the generalized differential Riccati equation for the matrix wave coefficient of reflection and associated differential equation for the matrix wave coefficient of transmission through a periodic surface as well as the differential equation for the local fields in the transition region. The imbedding parameter in these differential equations with the "initial" (or "final") conditions is chosen along the z axis.

The Riccati equation written for the matrix coefficient $R_{\mu\nu}(k_{ox})$ of reflection from a truncated transition region, i.e., the coefficient of reflection from an auxiliary surface $\xi = f(x, z)$, which is the cross section of the transition region by the plane z = const, has the form

$$\frac{dR}{dz} - i(\sigma R + R\sigma) = aF(z) + aF(z)R + RaF(z) + RaF(z)R \quad (1)$$

for $0 < z < h$. The "initial" condition for the Riccati equation is the substrate reflection coefficient $R(k_{ox})$, i.e., the condition $R_{\mu\nu}(k_{ox}) \longrightarrow R(k_{ox} + 2\pi\nu/\Lambda)\delta_{\mu\nu,0}$ for $z \longrightarrow 0$. Here, h is the maximal depth of the surface profile and $\delta_{\mu\nu}$ is the Kronecker delta symbol. Indices μ, $\nu = 0, \pm1, \pm2, \ldots$ label the diffracted and incident spectral orders. In particular, $\nu = 0$ corresponds to the incidence of a plane wave, while $\mu = 0$ corresponds to the direction of specular reflection.

The Riccati equation (1) contains two coefficients σ and a, which are diagonal matrices with elements of the form

$$\sigma(\mu) = [k_o^2 - (k_{ox} + 2\pi\mu/\Lambda)^2]^{1/2},$$
$$2\pi a(\mu) = -iV_o/\sigma(\mu).$$

Here, Λ is the modulation period of the surface, $V_o = -k_o^2(\varepsilon - \varepsilon^{bac})/\varepsilon^{bac}$ is the dielectric potential, $k_o = (\omega/C)(\varepsilon^{bac})^{1/2}$ is the wave number of optical radiation, and ω and C are the cyclic frequency of radiation and the velocity of light in vacuum.

The physical meaning of the coefficient $\sigma(\mu)$ is that the diagonal of this matrix is composed of z components of the wave vectors $\mathbf{k}_\mu^+$ of radiation scattered to the μth diffraction order. Vector $\mathbf{k}_\mu^+$ has the x component $k_{\mu x}^+ = k_{ox} + 2\pi\mu/\Lambda$, where $k_{ox} = -k_o\sin\alpha$. Here, α denotes the angle of incidence of the wave.

Matrix $F_{\mu\nu}(z) \equiv F_{\mu-\nu}(z)$ with elements

$$F_\mu(z) = \frac{1}{\mu}\sin\frac{2\pi\mu X_o(z)}{\Lambda} \quad (2)$$

describes in Eq. (1) the mutual transformation of propagating and evanescent waves, occurring during their

multiple scattering from a periodic surface. In matrix elements (2), $X_o(z)$ denotes the root of the equation $f(x) = f(-x) = z$ for $0 \le x \le \Lambda/2$.

Among all elements of the matrix reflection coefficient $R_{\mu\nu}$ corresponding to $z = h$, only elements of the first column $R_{\mu o}$ have physical meaning. Indeed, the expression for the y component of the electric field strength of the wave in the upper half-space $z > h$,

$$E_y(\mathbf{r}) = \exp[i\mathbf{k}_o(\mathbf{r} - h\hat{\mathbf{z}})] + \sum_{\mu=-\infty}^{\infty} R_{\mu o}(k_{ox}; h)\exp[i\mathbf{k}_\mu^+(\mathbf{r} - h\hat{\mathbf{z}})], \quad (3)$$

where $\hat{z}$ denotes the unit vector of the z axis, shows that the field reflected from a periodic surface is represented by the term in the form of the sum of products of the exponential factor and the partial reflection coefficient $R_{\mu o}$ for a wave incident on the grating.

Analogously to Riccati equation (1), we can write an associated equation for the matrix wave coefficient $\mathcal{T}_{\mu\nu}(k_{ox})$ of radiation transmission through a truncated periodic surface:

$$d\mathcal{T}/dz - i\mathcal{T}\sigma = \mathcal{T}aF(z) + \mathcal{T}aF(z)R \quad (4)$$

for $0 < z < h$ with the "initial" condition $\mathcal{T}_{\mu\nu}(k_{ox}) \longrightarrow \mathcal{T}(k_{ox} + 2\pi\nu/\Lambda)\delta_{\mu-\nu,0}$ for $z \longrightarrow 0$. Here, $\mathcal{T}(k_{ox})$ is the transmission coefficient for the substrate.

The y component of the electric field strength in the lower half-space $z + L_o < 0$ is given by

$$E_y(\mathbf{r}) = \sum_{\mu=-\infty}^{\infty} \mathcal{T}_{\mu o}(k_{ox}; h) \times \exp\{i\mathbf{k}_\mu^-[\mathbf{r} + (z + L_o)\hat{\mathbf{z}}]\}. \quad (5)$$

Here, $\mathbf{k}_\mu^-$ denotes the wave vector of the μth mode in the angular spectrum of transmitted radiation; the wave vectors have components $k_{\mu x}^- = k_{\mu x}^+$ and $k_{\mu z}^- = -\sigma(\mu)$. Formulas (3) and (5) describe the expansion of the electric field of a wave in spectral orders (propagating and attenuating along the z axis) of radiation reflected from the periodic surface and radiation transmitted through it. The substitution of these formulas into the definition of the Poynting vector $\mathbf{\Pi}$ leads to the expressions for the energy flux density of the reflected and transmitted electromagnetic field,

$$\Pi_z(z = h) = k_{oz} + \sum_\mu k_{\mu z}^+ |R_{\mu o}(k_{ox}; h)|^2, \quad (6)$$

$$\Pi_z(z = -L_o) = \sum_\mu k_{\mu z}^- |\mathcal{T}_{\mu o}(k_{ox}; h)|^2, \quad (7)$$

in which summation is carried out only over propagating spectral orders defined by the inequality $|k_{ox} + 2\pi\mu/\Lambda| \le k_o$. It should be noted that formulas (6) and (7) for nonabsorbing dielectric media must lead to identical values of the field energy flux density (Poynting theorem).

In numerical calculations, the infinite system of equations (1) was truncated into a system of 15×15 equations used for calculating the reflection coefficients $R_{\mu\nu}$ to the spectral orders $\mu, \nu = 0, \pm1, \ldots, \pm7$ with the help of the Runge–Kutta method.

3. WOOD ANOMALIES

It is known from the literature (see, for example, [21]) that the coefficient of reflection of a *TE* polarized wave from a diffraction grating may sharply increase for some values of the grating period or, on the contrary, may abruptly decrease for some values of its profile depth. The resonant increase in reflectance (so-called Rayleigh type of Wood anomalies [34]) was considered for a metallic grating to demonstrate the effectiveness of numerical count using the transfer relation method [33]. Here, we will describe quantitatively for the first time the resonant grating transparency effect known as "parallel" Wood anomalies [34] investigated in [18, 35].

Figure 2 shows the dependence of the modulus of reflectance, calculated on the basis of the Riccati equation (1) for several spectral orders, on the profile depth h/λ for a triangular silver grating with a period $\lambda/\Lambda = 1/3 + 10^{-7}$. In the case of a triangular grating profile (see Fig. 1), function $X_o(z)$ in the transformation matrix for waves has the form $X_o(z) = (\Lambda/2)(1 - z/h)$. For the normal incidence of a wave ($\alpha = 0$), propagating spectral orders are defined by the inequality $1/|\mu| \ge \lambda/\Lambda = 1/3 + 10^{-7}$; i.e., $|\mu| < 3$.

Let us consider the results depicted in Fig. 2. First, the three minima on the curve labeled by zero indicate partial suppression of specular reflection [12]. In addition, the minima on the curves corresponding to different spectral orders coincide with the experimental results [35] generalized in [18] in the form of the inequalities

$$n/2 > h/\lambda > (2n - 1)/4, \quad n = 1, 2, \ldots \quad (8)$$

and represented in Fig. 2 by the hatched rectangles on the abscissa axis. Second, the minimum on the curve corresponding to the propagating mode (three arrows in the figure) corresponds to the maximum on the curve corresponding to the decaying mode, indicating a considerable redistribution of the energy of incident radiation among various spectral orders of the reflected field. Third, the reflectance for the spectral order $\mu = 3$ is close to unity, except for a grating with a shallow profile ($h/\lambda < 0.75$). This result is of interest for near field spectroscopy since the third mode, being a gliding (with a glancing angle of 0.04°) and damped mode (along the z axis), does not carry away the field energy from the sur-

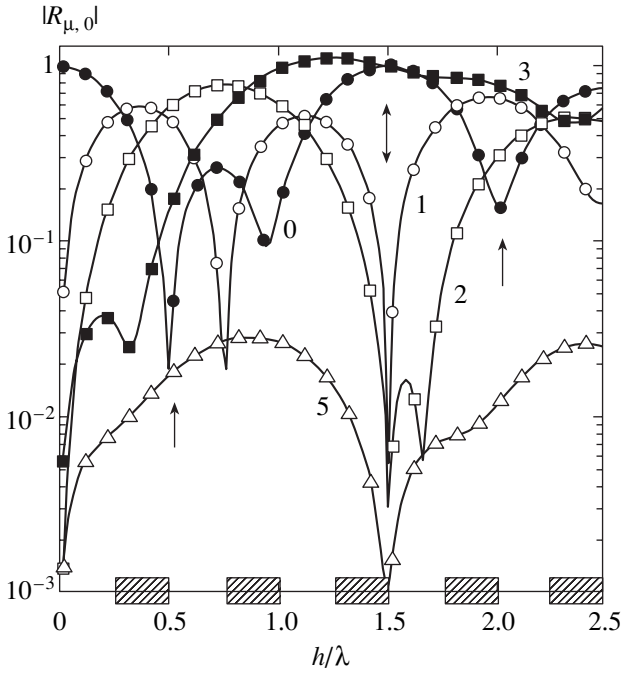


Fig. 2. Modulus of the partial coefficient $|R_{\mu,0}|$ of reflection to the μth spectral order of a *TE* polarized plane electromagnetic wave ($\lambda = 632.8$ nm in a medium with $\varepsilon^{bac} = 1$) from a silver grating ($\varepsilon = -17.5 + i \times 0.7$, $\lambda/\Lambda = 1/3 + 10^{-7}$) as a function of the grating profile depth h/λ for the normal incidence of the wave ($\alpha = 0$) and for a semi-infinite substrate with $L_o \longrightarrow \infty$. The curves are labeled by the number of corresponding diffraction orders. The rectangles on the abscissa axis illustrate inequalities (8).

face. Finally, additional computations show that each minimum on the frequency dependence of the reflectance can be enhanced additionally due to oblique incidence of the wave for a fixed value of the grating profile depth. However, for an angle of incidence $\alpha > 70°$, the reflectance (its modulus and phase) is practically independent of the grating profile depth.

Thus, the calculations made for a metallic grating (see Fig. 2) and, in addition, for a dielectric grating (Si_3N_4, $\varepsilon = 4.41$) confirm the assumption made in [35] that the reason for parallel anomalies of the coefficient of reflection of an electromagnetic wave from a grating is multiple reflection of the wave between the slopes of two adjacent crests of the grating, i.e., the effect of its profile depth.

4. TWO-DIMENSIONAL PHOTONIC CRYSTALS

Riccati equation (1) and Eq. (4) associated with it can be used directly for calculating the coefficients of reflection and transmission of a *TE* polarized plane wave from and through a 2D photonic crystal formed by layers of rods oriented in each layer along the *y* axis and forming a triangular [25, 31] or rectangular [26] cell in the *xz* cross section. In this case, only the expression for X_o in the function of mutual transformation of modes (2) is changed. In the particular case of a square grating with period Λ formed by cylinders of radius R (see Fig. 1b), function X_o is given by $X_o(z) = [R^2 - (z - n\Lambda)^2]^{1/2}$ if the inequality $-R + n\Lambda < z < R + n\Lambda$ is satisfied; i.e., the cut plane $z = \text{const}$ lies within a layer of cylinders with number $n = 0, \pm1, \pm2, \ldots$. If, however, the cut plane lies between layers of cylinders, matrix (2) has zero elements.

In the case of a nonabsorbing material ($\text{Im}\varepsilon = 0$), transparency (7) of a system of cylinders for the power of incident electromagnetic field can be calculated on the basis of Riccati equation (1) and the Poynting theorem, i.e., without solving Eq. (4). Figures 3 and 4 show the results of such a method for calculating the spectral power transparency for a system of infinitely long cylinders of radius $R = 0.6$ and permittivity $\varepsilon = (2.9)^2$, which are packed in air ($\varepsilon^{bac} = 1$) to form a square grating with period $\Lambda = 4$. The factor of filling the space with cylinders is $n^* = \pi R^2/\Lambda^2 \approx 7\%$. The values of parameters are given in dimensionless form [37].

The solid curve in Fig. 3a shows that the ordered system of cylinders for the chosen numerical values of parameters exhibit the properties of a photon structure. The power transmission frequency spectrum has two regions of virtually full reflection for the wave field with normal incidence (forbidden or opacity bands) separated by the transparency region. We will refer to the forbidden bands as the main ($8.2 \le \lambda \le 12.4$) and additional ($3 \le \lambda \le 6.5$) bands. The additional band displays power transmittance spikes. It should be noted that only the main forbidden band is shown in Fig. 4 from [37]. The range of shorter waves was not considered in [37]. At the same time, the theoretical spectrum of a 2D photon structure formed by solid cylinders in air ($\varepsilon = 10$, $n^* = 0.28\%$) shown in Fig. 3a in [38] has two forbidden bands; the opacity band lying in the range of shorter waves exhibits no transmittance spikes. Their absence can probably be due to the difference in the dielectric contrast and in geometrical parameters of the ordered system of cylinders, such as the number of cylinder layers, the cylinder radius, or the grating period. The values of R and Λ are not given in [38], while the values of the filling factors are ambiguous (the filling factor is 28% in the caption to Fig. 1, which also refers to Fig. 3a, while the factor given at the beginning of Section 3 is two orders of magnitude smaller).

We will interpret the transparency of a 2D photonic crystal on the basis of two effects: Bragg-type macroscopic multiple scattering from a periodic "potential" and Mie microscopic resonant scattering of optical radiation from a single "potential" (cylinder). For this purpose, Fig. 3a shows the frequency dependence of the total scattering cross section per unit length of an infinitely long single cylinder. The dependence calculated on the basis of formula (38) from [39] (dashed curve) has several peaks due to the multipolarity of scattering and labeled accordingly from 0 to 3. It should be noted that the monopole scattering peak (excitation of the

main zero eigenmode) from a cylinder of the given radius and dielectric contrast corresponds to $\lambda = 9.7$; the dipole (peak 1), quadrupole (peak 2), and octupole (peak 3) resonances are excited for $\lambda = 4.67$, 3.01, and 2.22, respectively.

Let us compare the solid and dashed curves in Fig. 3a. It can be seen that the monopole scattering maximum corresponds approximately to the middle of the forbidden band of the photonic crystal spectrum. This correspondence was noted as an unexpected fact in [40], where the method of linear combination of Mie resonances was developed for solving multiple scattering problems for random discrete and ordered 2D media [38]. Our calculations based on Riccati equation (1) indicate a correlation between the excitation frequency for the zero eigenmode of a single cylinder and the position of the main forbidden band in the transmission spectrum for an ordered system of cylinders on the frequency scale. It is important that the correlation is observed for the factor n^* of space filling with cylinders from the interval $3\% \le n^* \le 20\%$, which corresponds to the variation of the cylinder radius in the range $0.4 \le R \le 1$ for a fixed value of $\Lambda = 4$. An analysis shows that the monopole scattering peak (peak 0 in Fig. 3a) is shifted towards longer waves upon an increase in the cylinder radius. The main forbidden band also experiences a similar shift. Outside the interval $0.4 \le R \le 1$, the peak of monopole scattering from a cylinder is displaced from the wavelength range corresponding to the forbidden band to the short-wave ($R < 0.4$) to the long-wave ($R > 1$) region.

Thus, the position of the forbidden band for a 2D photon structure characterized by the filling factor from the range $3\% \le n^* \le 20\%$ on the frequency scale for a given dielectric contrast is determined by the excitation frequency for the zero resonance in a single cylinder. No correlation has been observed between the amplitude and width of the monopole scattering peak and the amplitude and width of the forbidden band. While the monopole scattering peak increases in amplitude and broadens upon an increase in the cylinder radius, the forbidden band in the spectrum of an ordered system of a cylinder is deepened and broadened, passing through the maximal values, and then becomes narrower and shallower.

The resonance scattering conditions $2\pi R/\lambda \ll 1$ and $2\pi R/\lambda_1 \sim 1$, where $\lambda_1 = \lambda/\sqrt{\varepsilon}$, for incident monochromatic radiation by a Mie volume mode excited in a single cylinder give the estimate $\lambda \sim 2\pi R\sqrt{\varepsilon} \sim 10.9$ (we can mention for comparison that the monopole scattering peak is observed for $\lambda \sim 9.7$; dashed curve in Fig. 3a). Let us detune the wavelength from the monopole resonance. Longer waves ($\lambda \ge 12.4$) do not lead to the excitation of cylinder eigenmodes; the cross section of scattering from a single cylinder decreases monotonically; and the grating formed by cylinders becomes virtually transparent (see Fig. 3a). A shift from the monopole resonance frequency towards shorter waves makes the scattering cross section for a cylinder nonmonotonic due to the excitation of higher eigenmodes of the cylinder. It can be seen from Fig. 3a that the spec-

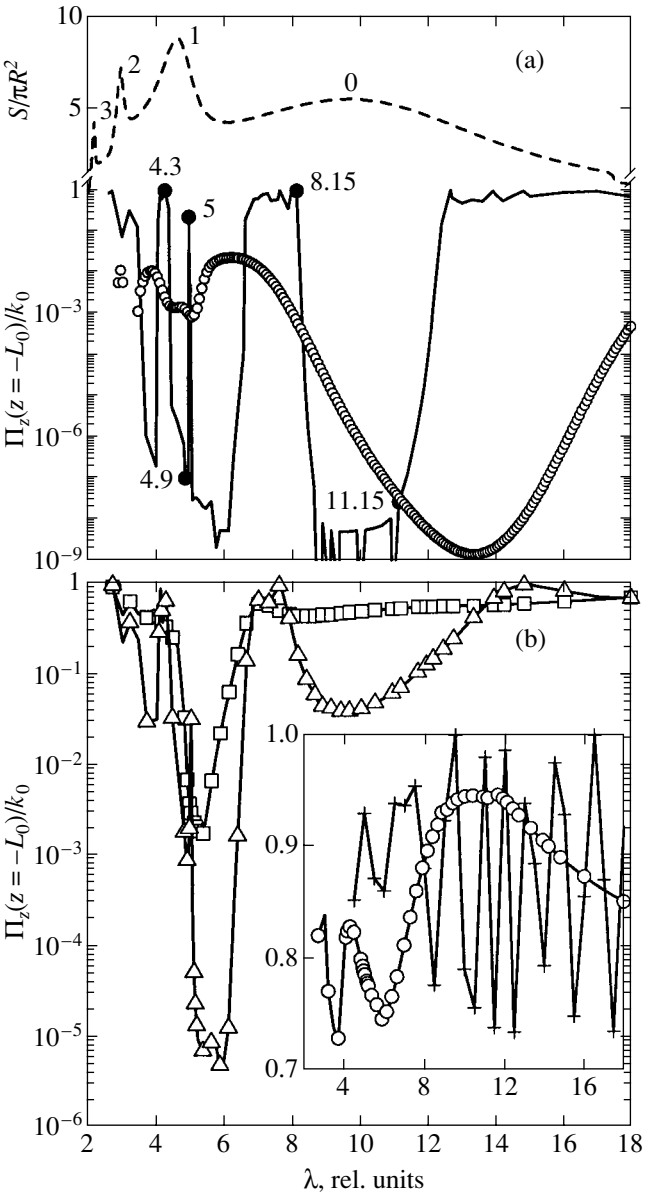


Fig. 3. Power transmission frequency spectra calculated using Eq. (1) and the Poynting theorem for a *TE* polarized wave incident along the normal ($\alpha = 0$) from the medium ($\varepsilon^{bac} = 1$) through N layers of cylinders ($R = 0.6$, $\varepsilon = 2.9^2$, $\Lambda = 4$) without a substrate ($L_o = 0$): (a) $N = 18$ (solid curve); (b) $N = 2$ (□) and $N = 4$ (△); in the inset: $N = 1$ (○) and $N = 18$ (+, evanescent spectral orders are excluded from calculation). The dashed curve in (a) is the frequency dependence of the total scattering cross section S per unit length of an infinitely long single cylinder; the curve with light circles corresponds to the transparency frequency spectrum for a layer of thickness $L = 18\Lambda$ of the medium ($\varepsilon^{bac} = 1$) filled with randomly arranged parallel cylinders with density $n = 1/\Lambda^2$ and with the scattering cross section $S(\lambda)$.

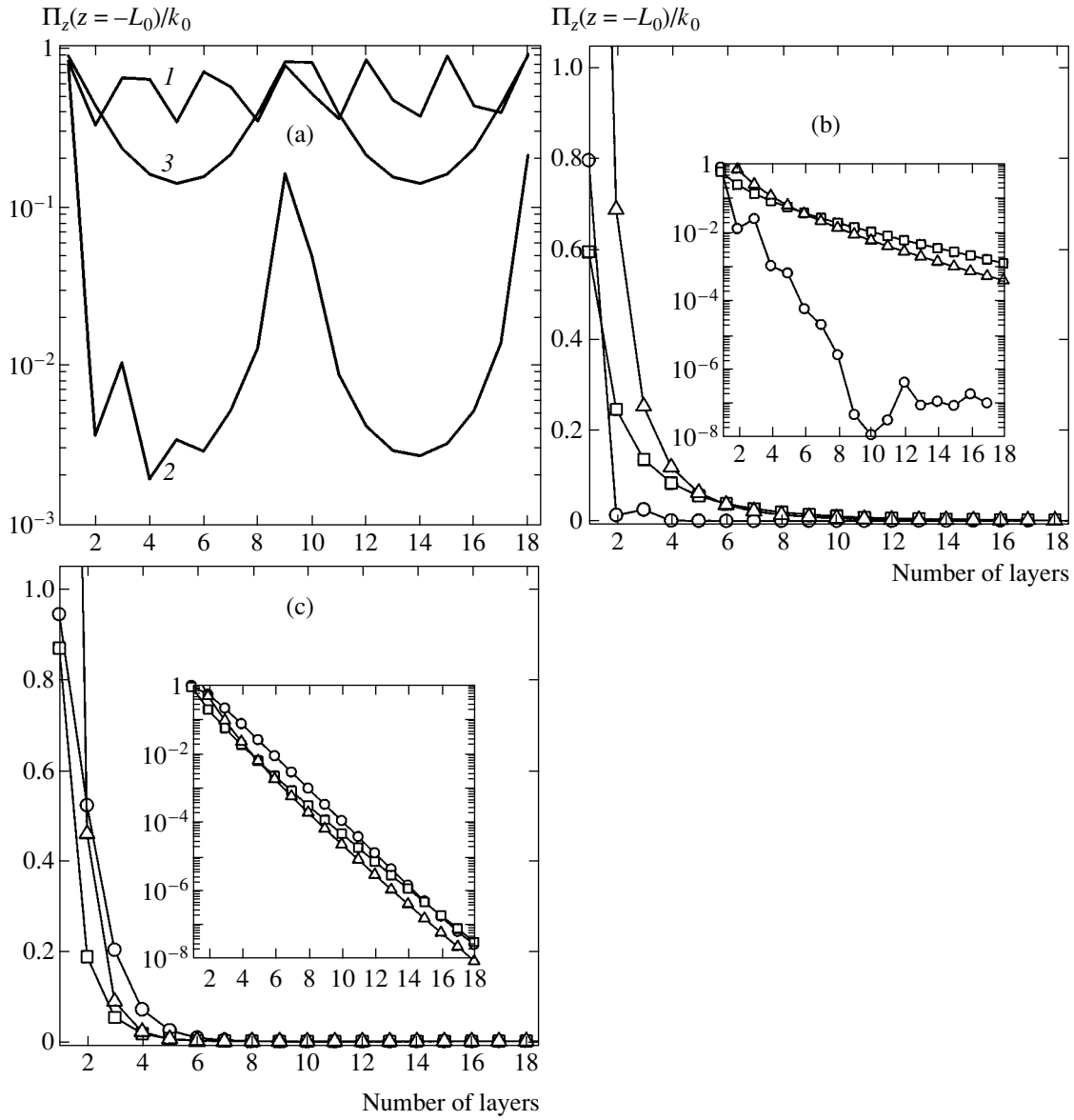


Fig. 4. Dynamics of formation of the frequency spectrum represented by the solid curve in Fig. 3a upon an increase in the number of layers of cylinders for certain characteristic wavelengths: (a) $\lambda = 4.3$ (curve *1*), $\lambda = 5$ (curve *2*), $\lambda = 8.15$ (curve *3*); (b) $\lambda = 4.9$ (○); (c) $\lambda = 11.15$ (○). The curves in (b) and (c) denoted by triangles and squares, respectively, are calculated by formulas (9) for values of parameters given in Fig. 5. The insets to (b) and (c) show the same data on the semilogarithmic scale.

tral width of the dipole scattering peak (peak 1) coincides with the widths of the additional forbidden band for the photon structure. The smooth variation of the scattering cross section for the cylinder between the dipole (peak 1) and monopole (peak 0) resonances ensures the monotonic shape of the right edge of the additional forbidden band. On the contrary, the quadrupole and octupole resonances (peaks 2 and 3 on the dashed curve) on the short-wave slope of the dipole scattering peak lead to a complex structure of the short-wave edge of the additional forbidden band for the photon structure. It should be emphasized that a comparison of the dipole scattering peak and scattering peaks of a higher multipolarity with the additional forbidden band and with the shape of its edges can be rightfully carried out for filling factors of a 2D photonic crystal from the range $3\% \le n^* \le 20\%$.

The above-mentioned correspondence between the excitation conditions for the single cylinder eigenmodes and the formation of forbidden bands in the frequency spectrum of an ordered system of cylinders is one of two reasons for which the two opacity bands for the photon structure were referred to as the main band (zero mode of the cylinder) and the additional band (higher modes). The other reason is associated with the mechanism of formation of these bands upon an increase in the number of layers. Nearly full reflection of waves from the range $8.2 \le \lambda \le 12.4$ is observed (see

Fig. 3b) only for an aggregate of a certain (in fact, not large) number of layers of cylinders; i.e., the main forbidden band is inherent just to a crystal. The additional forbidden band is formed starting from one layer (see the curve with circles in the inset and the curve with squares in Fig. 3b). It should be noted that the transparency spikes (for $\lambda = 4.3$ and 5 in Fig. 3a) in the additional band are observed even for a layer of cylinders (diffraction grating), frequency resolution being improved upon an increase in the number of gratings (see Fig. 3b).

The following four remarks can be made concerning the results shown in Figs. 3 and 4. First, two narrow peaks (at $\lambda = 4.3$ and 5 in Fig. 3a) resemble in appearance the transparency spikes in the forbidden band of photon structures, whose emergence is usually associated with violation of the periodicity of the structures, i.e., introduction of structural defects (see, for example, [40, 37]). In our case, the spectra in Fig. 3 correspond to a perfect 2D crystal. Nevertheless, the transparency spikes at $\lambda = 4.3$ and 5 can be attributed to the emergence of effective structural defects of the following origin. When higher and higher eigenmodes are excited in a cylinder by incident electromagnetic radiation, the field accumulated in the cylinder is concentrated in a thinner and thinner surface layer. This leads to radial separation of the accumulated field. The separation occurring for dipole and higher multipolarity scattering from a cylinder with a constant permittivity can be simulated for monopole scattering from a cylinder with a radially inhomogeneous dielectric function. The existence of inhomogeneous permittivity for some cylinders leads to the emergence of a sort of dynamic defects in the perfect photon structure in a certain frequency range of radiation diffracted from it. It should be noted that the variation of permittivity of scatterers is one of three ways of introducing disorder into photon structures [38].

Second, it can be seen from Fig. 4a that the transparency of a photon structure formed by cylinders at wavelengths $\lambda = 4.3$ and 5 is characterized by different dynamics of variation upon an increase in the number of cylinder layers. Namely, the transparency changes insignificantly at wavelength $\lambda = 4.3$ (curve *1*) and by more than an order of magnitude with a considerably higher periodicity in the number of layers at wavelength $\lambda = 5$ (curve *2*). The same figure shows for comparison a certain intermediate regime (curve *3*).

Third, the inset to Fig. 3b illustrates the role of evanescent spectral orders emerging due to diffraction of a *TE* polarized wave from a periodic 2D structure. The curve with crosses depicts the power transmission spectrum for 18 layers of cylinders in the absence of mutual transformation of propagating and evanescent waves. The mutual transformation of these waves was eliminated in calculations by carrying out summation in the Riccati equation (1) only over propagating spectral orders (as in the Poynting vector component (7)). The calculations demonstrate the decisive role of mutual energy redistribution of the field being diffracted among propagating and evanescent (along the z axis) diffraction orders in the formation of the frequency spectrum with a forbidden band. If the field energy is redistributed only among the propagating modes in the course of multiple scattering of electromagnetic radiation between the cylinders forming the photon structure, the amplitude of transparency variation for 18 layers of cylinders (curve with crosses in the inset to Fig. 3b) is on the order of the amplitude of variation of the transparency of a layer of cylinders (curve with circles), but calculated taking into account evanescent modes also.

The final, fourth, remark is that there is no need to prepare 18 layers of cylinders for the formation of a 2D photon structure in actual practice. Figure 3b illustrates the fact that a spectrally narrow opacity band with a transmission at a level of 0.1% of the power of incident electromagnetic radiation is formed in the transmission spectrum for just two layers (curve with squares); four layers of cylinders (curve with triangles) have a broader and deeper (on the order of 10^{-3}% of the incident radiation power) forbidden band in the wavelength range $5 \le \lambda \le 6$ and a shallow (at a level of ~5%) but broad band in the long-wave part of the spectrum.

Let us now consider the dynamics of formation of opacity bands for a 2D photon structure with increasing number of cylinder layers constituting the structure. The rate of transparency variation in the structure is comparable with the exponential-power decrease in the transmittance of scalar classical waves through a layer of a discrete random medium, which has been derived in the self-consistent theory of localization [36]. Figures 4b and 4c demonstrate the dynamics of formation of forbidden bands (○) in the frequency spectrum of a 2D photonic crystal (solid curve in Fig. 3a) for two waves whose wavelengths correspond to the additional ($\lambda = 4.9$) and main ($\lambda = 11.15$) forbidden bands. The same figures show the results of calculation of the transmission coefficient for waves passing through a layer of a random medium as a function of the number of layers (see formulas (23) and (24) from [36]),

$$T_{\text{loc}} = \begin{cases} (l_e/N)^2 \exp(-N/N_o), & \text{for } \xi > L, \\ (a/N)\exp(-N/N_o), & \text{for } \xi < L, \end{cases} \qquad (9)$$

where $L = N\Lambda$ is the layer thickness, ξ is the localization length, $N_o = \xi/\Lambda$, $a = (l_c - l_e)/\Lambda$, $l_c = \lambda$ is the localization threshold, $l_e = 1/nS$ is the mean free path for elastic scattering, and $n = 1/\Lambda^2$ is the number density of scatterers with the scattering cross section $S(\lambda)$. The transparency frequency spectrum for a random medium, calculated by formulas (9) for $L = 18\Lambda$, is shown by the curve with circles in Fig. 3a. A layer of the random medium is similar to a photonic crystal consisting of $N = 18$ layers of ordered cylinders as regards its thickness, number of

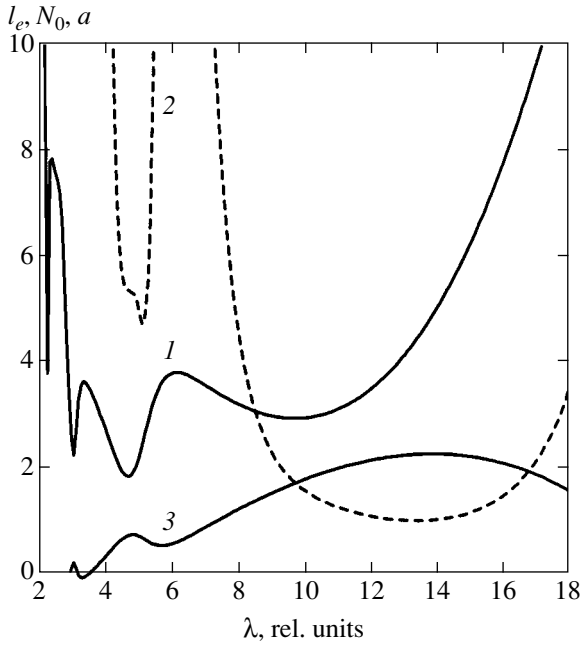


Fig. 5. Dependence of the mean free path l_e of the wave field for elastic scattering (curve *1*) and the localization length ξ (curve *2* corresponds to $N_o = \xi/\Lambda$) on the wavelength of radiation incident on a layer of a random medium, calculated by formulas (49) and (68) from [39] on the basis of the self-consistent theory of localization of acoustic waves. Curve *3* illustrates the variation of parameter $a = (\lambda - l_e)/\Lambda$.

cylinders, and the mean spacing between the cylinders. The dependences of quantities l_e, N_o, and parameter a on the wavelength of the incident wave are shown in Fig. 5.

A comparison of the solid curve and the curve with circles in Fig. 3a should be preceded by the following remark. Experiments with an ordered system of solid cylinders in air (see Fig. 2 from [41]) and calculations for cylindrical bubbles in water (see Fig. 2c from [42]) showed that the introduction of positional disorder into the system of cylinders narrows the forbidden band for the *TE* polarized radiation; this band becomes barely shallower in the central part in the case of solid cylinders and becomes slightly shallower in the opposite case. It can be seen from Fig. 3a that the results of calculation of transparency for a layer of a random medium, based on formulas (9) of the self-consistent localization theory, generally agree with the abovementioned results obtained in [41] as well as the calculations based on Riccati equation (1). In the transparency spectrum for a random medium layer, we can also single out two forbidden bands, while singularities of the type of local transparency spike can be indicated in the additional band.

Analyzing the dynamics of formation of the forbidden band in the spectrum of a layer of ordered and disordered 2D media, we note that the ordering of cylinders in the short-wave range ($\lambda = 4.9$) is responsible for a much more rapid decay of the transmitted power of electromagnetic field as compared to the exponential-power law (9) (see Fig. 4b). At longer waves, the agreement with the self-consistent theory of localization is much better and even perfect for some wavelengths (see Fig. 4c). It should be noted that the spectrum of the ordered structure in the frequency range corresponding to the main forbidden band (the reasons for using this term were given above) is formed due to resonant scattering at the fundamental Mie eigenmode of the cylinder with a considerable contribution from the resonant Bragg-type scattering from the periodic system of cylinders. It would be natural to expect the onset of localization of classical waves in a random medium exactly under the conditions for the existence of the main forbidden band after the introduction of disorder into the periodic system of cylinders. The conservation of a certain extent of long-range disorder after the violation of short-range ordering will ensure a Bragg-type scattering and suppression of coherent backward scattering of photons in certain Bragg directions. The "transfer relations" with Riccati equations following from them will make it possible (after a certain refinement) to calculate, with controllable accuracy, the frequency spectrum of transparency of a layer of an ordered 2D medium with any extent of disorder.

5. CONCLUSIONS

The method of transfer relations is applied to a unified theoretical treatment of resonant effects of multiple scattering of electromagnetic radiation from a periodic surface, and of transmission through a photonic crystal. For the first time, the effect of anomalous decrease in the reflection coefficient for a *TE* polarized wave in discrete ranges of the optical grating profile depth is demonstrated quantitatively. For a periodic system of parallel layers of solid cylinders with permittivity $\varepsilon = (2.9)^2$ in air, it is established that the formation of the short-wave forbidden band in the transmission spectrum of a photonic crystal (including narrow transparency spikes) is mainly due to multipole Mie resonances of a single cylinder. On the contrary, Bragg-type multiple scattering makes a significant contribution to the formation of the long-wave forbidden band; the position of the center of this band on the frequency scale correlates with the frequency of the monopole Mie resonance of a single cylinder in the range $3\% \le n^* \le 20\%$ of the factor of filling the space with the cylinders. The electromagnetic radiation decay in a photonic crystal with increasing number of its layers is satisfactorily described by an exponential-power formula from the self-consistent localization theory in the wavelength range corresponding to the long-wave forbidden band. However, ordering of the cylinders in the short-wave forbidden band leads to a much faster damping of radiation.

REFERENCES

1. P. Beckmann and A. Spizzichino, *The Scattering of Electromagnetic Waves from Rough Surfaces* (MacMillan, New York, 1963).
2. F. G. Bass and I. M. Fuks, *Wave Scattering from Statistically Rough Surfaces* (Nauka, Moscow, 1972; Pergamon, New York, 1978).
3. V. I. Tatarskii, Waves Random Media **3**, 127 (1993).
4. L. Tsang, I. A. Kong, and R. T. Shin, *Theory of Microwave Remote Sensing* (Wiley–Interscience, New York, 1985).
5. A. G. Voronovich, *Wave Scattering from Rough Surfaces* (Springer, Berlin, 1994).
6. M. L. Goldberger and K. M. Watson, *Collision Theory* (Wiley, New York, 1964; Mir, Moscow, 1967).
7. R. Bellman and G. M. Wing, *An Introduction to Invariant Imbedding* (Wiley–Interscience, New York, 1975).
8. V. I. Klyatskin, *The Embedding Method in the Theory of Wave Propagation* (Nauka, Moscow, 1986).
9. C. Barnes and J. B. Pendry, Proc. R. Soc. London, Ser. A **435**, 185 (1991).
10. R. W. Wood, Proc. Phys. Soc. London **18**, 396 (1902); Philos. Mag. **4**, 396 (1902).
11. H. Raether, *Surface Plasmons on Smooth and Rough Surfaces and on Gratings* (Springer, Berlin, 1988).
12. D. Maystre and R. Petit, Opt. Commun. **17**, 196 (1976).
13. E. Burstein, C. Y. Chen, and S. Lundqvist, *Light Scattering in Solids*, Ed. by J. L. Birman, H. Z. Cummis, and K. K. Rebane (Plenum, New York, 1979), p. 479.
14. I. Ursu, I. N. Mihăilescu, Al. Popa, *et al.*, Appl. Phys. Lett. **45**, 365 (1984).
15. J. A. Sanchez-Gil, A. A. Maradudin, Jun Q. Lu, *et al.*, Phys. Rev. B **50**, 15353 (1994).
16. D. Maystre, *Progress in Optics*, Ed. by E. Wolf (North-Holland, Amsterdam, 1984), Vol. 21, p. 1.
17. A. A. Maradudin, T. Michel, A. R. McGurn, and E. R. Mendez, Ann. Phys. (N.Y.) **203**, 255 (1990).
18. A. Hessel and A. A. Oliver, Appl. Opt. **4**, 1275 (1965).
19. G. M. Gandel'man and P. S. Kondatenko, Pis'ma Zh. Éksp. Teor. Fiz. **38**, 246 (1983) [JETP Lett. **38**, 291 (1983)].
20. K.-T. Lee and T. F. George, Phys. Rev. B **31**, 5106 (1985).
21. S. A. Akhmanov, V. N. Seminogov, and V. I. Sokolov, Zh. Éksp. Teor. Fiz. **93**, 1654 (1987) [Sov. Phys. JETP **66**, 945 (1987)].
22. F. Toigo, A. Marvin, C. Celli, and N. R. Hill, Phys. Rev. B **15**, 5618 (1977).
23. E. Yablonovitch, Phys. Rev. Lett. **58**, 2059 (1987); S. John, Phys. Rev. Lett. **58**, 2486 (1987).
24. K. M. Ho, C. T. Chan, and C. M. Soukoulis, Phys. Rev. Lett. **65**, 3152 (1990).
25. M. Plihal and A. A. Maradudin, Phys. Rev. B **44**, 8565 (1991).
26. K. Sakoda, Phys. Rev. B **51**, 4672 (1995).
27. N. Stefanou, V. Karathanos, and A. Modinos, J. Phys.: Condens. Matter **4**, 7389 (1992).
28. K. Ohtaka and Y. Tanabe, J. Phys. Soc. Jpn. **65**, 2265 (1996); J. Phys. Soc. Jpn. **65**, 2276 (1996).
29. S. Y. Tong, Prog. Surf. Sci. **7**, 1 (1975).
30. J. B. Pendry and A. Mackinnon, Phys. Rev. Lett. **69**, 2772 (1992).
31. M. Sigalas, C. M. Soukoulis, E. M. Economou, *et al.*, Phys. Rev. B **48**, 14121 (1993).
32. Y. Zhao, I. A. Avrutsky, and B. Li, Appl. Phys. Lett. **75**, 3596 (1999).
33. Yu. N. Barabanenkov, V. L. Kouznetsov, and M. Yu. Barabanenkov, *Progress in Electromagnetic Research, PIER*, Ed. by J. A. Kong (EMW, Cambridge, 1999), Vol. 24, p. 39; J. Electromagn. Waves Appl. **13**, 1335 (1999).
34. R. W. Wood, Philos. Mag. **23**, 315 (1912); Phys. Rev. **48**, 934 (1935).
35. C. H. Palmer, J. Opt. Soc. Am. **42**, 269 (1952).
36. D. Sornette, J. Stat. Phys. **56**, 669 (1989).
37. E. Centeno, B. Guizal, and D. Felbacq, J. Opt. A: Pure Appl. Opt. **1**, L10 (1999).
38. E. Lidorikis, M. M. Sigalas, E. N. Economou, and C. M. Soukoulis, Phys. Rev. B **61**, 13458 (2000).
39. C. A. Condat and T. R. Kirkpatrick, *Scattering and Localization of Classical Waves in Random Media*, Ed. by Ping Sheng (World Sci., Singapore, 1990), Vol. 8, p. 423.
40. E. Lidorikis, M. M. Sigalas, E. N. Economou, and C. M. Soukoulis, Phys. Rev. Lett. **81**, 1405 (1998).
41. M. Bayindir, E. Cubukcu, I. Bulu, *et al.*, Phys. Rev. B **64**, 195113 (2001).
42. Y.-Y. Chen and Z. Ye, Phys. Rev. E **65**, 056612 (2002).

Translated by N. Wadhwa

Journal of Experimental and Theoretical Physics, Vol. 96, No. 4, 2003, pp. 684–694.
Translated from Zhurnal Éksperimental'noĭ i Teoreticheskoĭ Fiziki, Vol. 123, No. 4, 2003, pp. 775–786.

PLASMA, GASES

Mechanism of Diffusion of Positively Charged Dust Particles in a Photoemission Cell under Microgravity Conditions

A. V. Filippov[a,*], V. E. Fortov[b], A. F. Pal'[a], and A. N. Starostin[a]

[a] *Troitsk Institute for Innovation and Fusion Research, Troitsk, Moscow oblast, 142190 Russia*

[b] *Institute of High Energy Densities (IVTAN), Russian Academy of Sciences, Moscow, 127412 Russia*

**e-mail: fav@triniti.ru*

Received October 31, 2002

Abstract—The decay of a dusty plasma in a photoemission cell under microgravity conditions is investigated on the basis of the method of nonlocal moments. It is founds that plasma decay in space experiments occurs in accordance with the mechanism of free electron diffusion followed by dust particle drift. An analytic solution is found for the evolution of radial distributions of the dust particle concentration and the electric field under the experimental conditions. The effect of abnormally high temperatures of dust particles is considered. The effect of axial magnetic fields on the decay of dusty plasma is investigated. It is shown that the plasma decay in a magnetic field is governed by the ambipolar diffusion mechanism, the decay being prolonged up to 10^3 s in a magnetic field on the order of 10^3–10^4 G in strength. © *2003 MAIK "Nauka/Interperiodica".*

1. INTRODUCTION

Abnormally high temperatures of dust particles in a complex plasma were reported in a number of recent publications [1–11]. It is well known [12] that the energy of heavy ions (and, accordingly, charged dust particles) can "break away" from the thermal energy in strong fields due to a drift with velocities much higher than thermal velocities. This is due to the fact that the mobility of dust particles is very high, for two reasons: (1) the smallness of reduced masses of a dust particle and a gas particle and (2) the large values of the dust particle charge. Consequently, the velocity of dust particles may exceed the thermal velocity even in weak electric fields. For example, bronze particles of radius 25 μm, which were used in experiments [1] and whose charge was estimated as $z_d = 4 \times 10^4 e$ (e is the electron charge), have equal drift and thermal velocities in Ne under a pressure of 40 Torr in a field $E \sim 0.1$ V/cm. As the field increases, the energy of dust particles increases quadratically and can easily become as high as tens of electronvolts. However, this is the kinetic energy of directional (and not random) motion; for this reason, the diffusion of, say, dust particles cannot be determined by this high kinetic energy. In almost all of the above-mentioned publications [2–9], the motion of particles was studied in the region of strong violation of plasma quasineutrality; consequently, high energies of particles are associated in all probability either with their drift or with vibrational motion due to the development of dust–acoustic instabilities.

Reports on abnormally high temperatures of dust particles in experiments staged on the Mir space station [1, 10, 11] proved to be most demonstrative. Under the action of solar radiation with a radiance temperature of 0.52 eV, the dust component was "heated" up to temperatures of 20–55 eV. Different temperatures were obtained in different directions. This study is devoted to an analysis of the experimental results [1] on the basis of a more complete model of plasma decay taking into account the nonlocal effects of the electron energy distribution function (EEDF). With a view to future experiments, the effect of external magnetic field transverse to the radial electric field of plasma polarization on the dusty plasma decay in such a cell was also analyzed.

2. DESCRIPTION OF THE MODEL

Let us briefly describe the conditions of experiments [1], analysis of which led to abnormally high temperatures of dust particles [10, 11]. Bronze dust particles of radius 25–50 μm, coated with a layer of cesium, were exposed to solar radiation on the Mir space station under microgravity conditions. Dust particles were in a cylindrical cell of radius R = 1.5 cm and height h = 6 cm. The most important results were obtained in Ne under a pressure p = 40 Torr. Henceforth, we will speak only of these experiments. The expansion of a dusty plasma consisting of heavy ions, positively charged dust particles with a charge $z_d \approx 4 \times 10^4 e$, and electrons was analyzed in [10, 11] on the basis of the ambipolar diffusion model with an electron temperature of 1–2 eV.

It should be noted that diffusion of dust particles makes a negligibly small contribution to the ambipolar diffusion coefficient

$$D_a = \frac{\mu_d}{e}\left(T_e + \frac{T_d}{z_d}\right) \approx \frac{\mu_d T_e}{e} \quad (1)$$

and cannot be determined from the ambipolar diffusion. Here, e is the absolute value of the electron charge; μ_d, ez_d, and T_d are the mobility, charge, and temperature of dust particles, respectively; and T_e is the electron temperature.

Allis and Rose [13], who studied the transition from free to ambipolar diffusion of charged particles in a steady-state plasma, showed that the equality of the electron Debye length and the diffusion length can be conditionally taken as the boundary between these regimes. It should be noted that, when this condition is satisfied, the electron concentration at the center is half the ion concentration. Under the experimental conditions [1], the electron Debye length is comparable to the diffusion length, and the validity of the assumption made in [10, 11] concerning the ambipolar regime of electron diffusion under the given conditions [1] is not obvious and requires verification.

Analysis of the form of the EEDF in the cell is complicated and requires separate experimental and theoretical studies. Here, we assume that the EEDF is of the Maxwell type. For the electron temperature T_e = 0.52 eV in neon with p = 40 Torr, the electron mean free path is $l_e \approx 38$ μm and the electron energy relaxation length is $l_u \approx 0.92$ cm. As the electron temperature decreases, these quantities increase monotonically, but not more than by a factor of 3 upon a transition from solar to room temperature. Consequently, the hydrodynamic description can be used for transfer processes in a photoemission plasma, but the nonlocal nature of the EEDF must be taken into account. For this purpose, we use here the nonlocal method of moments [14], which includes the balance equation for charged particles, the electron energy balance equation, and the Poisson equation for a self-consistent field. This model is exact in the case of the Maxwellian EEDF. The cell height in experiments was four times larger than cell radius; for this reason, in the first approximation, we solved the one-dimensional problem in the cylindrical system of coordinates. The characteristic time of dust particle charge stabilization under the action of solar radiation with the parameters given in [1] amounts to a value on the order of 0.1–1 μs. Consequently, analyzing plasma decay on a time scale of the order of seconds, we assume that the charge of dust particles remains unchanged; accordingly, the processes of production and loss of electrons and production and recombination of dust particle charge are in equilibrium. As a result, the self-consistent system of equations describing the decay of a dusty plasma in a photoemission cell, which is constructed on the basis of the nonlocal method of moments, assumes the form (see, for example, [15])

$$\begin{aligned}
\frac{\partial n_e}{\partial t} + \frac{1}{r}\frac{\partial (r j_{e,r})}{\partial r} &= 0,\\
\frac{\partial n_d}{\partial t} + \frac{1}{r}\frac{\partial (r j_{d,r})}{\partial r} &= 0,\\
\frac{\partial n_e \langle \varepsilon_e \rangle}{\partial t} + \frac{1}{r}\frac{\partial (r h_{e,r})}{\partial r} + e j_{e,r} E_r &= -n_e W_s + Q,\\
\frac{1}{r}\frac{\partial (r E_r)}{\partial r} &= 4\pi e (z_d n_d - n_e),
\end{aligned} \quad (2)$$

where n_e, n_d, $j_{e,r}$, and $j_{d,r}$ are the concentrations and densities of radial flows of electrons and charged dust particles, respectively; $\langle \varepsilon_e \rangle$ is the mean electron energy, which is equal to $(3/2)T_e$ for a Maxwellian EEDF; T_e is the electron temperature; $h_{e,r}$ is the density of the radial electron energy flux; W_s is the energy loss rate in elastic and inelastic collisions; E_r is the radial electric field strength; and Q is the source of electron heating.

In the diffusion-drift approximation, the fluxes are defined as

$$\begin{aligned}
j_{e,r} &= -\frac{\partial (D_T n_e)}{\partial r} - k_e n_e E_r,\\
j_{d,r} &= n_d \mu_d E_r - D_d \frac{\partial n_d}{\partial r} \approx n_d \mu_d E_r,\\
h_{e,r} &= -\frac{\partial (G n_e)}{\partial r} - \beta n_e E_r,
\end{aligned} \quad (3)$$

where k_e, μ_d, and D_T, D_d are the mobilities and the diffusion coefficients for electrons and dust particles, respectively; G is the electron energy diffusion coefficient; and β is a coefficient called the thermoelectric coefficient in [14] (in [16], the term "thermoelectric coefficient" is applied to the quantity β/ek_eT_e). For a Maxwellian EEDF, the electron transfer coefficients are connected through the Einstein relations:

$$\frac{D_T}{k_e} = \frac{G}{\beta} = \frac{T_e}{e}. \quad (4)$$

The mobility of dust particles was determined by the Stokes–Einstein formula with the Cunningham correction to the finiteness of the mean free path of the neon atom [17]:

$$\mu_d = ez_d \frac{1 + (l_g/r_d)(1.257 + 0.4\exp\{-1.1 r_d/l_g\})}{6\pi\eta r_d}, \quad (5)$$

where l_g is the mean free path for neon atoms, r_d is the dust particle radius, and η is the neon viscosity, which is virtually independent of pressure and is equal to 3.17×10^{-5} Pa s at room temperature [18]. Table 1 contains the mobilities and diffusion coefficients for dust particles with different radii. The same table contains

Table 1. Mobility $\mu_{d,z=1}$ of dust particles with unit charge, diffusion coefficient D_d, and distance $l_{100} = (2D_d t)^{1/2}$ traversed due to diffusion during time $t = 100$ s for different radii r_d for the dust component temperature $T_d = 300$ K

r_d, μm	25	37.5	50
$\mu_{d,z} = 1$, $cm^2/(V\,s)$	1.284×10^{-7}	8.086×10^{-8}	5.887×10^{-8}
D_d, cm^2/s	3.322×10^{-9}	2.092×10^{-9}	1.523×10^{-9}
l_{100}, μm	8.0	6.5	5.5
l_{100} ($T_d = 55$ eV), cm	0.038	0.030	0.025

the distances over which dust particles with temperature $T_d = 300$ K move due to diffusion over 100 s. In the experiment [1], dust particles during this time drift over a distance comparable to a cell radius of $R = 1.5$ cm. Consequently, we can disregard the diffusion of dust particles under the experimental conditions [1] under a pressure of 40 Torr. (Diffusion lengths turn out to be negligibly small as compared to drift lengths even for $T_d = 55$ eV.)

It was shown in [19] that a fairly wide charge distribution of dust particles is formed due to photoemission. In addition, polydisperse dust particles with radii of 25–50 μm were used in experiments [1, 10]. Let us prove that the decay of a dusty plasma in a photoemission cell with polydisperse particles can be described on the basis of the model for monodisperse particles.

Let $f_{a,z}$ be the distribution function for dust particles over charges z and sizes a, which is normalized by the condition

$$\sum_z \int f_{a,z}\,da = 1. \tag{6}$$

As the radius of dust particles changes in the range from 25 to 50 μm, the Cunningham correction in neon under a pressure of 40 Torr decreases monotonically from 1.2 to 1.1. Consequently, we can disregard the weak dependence of the mobility of dust particles on the ratio of the mean free path for atoms to the radius. In this approximation, the mobility of a dust particle of radius a and charge z is defined as

$$\mu_{a,z} = \frac{ez}{6\pi\eta a}. \tag{7}$$

We write the balance equation for the number of dust particles with a radius in the range from a to $a + da$ and charge z:

$$\frac{\partial(n_d f_{a,z})}{\partial t} + \frac{1}{r}\frac{\partial(r\mu_{a,z}E_r n_d f_{a,z})}{\partial r} = R_{a,z}, \tag{8}$$

where $R_{a,z}$ is the sum of the sources of formation and destruction of plasma particles of class (a, z). We disregard the coagulation of dust particles at the plasma decay stage. In this case, the number of particles of class (a, z) will change only due to the change in their charge. Summing Eqs. (8) over all possible charges and taking into account the fact that the number of particles of a given size remains unchanged, we obtain

$$\frac{\partial(n_d f_a)}{\partial t} + \frac{1}{r}\frac{\partial(r\mu_a E_r n_d f_a)}{\partial r} = 0, \tag{9}$$

where $f_a = \sum_z f_{a,z}$, μ_a is the mobility of dust particles of radius a and charge z_a equal to the mean charge of particles having the given size:

$$z_a = \frac{1}{f_a}\sum_z z f_{a,z}. \tag{10}$$

Integrating over the sizes of particles, we derive from Eq. (9) continuity equation (2) with flux (3), in which the mobility is defined as

$$\mu_d = \frac{eq_0}{6\pi\eta}. \tag{11}$$

Here, we have introduced the charge-to-radius ratio for a dust particle, averaged over the ensemble:

$$q_0 = \int \frac{z_a}{a} f_a\,da = \int \sum_z \frac{z}{a} f_{a,z}\,da. \tag{12}$$

The Poisson equation for a polydisperse system of dust particles can be reduced to the same form as in system (2) by introducing the averaged charge of dust particles:

$$z_d = \int \sum_z z f_{a,z}\,da = \int z_a f_a\,da. \tag{13}$$

For the sake of convenience, we introduce the effective radius of dust particles through the relation

$$r_d = z_d/q_0,$$

which, by the way, does not coincide with the averaged radius

$$\langle a \rangle = \int \sum_z a f_{a,z}\,da = \int a f_a\,da.$$

System of equations (2) was solved with the effective boundary conditions to the balance equations for

the electron number density and energy (see [15, 20, 21] for details):

$$\left[n_e + \gamma_e l_e \frac{\partial n_e}{\partial r}\right]_{r=R} = 0, \quad \left.\frac{\partial T_e}{\partial r}\right|_{r=R} = 0,$$
$$j_{e,r}|_{r=0} = 0, \quad j_{d,r}|_{r=0} = 0, \quad (14)$$
$$h_{e,r}|_{r=0} = 0, \quad E_r|_{r=0} = 0,$$

where l_e is the electron mean free path and γ_e is the Hopf constant. The initial conditions were specified in the form

$$n_e|_{t=0} = z_d n_{d,0}, \quad T_e|_{t=0} = T_{e,0}, \quad n_d|_{t=0} = n_{d,0}. \quad (15)$$

In experiments, the rate of electron departure to the walls,

$$\nu_{e,w} \sim D_e/5.7R^2 = (3\text{–}5) \times 10^3 \text{ s}^{-1},$$

was much lower than the rate of electron departure to dust particles,

$$\nu_{e,d} \sim \pi r_d^2 n_d v_{th,e} = (4\text{–}9) \times 10^5 \text{ s}^{-1},$$

where $v_{th,e}$ is the thermal velocity of electrons. Consequently, electrons escape to the walls as a result of diffusion from one dust particle to another. Having absorbed an electron that lost a part of its energy in collisions with buffer gas atoms, a dust particle exposed to light immediately emits another, "hot" electron. Consequently, the loss of electron energy under the condition $a \ll l_u$ is directly proportional to the square of the ratio a/l_u (the ratio of the mean distance $a = 1/\sqrt[3]{n_d}$ between particles to the electron energy relaxation length). In zero magnetic field, this condition is satisfied in the cell; for this reason, we disregarded the electron energy losses in collisions with atoms in our calculations for zero magnetic field. For this purpose, we specified an effective source of electron heating in balance equation (2),

$$Q = n_e W_s(T_{e,0}),$$

which ensured the preset electron temperature $T_{e,0}$ at the center of the photocell.

3. RESULTS OF SIMULATION OF DUSTY PLASMA DECAY IN ZERO MAGNETIC FIELD

Calculations were made with electron temperatures of $T_{e,0}$ = 1000, 3000, and 6000 K. The transfer coefficients and electron energy loss rate in Ne were calculated using a Maxwellian EEDF and cross sections from [22, 23]. The results of calculation of the evolution of dust particle concentration in the central region of the cell, averaged over radius in the region of $r <$ 0.5 cm, are given in Fig. 1. The data obtained in [10, 11] are also given in figures. It can be seen from Fig. 1 that the results obtained using our model are in good agreement with experiment. The only fitting parameter in calculations was the dust particle charge. The values of charge, which are in good agreement with the estimates made in [10, 11], are given in Table 2 for an ensemble of particles with effective radii of 25, 37.5, and 50 μm. A certain increase in the dust particle charge upon a decrease in the dust particle concentration can be due to a suppressed effect of neighbors on the electron concentration in the vicinity of a dust particle. It should be noted that the charge depends on the predicted value of the effective radius of dust particles so that it ensures the experimental decay rate for a dusty plasma (see relation (19) below).

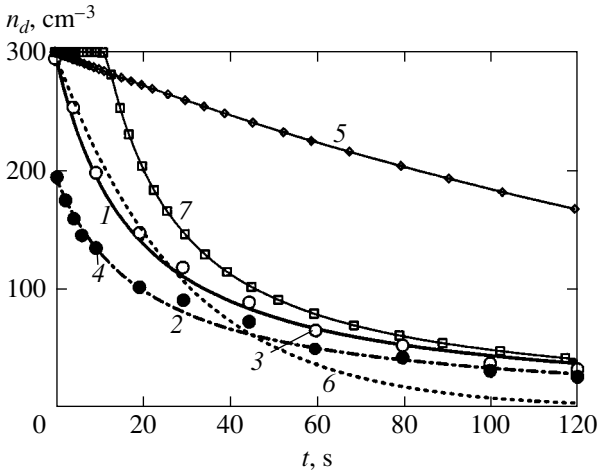


Fig. 1. Experimental and calculated time dependences of dust particle concentration in the central region of a photoemission cell: $T_{e,0}$ = 6000 K, r_d = 25 μm, $n_{d,0}$ = 300 cm^{-3}, $z_d = 2.9 \times 10^4$ (*1*); $T_{e,0}$ = 6000 K, r_d = 25 μm, $n_{d,0}$ = 195 cm^{-3}, $z_d = 3.2 \times 10^4$ (*2*); experimental results [10, 11]: $n_{d,0}$ = 300 (*3*) and 195 cm^{-3} (*4*); dependences calculated by formula (16) for $\nu_a = 4.84 \times 10^{-3}$ s^{-1} (*5*) and $\nu_a = 0.035$ s^{-1} (*6*); $T_{e,0}$ = 1000 K, r_d = 25 μm, $n_{d,0}$ = 300 cm^{-3}, $z_d = 2.9 \times 10^4$ (*7*).

Figure 1 also shows the plasma decay curve corresponding to the ambipolar mechanism:

$$n_d = n_{d,t=0} \exp(-\nu_a t), \quad (16)$$

where $\nu_a \approx 5.7 D_a/R^2 \approx 4.84 \times 10^{-3}$ s^{-1} is the rate of departure of charged particles due to ambipolar diffu-

Table 2. Average charge of an ensemble of dust particles with different effective radii obtained from simulation of the dusty plasma decay for two values of dust particle concentration

r_d, μm	25	37.5	50
$n_{d,0}$ = 300 cm^{-3}	2.9×10^4	3.6×10^4	4.3×10^4
$n_{d,0}$ = 195 cm^{-3}	3.2×10^4	4.0×10^4	4.8×10^4

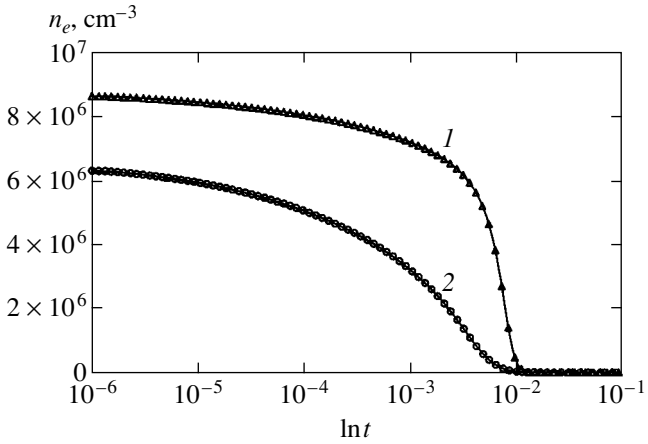


Fig. 2. Calculated time dependences of the electron concentration at the photoemission cell axis for different concentrations of dust particles: $T_{e,0}$ = 6000 K, r_d = 25 μm, $n_{d,0}$ = 300 cm^{-3}, $z_d = 2.9 \times 10^4$ (*1*) and $T_{e,0}$ = 6000 K, r_d = 25 μm, $n_{d,0}$ = 195 cm^{-3}, $z_d = 3.2 \times 10^4$ (*2*).

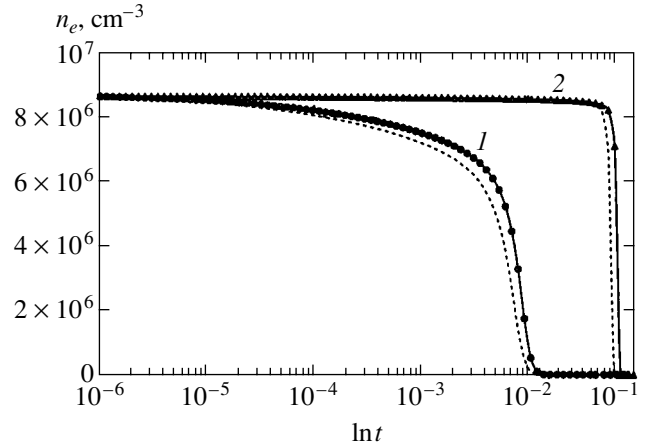


Fig. 3. Calculated time dependences of the electron concentration at the photoemission cell axis for different electron temperatures: $T_{e,0}$ = 6000 K, r_d = 25 μm, $n_{d,0}$ = 300 cm^{-3}, $z_d = 2.9 \times 10^4$ (*1*) and $T_{e,0}$ = 3000 K (*2*); dotted curves correspond to T_e = const.

sion with coefficient D_a defined by expression (1) with $T_{e,0}$ = 6000 K. It was mentioned in [10, 11] that it is possible to match expression (16) with experiment at the initial stage with $\nu_a = 0.035$ s^{-1}. Such a rate of departure of dust particles leads to a value of the product of $z_d T_e = 1.6 \times 10^5$ eV, which means that the electron temperature T_e = 4 eV for $z_d = 4 \times 10^4$ (this is deprived of any physical meaning). It should also be noted that the solution of the time-dependent equation for ambipolar diffusion in an infinitely long cylinder, which satisfies the initial condition

$$n_d|_{t=0} = \text{const},$$

contains not only the fundamental mode with the first zero of the zero-order Bessel function at the boundary, but also modes corresponding to the second, etc., zeros. Higher modes relax and, accordingly, the field penetrates the central region of the cell only by a time instant on the order of R^2/D_a. Consequently, the solution is described by the simple exponential function (3) only for large times (or provided that the profiles of charged particles are already in the form of a zero-order Bessel function with the first zero at the boundary for $t = 0$).

Figure 1 shows that, under conditions of experiments [1], the electron diffusion followed, in all probability, the free and not the ambipolar diffusion mechanism. It can be see from Fig. 2 that almost all electrons leave the cell volume at T_e = 6000 K by the time $t \sim 10^{-2}$ s. Even the decrease of electron temperature to 3000 K does not change the plasma decay pattern; only the electron departure time increases insignificantly to $t \sim 10^{-1}$ s (Fig. 3). It should be noted that, in calculations with a constant electron temperature, i.e., without solving the electron energy balance equation, the departure of electrons is slightly faster (see Fig. 3). During such periods of time, dust particles practically remain stationary. Further dispersion of dust particles is determined by the system of equations, comprising the balance equation for the number of dust particles and the Poisson equation:

$$\begin{aligned} \frac{\partial n_d}{\partial t} + \frac{1}{r}\frac{\partial}{\partial r}(r\mu_d n_d E_r) &= 0, \\ \frac{1}{r}\frac{\partial}{\partial r} r E_r &= 4\pi e z_d n_d. \end{aligned} \tag{17}$$

For the initial distribution of dust particles, which is independent of the radius, the solution of system (17) can be found easily by separating the variables:

$$\begin{aligned} E_r &= 2\pi e z_d n_d r, \\ n_d(r, t) &= \frac{\nu_d}{4\pi e z_d \mu_d}(1 + \nu_d t)^{-1}. \end{aligned} \tag{18}$$

It can be seen from these relations that a decrease in the dust particle concentration follows the hyperbolic law (it was noted in [1] that an equilateral hyperbola successfully describes the experimental curve for evolution of the dust particle concentration). Using initial condition (15), we can find from Eqs. (18) that

$$\nu_d = 4\pi e z_d \mu_d n_{d,0} = 4\pi e z_d^2 \mu_{d,z=1} n_{d,0}. \tag{19}$$

It can be seen that the rate of departure of dust particles is proportional to the square of the dust particle charge; consequently, the ensemble-averaged charge of dust particles can be determined to a high degree of accuracy from the experimental rate value.

Solution (18) shows that our problem of dispersion of dust particles in a photoemission cell is equivalent to

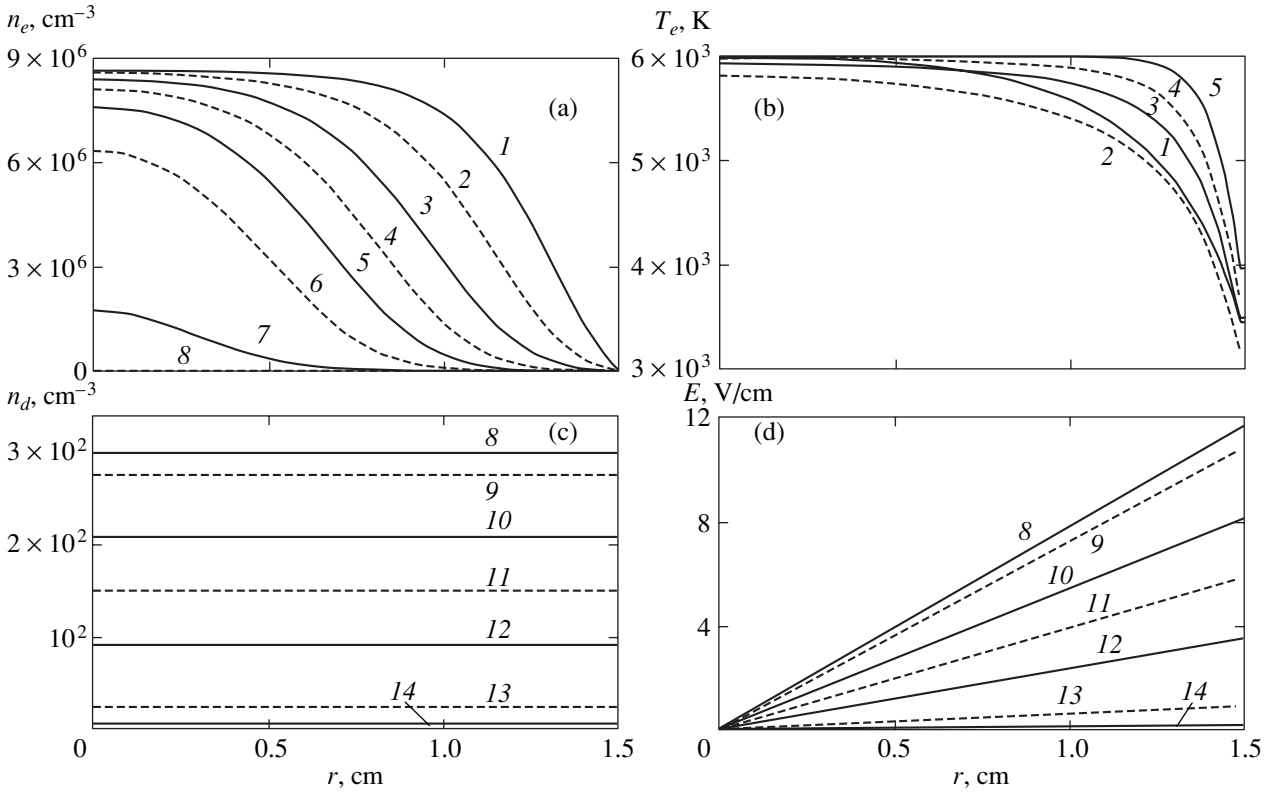


Fig. 4. Evolution of (a) concentration, (b) electron temperature, (c) dust particle concentration, and (d) electric field distributions over the photoemission cell radius for $T_{e,0}$ = 6000 K, r_d = 25 μm, $n_{d,0}$ = 300 cm^{-3}, $z_d = 2.9 \times 10^4$, and t = 1 (*1*), 5.3 (*2*), 28 (*3*), 150 μs (*4*), 0.8 (*5*), 4.1 (*6*), 9.4 (*7*), 22 ms (*8*), 1.5 (*9*), 7.6 (*10*), 17.2 (*11*), 40 (*12*), 200 (*13*), and 1000 s (*14*).

the problem of recombination decay of a plasma consisting of positive and negative ions with the same charge z_d. The recombination coefficient in the case when drift motion prevails over diffusion is determined by the Langevin theory, according to which [12]

$$\beta_L = 4\pi e z_d \mu_d.$$

The recombination decay of the plasma follows the law

$$n_d(r, t) = n_{d,0}(1 + \beta_L n_{d,0} t)^{-1},$$

analogously to Eq. (18).

Figure 4a shows the propagation of a diffusion wave annihilating the electron component in the photoemission cell. At the initial stage of dispersion, when the electron flux to the wall is large and is directed along the field, electrons are strongly cooled, the cooling occurring even in the central region (Fig. 4b). However, with decreasing electron flux, the electron temperature at the center is restored to the preset value, and cooling takes place only at the wall, where the field has the maximal value and the temperature is lower than at the center by approximately 2000 K. This cooling of electrons prolongs the electron dispersion process insignificantly. It should be noted that the evolution of the radial distributions of the dust particle concentration and of the electric field is, in accordance with Figs. 4c and 4d, in complete agreement with solution (18).

In calculations made for $T_{e,0}$ = 1000 K (curve *7* in Fig. 1), it was found that the dust plasma decay in the central region of the cell occurs with a certain delay. This is due to the fact that electrons in the central region for $T_{e,0}$ = 1000 K cannot escape to the walls since they do not possess the energy required for violating the quasineutrality condition. For this reason, diffusion of electrons at the initial stage follows the ambipolar mechanism, without any charge separation in the central region. When a part of dust particles from the periphery leave the cell, the field of the remaining particles cannot confine electrons and the electron diffusion follows the free diffusion mode. After this electrons rapidly escape from the cell volume, and the dispersion of dust particles occurs analogously to cases corresponding to a higher electron temperature (if we shift curve *7* by the magnitude of the initial delay, it almost coincides with curve *1*).

In calculations for the electron temperature $T_{e,0}$ = 3000 and 6000 K and for $n_{d,0}$ = 300 cm^{-3}, the maximal value of the wall potential equal to –8.4 V was obtained. The radial potential distribution in neon was measured in [24] in a tube of radius 1.7 cm under a pressure of 0.75 Torr. It was found that the wall potential $\phi_w \approx$

–20 V for an electron temperature at the tube axis of about 5 eV; i.e., the wall potential was $4T_e/e$. In calculations based on the nonlocal method of moments used here, the same potential of the wall was obtained for an electron temperature of 5.6 eV in [14] (i.e., $|\phi_w| \approx 3.5T_e/e$). In the steady-state regime, the wall potential is determined from the equality of the electron and ion currents; for this reason, the wall potential for heavy ions can be even higher [25]. In a glow discharge in neon under low pressures, Ne^+ is the main ion. Consequently, the wall potential in a photocell with heavy dust ions for $p = 40$ Torr may increase as compared to the value of the wall potential in a glow discharge by a factor of up to $\ln(k_i/\mu_d) \approx 10$, i.e., to $35T_e/e$ (here, k_i is the mobility of Ne^+ ions). For $T_{e,0} = 3000$ K = 0.26 eV, the wall potential under steady-state conditions may be as high as 8.75 V. This value is larger than the maximal value of the potential obtained in calculations; consequently, at such a temperature, electrons leave for the walls even in the free diffusion mode. At $T_{e,0} = 1000$ K, a segment corresponding to the ambipolar diffusion mode appears at the initial stage. After the system is rid of dust particles at the edge and the potential drops to approximately below 3 V, the departure of electrons is again governed by the free diffusion mechanism. In accordance with our estimates, for $T_{e,0} = 6000$ K = 0.52 eV, diffusion at the initial stage of plasma decay (provided that the mean charge of dust particles remains unchanged) is ambipolar for $n_{d,0} \geq 650$ cm^{-3} and $n_{e,0} \geq 2 \times 10^7$ cm^{-3}. In the central part of the photocell, the densities of charged particles remain unchanged until the arrival of the diffusion wave or until the diffusion passes to the free mode after the disposal of a part of the positive charge (if the dust particle concentration is not high enough).

Let us now estimate the drift velocity and the kinetic energy of dust particles in the central region of the photoemission cell. Using expressions (7), (18), and (19), we derive the relation defining the local drift velocity of a dust particle of radius a and charge z:

$$v_{dr}(t, r, a, z) = \frac{\nu_d r}{2} \frac{z}{z_d} \frac{r_d}{a} \frac{n_d}{n_{d,0}} = \frac{\nu_d r}{2} \frac{n_d}{n_{d,0}} \frac{q}{q_0}, \qquad (20)$$

where $q = z/a$ and $q_0 = z_d/r_d$. The velocities of directional motion and the thermal velocities of dust particles in a photoemission cell were determined in [26] by dividing the video image region into horizontal and vertical layers of width approximately equal to 0.1 cm. It was shown that the drift velocity distribution both in the radial direction and along the photocell axis in the chosen layers is a Maxwellian distribution. It can be seen from Eq. (20) that the drift velocity of dust particles at a preset distance from the cell axis is determined by relation $q = z/a$. Consequently, if the dusty plasma decay pattern under the condition of experiments made on the Mir space station is correct, we can conclude that the distribution function for dust particles over values of q is unambiguously connected with the "drift velocity distribution function." Namely, it is a Gaussian function,

$$f_q = \frac{1}{\sqrt{2\pi}\sigma} \exp\left[-\frac{(q - q_0)^2}{2\sigma^2}\right], \qquad (21)$$

with mathematical expectation q_0 and with the dispersion connected with the dust particle "temperature" determined in [10, 11, 26] via the relation

$$\sigma = \frac{2q_0}{\nu_d r} \sqrt{\frac{T_d}{m_d}}. \qquad (22)$$

Using the value $T_d \approx 51$ eV of dust particle temperature in the radial direction obtained in [26] for a vertical layer of $x = [0.7; 0.8]$ cm (taking into account that the center of the video image region is displaced to the left and has a coordinate of $x = 0.32$ cm), we find from relation (22) that $\sigma \approx 0.55q_0$ for $n_d = 300$ cm^{-3} and $\sigma \approx 0.65q_0$ for $n_d = 195$ cm^{-3}. Let us use the data obtained for deriving several estimates.

The drift velocity of dust particles averaged over a layer at the initial stage of dusty plasma decay in the photocell, when $n_d \approx n_{d,0}$, is defined as

$$\langle v_{dr} \rangle_k = \frac{\nu_d}{2q_0} \frac{\int_{r_k}^{r_{k+1}} \int f_q r q \, dq \, dr}{\int_{r_k}^{r_{k+1}} dr} = \frac{\nu_d (r_k + r_{k+1})}{4}, \qquad (23)$$

where r_k and r_{k+1} are the boundaries of the kth layer (it should be noted that averaging was carried out in [10, 11, 26] only over the particles falling into the laser knife plane; for this reason, the integrals for the planar geometry are used here). Figure 5 shows for comparison the velocities of dust particles calculated by formula (23) and determined in [28]. It can be seen that the results obtained using our model and by determining the velocities from the displacement over the known time are in satisfactory agreement.

In [10, 11, 26], the "thermal" velocity of dust particles was also determined:

$$v_{d,th} = \sqrt{\langle v_{dr}^2 \rangle - \langle v_{dr} \rangle^2}. \qquad (24)$$

Using relations (20) and (21), for the initial stage of

dusty plasma decay, we obtain

$$\langle v_{dr}^2 \rangle_k = \left(\frac{\nu_d}{2q_0}\right)^2 \frac{\int_{r_k}^{r_{k+1}} \int f_q r^2 q^2 dq dr}{\int_{r_k}^{r_{k+1}} dr} = \frac{\nu_d^2 (r_k^2 + r_k r_{k+1} + r_{k+1}^2)}{12}\left(1 + \frac{\sigma^2}{q_0^2}\right). \quad (25)$$

The temperature of the dust component was determined in [10, 11, 26] from the relation

$$T_d = m_d v_{d,th}^2. \quad (26)$$

For a vertical layer of $x = [0.7; 0.8]$ cm, we find from Eq. (26) that the temperature of dust particles of radius 37.5 μm is $T_d \approx 54.2$ eV for $n_d \approx 300$ cm^{-3} and $T_d \approx 51.7$ eV for $n_d \approx 195$ cm^{-3}, which is in good agreement with values $T_d \approx 51 \pm 5$ eV obtained for this layer in [26]. It can be seen that the temperature of dust particles may attain high values. It is clear from the above, however, that it is the kinetic energy of directional motion of dust particles. The reason for the breaking away of the dust particle energy from the thermal energy of buffer gas particles in an electric field is the same as for the breaking away of the electron energy; i.e., it is hampered energy exchange with the buffer gas due to the large difference in mass. In the case of electrons, their collisions with buffer gas particles leads to randomization of the direction of their motion with a rate exceeding considerably the energy loss rate; for this reason, the drift velocity of electrons is noticeably smaller than their thermal velocity. In the case of dust particles, the exchange of momentum with buffer gas particles is also hampered; consequently, the loss of directional motion of particles occurs more or less effectively only at thermal velocities.

The randomization of kinetic energy could occur during collisions between dust particles, but such events were quite rare in experiments [1]. This is due to the fact that dust particles in the photocell move along "smooth" electric field lines, which do not intersect. At present, an intense quest for the mechanisms of heating of the dust component continues; among these mechanisms, the most studied are those associated with charge fluctuations of dust particles and with electric field fluctuations. However, neither of these mechanisms can lead to the experimentally observed values of dust particle temperature (see, for example, [9]).

In [27, 28], a conclusion concerning the possibility of abnormally high dust particle temperatures was drawn from a theoretical analysis of the dust component on the basis of a chain of the BBGKY (Bogoliubov–Born–Green–Kirkwood–Yvon) equations. This conclusion is based on the erroneous assumption that the Landau length, or the Coulomb length of ions (the smallest spacing between an ion and a dust particle), can be much larger than the Debye radius. This leads to negative values of the Coulomb logarithm, which is positive by definition. In the applicability range for the theory developed in [27, 28], the dust particle temperature never exceeds three ionic temperatures.

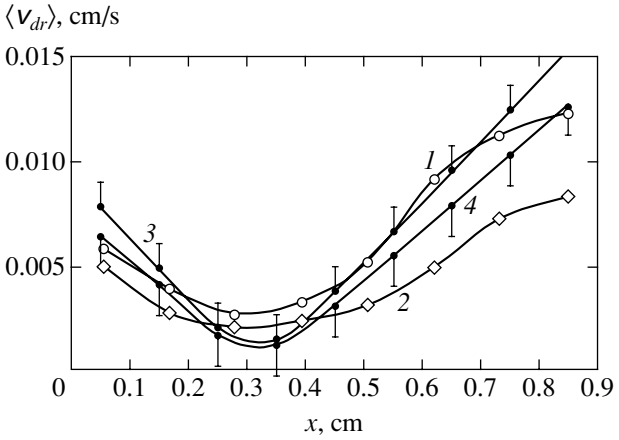


Fig. 5. Radial dependence of the layer-averaged drift velocity of dust particles: experimental data from [26] (*1*, *2*) and results of calculation by formula (23) (*3*, *4*); $n_{d,0} = 300$ (*1*, *3*) and 195 cm^{-3} (*2*, *4*). (For calculated data, the spread due to finite width of the layers is indicated.)

It was found in [26] that thermal velocities of dust particles in a photocell at the initial stage of dusty plasma decay are practically independent of the layer number. In our model, dependence on the layer number exists, which, in our opinion, is a consequence of the difference between the drift velocities calculated by us and the velocities determined in [26]; this difference is especially significant at the photocell center (see Fig. 5). The reason for the dependence of thermal velocities on the layer number may also be associated with the presence of acoustic vibrations of dust particles, which must be especially noticeable against the background of low directional velocities at the center of the photocell. The fact that less than 60 particles were contained in the video image region used for obtaining the above data, while some layers contained just 1–3 particles, could also be responsible for this dependence. Analysis of this question is beyond the scope of the present study; so, we pass to discussion of the effect of magnetic field on the dusty plasma decay process.

4. DIFFUSION OF CHARGED PARTICLES IN A PHOTOEMISSION CELL IN MAGNETIC FIELD

The analysis carried out in the previous section leads to the conclusion that the decay of the dusty plasma in the photoemission cell in experiments [1]

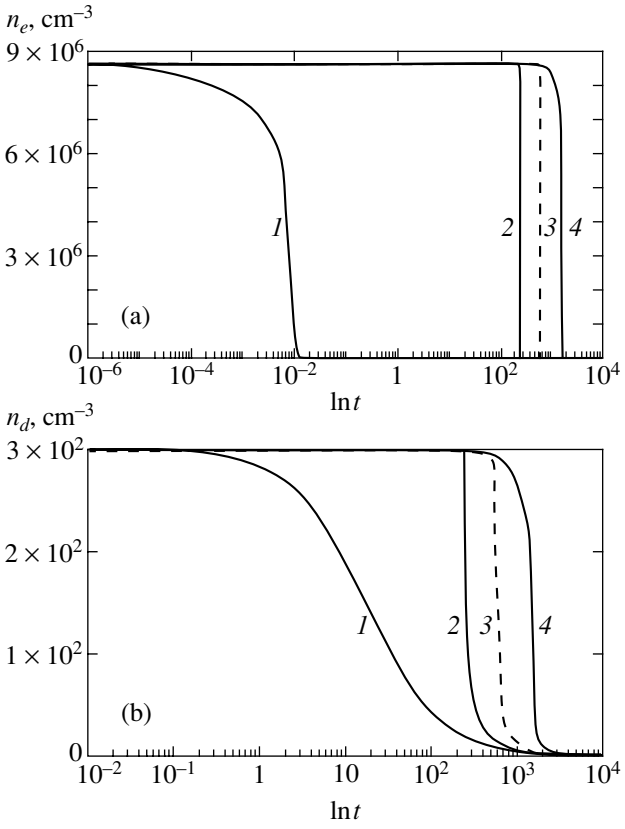


Fig. 6. Calculated time dependences of (a) electron concentration and (b) dust particle concentration in the central region of a photoemission cell in zero magnetic field for $T_{e,0}$ = 6000 K (*1*) and T_e = 300 K (*2*) and in a magnetic field $H = 10^3$ G (*3*) and $H = 10^4$ G (4); T_e = 300 K, r_d = 25 μm, $n_{d,0}$ = 300 cm^{-3}, $z_d = 2.9 \times 10^4$.

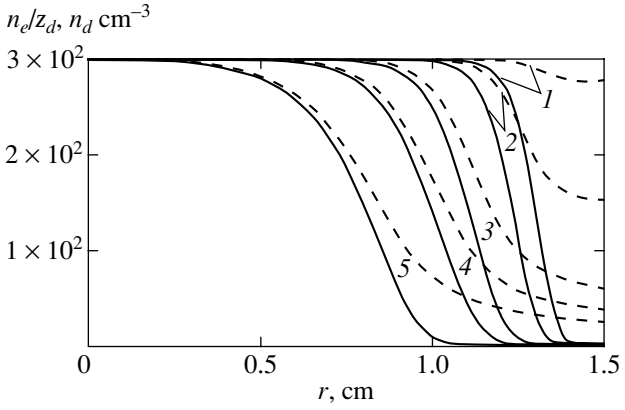


Fig. 7. Evolution of concentration distributions for electrons (solid curves) and dust particles (dashed curves) over the photoemission cell radius in a magnetic field $H = 10^3$ G, T_e = 300 K, r_d = 25 μm, $n_{d,0}$ = 300 cm^{-3}, $z_d = 2.9 \times 10^4$; t = 1.4 (*1*), 17 (*2*), 88 (*3*), 203 (*4*), and 467 s (*5*).

occurred in accordance with the free diffusion mechanism. In order to increase the dusty plasma confinement time in the cell, we must either considerably increase the electron concentration, which is difficult to realize, or considerably reduce the rate of diffusive departure of electrons. For this purpose, we must place the photoemission cell in a magnetic field directed along the cell axis; in order to suppress or prevent the departure of electrons along magnetic field lines, we can impart a strongly prolate or toroidal shape to the photoemission cell.

The application of a magnetic field perpendicular to the radial field of polarization of the dusty plasma leads to the following two effects. The first effect is a strong decrease in the electron mobility and diffusion coefficient (as well as in the electron energy diffusion coefficient and the thermoelectric coefficient) in crossed magnetic and electric fields [22]:

$$k_{e,\perp} = \frac{k_{e,0}}{1 + \omega_0^2/\nu^2}, \quad D_{e,\perp} = \frac{D_{e,0}}{1 + \omega_0^2/\nu^2} \qquad (27)$$

(here, ω_0 is the cyclotron frequency for electrons, ν is the frequency of elastic collisions, and $k_{e,0}$ and $D_{e,0}$ are the electron mobility and diffusion coefficient in zero magnetic field). For example, the frequency of elastic collisions of electrons in neon under a pressure of 40 Torr at room temperature is $\nu = 1.3 \times 10^8$ s^{-1}, while the cyclotron frequency in a magnetic field $H = 10^4$ G is $\omega_0 = 1.8 \times 10^{11}$ rad/s. For this reason, the transfer coefficients in such a field decrease by six orders of magnitude, while the transport coefficients for dust particles remain practically unchanged in this magnetic field.

The second effect is associated with a strong decrease in the electron energy relaxation length during the radial motion of electrons in spite of the fact that the electron energy loss rate in a magnetic field remains unchanged (in a magnetic field $H = 10^4$ G at T_e = 300 K, the energy length l_u = 0.6 μm against 2.5 cm in zero field). This leads to a strong increase in the electron energy loss per unit length; consequently, during diffusion of electrons from one dust particle to another towards the wall, the electrons have time to reach the thermodynamic equilibrium with the buffer gas. As a result, the electron temperature in the photoemission cell drops to room temperature. For such small lengths of electron energy stabilization, the local approach can be applied for calculating the transport coefficients, and the energy balance equation becomes unnecessary. In Fig. 6, the results of calculation of the evolution of dust particle concentration averaged over the central region with and without a magnetic field are compared. It can be seen that the application of a magnetic field leads to an increase in the dusty plasma confinement time up to values on the order of 10^3 s.

The curves in Fig. 7 describing the evolution of the radial distributions of concentrations of electrons and dust particles show that diffusion in magnetic field follows the ambipolar mechanism. This leads to an

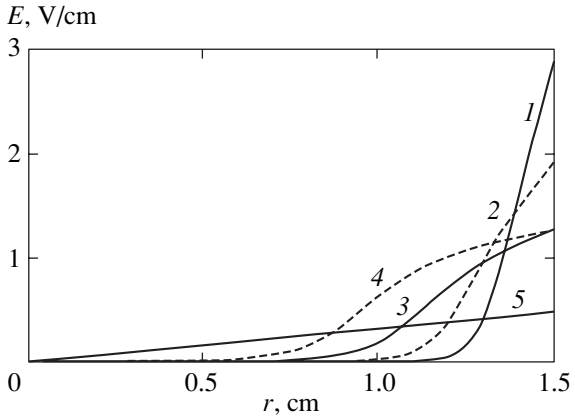


Fig. 8. Evolution of the electric field distribution over the photoemission cell radius in a magnetic field $H = 10^3$ G for $T_e = 300$ K, $r_d = 25$ μm, $n_{d,0} = 300$ cm^{-3}, $z_d = 2.9 \times 10^4$; $t =$ 1.4 (*1*), 40 (*2*), 203 (*3*), 467 (*4*), and 1000 s (*5*).

increase in the dusty plasma confinement time in the photoemission cell. For this reason, space experiments with a magnetic field should be staged for studying a dusty plasma with the ambipolar mechanism of departure of charged particles. Figure 8 shows how the electric field penetrates to the center of the cell in the course of evolution, being reduced simultaneously at the cell edges. In this case, the radial dependence of the field differs from the case of zero magnetic field (see Fig. 4d) and has a more complicated form.

The transition to the ambipolar mechanism of plasma decay in the photoemission cell was also observed in calculations for zero magnetic fields when the electron heating source was switched off. In this case, the electron temperature dropped from the initial temperature $T_{e,0} = 0.52$ eV to room temperature by the time instant $t = 10^{-5}$ s. Figure 6 shows the curves describing the evolution of the electron and dust particle concentrations in these calculations. It can be seen that thermalization of electrons in collisions with the buffer gas considerably prolongs the dusty plasma decay. These calculations, as well as calculations with a magnetic field, show that, as mentioned above, the concentrations of charged particles at the stage of formation of Bessel profiles at the center of the cell are frozen in the ambipolar diffusion regime. After the completion of this stage, practically all charged particles (first electrons and then dust particles) leave the cell approximately during the same time.

The results of calculations disregarding the electron heating source indicate another possible way for increasing the dusty plasma confinement time in a photocell, i.e., the replacement of neon by lighter helium, which, in addition, has a larger transport cross section, and the elevation of pressure to the atmospheric level. This will reduce the electron energy relaxation length to 100 μm (which is much smaller than the mean distance between dust particles in experiments [1]), which must lead to a decrease in the electron temperature to the gas temperature. An increase in pressure (we can expect that this will not lead to undesirable consequences under the microgravity conditions) will also reduce the electron mobility and diffusion coefficient. All this will make it possible to increase the dusty plasma decay time in helium to a value comparable with the plasma decay time in neon in a magnetic field.

5. CONCLUSIONS

The numerical experiments with a dusty plasma in a photoemission cell carried out in this study proved that the plasma decay under the experimental conditions [1] was governed, in all probability, by the mechanism of free electron diffusion followed by the drift of positively charged dust particles. The results of calculations are in good agreement both with the experimental curves describing the evolution of the dust particle concentration and with the experimental values of dust particle charge. It is shown that the radial decay of the dusty plasma in a magnetic field directed along the cell axis is strongly retarded, and the plasma decay occurs in accordance with the ambipolar diffusion mechanism.

ACKNOWLEDGMENTS

The authors express their gratitude to O.S. Vaulina and O.F. Petrov for fruitful discussions of the material and for valuable critical remarks.

This study was supported by the Russian Foundation for Basic Research (project nos. 01-02-16628 and 00-15-96539).

REFERENCES

1. V. E. Fortov, A. P. Nefedov, O. S. Vaulina, *et al.*, Zh. Éksp. Teor. Fiz. **114**, 2004 (1998) [JETP **87**, 1087 (1998)].
2. A. Melzer, A. Homann, and A. Piel, Phys. Rev. E **53**, 2757 (1996).
3. J. B. Pieper and J. Goree, Phys. Rev. Lett. **77**, 3137 (1996).
4. H. M. Thomas and G. E. Morfill, Nature **379**, 806 (1996).
5. V. E. Fortov, A. P. Nefedov, V. M. Torchinsky, *et al.*, Phys. Lett. A **229**, 317 (1997).
6. V. V. Zhakhovskiĭ, V. I. Molotkov, A. P. Nefedov, *et al.*, Pis'ma Zh. Éksp. Teor. Fiz. **66**, 392 (1997) [JETP Lett. **66**, 419 (1997)].
7. G. E. Morfill, H. Thomas, U. Konopka, and M. Zuric, Phys. Plasmas **6**, 1769 (1999).
8. R. A. Quinn and J. Goree, Phys. Plasmas **7**, 3904 (2000).
9. R. Quinn and J. Goree, Phys. Rev. E **61**, 3033 (2000).
10. O. S. Vaulina, A. P. Nefedov, O. F. Petrov, and V. E. Fortov, Zh. Éksp. Teor. Fiz. **119**, 1129 (2001) [JETP **92**, 979 (2001)].
11. O. S. Vaulina, A. P. Nefedov, O. F. Petrov, and V. E. Fortov, Phys. Rev. Lett. **88**, 035001 (2002).

12. B. M. Smirnov, *Ions and Excited Atoms in Plasma* (Atomizdat, Moscow, 1974).
13. W. P. Allis and D. J. Rose, Phys. Rev. **93**, 84 (1954).
14. J. H. Ingold, Phys. Rev. E **56**, 5932 (1997).
15. A. V. Filippov, N. A. Dyatko, A. F. Pal', and A. N. Starostin, Fiz. Plazmy **29**, 214 (2003) [Plasma Phys. Rep. **29**, 190 (2003)].
16. E. M. Lifshitz and L. P. Pitaevskiĭ, *Physical Kinetics* (Nauka, Moscow, 1979; Pergamon Press, Oxford, 1981).
17. N. A. Fuchs, *The Mechanics of Aerosols* (Pergamon, New York, 1964).
18. *Physical Quantities. Handbook*, Ed. by I. S. Grigor'ev and E. Z. Meĭlikhov (Énergoatomizdat, Moscow, 1991).
19. A. A. Sickafoose, J. E. Colwell, M. Horanyi, and S. Robertson, Phys. Rev. Lett. **84**, 6034 (2000).
20. A. F. Pal', A. O. Serov, A. N. Starostin, *et al.*, Zh. Éksp. Teor. Fiz. **119**, 272 (2001) [JETP **92**, 235 (2001)].
21. A. F. Pal', D. V. Sivokhin, A. N. Starostin, *et al.*, Fiz. Plazmy **28**, 32 (2002) [Plasma Phys. Rep. **28**, 28 (2002)].
22. L. G. H. Huxley and R. W. Crompton, *The Diffusion and Drift of Electrons in Gases* (Wiley, New York, 1974; Mir, Moscow, 1977).
23. A. V. Phelps, ftp://jila.colorado.edu/collision_data/eletrans.txt.
24. D. Uhrlandt and R. Winkler, J. Phys. D **29**, 115 (1996).
25. A. von Engel, *Ionized Gases*, 2nd ed. (Clarendon Press, Oxford, 1965; Fizmatgiz, Moscow, 1959).
26. A. P. Nefedov, O. S. Vaulina, O. F. Petrov, *et al.*, Fiz. Plazmy **29** (1), 37 (2003) [Plasma Phys. Rep. **29**, 31 (2003)].
27. A. G. Zagorodny, P. P. J. M. Schram, and S. A. Trigger, Phys. Rev. Lett. **84**, 3594 (2000).
28. P. P. J. M. Schram, A. G. Sitenko, S. A. Trigger, and A. G. Zagorodny, Phys. Rev. E **63**, 016403 (2001).

Translated by N. Wadhwa

Journal of Experimental and Theoretical Physics, Vol. 96, No. 4, 2003, pp. 695–703.
Translated from Zhurnal Éksperimental'noĭ i Teoreticheskoĭ Fiziki, Vol. 123, No. 4, 2003, pp. 787–797.

PLASMA, GASES

The Influence of the State of a Compressed Thermonuclear Substance on the Combustion of Inertial Fusion Targets under Direct Ignition by a Short Radiation Pulse

A. A. Andreev[a], S. Yu. Gus'kov[c,*], D. V. Il'in[b], A. A. Levkovskii[b], V. B. Rozanov[c], V. E. Sherman[b], and O. B. Vygovskiĭ[b]

[a]*Vavilov State Optical Institute, Birzhevaya liniya 12, St. Petersburg, 199034 Russia*
[b]*St. Petersburg Institute of Mechanical Engineering, St. Petersburg, 195108 Russia*
[c]*Lebedev Physical Institute, Russian Academy of Sciences, Leninskii pr. 53, Moscow, 119991 Russia*
**e-mail: guskov@sci.lebedev.ru*

Received October 31, 2002

Abstract—One-dimensional numerical calculations were performed to study the dependence of conditions for initiating thermonuclear combustion and of the target gain of direct-ignition inertial fusion targets ignited by a short radiation pulse on the initial temperature of a preliminarily compressed fuel and the initial heat energy distribution between plasma electrons and ions in the ignition region (igniter). The igniter parameters at which an effective thermonuclear target explosion with a $G \sim 10^3$ target gain occurred were shown to substantially depend on the initial temperature of the major fuel fraction and the initial heat energy distribution between igniter electrons and ions. The heat energy of the igniter passed a minimum as the size of the igniter decreased. The dependences of these minimum energies on the temperature of the major fuel fraction at various initial energy distributions between igniter electrons and ions were determined. An increase in the temperature of the major fuel fraction was shown to decrease the target gain. © *2003 MAIK "Nauka/Interperiodica".*

1. INTRODUCTION

One of the most promising directions in the area of inertial confinement fusion is based on the direct ignition concept [1, 2], according to which thermonuclear substance compression and heating under the action of two synchronized pulsed energy sources (drivers) is separated in time. The compressing driver, which acts first, effects slow compression of the target substance along a "cold" adiabat. The second, igniting driver must ensure a rapid heating of a small portion of the compressed thermonuclear fuel (igniter) in time not exceeding the time of the inertial confinement of the initial ignition region and provide the development of a self-sustaining thermonuclear combustion wave. The described ignition procedure allows the energy of a DT plasma to be minimized at a 20–50 kJ level in the attainment of the ignition threshold and at a 0.3–1 MJ level in the initiation of the combustion wave with a high target gain [1, 2].

The key direct ignition problem is that of heating the igniter by a firing pulse. We should mention two methods suggested for solving it. According to [1, 2], direct ignition can be effected with the use of inertial thermonuclear fusion targets capable of providing the internal introduction of igniting driver energy. Such targets may be spherical targets with one or several channels, for instance, conic in shape, or cylindrical targets with pinholes in one or both end surfaces. In [3], the formation of a channel for introducing igniting driver radiation directly during driver action on a spherical target was suggested. Accordingly, two laser pulses, one to produce a channel in the target as a result of ponderomotive action and the other to propagate along this channel and supply energy to the thermonuclear substance, had to be used. This method was named fast ignition.

Currently, various combinations of compressing and igniting drivers have been considered. From the point of view of energy requirements, the role of compressing drivers can be played by a pulse of short-wave laser radiation, a pulse of soft Z-pinched X-ray radiation, or a beam of heavy ions. The feasibility has been discussed of using igniting drivers such as a beam of "accelerator" heavy ions [1, 2], a beam of fast electrons formed under the action of short-wave [1–3] or long-wave [2] laser radiation on a substance, an X-ray radiation pulse [4], an accelerated substance microparticle [4], and a flow of light megavolt laser plasma ions [5] created under the action of a laser beam on a thin plane target-generator made of a substance of light elements and situated separately from the thermonuclear target. A beam of light laser plasma ions could prove to be the most promising igniting driver type. Indeed, the formation of a beam of accelerator ions and stable accelera-

tion of a macroparticle with the parameters necessary for direct ignition involves serious technical difficulties. The use of X-ray radiation requires solving the problem of generating a high-power nonequilibrium radiation pulse with an intensity of 10^{18}–10^{20} W/cm^2 and radiation quantum energy not higher than 500–800 eV [4, 5]. The main advantage of an ionic igniting driver compared with a beam of fast electrons is a higher efficiency of energy transfer to the thermonuclear substance virtually without scattering of particles, which accompanies energy transfer by fast electrons.

The key question for understanding the potentialities of various combinations of compressing and igniting drivers is the igniter firing parameters, namely, size, temperature, and energy, that allow an effective thermonuclear explosion in the target to be initiated. These parameters are determined exclusively by the hydrodynamics of thermonuclear plasma combustion after the action of both drivers. Calculations of igniter firing parameters by analytic and numerical methods were performed in a large number of works. A detailed bibliography can be found, e.g., in [6]. In order to minimize the energy of the compressing driver, which substantially (by an order of magnitude or more) exceeds the energy of the igniting driver, it was assumed in all these works that an igniter at the instant of its creation is surrounded by a dense and cold plasma of the major target fuel fraction in a state close to that of the completely degenerate Fermi gas. It should, however, be understood that preliminary compression of the thermonuclear fuel along the required cold adiabat, which corresponds to the degenerate final state of the substance in a spherical target, is a task far from simple. Fulfilling it requires the use of a temporarily shaped compressing driver pulse and very strict control of the action of sources which preliminarily heat the substance to be compressed and are related to energy transfer from the laser target crown by radiation and electronic heat conductivity.

We therefore consider it important to determine the degree to which the temperature of the fuel surrounding the igniter, which in real experiments can be higher than the Fermi energy, influences the parameters of igniter firing and the direct-ignition target gain. Clearly, an increase in the initial temperature of the major fuel fraction compared with the Fermi energy decreases the target gain compared with the direct ignition regime. In the limit of equal temperatures of the major fuel fraction and the igniter, the problem reduces to the combustion of a uniformly heated and compressed plasma; the target gain then does not exceed 400–500. Nevertheless, in the region of major fuel fraction temperatures much lower than the igniter temperature (5–7 keV), a study of the dependence of the target gain on the temperature of the major fuel fraction enables the conditions for directly igniting inertial fusion targets to be optimized at various ratios between the energies of the compressing and igniting drivers.

Another important characteristic related to the igniter parameters and, generally, the direct-ignition target gain is the initial distribution of heat energy between igniter electrons and ions. The energy of all igniting drivers specified above except for impact of a macroparticle is first transferred to electrons, whereas ions are only heated in electron–ion relaxation. The transfer of energy from an igniter to the surrounding cold fuel caused by electronic heat conductivity and radiative relaxation processes results in cooling the igniter before it begins to burn. Because of short igniter combustion times in direct-ignition targets (several dozen picoseconds), this cooling can require substantially higher ignition energies then with equal initial temperatures of igniter ions and electrons. In particular, precisely this effect is likely to be the main reason for a substantially higher ignition energy obtained in two-dimensional calculations of the direct ignition of inertial fusion targets by a beam of bismuth ions [7] compared with estimates and calculations in which relaxation energy exchange was ignored.

In this work, we study the influence of the two effects specified above, namely, the excess temperature of the major fuel fraction over the Fermi energy and relaxation energy exchange between igniter electrons and ions during primary igniting driver energy transfer to plasma electrons, on the conditions of initiating thermonuclear combustion and on the direct-ignition target gain in inertial confinement fusion. The results were obtained in numerical calculations for the model problem of combustion of a uniformly compressed spherical deuterium–tritium plasma with the central part (igniter) heated to a thermonuclear temperature significantly above that of the major part of the fuel surrounding the igniter. The computations were performed with the TERA program, which included a procedure for Monte Carlo calculations of energy transfer by thermonuclear particles.

2. AN ANALYSIS OF PHYSICAL PROCESSES

The inertial fusion target gain is defined as the ratio between the energy released in fusion reactions and the internal energy of the whole target E_p,

$$G_p = \frac{QgM}{E_p}. \quad (1)$$

Here, Q is the energy released in complete combustion of the unit mass of the thermonuclear fuel in the fusion reaction (for the DT reaction, $Q = 3.34 \times 10^{11}$ J/g); g is the degree of thermonuclear substance burnup in fusion reactions, that is, the ratio between the number of plasma nuclei that entered the fusion reaction to the initial number of nuclei; M is the mass of the plasma; and E_p is the internal energy of the target at the initial combustion time.

It is expedient to begin a qualitative analysis of the combustion of direct-ignition targets taking into consideration the general laws of combustion of uniformly heated and compressed targets, because simulating combustion in uniform conditions makes it possible to determine the ranges of variations in the initial igniter and major fuel fraction parameters for direct ignition. In addition, calculations show that the plasma characteristics during combustion approach uniform plasma characteristics for any effectively burning target in which a thermonuclear explosion occurs. The target gain of a uniform spherical DT plasma is

$$G_p = \frac{Qg}{c_V T_p}, \tag{2}$$

where c_V is the specific heat capacity [for a DT mixture of equal numbers of tritium and deuterium nuclei, $c_V = 1.15 \times 10^{15}$ erg/(g keV)] and T_p is the initial temperature of plasma electrons and ions ($T_e = T_i = T_p$).

The combustion of uniform targets has been studied in many works. In particular, a large cycle of numerical calculations of the combustion of uniform targets were performed by us with the TERA program in [8–11]. The main characteristics of combustion of uniform targets are as follows. The lower limit of ignition temperature is $T_p \sim 5$–7 keV; it is mainly determined by the condition that the rate of thermonuclear energy release should exceed the rate of radiative energy loss. According to (2), the highest target gain for a uniformly heated plasma, which corresponds to complete burnup of plasma nuclei in fusion reactions ($g = 1$), is 500–400 for the specified plasma temperatures. The degree of plasma burnup increases with the ρR parameter (ρ is the density and R the size of the igniter) and amounts to considerable values of 0.2–0.3 at $\rho R \geq 1$ g/cm^2.

For targets with $\rho R \geq 1$ g/cm^2, a sharp increase in energy release corresponding to the development of a thermonuclear explosion occurs in a narrow temperature interval. The target gain G values then change from $G \ll 1$ to $G \geq 100$ virtually in a jump. The minimum, critical ignition temperatures for such targets (5–7 keV) are substantially lower than the temperatures corresponding to maximum DT reaction rates ($T \sim 20$ keV). Indeed, at the initial stage of combustion, such targets are nontransparent to fast charged particles, that is, particles that transfer a substantial part of released thermonuclear energy to a plasma and heat it to temperatures of several dozen keV, which substantially exceed ignition temperatures. In targets with $\rho R < 1$ g/cm^2, heating by charged thermonuclear particles is ineffective, which results in comparatively low target gains ($G \leq 10$) and their smooth dependence on T_p in a wide range of initial temperatures.

The numerical calculations [8–11] gave simple approximation formulas for the degree of plasma burnup and the uniform target gain at high parameter values

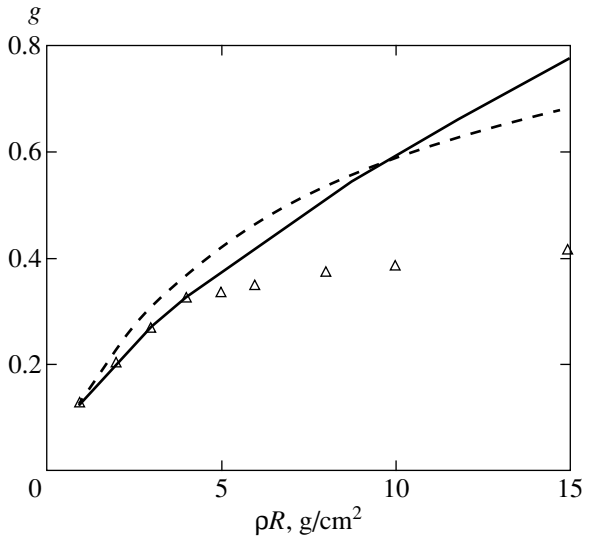


Fig. 1. Degree of burnup of a uniform DT plasma g as a function of plasma thickness ρR. The solid line corresponds to (4), and the dashed line, to (5); △ are the numerical simulation results.

attainable in practice, 5 g/cm^2 ≥ ρR ≥ 1 g/cm^2. These formulas read

$$G = \frac{370(\rho R)^{2/3}}{T_p}, \tag{3}$$

$$g = \frac{(\rho R)^{2/3}}{7.9}, \tag{4}$$

where ρR is in g cm^{-2} and T_p in keV.

The approximate equation [12]

$$g = \frac{\rho R}{\rho R + 7 \text{ g/cm}^3} \tag{5}$$

is very accurate and convenient for analytic calculations of the degree of burnup.

The dependences of the degree of DT plasma burnup g on the ρR parameter constructed using analytic formulas (4) and (5) and based on uniform target gain TERA numerical calculations are shown in Fig. 1. Equation (5) takes into account a decrease in plasma density (accordingly, in the thermonuclear reaction rate) during combustion and therefore gives a correct passage to the limit of large ρR parameter values, $g \longrightarrow 1$ as $\rho R \longrightarrow \infty$. On the other hand, (4) gives results that are closer to numerical calculation data in the range of ρR values that is most interesting for practical applications. It follows that the limits of the applicability of (3) and (4) are determined by the inequalities $G \gg 1$ and $g \leq 0.35$. The corresponding range of target parameters (1 g/cm^2 ≤ ρR ≤ 5 g/cm^2, $T_p < 100$ keV) covers the

region of parameters of interest for considering direct ignition.

According to (3), at a $T_p = 7$ keV ignition temperature, the target gain of a uniform DT plasma is $G = 110$ at $\rho R = 3$ and $G = 134$ at $\rho R = 4$.

Next, consider target combustion under direct ignition conditions. In direct ignition, an igniter is created in a short time, substantially shorter than the characteristic time of hydrodynamic processes. Target combustion calculations can therefore be performed using the isochoric model [13–15], in which the densities of the hot initial initiation region and the major cold fuel fraction are equal at the initial instant. The target gains and the energies of the igniting and compressing drivers are usually estimated by selecting the initial igniter parameters [its mean temperature T_f and size $(\rho R)_f$] based on the well-known criterion of a thermonuclear explosion in an isolated plasma bunch caused by heating with alpha particles, that is, $T_f \sim 5$–10 keV and $(\rho R)_f \sim 0.3$–0.4 g cm^{-2}. The igniter energy is

$$E_f \approx \frac{46 T_f (\rho R)_f^3}{(\rho/100 \text{ g cm}^{-3})^2} \text{ [kJ]}. \quad (6)$$

Here and throughout, igniter characteristics are labeled by index *f*, and the characteristics of the cold compressed fuel, by index 0. If the state of the major fuel fraction is completely degenerate, the mass of this fraction is characterized by two parameters, size $(\rho R)_0$ and density, which, in the isochoric model, equals the density of the igniter, $\rho_0 = \rho_f = \rho$. The energy of the major fuel fraction is then

$$E_0 \approx \frac{3(\rho R)_0^3}{(\rho/100 \text{ g cm}^{-3})^{4/3}} \text{ [kJ]}. \quad (7)$$

As mentioned, effective target combustion corresponds to high values of the $(\rho R)_0$ parameter of the major fuel fraction, $(\rho R)_0 = 3$–5 g/cm^2. For the igniter and major fuel fraction parameters specified above, the energies of the two target components are

$$E_f \approx \frac{15}{(\rho/100 \text{ g cm}^{-3})^2} \text{ [kJ]},$$

$$E_0 \approx \frac{200}{(\rho/100 \text{ g cm}^{-3})^{4/3}} \text{ [kJ]}.$$

It follows that, if the bulk of the fuel is in the degenerate state, the ratio between the igniter energy and the energy of the major fuel fraction decreases as density increases by the law $E_f/E_0 \propto \rho^{-2/3}$ and equals 0.08 at density $\rho = 100$ g/cm^3 ($E_f = 15$ kJ, $E_0 = 200$ kJ) and 0.04 at density $\rho = 300$ g/cm^3 ($E_f \approx 1.5$ kJ, $E_0 = 40$ kJ).

As the energy and mass of the cold fuel are much larger than those of the igniter, the direct-ignition target gain is determined to a high accuracy by the combustion of precisely the major part of the target,

$$G_d = \frac{QgM_0}{E_0}. \quad (8)$$

Substituting (7) and (4) for the energy and the degree of burnup of the major fuel fraction into (8) yields

$$G_d = 1.25 \times 10^5 \frac{(\rho R)_0^{2/3}}{\rho^{2/3}}. \quad (9)$$

At equal ρR parameters of a uniform target and the major direct-ignition target fraction (that is, at equal burnup degrees), the ratio between the direct-ignition target gain and the target gain of a uniform plasma is determined by the ratio between ignition temperature and Fermi energy,

$$\frac{G_d}{G_p} = \frac{5}{2}\frac{Z+1}{Z}\frac{T_p}{\varepsilon_F}. \quad (10)$$

Here, the Fermi energy is $\varepsilon_F \approx 1.4 \times 10^{-2} \rho^{2/3}$ keV. At a $\rho = 300$ g/cm^3 density of the direct-ignition target plasma and a 7 keV ignition temperature of a uniform target, this ratio is about 60. The direct-ignition target gain at the density specified above is estimated at $G_d \approx 5.8 \times 10^3$ for $(\rho R)_0 = 3$ g/cm^2 and $G_d \approx 2.7 \times 10^3$ for $(\rho R)_0 = 1$ g/cm^2.

Clearly, the effect of superheating the major direct-ignition target fraction in comparison with the completely degenerate state can be estimated based on the above reasoning and Eq. (10), in which the temperature of the superheated major plasma fraction of the direct-ignition target should be substituted for T_p. For this reason, the superheated direct-ignition target gain can be estimated by (3) after substituting the T_0 temperature of the major fraction of the direct-ignition target plasma for T_p,

$$G_d = \frac{370(\rho R)_0^{2/3}}{T_0}. \quad (11)$$

Calculations of the superheated direct-ignition target gain by (11) at $(\rho R)_0 = 3$ g/cm^2 give $G_{d*} \approx 778$ at a $T_0 = 1$ keV temperature of the major plasma fraction and $G_{d*} \approx 1.23 \times 10^3$ at $T_0 = 0.6$ keV. Note that, apart from the temperature of a superheated target plasma being higher than the Fermi energy, a considerable contribution to decrease in the superheated target gain compared with a target with a completely degenerate plasma is made by an increase in the number of heated

plasma particles [the $(Z + 1)/Z$ coefficient in (10)], because the energy of ions is zero in the degenerate state.

The following conclusions about the special features of combustion of superheated direct-ignition targets can be drawn from the above consideration. In spite of the presence of the additional T_0 parameter, the dependence of the superheated target gain on the plasma parameters becomes simpler in a sense. At fixed (ρR) and igniter and major target fraction temperatures, density ρ only plays the role of a characteristic scale for the other physical values. For instance, mass $M \sim (\rho R)_0^3/\rho^2$; time of target expansion $\Delta t \sim R/v \sim (\rho R)_0/\rho T^{1/2}$, where v is the mean velocity of sound; $E_0 \sim MT \sim \rho^{-2}$; etc. Thermonuclear combustion can be described as an approximately scale-invariant process in the variables $r^* = \rho R$, $t^* = \rho t$, $E^* = \rho^2 E_0$, and $M^* = \rho^2 M_0$. In particular, as distinguished from the direct-ignition target with a completely degenerate plasma, the superheated target gain at given $(\rho R)_0$ and T_0 does not depend on ρ and, accordingly, on the initial energy E_0. The exact scale invariance is only violated because of a weak density dependence of the Coulomb logarithm. This conclusion is also confirmed by the numerical calculations considered below.

3. NUMERICAL CALCULATION RESULTS

Numerical calculations of direct-ignition target gains have been performed starting with the creation of an isochoric igniter. The mathematical model of thermonuclear combustion of a nonuniform spherically symmetrical plasma is described in the TERA program by a system of continuity equations and equations of motion, energy exchange, and plasma state combined with the kinetic equations of fast thermonuclear particles. The hydro- and thermodynamic processes are described in the approximation of a one-liquid two-temperature (T_e, T_i) plasma taking into account electronic and ionic heat conductivities and electron–ion energy exchange.

The kinetic processes with the participation of fast thermonuclear particles and thermal radiation, which give the major contribution to the propagation of the combustion wave, are characterized by high density and temperature gradients along the mean free path of fast thermonuclear particles, spatial anisotropy of the distribution function of these particles, a complex energy dependence of the Coulomb collision loss, and the existence of several connected thermonuclear reaction channels. In these conditions, the most correct method for simulating the kinetics of fast thermonuclear particles is the Monte Carlo method. As the flight time of fast thermonuclear particles is considerably shorter than the characteristic time of changes in plasma hydrodynamic parameters, the kinetics can be stochastically simulated by solving quasi-stationary kinetic equations at each time step of a nonstationary difference scheme based on continuum equations. This scheme for simultaneously solving hydrodynamic and kinetic equations is implemented in the modified TERA package that we use. A more detailed description of the physicomathematical model is given in works [10, 16–18], concerned with numerically simulating the combustion of thermonuclear targets with a uniformly heated and compressed plasma and targets with an isobaric distribution of parameters formed under purely hydrodynamic compression and heating of a spherical target.

It is assumed that, at the instant of maximum compression, the target is uniform except a comparatively small high-temperature region (igniter) in the center. The initial conditions for an isochoric igniter ($\rho_f = \rho$) are described by two parameters, namely, characteristic thickness $(\rho R)_f$ and temperature T_f. Further, we assume that the major fuel is in a nondegenerate state. Its temperature T_0 is one of the free parameters of the problem that determine the critical energy of the igniter,

$$E_f \geq \frac{F(T_0)}{(\rho/100\ \text{g cm}^{-3})^2}\ [\text{kJ}].$$

To determine the influence of energy distribution between igniter ions and electrons on the igniter critical parameters, two limiting cases were analyzed. In one variant, the temperatures of igniter ions and electrons at the initial instant were considered equal ($T_i = T_e = T_f$). In the other, it was assumed that igniter electrons were only heated at the initial instant, whereas the temperature of igniter ions coincided with the major fuel temperature ($T_i = T_0 \ll T_f$).

Calculations show that, if an effective thermonuclear explosion occurs in a DT target plasma, the degree of burnup weakly depends on the ignition method. A study of combustion efficiency therefore reduces to the determination of the critical (minimum) igniter parameter values that ensure stable target ignition.

The ranges of igniter and major fuel parameter variations were selected based on the analysis and estimates of the preceding section. The igniter parameters satisfied the conditions $T_f \geq 5$–10 keV and $(\rho R)_f \geq 0.3$–0.4 g/cm^2. The ρR parameter of the major fuel fraction corresponded to a high degree of burnup, $(\rho R)_0 > 1$ g/cm^2. The range of temperature variations, $T_0 \sim 0.5$–1 keV, was selected to satisfy two requirements. First, the temperature of the major fuel fraction had to be much lower than the critical ignition temperature of a uniform plasma equal to 5–7 keV. On the other hand, it had to be higher than the Fermi energy, which amounted to several hundreds eV at densities of $\rho \sim 10^2$ g/m^3.

At fixed $(\rho R)_0$ and T_0 major fuel parameters, series of model isochoric configurations with different T_f and $(\rho R)_f$ central igniter parameters were constructed. At

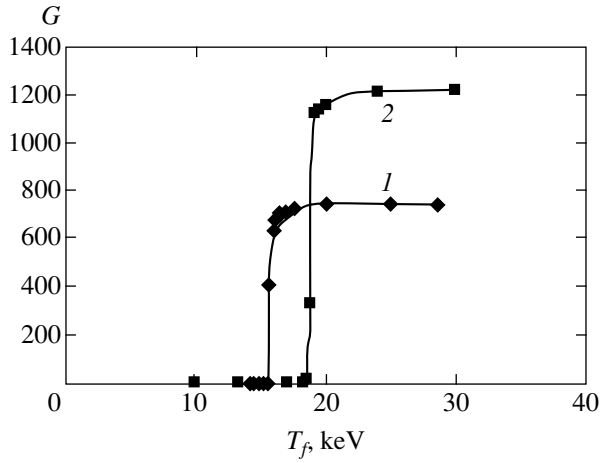


Fig. 2. Target gain G for a target with $(\rho R)_0 = 3$ g/cm^2 as a function of igniter temperature T_f. The igniter ion and electron temperatures at the initial time are $T_i = T_e = T_f$. Igniter thickness $(\rho R)_f = 0.3$ g/cm^2, major fuel temperature $T_0 =$ (*1*) 1 and (*2*) 0.6 keV.

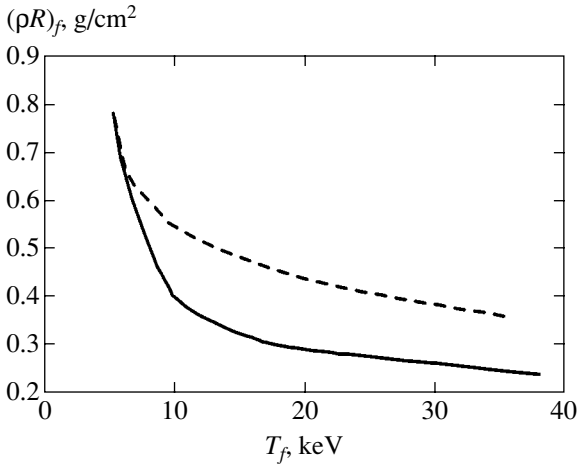


Fig. 3. Critical isochoric igniter parameters for targets with major fuel temperature $T_0 = 0.6$ keV. The solid line corresponds to equal energy distribution between igniter ions and electrons ($T_i = T_e = T_f$), and the dashed line, to the situation when the whole igniter energy is concentrated in the electronic component ($T_i = T_0$, $T_f = T_e/2$).

each selected set of parameters, two limiting initial igniter energy distributions between plasma electrons and ions were considered, namely, (a) $T_i = T_e = T_f$ and (b) $T_e = T_f$ and $T_i = T_0 \ll T_f$. For each configuration, the evolution of the target up to its complete expansion was simulated and the target gain calculated using the TERA package.

The results of combustion calculations for a target with $(\rho R)_0 = 3$ g/cm^2 at various initial major plasma temperatures T_0 can be considered typical. The dependences of the calculated target gains on igniter temperature T_f ($T_i = T_e = T_f$) at a fixed igniter thickness ($(\rho R)_f = 0.3$ g/cm^2) and two T_0 temperatures are shown in Fig. 2. According to this figure, there exist T_0-dependent critical igniter temperatures T_f near which a sharp increase in the efficiency of thermonuclear combustion occurs in a narrow temperature interval ($\Delta T_f \sim 0.2$ keV). Dependences of the same type were obtained for other $(\rho R)_0$ values, but the interval of critical values narrowed as target thickness increased. A similar result is obtained if calculation data are represented as the dependence of G on $(\rho R)_f$ at a fixed igniter temperature T_f [10, 19]. Changes of several percent in E_f near critical values change the target gain from $G \leq 1$ to $G \sim 10^3$.

The physical nature of the phenomenon is as follows. At near critical parameter values, the time of thermonuclear combustion wave propagation to the outside plasma boundary approximately coincides with the time of target expansion. At lower $(\rho R)_f$ and T_f values, there is sufficient time for plasma to expand hardly without a thermonuclear explosion, and the low efficiency of burning corresponds to "smoldering" of a uniform target without an igniter. If these values are high, a thermonuclear plasma explosion accompanied by a high burnup with a $G \sim 10^2$–10^3 target gain occurs. If an effective thermonuclear explosion does occur, the target gain reaches approximately the same maximum value irrespective of the configuration. In other words, the prehistory of the process weakly influences the effectiveness of burnup. The calculated target gains at igniter parameter values above critical tend to values given by (11) with an obvious but small correction to the additional igniter energy.

4. THE CRITICAL DIRECT-IGNITION TARGET IGNITER PARAMETERS

The critical igniter parameters were determined for targets of thicknesses $(\rho R)_0 = 3$, 4, and 6 g/cm^2. The calculations showed that the critical igniter parameters were virtually independent of target thickness $(\rho R)_0$ in the region of problem parameter values under consideration. The calculation results corresponding to various energy distributions between igniter electrons and ions can conveniently be compared by using the mean value

$$T_f = \frac{T_e + T_i}{2} \approx \frac{T_e}{2}$$

as igniter temperature T_f at $T_e \gg T_i$. Equal temperatures T_f then correspond to equal igniter energies in both limiting cases. The critical igniter parameters $(\rho R)_f$ and T_f calculated at a $T_0 = 0.6$ keV target temperature are shown in Fig. 3. For fairly small igniters, the critical ignition temperature and, accordingly, energy are several times larger when igniting driver energy is only transferred to plasma electrons than when energy is

equally distributed between electrons and ions. At $(\rho R)_f \geq 0.6$ g/cm^2, the corresponding critical parameter curves, however, virtually coincide.

To gain insight into the reasons for this phenomenon, it is expedient to pass to other variables. The most interesting question is that of the behavior of critical igniter heat energy $E_f \sim (\rho R)_f^3 T_f$, because it is directly related to the igniting driver energy. According to [10, 11], E_f decreases as the igniter size diminishes to some limiting value $(\rho R)_f \sim 0.3$–0.4 g/cm^2. When the mass of the igniter decreases further to less than 1% of the mass of the target, the critical E_f value becomes virtually independent of $(\rho R)_f$. In other words, some minimum ignition energy corresponds to each target, and a limiting igniter smallest in size corresponds to this energy. A similar result for targets with a degenerate plasma of the major fuel fraction was obtained in the calculations performed in [6].

The existence of a limiting igniter becomes still more obvious if the dependence of the critical ignition energy on igniter temperature rather than size is considered. As the absolute E_f value at fixed $(\rho R)_f$ and T_f depends on the density of the plasma, an analysis can conveniently be performed using the dimensionless relative igniter energy E_f/E_0. The dependences of the E_f/E_0 ratio on temperature T_f obtained by recalculating the critical igniter parameters shown in Fig. 3 to $T_i = T_e = T_f$ are plotted in Fig. 4, according to which the energy of the igniter becomes almost constant at $T_f >$ 10–12 keV. The minimum igniter energy E_f does not exceed several percent of the total target internal energy E_0, which is in agreement with the estimates made in Section 2.

The existence of the limiting igniter is related to the character of the temperature dependence of DT reaction rates. As long as the critical igniter temperature is below the values corresponding to a maximum reaction rate ($T \sim 15$–20 keV), a substantial decrease in the critical igniter size can be balanced by a small increase in temperature, because, in this region, the rate of thermonuclear reactions sharply increases as temperature grows. A temperature increase, however, gives no additional gain as the size of the igniter decreases further, because the rate of thermonuclear reactions is then virtually independent of temperature. Electronic heat conductivity smears the absorbed energy over the region corresponding to the critical temperature, which is situated close to a thermonuclear reaction rate maximum. Precisely this smearing determines the limiting igniter size.

The limiting igniter size is close to $(\rho R)_f \sim 0.4$ g/cm^2 for igniters with equal initial temperatures of electrons and ions. However, if the whole absorbed igniting driver energy is concentrated in the electronic igniter component, then, owing to electronic and radiative heat conductivity, part of this energy is distributed over the target volume while the ionic and electronic temperatures are equalized. This substantially increases the size of the hot spot and, accordingly, the energy of ignition. By way of illustration, the distributions of ionic and electronic temperatures in a target with density ρ = 100 g/cm^3 and the initial major fuel temperature T_0 = 1 keV at the initial time and at the instant when the ionic and electronic temperatures become equal, $t \approx$ 20 ps, are shown in Fig. 5. The initial igniter size was selected close to the limiting size of a one-temperature igniter with $(\rho R)_f$ = 0.4 g/cm^2. By the time of electron and ion temperature equalization, the size of the igniter

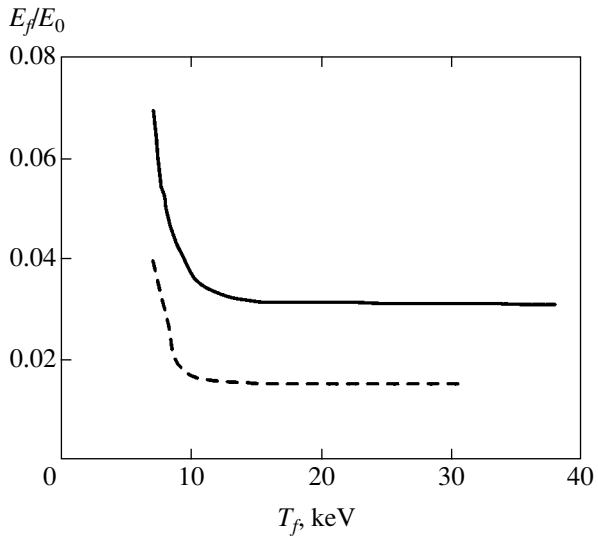


Fig. 4. Relative energy fraction $\Delta E = E_f/E_0$ of an igniter with critical parameter values as a function of igniter temperature T_f ($T_i = T_e = T_f$) for targets with T_0 = 1 and 0.6 keV major fuel temperatures (the solid and dashed lines, respectively).

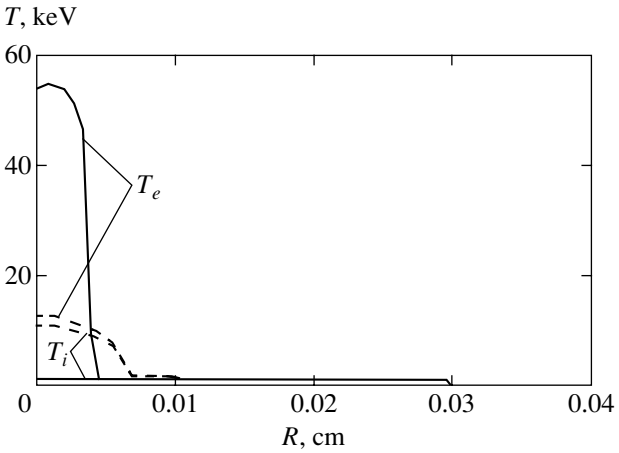


Fig. 5. Radial distributions of ionic and electronic temperatures in a target with density ρ = 100 g/cm^3 at the initial time (solid lines) and at the time of equalizing ionic and electronic temperatures ($t \approx 1.9 \times 10^{-2}$ ns, dashed lines).

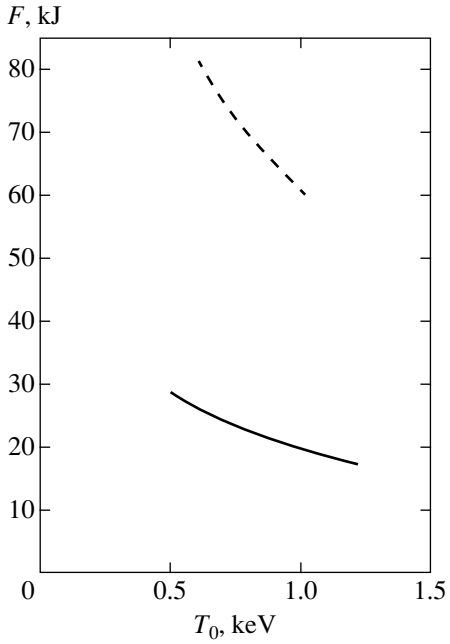


Fig. 6. Minimum normalized ignition energies $F(T_0) = E_f[(\rho/100\ \mathrm{g\ cm^{-3}})]^2$ at various major fuel temperatures T_0. The solid curve corresponds to equal additional energy distribution between igniter ions and electrons ($T_i = T_e = T_f$), and the dashed curve, to the situation when the whole additional energy is concentrated in electrons ($T_i = T_0$, $T_f = T_e/2$).

amounts to $(\rho R)_f \sim 0.6$ g/cm^2. Similar results are obtained for other initial conditions. Because of the smearing of the hot spot during temperature equalizing in a two-temperature igniter, the limiting size corresponding to a minimum ignition energy is close to $(\rho R)_f \approx 0.6$ g/cm^2. Precisely for this reason, energy values and, accordingly, critical temperatures coincide in the two limiting cases at $(\rho R)_f \geq 0.6$ g/cm^2 (see Fig. 3), and the minimum ignition energy in the second case increases three to four times because of an increase in the limiting igniter size.

The exact limiting igniter parameter values and, accordingly, minimum ignition energies depend on the T_0 temperature of the major fuel. The minimum igniter energies E_f for direct-ignition targets at various major fuel temperatures T_0 were determined by numerical calculations. The E_f value depends on target density ρ. As mentioned, thermonuclear combustion is an approximately scale-invariant process in the $r^* = \rho R$ and $t^* = \rho t$ variables. In other words, the $(\rho R)_f$ and T_f critical igniter parameters are independent of the density of the plasma. At fixed $(\rho R)_f$ and T_f, we have $E_f \sim \rho^{-2}$. (We checked this relation by numerically simulating the combustion wave for several isochorically compressed targets at various densities from $\rho = 10$ g/cm^3 to $\rho = 100$ g/cm^3. A tenfold density increase changed the scaled energy variable $E_f^* = \rho^2 E_f$ by less than 5%.) In order that our calculations can conveniently be compared with the results obtained by other authors, the ignition energy will be represented in the form

$$E_f = \frac{F(T_0)}{(\rho/100\ \mathrm{g\ cm^{-3}})^2}\ [\mathrm{kJ}].$$

The calculated temperature dependences of the minimum ignition energy $F(T_0) = E_f(\rho/100\ \mathrm{g/cm^3})^2$ for both limiting energy distributions between igniter electrons and ions are shown in Fig. 6.

When the temperatures of ions and electrons are equal, $F(T_0)$ decreases as temperature increases from $F = 30$ kJ at $T_0 = 0.5$ keV to $F = 15$ kJ at $T_0 = 1.2$ keV. These results are in agreement with the numerical calculations performed in [18] and the analytic estimates made in [19] for the minimum energy of an isochorically compressed target igniter in a cold degenerate plasma, $(rR)_f \geq 0.47$ g/cm^2, $T_f \geq 10$ keV and, accordingly, $E_f \geq 50/(\rho/100\ \mathrm{g/cm^3})^2$ kJ [$E_f \geq 43/(\rho/100\ \mathrm{g/cm^3})^2$ kJ].

In the other limiting case, when all the additional energy is only transferred to plasma electrons, igniter smearing during ionic and electronic temperature equalization substantially increases the minimum ignition energy ($F = 80$ kJ at $T_0 = 0.5$ keV and $F = 60$ kJ at $T_0 = 1.2$ keV). This result is also in reasonable agreement with the calculations of igniting a cold plasma by ion beams [6, 7, 20]. For instance, the results obtained in [7] in simulating the ignition of isochorically compressed targets with densities of $\rho = 100$ and 200 g/cm^3 by bismuth ion beams of a 15 GeV energy can be represented in the form $E_f \geq 160/(\rho/100\ \mathrm{g/cm^3})^2$ kJ. More recent calculations in a broader range of densities, $50 \leq \rho \leq 3000$ g/cm^3, gave a minimum of $E_f = 140/(\rho/100\ \mathrm{g/cm^3})^{1.85}$ kJ [20]. The small deviation of the last estimate from the usual similarity relation $E_f \sim \rho^{-2}$ is caused by the density dependence of the Coulomb logarithm ($\ln\Lambda \sim \ln\rho^{-1/2}$) [6] and the energy of the degenerate electron gas of the major fuel.

In conclusion, let us estimate the igniting driver energy required to provide the necessary direct-ignition target igniter energy for the example of one of the promising types of such drivers mentioned in the Introduction, namely, for a beam of light ions formed under the action of a narrow laser pulse with an energy of $I_L \sim 10^{20}$ W/cm^2 on a thin target. The degree of laser pulse energy conversion into the energy of light ions can be close to 10% [5]. Taking into account the calculation data shown in Fig. 6, we arrive at the conclusion that the laser pulse energy should be 100–150 kJ for a direct-ignition target with a nondegenerate major fuel.

5. CONCLUSION

The more fuel temperature exceeds the Fermi energy, the lower the ratio between the direct-ignition target gains obtained with nondegenerate and degenerate major fuel plasma states. Nevertheless, direct-ignition target gains, $G \sim 10^3$, remain fairly high at a $T_0 \geq 1$ keV temperature of the major fuel and a $(\rho R)_0 \geq 1$ g/cm^2 target thickness. They are several times higher than target gains with uniform plasma combustion. Direct target ignition occurs in a very narrow interval of igniter parameters close to critical (changes in igniter energy smaller than 5%). The critical minimum $(\rho R)_f$ and T_f parameters of an igniter in an isochorically compressed target, the attainment of which results in the formation of a thermonuclear combustion wave in time substantially shorter than the fly-off time, only depend on the major fuel temperature T_0 and are virtually independent of the thickness of the target.

If the major target fuel is in a nondegenerate state, thermonuclear combustion can be considered approximately scale-invariant in the variables $r^* = \rho R$, $t^* = \rho t$, $E^* = \rho^2 E$, and $M^* = \rho^2 M$. The exact scale invariance is only violated by a weak density dependence of the Coulomb logarithm. The plots given above allow igniter parameters to be estimated at arbitrary target densities.

The E_f critical igniter energy is a simple function of the T_f and $(\rho R)_f$ critical igniter parameters. For fairly large igniters, E_f decreases as the size of the igniter becomes smaller. Nevertheless, there exists a limiting igniter size starting with which E_f becomes size-independent. This limiting igniter size is determined by the existence of the optimal ignition temperature, $T_f \sim 10$ keV. The corresponding minimum ignition energies E_f depend only on the initial temperature of the major fuel fraction.

The critical parameters and energy essentially depend on the initial energy distribution between igniter electrons and ions. At equal initial temperatures ($T_e = T_i$), the limiting igniter size is around $(\rho R)_f \sim 0.4$ g/cm^2 (the exact size depends on the major fuel temperature). If the whole absorbed igniting driver energy is contained in the electronic igniter component, the size of the hot spot substantially increases during equalization of the temperatures of ions and electrons, and the limiting igniter size is then $(\rho R)_f \sim 0.6$ g/cm^2. Accordingly, the minimum ignition energy increases three to four times. The critical temperature and energy for igniters with $(\rho R)_f \gg 0.6$ g/cm^2 are virtually independent of energy redistribution between ions and electrons at the initial time.

ACKNOWLEDGMENTS

The authors are deeply indebted to A.A. Samarskii for support and N.V. Zmitrenko for valuable discussions. This work was financially supported by the Russian Foundation for Basic Research (project no. 99-02-16100), the Ministry of Education of the Russian Federation (grant "Thermonuclear and Nuclear Reactors 2001"), and the Universities of Russia Program 015 (project no. 03.01.05).

REFERENCES

1. N. G. Basov, S. Yu. Gus'kov, and L. P. Feoktistov, in *Proceedings of the 21st European Conference on Laser Interaction with Matter, Warsaw*, Ed. by H. Fieodorowich *et al.* (1991), p. 189.
2. N. G. Basov, S. Yu. Gus'kov, and L. P. Feoktistov, J. Sov. Laser Res. **13**, 396 (1992).
3. M. Tabak, J. Hammer, M. E. Glinsky, *et al.*, Phys. Plasmas **1** (5), 1626 (1994).
4. A. Caruso, in *Proceedings of I. A. E. A. Technical Committee Meeting on Drivers for Inertial Confinement Fusion, Paris, 1994*, Ed. by J. Coutant (1995), p. 325.
5. A. Caruso, S. Yu. Gus'kov, V. B. Rozanov, and C. Strangio, in *Abstracts of the 26th European Conference on Laser Interaction with Matter, Prague* (2000), p. 56.
6. S. Atzeny, Phys. Plasmas **6**, 3316 (1999).
7. A. Caruso and V. A. Pais, Nucl. Fusion **36**, 745 (1996).
8. A. A. Andreev, D. V. Il'in, A. A. Levkovskii, *et al.*, Izv. Ross. Akad. Nauk, Ser. Fiz. **63**, 1182 (1999).
9. A. A. Andreev, D. V. Il'in, O. B. Vygovsky, *et al.*, Proc. SPIE **3683**, 170 (1999).
10. A. Levkovskii, A. Andreev, D. Il'in, *et al.*, Proc. SPIE **3886**, 448 (2000).
11. A. A. Andreev, S. Yu. Gus'kov, D. V. Il'in, *et al.*, Zh. Éksp. Teor. Fiz. **119**, 80 (2001) [JETP **92**, 69 (2001)].
12. J. J. Duderstadt and G. A. Moses, *Inertial Confinement Fusion* (Wiley, New York, 1982).
13. K. A. Brueckner and S. Jorna, Rev. Mod. Phys. **46**, 325 (1974).
14. S. Yu. Gus'kov, O. N. Krokhin, and V. B. Rozanov, Nucl. Fusion **16**, 957 (1976).
15. R. Kidder, Nucl. Fusion **16**, 405 (1976).
16. A. A. Levkovskii, Preprint No. 73, FIAN SSSR (Physical Inst., USSR Academy of Sciences, Moscow, 1990).
17. S. Yu. Gus'kov, N. V. Zmitrenko, D. V. Il'in, *et al.*, Zh. Éksp. Teor. Fiz. **106**, 1069 (1994) [JETP **79**, 581 (1994)].
18. S. Atzeny, Jpn. J. Appl. Phys., Part 1 **34**, 1980 (1995).
19. A. R. Piriz and M. M. Sanchez, Phys. Plasmas **5**, 2721 (1998).
20. S. Atzeny, M. L. Ciampi, and A. R. Piriz, in *Proceedings of the 23rd ECLIM Conference* (World Sci., Singapore, 1997), p. 275.

Translated by V. Sipachev

Journal of Experimental and Theoretical Physics, Vol. 96, No. 4, 2003, pp. 704–718.
Translated from Zhurnal Éksperimental'noĭ i Teoreticheskoĭ Fiziki, Vol. 123, No. 4, 2003, pp. 798–815.

PLASMA, GASES

Dynamics of Macroparticles in a Dusty Plasma under Microgravity Conditions (First Experiments on Board the ISS)

V. E. Fortov[a], O. S. Vaulina[a], O. F. Petrov[a,*], V. I. Molotkov[a], A. V. Chernyshev[a], A. M. Lipaev[a], G. Morfill[b], H. Thomas[b], H. Rotermell[b], S. A. Khrapak[b], Yu. P. Semenov[c], A. I. Ivanov[c], S. K. Krikalev[c], and Yu. P. Gidzenko[d]

[a]*Institute of High Energy Densities, IVTAN, Russian Academy of Sciences, Moscow, 127412 Russia*
[b]*Institute of Extraterrestrial Physics, M. Planck Society, Garhing, 85740 Germany*
[c]*Korolev Energiya Rocket-Building Corporation, Korolev, Moscow oblast, 141160 Russia*
[d]*Gagarin Center for Cosmonaut Training, Moscow oblast, 141160 Russia*
**e-mail: industpl@redline.ru*

Received November 6, 2002

Abstract—The results of experimental investigation of macroparticle transport in the dusty plasma of a capacitive high-frequency discharge under microgravity conditions are considered. Experimental data were obtained for monodisperse polymer particles of radius $a_p = 1.7$ μm in a wide range of plasma parameters on the International Space Station. Analysis of macroparticle dynamics for a strongly nonideal dusty plasma, including diffusion and dust vortex formation processes, is carried out. *© 2003 MAIK "Nauka/Interperiodica".*

1. INTRODUCTION

A dusty plasma is a strongly ionized gas containing charged particles of a substance (dust) of micrometer size. Such plasmas are good experimental models for studying various transport effects in systems of interacting particles which are of utmost importance for the physics of a nonideal plasma as well as for other fields of natural science such as the physics of condensed media, chemistry, physics of the atmosphere, and astronomy.

Owing to their considerable size, dust particles in a laboratory plasma can be recorded by a video camera, which simplifies the application of direct contactless methods for their diagnostics. The main mechanism of dust particles charging in a gas-discharge plasma is associated with flows of electrons and ions. Due to higher mobility of electrons, micrometer-size macroparticles can acquire a considerable negative charge (on the order of $(10^3–10^5)e$, where e is electron charge) and may interact electrostatically with one another.

The main source of the kinetic energy dissipation for macroparticles in a weakly ionized plasma of gas discharges is their collisions with neutrals of a buffer gas. It should be noted that a laboratory dusty plasma is an open dissipative system on account of external fields (electric, gravity, etc.) and forces (ion drag, thermophoresis, etc.) exerted on particles by the surrounding plasma, as well as processes of macroparticle charging determined by the flows of the surrounding plasma to the particle surface. When plasma parameters change, the charge of macroparticles can be a function of time and the position of a particle. The joint action of the forces of interaction between particles and dissipative processes in such a plasma may lead to the formation of steady-state dust structures (resembling a liquid or a solid) as well as to complex vibrational or stochastic modes [1–10]. Under standard laboratory conditions, the observed dust structures are confined in the Earth's gravity field by the electric field of the strata (in a d.c. glow discharge) or by the electrode layer (in an rf discharge) [1–3], while the gravity field sets a limit on the experimental results.

In recent years, considerable attention is paid to experimental studied of dusty plasmas under microgravity conditions [6–11]. Such experiments make it possible to study a wide range of phenomena (photoemissive charging of atmospheric aerosols, ambipolar diffusion, dynamics of massive dust particles for a large size (>100 μm), and so on), which cannot be observed in terrestrial laboratory conditions [6–8, 11]. One of the important advantages of experiments under microgravity conditions is the possibility of operating in a wide range of dusty plasma parameters, which is not limited by the necessity of ensuring levitation of particles in the gravity field. In recent experiments with rf-discharge dusty plasma made on board the International Space Station (ISS), a group of Russian and German scientists discovered a number of new effects (formation of complex crystal lattices, nonlinear waves, unsimilar charg-

1063-7761/03/9604-0704$24.00

ing of macroparticles, etc.), which have no analogs in ordinary laboratory conditions [10]. This experiment is known as Plasma Crystal-3 (PC-3) since it was a continuation of the research work devoted to study of dusty plasmas under microgravity conditions, which was started in 1997 on board the Mir orbital complex by a group of scientists from the Institute of High Energy Densities together with the Energiya Rocket-Building Corporation [6–8, 11].

In this paper, we report on a part of the results obtained in the PC-3 experiments and pertaining to analysis of transport processes such as macroparticle diffusion in a strongly nonideal dusty plasma and the dynamics of formation of dust vortices. In Section 2, various transport characteristics of dust particles in a plasma (such as charge, diffusion coefficient, and pair correlation function) are considered; the results of numerical simulation of macroparticle dynamics in Yukawa systems are given; and one of the possible mechanisms for dust vortex formation is considered. Section 3 deals with the results and analysis of experiments carried out for monodisperse polymer particles of radius $a_p \approx 1.7$ μm (density $\rho_p \approx 1.5$ g cm^{-3} and mass $m_p \approx 3.1 \times 10^{-11}$ g) in a wide range of plasma parameters, which was ensured by variation of pressures P = 0.36–0.98 mbar and power W = 0.15–0.98 W of a capacitive rf discharge in argon. The number density n_p of macroparticles in a dust cloud varied from 0.95×10^5 to 1.25×10^5 cm^{-3}.

The PC-3 experiment is the first physical experiment made on board the ISS. The experimental equipment was delivered to the ISS in February 2001; the first series of experiments was carried out by S. Krikalev and Yu. Gidzenko with the participation of American astronaut W. Sheppard in March 2001.

Distinguished service in staging and implementation of the PC-3 experiment was rendered by Prof. A.P. Nefedov, who passed away a short time before the experiment was launched. The researchers participating in this experiment devote it to fond memories of him.

2. TRANSPORT CHARACTERISTICS OF MACROPARTICLES IN DUSTY PLASMA

2.1. Macroparticle Charge

The kinetics of dust particle charging in a plasma is described by the equation

$$\frac{dZ_p}{dt} = \sum_j I_j, \qquad (1)$$

where summation is carried out over all charged particle fluxes I_j absorbed or emitted by a dust particle. In the steady state, $dZ_p/dt = 0$, which determines the equilibrium charge $\langle Z_p \rangle$ of dust particles. For a spherical particle, the relation between the dust particle charge Z_p and potential φ_s of its surface is determined by the relation

$$Z_p e \approx a_p \varphi_s, \quad a_p \ll \lambda,$$

where λ is the radius of screening the dust particle by ions and electrons of the surrounding plasma.

Most theoretical models of charging particles with a size $a_p < l_{e(i)}$ ($l_{e(i)}$ is the mean free path of electrons (ions) in a plasma) are based on the orbital motion limited approximation (OML theory). The calculation of electron (ion) fluxes $I_{e(i)}$ with a Maxwellian velocity spectrum and temperature $T_{e(i)}$ to a negatively charged spherical particle in the OML theory gives the following expression for the equilibrium charge of particles [12, 13]:

$$\exp(-z) = \frac{v_{Ti}}{v_{Te}}\left(1 + z\frac{T_e}{T_i}\right)(1 + z\chi); \qquad (2)$$

here,

$$v_{Te(i)} = \left(\frac{8T_{e(i)}}{\pi m_{e(i)}}\right)^{1/2}$$

is the mean thermal velocity of electrons (ions); $m_{ie(i)}$ and $n_{e(i)}$ are their mass and concentration, respectively; the value of parameter χ is given by the relation

$$\chi = \frac{|\langle Z_p \rangle| n_p}{z n_e}, \qquad (3)$$

and parameter z is proportional to the ratio of the surface potential φ_s of a dust particle to the electron temperature:

$$z = \frac{|\langle Z_p \rangle| e^2}{a_p T_e}. \qquad (4)$$

The value of parameter z is determined by the buffer gas ions; for argon, $z \approx 2$–4 for most experimental conditions in a gas-discharge plasma [12–14]. Numerical calculations show that the value of parameter z for thermal fluxes of electrons (ions) for $\chi < 1$ changes insignificantly and is close to the value $z \approx 2.9$ for a solitary particle. In a dense dust cloud with $\chi > 1$, an increase in the values of z and φ_s takes place [12, 14], which can be partly compensated in some cases due to an increase in the electron temperature under the conditions of variation of equilibrium ionization processes in the gas discharge [11].

Since discharges in noble gases are usually controlled by ambipolar diffusion (plasma recombination at the walls of the gas-discharge tube), the effect of dust particles on equilibrium ionization processes is significant only if the electron loss rate ν_{ep} at particles in the

dust cloud is comparable with or much higher than the rate ν_{ab} of diffusion losses of electrons [11, 15]:

$$\nu_{ab} \approx D_a/\Lambda_d^2, \tag{5}$$

where $D_a \approx \mu_i T_e/e$ is the ambipolar diffusion coefficient for $T_e \gg T_i$, μ_i is the mobility of ions ($P\mu_i \approx$ 1250 Torr cm^2/V s for argon [15]), and Λ_d is a certain characteristic scale. For a cylinder of radius R and length L, we have [15]

$$\Lambda_d^2 \approx ((2.4/R)^2 + (\pi/L)^2)^{-1}.$$

The electron loss rate ν_{ep} at particles in the dust cloud can be estimated as

$$\nu_{ep} \approx \pi a_p^2 n_p v_{Te} \exp(-z), \tag{6}$$

whence we find that electron loss rate ν_{ep} at particles in the dust cloud under typical experimental conditions for an rf discharge in argon ($P \approx 0.25$–1 Torr, $T_e \approx$ 1–3 eV) with particles of radius $a_p \approx 2$ μm for $\Lambda_d \approx$ 1 cm and $z = 2$–4 is comparable with or higher than the rate ν_{ab} of diffusion loss ($\nu_{ep} > \nu_{ab}$) for dust concentrations $n_p > 10^4$ cm^{-3}. Thus, we can assume that the operation of a discharge under the experimental conditions (see Introduction, $n_p \approx 10^5$ cm^{-3}) may be determined to a considerable extent by recombination of plasma at the surface of dust particles.

2.2. Particle Interaction Potential

It is generally assumed that dust particles in a weakly ionized plasma interact with one another through the screened Coulomb potential (Yukawa potential)

$$\varphi = eZ_p \exp(-l/\lambda)/l, \tag{7}$$

where l is distance. This assumption contradicts the results of measurements of radial (perpendicular to the force of gravity on the Earth) forces of interaction between two particles, which were made in [16]. Nevertheless, we can indicate at least two reasons for which the form of the potentials determining the particle interaction forces may differ significantly from that given above [16–22]. The first reason is associated with the existence of attractive forces between dust particles due to polarization of the surrounding plasma or other effect caused by "dimming out" of directional plasma flows [17–20]. The existence of attractive forces in dust systems was actively studied in a number of theoretical and experimental works. However, convincing experimental proofs confirming the existence of attractive forces in dust systems have not been obtained as yet.

The second reason is that the floating potential φ_s at the particle surface is equal in order of magnitude to the electron energy and is considerably higher than the ion energy; for this reason, the screening of particles is essentially nonlinear. In addition, the electron–ion recombination takes place at the particle surface; as a result, there us no back current of ions (electrons) in the vicinity of a dust particle and their distribution functions are not Maxwellian. Consequently, the asymptotic behavior of a potential at large distances from a dust particle does not obey relation (7) any longer, but exhibits a power dependence on the distance:

$$\varphi \approx eZ_p a_p/l^2.$$

This effect has been comprehensively studied in the theory of spherical electric probes [22]. The structure of the screening cloud in the collisionless mode was calculated in [21] taking into account the nonlinearity of the Poisson equation and non-Maxwellian velocity spectra for plasma particles. It was proved that the screening of solitary particles is determined by the value of λ_e for a particle size of $a_p \approx 2\lambda_i - \lambda_e$ (λ_i and λ_e are the ionic and electronic Debye radii) and by the value of λ_i for particles with a size $a_p \ll \lambda_i$ at distances $l < l_D \approx (5–7)\lambda_i$ from the particle surface. As the particle size increases to $a_p > \lambda_e$, the effective screening length λ may become much larger than λ_e. The existence of a non-Debye screening in a dust cloud with concentration n_p at mean distances $l_p = n_p^{-1/3} > l_D$ may suppress the effect of the concentration of the surrounding plasma on the particle interaction in the dust cloud and, as a result, enhance the repulsion between particles.

2.3. Simulation of Macroparticle Dynamics in a Dusty Plasma

Correct simulation of the transport of macroparticles in a dusty plasma requires the application of the molecular dynamics method based on the solution of a system of ordinary differential equation with the Langevin force F_{br} taking into account random impacts of surrounding gas molecules or other stochastic processes leading to an increase in their kinetic temperature T_p [3, 5, 14, 23] over the temperature T_n of the surrounding gas. In simulation of microscopic transport processes in homogeneous extended clouds of interacting macroparticles, pair particle interaction forces F_{int} are taken into account in the system of N_p equations of motion (N_p is the number of particles) in addition to random forces F_{br}, which are sources of stochastic motion of particles,

$$m_p \frac{d\mathbf{l}_k}{dt^2} = \sum_j F_{\text{int}}(l)\big|_{\mathbf{l}=|\mathbf{l}_k-\mathbf{l}_j|} \frac{\mathbf{l}_k - \mathbf{l}_j}{|\mathbf{l}_k - \mathbf{l}_j|} - m_p \nu_{fr} \frac{d\mathbf{l}_k}{dt} + \mathbf{F}_{br}; \tag{8}$$

in this case, periodic boundary conditions are used.

Here,

$$F_{\text{int}}(l) = -eZ_p\frac{\partial\varphi}{\partial l},$$

$l = |\mathbf{l}_k - \mathbf{l}_j|$ is the particle spacing; m_p is the particle mass; and ν_{fr} is the friction coefficient for dust particles, which is determined by the frequency of their collisions with neutrals of the surrounding gas and can be derived in the framework of the free molecular approximation if the particle size is smaller than the mean free path l_n of the buffer gas neutrals ($a_p \ll l_n$) [24]. For argon at room temperature, l_n [μm] ≈ 62/P [mbar], while the friction coefficient for a spherical particle can be written in the form ν_{fr} [s^{-1}] ≈ 860P [mbar]/(a_p [μm]ρ_p [g cm^{-3}]).

Under the conditions of local thermodynamic equilibrium of the dust system, the mean value of the random force is

$$\langle F_{br}\rangle = 0,$$

while the autocorrelation function [25, 26]

$$\langle F_{br}(0)F_{br}(t)\rangle = 6T_p m_p \nu_{fr}\delta(t)$$

describes a delta-correlated Gaussian process. Here, $\delta(t)$ is the delta function and the angle brackets denote time averaging. In simulating equilibrium random processes in a dust cloud, use can be made of random increments of the macroparticle momentum,

$$p_{br}^x = m_p(2T_p\nu_{fr}\Delta t/m_p)^{1/2}\psi_j,$$

where p_{br}^x is the momentum increment per degree of freedom and ψ_j is a random quantity distributed in accordance with the normal law with the standard deviation equal to unity. In a correct simulation of random forces, the step Δt of integration with respect to time in Eqs. (8) must satisfy the condition

$$\Delta t \ll \max\{\nu_{fr}, \omega^*\},$$

where the quantity ω^* characterizes the frequency of collisions between charged macroparticles.

For liquid Yukawa systems with screening parameter

$$\kappa = l_p/\lambda < 6\text{–}7,$$

the characteristic dust frequency ω^* can be represented in the form [27]

$$\omega^* = eZ_p^*(n_p/\pi m_p)^{1/2}, \quad (9)$$

where Z_p^* is the effective charge of macroparticles:

$$Z_p^* = Z_p\{(1+\kappa+\kappa^2/2)\exp(-\kappa)\}^{-1/2}. \quad (10)$$

The relation between the particle interaction and dissipation in the dust system is determined by the ratio

$$\xi = \omega^*/\nu_{fr}, \quad (11)$$

which is the scaling parameter for various dynamic processes in dissipative dust structures, e.g., for diffusion of macroparticles and for the dynamics of formation of vortex motion [23, 26–28]. Analysis of the experimental conditions for particles of radius $a_p \approx 2$ μm upon a change in their concentration n from 10^3 to 10^6 cm^{-3} and for variation of the gas pressure P from 1 to 0.01 Torr leads to the following range for the scaling parameter:

$$\xi \approx 0.02\text{–}5.$$

The dynamics of macroparticles in a dissipative ($\nu_{fr} \neq 0$) extended Yukawa system was studied by several authors for parameters close to the experimental conditions in a gas-discharge plasma ($\xi \approx 0.02$–4) [26–30]. The results of simulation show that the buffer gas viscosity practically does not affect the degree of correlation of macroparticles in such systems, and analysis of the phase state of the three-dimensional systems under investigation can be based on the effective nonideality parameter

$$\Gamma^* = (Z_p^* e)^2/T_p l_p. \quad (12)$$

The normalized parameter Γ^* completely determines the correlation of macroparticles (the emergence of long-range and short-range orders in a dust system) in Yukawa dissipative systems starting from $\Gamma^* < 1$ to the point of system crystallization. As the parameter increases so that $\Gamma^* \longrightarrow \Gamma_m^* = 106$ (for $\kappa < 6$), a body-centered crystalline structure is formed [29–32]. It is proposed that a conversion of a nonideal system into a strongly correlated liquid occurs for values of parameter $\Gamma^* \approx 22$–24, for which the formation of ordered groups of macroparticles (clusters), which is accompanied by a jumplike decrease in the coordinate of the first peak of the pair correlation function and the diffusion coefficient for macroparticles, is observed in the course of a numerical experiment [29, 30].

The dependences of the heights of the first peaks $g_{\max}$ of the pair correlation function $g(l)$ and their positions ($l = d_{\max}$) on parameter Γ^* (12) are shown in Fig. 1a for different parameters of the Yukawa system. A sharp increase (jump) in the value of $g_{\max}$ from 2.65 to 3.1 is observed in the range of the normalized nonideality parameter Γ^* from the point of crystallization $\Gamma_c^* \approx 102$–104 to the melting point $\Gamma_m^* \approx 106$–107 of the system, where the position of the first peak of function $g(l)$ changes from $l \approx 1.075l_p$ to the characteristic body-centered lattice spacing $d_{\max} \approx (3\sqrt{3}/n_p 4)^{1/3} \approx 1.092l_p$ (see Fig. 1b) [29, 30]. In the same range of non-

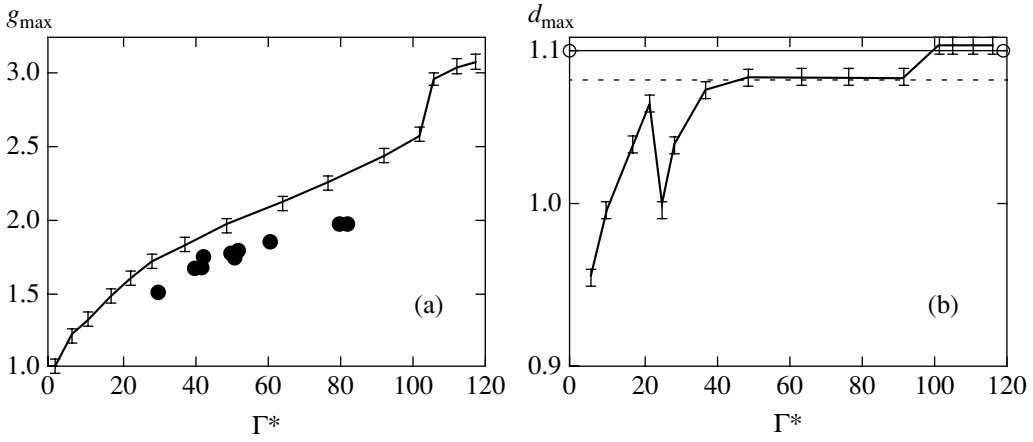


Fig. 1. Dependences of (a) the first maxima g_{max} of the pair correlation function and (b) their relative position d_{max}/l_p on Γ*. Intervals of absolute deviations of values for various cases of calculation (ξ = 0.05–1.22) are indicated. Dark circles show the results of determining g_{max} for experimental correlation functions depending on the values of Γ* reconstructed from the measurements of the macroparticle diffusion coefficient *D*.

ideality parameters (Γ* ≈ 102–107), the diffusion coefficient of particles, which was obtained by simulation, decreases abruptly (by more than two orders of magnitude) [16, 29, 30].

The diffusion coefficient for interacting macroparticles can be derived from the relation [26]

$$D(t) = (\langle \mathbf{l}(t) - \mathbf{l}(0)\rangle_N^2)_t/6t, \qquad (13)$$

where $\mathbf{l}(t)$ is the displacement of an individual particle and $\langle\rangle_N$ and $\langle\rangle_t$ denote the averaging over ensemble *N* and time *t*, respectively. This relation follows from analysis of the diffusion transfer of particles through a unit area element in a homogeneous medium and is similar to the well-known Green–Kubo formula [33]. While deriving these formulas, no assumptions concerning the type of thermal motion have been made; consequently, these expression are valid for gases as well as for liquids and solids.

Figure 2 shows the time dependence of the ratio of the diffusion coefficient $D(t)$ for interacting particles to the diffusion coefficient for noninteracting Brownian particles,

$$D_0 = \frac{T_p}{\nu_{fr} m_p}$$

(in reciprocal drag times ν_{fr}^{-1}) for Yukawa systems with different parameters $\xi = \omega^*/\nu_{fr}$ and Γ*. Curve *6* depicts the exact solution of the Langevin equation for noninteracting particles [26]:

$$\frac{D_0(t)}{D_0} = 1 - \frac{1 - \exp(-\nu_{fr} t)}{\nu_{fr} t}; \qquad (14)$$

consequently, we have $D_0(t) = D_0$ for large values of time as compared to the reciprocal frequency of friction ($\nu_{fr} t \gg 1$), while the ballistic nature of motion of a particle is manifested for small values of time ($\nu_{fr} t \ll 1$),

$$\langle \Delta^2 l(t)\rangle = \langle \langle \mathbf{l}(t) - \mathbf{l}(0)\rangle_N^2\rangle_t \approx 3 v_{Tp} t^2$$

and $D_0(t) \propto t$. In the presence of interaction (see Fig. 2), the behavior of the quantity $D(t)$ for small values of time remains unchanged. As the time increases, this quantity attains its maximum value, D_{max}, i.e., the value which is expedient for determining the short-term diffusion coefficient. It should be noted that coefficient D_{max} is smaller than D_0 and tends to the latter upon an

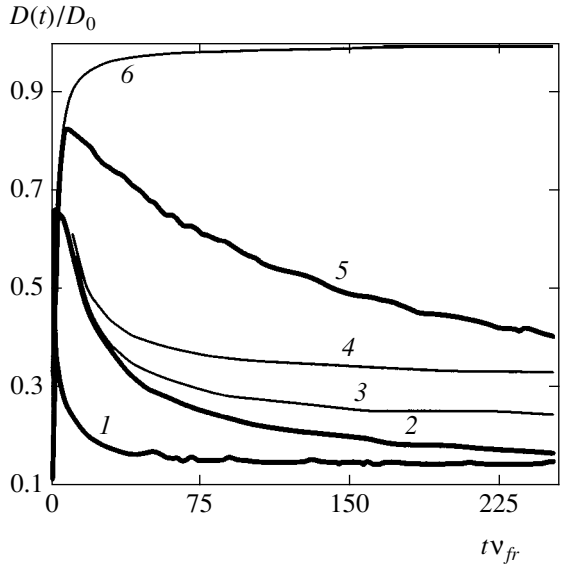


Fig. 2. Dependences of the ratio of the diffusion coefficient $D(t)$ of charged particles to D_0 on $t\nu_{fr}$ for various values of parameters ξ and Γ*: ξ = 1.22, Γ* = 80 (*1*); ξ = 0.14, Γ* = 80 (*2*); ξ = 0.14, Γ* = 60 (*3*); ξ = 0.14, Γ* = 30 (*4*); ξ = 0.04, Γ* = 80 (*5*); $D_0(t)/D_0$ (*6*).

increase in the medium viscosity. The dependence of the maximum of the ratio $D(t)/D_0$ and its position $(t\nu_{fr})_{\max}$ on parameter ξ is shown in Fig. 3. It should be noted that neither $D_{\max}$ nor the position of the maximum $(t\nu_{fr})_{\max}$ depends on the value of Γ^*, and the behavior of function $D(t)/D_0$ for values of time

$$t < \frac{(t\nu_{fr})_{\max}}{2\nu_{fr}}$$

is determined by relation (14) for noninteracting particles. Analysis of the behavior of $D(t)/D_0$ for small values of observation time can be useful for reconstructing coefficient D_0 or macroparticle temperature T_p (for a given ν_{fr} from the results of measurements of $D(t)$ if the resolution of the measuring instrument does not permit correct determination of the velocity spectrum of macroparticles. For example, if the displacements of macroparticles is recorded by a video camera with a recording frequency smaller than the friction coefficient ν_{fr}, the measurements of the instantaneous velocity distributions for macroparticles give too low a value of their temperature [34].

With increasing time, the diffusion coefficient (13) tends to a constant value $D = \lim_{t\to\infty} D(t)$, which can be easily determined experimentally while studying a dusty plasma [7, 11]. The ratio of the diffusion coefficient D for interacting macroparticles multiplied by $(1 + \xi)$ to coefficient D_0 is shown in Fig. 4. It can be seen that function $D(1 + \xi)/D_0$ depends only on the effective parameter Γ^* for weakly correlated as well as strongly nonideal systems. In the latter case (for $\Gamma^* > 50$), the diffusion coefficient for interacting macroparticles has the form [27, 30]

$$D \approx \frac{T_p\Gamma^*}{12\pi(\omega^* + \nu_{fr})m_p}\exp\left(-c_1\frac{\Gamma^*}{\Gamma_c^*}\right), \qquad (15)$$

where $c_1 = 2.9$ for $\xi > 0.3$ and $c_1 = 3.15$ for $\xi < 0.3$. Approximation of the results of calculations by formula (15) taking into account the difference in coefficients c_1 is also shown in Fig. 4. Thus, we can obtain the following two simple relations for the diffusion coefficient D in two limiting cases:

$$D \approx \frac{l_p}{12}\sqrt{\frac{T_p\Gamma^*}{\pi m_p}}\exp\left(-2.9\frac{\Gamma^*}{\Gamma_c^*}\right), \quad \omega^* \gg \nu_{fr}, \qquad (16a)$$

$$D \approx D_0\frac{\Gamma^*}{12\pi}\exp\left(-3.15\frac{\Gamma^*}{\Gamma_c^*}\right), \quad \omega^* \ll \nu_{fr}. \qquad (16b)$$

These relations make it possible to easily determine the effective parameter Γ^* from the results of measurements of the mean particle spacing, temperature, and diffusion coefficient of particles in liquid systems and can be useful for experimental analysis of particle parameters such as their charge and screening radius.

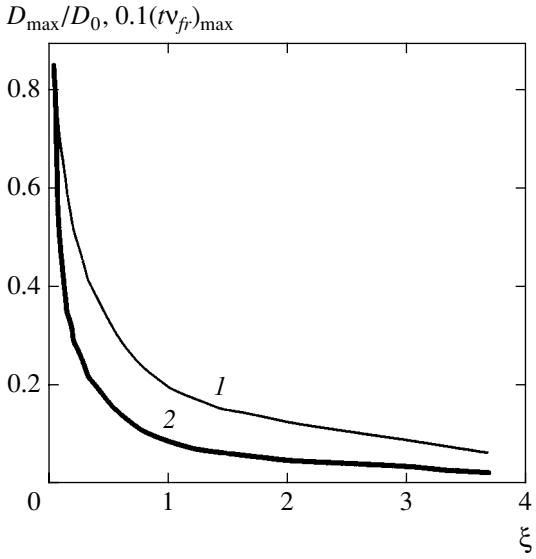


Fig. 3. Maximum of the ratio D/D_0 (*1*) and its position $(t\nu_{fr})_{\max}$ (*2*) as functions of parameter ξ.

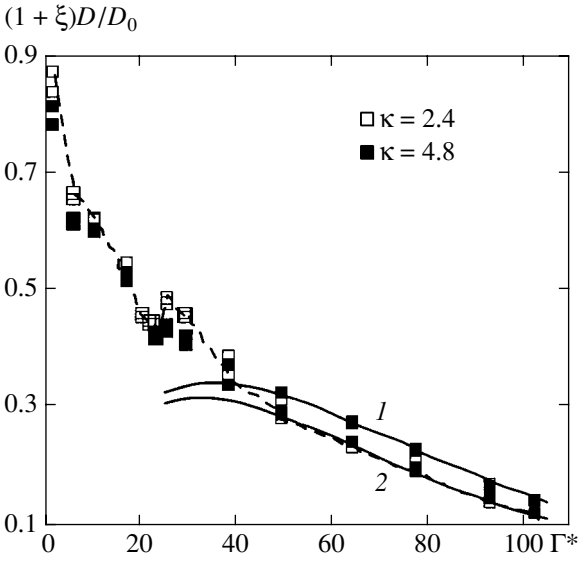


Fig. 4. Dependence of function $(1 + \xi)D/D_0$ on Γ^* for various values of ξ and screening parameters κ as well as the approximation of this function by relation (15) for $\xi > 0.3$ (*1*) and $\xi < 0.3$ (*2*).

2.4. Vortex Motion of Macroparticles in a Heterogeneous Plasma

Owing to the high mobility of electrons, nonemitting dust particles acquire a negative equilibrium charge matching the parameters of surrounding plasma (see Section 2.1). Upon variation of plasma parameters,

this charge can exhibit a dependence on time and on the position of a macroparticle. The space and time variation of dust charges is one of the possible factors permitting the conversion of the energy of the electric field into the kinetic energy of macroparticles [3, 5, 14, 23]. Random fluctuations of the macroparticle charge may cause an "anomalous heating" of dust, but fail to explain the self-excitation of regular movements of macroparticles and the formation of dynamic dissipative structures from them. Theoretical analysis and simulation show that such regular dust self-sustained oscillations can be formed in the presence of a gradient $\boldsymbol{\beta} = e\nabla Z_p$ of the dust particle charge and nonelectrostatic forces $\mathbf{F}_{\text{non}}$ (such as gravity force, thermophoretic force, or ion drag force) orthogonal to gradient $\boldsymbol{\beta}$ and acting on macroparticles of the dust cloud [23]. Spatial variations of macroparticle charges in a plasma may appear due to inhomogeneity of the background plasma surrounding a dust cloud, e.g., gradients of concentration $n_{e(i)}$ or temperatures $T_{e(i)}$ of electrons (ions). Such conditions are often created in the plasma of an inductive rf or glow discharge [35, 36]. For example, the charge gradient $\boldsymbol{\beta}$ of macroparticles in a discharge controlled by ambipolar diffusion may attain the values

$$\beta \approx (0.1\text{–}0.3)e\langle Z_p\rangle \ \text{cm}^{-1}$$

only due to weak violation of electroneutrality

$$\delta n = |n_i - n_e| \ll n_i \approx n_e$$

of the surrounding plasma [28].

Let us consider the motion of N_p macroparticles with a charge $Z(\rho, y) = Z_0 + \Delta Z(\rho, y)$ (where $\rho = (x^2 + z^2)^{1/2}$ is the radial coordinate of a particle and $Z_0 = Z(0, 0)$ in an electric field $\mathbf{E}(\rho, y)$ of a cylindrical trap, taking into account the pair interaction, random motion of particles (F_{br}), and a certain constant nonelectrostatic force

$$|\mathbf{F}_{\text{non}}| = F_{\text{non}}(y) \equiv \text{const},$$

acting along the y axis in the cylindrical system in question:

$$m_p\frac{d^2\mathbf{l}_k}{dt^2} = -m_p\nu_{fr}\frac{d\mathbf{l}_k}{dt} + \mathbf{F}_\Sigma + \mathbf{F}_{\text{non}} + \mathbf{F}_{br}. \qquad (17)$$

Here,

$$\mathbf{l}_k(\rho, y) = \mathbf{i}y + \mathbf{j}\rho,$$

$$\mathbf{F}_\Sigma = \mathbf{F}_{\text{int}} + \mathbf{F}_{\text{ext}}$$

is the resultant electric force, where

$$\mathbf{F}_{\text{int}} = \sum_j eZ(\rho, y)\frac{\partial\varphi}{\partial l}\bigg|_{\mathbf{l} = |\mathbf{l}_k - \mathbf{l}_j|}\frac{\mathbf{l}_k - \mathbf{l}_j}{|\mathbf{l}_k - \mathbf{l}_j|}$$
$$\equiv F^y_{\text{int}}(\rho, y)\mathbf{i} + F^\rho_{\text{int}}(\rho, y)\mathbf{j}$$

is the force exerted on an individual macroparticle by all other macroparticles of the dust cloud, while the external force is defined as

$$\mathbf{F}_{\text{ext}} = eZ(\rho, y)\mathbf{E}(\rho, y) \equiv \mathbf{i}F^y_{\text{ext}}(\rho, y) + \mathbf{j}F^\rho_{\text{ext}}(\rho, y).$$

When the curl of forces acting on a macroparticle differs from zero, system (17) can perform positive work in compensating dissipative energy losses. This means that infinitely small perturbations emerging in the system due to thermal and other fluctuations may be augmented. A detailed analysis of linearized equations (17) is given in [23], where two main cases of instability of the dust system are considered. Under certain conditions, one of the instabilities considered above ("dissipative" or "absolute" instability [24, 37]) may lead to the formation of large-scale vortex movements (dissipative dust structures). When

$$|Z_0 \approx \langle Z_p\rangle| \gg |\Delta Z(\rho, y)|,$$

the emergence of dissipative-type instability is determined by the condition

$$\omega_c^4 < \left|\frac{\gamma_0\beta_\rho F_{\text{non}}}{m_p}\frac{1}{Z_0}\right|, \qquad (18)$$

where

$$\beta_\rho = \frac{\partial Z(\rho, y)}{\partial\rho},$$

$$\gamma_0 = \frac{1}{m_p}\frac{\partial F^y_\Sigma}{\partial\rho}\bigg|_{\rho_0, y_0} \equiv \frac{1}{m_p}\frac{\partial F^\rho_\Sigma}{\partial y}\bigg|_{\rho_0, y_0}$$

is the shift parameter determining the system reaction to transverse perturbations and quantity ω_c can be regarded as a certain resonance frequency of the system, defined as [24, 28]

$$\omega_c^4 \approx \frac{1}{m_p^2}\left(\frac{\partial F^\rho_\Sigma}{\partial\rho}\frac{\partial F^y_\Sigma}{\partial y}\right)_{\rho_0, y_0}.$$

For particles with a Coulomb pair interaction potential, we have

$$\gamma_0, \omega_c^2 \propto \frac{e^2\langle Z_p\rangle^2 n_p}{m_p}.$$

Since the curl $\boldsymbol{\Omega} = \text{curl}\,\mathbf{V}$ of the macroparticle velocity $\mathbf{V}$ differs from zero, condition (18) describes the emergence of a vortex motion along a certain closed curve. The direction of this rotation in a plane parallel to $\mathbf{F}_{\text{non}}$ for monotonic spatial dependences $\mathbf{E}(\rho, y)$ and $Z(\rho, y)$ can be determined from the sign of Ω. In the field of a constant nonelectrostatic force ($F_{\text{non}}(y) \equiv \text{const}$), a macroparticle moves in the direction of action of this force

in the region where its charge has the maximum value. The amplitude of the vortex motion is determined by various nonlinear effects and/or by the boundary conditions of the problem. The frequency ω of the steady-state rotation can be estimated since an important feature of unstable systems is their tendency to act predominantly through the resonance mode ω_c [37, 38], characterizing the bifurcation point for system (17). Assuming that

$$\omega_c^4 \approx \left|\frac{\gamma_0 \beta_\rho F_{\text{non}}}{m_p}\frac{1}{Z_0}\right|,$$

we find that, in the case of a circular motion for which

$$2\omega \approx |\Omega| = \left|\frac{\beta_\rho F_{\text{non}}}{m_p}\frac{1}{\langle Z_p\rangle \nu_{fr}}\right|,$$

the product of the circular frequency ω and the friction coefficient ν_{fr} for particles is proportional to the square of the resonance frequency of the system:

$$\omega\nu_{fr} \propto \omega_c^2.$$

We assume that the sought nonelectric force $\mathbf{F}_{\text{non}}$ required for the formation of the vortex motion of particles is induced by the directional motion of ions at velocity u relative to a dust particle, i.e., an ion drag force. The latter force is due to momentum transfer from the ions of a dust particle and acts in the direction of the relative motion of ions. The momentum transfer is associated with two processes: inelastic collisions (absorption) of ions with a dust particle and elastic scattering of ions from the particle potential. It was shown in [39] that the ion drag force for a weakly anisotropic plasma ($u \ll v_{Ti}$) in the case of not very large particles prevails for elastic scattering. The ion drag force can be estimated as

$$F_I \approx \frac{2\sqrt{2\pi}}{3} a_p^2 n_i m_i v_{Ti} (z\tau)^2 \Lambda u, \qquad (19)$$

where

$$\tau = T_e/T_i, \quad v_{Ti} = \sqrt{T_i/m_i}$$

and Λ is a modified Coulomb logarithm integrated over the shifted Maxwellian velocity distribution function for ions. For the experimental conditions ($a_p = 1.7$ μm, $n_i \approx 10^9$ cm^{-3}, $z\tau \approx 200$–300, and $\lambda_D/a \approx 20$), the estimation of the modified Coulomb logarithm gives (see [39])

$$\Lambda \sim \frac{4}{z\tau}\frac{\lambda_D}{a}.$$

Thus, for $v_{Ti}/u \approx 4$–8, the ion drag force is $F_I \approx (0.42$–$1.25) \times 10^{-8}$ dyne (or $F_I \approx (0.15$–$0.4) m_p g$ in units of $m_p g$). The comparison of the ion drag force and gravity force acting on particles on the surface of the Earth is convenient from the viewpoint of estimating the significance of this force in land-based experiments.

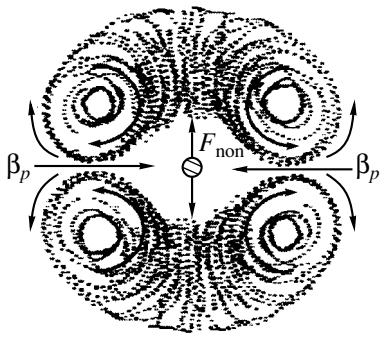


Fig. 5. Illustration of simulation of rotation of macroparticles in the field of orthogonal vectors of nonelectric force F_{non} and charge gradient β_p (cross section along the axis of the cylindrical system).

The value of the regular velocity $u \approx (0.12$–$0.25) v_{Ti}$ assumed for estimates corresponds to directional motion of ions in electric fields $E \approx 1$–2 V/cm. Such fields may be consequences of plasma polarization in discharges controlled by the ambipolar diffusion of plasma particles to the working chamber walls. The velocity of the ambipolar transport of ions can be estimated as

$$u \approx \frac{T_e}{m_i \nu_{ni} \Lambda_d},$$

where ν_{ni} is the frequency of collisions of ions with buffer gas neutrals and Λ_d is the characteristic diffusion length (see Section 2.1) [15]. For argon ($\nu_{ni} \approx 8 \times 10^6$ s^{-1} for $T_e = 1$–2 eV and $\Lambda_d \approx 1$ cm), the value of $u \approx (3.3$–$6.6) \times 10^3$ cm/s $\approx (0.12$–$0.25) v_{Ti}$. Figure 5 illustrates the results of simulation of problem (17) for $N_p = 1000$, $\beta_\rho/e\langle Z_p\rangle = -0.3$ cm^{-1}, $m_p \approx 3 \times 10^{-11}$ g, $\nu_{fr} = 200$ s^{-1}, and $\langle Z_p\rangle = 6000$ for the force $F_{\text{non}} = 0.4 m_p g$ acting from the center of the system along the cylinder axis. The external electric fields ($E(y)$, $E(\rho)$) in directions y and ρ from the center of the system were assumed to be linear fields with equal gradients. The mean kinetic energy of the directional motion (rotation) of particles in this case was about 0.2 eV.

3. EXPERIMENT

3.1. Experimental Setup

The setup intended for experiments on the ISS was similar to that used for studying dusty plasmas under microgravity conditions during parabolic flights and described in detail in [9]. The main element of the

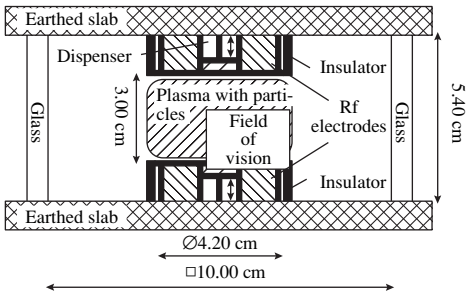


Fig. 6. Vacuum plasma chamber.

experimental setup was a vacuum plasma chamber (Fig. 6) consisting of two square slabs and a square glass insert. On each slab, disk electrodes creating an rf discharge were mounted. Each electrode contained a unit for injecting dust particles into the plasma (dispenser). A high-frequency (13.56 MHz) ac voltage was supplied to the electrodes. The voltage could be controlled in wide limits, thus varying the discharge power W from 0.15 to 0.98 W. The value of the working pressure P of the buffer gas (argon) varied from 0.36 to 0.98 mbar. The experiments were made for monodisperse particles of radius $a_p = 1.7$ μm and density $\rho_p = 1.5$ g/cm^3.

Two CCD chambers with lenses and a diode laser were installed for observing plasma-dust structures in the system. The field of vision of a chamber is shown in Fig. 6 and had a size of 2.8×2.1 cm. The laser beam formed into a plane beam having a thickness of 150 μm and a height of about 25 mm. The control of experimental parameters and the recording of video and digital information obtained in the course of experiments was ensured by a Telescience computer. Video recording was then processed with a help of a special program, which made it possible to identify the positions and to calculate the displacements of individual particles. This program was used for determining the velocity spectra for macroparticles in various regions of the observed dust structures.

The observed dust structures are shown in Figs. 7a and 7b. In all cases, a region free of particles (void) was formed at the center of the dust structure, while quasi-stationary liquid dust structures were observed below and above the void on the axis of the cylindrical discharge flask. The motion of particles in this region was almost "thermal" (the velocity spectrum of particles

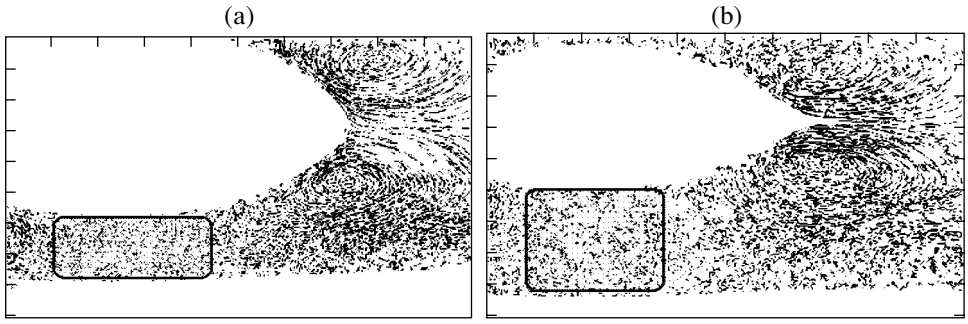


Fig. 7. Illustration of observed dust structures in the field of vision of a video camera (2.8×2.1 cm) for discharge power $W = 0.25$ W under various pressures $P = 0.49$ (a) and 0.98 (b) mbar. The regions of measurement of transport parameters of dust structures are indicated.

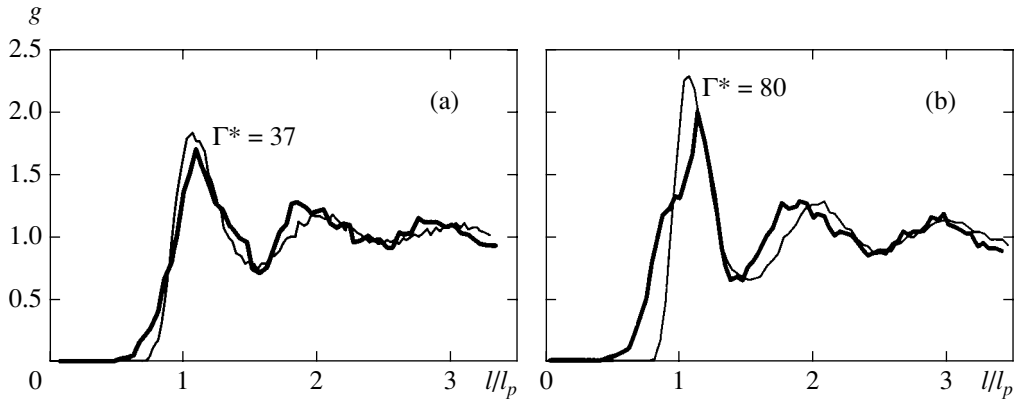


Fig. 8. Measured (bold curve) pair correlation functions $g(l/l_p)$ for $W = 0.25$ W and pressures $P = 0.49$ (a) and 0.98 (b) mbar. Fine curves describe the functions $g(l/l_p)$ obtained from simulation for the values of Γ* indicated in the figure.

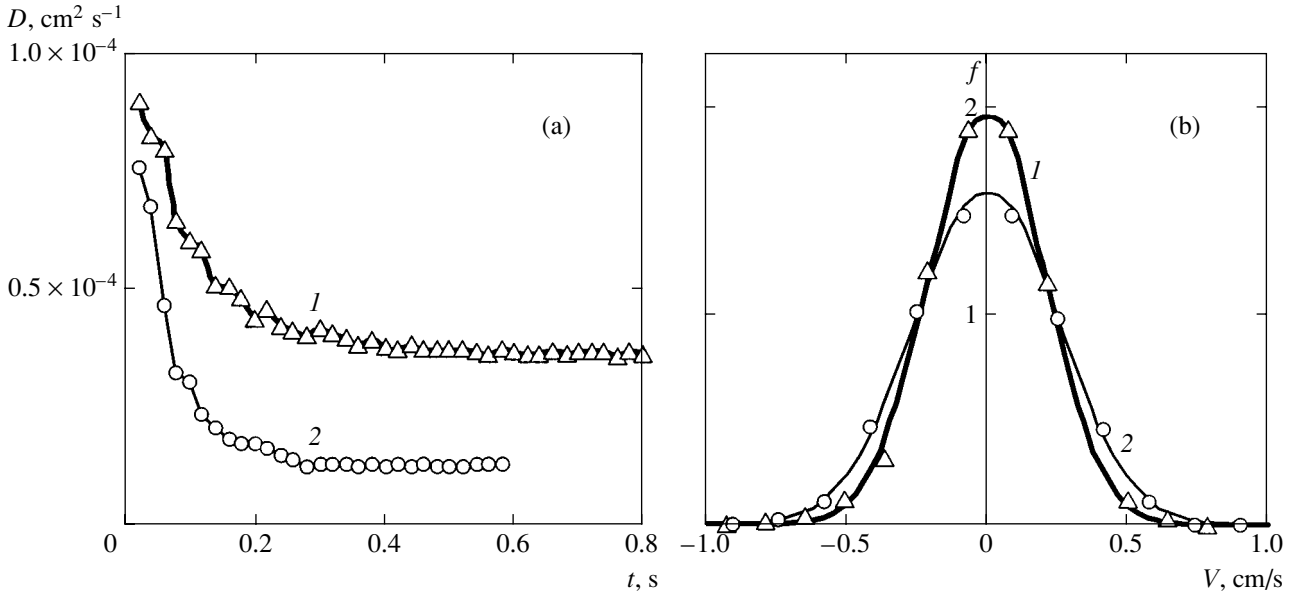


Fig. 9. (a) Time evolution of the diffusion coefficient D and (b) velocity distribution $f(V)$ of dust particles for experiments (W = 0.25 W): P = 0.49 mbar (curve *1*, triangles) and P = 0.98 mbar (curve *2*, circles) (the symbols correspond to experimental results, and the solid curve is the Maxwell distribution function with the temperature of particles given in the table).

was close to the Maxwellian spectrum). In the radial boundary regions of the cloud, symmetric large-scale vortex flows of macroparticles were formed.

3.2. Experimental Analysis of Transport Characteristics of Macroparticles

Analysis of transport characteristics of macroparticles, such as their concentrations, diffusion coefficients, pair correlation functions, and effective charges, was carried out in the regions of dust structures (under the void) without regular motion of macroparticles. The mean particle spacing $l_p = n_p^{-1/3}$ and the concentration n_p of macroparticles were determined from the position of the peak of pair correlation functions and were found to be $l_p \approx 200$ μm and $n_p \approx 1.25 \times 10^5$ cm^{-3} for $P <$ 0.86 mbar and $l_p \approx 220$ μm, $n_p \approx 0.95 \times 10^5$ cm^{-3} for $P \geq$ 0.86 mbar (see Figs. 8a and 8b). To determine the temperature T_p of macroparticles, which characterizes the kinetic energy of their random motion, we measured the velocity spectra of particles. For correcting the results of measurements and for determining the macroparticle temperature from the variance of their velocity distribution, the procedure of the best matching between the results of measurement of macroparticle

Temperature T_p and diffusion coefficient D of macroparticles, maximal g_{max} and minimal g_{min} values of pair correlation function, and values of Γ^*, D/D_0, and ξ for various discharge parameters W and P

P, mbar	W, mW	T_p, eV	$D \times 10^5$, cm^2 s^{-1}	D/D_0	g_{max}/g_{min}	g_{max}	Γ^*	ξ
0.36	25	0.48	7.45	0.34	1.7	1.52	25	0.18
0.49	25	0.40	3.4	0.27	2.4	1.68	40	0.16
0.61	25	0.27	1.0	0.15	3.0	1.97	82	0.15
0.73	25	0.37	1.75	0.22	3.0	1.86	61	0.12
0.86	25	0.4	2.0	0.29	2.4	1.75	42	0.07
0.98	25	0.45	1.25	0.165	3.1	1.97	78	0.12
0.49	15	0.34	2.85	0.265	2.3	1.68	42	0.15
0.49	19	0.40	3.0	0.235	2.7	1.80	52	0.17
0.49	49	0.24	1.85	0.25	2.5	1.78	50	0.13
0.49	98	0.38	2.8	0.24	2.6	1.75	51	0.17

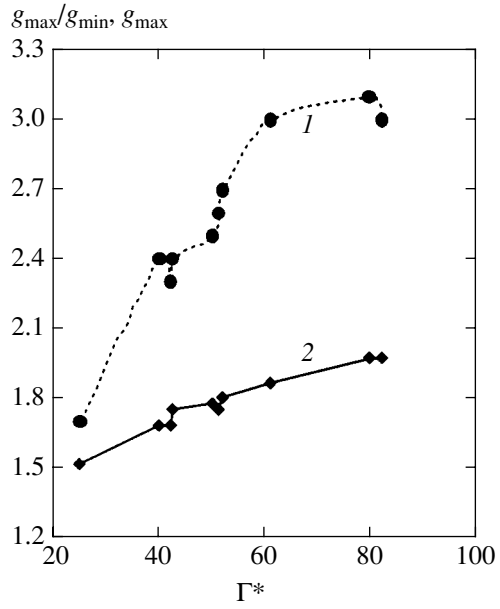


Fig. 10. Dependence of g_{max}/g_{min} (*1*) and g_{max} (*2*) on the reconstructed parameter Γ* (see table).

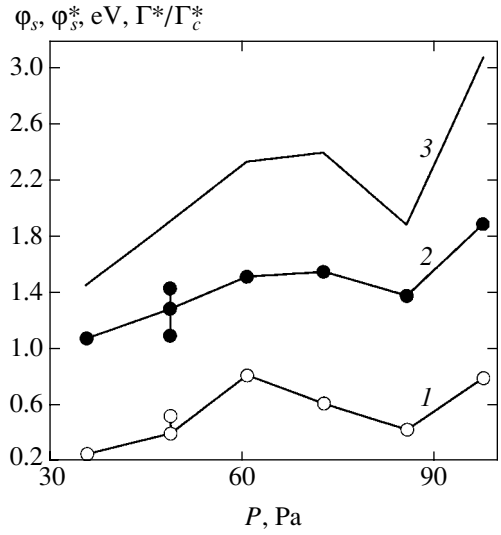


Fig. 11. Values of Γ^*/Γ_c^* reconstructed for different discharge parameters (*1*) and the values of the surface potential φ_c^* (*2*) and φ_s (*3*) obtained from relation (10) for $n_i \approx Z_p n_p$.

diffusion and the numerical data for small observation times was used (see Section 2.3). Figures 9a and 9b illustrate the experimental values of diffusion coefficients and the velocity spectra of macroparticles.

The results of measurements of all working parameters of the discharge are given in the table containing the temperature T_p of macroparticles, their diffusion coefficient D, ratio D/D_0, the peak values of the pair correlation function $g(l)$, and the extent of their correlation (the ratio of the maximal g_{max} and the minimal g_{min} values of $g(l)$ for $l \neq 0$). The same table contains estimates of the nonideality parameter Γ* (12) and the scaling parameter $\xi = \omega^*/\nu_{fr}$ (11) reconstructed from the results of measurements of the diffusion coefficient D of macroparticles (see Section 2.3) using the value of the friction coefficient in the free molecular approximation [24].

The dependences of the ratio g_{max}/g_{min} and g_{max} on the effective parameter Γ* reconstructed from the macroparticle diffusion measurements are given in Fig. 10. Good agreement between the extent of correlation (g_{max}/g_{min}, g_{max}) of particles in the system and the value of the nonideality parameter Γ* obtained is worth noting. Nevertheless, the recorded values of g_{max} for the experimental correlation functions $g(l)$ were slightly lower than the values obtained in simulation (see Fig. 1a), and functions $g(l)$ themselves had a noticeable broadening in the region of small particle spacings. A comparison of the experimental pair correlation functions and the functions $g(l)$ obtained as a result of simulation for close values of Γ* is illustrated in Figs. 8a and 8b. The discrepancy between the experimental and numerical results may be associated either with the procedure of experimentally determining $g(l)$ for a bounded volume containing a finite number of particles or due to spatial inhomogeneity of the dust cloud.

The effective surface potential

$$\varphi_s^* = eZ_p/a_p$$

(see Eq. (10)) and the ratio of parameter Γ* to its value $\Gamma_c^* \approx 102$ at the crystallization point, reconstructed for various discharge parameters, are shown in Fig. 11. For the screened interaction between particles, relation (10) makes it possible to estimate either the macroparticle charge Z_p, or the value of $\kappa = l_p/\lambda$ for a given value of one of the parameters.

Under additional assumptions, we can obtain information on both the macroparticle charge and the screening parameter κ. For example, if the ionization processes in the discharge are controlled by the plasma absorption by dust cloud particles (see Section 2.1) and if the ion concentration $n_i \approx Z_p n_p$, the screening radius can be found from the relation

$$\lambda^2 \approx \frac{T_i}{4\pi e^2 Z_p n_p},$$

where $T_i \approx 0.027$ eV. The surface potential of macroparticles,

$$\varphi_s = eZ_p/a_p,$$

obtained on the basis of this assumption, is shown in Fig. 11, while Fig. 12a gives the values of $\kappa = l_p/\lambda$ and $n_i \approx Z_p n_p$ for this case.

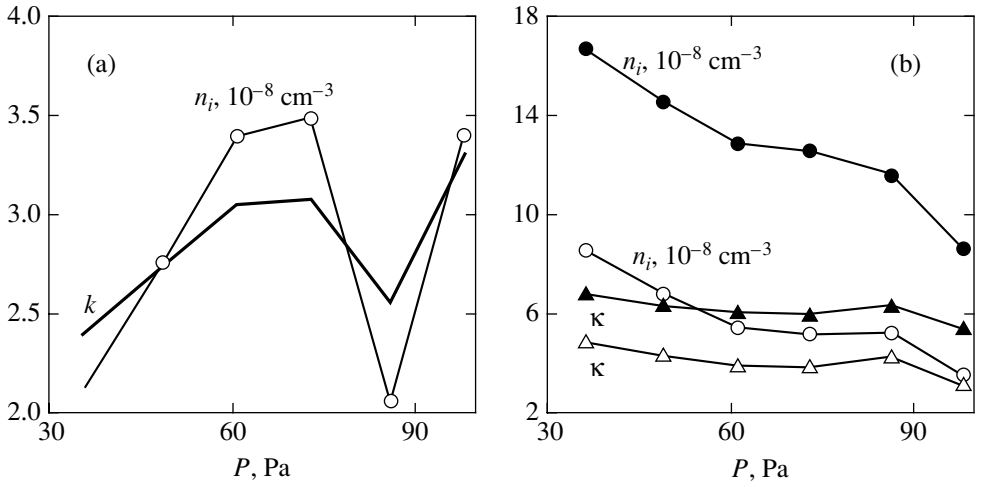


Fig. 12. Reconstructed values of the screening parameter $\kappa = l_p/\lambda$ and ion concentration n_i (a) under the assumption that $n_i \approx Z_p n_p$ and (b) for fixed values of the surface potential $\varphi_s = 2.9$ (light symbols) and 5.8 (dark symbols) eV.

Numerical calculations show that, for

$$\chi = \frac{|\langle Z_p \rangle| n_p}{z n_e} < 1 \quad \left(n_e > \frac{Z_p n_p}{3} \right),$$

the surface potential φ_s of macroparticles can be estimated using the relation $\varphi_s \approx 2.9 T_e/e$ which describes the charging of a solitary particle (see Section 2.1). Figure 12b shows the value of the screening parameter $\kappa = l_p/\lambda$ and the values of ion concentration $n_i \approx T_i/4\pi e^2 \lambda^2$ corresponding to the reconstructed parameter κ for fixed values of the surface potential $\varphi_s = 2.9$ and 5.8 eV ($T_e \approx 1$–2 eV, $Z_p \approx 3400$–6800). Thus, the obtained ion concentration $n_i > Z_p n_p$ ($\chi < 1$, $n_p \approx 10^5$ cm^{-3}) lies in the range from approximately 5×10^8 cm^{-3} to 2×10^9 cm^{-3}; this is in accordance both with the results of measurements and with the numerical data on the plasma concentration in an rf discharge with parameters close to experimental values [15, 41], which give $n_i \approx 10^9$ cm^{-3}. Estimation of macroparticle charges for a fixed value of $\lambda = \lambda_i \approx 40$ μm ($n_i \approx 10^9$ cm^{-3}) gives $Z_p \approx 3600$–5200 for all pressures below $P = 0.98$ mbar and $Z_p \approx 7700$ for $P = 0.98$ mbar. This, in turn, makes it possible to estimate the range of electron temperatures at $T_e \approx 1$–2.5 eV.

3.3. Experimental Analysis of Vortex Motion of Particles

For an analysis of the vortex motion of particles (see Fig. 7), we singled out (in the working range) the field of velocities directed along the normal to the radius of rotation of particles. This field was divided into individual "annular" segments of width $\Delta R = R_{max}/20$, where R_{max} is the maximal radius of identified rotation of macroparticles, which was determined from the area S involved in rotation ($R_{max} = \sqrt{S/\pi}$). Figure 13 illustrates such a division. Then we determined the mean time taken by a macroparticle to traverse a small segment of the arc of the "ring," which was then summed for calculating the mean period T and the rotational frequency $\omega = 2\pi/T$.

The results of measurements of the radial distribution of the rotational frequency ω in vortices for different discharge modes are given in Figs. 14–16. The rotational frequency of macroparticles decreased with increasing distance R from the vortex center (see Fig. 14). For an analysis of this decay at the initial segments of vortices (for rotational radii $R < 2$ mm), we approximated the experimental data with the exponential function $\omega = \omega_0 \exp(-\sigma R)$, where $\omega_0 \equiv \omega(R = 0)$, and coefficient σ can be regarded as the radial diffusion probability for macroparticles. This approximation is illustrated in Fig. 14. The dependences of coefficient σ on the maximal rotational frequency ω_0 on the working

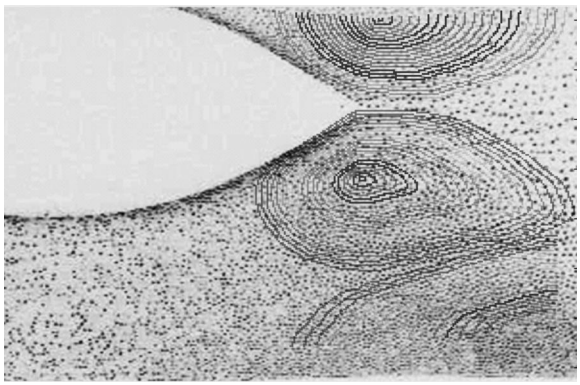


Fig. 13. Illustration of the procedure of processing of video recordings of the vortex motion of macroparticles.

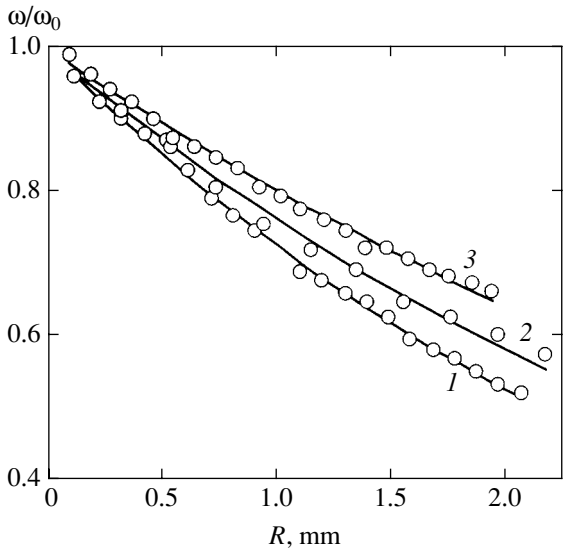


Fig. 14. Experimental (circles) dependence $\omega(R)/\omega_0$ and the approximation of this dependence by exponential functions (solid curves) for $W = 0.25$ W and for different values of P = 0.36 (*1*), 0.49 (*2*), and 0.98 (*3*) mbar.

parameters of the discharge (pressure P and power W) are shown in Figs. 15a and 15b.

Let us consider the probability of vortex formation due to the existence of orthogonal charge gradient $\boldsymbol{\beta}$ and a certain nonelectrostatic force F_{non}, e.g., the ion drag force $F_I \approx 0.3 m_p g$, on the basis of the model described in Section 2.4. Assuming that

$$\omega_0 \approx \frac{\Omega}{2} = \frac{\beta F_{\text{non}}}{m_p} \frac{1}{2\langle Z_p \rangle \nu_{fr}},$$

we find that inhomogeneous conditions ensuring the relative charge gradients $\beta/e\langle Z_p \rangle$ in the range from 0.07 to 0.2 cm^{-1} is sufficient for evolution of rotation with detectable frequencies $\omega_0 \approx 0.1$–0.16 s^{-1} for $\nu_{fr} \approx 120$–330 s^{-1}. It was noted above (see Section 2.4) that such values of $\beta/e\langle Z_p \rangle$ are quite attainable in an ambipolar-diffusion-controlled discharge and may appear due to slight violation of electroneutrality of the surrounding plasma.

Concluding the section, let us briefly analyze the detected radial distributions $\omega(R)$ of angular velocities of rotation of macroparticles in dust vortices, proceeding from the exponential dependence

$$\omega(R) \propto \omega_0 \exp(-\sigma R),$$

obtained in the course of experiments. It should be noted that the theory of vortices in strongly nonideal systems, in which the mean free path of particles is comparable with or smaller than the particle spacing, has not been developed as yet [41]. Most of the proposed models pertain to the case when collisions at distances close to the Onsager radius, for which the potential energy of particle interaction is comparable with the thermal energy of macroparticles, play a decisive role in the system. It should also be noted that simple analytic models for describing rotations in dissipative systems were developed only for plane potential flows, for which the circulation $\oint_C \mathbf{V} d\mathbf{s}$ of the velocity vector $\mathbf{V}$ (where $\mathbf{s}$ is a length element of contour C) is equal to zero or is independent of the choice of contour [41]. In the latter case, we are dealing with a potential flow with circulation. Thus, the experimental investigations of rotation of dust particles are of considerable interest for the development of appropriate theoretical models for describing vortices in strongly nonideal dissipative systems.

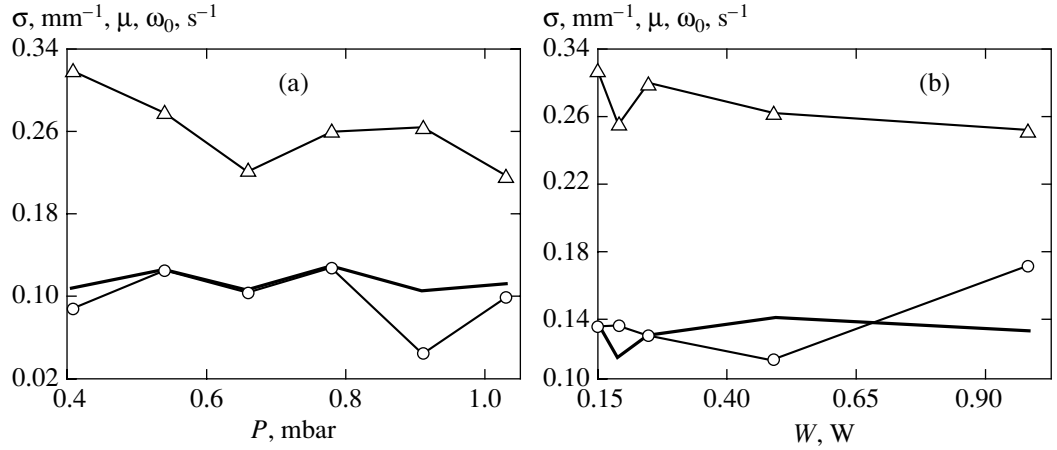


Fig. 15. Dependences of coefficient σ (triangles), maximum rotational frequency $\omega_0 = \omega(0)$ (circles), and parameter $\mu = (\omega_0 \nu_{fr})^{1/2}/\omega^*$ (bold curve) for macroparticles in a vortex on pressure P for $W = 0.25$ W (a) and on the discharge power W for P = 0.49 mbar (b).

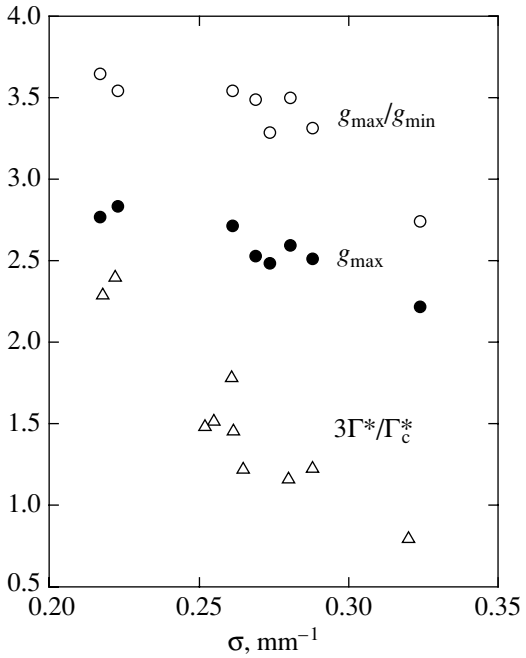


Fig. 16. Dependences of particle correlation g_{max}/g_{min} (○), g_{max} (●), and Γ^*/Γ_c^* (△) on coefficient σ for various discharge parameters.

For an analysis of the dependences of coefficients ω_0 and σ for the proposed exponential approximation

$$\omega(R) = \omega_0 \exp(-\sigma R),$$

information on the parameters of the medium in the zone of dust rotation is required. However, it is practically impossible to determine correctly the pair correlation function or the diffusion coefficients of macroparticles, which would ensure a qualitative estimate of the extent of correlation between particles in the vortex region, because of considerable regular motion of particles as well as due to strong inhomogeneity of the region in question. Nevertheless, we can assume that the effective parameters Γ* and ω for macroparticles have either close values in the entire volume of the dust cloud or change in proportion to the discharge parameters *P* and *W*. In this case, a simple approximation of the maximal rotational frequency ω_0 can be obtained on the basis of the theory of evolution of instabilities (see Section 2.4) if we assume that the characteristic frequency ω_c of the dust is proportional to the effective frequency ω* (9). In this case, the relation $\mu^2 = \omega\nu_{fr}/\omega^{*2}$ must be preserved (or be nearly the same) for all discharge parameters since $\omega_c^2 \propto \omega\nu_{fr}$. This assumption is confirmed by the results of measurements of the parameter

$$\mu = (\omega_0 \nu_{fr})^{1/2}/\omega^* \equiv (\omega_0/\omega^*\xi)^{1/2}$$

(see Fig. 15), where ω* is the effective frequency (9) in the region under the void. It can easily be seen that $\mu \approx 0.12 \pm 0.01$ in all experiments.

Since there is no radial diffusion of macroparticles in rotating solids (for a planar circular motion, $\omega(R) = \omega_0$, $\sigma = 0$), we can assume that the value of coefficient σ will decrease upon an increase in the effective nonideality parameter Γ*. The relation between the particle correlation in the space under the void and the rate of decay of dust rotation is shown in Fig. 16. It can easily be seen that the rate of decay of vortex motion decreases upon an increase in the particle interaction. Thus, parameters ξ and Γ* describing microscopic processes in nonideal dissipative systems play a decisive role in the dynamics of the dust vortices observed.

4. CONCLUSIONS

The results of investigation of transport processes such as diffusion of macroparticles and formation of dust vortices in the strongly nonideal dusty plasma of an rf discharge under microgravity conditions are considered. Experimental results were obtained for monodisperse particle of radius $a_p = 1.7$ μm in a wide range of plasma parameters. Pair correlation functions, velocity spectra, and diffusion coefficients of macroparticles were measured. On the basis of the results of measurements, the concentrations and temperature (kinetic energy of random motion of dust structures) of macroparticles were obtained for the regions of dust structures in which there was no regular motion of dust; angular velocity distributions of macroparticles in the region of their vortex motions were also determined. The effective parameters φ_s^* and Γ* of dust structures were reconstructed for various discharge parameters. It was found that the degree of correlation between particles in the system, which is determined from the shape of the pair correlation function, is in good agreement with the value of effective nonideality parameter Γ* reconstructed from the measurements of particle diffusion. The observed dust structures were strongly correlated dust liquids (Γ* ≈ 30–80).

Under the assumption of screened interaction between particles, macroparticle charges Z_p and screening parameters $\kappa = l_p/\lambda$ were estimated. Analysis of these data shows that the screening by ions plays a significant role for dust structures under the given experimental conditions. For the plasma parameters ($T_e \approx 1$–3 eV, $n_i \approx 10^9$ cm^{-3}) typical of the given conditions, the charge Z_p of dust particles can be in the range from 3500 to 7000*e*.

The dynamics of dust vortices was analyzed experimentally. An empirical approximation was obtained for the radial distribution of angular velocities of macroparticles. It was shown that, in the presence of a considerable ion drag force, a slight variation of the charge of macroparticles is sufficient for the formation of their vortex motion.

The results of experiments discussed here show that both the microscopic motion of macroparticles and the large-scale rotation of dust in the plasma under investigation can be described using two basic parameters, namely, the effective nonideality parameter Γ^* and the scaling parameter ξ.

The experiment Plasma Crystal-3 was realized owing to the support and participation of the Russian Aviation and Space Agency, Energiya Rocket-Building Corporation, the German Space Agency, the Russian Ministry of Industry, Science and Technology, the German firm Keiser Trede, the Gagarin Center for Cosmonaut Training, and the Space Flight Control Center.

ACKNOWLEDGMENTS

The authors are grateful to a large group of specialists participating in preparation of the experiment: M. Belyaev, H. Binnenbruk, V. Blagov, L. Deputatova, J. Goree, S. Goryainov, Yu. Grigor'ev, V. Griete, T. Hagl, A. Kelling, R. Klett, K. Kerner, U. Konopka, M. Kudashkina, R. Kuhl, S. Kuznetsov, A. Lebedev, V. Nikitskii, H. Pfeuffer, M. Pronin, I. Roslavtseva, T. Rostopirov, M. Root, P. Saburov, A. Shcherbak, A. Shurov, and M. Tsutsik.

This study was partly financed by the Russian Foundation for Basic Research (project no. 01-02-16658) and INTAS (project nos. 01-0391 and 00-0522).

REFERENCES

1. H. Thomas, G. E. Morfill, V. Demmel, *et al.*, Phys. Rev. Lett. **73**, 652 (1994).
2. A. Melzer, T. Trottenberg, and A. Piel, Phys. Lett. A **191**, 301 (1994).
3. V. V. Zhakhovskiĭ, V. I. Molotkov, A. P. Nefedov, *et al.*, Pis'ma Zh. Éksp. Teor. Fiz. **66**, 392 (1997) [JETP Lett. **66**, 419 (1997)].
4. S. Nunomura, T. Misawa, N. Ohno, and S. Takamura, Phys. Rev. Lett. **83**, 1970 (1999).
5. O. Vaulina, S. Khrapak, A. Nefedov, *et al.*, Phys. Rev. E **60**, 5959 (1999).
6. V. E. Fortov, A. P. Nefedov, O. S. Vaulina, *et al.*, Zh. Éksp. Teor. Fiz. **114**, 2004 (1998) [JETP **87**, 1087 (1998)].
7. O. S. Vaulina, A. P. Nefedov, O. F. Petrov, *et al.*, Phys. Rev. Lett. **88**, 035001 (2002).
8. V. E. Fortov, A. P. Nefedov, V. I. Molotkov, *et al.*, Phys. Rev. Lett. **87**, 205002 (2001).
9. G. Morfill, H. Thomas, U. Konopka, *et al.*, Phys. Rev. Lett. **83**, 1598 (1999).
10. T. Stuffler, G. Schmitt, H. Pfeuffer, *et al.*, in *Proceedings of the 52nd International Astronautical Congress, Toulouse, France* (2001).
11. A. P. Nefedov, O. S. Vaulina, O. F. Petrov, *et al.*, Zh. Éksp. Teor. Fiz. **122**, 778 (2002) [JETP **95**, 673 (2002)].
12. T. Nitter, Plasma Sources Sci. Technol. **5**, 93 (1996).
13. J. Goree, Plasma Sources Sci. Technol. **3**, 400 (1994).
14. O. S. Vaulina, S. A. Khrapak, A. A. Samarian, and O. F. Petrov, Phys. Scr. **84**, 229 (2000).
15. Yu. P. Raĭzer, *Physics of the Gas Discharge* (Nauka, Moscow, 1987).
16. U. Konopka, G. E. Morfill, and L. Ratke, Phys. Rev. Lett. **84**, 891 (2000).
17. V. N. Tsytovich, Usp. Fiz. Nauk **167**, 57 (1997) [Phys. Usp. **40**, 53 (1997)].
18. S. A. Khrapak, A. V. Ivlev, and G. Morfill, Phys. Rev. E **64**, 046403 (2001).
19. D. P. Resendes, J. T. Mendonca, and P. K. Shukla, Phys. Lett. A **239**, 181 (1998).
20. A. S. Ivanov, Phys. Lett. A **290**, 304 (2001).
21. J. E. Daugherty, R. K. Porteous, M. D. Kilgore, *et al.*, J. Appl. Phys. **72**, 3934 (1992).
22. J. E. Allen, Phys. Scr. **45**, 497 (1992).
23. O. S. Vaulina, A. P. Nefedov, O. F. Petrov, and V. E. Fortov, Zh. Éksp. Teor. Fiz. **118**, 1325 (2000) [JETP **91**, 1147 (2000)].
24. E. M. Lifshitz and L. P. Pitaevskiĭ, *Physical Kinetics* (Nauka, Moscow, 1979; Pergamon Press, Oxford, 1981).
25. A. A. Ovchinnikov, S. F. Timashev, and A. A. Belyĭ, *Kinetics of Diffusion-Controlled Chemical Processes* (Khimiya, Moscow, 1986).
26. O. S. Vaulina and S. A. Khrapak, Zh. Éksp. Teor. Fiz. **119**, 264 (2001) [JETP **92**, 228 (2001)].
27. O. S. Vaulina and S. V. Vladimirov, Phys. Plasmas **9**, 835 (2002).
28. O. S. Vaulina, A. P. Nefedov, O. F. Petrov, *et al.*, Zh. Éksp. Teor. Fiz. **120**, 1369 (2001) [JETP **93**, 1184 (2001)].
29. O. S. Vaulina, S. V. Vladimirov, O. F. Petrov, *et al.*, Phys. Rev. Lett. **88**, 245002 (2002).
30. O. S. Vaulina, Zh. Éksp. Teor. Fiz. **121**, 35 (2002) [JETP **94**, 26 (2002)].
31. S. Hamaguchi, R. T. Farouki, and D. H. E. Dubin, Phys. Rev. E **56**, 4671 (1997).
32. M. J. Stevens and M. O. Robbins, J. Chem. Phys. **98**, 2319 (1993).
33. H. Ohta and S. Hamaguchi, Phys. Plasmas **7**, 4506 (2000).
34. A. Melzer, A. Homann, and A. Piel, Phys. Rev. E **53**, 2757 (1996).
35. Yu. V. Gerasimov, A. P. Nefedov, V. A. Sinel'shchikov, and V. E. Fortov, Pis'ma Zh. Tekh. Fiz. **24** (19), 62 (1998) [Tech. Phys. Lett. **24**, 774 (1998)].
36. V. E. Fortov, V. I. Molotkov, A. P. Nefedov, *et al.*, Phys. Plasmas **6**, 1759 (1999).
37. R. K. Dodd, J. C. Eilbeck, J. Gibbon, and H. C. Morris, *Solitons and Nonlinear Wave Equations* (Academic, New York, 1982; Mir, Moscow, 1988).
38. T. Akhromeeva, S. Kurdyumov, and G. Malinetskiĭ, *Computers and Nonlinear Phenomena* (Nauka, Moscow, 1988).
39. S. A. Khrapak, A. V. Ivlev, G. Morfill, and H. Thomas, Phys. Rev. E **66**, 046414 (2002).
40. M. R. Akdim and W. J. Goedheer, Phys. Rev. E **65**, 015401 (2002).
41. V. V. Kozlov, *General Theory of Vortices* (Udmurtskiĭ Univ., Izhevsk, 1998).

Translated by N. Wadhwa

Journal of Experimental and Theoretical Physics, Vol. 96, No. 4, 2003, pp. 719–729.
Translated from Zhurnal Éksperimental'noĭ i Teoreticheskoĭ Fiziki, Vol. 123, No. 4, 2003, pp. 816–827.

FLUIDS

Longwave Marangoni Convection in an Inhomogeneously Heated Liquid

V. P. Shilov

Institute of Mechanics of Continua, Urals Division, Russian Academy of Sciences, Perm, 614013 Russia
e-mail: v_p_shilov@icmm.ru

Received September 10, 2002

Abstract—The supercritical Marangoni convection has been studied in a plane-parallel liquid layer, bounded by a free deformable gas–liquid interface from above and by a low-heat-conductivity wall from below, occurring under conditions of inhomogeneous heating in the horizontal plane. In a longwave approximation with a small inhomogeneity of heat flux, the process is described by a system of two-dimensional nonlinear equations for the temperature perturbations, vorticity, and free surface deformation. The concept of quasiequilibrium, implying stability of long-range flows, is introduced, which allows the inhomogeneous heat flux to be modeled by a step function. The linear stability is analyzed in the cases of planar and axisymmetric heat fluxes. The boundaries of stability of the convection regimes are determined on the plane of parameters characterizing the degree of supercriticity inside a heated spot and the depth of damping outside the spot. For an axisymmetric spot, the domains of stability with respect to perturbations for various azimuthal numbers are established. © *2003 MAIK "Nauka/Interperiodica".*

1. INTRODUCTION

In the 19th century, the Italian winemaker Marangoni paid attention to a special kind of flows in wine that were caused by inhomogeneous surface tension. Representing the so-called thermocapillary or capillary-concentration convection, this phenomenon was later termed after the name of its discoverer. The wish to decrease the role of capillary flows in technological processes (such as obtaining ultrahigh-purity materials on board space vehicles, laser doping in metallurgy, etc.), as well as a basic interest in the new mechanisms of instability development and the formation of spatiotemporal structures, have inspired extensive investigation in this direction.

An original study into the thermocapillary instability of a planar layer of liquid heated from below was undertaken by Pearson [1]. Subsequent investigations into this phenomenon were summarized in monograph [2]. Despite a two-century history of research, the Marangoni flows experimentally observed under conditions of inhomogeneous heating or inhomogeneous surfactant application are still insufficiently understood and have not been given adequate explanation.

Experimental investigations of the Marangoni convection were initiated in 1974 with a pioneering work by Pshenichnikov and Yatsenko (reproduced in [3]) devoted to the motion of a liquid in the presence of a point source of surfactant on the free surface. The experiments were performed with distilled water as the liquid medium and ethyl alcohol as the surfactant. A 10% drop in the ethanol concentration in the surface layer changed the surface tension of water approximately by 20 dyn/cm. The experiments showed that, instead of a radially symmetric flow distribution, a multilobe pattern of motion takes place with the number of lobes depending on the integral mass flow rate of the surfactant (ethanol). At a constant supply of the surfactant, the flow is steady and stable, with a sharp peak in the radial velocity component and one vortex on each side. As the surfactant supply increases, the intensity of motion grows and, at a certain value of the mass flow rate, the two-vortex motion ceases to be stationary: the third and then the forth vortex arise from time to time on the side opposite to the radial velocity maximum. In this nonstationary regime, the two- and four-vortex stages alternate unless, with further growth in the surfactant supply, another steady state with four vortices at the source is established in the system. At still higher values of the surfactant mass flow rate, a stable six-vortex motion takes place, followed by an eight-vortex stage and so on. Each transition from a stage with $2m$ vortices (where m is an integer) to that with $2(m + 1)$ vortices corresponds to a certain value of the surfactant flux. The mass flow rate interval in which a stationary $2m$-vortex pattern of motion exists rapidly decreases with increasing m, so that steady states featuring no more than ten vortices could be experimentally observed. Note that, in the entire range of the flow parameters studied, the observed flows exhibited a long-range character.

Ezersky *et al.* [4] studied the structure of a thermocapillary flow from a localized heat source situated at the bottom of a plane-parallel liquid layer in a round cell. The experiments were performed with 2- to 8-mm-thick layers of silicon oil. In contrast to the system stud-

ied in [3], thin (3-mm-thick and below) layers at small temperature gradients always exhibited radially symmetric flows directed from the center of a heated spot toward cold walls of the cell. An increase in the temperature difference between the spot and periphery led to the formation of steady-state concentric billows in the entire cell space. Further increase in the source temperature gave rise to numerous defects over the heated spot and to running waves outside, superimposed on the main flow. At a large liquid layer thickness (above 3 mm), no steady-state pattern of billows was observed. With increasing heat source intensity, the pattern of radially symmetric flow converted into that of running waves. It was suggested [4] that the observed instability has a shortwave character related to the combined action of the thermocapillary and gravitational mechanisms of convection.

Flow patterns analogous to those observed in [4] were also reported in [5], where a thermocapillary flow controlled by a lumped heat source on the surface of a thin layer of silicon oil (with a thickness to lateral size ratio not exceeding 0.1). This system also exhibited a radially symmetric flow toward the cell walls. Even at relatively small temperature differences, this flow lost stability. The secondary structure represented concentric billows arising at the boundaries of opposite flows: the main outgoing flow and the counterflow directed toward the heat source. Further increase in the source temperature led to the formation of radial billows, so that the instability acquired a three-dimensional character.

Recently, Bratukhin *et al.* [6] experimentally and theoretically studied the Marangoni convection in a liquid (decane) filling a deep pool with a flat circular (5–7 mm in diameter) heat source placed at various depths. With the source situated at a large depth, the system exhibited predominantly a thermogravitational flow with a thermal torch. The torch spreading from the source toward the gas–liquid interface distorted the free liquid surface, making it convex toward the gas phase. As the source depth decreased, the capillary effects became more pronounced, leading to the formation of a funnel at the gas–liquid interface. At a small depth of the heat source (1 mm and below), the liquid surface featured spiral waves turning either clockwise or anticlockwise. The direction was established randomly, right- and left-hand twist being equiprobable. As the source approached the surface, the number of spiral arms increased from one to ten. When the heat source escaped on the surface, the running waves decayed to leave a beamlike standing-wave structure analogous to that observed in [5]. It was pointed out that the convection exhibited a long-range character [6].

This study aimed at obtaining and solving a system of model equations reflecting the main features of experiments described above. All systems studied previously had in common the presence of a source of inhomogeneous heating or surface tension (surfactant), which gave rise to a region of inhomogeneous temperature or concentration distribution at the source boundary. Experimental data reported in [3–6] differ in the character of convective motions. In particular, the experiments described in [4, 5] require taking into account the three-dimensional character of flows, which is a rather complicated task. For this reason, the theoretical model constructed below is restricted to the asymptotic longwave expansion, which allows the problem of long-range convection [3, 6] to be reduced to a system of two-dimensional equations. Originally suggested by Sorokin [7], this approach to description of a finite-amplitude convection is valid in the region of small supercriticity. Describing the convection in a plane- parallel layer of liquid featuring a continuous spectrum of motions, this approach provides an answer to the main question as to which motion is favorable.

A system of nonlinear equations of supercritical thermogravitational convection, describing the slow long-range dynamics of the average temperature field, was obtained by Shtilman and Sivashinsky [8]. In application to the Marangoni convection, this problem was solved by Knobloch [9] in a system with nondeformable surface and infinitely large Prandtl numbers. Finite Prandtl numbers lead to the predominant influence of inertial effects in the formation of a nonpotential averaged flow. Such effects were previously studied in the case of gravitational convection [10] and were then taken into account for capillary convection as well [8]. A generalization of the Knobloch equations to the case of a two-layer system with deformable surface was formulated by Golovin *et al.* [11] together with dimensionless physical criteria of a longwave instability development in the system.

A characteristic feature of the aforementioned theoretical investigations is the assumption concerning homogeneous heating of the liquid layer. Inhomogeneity of the heat flux was taken into account only in the case of thermogravitational convection [12]. The problem of interaction between inhomogeneous heating (or inhomogeneous surfactant flow) and long-range supercritical Marangoni convection remains unsolved. In this study, the problem will be solved using an asymptotic longwave expansion [12, 13].

In Section 2, the problem is formulated, the dimensionless quantities are introduced, and a system of two-dimensional model equations is derived that describes slow, long-range thermocapillary Marangoni convection in an inhomogeneously heated plane-parallel liquid layer with allowance for surface deformation. Physical limitations for proposed model are indicated and the possibility of application to the concentration Marangoni convection is considered.

Section 3 introduces the concept of mechanical quasiequilibrium for a long-range loss of stability, which allows the inhomogeneous heat flux to be modeled by a step function. It will be demonstrated that all normal perturbations vary monotonically (increase or decrease) with time. Section 3.1 is devoted to the linear

stability of this quasiequilibrium for a one-dimensional inhomogeneity of the heat flux. Neutral curves for the symmetric and antisymmetric modes are constructed with allowance for the surface deformation. Section 3.2 considers flow stability in the case of an axisymmetric inhomogeneity of the heat flux and presents the neutral curves for various values of the azimuthal "wavenumber." It will be shown that, under certain conditions, a dipole mode (rather than axisymmetric) is most dangerous from the standpoint of instability development. In the Conclusion, the results are discussed and compared to experiment [3, 6]. The Appendix presents detailed derivation of the system of model equations.

2. FORMULATION OF THE PROBLEM

Consider a convective Marangoni flow in a plane-parallel liquid layer of thickness L, bounded with a free deformable surface from above and by a low-heat-conductivity wall from below. The upper halfspace is filled by a gas. The liquid and gas are heated from below, conducting an external inhomogeneous steady-state heat flux. Therefore, the liquid is featuring a flow induced by this inhomogeneous heating. Let us introduce a coordinate system with the z axis pointing upward and the x and y axes lying in the bottom plane. The character and stability of convection in this system will be studied under the following conditions: (i) the dynamic viscosity and heat conductivity of the gas being lower than those of the liquid, the gas phase characteristics do not cause the flow instability; (ii) the heat flux inhomogeneity is long-range as compared to the liquid layer thickness and is small in the sense that a local heat flux differs only slightly from the average flux Q. All quantities are rendered dimensionless, being expressed in units of length (L), time (L^2/χ), velocity (χ/L), pressure ($\eta\chi/L^2$), and temperature LQ/κ, where κ, χ, and η are the coefficients of heat conductivity, thermal diffusivity, and dynamic viscosity of the liquid. For the sake of convenience, with a view to applying a longwave approximation, let us introduce the horizontal ($\mathbf{u}$) and vertical (w) velocity components; differentiation with respect to the vertical coordinate (∂_z) will be denoted by the prime, and that with respect to the horizontal coordinates will be denoted by operator $\nabla = \mathbf{i}\partial_x + \mathbf{j}\partial_y$, where $\mathbf{i}$ and $\mathbf{j}$ are the unit vectors along the x and y axes. In these terms, the system of dimensionless equations of hydrodynamics [14], including the Navier–Stokes equation and the heat conductivity equation, takes the following form:

$$\begin{aligned}
\frac{1}{\mathrm{Pr}}(\partial_t\mathbf{u} + \mathbf{u}\cdot\nabla\mathbf{u} + w\mathbf{u}') &= -\nabla P + \nabla^2\mathbf{u} + \mathbf{u}'',\\
\frac{1}{\mathrm{Pr}}(\partial_t w + \mathbf{u}\cdot\nabla w + ww') &= -P' + \nabla^2 w + w'' - \mathrm{G},\\
\partial_t T + \mathbf{u}\cdot\nabla w + wT' &= \nabla^2 T + T'',
\end{aligned} \tag{1}$$

where $\mathrm{G} = gL^3/\nu\chi$ and $\mathrm{Pr} = \nu/\chi$ are the Galilean and Prandtl numbers, respectively; ν is the kinematic viscosity coefficient; and T is the temperature relative to the value averaged over the liquid layer. As for the boundary conditions, the bottom surface ($z = 0$) is characterized by sticking (for the velocity) and a preset heat flux. The upper interface, $z = 1 + h(x, y)$, is characterized by the kinematic condition and the heat flux $\mathbf{Q} = \mathbf{e}_z(-1 + q(x, y))$ along the vertical coordinate axis, the projection of which onto the normal equals the normal component of the heat flux per unit surface area:

$$\mathbf{n}\cdot\nabla T = \mathbf{n}\cdot\mathbf{e}_z[-1 + q(x, y)].$$

In addition we require the continuity of stresses at $z = 0$,

$$\mathbf{u} = 0, \quad w = 0, \quad T' = -1 + q(x, y), \tag{2a}$$

and at $z = 1 + h(x, y)$:

$$\begin{aligned}
\partial_t h + \mathbf{u}\cdot\nabla h &= w,\\
T' &= \nabla h\cdot\nabla T - 1 + q(x, y),\\
\hat{\sigma} : \mathbf{n}\cdot\mathbf{n} &= (-\mathrm{Ca} + \mathrm{Ma}T)\mathrm{div}\mathbf{n},\\
\hat{\sigma} : \mathbf{n}\cdot\boldsymbol{\tau}_\alpha &= -\mathrm{Ma}\boldsymbol{\tau}_\alpha\nabla T,
\end{aligned} \tag{2b}$$

where $\mathrm{Ca} = \sigma_s L/\eta\chi$ is the capillary number, $\mathrm{Ma} = (\partial\sigma_s/\partial T)QL^2/\eta\chi\kappa$ is the Marangoni number, σ_s is the surface tension; $\hat{\sigma}$ is the tensor of viscous stresses, $h(x, y)$ is the amplitude of the free surface deviation from the $x = 1$ plane, $q(x, y)$ is the inhomogeneous component of the heat flux, $\mathbf{n} = (-\nabla h; 1)/\sqrt{1 + (\nabla h)^2}$ is the unit vector of the normal to the free surface, $\boldsymbol{\tau}_\alpha = (\delta_{\alpha x}, \delta_{\alpha y}, \nabla h)/\sqrt{1 + (\nabla_\alpha h)^2}$ are the tangential unit vectors, and the colon indicates the twofold convolution of the viscous stress tensor.

The role of the surface tension and the gravity force in the free surface deformation is characterized by the capillary number and the Galilean number, respectively. In the problem under consideration, these parameters are sufficiently large (for a 1-mm-thick water layer at 20°C and atmospheric pressure, $\mathrm{G} = 6.8\times10^4$ and $\mathrm{Ca} = 5.1\times10^5$); for this reason, the free surface deformation is very small and does not affect the linear stability. In what follows, we will assume that a heat flux is close to critical throughout the liquid layer, differing from this level to within the second-order term in $q(x, y)$, a slowly varying function of lateral coordinates. A small parameter is defined as $\varepsilon^2 = \mathrm{Ma}/\mathrm{Ma}^* - 1$, where Ma^* is the critical Marangoni number defined below. For a longwave character of the flow, the derivatives of any quantities in the lateral directions are much smaller than those in the vertical direction. This implies that there is another small parameter in the problem, which is generally not related to ε. However, we can obtain a closed system of equations by imposing the following limitation: parameters of different physical natures are considered as quantities of the same

order of smallness, so that $\varepsilon \sim |\partial f/\partial x|/|\partial f/\partial z| \sim L/R$, where R is a characteristic scale of the flow or the heat flux inhomogeneity $q(x, y)$ in the lateral directions. For convenience, let us calibrate the capillary number and Galilean number as

$$G = \frac{\tilde{G}}{\varepsilon^2}, \quad Ca = \frac{\tilde{Ca}}{\varepsilon^4}, \quad \tilde{G} \sim 1, \quad \tilde{Ca} \sim 1$$

(below, the tilde is omitted). Consider the case of low-intensity motions in which the velocity $\mathbf{u}$ is of the first order of smallness in ε (a square root of supercriticity). Under these conditions, the vertical component of velocity is of the second order of smallness, while the evolution operator is on the order of ε^4. Let us perform rescaling of the horizontal coordinate, the time, and the heat flux inhomogeneity by substituting

$$x \longrightarrow \varepsilon x, \quad y \longrightarrow \varepsilon y, \quad t \longrightarrow \varepsilon^4 y, \quad q \longrightarrow \varepsilon^2 q.$$

A remarkable feature of system (1) upon this rescaling is that the unit of length for the lateral coordinates is determined by the characteristic scale (R) of the heat flux inhomogeneity.

Let us seek a solution to system (1) in the form of expansion with respect to a small parameter represented by ε, the square root of supercriticity:

$$\begin{aligned} \mathbf{u} &= \varepsilon(\mathbf{u}_0 + \varepsilon^2\mathbf{u}_2 + \varepsilon^4\mathbf{u}_4 + \ldots), \\ w &= \varepsilon^2(w_0 + \varepsilon^2 w_2 + \ldots), \\ P &= \varepsilon^{-2}P_{-2} + P_0 + \varepsilon^2 P_2 + \ldots, \\ h &= \varepsilon^2(H_0 + \varepsilon^2 H + \ldots), \\ T &= T_0 + \varepsilon^2 T_2 + \varepsilon^4 T_4 + \ldots \end{aligned} \tag{3}$$

Detailed derivation of the system of evolution equations is presented in the Appendix. Here, let us only briefly consider the main stages of this procedure.

Since the solid boundary of the system is low-heat-conducting, the long-range motions are most dangerous in the lateral directions [2]. For this reason, perturbations of the vertical temperature gradient are small (on the order of ε^2); in lower orders of perturbation theory (see Appendix), the temperature depends only on the lateral coordinates x, y. The free surface deformation is also small and the velocity field is proportional to the gradient of the average temperature field $\phi(x, y, t)$. With allowance for the continuity equation, this yields

$$\mathbf{u} \propto \nabla\phi, \quad w \propto \nabla^2\phi.$$

The critical Marangoni number is determined in second-order perturbation theory and coincides with the value determined by linear analysis [2] for a system with a nondeformable boundary: Ma* = 48 (see Appendix). Vortex flows can be described (see Appendix) in terms of the average vertical vorticity potential ψ, for which $\mathbf{e}_z\nabla^2\psi \propto \langle\nabla \times \mathbf{u}_2\rangle$ (angle brackets $\langle\ldots\rangle = \int_0^1 \ldots dz$ denote averaging over the coordinate z).

The system of closed evolution equations, relating the average temperature field $\phi(x, y, t)$, the amplitude $H(x, y, t)$ of the free surface deviation from the plane, and the potential $\psi(x, y, t)$ of the average vertical vorticity is obtained in the fourth order in ε (see Appendix):

$$\begin{aligned} &\partial_t\phi + \nabla\phi\nabla\times(\mathbf{e}_z\psi) + \nabla^4\phi - \nabla^2 H \\ &+ \frac{\sqrt{7}}{6}\nabla(H\nabla\phi) - \nabla(q\nabla\phi) + \frac{15}{8}\nabla^2 q \\ &- \nabla(|\nabla\phi|^2\nabla\phi) + \frac{\sqrt{7}}{8}\left(1 + \frac{2}{\mathrm{Pr}}\right)\nabla(\nabla^2\phi\nabla\phi) \\ &+ \frac{3\sqrt{7}}{4}\left(1 + \frac{1}{6\mathrm{Pr}}\right)\nabla^2|\nabla\phi|^2 = 0, \\ &\mathbf{e}_z\nabla^2\psi = \frac{13}{2\mathrm{Pr}}\nabla(\nabla^2\phi)\times\nabla\phi - \frac{7}{6}\nabla H\times\nabla\phi, \\ &\delta\nabla^2 H - \nabla^4 H = -c\nabla^2\phi, \end{aligned} \tag{4}$$

where

$$\delta = \frac{2}{15}\frac{G}{Ca}, \quad c = \frac{48}{5Ca}.$$

For $q(x, y) = -1$, which corresponds to a homogeneously exceeded longwave instability threshold in the entire layer, Eqs. (4) transform into the system obtained in [11]. Two additional terms (proportional to $\nabla(q\nabla\phi)$ and $\nabla^2 q$) in the first equation (4) characterize inhomogeneity of the heat flux. These terms are qualitatively similar to the terms proportional to $\nabla(H\nabla\phi)$ and $\nabla^2 H$, describing the influence of the amplitude of the free surface deviation from the plane on the dynamics of the average temperature field (see the first equation in system (4)). This is quite natural, since the local Marangoni number can be modified by two methods: via increasing the heat flux or changing the liquid layer thickness. In addition, deformation of the surface leads to a new mechanism of vorticity generation, this contribution being proportional to $\nabla H \times \nabla\phi$ (see the second equation (4)).

The physical characteristics of the gas do not influence the loss of stability in the liquid layer under the following conditions [11]:

$$\tilde{\eta}/\tilde{L} \ll 1, \quad \tilde{L}^2/\tilde{\chi} \ll 1, \quad \tilde{L}\tilde{\kappa} \ll 1, \quad \tilde{L}^3\tilde{\kappa} \ll 1, \tag{5}$$

where $\tilde{\eta} = \eta_{gas}/\eta$ is the ratio of the dynamic viscosities in the gas and liquid; $\tilde{\kappa} = \kappa_{gas}/\kappa$ and $\tilde{\chi} = \chi_{gas}/\chi$ are the ratios of the heat and conductivity and thermal diffusivity coefficients, respectively; and $\tilde{L} = L_{gas}/L$ is the ratio of thicknesses of the layers occupied by the gas and liquid phases. The first condition (5) implies that the tensor of viscous stresses in the gas phase is much smaller

than in the liquid; the second condition indicates that the characteristic temperature diffusion time in the gas phase is much smaller than that in the liquid; and the third condition reflects the small rate of heat transfer in the gas as compared to that in the liquid. The latter condition determines the regions of longwave instability for a sufficiently thin gas layer (the minimum Marangoni number falls within the longwave region). For a two-layer air–vapor system [11] at 20°C and a pressure of 1 atm, we obtain $\tilde{\eta} = 1.8 \times 10^{-2}$, $\tilde{\chi} = 1.6 \times 10^{2}$, and $\tilde{\kappa} = 4.3 \times 10^{-2}$. As can be seen, conditions (5) are satisfied provided that $0.01 \ll \tilde{L} \ll 10$.

The system of equations (4) is also applicable to the concentration Marangoni convection. In this case, the average temperature field $\phi(x, y, t)$ is replaced by the average concentration field and the thermal diffusivity coefficient, by the surfactant diffusion coefficient. Note that the concentration convection in thin layers [3] always possesses a long-range character, whereby the surfactant diffusion from liquid to gas can be ignored.

3. QASIEQUILIBRIUM: LINEAR ANALYSIS

As was pointed out above, the liquid layer features a flow induced by inhomogeneous heating in the horizontal plane. However, under certain conditions this motion may play not the main role in the formation of convective flow structures. Such a case is realized, in particular, when both inhomogeneity of the heat flux in the lateral directions and the related convective flow are small in comparison to the heat flux passing through the layer and the related long-range convection.

First, consider the heat flux inhomogeneity in the form of a step function:

$$q = \begin{cases} \pm a^2, & |\mathbf{r}| \le 1, \\ \pm b^2, & |\mathbf{r}| > 1, \end{cases} \tag{6}$$

where $\mathbf{r}$ is the radius vector in the horizontal plane. The heat flux discontinuity gives rise to short-range motions at the spot boundaries (not featured by the system at the longwave instability threshold). Therefore, these motions are insignificant in the analysis of long-range convection and we can speak of a mechanical quasiequilibrium. For this reason, $\nabla q = 0$, $\nabla^2 q = 0$ and the solution $\phi(x, y) = 0$, $H(x, y) = 0$, $\psi(x, y) = 0$ describes the state of quasiequilibrium with respect to long-range convection, which can be either stable or unstable.

The quasiequilibrium is stable, provided that all long-range perturbations tend to decay. Should one or several long-range perturbations grow with time, the quasiequilibrium becomes unstable with respect to these perturbations. Therefore, in order to judge the stability of a quasiequilibrium, it is necessary to consider the development of various long-range perturbations with time. Indeed, the form of the function (6) and the concept of quasiequilibrium allow us to decline from description of the main flow and to describe behavior of the system in terms of perturbation theory. At the same time, it is clear that flow structures arising in the presence of a thermal inhomogeneity will be qualitatively described by two parameters: deviations from the critical heat flux inside ($\pm a^2$) and outside ($\pm b^2$) the spot. Here, the minus and plus signs correspond to the heat flux above and below the critical value, as determined by the z axis direction selected. This formally implies that there are four different regions on the plane of parameters (b^2, a^2), of which we will consider the two corresponding to the first and fourth quadrants. In both these regions, the heat flux outside is always understated. In the first quadrant, there is an excess heat flux inside the spot. In the fourth quadrant, the heat flux inside the spot is also below critical, but the reduction is different inside and outside the spot. These very situations take place in experiments [3, 6].

Let us consider small nonstationary perturbations of the quasiequilibrium (linear theory of stability). Under these conditions, we may neglect the terms quadratic and cubic in perturbations in Eqs. (4). Taking into account (6) and using subscripts i and e to indicate the quantities in the internal (inside the spot) and external (outside the spot) regions, we obtain a system of linear equations

$$\begin{aligned} \partial_t \phi_i + \nabla^4 \phi_i - \nabla^2 H_i + a^2 \nabla^2 \phi_i &= 0, \\ \delta \nabla^2 H_i - \nabla^4 H_i &= -c \nabla^2 \phi_i, \\ \partial_t \phi_e + \nabla^4 \phi_e - \nabla^2 H_e - b^2 \nabla^2 \phi_e &= 0, \\ \delta \nabla^2 H_e - \nabla^4 H_e &= -c \nabla^2 \phi_e. \end{aligned} \tag{7}$$

Eliminating the surface perturbations, we obtain

$$\begin{aligned} &\partial_t (\nabla^2 \phi_i - \delta \phi_i) + \nabla^6 \phi_i + (a^2 - \delta) \nabla^4 \phi_i \\ &\quad - (c + \delta a^2) \nabla^2 \phi_i = 0, \\ &\partial_t (\nabla^2 \phi_e - \delta \phi_e) + \nabla^6 \phi_e - (b^2 + \delta) \nabla^4 \phi_e \\ &\quad - (c - \delta b^2) \nabla^2 \phi_e = 0. \end{aligned} \tag{8}$$

The boundary conditions at $|\mathbf{r}| = 1$, reflecting continuity of the average temperature field $\phi(x, y, t)$, the velocity field $\mathbf{u} \propto \nabla\phi$ and $w \propto \nabla^2\phi$, the amplitude $H(x, y, t)$ of the free surface deviation from the plane, and the Laplace pressure field (proportional to $\nabla^2 H$) are as follows:

$$\begin{aligned} &\phi_i - \phi_e = \phi_i' - \phi_e' = \nabla^2 \phi_i - \nabla^2 \phi_e \\ &= \nabla^2 \phi_i' - \nabla^2 \phi_e' = \nabla^4 \phi_i - \nabla^4 \phi_e = 0, \end{aligned} \tag{9}$$

$$\nabla^4 \phi_i' + a^2 \nabla^2 \phi_i' - a^2 \delta \phi_i' - \nabla^4 \phi_e' + b^2 \nabla^2 \phi_e' - b^2 \delta \phi_e' = 0.$$

The latter condition is supplementary and can be obtained using the Green formulas [15].

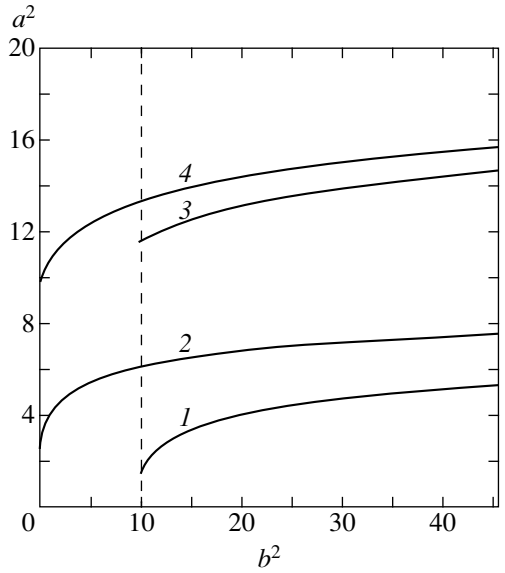


Fig. 1. The first levels of instability in the case of a one-dimensional inhomogeneity of the heat flux for (*1*, *2*) a symmetric mode $\phi(-x) = \phi(x)$ described by Eqs. (12) and (*3*, *4*) an antisymmetric mode $\phi(-x) = -\phi(x)$ described by Eqs. (13) with the parameters (*1*, *3*) $c = 10$, $\delta = 1$ and (*2*, *4*) $c = 0$, $\delta = 0$. The vertical dashed line $b^2 = c/\delta$ separates the domains of local and global instability. The domains of stability are situated under neutral curves *2* and *4*.

Let us consider the normal perturbations of the $\exp(-\lambda t)$ type. Multiplying Eqs. (8) by ϕ^*, integrating over the space with allowance for the boundary conditions (9), and using the Green formulas [15], we obtain, upon simple but cumbersome transformations,

$$\lambda = \left[\int|\nabla\nabla^2\phi|^2 dV + (q+\delta)\int|\nabla^2\phi|^2 dV + (\delta q - c)\int|\nabla\phi|^2 dV\right]\left(\int|\nabla\phi|^2\, dV + \delta\int|\phi|^2 dV\right)^{-1}, \quad (10)$$

where $q(x, y)$ is given by formula (6) and $\phi = (\phi_i, \phi_e)$. Equation (10) is indicative of a monotonic character of perturbations. This expression is conveniently presented in the dimensional form:

$$\lambda = \left[\frac{1}{R^6}\int|\nabla\nabla^2\phi|^2 dV + \frac{q+\delta}{R^4}\int|\nabla^2\phi|^2 dV + \frac{\delta q - c}{R^2}\int|\nabla\phi|^2 dV\right]\left(\frac{1}{R^2}\int|\nabla\phi|^2 dV + \delta\int|\phi|^2 dV\right)^{-1}. \quad (11)$$

At $q < c/\delta$, the third term in expression (11) is negative. For a sufficiently large spot size, this term becomes predominant, so that the decrement is negative and the flow is unstable. Thus, long-range inhomogeneity leads to instability of the induced flow. This condition determines the boundary of the domain of instability in the first and fourth quadrants on the (b^2, a^2) plane: the regions of $b^2 > c/\delta$ and $a^2 > -c/\delta$ correspond to a local instability caused by the spot. For $b^2 < c/\delta$, the loss of stability is related to the region outside the spot and possesses a global character. We will dwell on the local mode, since this very type of flow is observed in experiment [3, 6]. It should be noted that, as $R \longrightarrow 0$, the decrement is always positive and an equilibrium does exist. A small region of heating does not influence the long-wave loss of stability. This is also related to the fact that, in selecting the form of function (6), we ignore the heat flux along the liquid layer and the related convective motion. This is just what allows us to speak of a quasiequilibrium.

3.1. One-dimensional Inhomogeneity of Heat Flux

In this case, the role of coordinate **r** in the heat flux inhomogeneity described by step function (6) is played by the Cartesian coordinate x. In order to find the neutral curves ($\partial_t = 0$), let is use the following set of functions:

$$\phi_i = A_1 + A_2\cos(a_1 x) + A_3\cosh(a_2 x), \quad \phi_e = B_1\exp(-b_1|x|) + B_2\exp(-b_2|x|), \quad (12)$$

for a symmetric mode [$\phi(-x) = \phi(x)$], and

$$\phi_i = A_1 x + A_2\sin(a_1 x) + A_3\sinh(a_2 x), \quad |x| < 1,$$
$$\phi_e = \begin{cases} B_1\exp(-b_1|x|) + B_2\exp(-b_2|x|), & x > 1, \\ -B_1\exp(-b_1|x|) - B_2\exp(-b_2|x|), & x < -1, \end{cases} \quad (13)$$

for an antisymmetric mode $\phi(-x) = -\phi(x)$. Substituting expressions (12) and (13) into (8) yields a system of algebraic equations for determining the roots a_j and b_j:

$$a_j^4 \pm (\delta - a^2)a_j^2 - \delta a^2 - c = 0, \quad b_j^4 - (\delta + b^2)b_j^2 - \delta b^2 - c = 0. \quad (14)$$

Upon solving system (14), we obtain

$$a_{1,2} = \frac{1}{\sqrt{2}}\sqrt{\pm\left(a^2 - \delta \pm \sqrt{(\delta - a^2)^2 + 4(\delta a^2 + c)}\right)}, \quad b_{1,2} = \frac{1}{\sqrt{2}}\sqrt{b^2 + \delta \pm \sqrt{(\delta + b^2)^2 - 4(\delta b^2 - c)}}, \quad (15)$$

where the plus and minus signs correspond to the roots a_1 (b_1) and a_2 (b_2), respectively.

The neutral curves in the (b^2, a^2) plane are determined from boundary conditions (9). Figure 1 presents the neutral curves for $c = 0$, $\delta = 0$ in a system with a nondeformable upper boundary and for $c = 10$, $\delta = 1$ in

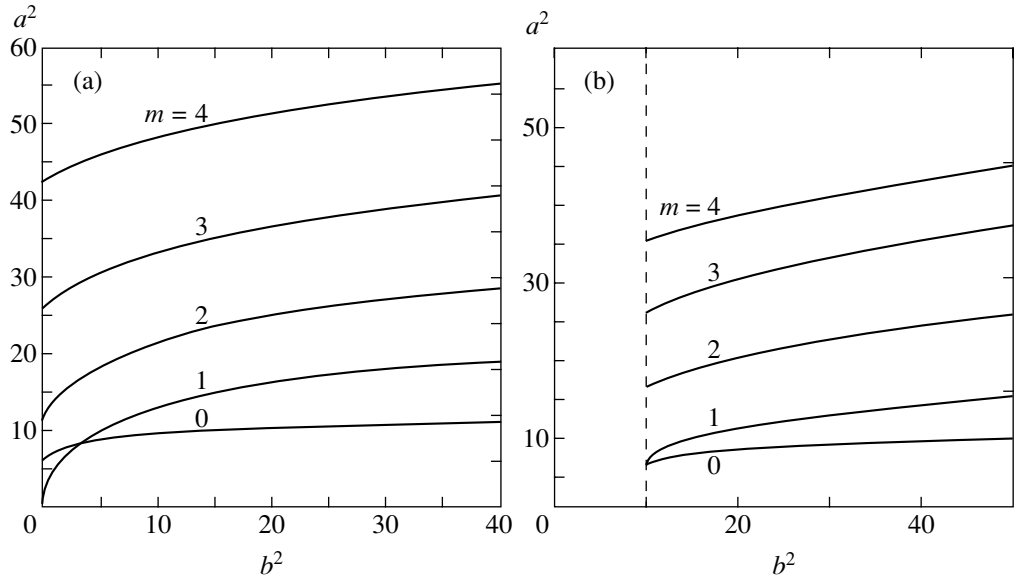


Fig. 2. The first levels of instability for the modes described by system (18) with various values of the azimuthal number m for the parameters (a) $c = 0$, $\delta = 0$ and (b) $c = 10$, $\delta = 1$. The domains of stability are situated under neutral curves; the vertical dashed line $b^2 = c/\delta$ separates the domains of local and global instability.

a system with free deformable surface. These curves correspond to the first levels of instability for the symmetric and antisymmetric modes. The domain of stability is situated under the neutral curves. The vertical dashed line shows the boundary ($b^2 = c/\delta$) separating the domains of local and global instability. As can be seen, the boundary of stability with respect to parameter a^2 (characterizing a supercritical heat flux inside the heat spot) is determined by the damping parameter (b^2) outside the spot and approaches an asymptotic level at sufficiently large b^2 values. The influence of the parameters c and δ is manifested by a change in the values $b^2 = c/\delta$ corresponding to the boundaries of local instability (see expression (11)) determined by the spot ($b^2 > c/\delta$) and global instability caused by the region outside the spot ($b^2 < c/\delta$). The free surface deformation leads to a decrease in the stability threshold for all levels of both symmetric and antisymmetric modes. In the entire region of parameters, the stability threshold for symmetric modes is lower than that for antisymmetric ones.

In the case of $c = 0$ and $\delta = 0$, the neutral curves admit analytical determination. Indeed, the symmetric model is described by the relation

$$b\sin a + a\cos a = 0, \tag{16}$$

and the antisymmetric mode corresponds to

$$b^2 a\cos a - (a^2 b + a^2 + b^2)\sin a = 0. \tag{17}$$

The results of calculations show that the neutral curves are qualitatively similar for any values of c and δ, whereby the predominant quantity determining the pattern is the ratio $c/\delta \sim G^{-1}$. This conclusion follows from relation (11) for large heat flux inhomogeneities as compared to the liquid layer thickness.

3.2. Axisymmetric Inhomogeneity of Heat Flux

In order to find the neutral curves ($\partial_t = 0$) in this case, let is use a set of eigenfunctions of the Laplace operator in a cylindrical coordinate system determined by the symmetry of the spot:

$$\begin{aligned}\phi_i &= [A_1 J_m(a_1 r) + A_2 I_m(a_2 r) + A_3 r^m]e^{im\varphi} + \ldots,\\ \phi_e &= [B_1 K_m(b_1 r) + B_2 K_m(b_2 r) + B_3 r^{-m}]e^{im\varphi} + \ldots,\end{aligned} \tag{18}$$

where J_m, I_m, and K_m are the Bessel functions of the first kind, the modified Bessel functions, and the Hankel functions, respectively; the azimuthal number m determines the number of lobes [3] or spiral arms [6] in the convective flow pattern.

Under the action of operator ∇, Eqs. (18) yield

$$\begin{aligned}&\nabla^2\phi_i = -a_j^2\phi_i, \quad \nabla^4\phi_i = a_j^4\phi_i, \quad \nabla^6\phi_i = -a_j^6\phi_i,\\ &\nabla^2\phi_e = b_j^2\phi_e, \quad \nabla^4\phi_e = b_j^4\phi_e, \quad \nabla^6\phi_e = b_j^6\phi_e.\end{aligned} \tag{19}$$

Substituting expressions (18) into system (8) and taking into account relations (19), we obtain expressions for the roots a_j, b_j that coincide with formulas (15). The neutral curves in the (b^2, a^2) plane are determined from boundary conditions (9).

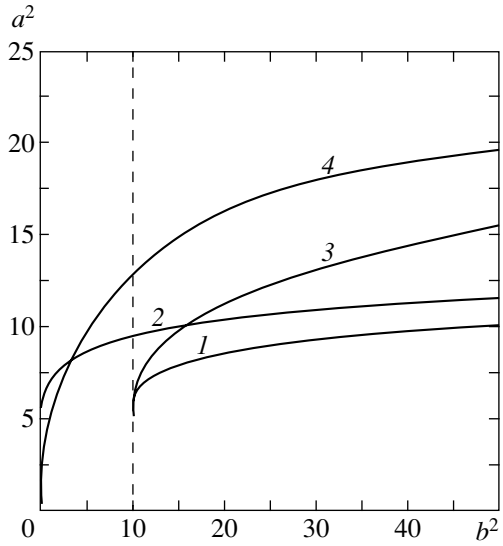


Fig. 3. The first levels of instability for (*1*, *2*) an axisymmetric mode with $m = 0$ and (*3*, *4*) a dipole mode with $m = 1$ for the parameters (*1*, *3*) $c = 10$, $\delta = 1$ and (*2*, *4*) $c = 0$, $\delta = 0$. The vertical dashed line $b^2 = c/\delta$ separates the domains of local and global instability. The domains of stability are situated under neutral curves.

Figure 2 shows neutral curves for the convective modes (see Eqs. (18)) at various values of the parameters c and δ. Here, the domain of stability of the convective mode is situated under the corresponding neutral curve. The vertical dashed line represents the boundary ($b^2 = c/\delta$) separating the domains of local and global instability. As can be seen, deformation of the free surface leads to a decrease in the stability threshold for the modes with any azimuthal number $m > 1$. However, which type of the modes (axisymmetric with $m = 0$ or dipole with $m = 1$) is most dangerous depends on the parameters c and δ. In the region of $c/\delta > 10$, more dangerous is the axisymmetric mode; at lower values of this ratio, there is a region in the (b^2, a^2) plane in which the dipole mode ($m = 1$) is more dangerous. Figure 3 illustrates a change in the symmetry of modes upon symmetry breakage depending on the parameters characterizing deformation of the free liquid surface.

Similarly to the one-dimensional case, the pattern of neutral curves is sensitive with respect to the parameter $c/\delta \sim G^{-1}$ and the boundary of stability with respect to parameter a^2 (characterizing a supercritical heat flux inside the heat spot) is determined by the damping parameter (b^2) outside the spot and approaches an asymptotic level at sufficiently large b_2 values.

In the case of $c = 0$ and $\delta = 0$, the neutral curves also admit analytical determination depending on the azimuthal number:

$$\frac{J_{m+1}(a)}{aJ_m(a)} + \frac{K_{m+1}(b)}{bK_m(b)} = \frac{2m}{b^2}\frac{a^2 + b^2}{a^2 - b^2}. \qquad (20)$$

4. CONCLUSIONS

We have studied the Marangoni convection under weakly supercritical conditions in a layer of liquid with a free deformable upper surface and an inhomogeneous heating source. In a longwave approximation with weakly inhomogeneous heat flux, the problem reduces to a system of two-dimensional nonlinear equations in terms of the amplitudes of perturbations in the temperature, vorticity, and free surface deformation. After introduction of the concept of qasiequilibrium (in which the system no longer features short-range motions), the inhomogeneity was modeled by a step function. This simplification opens way to studying a still rather complicated system and characterizing instabilities depending on the supercritical heat flux exceeding the stability threshold inside the heating spot and reduced heat flux (damping) outside the spot.

In the cases of planar and axisymmetric inhomogeneities of the heat flux, we determined the boundaries of stability of the convective flows on the plane of parameters characterizing supercritical heat flux inside the spot and the degree of its damping outside the spot for various values of the capillary and Galilean numbers. For an axisymmetric spot, the domains of stability with respect to perturbations were established for various azimuthal numbers. The results of a linear analysis of the problem in the case of a stepwise heating inhomogeneity lead to the following conclusions. The type of longwave instability of the equilibrium (global versus local mode) is determined by the degree of reduction in the heat flux outside the local source: the global instability changes to local at a certain value of the parameter $b^2 = c/\delta$. This threshold value, determining the boundary of a neutral stability curve, is zero for a nondeformable surface and increases for the other parameters c and δ characterizing deformation of the free surface.

In the case of an axisymmetric heating spot, the boundary of the local instability of equilibrium and the type of most dangerous instability mode (axisymmetric versus azimuthal) are determined by the dimensionless parameters $a^2 + c/\delta$, $b^2 - c/\delta$ characterizing the heat source intensity, dimensions, and the effects of the surface tension and force of gravity. There is a certain region in the space of parameters in which the dipole mode ($m = 1$) is more dangerous than the axisymmetric mode ($m = 0$). Deformation of the free liquid surface provides for a new pathway for the heat dissipation, thus changing the symmetry of modes capable of breaking the stability of equilibrium. In the range of control parameters $c/\delta < 10$, there is a region on the plane of parameters (b^2, a^2) in which the dipole mode ($m = 1$) is most dangerous. As the control parameter (and, hence, the surface deformation amplitude) increases, the axisymmetric mode becomes most dangerous in the entire range of heat flux discontinuities. The instability threshold with respect to the modes with

$m > 1$ is always higher in the entire region of parameters studied.

A comparison with the experimental data reported in [3] shows that the above theoretical results qualitatively describe the multilobe convection structures. The neglect of the horizontal heat flux component (and, hence, of the induced flow as well) leads to a monotonic character of perturbation development with time, which hinders the description of complicated oscillatory regimes observed in [6]. A numerical analysis of system (4) has been reported in [16].

ACKNOWLEDGMENTS

The author is grateful to A.F. Pshenichnikov, K.I. Morozov, B.I. Myznikova, A.I. Mizev, and I.I. Vertgeim for fruitful discussions and useful remarks.

This study was supported by the Russian Foundation for Basic Research (RFBR–Urals project no. 02-01-96407) and by the U.S. Civilian Research and Development Foundation for Independent States of the Former Soviet Union (CRDF Award No. RE-009-0).

APPENDIX

The method of nonlinear perturbation theory employed below is analogous to the procedure described in [8, 11]. For the sake of convenience, the boundary conditions (2) are transferred from the surface $z = 1 + h(x, y)$ to the plane $z = 1$ by using expansions of the velocity, temperature, and pressure fields into Taylor series in small deviations of the surface from the plane ($|h| \ll 1$). In the lowest order with respect to the supercriticity parameter, this yields the following equations and boundary conditions:

$$P'_{-2} + \mathrm{G} = 0, \quad z = 1\colon P_{-2} = 0, \tag{A.1}$$

$$P'_0 = 0, \quad T''_0 = 0, \quad -\nabla P_0 + \mathbf{u}''_0 = 0, \quad -\nabla \mathbf{u}_0 = w'_0, \tag{A.2}$$

$$z = 0\colon \mathbf{u}_0 = 0, \quad w_0 = 0, \quad T'_0 = -1, \tag{A.3}$$

$$z = 1\colon P_0 + HP'_{-2} = -\mathrm{Ca}\nabla^2 H, \quad w_0 = 0, \quad \mathbf{u}'_0 = -\mathrm{Ma}^*\nabla T_0, \quad T'_0 = -1. \tag{A.4}$$

The solution of Eq. (A.1) determines the pressure field

$$P_{-2} = \mathrm{G}(1 - z), \quad P_0 = Q_0 = \mathrm{G}H - \mathrm{Ca}\nabla^2 H, \tag{A.5}$$

where $Q_0(x, y, t)$ is the pressure as a function of the lateral coordinates and $H(x, y, t)$ is the surface deformation amplitude, which are still unknown functions of recalculated coordinates x, y and the current time t.

The solution of the second equation in (A.2) is

$$T_0 = -z + \phi(x, y, t), \tag{A.6}$$

where $\phi(x, y, t)$ is an unknown temperature field, which will be determined below in the fourth order of perturbation theory.

Integrating the continuity equation (third equation in (A.2)) with respect to coordinate z, we obtain

$$\nabla \int_0^{1+h} \mathbf{u}_0 dz \approx \nabla \int_0^1 \mathbf{u}_0 dz + O(h) = 0 \longrightarrow \langle \mathbf{u}_0 \rangle = 0, \tag{A.7}$$

where $\langle \ldots \rangle = \int_0^1 \ldots dz$. In this order of perturbation theory, a shift of the surface ($h \sim \varepsilon^2$) does not influence the results of averaging through the layer.

In order to take into account the solenoidal component of the velocity field, let us introduce the vorticity integral

$$\int_0^1 2z\mathbf{e}_z \nabla \times \psi_0 dz = \nabla \times (\mathbf{e}_z \psi_0).$$

Applying the operations of rotor and divergence to the Marangoni boundary conditions (third equation in (A.4)), we obtain

$$\nabla^2 \psi_0 = 0, \quad \nabla^2 Q_0 = -\frac{3}{2}\mathrm{Ma}^*\nabla^2 \phi. \tag{A.8}$$

Since there are no mass forces creating the vorticity, we set $\psi_0 = 0$. Then, the solution for the velocity is

$$\mathbf{u}_0 = -\frac{3}{2}\left(\frac{z^2}{2} - \frac{z}{3}\right)\mathrm{Ma}^*\nabla^2 \phi, \quad w_0 = \frac{1}{4}(z^3 - z^2)\mathrm{Ma}^*\nabla^2 \phi. \tag{A.9}$$

In the second order of perturbation theory,

$$P'_2 = w''_0, \quad z = 1\colon P_2 + HP'_0 = -\mathrm{Ca}\nabla^2 H_2 + 2w'_0, \tag{A.10}$$

$$\mathbf{u}''_2 = \nabla P_2 - \nabla^2 \mathbf{u}_0 + \frac{1}{\mathrm{Pr}}[(\mathbf{u}_0\nabla)\mathbf{u}_0 + w_0\mathbf{u}'_0], \tag{A.11}$$

$$T''_1 = -\nabla^2 \phi + w_0 T'_0 + \mathbf{u}_0 \nabla T_0, \tag{A.12}$$

$$\nabla \mathbf{u}_2 = -w'_2, \tag{A.13}$$

$$z = 0\colon \mathbf{u}_2 = 0, \quad w_2 = 0, \quad T'_2 = q(x, y), \tag{A.14}$$

$$z = 1\colon w_2 = \mathbf{u}_0 \nabla H - Hw'_0, \quad T'_2 = q(x, y), \tag{A.15}$$

$$\mathbf{u}'_2 + \mathrm{Ma}^*\nabla T_2 = -H\mathbf{u}''_0 + \mathrm{Ma}^*\nabla H. \tag{A.16}$$

The search for a solution begins with the temperature field T_2. Integrating Eq. (A.12) with respect to z and using the boundary conditions on the lower interface, we determine the integration constant $C(x, y) = q(x, y)$, while the boundary conditions on the upper interface yield the critical Marangoni number Ma* = 48. Equations (A.5) and (A.8) lead to the following relation between the free surface deformation and the temperature field:

$$\nabla^2(GH - \mathrm{Ca}\nabla^2 H) = -72\nabla^2\phi. \tag{A.17}$$

In this order of perturbation theory, integration of the continuity equation (A.13) with respect to z must be performed taking into account deformation of the free liquid surface:

$$\nabla\int_0^{1+h}\mathbf{u}_2 dz = -w_2(1+h)$$

$$\longrightarrow \nabla\int_0^1\mathbf{u}_2 dz + h\nabla\mathbf{u}_2 = -w_2 - w_2' h + O(h^2),$$

which leads to the relation

$$\nabla\int_0^1\mathbf{u}_2 dz = -w_2(1). \tag{A.18}$$

Using the kinematic condition (A.15), we obtain

$$\nabla\langle\mathbf{u}_2\rangle = 12\nabla(H\nabla\phi). \tag{A.19}$$

In order to describe the solenoidal velocity field, we introduce the vorticity potential ψ_2 determined in the same way as in zero-order perturbation theory. This potential is related to the surface deformation as

$$\langle\mathbf{u}_2\rangle = \mathbf{e}_z\times\nabla\psi_2 + 12H\nabla\phi. \tag{A.20}$$

Upon introducing the integral operator

$$\hat{\mathbf{I}}[\ldots] = \int_0^z \ldots dz,$$

the solution to the system of equations (A.10)–(A.16) obtained using relations (A.7) and (A.18)–(A.20) is as follows:

$$\mathbf{e}_z\nabla^2\psi_2 = -\frac{72}{\mathrm{Pr}}\{\langle ff''\rangle - 2\langle\hat{\mathbf{I}}^2(ff'')\rangle\} \times\nabla(\nabla^2\phi)\times\nabla\phi - 24\nabla H\times\nabla\phi, \tag{A.21}$$

$$T_2 = \left\{-\frac{z^2}{2} + 12\hat{\mathbf{I}}^2[f]\right\}\nabla^2\phi + 12\hat{\mathbf{I}}[f]|\nabla\phi|^2 + zq + \phi_2, \tag{A.22}$$

$$P_2 = -12f'\nabla^2\phi + Q_2(x, y), \tag{A.23}$$

$$\mathbf{u}_2 = -24\{\hat{\mathbf{I}}[f] - 2z\langle\hat{\mathbf{I}}[f]\rangle\}\nabla\nabla^2\phi + 2z\mathbf{e}_z\times\nabla\psi_2(x, y) + \left(\frac{z^2}{2} - \frac{z}{3}\right)\nabla Q_2 + \frac{72}{\mathrm{Pr}}\{\hat{\mathbf{I}}^2[(f')^2] - 2z\langle\hat{\mathbf{I}}^2[(f')^2]\rangle\}\nabla|\nabla\phi|^2 + 24zH\nabla\phi + \frac{144}{\mathrm{Pr}}\{\hat{\mathbf{I}}^2[(ff'')] - 2z\langle\hat{\mathbf{I}}^2[(ff'')]\rangle\}\nabla\phi\nabla^2\phi, \tag{A.24}$$

$$w_2 = \frac{f}{6}\nabla^2 Q_2 + 24\{\hat{\mathbf{I}}^2[f] - z^2\langle\hat{\mathbf{I}}[f]\rangle\}\nabla^2\nabla^2\phi - 12z^2\nabla(H\nabla\phi) - \frac{72}{\mathrm{Pr}}\{\hat{\mathbf{I}}^3[(f')^2] - z^2\langle\hat{\mathbf{I}}^2[(f')^2]\rangle\}\nabla^2|\nabla\phi|^2 + \frac{144}{\mathrm{Pr}}\{\hat{\mathbf{I}}^3[(ff'')] - z^2\langle\hat{\mathbf{I}}^2[(ff'')]\rangle\}\nabla(\nabla\phi\nabla^2\phi), \tag{A.25}$$

$$\nabla Q_2 = -936\langle\hat{\mathbf{I}}[f]\rangle\nabla\nabla^2\phi + 72\{\nabla H - \nabla|\nabla\phi|^2 + H\nabla\phi + \nabla(q + \phi_2)\} + \frac{216}{\mathrm{Pr}}\{\langle ff''\rangle - 2\langle\hat{\mathbf{I}}^2[(ff'')]\rangle\}\nabla\phi\nabla^2\phi - \frac{108}{\mathrm{Pr}}\{\langle(f')^2\rangle - 2\langle\hat{\mathbf{I}}^2[(f')^2]\rangle\}\nabla|\nabla\phi|^2 + 36\nabla\nabla^2\phi, \tag{A.26}$$

where $f(z) = z^2 - z^3$. Below, we will omit subscript 2 in writing the potential ψ_2 determined by Eq. (A.21).

In order to obtain an equation describing evolution of the temperature field ϕ, let us pass to the fourth order of perturbation theory. Using the continuity condition,

$$\phi(\nabla\mathbf{u}_2 + w_2') = 0, \quad T_2(\nabla\mathbf{u}_0 + w_0') = 0, \tag{A.27}$$

we obtain (in this order of the theory) the following equation with the boundary conditions

$$\partial_t\phi + \nabla(\mathbf{u}_2\phi + \mathbf{u}_0 T_2) + (w_2\phi + w_0 T_2)' - w_2 - \nabla^2 T_2 = T_4'', \tag{A.28}$$

$$z = 0: \ T_4' = 0, \tag{A.29}$$

$$z = 1: \ T_4' + T_2''H - \nabla H\nabla\phi = 0, \tag{A.30}$$

which determine the dynamics of the temperature perturbations.

Averaging with respect to coordinate z with allowance for Eqs. (A.29) and (A.30) yields

$$\partial_t\phi + \nabla\langle\mathbf{u}_2\phi + \mathbf{u}_0 T_2\rangle + w_2\phi|_{z=1} - \langle w_2 + \nabla^2 T_2\rangle = \nabla(H\nabla\phi) + 12H|\nabla\phi|^2. \tag{A.31}$$

Upon changing the scale,

$$(x, y) \longrightarrow \sqrt{\frac{2}{15}}(x, y), \quad t \longrightarrow \frac{4}{15}t,$$

$$\phi \longrightarrow \frac{\sqrt{7}}{12}\phi, \quad H \longrightarrow \frac{\sqrt{7}}{24}H,$$

$$\psi \longrightarrow \frac{1}{2}\psi, \quad q \longrightarrow \frac{1}{2}q,$$

and using relation (A.18) for the calculation of (A.31), we arrive at the system of equations (5). Note that the fields ϕ_2 (A.22) and H_2 (A.10) do not influence the dynamics of the average temperature field ϕ.

REFERENCES

1. J. K. A. Pearson, J. Fluid Mech. **4**, 489 (1958).
2. G. Z. Gershuni and E. M. Zhukhovitskiĭ, *Convective Stability of an Incompressible Liquid* (Nauka, Moscow, 1972).
3. A. F. Pshenichnikov and S. S. Yatsenko, in *Hydrodynamics* (Permsk. Gos. Univ., Perm, 1974), No. 5, p. 33.
4. A. B. Ezersky, A. Garcimartin, J. Burguete, *et al.*, Phys. Rev. E **47**, 1126 (1993).
5. E. Favre, L. Blumenfeld, and F. Daviaud, Phys. Fluids **9**, 1473 (1997).
6. Yu. K. Bratukhin, S. O. Makarov, and A. I. Mizev, Izv. Akad. Nauk, Mekh. Zhidk. Gaza, No. 2, 92 (2000).
7. V. S. Sorokin, Prikl. Mat. Mekh. **18**, 197 (1954).
8. L. Shtilman and G. Sivashinsky, Physica D (Amsterdam) **52**, 477 (1991).
9. E. Knobloch, Physica D (Amsterdam) **41**, 450 (1990).
10. L. M. Pismen, Phys. Lett. A **116**, 241 (1986).
11. A. A. Golovin, A. A. Nepomnyashchy, and L. M. Pismen, Physica D (Amsterdam) **81**, 117 (1995).
12. D. V. Lyubimov and A. A. Cherepanov, in *Convective Flows* (Permsk. Gos. Univ., Perm, 1991), p. 17.
13. V. L. Gertsberg and G. I. Sivashinsky, Prog. Theor. Phys. **6**, 1219 (1981).
14. L. D. Landau and E. M. Lifshitz, *Course of Theoretical Physics*, Vol. 6: *Fluid Mechanics*, 4th ed. (Nauka, Moscow, 1988; Pergamon, New York, 1987).
15. G. A. Korn and T. M. Korn, *Mathematical Handbook for Scientists and Engineers* (McGraw-Hill, New York, 1968; Nauka, Moscow, 1984).
16. B. I. Myznikova and I. I. Vertgeĭm, in *Hydrodynamics* (Permsk. Gos. Univ., Perm, 2002), No. 13, p. 18.

Translated by P. Pozdeev

Journal of Experimental and Theoretical Physics, Vol. 96, No. 4, 2003, pp. 730–746.
Translated from Zhurnal Éksperimental'noĭ i Teoreticheskoĭ Fiziki, Vol. 123, No. 4, 2003, pp. 828–845.

FLUIDS

Screening of Strongly Charged Macroparticles in Liquid Electrolyte Solutions

N. F. Bunkin* and F. V. Bunkin

Research Center for Wave Studies, Prokhorov Institute of General Physics, Russian Academy of Sciences, ul. Vavilova 38, Moscow, 119991 Russia

e-mail: nbukin@orc.ru

Received September 25, 2002

Abstract—A modified Poisson–Boltzmann model has been proposed which makes it possible to describe the screening of strongly charged macroparticles in liquid electrolyte $Z:Z$ solutions in the case when parameter $B = ZeQ_0/\varepsilon RT \gg 1$ (Q_0 is the surface electric charge, T is the temperature, ε is the solution permittivity, and Z is the valence of ions) provided that the solution is dilute: $\kappa R \equiv (8\pi Z^2 e^2 n_{i0}/\varepsilon T)^{1/2} R \ll 1$ (n_{i0} is the equilibrium number density of ions). It is assumed that the charge Q_0 of a macroparticle appears as a result of adsorption of ions of a certain polarity on its surface. Quantitative criteria of division of dissolved ions into capable and incapable of adsorption are formulated. For aqueous solutions, the adsorption mechanism always leads to values of $B \gg 1$. It is shown that the charge inversion effect predicted by other authors on the basis of different models must be observed for such solutions for all $Z \geq 1$. The effect of Brownian movement of macroparticles on their screening is considered. It is shown that viscous forces emerging during such movement lead to peripheral destruction ("washing out") of the screening ionic shell of macroparticles and, as a result, to violation of their electroneutrality. This results in the emergence of two types of oppositely charged compound particles with small radii close to R and with radii much larger than R, the charge polarity of the latter being opposite to the polarity of Q_0. It is found that both types of ions of compound particles obey the "law of distribution" of the mean energy of their electric field, expressed by formula (29). The problem of ionic screening of gas bubbles accompanied by the formation of bubstons (bubbles stabilized by ions) is considered separately. It is shown that the bubston radius R in pure water and in aqueous solutions of electrolytes is equal to 14 nm irrespective of the ion number density n_{i0}. The value of n_{i0} determines the number density n_b of bubstons themselves, which are formed spontaneously under equilibrium conditions. © *2003 MAIK "Nauka/Interperiodica".*

1. INTRODUCTION

A macroparticle (in particular, a gas bubble) in a liquid solution of electrolyte acquires an electric charge Q_0 owing to certain physicochemical processes on its surface. Under equilibrium conditions, this charge is screened by the ionic shell with electric charge density $\delta(\mathbf{r})$. Henceforth, we will consider only spherical particles of radius R and, accordingly, assume that charge density $\delta(\mathbf{r})$ is spherically symmetric. The screening (quasineutrality) condition can be written in the form

$$Q_0 + 4\pi\int_R^\infty r^2\delta(r)dr = 0 \qquad (1)$$

(r is the distance from the particle center). Depending on the value of parameter $B = Z_e Q_0/\varepsilon RT$ (here and below, we assume that $Q_0 > 0$ and consider a binary $Z:Z$ solution, where Z is the valence of dissolved ions, e is the elementary charged, T is the temperature in energy units, and ε is the permittivity of the liquid; the Gaussian system of units is used), particles can be divided into weakly charged when $B \ll 1$ and strongly charged when $B \gg 1$.

In the case when $B \ll 1$, using the Poisson–Boltzmann equation in the Debye–Hückel approximation, we obtain the well-known result [1] for the charge density distribution $\delta(r)$, satisfying condition (1):

$$\delta(r) = -\frac{Q_0\kappa^2}{4\pi(1+\kappa R)}\frac{e^{-\kappa(r-R)}}{r}, \quad r \geq R, \qquad (2)$$

$$\kappa = a_D^{-1} = \sqrt{8\pi Z^2 e^2 n_{i0}/\varepsilon T} = Z\sqrt{8\pi l_B/n_{i0}}. \qquad (2a)$$

Here, $l_B = e^2/\varepsilon T$ is the Bjerrum length ($l_B = 7$ Å for water at room temperature), a_D is the Debye radius, and n_{i0} is the equilibrium density of ions (for $r = \infty$). A distinguishing feature of weakly charged particles is that, in accordance with Eq. (2), the sign of density $\delta(r)$ is opposite to the sign of the surface charge Q_0 for all values of $r \geq R$; in other words, counterions always prevail in the screening ionic shell (ions adsorbed at the surface of a particle will be referred to as auxiliary and charge-forming ions). As a result, ionic screening of a charged macroparticles is impossible, in accordance with Eq. (1), for finite values of radius r. For a long time, this circumstance was indisputable for specialists in colloi-

dal systems irrespective of values of parameter B. However, some experimental data that have been obtained in recent years obviously have contradicted the concept of a screening ionic shell with absolute prevalence of counterions. Above all, this concerns the experimental proof of attraction of likely charged macroparticles in aqueous solutions of electrolytes under certain conditions [2–4], which contradicts the Deryagin–Landau–Verwey–Overbeck (DLVO) theory [5, 6] based on the Debye screening (2). Such contradictory facts also include recent experimental results on electrophoresis of macroparticles in solutions of various electrolytes, indicating an ambiguous dependence of the electrophoretic mobility of particles on the polarity of their surface charge [7–9]. An example of such a contradiction is the situation with stable gas bubbles (bubstons) in pure water. According to [10, 11], such bubbles must carry a positive surface charge of adsorbed H^+ ions, which is screened by both types of intrinsic OH^- and H^+ water ions. At the same time, experiments on electrophoresis of bubbles in pure water indicate that their charge is negative. In contradiction to [10, 11], it is concluded from these data that OH^- ions are adsorbed on the surface of bubbles in pure water.

These contradictions, as well as other facts, stimulated a large number of theoretical publications [11, 12], in which the so-called inversion (or superscreening) effect of the charge of macroparticles in aqueous solutions of electrolytes is predicted. The effect is that the charge density distribution $\delta(r)$ for strongly charged particles in the ionic shell can reverse its sign upon an increase in r. In this case, the sign of the total charge within a sphere with a certain finite value of radius $r > R$,

$$Q(r) = Q_0 + 4\pi\int_R^r x^2\delta(x)dx, \qquad (3)$$

can be inverted; i.e., $\operatorname{sgn} Q(r) = -\operatorname{sgn} Q_0$. The existence of this effect opens wide possibilities in interpreting the experimental facts mentioned above. In the cited publications, theoretical concepts are developed according to which condensation of counterions can occur on the surface of a macroparticle with the formation of a two-dimensional strongly correlated liquid for large values of B. Under certain conditions, the charge density (per unit area) of such a liquid may exceed in absolute value the unscreened (initial) surface charge density $Q_0/4\pi R^2$, which is the main reason for charge inversion. According to the results obtained by these authors, the screening of the inverted charge can be successfully described on the basis of the Poisson–Boltzmann equation.

An important step in explaining the details of the charge inversion effect was publication [13], which described the results of numerical counting aimed at studying the structure of the screening ionic shell of a strongly charged macroparticle in an aqueous salt solution. It was assumed that charge $Q_0 = 20e$ and that the ionic salt solution is a binary 2 : 2 or 1 : 1 solution. In this case, parameter $B = ZeQ_0/\varepsilon RT \equiv 20Z(l_B/R)$; consequently, parameter B for 2 : 2 and 1 : 1 solutions was equal to $280/R(\text{Å})$ and $140/R(\text{Å})$, respectively. In calculations, only the Coulomb interaction between dissolved ions and between ions and charged macroparticles was taken into account for $T = 300$ K. For various concentrations of both types of solutions, only three values of parameter B were taken into account: $B = 14$ (2 : 2 solution, $R = 20$ Å, and 1 : 1 solution, $R = 10$ Å), $B = 22.4$ (2 : 2 solution, $R = 12.5$ Å), and $B = 28$ (2 : 2 solution, $R = 10$ Å). The main result obtained in [13] is as follows: inversion is absent for $B = 14$ and inversion is present for $B = 22.4$ and 28. It is significant that the sign reversal of charge $Q(r)$ always occurs not jumpwise, as follows from the concepts of the formation of a two-dimensional liquid of counterions, but as a result of a smooth tendency of $Q(r)$ to zero for $r \longrightarrow R + \Delta r$, where Δr is the scale of a considerable decrease (approximately, to one-third of the initial value) in the volume density of counterions near the particle surface (in the immediate vicinity of the surface, their density is maximal, but finite). Thus, the results obtained in [14] indicate that, first, the inversion effect is indeed of the threshold type in parameter B. Second, charge inversion occurs not at the particle surface proper, but as a result of gradual mixing of counterions with secondary ions upon an increase in $(r - R)$. For large values of B, such mixing leads to prevalence of secondary ions, i.e., to sign reversal of the charge density $\delta(r)$ and to the emergence of inversion of charge $Q(r)$. These conclusions make it possible to refute the model of the inversion effect, which stems from ideas on surface condensation of counterions (formation of a two-dimensional liquid) and consider another model, which is free of such concepts. This model is proposed here as applied to dilute solutions of electrolytes, for which parameter $\zeta \equiv \kappa R \ll 1$. We proceed from the fact that the charge Q_0 of a macroparticle emerges on its surface as a result of adsorption of one of two types of ions dissolved in the liquid ($Z : Z$ solution). The possibility of such an ion-adsorption mechanism of emergence of Q_0 for gas bubbles was considered in our earlier publication [11]. Here, we consider the possibilities of the ion-adsorption mechanism for spherical macroparticles formed from an arbitrary dielectric with permittivity $\varepsilon_p \ll \varepsilon$ (e.g., solid dielectric spheres, oil drops, or gas bubbles in aqueous solutions) on another theoretical basis (as compared to [11]). The term "macroscopic" will be henceforth applied to particles for which $R \gg 3\delta_l \sim (0.3–1.0)$ nm (δ_l is the radius of a liquid molecule). In the course of model construction, this concept will be refined, but in all cases we will speak of nanometer-size and coarser particles. The main result of this analysis (Section 2) is as follows: macroparticles present in a liquid ionic solution and capable of stable adsorption of ions on their surface are always strongly charged ($B \gg 1$) and, hence, their ionic screening

always differs from the Debye screening (2). Quantitative conditions for the selection of dissolved ions in respect to the possibility of their strong adsorption are formulated. Section 3 is devoted to quantitative analysis of the proposed model of screening of strongly charged particles; the conditions for the emergence of the charge inversion effect are considered. In Section 4, we consider the effect of thermal (Brownian) movement of macroparticles on their screening. It should be emphasized that, although we consider liquid solutions of electrolytes, the results obtained pertain equally to pure liquids with ion-type conductivity. In particular, this also applies to pure (e.g., distilled) water containing only its own ions H^+ and OH^-, appearing due to thermal electrolytic dissociation of H_2O molecules.

2. STRONGLY CHARGED MACROPARTICLES: RESULTS OF ADSORPTION OF DISSOLVED IONS OF CERTAIN POLARITY ON THEIR SURFACE

If the surface charge Q_0 of a particle is due to adsorption of ions with charge $Ze > 0$ on its surface, parameter

$$B = ZeQ_0/\varepsilon RT = 4\pi Z^2 l_B R\gamma_i^{AD} \approx 9Z^2\bar{R}\gamma_i^{AD}\ (\text{nm})^{-2}, \quad (4)$$

where γ_i^{AD} is the surface density of adsorbed ions and $\bar{R} \equiv R$ [nm] is (here and below) the radius of a particle in nanometers. The last equality in relation (4) corresponds to aqueous solutions at room temperature. It shows that, for $\gamma_i^{AD} \geq 10^{14}$ cm^{-2} = 1 nm^{-2} and $\bar{R} > 1$, parameter $B \gg 1$ irrespective of valence Z.[1]

The possibility of adsorption of an ion at the surface of a liquid coating a macroparticle is due to the existence of attraction energy between the ion and a neutral macroparticle (atom or molecule), which is known to be equal to $(-\beta q^2/2r^4)$, where β is the electronic polarizability of the neutral particle, q is the ion charge, and r is the distance between the ion and the particle. In our case, we must set $q = Ze$ and $r = (\delta_l + \delta_i)$, where δ_i is the ion radius. The adsorption energy U_{AD} in this case is determined by the maximal (absolute) value of the energy corresponding to simultaneous attraction of the ion to three surface molecules of the liquid with the formation of a densely packed tetrahedron.[2] Thus, we can assume that $U_{AD} = (3/2)Z^2e^2\beta/\delta_0^4$, where $\delta_0 = (\delta_l + \delta_i)$. For $\delta_0 \sim 3$ Å and $\beta \sim 10^{-24}$ cm^3, we have $U_{AD} \sim (Z/2)^2$ eV/ion $\approx 5Z^2$ kcal/mol, which is a typical value of adsorption energy at the liquid surface. The desorption energy $U_D = U_{AD} - U_L$, where U_L is the adsorption activation energy determined by the energy of lateral Coulomb interaction of adsorbed ions located at adsorption centers. This energy should be set equal to $U_L = Z_0(Ze)^2/[(\varepsilon/2)d]$, where Z_0 is the coordination number of such centers and d is the spacing between the nearest centers. The fact that U_L is doubled as compared to the conventional energy of Coulomb interaction (term $(\varepsilon/2)$ appears instead of ε) is due to the location of interacting charges (ions) on the liquid surface proper.[3] The values of Z_0 and d are determined here from the condition that the packing of molecules of liquid on the surface of a macroparticle is close to dense; consequently, the molecules form a two-dimensional hexagonal lattice. Naturally, this condition idealizes the surface of the macroparticle itself, presuming that it has no microscopic structure. According to such concepts, uniform distribution of adsorption centers with the maximal surface density must correspond to their arrangement at sites (craters) of such a lattice with spacing $d = 2\sqrt{3}\,\delta_l$ between the centers. This value of d is the minimal distance for which each of the three liquid molecules in a tetrahedron can interact only with a single ion, thus ensuring the maximal energy U_{AD}. Each adsorbed ion (adsorption center) in this case is at the center of its own Wigner–Seitz cell with the minimal possible area $S_{WS} = (\sqrt{3}/2)d^2 = 6\sqrt{3}\,\delta_l^2$. For a hexagonal lattice, $Z_0 = 6$ and, hence, energy $U_L = 2\sqrt{3}\,Z^2e^2/(\varepsilon d_l)$; for the desorption energy, we obtain

$$U_D = U_{AD} - U_L = U_D^0\left(1 - \frac{4A}{\sqrt{3}}\alpha_0^4\right), \quad (5)$$

where $U_D^0 = [(3/2)Z^2e^2/(\varepsilon\delta_l)](1/A\alpha_0^4)$, $A = \delta_l^3/\varepsilon\beta$, and $\alpha_0 = (1 + \delta_i/\delta_l)$. The necessary condition for adsorption is $U_D > 0$ (i.e., a potential barrier must exist). Adsorption is stable if the ratio (U_D/T) is quite large. Form condition $U_D > 0$, we obtain the condition that must be imposed on the ratio δ_i/δ_l:

$$\frac{\delta_i}{\delta_l} < \left(\frac{\sqrt{3}}{4A}\right)^{1/4} - 1. \quad (6)$$

Thus, in accordance with the model we are dealing with, only dissolved ions whose radii δ_i satisfy condition (6) are capable of adsorption in principle. The

[1] It should be noted that all numerical estimates will henceforth be referred to aqueous solutions at room temperature without special stipulations.

[2] In the rigid sphere model, such an arrangement of an ion corresponds to its location in a "crater" of the surface layer of spherical molecules.

[3] Here, we use the well-known result from electrostatics [15], according to which the energy of interaction between two charges q located on the plane surface of half-spaces having permittivities ε and ε_p and separated by distance d is $2(q^2/d)/(\varepsilon + \varepsilon_p)$. The applicability of this result to the case in question is ensured by the general requirement $R \gg d$ of the model and the assumption that $\varepsilon_p \ll \varepsilon$.

more stringent this condition, the higher the energy U_D and, accordingly, the more stable the adsorption. It should be noted that condition (6) does not contain the valence Z of ions; only energy U_D^0 is a function of valence. For aqueous solutions, $\delta_l = 1.23$ Å, $\beta = 1.47 \times 10^{-24}$ cm^3, $\varepsilon = 81$, parameter $A = 1/64$, and condition (6) assumes the form

$$\delta_i/\delta_l < 1.29; \tag{6a}$$

i.e., only ions with $\delta_i < 1.29\delta_l = 1.59$ Å are capable of adsorption in principle. For example, in salt solutions of NaCl (Na^+ and Cl^- ions) and KCl (K^+ and Cl^- ions), only cations Na^+ and K^+ are capable of adsorption since the anion radius $\delta_{Cl^-} = 1.81$ Å, while $\delta_{Na^+} = 0.98$ Å and $\delta_{K^+} = 1.33$ Å. In accordance with relation (5), energy U_D for these cations is equal to 0.84 and 0.24 eV, respectively. At room temperatures, this corresponds to a ratio $(U_D/T) \gg 1$; i.e., adsorption of these cations in water is always stable, and charge Q_0 emerging on a macroparticle is always positive.

It would be interesting to analyze the possibility of adsorption of intrinsic H^+ and OH^- ions. The ion radii $\delta_{H^+} = 0.62$ Å and $\delta_{OH^-} = 1.53$ Å,[4] i.e., not only H^+ ions, but also OH^- ions can be adsorbed in principle. However, the ratio $(U_D/T) \gg 1$ for H^+ ions, while for OH^- ions, in accordance with relation (5), ratio $((U_D/T) \approx 1.8$; i.e., the adsorption of these ions is unstable, and they cannot compete with H^+ ions in relation of adsorption. This means that macroparticles in pure water must be positively charged. It should be noted that the conclusions concerning the specific examples considered by us here coincide with the conclusions given in [11].

The assumptions concerning the arrangement of ion adsorption centers make it possible to determine immediately the surface density γ_i^{AD} of stably adsorbed ions: since Wigner–Seitz cells densely cover the entire surface of a macroparticle, we have $\gamma_i^{AD} = 1/S_{WS}$, where $S_{WS} = 6\sqrt{3}\delta_l^2$ is the area of a cell. Thus,

$$\gamma_i^{AD} = 1/6\sqrt{3}\delta_l^2. \tag{7}$$

It should be emphasized once again that this result corresponds only to stable adsorption, for which $(U_D/T) \gg 1$, and is hence independent of the solution temperature and of the valence Z of ions. In other words, if ions capable of stable adsorption at a given temperature are dissolved in a liquid with the molecular radius δ_l, the surface density of adsorption in this case is determined by formula (7) irrespective of the ion density n_{i0}. The latter statement is valid in all cases when dissolved ions are intrinsic ions of the liquid; consequently, their density is determined by temperature T (and by the energy of electrolytic dissociation of liquid molecules) and does not depend on their possible adsorption at particles. If, however, ions are impurities (the density of intrinsic ions is negligibly small), only their total number per unit volume is conserved during adsorption at particles and, hence, the density of free ions decreases. The above assumptions are valid only provided that the density of particles in the solution is $n_p \ll n_{i0}/(4\pi R^2\gamma_i^{AD}) \sim n_{i0}(\delta_l/R)^2$, i.e., provided that adsorption does not affect the density n_{i0} of dissolved ions. It should also be noted that formula (7) is not valid for gas bubbles (see Section 5). For aqueous solutions, $\delta_l = 1.23$ Å and, hence, $\gamma_i^{AD} = 6.4 \times 10^{14}$ cm^{-2}, and the general requirement $R \gg 2\sqrt{3}\delta_l$ of the model leads to the universal limitation $R \geq 4$ nm on admissible sizes of macroparticles (since $2\sqrt{3}\delta_l = 0.43$ nm). The above value of γ_i^{AD} is typical of most molecular liquids. The only exception is organic liquids with "giant" molecules for which the value of γ_i^{AD} can be considerably smaller.

The results of this section confirm its title: if a solution contains ions capable of adsorption at the surface of a macroparticle, i.e., satisfying conditions (6) and $(U_D/T) \gg 1$, these ions always charge these particles strongly ($B \gg 1$) in accordance with relation (4) irrespective of the concentration of ions themselves (the latter quantity determine only the stabilization time of adsorption).

3. SCREENING OF STRONGLY CHARGED PARTICLES

Quasineutrality condition (1), which must be observed for arbitrary values of B, implies that the ionic charge $Q(r)$ contained within a sphere of radius $r > R$ (see relation (3)) tends to zero with increasing $(r - R)$. This leads to the situation when the value of $Ze|Q(r)|/\varepsilon rT$ becomes smaller than unity for large values of $(r - R)$ and, hence, the density $\delta(r)$ of the screening ionic charge for such values of $(r - R)$ must have the Debye form; i.e., $(C/r)\exp[-\kappa(r - R)]$, where κ is defined by formula (2a)). On the other hand, it was shown in the Introduction that, in accordance with numerical counting [14], the decrease in $\delta(r)$ from the

[4] The values of δ_{H^+} and δ_{OH^-} were determined by us on the basis of experimental data on hydration energies ΔG of H^+ and OH^- ions and using the Born model of hydration. In accordance with this model, $\Delta G_i = (Z^2e^2/2\delta_i)(1 - 1/\varepsilon)$. For H^+ ions, $\Delta G_{H^+} = 264$ kcal/mole [16]. In water, $1/\varepsilon \ll 1$ and, hence, we can assume that $\delta_{H^+} = e^2/2\Delta G_{H^+} = 0.62$ Å for the effective radius of H^+ ions in water. The hydration energy of OH^- ions in water is $\Delta G_{OH^-} = 107$ kcal/mole. Consequently, $\delta_{OH^-} = \delta_{H^+}/\Delta G_{OH^-} = 2.46\delta_{H^+} = 1.53$ Å.

maximal value $\delta(R)$ for small values of $(r - R)$ must be much more rapid, but smooth. On the basis of these two features of the behavior of density $\delta(r)$, our construction of the model of screening of strongly charged macroparticles stems from the representation of $\delta(r)$ in the entire range of values of $(r - R)$ in the form

$$\delta(r) = C_1\frac{e^{-b(r-R)}}{r} + C_2\frac{e^{-\kappa(r-R)}}{r}, \quad r \geq R, \tag{8}$$

where C_1 and C_2 are the sought parameters independent of r. It was mentioned above that we will consider only dilute solutions for which $\zeta \equiv \kappa R \ll 1$. It will be clear from subsequent analysis that parameter $\xi \equiv bR$ is on the order of or greater than unity in accordance with what has been said concerning the behavior of $\delta(r)$ for small $(r - R)$; consequently, $\zeta/\xi \ll 1$ in all cases.

The substitution of Eq. (8) into relation (1) and the inclusion of the boundary condition $C_1 + C_2 = R\delta(R)$ gives a system of two equations in C_1 and C_2. Its solution, accurate to terms on the order of $\sim\zeta^2$, has the form

$$C_1 = -\frac{Q_0 F}{4\pi R^2}\left(1 + \frac{f}{F}\zeta^2\right), \quad C_2 = \frac{Q_0 f}{4\pi R^2}\zeta^2, \tag{9}$$

where

$$F = -\frac{4\pi R^3\delta(R)}{Q_0}, \quad f = F\frac{\xi + 1}{\xi^2} - 1. \tag{9a}$$

For strongly charged particles, the density of charge $\delta(R)$ on the surface of a particle must be completely determined by the density $n_i(R)$ of counterions having the charge $(-Ze)(\operatorname{sgn}\delta(R) = -\operatorname{sgn}Q_0)$. Consequently, $\delta(R) = -Zen_i(R)$, and parameter $F > 0$ is defined by the formula

$$F = \frac{4\pi R^3 Zen_i(R)}{Q_0}. \tag{10}$$

Equations (9) and (10) contain two as yet unknown parameters, viz., the number density $n_i(R)$ of counterions at the surface $n_i(R)$ and $\xi = bR$.

In order to determine $n_i(R)$, we use the Boltzmann equation

$$n_i(R) = n_{i0}\exp\left[\frac{Ze\varphi(R)}{T}\right]. \tag{11}$$

Here, $\varphi(R)$ is the potential of the electric field produced by the charge density distribution $\delta(r)$ in form (8). The number density $n_i(R)$ of counterions defined in this way (and, hence, potential $\varphi(R)$) increases indefinitely with B. It is this circumstance that mainly stimulated other authors to develop the concept of counterion condensation at the surface of a particle, leading to the formation of a two-dimensional liquid mentioned in the Introduction. Remaining in the framework of our model, we proceed from the fact that the counterion density $n_i(R)$ for $B \gg 1$ attains its maximal but finite value $n_i^{\max}(R)$ that depends only on the surface density γ_i^{AD} of adsorbed ions. In this case, Eq. (11) should be regarded as the equation determining the value of $\varphi(R)$ which, in turn, depends on ξ, C_1, and C_2 in accordance with Eq. (8) and, ultimately, only on ξ and $n_i^{\max}(R)$ in accordance with Eqs. (9) and (10). As a result, parameter ξ can be defined in terms of $n_i^{\max}(R)$ and other given parameters of the problem.

First of all, we determine the dependence of $\varphi(R)$ on ξ and $n_i^{\max}(R)$. The solution of the Poisson equation $\nabla^2\varphi = -(4\pi/\varepsilon)\delta(r)$ for an arbitrary distribution $\delta(r)$ satisfying quasineutrality condition (1) with the boundary conditions $\varphi(\infty) = 0$, $\varphi'(R) = -Q_0/\varepsilon R^2$ can be represented in the form

$$\varphi(r) = -\frac{4\pi}{\varepsilon r}\int_r^\infty (x^2 - xr)\delta(x)dx. \tag{12}$$

Substituting expression (8) into this formula, integrating, setting $r = R$, and taking into account expressions (9), we obtain

$$\varphi(R) = \frac{Q_0}{\varepsilon R}\left(1 - \frac{F_0}{\xi}\right), \tag{12a}$$

where, in accordance with Eq. (10),

$$F_0 = \frac{4\pi R^3 Zen_i^{\max}(R)}{Q_0} = \frac{Rn_i^{\max}(R)}{\gamma_i^{AD}}. \tag{13}$$

Substituting now $n_i(R) = n_i^{\max}(R)$ into the left-hand side of Eq. (11) and expression (12a) into its right-hand side, we obtain the following relation between ξ and $n_i^{\max}(R)$:

$$\xi = \frac{F_0}{1 - \alpha} = \frac{R}{1 - \alpha}\frac{n_i^{\max}(R)}{\gamma_i^{AD}}, \tag{14}$$

where $\alpha = L/B$, $L = \ln[n_i^{\max}(R)/n_{i0}]$. In this case, quantity f (see Eq. (9a)) assumes the form

$$f = \frac{1 - \alpha}{\xi} - \alpha. \tag{14a}$$

It remains for us to determine the value of $n_i^{\max}(R)$. Counterions are not capable of adsorption (they do not satisfy condition (6), and if this condition is satisfied,

$U_D/T \le 1$ for these ions), but being near the surface of a macroparticle, they are attracted to it by the Coulomb forces exerted by adsorbed ions. The latter are located at the centers of their Wigner–Seitz cells covering the entire surface of the particle. The maximal surface density $\Gamma_i^{\max}$ of counterions obviously corresponds to their arrangement at the same distance $2\delta_l$ from three adsorbed ions, each of which mainly interacts with only one counterion (Fig. 1). It is this arrangement that corresponds to a stable state. Since each counterion in this case corresponds to three individual Wigner–Seitz cells, we have $\Gamma_i^{\max} = 1/3S_{WS} = 1/8\sqrt{3}\delta_l^2$. Assuming further that $\Gamma_i^{\max} = [n_i^{\max}(R)]^{2/3}$, we obtain

$$n_i^{\max}(R) = \frac{\delta_l^{-3}}{162(4/3)^{1/4}},$$
$$\frac{n_i^{\max}(R)}{\gamma_i^{AD}} = \frac{\delta_l^{-1}}{9(12)^{1/4}}. \qquad (15)$$

For aqueous solutions, the number density $n_i^{\max}(R) = 3.1 \times 10^{21}$ cm^{-3}.

It can be seen from Eq. (8) that distribution $\delta(r)$ must vanish at a certain point $r = r_0$ if quantities C_1 and C_2 have opposite signs. In accordance with relations (9), this is observed for $f > 0$ or, in accordance with Eq. (14a), for $\xi < (1 - \alpha)/\alpha$. Introducing a new variable,

$$\lambda = \frac{R/\delta_l}{18(12)^{1/4}} \equiv \frac{F_0}{2}, \qquad (16)$$

we have

$$\xi = \frac{2\lambda}{1-\alpha}; \qquad (16a)$$

consequently, condition $f > 0$ assumes the form

$$\frac{\alpha}{(1-\alpha)^2} < \frac{1}{2\lambda}. \qquad (17)$$

When this inequality holds, we can easily obtain the following expression for the position of point r_0:

$$r_0 = R\left[1 + \frac{1}{\xi}\ln\frac{F_0}{f\zeta^2}\right]$$
$$\equiv R\left\{1 + \frac{1}{\xi}\ln\frac{(\xi/\zeta)^2}{\left[1 - 2\lambda\dfrac{\alpha}{(1-\alpha)^2}\right]}\right\}. \qquad (18)$$

At this point, the charge density $\delta(r)$ of the ionic shell changes its sign: $\operatorname{sgn}\delta(r < r_0) = -\operatorname{sgn}\delta(r > r_0)$ (Fig. 2).

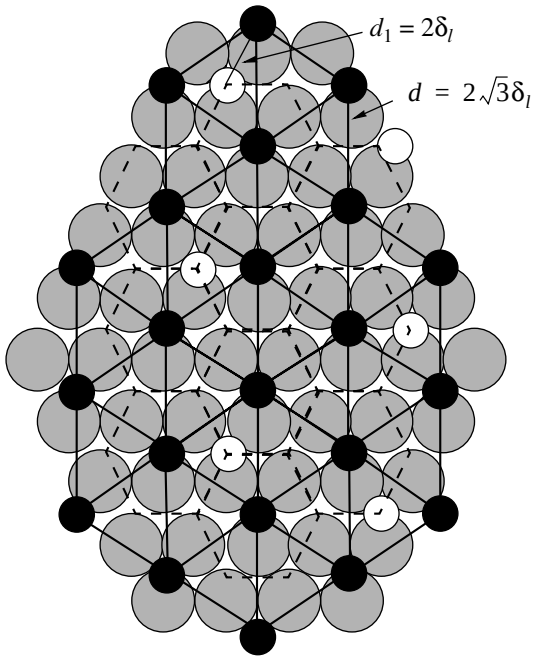


Fig. 1. Ion adsorption centers (dark circles) form a two-dimensional hexagonal lattice with spacing $d = 2\sqrt{3}\,\delta_l$ between centers (δ_l is the radius of liquid molecules depicted by large circles). Each center (adsorbed ion) is at the center of its Wigner–Seitz cell, having an area of $S_{WS} = 6\sqrt{3}\,\delta_l^2$ (hatched circles). The density of adsorbed ions is $\gamma_i^{AD} = 1/S_{WS}$. Small light circles denote the positions of surface counterions, each of which corresponds to three individual Wigner–Seitz cells. Consequently, $\Gamma_i = \gamma_i^{AD}/3$. The distance between each of the three adsorbed ions and the surface counterion is $d_1 = 2\delta_l$.

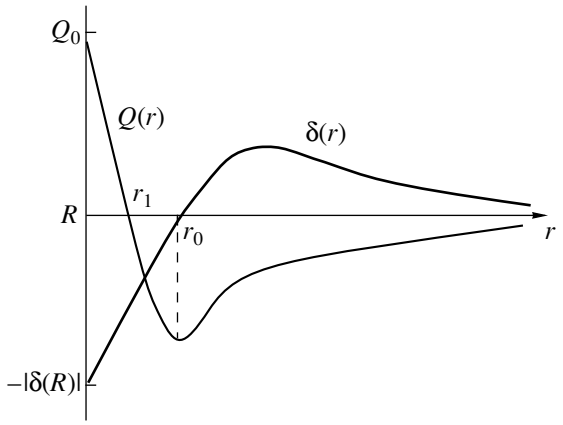


Fig. 2. The charge density in the ionic shell $\delta(r)$; $\delta(R) = -eZn_i(R)$ is the charge density on the surface ($n_i(R)$ is the density of counterions at the surface); Q_0 is the surface charge; $Q(r) = Q_0 + 4\pi\int_R^r x^2\delta(x)dx$, is the charge in a sphere of radius $r > R$.

This must also lead to sign inversion of the total charge $Q(r)$, occurring at a certain value of radius $r_1 \neq r_0$. Let us determine r_1. On the basis of Eqs. (1), (3), (8), and (9), we have

$$Q(r) = -4\pi\int_r^\infty x^2\delta(x)dx = Q_0\left\{\frac{F_0}{\xi^2}(1+br)e^{-b(r-R)} - f(1+\kappa r)e^{-\kappa(r-R)}\right\}, \quad (19)$$

$$\frac{dQ}{dr} = 4\pi r^2\delta(r). \quad (19a)$$

It follows from relation (19) that the necessary condition for $Q(r)$ vanishing is, as before, the inequality $f > 0$, i.e., inequality (17). In this case, the equation for $y_1 \equiv r_1/R$ has the form

$$y_1 = 1 + \frac{1}{\xi}\ln\frac{1+\xi y_1}{1-\dfrac{\alpha\xi}{1-\alpha}} \equiv \frac{r_0}{R} - \frac{1}{\xi}\ln\frac{(\xi/\zeta)^2}{1+\xi y_1}. \quad (20)$$

Since $(\xi/\zeta)^2 \gg 1$, it follows from expression (20) that radius $r_1 < r_0$; i.e., the sign inversion of the total charge $Q(r)$ ($\operatorname{sgn}Q(r < r_1) = \operatorname{sgn}Q_0 = -\operatorname{sgn}Q(r > r_1)$) occurs at such distances from the particle surface where the sign of the charge density remains unchanged ($\operatorname{sgn}\delta(r_1) = \operatorname{sgn}\delta(R_0) = -\operatorname{sgn}Q_0$) (see Fig. 2). In accordance with formula (19a), charge $Q(r)$ attains its extreme value at point $r = r_0$. On the basis of Eqs. (19) and (18), we obtain the following relation (accurate to terms on the order of $(\zeta/\xi)^2$) for this value of $Q(r_0)$:

$$Q(r_0) = -Q_0 f(1 + \xi r_0/R). \quad (21)$$

As the value of r increases further, charge $Q(r)$ tends to zero.

Let us now consider inequality (17), viz., the condition for the existence of the charge inversion effect in dilute solutions ($\zeta \ll 1$). This condition can be written in the form

$$B\alpha_1 > L, \quad (22)$$

where α_1 is the smaller of two positive roots of the equation

$$\alpha^2 - 2(1+\lambda)\alpha + 1 = 0,$$

i.e.,

$$\alpha_1 = 1 + \lambda(1 - \sqrt{1 + 2/\lambda}). \quad (22a)$$

Substituting this value into inequality (22) and considering that, in accordance with relations (2a), (4), (7), (15), and (16), parameters B and L can be represented in the form

$$B = \frac{24\pi}{(12)^{1/4}}\left(\frac{l_B}{\bar{\delta}_l}\right)Z^2\lambda,$$

$$L \equiv \ln\left[\frac{n_i^{\max}(R)}{n_{i0}}\right] = \ln\left(\frac{l_B}{\bar{\delta}_l}\right) + 2\ln\lambda + 2\ln\left(\frac{Z}{\bar{\zeta}}\right) + 5,$$

we arrive at condition

$$\frac{24\pi}{(12)^{1/4}}Z^2\left(\frac{l_B}{\bar{\delta}_l}\right)\Psi(\lambda) > \ln\left(\frac{l_B}{\bar{\delta}_l}\right) + 2\ln\lambda + 2\ln\left(\frac{Z}{\bar{\zeta}}\right) + 5, \quad (23)$$

where $\Psi(\lambda) = \lambda[1 + \lambda(1 - \sqrt{1 + 2/\lambda})]$ is a monotonically increasing function of λ. Let us analyze the fulfillment of this inequality for aqueous solutions ($l_B = 7$ Å, $\delta_l = 1.23$ Å), for which $l_B/\delta_l = 5.7$ and λ is defined, in accordance with Eq. (16), by the formula $\lambda = \bar{R}/4$ with an accuracy admissible for our model. It follows hence that the range of possible values of λ is determined by the range of values of $\bar{R}$ permissible in the model. The minimal value $\bar{R}_{\min}$ is determined by the minimal macroparticle radius $R = 4$ nm for an aqueous solution, which was estimated by us earlier. Hence, we can assume that $\bar{R}_{\min} = 4$ and, accordingly, $\lambda_{\min} = 1$. The maximal value $\bar{R}_{\max}$ must correspond to the condition $\zeta = \kappa R \ll 1$ for dilute solutions. Setting $\zeta = 0.1$, we obtain, in accordance with Eq. (2a), $\bar{R}_{\max} \approx (8/Z) \times 10^8/n_{i0}^{1/2}$ [cm^{-3}]. Since the value of $\bar{R}_{\max}$ decreases upon an increase in density n_{i0}, while $\bar{R}_{\min}$ is independent of density, inequality $n_{i0} < n_{i0}^{\max}$ must be satisfied in the framework of the model, where $n_{i0}^{\max}$ is determined from the condition $\bar{R}_{\max} = \bar{R}_{\min}$. These formulas lead to $n_{i0}^{\max} = (4/Z^2) \times 10^{16}$ cm^{-3} $= (7/Z^2) \times 10^{-5}$ M. Thus, the above results are valid for aqueous solutions only for ionic densities $n_{i0} < 4 \times 10^{16}$ cm^{-3} $= 7 \times 10^{-5}$ M (and only for $Z = 1$). If we are dealing with pure water with pH = 7, and the density of adsorbed H^+ ions is $n_{i0} = 6 \times 10^{13}$ cm^{-3} $= 10^{-7}$ M, we have $\bar{R}_{\max} \approx 10^2$ and, accordingly, $\lambda_{\max} = 25$. It can easily be verified that inequality (23) holds for all aqueous solutions satisfying the condition of their being dilute ($n_{i0} < n_{i0}^{\max}$), i.e., in the entire range of possible values of λ ($1 < \lambda < 25$) irrespective of valence Z.

Let us consider quantitative estimates for a particle of radius $R = 4$ nm ($\bar{R} = 4$, $\lambda = 1$ in a 1 : 1 aqueous solution (we assume that $\zeta = 0.1$ and, accordingly, $n_{i0} = n_{i0}^{\max} = 4 \times 10^{16}$ cm^{-3} $= 7 \times 10^{-5}$ M). For such a particle, the adsorbed charge $Q_0 = 1300e$, $B = 224$, $L = 11.2$, $\alpha = L/B = 0.05$, $\xi = 2$, $\alpha_1 = 0.27$, and $B\alpha_1 = 60 > L = 11.2$ (i.e., condition (22) is satisfied). Radii r_0 and r_1 are equal respectively to 16 and 7 nm, while the maximum (absolute) value of inverted charge $Q(r_0) = -900e$.

4. EFFECT OF BROWNIAN MOTION OF MACROPARTICLES ON THEIR SCREENING

While considering the screening of macroparticles by the ionic shell, we tacitly assumed that a particle in a liquid solution is at rest and, hence, its ionic shell is not subjected to hydrodynamic action (such a shell will be referred to as an equilibrium shell). As a matter of fact, macroparticles perform thermal Brownian movement, and such an action is inevitable. In this case, a compound particle with total charge Q_c, which consists of the initial macroparticle of radius R and surface charge Q_0 and a certain layer of the liquid with the "frozen-in" ionic shell distorted by motion, is involved in motion. If the velocity of the particle is small, we can assume that the distortion of the ionic shell under steady-state conditions boils down to washing out (vanishing) of the equilibrium charge density distribution $\delta(r)$ due to viscous forces of a peripheral layer $r > a_c$, while the remaining part of this distribution ($r < a_c$) is spherically symmetric as before. Such a pattern is widely used in theoretical treatment of electrophoresis (see, for example, [17]); in this case, the radius a_c of a compound particle is usually referred to as the hydrodynamic radius of a colloidal particle and the spherical surface with such a radius is called the glide surface. At the same time, there are no publications (to our knowledge) in which this concept is used for clarifying the role of thermal motion of a particle. Our task is to establish the relation between parameters Q_c and a_c of a compound particle and the initial parameters Q_0 and R of the macroparticle and of the solution itself. As before, we assume that the solution is dilute ($\zeta \ll 1$).

The physical foundation for solving the formulated problem is that the pressure of viscous forces destroying the ionic shell during its motion is always opposed by the electrostriction-induced pressure p_{str} compressing the compound particle and associated with its charge Q_c, which coincides in the given model with the charge $Q(r)$ considered above for $r = a_c$ (see Eq. (3)), i.e.,

$$Q_c = Q_0 + 4\pi\int_R^{a_c} x^2\delta(x)dx = -4\pi\int_{a_c}^{\infty} x^2\delta(x)dx, \quad (24)$$

and $\delta(r)$ is the equilibrium charge density distribution defined in the previous section. The electric field strength at the surface of a compound particle ($r = a_c$) is $E_c = Q_c/\varepsilon a_c^2$, while the pressure is given by

$$p_{str} = \frac{\varepsilon E_c^2}{8\pi} = \frac{Q_c^2}{8\pi\varepsilon a_c^4}. \quad (25)$$

The steady-state (stationary) state corresponds to the equality of the spherically symmetric pressure p_{str} to the maximal absolute value of the negative pressure of viscous forces (i.e., the pressure of detachment from the rear side relative to motion), which sets in for a certain value of radius a_c of the glide surface. Under the laminar flow conditions for a particle with velocity u, the maximal pressure of detachment is given by [18]

$$p_v = \frac{3\eta u}{2a_c} \quad (26)$$

(η is the viscosity of the liquid); consequently, for the given velocity u, parameters Q_c and a_c can be defined, in compliance with our model, from the equation

$$\frac{3\eta u}{2a_c} = \frac{Q_c^2}{8\pi\varepsilon a_c^4} \quad (26a)$$

and Eq. (24).

In thermal motion of particles, their velocity $u = u(t)$ (here and below, u stands for the velocity component along a certain fixed direction **s**) is a steady-state random process with the mean value $\langle u\rangle = 0$ and the spectral intensity

$$g(\omega) = \frac{2\tau_0}{\pi}\frac{T/m_c}{1 + (\omega\tau_0)^2}, \quad \int_0^{\infty} g(\omega)d\omega = \langle u^2\rangle = \frac{T}{m_c}. \quad (27)$$

Here, m_c is the mass of a compound particle, which can be set equal to $(4\pi/3)\rho a_c^3$ (if the density of the macroparticle itself is close to the density ρ of the liquid or if $R \ll a_c$), and $\tau_0 = m_c/5\pi\eta a_c$ is the correlation time of process $u(t)$.[5] Obviously, it would be incorrect to substitute the quantity $\langle u^2\rangle^{1/2} = (T/m_c)^{1/2}$ for u into Eq. (26). As a matter of fact, the motion of a Brownian compound particle consists of fast "vibrations" with amplitudes smaller than its radius a_c (the relative amplitude of such vibrations is $\langle u^2\rangle^{1/2}(\tau_0/a_c) \sim 0.1(T/\rho\nu^2 a_c)^{1/2} \sim 10^{-2}/a_c^{1/2}$ [nm] $\ll 1$, $\nu = \eta/\rho \sim 10^{-2}$ cm^2/s) and "smooth" displacements with amplitudes comparable with or greater than a_c. For our problem, the second type of

[5] While defining τ_0, we use for the Stokes frictional force the formula $F = 5\pi\eta a_c u$ corresponding to the motion of a liquid drop in a liquid with the same viscosity η.

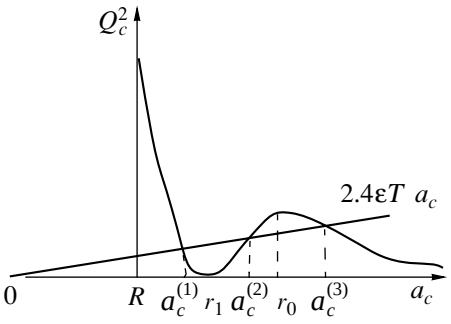


Fig. 3. a_c is the radius of a compound particle and $Q_c \equiv Q(a_c)$ is its charge.

motion is of interest since only such movements of a compound particle may lead to viscous washing out of the peripheral layer of the ionic shell. The smooth velocity process $u_1(t)$ of interest to us corresponds to the low-frequency part of spectrum $g(\omega)$ in the interval $(0, \tau_1^{-1})$, where $\tau_1 \gg \tau_0$, and can be represented as process $u(t)$ averaged over interval τ_1:

$$u_1(t) = \frac{1}{\tau_1}\int_{-\infty}^{t} e^{-(t-t')/\tau_1} u(t')dt', \quad \tau_1 \gg \tau_0. \tag{28}$$

The spectral intensity $g(\omega)$ of this process is given by

$$g_1(\omega) = \frac{g(\omega)}{1 + (\omega\tau_1)^2},$$
$$\int_0^\infty g_1(\omega)d\omega = \langle u_1^2\rangle = \langle u^2\rangle\frac{\tau_0}{\tau_1}, \tag{28a}$$

and $u(t)$ for the complete process can be written in the form of the sum

$$g(\omega) = g_1(\omega) + g_2(\omega), \tag{28b}$$
$$g_2(\omega) = g(\omega)\frac{(\omega\tau_1)^2}{1 + (\omega\tau_1)^2}.$$

Spectrum $g_2(\omega)$ vanishes for $\omega = 0$ and has a width of $\Delta\omega_2 \approx \tau_0^{-1}$. It is responsible only for small vibrations of the particles mentioned above and makes zero contribution even to their diffusive movements (since the diffusion coefficient $D_2 = \pi g_2(0) = 0$). Smooth displacement of particles are controlled by spectrum $g_1(\omega)$, which, in accordance with Eqs. (28a), has a width of $\Delta\omega_1 \approx \tau_1^{-1}$; accordingly, the correlation time of process $u_1(t)$ is equal to τ_1. Washing out of the peripheral layer is the most effective on intervals of unidirectional motion of a particle, i.e., in the intervals between successive extrema of process $u_1(t)$. It is well known (see, for example, [19]) that the mean value of such intervals coincides in order of magnitude with the correlation time of a random process, i.e., with time τ_1 in our case. During this time, the particle is displaced on the average by $\langle \Delta S^2\rangle^{1/2} = \langle \Delta u_1^2\rangle^{1/2}\tau_1$. We use the condition $\langle \Delta S^2\rangle^{1/2} = a_c$ to determine the value of interval τ_1 (spectral width $\Delta\omega_1$) for which the washing out is the most effective: $\tau_1 = a_c/\langle u_1^2\rangle^{1/2}$. However, $\tau_1 = \langle u^2\rangle\tau_0/\langle u_1^2\rangle$ in accordance with Eqs. (28a). This leads to the expression $\langle u_1^2\rangle^{1/2} = \langle u^2\rangle\tau_0/a_c = T/5\pi\eta a_c^2$, which does not depend on mass m_c, required for the substitution into formula (24). The substitution of this expression for $\langle u_1^2\rangle^{1/2}$ into Eq. (26a) gives

$$\frac{Q_c^2}{2\varepsilon a_c} = 0.8\left(\frac{3T}{2}\right) = 0.8 \text{ K}, \tag{29}$$

where $K = (3/2)T$ is the average thermal energy of the particle; the left-hand side of this formula defines the energy of the electric field created by the compound particle. It is important to note that the expression establishing the relation between a_c and Q_c was derived on the basis of Eqs. (25), (26), and (27), i.e., universal results of macroscopic electrodynamics, hydrodynamics, and statistical physics (the theory of Brownian motion). Thus, it is not associated in any way with our model (including the condition $\kappa R \ll 1$) and, being independent of viscosity η, is of the thermodynamic nature: in the steady state, the energy of the electric field of a compound particle for an arbitrary value of the adsorbed charge Q_0 is a function of the temperature of the medium only (and is close to its mean kinetic energy).

It should also be noted that this result is not associated in any way with the sign inversion effect and is also valid for weakly charged macroparticles, for which $B \ll 1$ and the Debye screening (2) takes place. The charge of a compound particle in this case is given by

$$Q_c = Q_0\frac{1 + \kappa a_c}{1 + \kappa R}e^{-\kappa(a_c - R)}. \tag{30}$$

Here, for all values of a_c, charge Q_c has the same sign as charge Q_0 and tends monotonically to zero with increasing a_c. As a result, the system of equations (29) and (30) has only one pair of roots (Q_c, a_c) for all possible temperatures, which corresponds to a single type of compound particles. We will not dwell on detailed analysis of this exotic case (from the viewpoint of our model) and we return to strongly charged particles. In accordance with Eqs. (19) and (24), the quantity Q_c^2 as a function of a_c has a two-hump shape (Fig. 3), and the system of equations (24) and (29) for $T < T^*$ (T^* will

be defined below) has formally three pairs of roots $(Q_c^{(i)}, a_c^{(i)})$ $(i = 1, 2, 3)$, the relation between Q_c and a_c in each pair being

$$(Q_c^{(i)}/e)^2 = 2.4(a_c^{(i)}/l_B) \quad (i = 1, 2, 3). \tag{31}$$

It can easily be seen, however, that only two pairs of roots $(Q_c^{(1)}, a_c^{(1)})$ and $(Q_c^{(3)}, a_c^{(3)})$ correspond to stable states of compound particles; i.e., only two types of compound particles can exist and have such parameters. Parameters $(Q_c^{(2)}, a_c^{(2)})$ are found to be impracticable. This follows from the fact that, for radii $a_c < a_c^{(2)}$ (but for $a_c > a_c^{(1)}$), pressure $p_{str} < p_v$ (see Fig. 3); consequently, the particle size must continue to decrease in the course of Brownian motion (ionic shell is washed out) down to value $a_c^{(1)}$. For $a_c > a_c^{(2)}$ (but $a_c < a_c^{(3)}$), pressure $p_{str} > p_v$ and, hence, motion does not hamper the growth of the ionic shell of a compound particle up to value $a_c^{(3)}$. At the same time, radii $a_c^{(1)}$ and $a_c^{(3)}$ correspond to stable states since the condition $\mathrm{sgn}(a_c - a_c^{(1,3)}) = -\mathrm{sgn}(p_{str} - p_v)$ holds for small deviations $(a_c - a_c^{(1,3)})$.

Thus, compound particles can be divided into two types: "small" particles with radii $a_c^{sm} \equiv a_c^{(1)}$ and charges $(Q_c^{sm}/e) = (2.4a_c^{sm}/l_B)^{1/2}$ $(\mathrm{sgn}\, Q_c^{sm} = \mathrm{sgn}\, Q_0)$ and "coarse" particles with radii $a_c^l \equiv a_c^{(3)}$ and charges $-Q_c^l/e = (2.4a_c^l/l_B)^{1/2}$ $(\mathrm{sgn}\, Q_c^l = -\mathrm{sgn}\, Q_0$; it should be recalled that $Q_0 > 0$ by hypothesis). Such a division of compound particles into two types with opposite charge polarities is a peculiar manifestation of the charge inversion effect; its presence follows from independent data on parameters of compound particles.

In accordance with Eqs. (21) and (30), temperature T^* is defined by the formula

$$T^* = \frac{Q_0^2 f^2 (1 + \zeta r_0/R)^2}{2.4\varepsilon r_0}. \tag{32}$$

Using formulas (14), (14a), and (18) and taking into account the fact that $\alpha \sim 0.1$ and $\zeta \le 0.1$, we obtain the approximate formula $(T^*/T) \approx 10^2 Z^2 (R l_B \gamma_i^{AD})$. For aqueous solutions (for $Z = 1$), this gives $(T^*/T) \sim 10^2 \bar{R}\gamma_i^{AD}$ $[\mathrm{nm}^{-2}] \ge 10^3$; i.e., for any temperature T, all three pairs of roots $(Q_c^{(i)}, a_c^{(i)})$, i.e., both types of compound particles, must exist. Since $(T/T^*)^{1/2} \ll 1$, for small compound particle we have[6]

$$a_c^{sm} = r_1 \approx R\left(1 + \frac{2}{\bar{R}}\right),$$
$$\frac{Q_c^{sm}}{e} = \sqrt{2.4\frac{R}{l_B}\frac{r_1}{R}} \approx \sqrt{\frac{24}{7}(\bar{R} + 2)}. \tag{33}$$

Parameters $(Q_c^{(i)}, a_c^{(i)})$ for coarse compound particles can be determined from Eq. (29), into which we must substitute only the second term of formula (19) (the first term is exponentially small since $((a_c^b - R)b \gg 1)$. Using formula (32) and the above approximate formulas for (T^*/T), we can easily obtain the following approximate results (we assume that $(r_0/R) \sim 3$, $(T^*/T) \ge 10^3$):

$$\frac{a_c^l}{R} = \frac{1}{2\zeta}\ln\left[\frac{Q_0^2 f^2}{2.4\varepsilon r_0 T}\frac{r_0}{R}\frac{(1 + \zeta a_c^l/R)^2}{a_c^l/R}\right]$$
$$\approx \frac{1}{2\zeta}\ln\left(\frac{3T^*}{T}\right). \tag{34}$$

For $\ln(3T^*/T) \approx 10$ and $\zeta = 0.1$, we obtain

$$a_c^l \approx 50R, \quad \frac{Q_c^l}{e} = -12\bar{R}^{1/2}$$
$$= -\left(\frac{42}{1 + 2/\bar{R}}\right)^{1/2}\left(\frac{Q_c^{sm}}{e}\right). \tag{34a}$$

It is worth noting that parameters a_c^{sm} and Q_c^{sm} do not depend on the adsorbed charge Q_0 to within terms on the order of $(T/T^*)^{1/2}$, while a_c^l and Q_c^l depend on it only logarithmically.

It would be interesting to analyze the electrophoretic mobility of compound particles. This process is linear; i.e., mobility μ_c is independent of the external electric field E only if the field in the solution is $E/\varepsilon \ll E_c \equiv Q_c/\varepsilon a_c^2 = (1/\varepsilon)\sqrt{2.4\varepsilon a_c T}/a_c^2$. In order to satisfy this condition for both types of compound particles, it is necessary that

$$E \ll \sqrt{\frac{2.4\varepsilon T}{(a_c^l)^3}} \approx \sqrt{\frac{2.4\varepsilon T}{(50R)^3}}. \tag{35}$$

[6] It can easily be verified that $a_c^{(1,2)} = r_1[1 \mp O(T/T^*)^{1/2}]$. The approximate expression for $r_1 \approx R(1 + 2/\bar{R})$ follows from Eq. (20), taking into account the fact that $\xi \approx \bar{R}/2$.

It can easily be verified that this condition is equivalent to the condition $u_c \ll \langle u_1^2 \rangle^{1/2} = (T/5\pi\eta a_c^2)$ imposed on the velocity of compound particles for linear electrophoresis and, hence, does not affect the Brownian mechanism of their formation. For aqueous solution, relation (35) leads to the condition $E \ll 7.4 \times 10^4/\bar{R}^{3/2}$ [V/cm], which can be satisfied in traditional experiments on electrophoresis up to $R \sim 10^3$ nm = 1 μm.

The electrophoretic mobility μ_c can be determined from the equality of the ponderomotive force $Q_c(E/\varepsilon)$ acting on a compound particle to the Stokes force $5\pi\eta a_c u_c$. It follows hence that $\mu_c = u_c/E = Q_c/5\pi a_c \eta\varepsilon$ and, in accordance with Eq. (30), we obtain

$$\mu_c = \frac{1}{2\pi\eta}\frac{T}{Q_c}. \tag{36}$$

Mobilities μ_c^{sm} and μ_c^{l} have opposite signs and their ratio is given by

$$\frac{\mu_c^{sm}}{\mu_c^{l}} = \frac{Q_c^{l}}{Q_c^{sm}}. \tag{36a}$$

Thus, the positive mobility of small compound particles (for $Q_0 > 0$) is always higher in absolute value than the negative mobility of coarse compound particles.

Let us now determine the electrical conductivity of a colloidal solution, which is due to the electrophoretic effect in question (molar-ionic conductivity). Since the densities of both types of compound particles are identical and equal to the density n_p of macroparticles themselves, we have the following expression for the current density j_p:

$$j_p = n_p(\mu_c^{sm}Q_c^{sm} + \mu_c^{l}Q_c^{l})E = \sigma_p E.$$

On the basis of Eqs. (36) and (36a), we derive the following expression for the molar-ionic conductivity σ_p:

$$\sigma_p = \frac{n_p}{2\pi\eta}\frac{T}{}. \tag{37}$$

This formula contains the only characteristic of colloidal particles as charge carriers, namely, their number density n_p; conductivity σ_p does not depend on radius R or on the adsorbed charge Q_0. At room temperature and viscosity $\eta \sim 10^{-2}$ g/ cm s = 1 cP, conductivity $\sigma_p \sim 10^{-24} n_p$ cm^{-3} Ω^{-1} cm^{-1}, which corresponds to values of conductivity for nonpolar liquids for $n_p > 10^6$ cm^{-3}. This indicates that the molar-ionic conductivity we are dealing with may be the main conductivity in such liquids; i.e., it can be much higher than the ionic impurity conductivity σ_i (in nonpolar liquids, it usually prevails over the intrinsic ionic conductivity). The latter is defined by the formula

$$\sigma_i = \frac{n_{i0}Z^2e^2(D_{\text{cat}} + D_{\text{an}})}{T}, \tag{38}$$

where D_{cat} and D_{an} are the diffusion coefficients for impurity cations and anions, n_{i0} is their density, Z is their valence, and the conductivity ratio is given by

$$\frac{\sigma_i}{\sigma_p} = \frac{n_{i0}}{n_p}\frac{Z^2e^2\pi\eta(D_{\text{cat}} + D_{\text{an}})}{T^2}. \tag{38a}$$

For typical values of $(D_{\text{cat}} + D_{\text{an}}) \sim 3 \times 10^{-5}$ cm^2/s, the ratio $\sigma_i/\sigma_p \sim (10Z)^2(n_{i0}/n_p)$ and, hence the molar-ionic conductivity prevails when $n_p > (10Z)^2 n_{i0}$. In nonpolar liquids, ionic concentrations are relatively low, and the fulfillment of this condition is quite feasible (especially for $Z = 1$). In addition, owing to small values of n_{i0}, the condition $\kappa R \ll 1$ for dilute solutions holds for a wide range of particle sizes. This in turn extends the limits of applicability for our model so that we can assume on its basis that colloidal solutions in nonpolar liquids mainly have the form of an ensemble of two types of oppositely charged compound particles. Among other things, this conclusion explains the fact that, in order to improve the insulating properties of nonpolar liquids, we must first of all try to purify a liquid from colloidal particles and only at the last stage to reduce the impurity ion density n_{i0}. The purification from particles itself can be carried out through electrophoresis with adsorption of oppositely charged compound particles at the surface of the electrodes.

The construction of the model and results obtained on its basis were applied so far to the case of macroparticles with a preset radius R limited only by two conditions: $R \gg 2\sqrt{3}\,\delta_l$ ($R \geq 4$ nm for aqueous solution) and $\kappa R \ll 1$ (condition for dilute solutions). In the next section, we will consider the problem of screening of gas bubbles in an electrolyte solution. In this case, radius R of the stable macroparticles formed is no longer a preset quantity, but is determined by the parameters of the liquid itself and by the charge of dissolved ions.

5. BUBSTONS

The model of ion adsorption at the surface of macroparticles and their screening by an ionic shell used in the above analysis can be extended to the case of gas bubbles, which makes it possible to consider their stability on a theoretical basis differing from that in [10, 11] (we will henceforth consider 1 : 1 aqueous solutions). The concept of a bubston (bubble stabilized by ions) was introduced by us for the first time in [10]. This concept reflects the possibility of the existence of nanome-

ter-size stable gas bubbles, associated with adsorption of dissolved ions of a definite polarity on their surface, in liquid solutions of electrolytes (in particular, in pure water displaying ionic conductivity due to intrinsic thermal H^+ and OH^- ions). The necessary condition for the stability is the mechanical equilibrium condition, under which the surface tension pressure $2\sigma/R$ (σ is the surface tension; $\sigma = 73$ erg/cm^2 for aqueous solutions under normal conditions) is balanced by the negative (tensile) ponderomotive pressure p_E associated with the presence of the surface electric charge Q_0. We assume that an ionic solution together with equilibrium gas bubbles present in it is in equilibrium with the external gaseous medium (e.g., atmospheric air) under a pressure p_0; accordingly, the concentration c_g of the gas dissolved in it obeys the Henry's law: $c_g = \Gamma(T)p_0$. However, the condition of equilibrium between the dissolved gas and the gas in a bubble under pressure p_b also implies that $c_g = \Gamma(T)p_b$ (with the same value of the Henry coefficient $\Gamma(T)$). It follows hence that, in complete equilibrium, the pressure in the bubbles $p_b = p_0$ (i.e., it is balanced by the external pressure) and, hence, is not reflected in the necessary stability condition $2\sigma/R = p_E$. It is important that pressure p_E depends only on the surface charge Q_0 and not on the entire charge distribution for $r \geq R$ and is equal to $p_E = (1/8\pi\varepsilon)Q_0^2/R^4 = (2\pi/\varepsilon)e^2(\gamma_i^{AD})^2$. This follows from the general definition of pressure $p = (\partial\Phi/\partial V)_T$, where Φ is the free energy of the system under investigation. In our case, the system is a region of the medium with a radius $r \geq R$, in which the electric field differs from zero (the field in the bubble is equal to zero in view of the central symmetry). For this reason, energy $\Phi = (1/2\varepsilon)\int_R^r [Q^2(x)/x^2]dx$, and volume $V = (4\pi/3)(r^3 - R^3)$. The pressure exerted by a bubble on the system (i.e., for $r = R$) is $p(R) = -(1/4\pi R^2)(d\Phi/dR) = (1/8\pi\varepsilon)Q^2(R)/R^4 \equiv (1/8\pi\varepsilon)Q_0^2/R^4$. This is obviously just the pressure p_E stretching the bubble. The same definition implies that the pressure exerted on the outer surface of the region is $p(r) = (1/4\pi r^2)(d\Phi/dr) = (1/8\pi\varepsilon)Q^2(r)/r^4$. For $r = a_c$, this expression coincides with the expression used by us for p_{str}.

The equality $2\sigma/R = p_E$ leads to

$$R = \sigma\varepsilon/[\pi e^2(\gamma_i^{AD})^2]. \tag{39}$$

A question arises: what value of γ_i^{AD} should be substituted into this formula? The value $\gamma_i^{AD} = 6.4 \times 10^{14}$ cm^{-2} used by us earlier in accordance with formula (7) was obtained for macroparticles with preset values of radii $R \geq 4$ nm (these can be solid particles or drops of an alien liquid). The substitution of this value into Eq. (39) gives $R = 2$ Å, which is a physically meaningless result. The reason is that radius R of a bubston is not a preset quantity, but is determined from the self-consistent mechanical equilibrium condition $2\sigma/R = p_E$ for the maximal density of stably adsorbed ions, which satisfies the general condition $R \gg d$ of the model, where d is the spacing between ion adsorption centers on their surface. As applied to gas bubbles, the physical meaning of the condition $R \gg d$ is that the discrete distribution of adsorbed ions at the lattice sites must be such that the tensile pressure p_E associated with these ions must differ from the spherically symmetric pressure ensuring the bubston stability only insignificantly. However, it is necessary that the total number N_i of adsorbed ions be large enough; for $R \gg d$, the number $N_i = 4\pi R^2/S_{WS} = (8\pi/3^{1/2})(R/d)^2 \geq 10^3$ (we assume that $S_{WS} = (3^{1/2}/2)d^2$; see below).

As before, we assume that the surface arrangement of liquid molecules for gas bubbles also forms a densely packed hexagonal lattice; consequently, the ion adsorption energy in this case is defined by the formula $U_{AD} = (3/2)e^2\beta/\delta_0^4$ derived in Section 2 (for $Z = 1$). However, by virtue of condition $R \gg d$, the admissible distances d between adsorption centers in this case exceed the minimal distance $d = d_{\min} = 2 \times 3^{1/2}\delta_l$ (corresponding to density $\gamma_i^{AD} = 6.4 \times 10^{14}$ cm^{-2}; see below). For such a lattice, the complete set of possible distances d is defined by the formula $d = k\delta_l$, where coefficient k can assume the following discrete values (in increasing order):

$$2\sqrt{3},\ 4,\ 6,\ 4\sqrt{3},\ 8,\ 10,\ 6\sqrt{3},\ 12,\ 8\sqrt{3},\ 14,\ 16,\ 10\sqrt{3},\ 18,\ 20,\ 12\sqrt{3},\ \ldots \tag{40}$$

Remaining in the framework of the model, we must assume that $\gamma_i^{AD} = 1/S_{WS}$, where S_{WS} is the area of a Wigner–Seitz cell corresponding to $d = k\delta_l$ with a certain value of k from series (40). For a densely packed hexagonal lattice, $S_{WS} = (3^{1/2}/2)d^2$ and, hence, $\gamma_i^{AD} = (2/3^{1/2})(k^2\delta_l^2)^{-1}$. Substituting this value into Eq. (39), we obtain

$$R = \frac{3}{4\pi}\frac{\sigma\varepsilon}{e^2}\delta_l^4 k^4.$$

The requirement $R \gg d = k\delta_l$ leads to the following condition for admissible values of k: $k^3 \gg (4\pi/3)e^2/(\sigma\varepsilon\delta_l^3)$; for aqueous solutions, this means that $(k/4.4)^3 \gg 1$. From the series (40) of possible values of k, this condition holds only for $k \geq 10$. However, since $\gamma_i^{AD} \propto 1/k^2$, the requirement that γ_i^{AD} must have the maximal value dictates that only one value, $k = 10$, should be chosen.

In this case, $(k/4.4)^3 \approx 12$, the surface density of adsorbed ions is

$$\gamma_i^{AD} = \frac{2/\sqrt{3}}{(10\delta_l)^2} = 7.6\times 10^{13}\ \text{cm}^{-2}, \tag{40a}$$

the volume density of counterions on the surface of bubstons is

$$n_i^{\max}(R) = \left(\frac{\gamma_i^{AD}}{3}\right)^{3/2} = \frac{2(4/3)^{1/4}}{9(10\delta_l)^3} = 1.3\times 10^{20}\ \text{cm}^{-3}, \tag{40b}$$

the total number of adsorbed ions is

$$N_i = 4\pi R^2\gamma_i^{AD} = \frac{3\sqrt{3}}{2\pi}\times 10^6\frac{(\sigma\varepsilon)^2}{e^4}\delta_l^6 = 1.9\times 10^3, \tag{40c}$$

and the bubston radius is

$$R = \frac{3}{4\pi}\times 10^4\frac{\sigma\varepsilon}{e^2}\delta_l^4 = 14.0\ \text{nm}. \tag{40d}$$

Thus, our model of adsorption permits only one value of the bubston radius, which depends neither on the species of dissolved ions nor on their density n_{i0}. The equilibrium density of bubstons themselves in the solution must naturally depend on the latter quantity. In other words, our model predicts that the bubston radius $R = 14$ nm is universal for all (1 : 1) aqueous solutions of electrolytes, including pure water. Bubstons themselves turn out to be strongly charged since parameter B for them is $B = 4\pi l_B R\gamma_i^{AD} = 94$. It should be emphasized that this conclusion holds only for pure water and for (1 : 1) aqueous solutions. It can easily be seen that, for (Z : Z) aqueous solutions, the bubston radius $R = 14\times 10^{-4}(k^4/Z^2)$ (nm), where the value of k must satisfy the condition $(k/4.4Z^{2/3})^3 \gg 1$. For $Z = 2$, this condition holds for the value of $k = 16$ from series (40) and, hence, $R = 23$ nm. For $Z = 3$, the value of $k = 12\sqrt{3}$ must be chosen from series (40); consequently, $R = 29$ nm.

Let us now prove that the maximal density $\gamma_i^{AD} = 7.6\times 10^{13}\ \text{cm}^{-3}$ of adsorbed ions (which determines the universal value of the bubston radius of 14 nm) can indeed set in for $k = 10$, when the distance d between adsorption centers is equal to $10\delta_l$. The activation energy of ion adsorption for $d = k\delta_l$ is $U_L = 12e^2/\varepsilon d = (12/k)e^2/\varepsilon\delta_l$ and, accordingly, the desorption energy is

$$U_D = U_{AD} - U_L = \frac{3}{2}\frac{e^2}{\varepsilon\delta_l}\frac{1}{A\alpha_0^4}\left(1-\frac{8}{k}A\alpha_0^4\right) = 96\frac{e^2}{\varepsilon\delta_l\alpha_0^4}\left(1-\frac{\alpha_0^4}{8k}\right) \tag{41}$$

(in the latter formula, it is assumed that $A = 1/64$). It follows hence that, in contrast to conditions (6) and (6a), the necessary condition $U_D > 0$ for adsorption of bubstons assumes the form

$$\delta_i/\delta_l < [(8k)^{1/4} - 1]_{k=10} = 1.98 \tag{42}$$

or $\delta_i < 1.98$, $\delta_l = 2.44$ Å. It holds for a wide class of singly charged ions, in particular, for all ions emerging during dissolution of alkali metal salts such as bromides, fluorides, or iodides. It can easily be seen, however, that from two species of ions in the solution, only the ions with smaller radii δ_i can experience stable adsorption. However, in contrast to the case with $d = d_{\min} = 2\times 3^{1/2}\delta_l$, when the division of ions into those capable and incapable of adsorption is determined precisely by the fact that $U_D < 0$ (or $U_D/T \le 1$) for coarser ions, in the present case such a division is based on the difference in the desorption rates for ions with different values of δ_i. We will mark ions with larger and smaller radii by superscripts "1" and "2," respectively ($\delta_i^{(1)} < \delta_i^{(2)}$). Let us compare the desorption probabilities for such ions during the same time interval t (provided that ions were initially in the adsorbed state). These probabilities are given by

$$w_n = [1 - \exp(-\lambda_n t)], \quad n = 1, 2,$$

where

$$\lambda_n = \tau_n^{-1} = \nu_n\exp(-U_D^{(n)}/T)$$

is the desorption probability for the nth ion per unit time, τ_n is the mean lifetime for such an ion in the adsorbed state, and ν_n is the vibrational frequency of this ion in the adsorbed state. Assuming that $w_1 \ll 1$, we have $t = \tau_1 w_1 = (w_1/\nu_1)\exp(U_D^{(1)}/T)$. Then we obtain the following expression for probability w_2:

$$w_2 = 1 - \exp\left(-\frac{\tau_1 w_1}{\tau_2}\right) = 1 - \exp\left[-\frac{\nu_2 w_1}{\nu_1}\exp\left(\frac{U_D^{(1)} - U_D^{(2)}}{T}\right)\right]. \tag{43}$$

It follows hence that, for $\tau_1 w_1/\tau_2 > 1$ or $(U_D^{(1)} - U_D^{(2)})/T > \ln(\nu_1/\nu_2 w_1)$, the desorption probability for coarse ions is $w_2 \approx 1$, while that for small ions (during the same time interval) is $w_1 \ll 1$. Using relation (41), we can write these inequalities in the form

$$\frac{1}{(\alpha_0^{(1)})^4} - \frac{1}{(\alpha_0^{(2)})^4} > \frac{\varepsilon\delta_l T}{96e^2}\ln\frac{\nu_1}{\nu_2 w_1}, \tag{44}$$

where $\alpha_0^{(n)} = (1 + \delta_i^{(n)}/\delta_l)$, $\varepsilon\delta_l T/96e^2 = 1.8 \times 10^{-3}$. When this inequality is satisfied, and the lifetime $\tau_1 > \tau_2/w_1 \gg \tau_2$, coarse ions cannot compete with small ions in relation to adsorption, although both species satisfy the necessary condition (42). While drawing this conclusion, we naturally assume that the fluxes of ions of both species to the bubble surface are equal (for 1 : 1 solution) and, hence, vacant adsorption centers are filled with small ions with a probability $p \approx 1$ for $\tau_1/\tau_2 \gg 1$. Quantity p is the probability of filling, in some way or another, a vacancy with a small ion over time $\tau_1 \gg \tau_2$. Such an event may occur via one of a multitude of statistically independent events: (1) a small ion is adsorbed immediately (probability $p_{10} = 1/2$); (2) it is adsorbed after the adsorption of a coarse ion, followed by desorption ($p_{11} = (1/2)(1/2)$); (3) it is adsorbed after double adsorption followed by desorption of coarse ions ($p_{12} = (1/2)(1/2)^2$); etc., to the case when the adsorption of a small ion follows N-fold adsorption and desorption of coarse ions ($p_{1N} = (1/2)(1/2)^N$), where $N = [\tau_1/\tau_2]$ is an integral part of (τ_1/τ_2). Here, we assume that the time interval between acts of desorption followed by adsorption is $t \ll \tau_2$, which is quite admissible since $t \sim 1/n_{i0}\overline{v}\pi\delta_l^2 \leq 1$ ms ($\overline{v} \sim 10^5$ cm/s is the mean thermal velocity of ions, while lifetimes $\tau_{1,2}$ are measured in hours. Consequently, the probability is given by

$$p = \sum_{j=0}^{N} p_{1j} = \frac{1}{2}\left(1 + \sum_{j=1}^{N}\left(\frac{1}{2}\right)^j\right)$$

$$\approx \frac{1}{2}\left(1 + \sum_{j=1}^{\infty}\left(\frac{1}{2}\right)^j\right) = \frac{1/2}{1-1/2} = 1.$$

Considering ion adsorption in aqueous solutions of electrolytes, we assumed that these solutions are binary (with only two species of ions), which apparently contradicts the actual situation. Indeed, in addition to dissolved ions, solutions always contain intrinsic H^+ and OH^- ions with a concentration approximately equal to 10^{-7} M, and competition between the adsorption of small dissolved ions and H^+ ions becomes significant. The case when the concentration of dissolved ions is much higher than the concentration of H^+ ions is of practical importance. In this case, we can easily prove that the probability of filling vacant adsorption centers with H^+ ions is equal to the ratio of their concentration to the concentration of dissolves ions, i.e., smaller than unity. This means that intrinsic ions in aqueous solutions cannot compete in respect to adsorption with dissolved ions. Naturally, this conclusion is valid not only for bubstons, but also for any macroparticles: when the concentration of aqueous solutions of electrolytes is much higher than 10^{-7} M, the macroparticles (including bubstons) present in these solutions are charged by small dissolved ions.

Let us demonstrate the validity of inequality (44) for specific pairs of ions: (H^+, OH^-) (pure water), (Na^+, Cl^-) (NaCl salt solution), and (K^+, Cl^-) (KCl salt solution). The logarithmic dependence of the right-hand side of inequality (44) on the ratio (ν_2/ν_1) does not require high accuracy in determining its value, and we can use the approximate formula

$$\frac{\nu_1}{\nu_2} \approx \left(\frac{\alpha_0^{(2)}}{\alpha_0^{(1)}}\right)^3\left(\frac{m_2}{m_1}\right)^{1/2}, \tag{45}$$

where $m_{1,2}$ are the masses of the ions. The "exact" formula derived by us on the basis of classical mechanics has the form

$$\frac{\nu_1}{\nu_2} = \left(\frac{\alpha_0^{(2)}}{\alpha_0^{(1)}}\right)^3\left(\frac{m_2}{m_1}\right)^{1/2}\left\{\frac{1 - \dfrac{(\alpha^{(2)})^4}{8k}\,1 - \dfrac{4}{3(\alpha_0^{(1)})^2}}{1 - \dfrac{(\alpha^{(1)})^4}{8k}\,1 - \dfrac{4}{3(\alpha_0^{(2)})^2}}\right\}.$$

For $k = 10$, the last factor is on the order of unity. For the above three pairs of ions, we obtain, respectively, $\nu_1/\nu_2 = 13$, 3.7, and 1.6. Assuming that $w_1 \approx 0.1$, we obtain the following values for the right-hand side of inequality (44): 9×10^{-3}, 6.5×10^{-3}, and 0.9×10^{-3}. The left-hand sides are equal respectively to 1.4×10^{-1}, 6.8×10^{-2}, and 2.7×10^{-2}; consequently, inequality (44) indeed holds, and the surface charge Q_0 is formed due to adsorption of H^+, Na^+, and K^+ ions, respectively.

In addition to the features of bubstons considered above (distinguishing them from other macroparticles), they have one more peculiarity associated with the behavior of compound bubstons. The latter term is applied to particles consisting of bubstons proper (ion-stabilized gas bubbles) with a surface charge Q_0 and a layer of liquid surrounding them, into which the ionic shell of a bubston distorted by its motion remains frozen in. As before, we denote the total charge of such a particle by Q_c and its radius by a_c. As compared to the compound particles considered in Section 4, compound bubstons not only participate in Brownian movement,

but also display regular tendency to emersion at the free surface of the liquid as a result of action of buoyancy forces. The velocity u_c of such emersion can be determined from the condition of equality of the Stokes force $F_{ST} = 5\pi\eta a_c u_c$ (see Footnote 5) and the buoyant force $(F_A - F_g)$, where $F_A = (4\pi/3)\rho g a_c^3$ is the buoyancy force and $F_g = (4\pi/3)\rho g(a_c^3 - R^3)$ is the force of gravity (g being the acceleration due to gravity). This gives

$$u_c = \frac{4}{15}\frac{\rho g R^3}{\eta a_c}. \quad (46)$$

Let us compare this velocity with the root-mean-square value of the smooth component of the velocity of Brownian motion (see Section 4) $\langle u_1^2\rangle^{1/2} = T/5\pi\eta a_c^2$. The velocity ratio $u_c/\langle u_1^2\rangle^{1/2} = (4\pi/3)\rho g R^2 a_c/T$. For $u_c/\langle u_1^2\rangle^{1/2} \ll 1$, i.e., for

$$4\rho g R^3 a_c/T \ll 1, \quad (47)$$

regular emersion of compound bubstons should not affect the influence of Brownian movement on the ionic shell, which was established in Section 4. Since the bubston radius $R = 14$ nm, the left-hand side of inequality (47) does not exceed $3\times10^{-8}a_c$ nm; i.e., condition (47) obviously holds. This, in turn, means that quantities Q_c and a_c must be connected through the universal relation (29) (see comments on formula (29) in Section 4). For dilute solutions with $\kappa R \le 0.1$ (or $n_{i0} \le 4\times10^{-4}/l_B R^2 \approx 3\times10^{15}$ cm^{-3} $= 5\times10^{-6}$ M), when our screening model gives another independent relation between Q_c and a_c in the form of a two-hump dependence of Q_c^2 on a_c (see Fig. 3), we arrive at the division of compound bubstons into two types described in Section 4, viz., small compound bubstons with parameters (Q_c^{sm}, a_c^{sm}) and large ones with parameters (Q_c^l, a_c^l). However, formulas (33) and (34) turn out to be inapplicable for determining these parameters since they were derived for macroparticles with preset values of radii $R \ge 4$ nm and for the surface density of ions $\gamma_i^{AD} = 6.4\times10^{14}$ cm^{-2} (in this case, we could assume, in particular, that $\xi \approx \bar{R}/2$). In the case of bubstons, these formulas need refining. We will do this below as applied to bubstons (and, accordingly, compound bubstons) in pure water ($n_{i0} = 6\times10^{13}$ cm^{-3}). In this case, in accordance with formulas (2a), (14), (14a), (18), (20), and (32), we have the following values of parameters: $\zeta = \kappa R = 1.4\times10^{-2}$, $L = 14.5$, $\alpha = L/B = 0.15$, $\xi = 2.8$, $F_0 = 2.4$, $f = 0.15$, $y_0 = r_0/R = 5.0$, $y_1 = r_1/R = 1.9$, and $(T^*/T) = 390$. Using these values, we obtain for parameters a_c and Q_c

$$a_c^{sm} \approx r_1 = 27 \text{ nm}, \; Q_c^{sm}/e = (24 r_1/l_B)^{1/2} = 9.6.$$

For $a_c^l/R \equiv x$, we have the equation $x = (2\zeta)^{-1}\ln[(5T^*/T)\times(1+\zeta x)^2/x]$, whose solution gives

$$a_c^l/R = 175 \text{ or } a_c^l = 92\, a_c^{sm} = 2.5\ \mu\text{m},$$

$$Q_c^l/e = (a_c^l/a_c^{sm})^{1/2}(Q_c^{sm}/e) = 9.6(Q_c^{sm}/e) = 9.2.$$

Thus, we find that radii a_c^l and a_c^{sm} for pure water differ approximately by two orders of magnitude, while charges Q_c^l and Q_c^{sm} differ, accordingly, by an order of magnitude. The electrophoresis rates u_c^{sm} and u_c^l also differ approximately by an order of magnitude, while condition (35) of electrophoresis linearity in the present case, when $a_c^l = 2.5$ μm, is $E \ll 200$ V/cm. The molar-ionic conductivity σ_b of water associated with bubstons is defined by formula (37), where the bubston density n_b is substituted for n_p. This value of conductivity is always smaller than the intrinsic ionic conductivity σ_i of pure water, which is known to be $\sim10^{-8}$ Q^{-1} cm^{-1} (millipores).

At the same time, the presence of compound bubstons in water must impart the latter a number of specific properties that have not been studied comprehensively as yet. These properties are primarily due to the fact that such formations contain in their cores (bubstons themselves of radius $R = 14$ nm) a free (bubston) gas whose pressure in equilibrium coincides with the external gas pressure p_0 (see above). Pressure p_0 itself does not affect the mechanical equilibrium condition for bubstons and their radius, but naturally determines the number of gas molecules $N_g = (4\pi/3)R^3(p_0/T)$ in a bubston; for $p_0 = 1$ atm, number $N_g \approx 300$. Under conditions of complete thermodynamic equilibrium between the solution with bubstons contained in it and the surrounding gaseous medium, the bubston number density n_b must also be a function of pressure p_0. This follows from the fact that continuous spontaneous (fluctuation-induced) nucleation of microscopic bubbles takes place in the liquid under such conditions, followed by their growth to the state of bubstons. Microscopic bubbles themselves are formed at molecules of the dissolved gas (or their associates), which play the role of inhomogeneities in the liquid. The specific rate (per unit volume) of bubston formation must increase with the concentration c_g of the dissolved gas. In the equilibrium state, the increase in the number of bubstons is compensated by their destruction as a result of emersion at the liquid surface. As a result, the equilibrium density n_b must vanish for $c_g = 0$ and must increase

in proportion to c_g^2, i.e., in proportion to p_0^2, for small c_g. Thus, with increasing external gas pressure p_0, the specific concentration $N_g n_b$ of the free bubston gas in the liquid increases in proportion to p_0^3.

Another aspect of specific properties of water associated with the presence of compound bubstons in it is that the interaction between them must lead to the formation of bubston clusters, viz., peculiar mesoscopic structure, in water. Experimental data on the optical (laser) breakdown of pure water in its transparency region may serve as a proof of the existence of such a structure. The physical mechanism of this effect, which was proposed and studied theoretically in [20] on the basis of the concepts of bubston clusters, was subsequently confirmed by many authors [21–27]. The existence of a bubston cluster structure in water in equilibrium with the external gaseous (air) medium may play a certain role in active life of fishes and other organisms.

6. CONCLUSIONS

Analysis of the screening of strongly charged particles shows that the charge inversion effect can be interpreted in the framework of a modified Poisson–Boltzmann model. The quantitative basis of this model is formed by Eqs. (8), (11), and (12) combined with the quasineutrality condition (1) and the assumption that liquid molecules are arranged on the surface of macroparticles in the form of a densely packed two-dimensional hexagonal lattice. This assumption makes it possible to connect the surface density γ_i^{AD} of adsorbed ions and the volume density $n_i^{\max}(R)$ of counterions on the surface of a macroparticle with the radius δ_l of liquid molecules through Eqs. (7) and (15) for particles with a preset radius or Eqs. (40a) and (40b) for bubstons. The proposed model provides a quantitative description of selective adsorption of ions at the macroparticle surface, leading to the emergence of surface charge Q_0. The sign of charge Q_0 in this case is determined by the sign of ions capable of adsorption. Quantitative criteria for the division of dissolved ions into those capable and incapable of stable adsorption have been formulated. Our analysis of the effect of the Brownian motion of macroparticles on screening shows that the charge inversion effect is manifested in the formation of two types of compound particles of opposite polarities, which strongly differ in their radii and in the absolute values of charges. In this case, the equipartition law expressed by formula (30) is observed for the mean energy of the electric field created by both types of compound particles. This result brings our concepts concerning the interaction of macroparticles in liquid (including aqueous) solutions of electrolytes to a new qualitative level. For example, according to these concepts, coagulation of identical macroparticles in a solution should be attributed not to the interaction of likely charged particles (as was done in numerous theoretical publications), but to the interaction of oppositely charged compound particles of different sizes and absolute values of the charges.

Naturally, the proposed model requires further refinement, primarily in respect of the suspension of the requirement that the solution must be dilute ($\zeta = \kappa R \ll 1$). The charge inversion effect should naturally be preserved after abandoning this requirement. It can obviously be stated that the effect of Brownian movement on screening must also be preserved qualitatively, but the quantitative parameters of the compound particles formed in this case may change.

ACKNOWLEDGMENTS

The authors are grateful to G.A. Lyakhov for the discussion of crucial questions related to the subject matter of this paper and to A.V. Krasnoslobodtsev for his help at the initial stage of this research.

This study was supported financially by the Russian Foundation for Basic Research (project nos. 01-02-17116 and 00-15-96636).

REFERENCES

1. P. Debye and E. Hückel, Phys. Z. **24**, 185 (1923).
2. G. M. Kepler and S. Fraden, Phys. Rev. Lett. **73**, 356 (1994).
3. J. C. Crocker and D. G. Grier, Phys. Rev. Lett. **73**, 352 (1994); Phys. Rev. Lett. **77**, 1897 (1996).
4. A. E. Larsen and D. G. Grier, Nature **385**, 230 (1997).
5. B. V. Deryagin and L. D. Landau, Zh. Éksp. Teor. Fiz. **11**, 802 (1941).
6. E. J. Verwey and J. T. G. Overbeck, *Theory of the Stability of Lyophobic Colloids* (Elsevier, Amsterdam, 1948).
7. R. H. Yoon and J. L. Yordan, J. Colloid Interface Sci. **113**, 430 (1986).
8. C. Li and P. Somasundaran, J. Colloid Interface Sci. **146**, 215 (1991).
9. C. Li and P. Somasundaran, Colloids Surf. A **81**, 13 (1993).
10. N. F. Bunkin and F. V. Bunkin, Zh. Éksp. Teor. Fiz. **100**, 512 (1992) [Sov. Phys. JETP **74**, 271 (1992)].
11. N. F. Bunkin and F. V. Bunkin, Z. Phys. Chem. (Munich) **215**, 111 (2001).
12. E. M. Matescu, C. Jeppersen, and P. Pincus, Europhys. Lett. **46**, 454 (1999); S. Y. Park, R. F. Bruinsma, and W. M. Gelbart, Europhys. Lett. **46**, 493 (1999); J. F. Joanny, Eur. J. Phys. B **9**, 117 (1999).
13. V. I. Perel and B. I. Shklovskii, Physica A (Amsterdam) **274**, 446 (1999); B. I. Shklovskii, Phys. Rev. E **60**, 5802 (1999); T. T. Nguyen, F. Yu. Grosberg, and B. I. Shklovskii, J. Chem. Phys. **113**, 1110 (2000); Phys. Rev. Lett. **85**, 1568 (2000).
14. T. Terao and T. Nakayama, Phys. Rev. E **63**, 041401 (2001).

15. L. D. Landau and E. M. Lifshitz, *Course of Theoretical Physics*, Vol. 8: *Electrodynamics of Continuous Media*, 3rd ed. (Nauka, Moscow, 1982; Pergamon, New York, 1984).
16. B. E. Conway, J. Solution Chem. **7**, 721 (1978).
17. S. S. Dukhin and B. V. Deryagin, *Electrophoresis* (Nauka, Moscow, 1976).
18. L. D. Landau and E. M. Lifshitz, *Course of Theoretical Physics*, Vol. 6: *Fluid Mechanics*, 4th ed. (Nauka, Moscow, 1988; Pergamon Press, Oxford, 1987).
19. V. I. Tikhonov, *Outliers in Random Processes* (Nauka, Moscow, 1970).
20. N. F. Bunkin and F. V. Bunkin, Laser Phys. **3**, 63 (1993).
21. N. F. Bunkin and A. V. Lobeev, Pis'ma Zh. Éksp. Teor. Fiz. **58**, 91 (1993) [JETP Lett. **58**, 94 (1993)].
22. N. F. Bunkin, G. A. Lyakhov, and O. V. Umnova, Zh. Éksp. Teor. Fiz. **104**, 3287 (1993) [JETP **77**, 549 (1993)].
23. N. F. Bunkin and A. V. Lobeev, Kvantovaya Élektron. (Moscow) **21**, 319 (1994).
24. N. F. Bunkin and A. V. Lobeyev, Phys. Lett. A **229**, 327 (1997).
25. N. F. Bunkin, O. A. Kiseleva, A. V. Lobeyev, *et al.*, Langmuir **13**, 3024 (1997).
26. N. F. Bunkin, A. V. Lobeyev, G. A. Lyakhov, and B. W. Ninham, Phys. Rev. E **60**, 1681 (1999).
27. N. F. Bunkin, A. V. Lobeyev, and V. G. Mikhalevich, Phys. Vibr. **7**, 205 (1999).

Translated by N. Wadhwa

Journal of Experimental and Theoretical Physics, Vol. 96, No. 4, 2003, pp. 747–756.
Translated from Zhurnal Éksperimental'noĭ i Teoreticheskoĭ Fiziki, Vol. 123, No. 4, 2003, pp. 846–856.

SOLIDS
Structure

The Structural Properties of $RbMnX_3$ (X = F, Cl, Br) Halides

V. I. Zinenko*, N. G. Zamkova, and S. N. Sofronova

Kirenskii Institute of Physics, Siberian Division, Russian Academy of Sciences, Krasnoyarsk, 660036 Russia

**e-mail: zvi@iph.krasn.ru*

Received November 27, 2002

Abstract—The results of nonempirical calculation of energies of three polytypes (cubic, two-layer hexagonal, and six-layer hexagonal) are given for $RbMnX_3$ (X = F, Cl, Br) crystals. The calculation is performed using an ionic crystal model with regard for the deformability and the dipole and quadrupole polarizabilities of ions. The behavior of these crystals under the action of hydrostatic pressure is studied. It is demonstrated that, at normal pressure, the $RbMnCl_3$ and $RbMnBr_3$ crystals have a six-layer hexagonal structure. At pressures above 11 kbar, $RbMnCl_3$ passes to a phase with a cubic structure; $RbMnBr_3$ at pressures above 90 kbar passes to a phase with a two-layer hexagonal structure. The $RbMnF_3$ crystal under normal conditions has a cubic structure and experiences no phase transformations under the effect of pressure. The obtained results are in satisfactory agreement with the known experimental data. *© 2003 MAIK "Nauka/Interperiodica".*

1. INTRODUCTION

$RbMnX_3$ (X = F, Cl, Br) crystals belong to the family of perovskite-like crystals of the general formula ABX_3. The structures of these compounds may be represented as a three-dimensional packing of rhombohedral layers of AX_3, where A is a large cation and X is an anion. Small cations B are located between layers and occupy the centers of octahedrons formed by anions. The packing of layers may be cubic (Fig. 1a) when the anion octahedrons are bound by corners, or hexagonal (*h* packing) (Figs. 1b and 1c) when the octahedrons are bound by faces. The majority of ABO_3 oxide compounds are crystallized in a perovskite structure (*c* packing); this structure and the physical properties of these substances have been fairly well studied by numerous researchers using both experimental and theoretical (including *ab initio*) methods. In compounds in which X is halogen (F, Cl, Br), both *c* packing and *h* packing may be realized, as well as mixed *ch* packings. For example, $RbMnF_3$ has a cubic perovskite structure, and $RbMnBr_3$ apparently has a hexagonal structure with a two-layer *h* packing. The $RbMnCl_3$ crystal also has a hexagonal structure, but with a six-layer packing of $RbCl_3$ layers (Fig. 1c). Here, pairs of face-bound octahedrons are interconnected via intermediate octahedrons bound to them by corners. Such a six-layer packing is designated as *hcc* packing. Almost no theoretical calculations of such hexagonal structures and of the physical properties of such compounds were performed because of the fairly complex structure and large number of atoms per unit cell.

From the standpoint of a rigid ion model, the formation of hexagonal structures is energetically disadvantageous because, in the case of the *h* packing of AX_3 layers, the ions B come very close together, which leads to a loss in the Madelung energy. Note, however, that hexagonal packings are formed, as a rule, in compounds in which X is an easily polarizable anion (Cl, Br, I). The environment of some ions in hexagonal structures is not centrally symmetric; consequently, in calculating the energy of hexagonal structures, one must take into account the polarization energy arising due to the emergence of induced electric moments (dipole, quadrupole, and so on) in such structures. The importance of inclusion of the polarization energy was emphasized in [1–4], where a number of structures (fluorite, rutile, layer structures, and the like) in compounds of the general formula MX_2 were investigated. Wilson and Madden [1] demonstrated that experimentally observed structures with a symmetry lower than cubic are stabilized only if the total crystal energy includes the contribution by the polarization energy associated with the emergence of induced dipole moments on ions of low-symmetry structures. However, all of the short-range interactions in [1] are written in a parametric form, and the number of parameters is quite large (from six to ten). Wilson *et al.* [2], who used the same parametric model of polarizable ions [3] in the investigation of different phases of ZrO_2, likewise took into account the polarization energy due to induced quadrupole moments and demonstrated that this energy played an important part in low-symmetry structures. However, Wilson *et al.* [2] themselves indicate that the

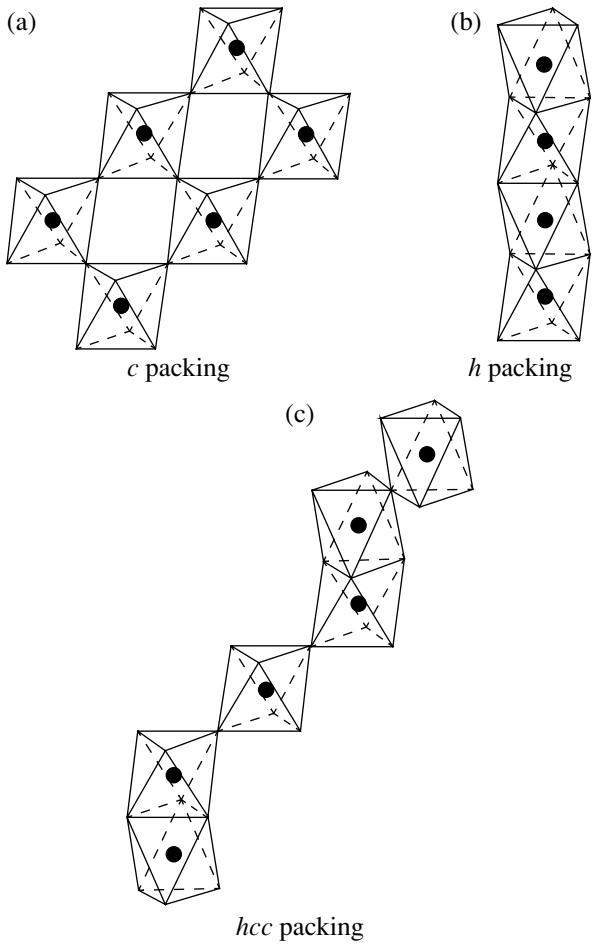


Fig. 1. The arrangement of octahedrons in different polytypes of ABX_3: (a) cubic packing (perovskite structure), (b) two-layer hexagonal packing, (c) six-layer hexagonal packing.

parametrization of the energy associated with quadrupoles is poorly validated.

The electrostatic lattice energies (Madelung energy and dipole energy) for five ideal structures with different sequences of close-packed layers of AX_3 (X = F, Cl, Br, I, O, S) and ABX_3 compounds were calculated by Weenk and Harwig [4]. They assumed the same energy of short-range spherically symmetric ion–ion interactions in different polytypes and ignored short-range dipole–dipole interactions. It was found in [4] that a two-layer hexagonal packing (*h*) turned out to be the stablest structure in ideal structural polytypes of ABX_3 compounds (including oxide compounds).

We used the nonparametric generalized Gordon–Kim model [5] to calculate the energetics of different polytypes for $RbMnX_3$ crystals, where X = F, Cl, Br. In calculating the total energy of these crystals, allowance was made for induced moments, both dipole and quadrupole. The model and method of calculation are described in Section 2. Section 3 gives the results of calculation of the total energy of three crystals, and Section 4 deals with the investigation of the effect of hydrostatic pressure. And, finally, our main results are described in Section 5.

2. MODEL: METHOD OF CALCULATION

In the Gordon–Kim model for ionic crystals, the total electron density is written as the sum of electron densities of separate ions making up a crystal,

$$\rho(\mathbf{r}) = \sum_i \rho(\mathbf{r} - \mathbf{R}_i). \tag{1}$$

The electron densities of separate ions are calculated with regard to the crystalline potential approximated by a charged Watson sphere,

$$V(r) = \begin{cases} -Z_{\text{ion}}/R_w, & r < R_w, \\ -Z_{\text{ion}}/r, & r > R_w. \end{cases} \tag{2}$$

The radii of Watson spheres in separate ions were found from the condition of minimal total energy of crystal.

In the original Gordon–Kim model, the electron density of ions was taken to be spherically symmetric; however, as was observed in [5], distortions of the electron density of any multipole symmetry are possible in an actual crystal. Ivanov and Maksimov [5] suggested a generalization of the Gordon–Kim model, enabling one to take into account multipole density distortions of any order. Here, we allowed for the dipole and quadrupole distortions of electron density,

$$\begin{aligned} \rho(\mathbf{r}) &= \sum_{l=0}^{2} \rho^{(l)}(\mathbf{r}), \\ \rho^{(l)}(\mathbf{r}) &= \sum_{m=-l}^{l} \rho^{(l)}(\mathbf{r}) Y_{lm}(\theta, \phi). \end{aligned} \tag{3}$$

Pair interactions are calculated within the theory of density functional,

$$\begin{aligned} \Phi_{ij}^{ll'} &= F\{\rho_i^{(l)}(\mathbf{r}' - \mathbf{R}_i) + \rho_j^{(l')}(\mathbf{r} - \mathbf{R}_i)\} \\ &- F\{\rho_i^{(l)}(\mathbf{r} - \mathbf{R}_i)\} - F\{\rho_j^{(l')}(\mathbf{r} - \mathbf{R}_i)\}. \end{aligned} \tag{4}$$

The total crystal energy has the form

$$E = E_0 + E_{d-d} + E_{q-q} + E_{d-q} + E_{\text{self}},$$

$$E_0 = -\frac{1}{2} \sum_{i,j=1}^{N_a} Z_i C_{ij}^{(0)} Z_j + \sum_{i,j=1}^{N_a} \Phi_{ij}^{(00)}(V_i, V_j, |\mathbf{R}_i - \mathbf{R}_j|),$$

$$E_{d-d} = \frac{1}{2}\sum_{i,j=1}^{N_a}\sum_{\alpha,\beta=1}^{3} d_i^{\alpha}\Big(\frac{\delta_{ij}}{\alpha_i^d(V_i)} + \Phi_{ij,\alpha\beta}^{(11)}(V_i, V_j, |\mathbf{R}_i-\mathbf{R}_j|) - C_{ij,\alpha\beta}^{(2)}\Big)d_j^{\beta} + \sum_{i,j=1}^{N_a}\sum_{\alpha=1}^{3} d_i^{\alpha}(\Phi_{ij,\alpha}^{(10)}(V_i, V_j, |\mathbf{R}_i-\mathbf{R}_j|) - C_{ij,\alpha}^{(1)}Z_j),$$

$$E_{q-q} = \frac{1}{2}\sum_{i,j=1}^{N_a}\sum_{\alpha,\beta,\gamma,\delta=1}^{3} q_i^{\alpha\beta}\Big[\frac{\delta_{ij}}{\alpha_i^q(V_i)} - \frac{1}{36}(\Phi_{ij,\alpha\beta\gamma\delta}^{(22)}(V_i, V_j, |\mathbf{R}_i-\mathbf{R}_j|) - C_{ij,\alpha\beta\gamma\delta}^{(4)})\Big]q_j^{(\gamma\delta)} - \frac{1}{6}\sum_{i,j=1}^{N_a}\sum_{\alpha,\beta=1}^{3} q_i^{\alpha\beta}(\Phi_{ij,\alpha\beta}^{(20)}(V_i, V_j, |\mathbf{R}_i-\mathbf{R}_j|) - C_{ij,\alpha\beta}^{(2)}Z_j), \quad (5)$$

$$E_{d-q} = -\frac{1}{6}\sum_{i,j=1}^{N_a}\sum_{\alpha,\beta,\gamma=1}^{3} q_i^{\alpha\beta}\times(\Phi_{ij,\alpha\beta\gamma}^{(21)}(V_i, V_j, |\mathbf{R}_i-\mathbf{R}_j|) - C_{ij,\alpha\beta\gamma}^{(3)})d_j^{\gamma},$$

where E_0 is the interaction energy of spherically symmetric ions; E_{d-d}, E_{q-q}, and E_{d-q} are, respectively, the energies associated with the interaction of dipole and quadrupole moments; $C_{ij}^{(n)} = \nabla^n(|\mathbf{R}_i-\mathbf{R}_j|)^{-1}$ is the long-range part of interactions; $E_{\text{self}} = \sum_{i=1}^{N_a} E_i^{\text{ion}}(V_i)$ is the ion self-energy; and N_a is the number of atoms per unit cell. In calculating the short-range interactions given by Eq. (4) for the kinetic energy, the Thomas–Fermi approximation [6] was used, and for the exchange-correlation energy, the Hedin–Lundqvist approximation [7]. The long-range interactions $C_{ij}^{(n)}$ were calculated by the Ewald method. The calculation for ion was performed using Liberman's codes [8]. The modified Sternheimer equation [9] was used to calculate the dipole α^d and quadrupole α^q polarizabilities and the respective components of electron density.

The dipole d_i^{α} and quadrupole $q_i^{\alpha\beta}$ moments were found from the condition of minimal energy with respect to the relevant moment,

$$\frac{\partial E}{\partial d_i^{\alpha}} = 0, \quad \frac{\partial E}{\partial q_i^{\alpha\beta}} = 0,$$

$$d_i^{\alpha} = \sum_{j=1}^{N_a}\sum_{\beta=1}^{3} A_{ij}^{\alpha\beta} \times \sum_{k=1}^{N_a}(\Phi_{jk,\alpha\beta}^{(10)}(V_j, V_k, |\mathbf{R}_j-\mathbf{R}_k|) - C_{jk,\beta}^{(1)}Z_k + q_i^{\alpha\beta}(\Phi_{ij,\alpha\beta\gamma}^{(21)}(V_i, V_j, |\mathbf{R}_i-\mathbf{R}_j|) - C_{ij,\alpha\beta\gamma}^{(3)})), \quad (6)$$

$$q_i^{\alpha\beta} = \frac{1}{6}\sum_{j=1}^{N_a}\sum_{\gamma\delta=1}^{3} \tilde{B}_{ij}^{\alpha\beta,\gamma\delta} \times \sum_{k=1}^{N_a}(\Phi_{jk,\gamma\delta}^{(20)}(V_j, V_k, |\mathbf{R}_j-\mathbf{R}_k|) - C_{jk,\gamma\delta}^{(1)}Z_k + \tilde{A}_{jk,\gamma\delta}),$$

$A_{ij}^{\alpha\beta}$ are elements of the matrix reciprocal to the matrix of dipole–dipole interaction in expression (5);

$$\tilde{B}_{ij}^{\alpha\beta,\gamma\delta} = \Big[\frac{\delta_{ij}}{\alpha_i^q(V_i)} - \frac{1}{36}(\Phi_{ij,\alpha\beta\gamma\delta}^{(22)}(V_i, V_j, |\mathbf{R}_i-\mathbf{R}_j|) - C_{ij,\alpha\beta\gamma\delta}^{(4)}) + \sum_{\lambda,\mu=1}^{3}(\Phi_{ij,\alpha\beta\lambda}^{(21)}(V_i, V_j, |\mathbf{R}_i-\mathbf{R}_j|) - C_{ij,\alpha\beta\lambda}^{(3)}) \times A_{ij}^{\lambda\mu}(\Phi_{ij,\mu\gamma\delta}^{(21)}(V_i, V_j, |\mathbf{R}_i-\mathbf{R}_j|) - C_{ij,\mu\gamma\delta}^{(3)})\Big]^{-1},$$

$$\tilde{A}_{jk,\gamma\delta} = \sum_{\lambda,\mu=1}^{3}(\Phi_{ij,\lambda}^{(10)}(V_i, V_j, |\mathbf{R}_i-\mathbf{R}_j|) - C_{ij,\alpha}^{(1)}Z_j) \times A_{ij}^{\lambda\mu}(\Phi_{ij,\mu\gamma\delta}^{(21)}(V_i, V_j, |\mathbf{R}_i-\mathbf{R}_j|) - C_{ij,\mu\gamma\delta}^{(3)}).$$

3. RESULTS OF CALCULATION OF TOTAL ENERGY

As was mentioned in the Introduction, ABX_3 halides may have both cubic and hexagonal structures. We will restrict ourselves to the discussion of structures of three types, namely, cubic with a perovskite structure (*c* packing) and two hexagonal with two-layer and six-layer packings (*h* and *hcc* packings, respectively), and calculate the energies for three crystals, $RbMnF_3$, $RbMnCl_3$, and $RbMnBr_3$, in these structures.

$RbMnF_3$ crystal has a structure of ideal perovskite with an O_h^1 space group and one molecule per unit

cell [10]. In this structure, the coordinates of all atoms in a unit cell are fixed,

$$\begin{array}{cccccc} \text{A(Rb)} & 1(b) & 1/2 & 1/2 & 1/2 \\ \text{B(Mn)} & 1(a) & 0 & 0 & 0 \\ \text{X} & 3(c) & 1/2 & 0 & 0 \end{array} \quad (7)$$

$RbMnBr_3$ crystal is characterized by a sequence of structural phase transitions, and, in the opinion of Kato *et al.* [11], the crystal structure in all phases has a distorted form of two-layer hexagonal packing. Apparently, no highly symmetric hexagonal phase with a D_{6h}^4 space group and two molecules per unit cell (Fig. 1b) is observed in this crystal up to the melting point [11]. The positions of all cations in a two-layer hexagonal packing are fixed, and the anions have one free parameter,

$$\begin{array}{cccccc} \text{A(Rb)} & 2(c) & 2/3 & 1/3 & 1/4 \\ \text{B(Mn)} & 2(a) & 0 & 0 & 0 \\ \text{X} & 3(h) & x & 2x & 1/4 \end{array} \quad (8)$$

In an ideal structure, $x = 1/6$.

The structure of $RbMnCl_3$ crystal is the most complex of the structures treated by us. In this crystal, a six-layer *hcc* packing with six molecules per unit cell is realized in a highly symmetric phase [12]. All ions in this structure have two crystallographically nonequivalent positions each,

$$\begin{array}{cccccc} \text{A}_1(\text{Rb}_1) & 2(b) & 0 & 0 & 1/4 \\ \text{A}_2(\text{Rb}_2) & 4(f) & 1/3 & 2/3 & z_1 \\ \text{B}_1(\text{Mn}_1) & 2(a) & 0 & 0 & 0 \\ \text{B}_2(\text{Mn}_2) & 4(f) & 1/3 & 2/3 & z_2 \\ \text{X}_1 & 6(h) & y_1 & 2y_1 & 1/4 \\ \text{X}_2 & 12(k) & y_2 & 2y_2 & z_3 \end{array} \quad (9)$$

In this case, five free parameters are observed. In an ideal hexagonal structure, i.e., in a structure in which anions form regular octahedrons, these parameters assume the values

$$z_1 = -1/12, \quad z_2 = 1/6, \quad z_3 = 1/12,$$

$$y_1 = 1/2, \quad y_2 = 1/6.$$

We will first discuss the case of ideal structures with close-packed layers of MnX_3 (X = F, Cl, Br). In this case, the unit cell parameters *a*, *b*, and *c* for the structures being treated are related by stringent relations:

c (cubic perovskite)

$$a_c = b = c = a_0\sqrt{2};$$

h (two-layer hexagonal)

$$a_h = b = 2a_0, \quad c_h = 2a_0\sqrt{6}/3;$$

hcc (six-layer hexagonal)

$$a_h = b = 2a_0, \quad c_h = 2a_0\sqrt{6},$$

where a_0 is the distance Rb–X (X = F, Cl, Br).

Table 1. The values of the Watson sphere radii, dipole and quadrupole polarizabilities of ions, and self-energy of $RbMnX_3$ crystals

Crystal		R_w, Å	α^d, Å^3	α^q, Å^5	E_{self}, eV
$RbMnF_3$	Rb	1.98	1.12	1.97	–120521.7921
	Mn	2.88	0.78	0.85	
	F	1.32	0.79	1.01	
$RbMnCl_3$	Rb	1.75	1.14	2.02	–149971.5256
	Mn	2.22	0.84	0.96	
	Cl	1.39	3.12	7.52	
$RbMnBr_3$	Rb	1.72	1.15	2.04	–324847.5789
	Mn	2.22	0.84	7.52	
	Br	2.18	4.25	12.22	

Table 2. The calculated values (per molecule) of the total energies $E_{total} = E - E_{self}$ and of individual contributions (E^c, Madelung energy; E^s, energy of short-range spherically symmetric ion–ion interactions; E^c_{d-d}, E^c_{q-q}, and E^c_{d-q}, energies of long-range dipole–dipole, quadrupole–quadrupole, and quadrupole–dipole interactions, respectively; and E^s_{d-d}, E^s_{q-q}, and E^s_{d-q}, short-range parts of these interactions, respectively) for ideal close-packed structures

E, eV	$RbMnF_3$ a_0 = 3.11 Å			$RbMnCl_3$ a_0 = 3.63 Å			$RbMnBr_3$ a_0 = 3.85 Å		
	c	*h*	*hcc*	*c*	*h*	*hcc*	*c*	*h*	*hcc*
E^c	–40.5623	–37.6163	–39.6310	–34.6722	–32.1426	–33.8919	–32.6861	–30.3322	–32.0129
E^s	3.7548	3.8042	3.7675	2.9107	2.8909	2.8898	2.4462	2.4549	2.4604
E^c_{d-d}	0.0	–2.4809	–0.7853	0.0	–4.0991	–1.3888	0.0	–4.2000	–1.4459
E^s_{d-d}	0.0	1.8489	0.6014	0.0	2.9942	1.0591	0.0	3.0232	1.0873
E^c_{q-q}	–0.2286	–0.1249	–0.1963	–0.6357	–0.2520	–0.5269	–0.7159	–0.2804	–0.5950
E^s_{q-q}	0.2146	0.1049	0.1803	0.6341	0.2409	0.5225	0.6952	0.2747	0.5850
E^c_{d-q}	0.0	–0.0697	–0.0211	0.0	–0.1791	–0.1318	0.0	–0.2020	–0.1706
E^s_{d-q}	0.0	0.0568	0.0185	0.0	0.1547	0.1247	0.0	0.1865	0.1692
E_{total}	–36.8215	–34.4770	–36.0660	–31.7631	–30.3921	–31.3433	–30.2606	–29.0753	–29.9197

The total crystal energy given by expression (5) was minimized over the cell parameter of the cubic structure and over the radii of Watson spheres for all ions. These radii of Watson spheres were maintained for hexagonal structures as well, and, therefore, all contributions to the total crystal energy for different structures were calculated with the same values of self-energy, spherically symmetric electron density, and dipole and quadrupole polarizabilities of ions. The values of the Watson sphere radii, self-energy of ions, and dipole and quadrupole polarizabilities of ions for the crystals being treated are given in Table 1. The calculated values of individual contributions to the total energy for three structures being treated are given in Table 2.

A cubic perovskite structure (*c* packing) turns out to be more advantageous energetically for all three crystals in ideal structures with close-packed layers, although the difference between the energies of *c* packing and of two- and six-layer (*h* and *hcc*, respectively) hexagonal packing (as well as between the energies of *h* and *hcc* structures) decreases appreciably with increasing radius and polarizability of anion.

We will now discuss individual contributions to the crystal energy for different structures, as given in Table 2. One can see in this table a significant loss in the Madelung energy E^c in hexagonal *h* and *hcc* packing. At the same time, the energy of short-range interactions of the spherical part of the electron density of ions is almost the same for the structures being discussed. The main part in the stabilization of hexagonal structures is played by the polarization energy associated with the interaction between dipole distortions of the electron density of ions in noncentrally symmetric positions in hexagonal structures. Note that, if we take into account only the long-range contribution by pair interactions E^c_{d-d} to the crystal energy, a two-layer hexagonal packing turns out to be most advantageous for the $RbMnCl_3$ and $RbMnBr_3$ crystals. Energetically more advantageous for $RbMnF_3$ is a perovskite structure. The respective energies $E' = E^c + E^s + E^c_{d-d}$ (see Table 2) have the following values (in eV):

$RbMnF_3$ –36.8075(*c*), –36.2930(*h*), –36.6488(*hcc*)

$RbMnCl_3$ –31.7615(*c*), –33.3508(*h*), –32.3969(*hcc*)

$RbMnBr_3$ –30.2399(*c*), –32.0773(*h*), –30.9984(*hcc*)

Table 3. The unit cell parameters and the coordinates of ions in different structures

		$RbMnF_3$	$RbMnCl_3$	$RbMnBr_3$
		Cubic		
a_c	Calculation	4.4	5.1	5.5
	Experiment [10]	4.2		
		Two-layer hexagonal		
a_h	Calculation	5.9	6.7	6.8
	Experiment [14]			7.5
c_h	Calculation	6.0	6.9	7.2
	Experiment [14]			6.6
x/a_h	Calculation	0.1587	0.1727	0.1867
		Six-layer hexagonal		
a_h	Calculation	6.2	7.1	7.3
	Experiment [12]		7.1	
c_h	Calculation	15.7	19.0	20.8
	Experiment [12]		17.8	
y_1/a_h	Calculation	0.4880	0.5008	0.5088
	Experiment [12]		0.4928	
y_2/a_h	Calculation	0.1590	0.1456	0.1563
	Experiment [12]		0.1616	
z_1/c_h	Calculation	−0.1253	−0.1310	−0.133
	Experiment [12]		−0.0888	
z_2/c_h	Calculation	0.1510	0.1543	0.1376
	Experiment [12]		0.1603	
z_3/c_h	Calculation	0.0873	0.1000	0.106
	Experiment [12]		0.0820	

This result agrees with the findings of Weenk and Harwig [4] (in performing the calculations for fluorine compounds, they used values of a_0 and α_d different from those used by us and found almost the same energies for the *c* and *h* packings). However, one can see in Table 2 that the combined polarization energy of dipole–dipole interactions $E^c_{d-d} + E^s_{d-d}$ (note that the quantity E^s_{d-d} includes both the contributions by pair short-range interactions and the contributions by many-particle interactions, including long-range ones) is much lower than E^c_{d-d}. This is associated with the fact that the long-range field due to the lattice point charges and the field due to extended charges of an ion, which are induced on that ion when it is found in a noncentrally symmetric position, are opposite in sign and largely compensate each other. One can see in Table 2 that the contribution made to the total crystal energy by interactions associated with quadrupole distortions of the electron density of ions is small compared to the dipole–dipole energy and approximately the same for all structures being treated.

We will now turn to actual hexagonal structures. In this case, the total crystal energy given by expression (5) was minimized with respect to both the lattice parameters and all free parameters for the respective structure. In so doing, the Watson sphere radii given in Table 1 were preserved for the actual hexagonal structures as well, because our calculation results have demonstrated that, in the case of transition from one structure to another and during minimization over the lattice parameters, the Watson sphere radii either do not change at all or change insignificantly, even in the case

Table 4. The values of dipole and quadrupole moments of ions in different structures (in atomic units)

		$RbMnF_3$	$RbMnCl_3$	$RbMnBr_3$
		Cubic		
q_{zz}	X	0.0745	0.068	−0.309
		Two-layer hexagonal		
d	Rb	0.00	0.00	0.00
	Mn	0.00	0.00	0.00
	X	0.191	0.620	0.965
q_{xx}	Rb	0.025	0.004	−0.016
	Mn	−0.030	−0.032	−0.033
	X	0.044	−0.247	–1.183
q_{yy}	Rb	0.025	0.004	−0.016
	Mn	−0.031	−0.034	−0.035
	X	−0.059	−0.125	0.031
		Six-layer hexagonal		
d	Rb_1	0.0	0.0	0.0
	Rb_2	0.166	0.197	0.230
	Mn_1	0.0	0.0	0.0
	Mn_2	0.014	0.267	0.064
	X_1	0.102	0.372	0.520
	X_2	0.062	0.477	0.808
q_{xx}	Rb_1	0.029	0.065	0.077
	Rb_2	0.011	0.021	0.029
	Mn_1	0.059	0.136	0.146
	Mn_2	−0.024	−0.054	−0.056
	X_1	0.006	−0.077	−0.361
	X_2	−0.047	−0.318	0.299
q_{yy}	Rb_1	0.028	0.065	0.077
	Rb_2	0.011	0.021	0.029
	Mn_1	0.059	0.136	0.146

of a six-layer hexagonal structure with two nonequivalent positions of ions.

Table 3 gives the calculated parameters of structures along with all of the available experimental data. The predicted unit cell parameters agree with the experimental data within 1 to 8%. The greatest difference is observed in the case of determining the parameter c_h for hexagonal structures (6–8%). The calculated six-layer hexagonal structures are more extended along the z axis than the experimentally obtained structures. The agreement between the calculated coordinates of ions in a unit cell and the experimentally observed positions is within the same limits; in such a way, the ion shifts exhibit the same tendency as that observed experimentally.

The calculated values of dipole and quadrupole moments of ions for three structures are given in Table 4. Note that, in hexagonal structures, the total dipole moment of the unit cell is zero. The tensors of quadrupole moments were reduced to the principal axes, with two of three principal values being independent, $q_{zz} = -(q_{xx} + q_{yy})$. In a cubic structure, $q_{xx} = q_{yy} =$

Table 5. The calculated values (per molecule) of the total energies $E_{total} = E - E_{self}$ and individual contributions (designations are the same as in Table 2)

E, eV	$RbMnF_3$			$RbMnCl_3$			$RbMnBr_3$		
	c	*h*	*hcc*	*c*	*h*	*hcc*	*c*	*h*	*hcc*
E^c	–40.5623	–39.1902	–40.3444	–34.6722	–32.6301	–33.8317	–32.6861	–29.0245	–31.3275
E^s	3.7548	5.0400	3.9007	2.9107	3.1097	2.8239	2.4462	1.9673	2.4971
E^c_{d-d}	0.0	–2.4313	–0.2443	0.0	–4.1234	–0.8318	0.0	–4.4966	–1.3753
E^s_{d-d}	0.0	2.0824	0.1405	0.0	2.8193	0.1261	0.0	1.8115	–0.1070
E^c_{q-q}	–0.2286	–0.1247	–0.2044	–0.6357	–0.1133	–0.4500	–0.7159	–0.03979	–0.4177
E^s_{q-q}	0.2146	0.1167	0.1870	0.6341	0.0778	0.4072	0.6952	–0.2429	0.2511
E^c_{d-q}	0.0	–0.0812	–0.0095	0.0	–0.1207	–0.0850	0.0	–0.1175	–0.1515
E^s_{d-q}	0.0	0.0748	0.0073	0.0	0.1046	0.0563	0.0	0.0061	0.0591
E_{total}	–36.8215	–34.5136	–36.5671	–31.7631	–30.8761	–31.7850	–30.2606	–30.1364	–30.5716

$-q_{zz}/2$. The calculated unit cell parameters, the coordinates of ions, and the values of dipole and quadrupole moments were used to calculate individual contributions and the total energy for three crystals in two hexagonal structures; the results are given in Table 5.

One can see in Table 5 that, in the case of hexagonal structures, the crystal lattice relaxation to equilibrium values of the unit cell parameters and coordinates of ions in the lattice brings about a redistribution of the values of different contributions to the total crystal energy and brings about a still finer (compared to ideal packings) balance between these contributions. In a $RbMnF_3$ crystal, the cubic phase with a perovskite structure, in accordance with the experimental results of Copla *et al.* [10], remains more advantageous compared to nonideal hexagonal structures, although the values of the energies of the latter (compared to the energies of ideal structures) are much closer to the value of the energy of cubic structure.

In the $RbMnCl_3$ crystal, in accordance with the experimental results of Goodyear *et al.* [12], a structure with a six-layer *hcc* packing turns out to be stablest, although the energy of cubic *c* packing is very close to the energy of the latter hexagonal structure. Note that, in our calculation, the energetic advantage of a six-layer hexagonal structure compared to a cubic perovskite structure is caused by the contributions made to the total crystal energy by quadrupole–quadrupole and quadrupole–dipole interactions. The polarization energy associated with dipole distortions of the electron density of ions proves insufficient for the stabilization of the hexagonal structure in this crystal.

The dipole contribution to the total crystal energy stabilizes the six-layer hexagonal structure in $RbMnBr_3$. The energy of the *hcc* structure in this crystal is much lower than those of the *c* and *h* structures. The energetic advantage of the six-layer hexagonal structure in the $RbMnBr_3$ crystal is defined by two factors: first, the higher dipole polarizability of bromine ion results in a greater contribution by the dipole energy compared to such contribution in compounds with fluorine and chlorine; and, second, the shifts of ions in $RbMnBr_3$ are such that the difference between the Madelung energies of the *c* and *hcc* structures in this compound is much less than in $RbMnCl_3$.

Note that the results of this calculation demonstrate that, for all three crystals being treated, the two-layer hexagonal structure turns out to be energetically disadvantageous compared both to the cubic perovskite structure and to the six-layer hexagonal structure.

4. PHASE TRANSITIONS UNDER PRESSURE

Under the effect of hydrostatic pressure, many ABX_3 halogen compounds experience phase transitions between different polytypes, preferably to structures with a large fraction of cubic-packed layers [13]. Here, we give the results of calculation of enthalpy,

$$H(\Omega) = (E(\Omega) - E_{self}) + P\Omega$$

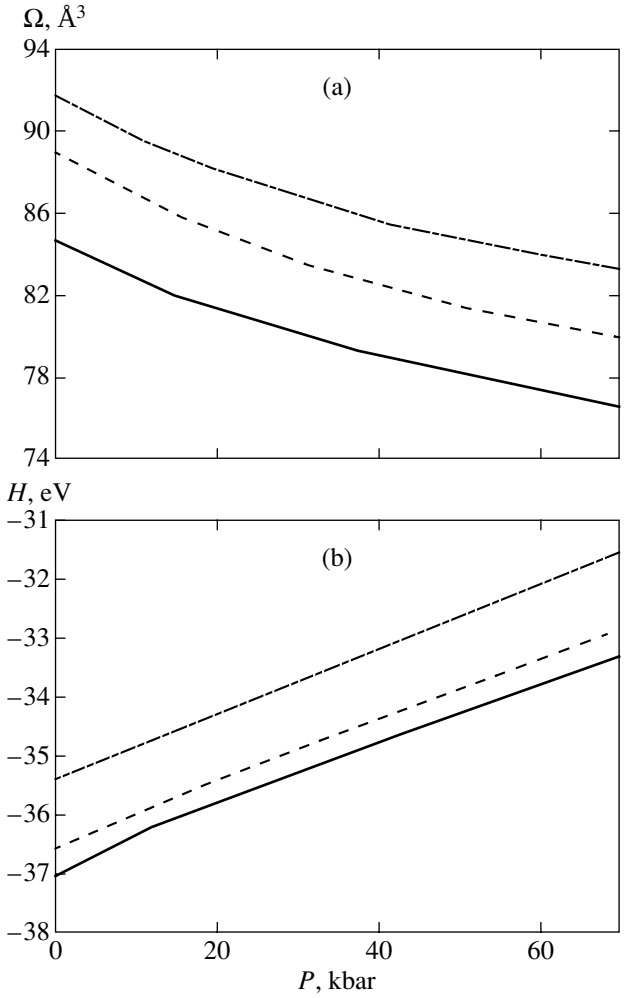


Fig. 2. (a) The equation of state and (b) the pressure dependence of enthalpy for the $RbMnF_3$ crystal. Solid curve, cubic structure; dashed curve, six-layer hexagonal structure; dot-and-dash curve, two-layer hexagonal structure.

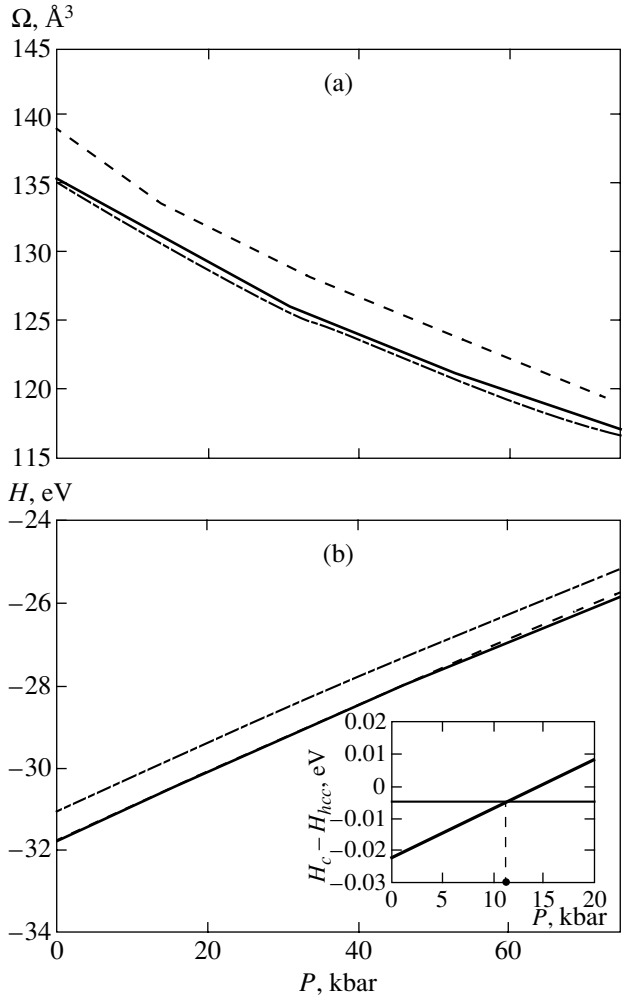


Fig. 3. (a) The equation of state and (b) the pressure dependence of enthalpy for the $RbMnCl_3$ crystal. Solid curve, cubic structure; dashed curve, six-layer hexagonal structure; dot-and-dash curve, two-layer hexagonal structure. The inset shows the pressure dependence of the difference between the enthalpies of the six-layer hexagonal and cubic structures.

($E(\Omega)$ is given by expression (5), P is the pressure, and Ω is the unit cell volume) for three crystals being treated in different structures. In deriving the equation of state, the enthalpy $H(\Omega)$ with the preassigned value of pressure P was minimized with respect to volume; in such a way, the ratio c/a between the unit cell parameters was maintained for all values of pressure. The $\Omega(P)$ equation of state and the $H(P)$ dependence of enthalpy are given in Figs. 2–4. One can see in these figures that the three crystals being treated behave differently under the effect of hydrostatic pressure. In $RbMnF_3$, the perovskite structure remains energetically advantageous under the effect of hydrostatic pressure as well. The $RbMnCl_3$ crystal at pressures above 11 kbar makes a transition from the phase with a six-layer hexagonal packing to the phase with a perovskite structure. The value of pressure $P = 11$ kbar obtained in this calculation agrees very well with the experimentally obtained value of $P = 7$ kbar [13]. The unit cell volume decreases during the $hcc \longrightarrow c$ transition, and the resultant value of the unit cell parameter of the cubic phase $a_0^{cal} = 5.094$ Å likewise agrees well with the experimentally obtained value of $a_0^{exp} = 5.058$ Å [13].

The behavior of the $RbMnCl_3$ and $RbMnBr_3$ crystals under the effect of hydrostatic pressure supports the statement made in the literature [10] about the stabilization, under pressure, of the phase with a perovskite structure in ABX_3 halides. However, in the case of the $RbMnBr_3$ crystal, the phase with a perovskite structure is not realized in this calculation up to pressures of 100 kbar (Fig. 4). Moreover, one can see in Fig. 4 that, at $P > 90$ kbar, the phase with a two-layer hexagonal packing becomes more advantageous energetically in this crystal. We are not aware of any experimental

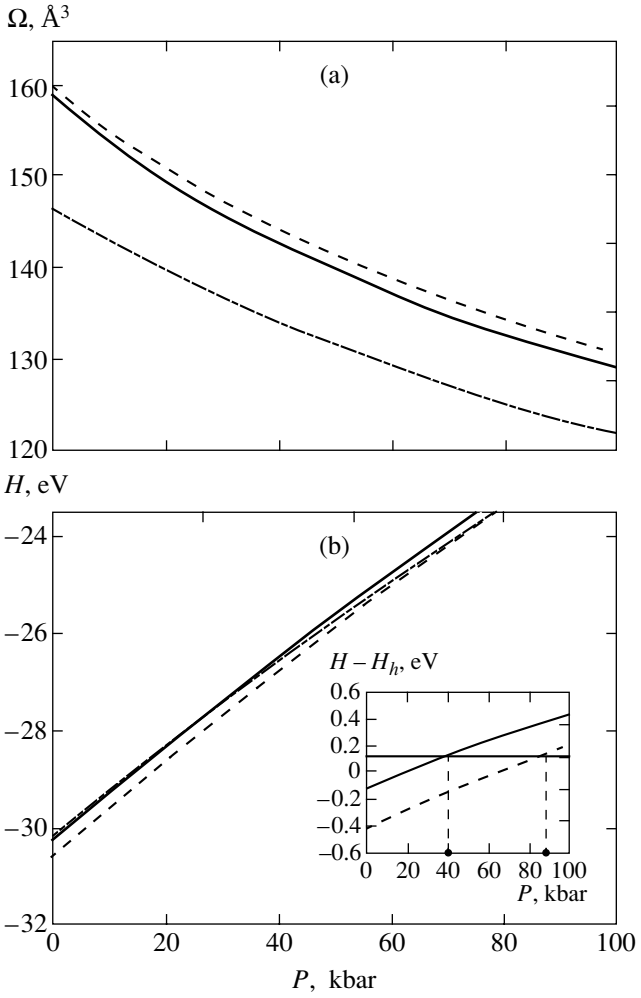


Fig. 4. (a) The equation of state and (b) the pressure dependence of enthalpy for the $RbMnBr_3$ crystal. Solid curve, cubic structure; dashed curve, six-layer hexagonal structure; dot-and-dash curve, two-layer hexagonal structure. The inset shows the pressure dependence of the difference between the enthalpies of the cubic and two-layer hexagonal structures (solid curve) and the six-layer hexagonal and two-layer hexagonal structures (dashed curve).

investigations of $RbMnBr_3$ under the effect of hydrostatic pressure.

5. CONCLUSIONS

We used a nonparametric model of ionic crystal with regard for the dipole and quadrupole polarizabilities to calculate the energies of three structures, namely, cubic (*c* packing), two-layer hexagonal (*h* packing), and six-layer hexagonal (*hcc* packing) for $RbMnF_3$, $RbMnCl_3$, and $RbMnBr_3$ crystals, and investigated the behavior of these crystals under the effect of hydrostatic pressure.

It was found that, in the $RbMnF_3$ crystal, the phase with a perovskite structure is stable both under normal conditions and under the effect of pressure; in such a way, the energy of this phase is significantly lower than the energies of the phases with *h* and *hcc* packings.

In the $RbMnCl_3$ crystal, the hexagonal *hcc* structure turns out to be stablest; under the effect of hydrostatic pressure, this crystal makes a transition to the phase with a perovskite structure. The calculated values of the phase transition pressure and of the unit cell parameter agree well with the experimental data.

In the case of the $RbMnBr_3$ crystal, we failed to observe an energetic advantage of a two-layer hexagonal structure under normal conditions. In our calculations, the energy of a six-layer hexagonal structure is always lower, in spite of the fact that the polarization energy of a two-layer hexagonal structure gives a greater negative contribution to the total crystal energy than the polarization energy in a six-layer hexagonal structure.

ACKNOWLEDGMENTS

This study was financially supported by the Russian Foundation for Basic Research–Yenisei (grant no. 02-02-97707).

REFERENCES

1. M. Wilson and P. A. Madden, J. Phys.: Condens. Matter **6**, 159 (1994).
2. M. Wilson, U. Schonberger, and M. W. Finnis, Phys. Rev. B **54**, 9147 (1996).
3. M. Wilson and P. A. Madden, J. Phys.: Condens. Matter **5**, 2687 (1993).
4. J. W. Weenk and H. A. Harwig, J. Phys. Chem. Solids **38**, 1055 (1977).
5. O. V. Ivanov and E. G. Maksimov, Zh. Éksp. Teor. Fiz. **108**, 1841 (1995) [JETP **81**, 1008 (1995)].
6. L. H. Thomas, Math. Proc. Cambr. Phil. Soc. **23**, 542 (1926); E. Fermi, Z. Phys. **48**, 73 (1928).
7. L. Hedin and B. I. Lundqvist, J. Phys. C **4**, 2064 (1971).
8. D. A. Liberman, D. T. Cromer, and J. J. Waber, Comput. Phys. Commun. **2**, 107 (1971).
9. G. D. Mahan, Phys. Rev. A **22**, 1780 (1980).
10. H. P. Copla, E. G. Sieverts, and R. H. van der Linde, Physica (Amsterdam) **51**, 573 (1971).
11. T. Kato, K. Machida, T. Ishii, and K. Iio, Phys. Rev. B **50**, 13039 (1994).
12. J. Goodyear, G. A. Steigmann, and E. M. Ali, Acta Cryst. B **33**, 256 (1977).
13. J. M. Longo and J. A. Kafalas, J. Solid State Chem. **3**, 429 (1971).
14. H. J. Seifert and E. Dan, Z. Anorg. Allg. Chem. **391**, 302 (1972).

Translated by H. Bronstein

Journal of Experimental and Theoretical Physics, Vol. 96, No. 4, 2003, pp. 757–765.
Translated from Zhurnal Éksperimental'noĭ i Teoreticheskoĭ Fiziki, Vol. 123, No. 4, 2003, pp. 857–866.

SOLIDS
Electronic Properties

Heat Capacity and Electric Resistance of $Sm_{0.55}Sr_{0.45}MnO_3$ Manganite near T_c in a Magnetic Field of up to 26 kOe: Fluctuation Effects and Colossal Magnetoresistance Development Scenario

Sh. B. Abdulvagidov*, I. K. Kamilov, A. M. Aliev, and A. B. Batdalov

Institute of Physics, Dagestan Scientific Center, Russian Academy of Sciences, Makhachkala, Dagestan, 367003 Russia

**e-mail: lowtemp@datacom.ru*

Received October 2, 2002

Abstract—The fluctuation effects in $Sm_{0.55}Sr_{0.45}MnO_3$ manganite are quantitatively characterized and the colossal magnetoresistance (CMR) development scenario in this manganite near T_c is refined based on the results of high-precision measurements of the heat capacity and electric resistance in a broad range of temperatures (77–300 K) and magnetic fields (up to 26 kOe). The experiments revealed a new type of hysteresis related to a jumplike change in the critical temperature T_c due to the phase transition from the ferromagnetic to paramagnetic state. The influence of the order parameter fluctuations on the heat capacity of $Sm_{0.55}Sr_{0.45}MnO_3$ was experimentally observed and quantitatively estimated for the first time. A fluctuation mechanism of CMR development near T_c is proposed. On approaching T_c in a strong magnetic field, the system exhibits a transition from one- to three-dimensional critical behavior. It is shown that, in the presence of a magnetic field, the behavior of microscopic ferromagnetic regions in $Sm_{0.55}Sr_{0.45}MnO_3$ changes near T_c from static into dynamic. It is established that the CMR in this manganite exhibits an anisotropy depending on the mutual orientation of the applied magnetic field and the electric current passing through the sample. The electric resistance of $Sm_{0.55}Sr_{0.45}MnO_3$ is adequately described within the framework of a double exchange model. *© 2003 MAIK "Nauka/Interperiodica".*

1. INTRODUCTION

The phenomenon of colossal magnetoresistance (CMR) in manganites, as well as the related effects and phase transitions, have been known and studied for a rather long time [1]. Nevertheless, the nature of CMR and related transformations still remains unknown. Apparently, new experimental data obtained by high-precision techniques are necessary. The thermal properties of manganites have also been insufficiently studied [2]. Moreover, problems such as the influence of fluctuations on the heat capacity and electric resistance of manganites, which are very important for elucidating the mechanism of CMR near T_c in these compounds, have not been experimentally investigated at all.

Manganites are characterized by a deep interrelation of the magnetic, lattice, and electron subsystems. In our opinion, this a rare phenomenon in solid state physics, where a crystal lattice is usually considered as a factor of the same kind as space and time, establishing rules governing the electric and magnetic phase transformations. For example, features of the crystal lattice structure influence the effective masses of charge carriers and their densities in semiconductors, effective masses of polarons in insulators, virtual photon energies in superconductors, and exchange integrals in ferro-, ferri, and antiferromagnets. It turned out that, in contrast to all these systems, the magnetic and exchange interactions in manganites are the most important factors, which determine the electron and lattice properties of a crystal. For this reason, the crystal lattice of manganites is referred to as "soft," pliable to the action of magnetic fields and exchange interactions. Indeed, a change in the unit cell volume of manganites in the course of a spontaneous magnetostriction near T_c is comparable to the values of analogous changes accompanying structural phase transitions.

As is known, the antiferromagnetic matrix in manganites behaves as a dielectric "parent rock" featuring nucleation of the ferromagnetic phase. However, the CMR in manganites is always induced by the magnetic phase transition from the ferromagnetic to paramagnetic state. This suggests that the CMR can probably take place in the absence of an antiferromagnetic matrix, that is, in a usual ferromagnet. Such a substance must behave as a metal in the ferromagnetic state and as a semiconductor or insulator in the paramagnetic state. In this case, CMR would allow a trivial explanation, being analogous to the electrical resistance vanishing in high-temperature superconductors as interpreted within the framework of the fluctuation theory of

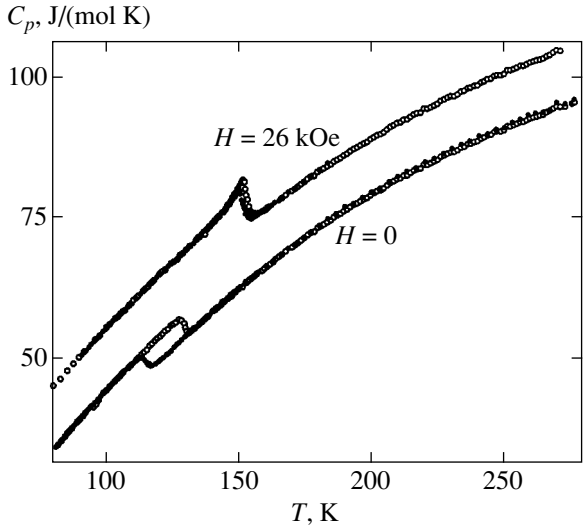


Fig. 1. Temperature variation of the heat capacity of $Sm_{0.55}Sr_{0.45}MnO_3$ manganite measured in the (○) heating and cooling (●) mode without an applied magnetic field and in a field of 26 kOe. For the sake of clarity, the latter curve is shifted by 10 J/(mol K) upward.

phase transitions. Alternation of the conductivity type in such substances is possible due to their giant magnetostriction [3, 4], which can under certain conditions lead to such changes in the energy band structure of a solid (appearance of the conduction band), that is, to a classical phase transformation of the Mott transition type [5].

This work aimed at quantitatively studying the fluctuation effects, refining the CMR development scenario near T_c, and elucidating the mechanisms of electric conductivity in ferromagnetic and paramagnetic states of manganites. The composition $Sm_{0.55}Sr_{0.45}MnO_3$ has proved to be a good choice (probably optimum among all manganites) for these purposes.

Technology used for the synthesis of $Sm_{0.55}Sr_{0.45}MnO_3$ ceramics is described in detail elsewhere [6]. According to X-ray diffraction data, the samples represented a single-phase orthorhombic perovskite with the lattice parameters a = 0.5424(1) nm, b = 0.7678(2) nm, c = 0.5434(2) nm. Scanning electron microscopy showed the material to be homogeneous with an average grain size of 2 μm, a good cleavage character, and a porosity of about 20%. Electron-probe microanalysis revealed homogeneous grain composition in accordance with the empirical formula $Sm_{0.55}Sr_{0.45}MnO_3$. The oxygen number determined by iodometric titration [7] amounted to 3.02.

Thermal behavior of the samples was studied in an automated setup for heat capacity measurements on small samples by ac calorimetry with a relative error not exceeding 0.1% [8]. The average temperature in the calorimeter was measured by a copper–constantan thermocouple with a wire diameter of 100 μm; the temperature oscillations were detected with a chromel–constantan thermocouple with a wire diameter of 25 μm. The temperature was varied at a rate not exceeding 0.01 K/min (in the vicinity of the phase transition, 0.005 K/min). The sample temperature oscillations at a frequency of 2 Hz were excited by alternating light flux from an incandescent lamp. The amplitude of these oscillations did not exceed 0.05 K, which is especially important for the measurements in the neighborhood of the phase transition. The electric resistance of samples was measured by the standard four-point-probe technique.

2. RESULTS AND DISCUSSION

2.1. Heat Capacity Measurements

Figure 1 shows the curves describing variation of the heat capacity of a $Sm_{0.55}Sr_{0.45}MnO_3$ sample in a temperature range from 77 to 300 K. The measurements were performed in the heating and cooling mode without applied magnetic field and in a field of 26 kOe. As can be seen, the heat capacity profiles of $Sm_{0.55}Sr_{0.45}MnO_3$ measured both with and without an applied field exhibit anomalies with hysteresis in the neighborhood of T_c. Moreover, the critical temperature T_c also depends both on the applied field strength H and on the temperature variation direction. The measurements at $H = 0$ in the heating mode (indicated by a plus superscript), which corresponds to the phase transition from the ferromagnetic to paramagnetic state, yield $T_c^+ = 128.6$ K; in the cooling mode (indicated by minus superscript), which corresponds to the reverse transition from the paramagnetic to ferromagnetic state, $T_c^- = 113.3$ K. In a magnetic field of H = 26 kOe, the corresponding characteristic values are $T_c^+ = 152.7$ K and $T_c^- = 150.6$ K. Thus, the critical temperature increases with the applied magnetic field strength. At the same time, the width of the hysteresis $\Delta T_c = T_c^+ - T_c^-$ decreases from $\Delta T_c(0) = 15$ K to $\Delta T_c(26\ \text{kOe}) = 2$ K. The magnitudes of jumps in the heat capacity, calculated by extrapolation of the experimental values obtained before and after the phase transition, were $\Delta C_p^+ = 4.3$ J/(mol K) and $\Delta C_p^- = 4$ J/(mol K) at $H = 0$ and became nearly equal $\Delta C_p^+ = \Delta C_p^- = 8.7$ J/(mol K) at H = 26 kOe. Taking into account that the hysteresis in magnetic materials can be related to a long-term relaxation or a first-order phase transition, the measurements were repeated at various heating and cooling rates and in different temperature cycling intervals. However, the T_c (and, hence, the ΔT_c) values obtained with and without applied field remained unchanged. An analysis of these results showed that the observed behavior can be given a consistent physical interpretation only assuming that the system exhibits a tempera-

ture-induced paramagnet–ferromagnet phase transition with a jumplike change in T_c, that is, the ferromagnetic and paramagnetic phases are characterized by different T_c values.

Previously [6], experimental results on the phase transition in $Sm_{0.55}Sr_{0.45}MnO_3$ were given a preliminary explanation in terms of the electronic phase separation model, according to which this manganite at $T < T_c$ is an antiferromagnetic dielectric with ferromagnetic inclusions (metal clusters). However, subsequent thorough analysis of the data reported in [6] showed that there is no need to use this model, since $Sm_{0.55}Sr_{0.45}MnO_3$ manganite has no antiferromagnetic dielectric matrix. Later, the validity of this approach was fully confirmed by the results of experiments with neutron diffraction and muon spin relaxation [9, 10]. According to these data, no phase separation at all takes place in $Sm_{0.55}Sr_{0.45}MnO_3$: below T_c, this manganite occurs in a homogeneous ferromagnetic state with a magnetic moment of manganese ions ($3.37\mu_B$ for saturation at $T = 4$ K) close to that in the completely ordered state ($3.55\mu_B$) and without any signs of an antiferromagnetic phase at $T < T_c$. In addition, according to the neutron diffraction data, the molar volume of $Sm_{0.55}Sr_{0.45}MnO_3$ near T_c exhibits a significant decrease (lattice contraction) upon the transition to the ferromagnetic phase. In our opinion, this change in the molar volume is caused by the giant spontaneous magnetostriction. Since the space symmetry group (*Pnma*) remains the same in the entire temperature range, no structural phase transition (according to the phase transition classifications of Ehrenfest and Landau) takes place in the system. The lattice parameters change in a very special manner: the rhombic base sharply contracts (which corresponds to a jump in the temperature dependence of *a* and *c* values), while the *b* value varies to a much lower extent. We believe that such changes in the interatomic distances and bond angles in the unit cell base plane lead to an increase in the energy of indirect exchange interaction between manganese ions and, hence, in the T_c^+ value for the ferromagnetic phase. Therefore, a transition from ferromagnetic to paramagnetic phase takes place in a crystal lattice with a different (according to our data, higher) value of T_c^+ that accounts for the observed hysteresis in the behavior of $Sm_{0.55}Sr_{0.45}MnO_3$. Therefore, an increase in T_c^+ and T_c^- and a decrease in the hysteresis width ΔT_c in the magnetic field are explained by the fact that the lattice parameters and the bond angles between magnetically active manganese ions in $Sm_{0.55}Sr_{0.45}MnO_3$ before and after the phase transition become closer due to the giant spontaneous magnetostriction.

Thus, $Sm_{0.55}Sr_{0.45}MnO_3$ manganite is characterized by the hysteresis behavior of a new type related to a jumplike change in T_c. This behavior has nothing in common, except a similar appearance in the curves, with hysteresis related to a phase transition of the first order. Hence, such a hysteresis cannot be considered as sufficient evidence for the first-order phase transition and it would be incorrect to judge the phase transition type based upon a single qualitative sign, the presence or absence of a hysteresis loop [11], as it is frequently done (see, e.g., [10, 12, 13]). Judging the phase transition type proceeding from the presence or absence of hysteresis (or, which is even worse, from the hysteresis width and the anomaly sharpness and magnitude) is essentially the same as doing this based on the latent heat of the transition (assuming that small and large values of this heat unambiguously correspond to phase transitions of the second and first order, respectively). Unfortunately, the aforementioned classifications of phase transitions according to Ehrenfest and Landau are almost not employed in the investigations of manganites. This opinion was recently supported by Salamon *et al.* [14], who indicated that the ferromagnetic transformation in $La_{0.7}Ca_{0.3}MnO_3$ manganite should be classified as a second-order transition in view of the absence of hysteresis, otherwise the sharpness of the heat capacity change could be evidence of the first-order transition. For correctly determining the phase transition type, it would be important to experimentally detect an infinite jump in the heat capacity. However, in practice we can only measure a sufficiently large heat capacity (two orders of magnitude greater than the value outside the transition neighborhood), which is related to random noise always present in the measurement tract.

As can be seen from Fig. 1, the heat capacity in a zero field smoothly varies from a value corresponding to the ferromagnetic phase to that characteristic of the paramagnetic state. According to the Landau classification, this corresponds to a continuous variation of the order parameter from unity in the ferromagnetic state to zero in the paramagnetic state and implies that the sample exhibits a second-order phase transition. The phase transition in $Sm_{0.55}Sr_{0.45}MnO_3$ at the Curie temperature in a zero magnetic field is adequately described by the Landau theory of phase transitions [15], the only special feature being that the ferromagnet–paramagnet transition takes place at T_c^+, and the reverse transition, at a different temperature T_c^-. In the absence of magnetic field, T_c^- is significantly lower than T_c^+, while in a field of 26 kOe, the critical temperatures approach one another: $\Delta T_c(26\ \text{kOe}) \approx 2$ K.

In addition, application of the magnetic field leads to a λ-shaped peak in the heat capacity, which is characteristic of the systems subject to a significant influence of fluctuations in the vicinity of T_c. This fact indicates that the phase transition in $Sm_{0.55}Sr_{0.45}MnO_3$ manganite should be described within the framework of the fluctuation theory of phase transitions [16]. Since such a behavior is usually not typical of ferromagnets, we

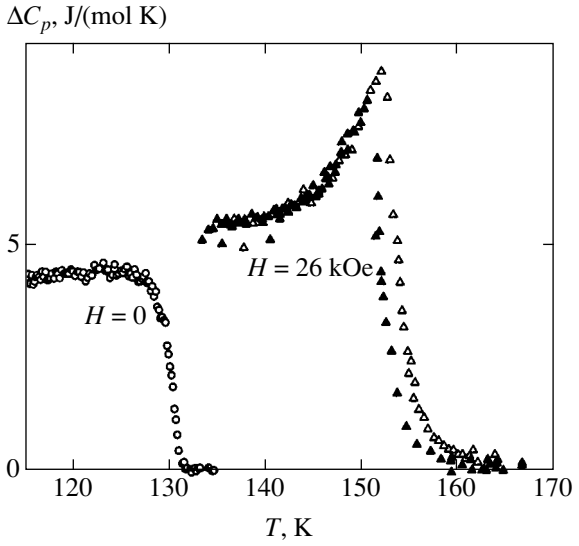


Fig. 2. Temperature variation of the anomalous part of the heat capacity of $Sm_{0.55}Sr_{0.45}MnO_3$ manganite, $\Delta C_p = C_p - C_B$, determined (○) for heating in the absence of a magnetic field and for (△) heating and (▲) cooling in a field of 26 kOe.

propose a simple explanation for the origin of a λ-shaped peak in the heat capacity profile of $Sm_{0.55}Sr_{0.45}MnO_3$ observed in the magnetic field. The width of the critical region in the vicinity of T_c is determined by the Ginzburg–Levanyuk relation

$$\varepsilon = \frac{1}{32\pi^2}\left(\frac{k_B}{\Delta C_p \xi^3}\right)^2, \quad (1)$$

where $\varepsilon = |T/T_c - 1|$ is the reduced temperature, ΔC_p is the heat capacity jump at T_c, and ξ is the correlation radius [17, 18]. Evidently, at $T < T_c$, a ferromagnetic phase dominates in which fluctuations of the paramagnetic phase take place when the temperature falls within the interval $[T_c, (\varepsilon + 1)T_c]$, while at $T > T_c$, a paramagnetic phase dominates with temperature fluctuations of the ferromagnetic phase taking place within the interval $[(\varepsilon - 1)T_c, T_c]$. In the presence of hysteresis, the appearance of paramagnetic (ferromagnetic) fluctuations in the ferromagnetic (paramagnetic) phase is possible provided that the temperature in the fluctuating region would spontaneously increase (decrease) by a value not smaller than the hysteresis width ΔT_c.

In the absence of magnetic field, $Sm_{0.55}Sr_{0.45}MnO_3$ is characterized by $\Delta T_c = 15$ K. This corresponds to a significant change in the internal energy $dU = Cp\xi^3\Delta T_c$ of the fluctuating regions and, according to thermodynamics and statistical physics, implies an extremely small probability of fluctuations. For this reason, the correlation radius of the temperature fluctuations is significantly smaller than the correlation radius of fluctuations in the magnetic order parameter. In contrast, $\Delta T_c = 2$ K for a field of $H = 26$ kOe corresponds to a significantly lower value of dU. In this case, the ΔT_c value becomes much lower than the usual width of the fluctuational ε-neighborhood ($\varepsilon = T_c \pm 10$ K for ferromagnets and antiferromagnets, ferroelectrics, and other substances). The correlation radius of the temperature fluctuations turns out to be comparable with the correlation radius of fluctuations in the magnetic order parameter and, according to formula (1), the probability of fluctuations sharply increases, which is manifested in experiment. Thus, the heat capacity of $Sm_{0.55}Sr_{0.45}MnO_3$ behaves in accordance with the Landau theory of phase transitions in the absence of magnetic field and obeys the fluctuation theory of phase transitions in a field of 26 kOe.

In order to quantitatively estimate the fluctuation effects near T_c, we separated the anomalous heat capacity contributions from the measured temperature dependences of ΔC_p (Fig. 1). This was achieved by subtracting a regular heat capacity contribution C_B, representing the heat capacity of $Sm_{0.55}Sr_{0.45}MnO_3$, measured in a zero field in the cooling mode, approximated by a cubic polynomial

$$C_B = a_0 + a_1 T + a_2 T^2 + a_3 T^3 \quad (2)$$

with the coefficients

$$a_0 = -62.67944 \text{ J/(mol K)},$$

$$a_1 = 1.51493 \text{ J/(mol K}^2\text{)},$$

$$a_2 = -0.00618 \text{ J/(mol K}^3\text{)},$$

$$a_3 = 1.08297 \times 10^{-5} \text{ J/(mol K}^4\text{)}.$$

Figure 2 shows the temperature dependence of the anomalous part of the heat capacity, $\Delta C_p = C_p - C_B$, in the absence of a magnetic field (for measurements in the heating mode) and in a field of 26 kOe (for both heating and cooling stages). As can be seen from these data, the ΔC_p values in the zero field change from one almost constant value, corresponding to the paramagnetic phase, to another almost constant value, characterizing the ferromagnetic phase, without any visible signs of fluctuation effects (no λ-shaped peak of heat capacity). A qualitatively different pattern of the phase transition is observed in the magnetic field of 26 kOe, whereby a λ-shaped peak (characteristic of the second-order phase transitions) is observed in both heating and cooling modes.

Above T_c, the fluctuational contribution (indicated by plus superscript) is [19, 20]

$$\Delta C_p = C^+ \varepsilon^{-\alpha}, \quad (3)$$

where $C^+ = k_B/8\pi\xi^3$ is the critical amplitude of heat capacity above T_c, $\alpha = 2 - d/2$ is the critical index of heat capacity, and d is the fluctuation space dimension.

In Fig. 3a, the anomalous part of the heat capacity in a zero magnetic field is plotted as ΔC_p versus $T/T_c - 1$ in a double logarithmic scale for $T_c = 127.5$ K (corresponding to the heat capacity peak). In the interval where $\log(T/T_c - 1)$ varies from –3 to –2, the plot is linear with a nearly zero critical index of heat capacity ($\alpha = 0.03$). This result confirms the above assumption concerning the absence of any significant influence of the order parameter fluctuations on the phase transition in a zero field (according to the Landau theory, $\alpha = 0$). In the narrow interval where $\log(T/T_c - 1)$ varies from –1.6 to –1.4, the linear plot with $\alpha = 10.26$ (far from any standard values) is evidence of the absence of a fluctuational smearing of the phase transition in $Sm_{0.55}Sr_{0.45}MnO_3$ heated in a zero field. Thus, the Curie point observed for $Sm_{0.55}Sr_{0.45}MnO_3$ in a zero field can be a classical example of the second-order phase transition, in full agreement with the Landau theory of phase transitions [15], which ignores the effect of the order parameter fluctuations on the phase transition.

According to the scaling theory [16], the T_c value is determined as a temperature at which the critical indices below and above this point are equal. However, in our case this scaling procedure is inapplicable because of a significant difference between critical temperatures of the ferromagnetic and paramagnetic phases of $Sm_{0.55}Sr_{0.45}MnO_3$. We have somewhat modified a procedure for determining T_c, so as to take into account the difference of critical temperatures for the ferromagnetic and paramagnetic manganite phases. The initial T_c values were taken equal to the temperatures of heat capacity maxima observed in the heating (T_c^+) and cooling (T_c^-) stages. Then, these values were varied with sequentially decreasing temperature steps (0.1, 0.01, and 0.001 K) until the critical indices α for the ferromagnetic and paramagnetic phases would coincide. This procedure yielded $T_c^+ = 153.425$ K and $T_c^- = 151.400$ K.

In order to provide for the most reliable evaluation of the effect of fluctuations on the heat capacity of $Sm_{0.55}Sr_{0.45}MnO_3$ in the vicinity of T_c, we have used T_c as a single fitting parameter. Analogous data processing is usually performed with several fitting parameters, including a regular part of the heat capacity and some other quantities.

Figures 3b and 3c show the anomalous part of the heat capacity determined in the heating and cooling modes for a magnetic field of 26 kOe, plotted as ΔC_p versus $T/T_c - 1$ in a double logarithmic scale for the corresponding values of T_c determined as described above. As can be seen, $\log\Delta C_p$ varies as a linear function of $\log(T/T_c - 1)$ in approximately the same temperature intervals. The heat capacity parameters of $Sm_{0.55}Sr_{0.45}MnO_3$ calculated according to the fluctuation theory of phase transitions are presented in the table.

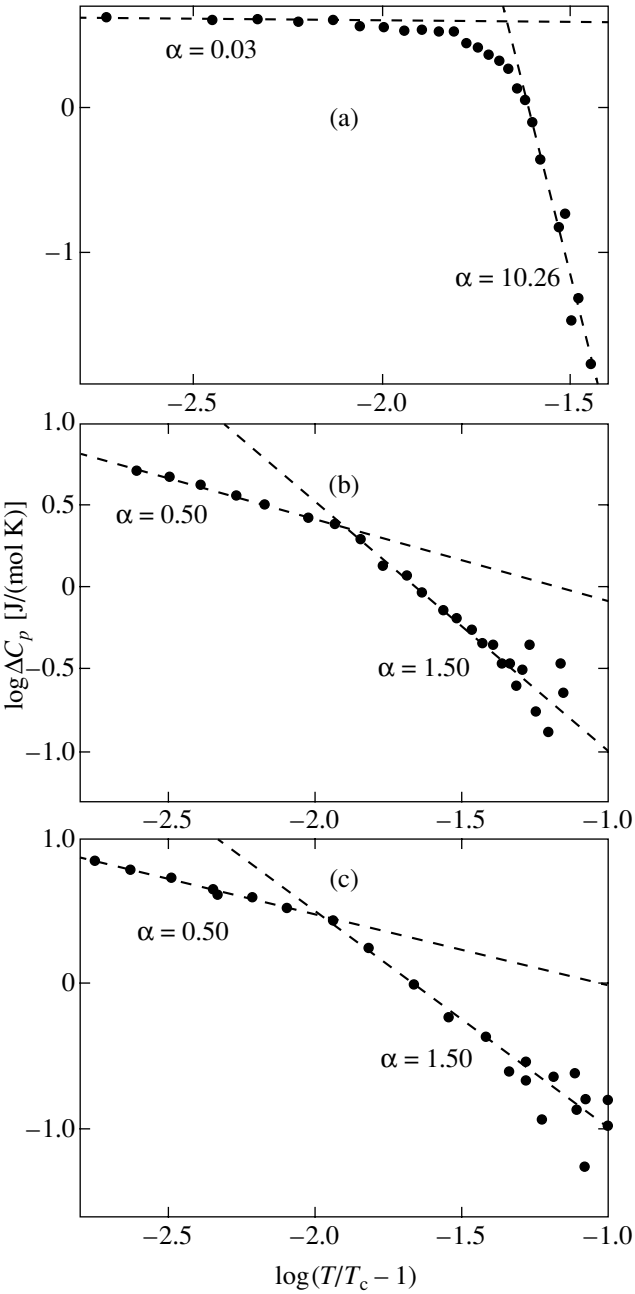


Fig. 3. Double logarithmic plot of ΔC_p versus reduced temperature $T/T_c - 1$ for $Sm_{0.55}Sr_{0.45}MnO_3$ manganite for (a) heating in a zero field ($T_c = 127.5$ K) and for (b) heating and (c) cooling in a field of 26 kOe ($T_c^+ = 153.425$ K and $T_c^- = 151.400$ K, respectively).

As can be seen from the data in the table, the critical indices α and heat capacity amplitudes C^+ for $Sm_{0.55}Sr_{0.45}MnO_3$ heated and cooled in a magnetic field of 26 kOe correspond to physically meaningful dimensions d of the fluctuation space and the correlation radii

The parameters of phase transitions in $Sm_{0.55}Sr_{0.45}MnO_3$ manganite in a magnetic field, calculated taking into account the fluctuation effects

T_c, K	H, kOe	Interval of logε	α	C^+, J/(cm^3 K)	d	ξ, Å
Heating stage						
153.425	26	[–2.8, –1.9]	0.50	0.0018725	3.00	6.64
		[–1.9, –1.3]	1.50	2.401×10^{-5}	1.00	28.39
Cooling stage						
151.400	26	[–2.8, –1.93]	0.50	0.0021748	3.00	6.32
		[–1.93, –1.45]	1.50	2.32×10^{-5}	1.00	28.72

ξ comparable with the crystal unit cell parameters of the manganite studied. Far from T_c, the dimensionality of fluctuations is $d = 1$. On approaching T_c (logε = –1.9), the system exhibits the transition to three-dimensional critical behavior ($d = 3$), whereby the size of ferromagnetic grains increases as well. Note that the fluctuation space dimensions equal, to within the second decimal digit, the real (integer) physical values. Another remarkable fact is that the correlation radii ξ of the magnetic order parameter for $Sm_{0.55}Sr_{0.45}MnO_3$ determined from our experimental heat capacity data (see table) agree well (or even coincide to within the experimental error) with the analogous values determined by De Teresa *et al.* [10] from the results of small-angle neutron scattering (SANS) experiments.

The SANS data [10] also showed that the lifetime of the so-called ferromagnetic clusters in $Sm_{0.55}Sr_{0.45}MnO_3$ exposed to a strong magnetic field never exceed 10^{-12} s. Therefore, it would be expedient to use a common terminology and refer these clusters as the magnetic order parameter fluctuations described within the framework of the well-known fluctuation theory of phase transitions. On the other hand, an analysis of the data in Fig. 3a shows that the microscopic ferromagnetic regions cease to fluctuate in a zero field and behave as static, so that the term "ferromagnetic cluster" would be more expedient under these conditions. Thus, the application of a magnetic field changes the pattern of coexisting ferromagnetic and paramagnetic phases in $Sm_{0.55}Sr_{0.45}MnO_3$ near T_c from static to dynamic.

The results of our investigation of the fluctuation effects in $Sm_{0.55}Sr_{0.45}MnO_3$ manganite suggest that the CMR has a fluctuational nature as well. In the manganite sample exposed to a magnetic field at $T > T_c$, fluctuations of the ferromagnetic phase in the form of droplets with a characteristic size on the order of $\xi = 7$ Å appear near T_c in the dominating paramagnetic phase volume. The magnetization vectors of these droplets are oriented predominantly along the applied field, thus favoring the mutual attraction of dipoles with the formation of something like ferromagnetic filamentary domains. This behavior resembles the orientation of iron filings on a sheet of paper along the lines of a magnetic field created by a U-shaped permanent magnet. As the magnetic field strength grows, the length of these filamentary domains increases to become comparable with the sample dimensions and the electric resistance sharply drops to produce the CMR effect. In addition, the ordering of magnetic moments of the ferromagnetic clusters in the magnetic field favors the conversion of bonds at the cluster boundaries from insulating to conducting, which is manifested by a shift of the effective percolation point upward in the temperature scale [14]. Thus, the appearance of a percolation threshold does not even require any direct contact between ferromagnetic clusters or droplets.

It should be noted that the formation of such a domain structure in a zero field is hindered by very strong forces of surface tension at the boundaries between ferromagnetic and paramagnetic phases. These forces are caused by the giant spontaneous magnetostriction of $Sm_{0.55}Sr_{0.45}MnO_3$ manganite [4]. In addition, the formation of this structure in a zero field is impossible because the fluctuations in $Sm_{0.55}Sr_{0.45}MnO_3$ are almost completely suppressed due to the large difference between critical temperatures of the ferromagnetic and paramagnetic phases.

If the above CMR scenario is valid, a sample of $Sm_{0.55}Sr_{0.45}MnO_3$ manganite in a magnetic field must exhibit an anisotropy in the electric resistance depending on the mutual orientation of the applied magnetic field and the electric current passing through the sample. This anisotropy is expected to be maximum at T_c and to decrease with temperature at $T < T_c$ while being absent in the paramagnetic state at $T > T_c$. This was confirmed by the results of our measurements of the anisotropy of the electric resistance of a $Sm_{0.55}Sr_{0.45}MnO_3$ sample in a magnetic field of 26 kOe. Indeed, a relative change in the resistance measured for the perpendicular and parallel or antiparallel mutual orientations of the current and the magnetic field directions was $\rho_\perp/\rho_\parallel - 1 = 0.102$, 0.191, and 0 at $T = 77$, 139, and 300 K, respectively. A relative change in the CMR

magnitude was $(CMR)_{\parallel}/(CMR)_{\perp} - 1 = 0.104$, 0.127, and 0, respectively.

Thus, $Sm_{0.55}Sr_{0.45}MnO_3$ manganite at a temperature and a magnetic field strength corresponding to the aforementioned filamentary domain structure formation exhibits a one-dimensional critical fluctuation behavior (see table). On approaching T_c from above (by decreasing the sample temperature or increasing the field strength at $T > T_c$), the intensity of fluctuations increases and they acquire a three-dimensional character, thus favoring realization of the finite-dimensional scaling in the magnetic field.

2.2. Electric Resistance

Figure 4 shows a family of temperature dependences of the electric resistance of $Sm_{0.55}Sr_{0.45}MnO_3$ measured at various magnetic field strengths (0. 6, 13, and 26 kOe) in the heating and cooling modes plotted in the logρ versus $1/T$ coordinates. As can be seen from these data, the electric resistance also exhibits hysteresis, which is suppressed by the magnetic field and shifted toward higher temperatures. The applied magnetic field strongly suppresses the electric resistance, thus leading to a CMR effect that is most pronounced at T_c. In addition, the transition from metallic to semiconductor conductivity in a zero field is sharp, while the same transition in a field of 26 kOe proceeds within a 10 K interval, which corresponds to the width of a fluctuation region in usual ferromagnets. This smearing of the transition region probably reflects the increasing role of fluctuations in the magnetic order parameter. In a zero field, the width of the critical temperature region for $Sm_{0.55}Sr_{0.45}MnO_3$ is virtually zero, which favors rapid change in the character of the electric conductivity type. In a field of 26 kOe, the fluctuation region significantly increases and microdomains of a ferromagnetic phase appear at $T > T_c$ in the (still predominant) paramagnetic phase. These domains produce a decrease in ρ on approaching T_c and lead to a percolation threshold of the fluctuation nature at $T = T_c$. Apart from the different character of the temperature dependences of ρ at $T > T_c$ and $T < T_c$ in manganites and high-T superconductors, the effect of fluctuations on the CMR in both media is much alike, except that the percolation threshold is not as clearly manifested in the ferromagnetic phase of manganites (because of finite ρ) as in high-T superconductors, where this phenomenon leads to complete vanishing of the nonzero electrical resistance.

Thus, an analysis of the results of electrical resistance measurements in $Sm_{0.55}Sr_{0.45}MnO_3$ depending on the temperature in a zero and nonzero magnetic field shows that smearing of the temperature interval of the metal–semiconductor transition and the appearance of

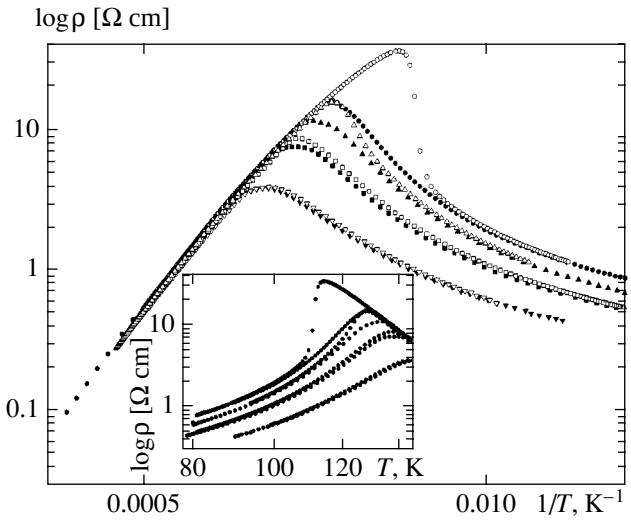


Fig. 4. Temperature dependences of the electric resistance ρ of $Sm_{0.55}Sr_{0.45}MnO_3$ manganite in various magnetic fields H = 0 (circles), 6 (triangles), 13 (squares), and 26 kOe (nablas) (black and open symbols correspond to heating and cooling modes, respectively). The inset shows the same curves plotted in different coordinates.

the percolation threshold in the magnetic field are related to the effect of thermodynamic fluctuations on the phase transition at T_c.

A sharp decrease in the electric resistance of $Sm_{0.55}Sr_{0.45}MnO_3$ measured in a zero field in the cooling mode and a less steep variation of $\rho(T)$ in the heating mode can be explained as follows. For a qualitatively different temperature behavior of the electric resistance of $Sm_{0.55}Sr_{0.45}MnO_3$ above and below T_c (increase in the ferromagnetic state against decrease in the paramagnetic one), the hysteresis observed upon the transition to the ferromagnetic state in the cooling mode leads (because of a decrease in T_c as compared to that in the paramagnetic state) to a greater difference in resistances of the ferromagnetic and paramagnetic phases at the transition temperature. This is manifested by a sharp drop in the electric resistance measured in the cooling mode. For the same reasons, measurements in the heating mode lead to the opposite effect: the hysteresis observed upon the transition to paramagnetic state leads (because of an decrease in T_c as compared to that in the ferromagnetic state) to a smaller difference in resistances of the ferromagnetic and paramagnetic phases at the transition temperature. This is manifested by a less steep anomaly in the electric resistance measured near T_c in the heating mode.

In addition, an analysis of the temperature dependences of the electric resistance of $Sm_{0.55}Sr_{0.45}MnO_3$ measured in various magnetic fields (Fig. 4) led us to the conclusion that the electric resistance of $Sm_{0.55}Sr_{0.45}MnO_3$ in a magnetic field is described well by a double exchange model, which is characterized by

a superexponential growth of $\rho(T)$ and $T < T_c$, a relatively high T_c value, and a semiconductor-type relation $\log\rho(T) \propto 1/T$ at $T > T_c$. According to the double exchange model, the charge carrier relaxation rate depends on the mutual orientation of the spins of ions in the neighboring lattice sites. In the absence of a long-range magnetic order (in the paramagnetic phase) or in the vicinity of T_c (where the thermodynamic fluctuations disordering the spins of the neighboring ions are strongly developed), the charge carriers can exhibit localization.

Therefore, according to the double exchange model, the fluctuation effects (leading to ordering or disordering of spins in microscopic regions with a characteristic size on the order of ξ) seem to significantly influence behavior of the electric resistance near T_c.

Recently, Izyumov and Skryabin [21] showed that the temperature dependence of the electric resistance in such a system is described by the formula

$$\rho(T) = \rho_0 \exp\left(\frac{1 - \langle S_0 S_1\rangle/S^2}{1 + \langle S^z\rangle/S}\frac{W}{4kT}\right), \qquad (4)$$

where $\langle S_0 S_1\rangle$ is the pairwise correlation function for spins in the neighboring sites and W is the seeding band width. Note that the exponent contains two factors exhibiting opposite variation with the temperature. At $T < T_c$, the first factor is small in the ferromagnetic region, thus making the electric resistance small as well. In the vicinity of T_c, this factor sharply grows and attains a constant level at $T > T_c$ (paramagnetic phase), where spins in the neighboring sites are no longer correlated. Here, the second factor in the exponent becomes predominant, since the metal–semiconductor transition of the Mott type at the Curie point must lead top a significant change in the mobility edge (E_c), the chemical potential (μ), and the seeding band width (W). The linear portions of the temperature dependence of the electric resistance at $T > T_c$ are well approximated by formula (4) with the seeding band width W = 110.79 meV. For $T < T_c$, the electric resistance of $Sm_{0.55}Sr_{0.45}MnO_3$ is also described by formula (4), where the first term in the exponent sharply increases on approaching T_c. As can be seen from Fig. 4, the electric resistance of $Sm_{0.55}Sr_{0.45}MnO_3$ at $T < T_c$ exhibits a superexponential growth with increasing deviation from the exponential law (see the inset in Fig. 4). Thus, the results of the electric resistance measurements for $Sm_{0.55}Sr_{0.45}MnO_3$ agree well with the double exchange model.

3. CONCLUSIONS

We have experimentally observed and interpreted a new type of the hysteresis behavior related to a jump-like change in T_c due to the phase transition from ferromagnetic to paramagnetic state. The effect of the order parameter fluctuations on the heat capacity and electric conductivity of $Sm_{0.55}Sr_{0.45}MnO_3$ manganite was experimentally observed and quantitatively estimated for the first time. A fluctuation mechanism of the CMR development near T_c is proposed. The heat capacity of $Sm_{0.55}Sr_{0.45}MnO_3$ exhibits a change from behavior according to the Landau theory in the zero magnetic field to behavior described by a fluctuation theory of phase transitions. On approaching T_c in a strong magnetic field, the system exhibits a transition from one- to three-dimensional critical behavior. It is shown that, in the presence of magnetic field, the behavior of microscopic ferromagnetic regions in $Sm_{0.55}Sr_{0.45}MnO_3$ changes near T_c from static into dynamic. It is established that the electric resistance ρ of $Sm_{0.55}Sr_{0.45}MnO_3$ manganite exhibits an anisotropy depending on the mutual orientation of the applied magnetic field and the electric current passing through the sample. The electric resistance of $Sm_{0.55}Sr_{0.45}MnO_3$ is adequately described within the framework of the double exchange model. In this manganite, unlike usual ferromagnets, the magnetic field increases the intensity of the order parameter fluctuations in the vicinity of T_c, rather than suppressing these fluctuations. A smooth change in the conductivity type observed in $Sm_{0.55}Sr_{0.45}MnO_3$ at T_c is related to the influence of fluctuations on the electric resistance. A fluctuational scenario of the CMR development is proposed, which is realized in $Sm_{0.55}Sr_{0.45}MnO_3$ manganite in the vicinity of T_c.

ACKNOWLEDGMENTS

The authors are grateful to O.Yu. Gorbenko for kindly providing a high-quality sample of $Sm_{0.55}Sr_{0.45}MnO_3$.

This study was supported by the Russian Foundation for Basic Research, project nos. 00-07-90241, 02-07-06048, 02-02-17895, and 00-15-96662).

REFERENCES

1. É. L. Nagaev, Usp. Fiz. Nauk **166**, 833 (1996) [Phys. Usp. **39**, 781 (1996)].
2. M. B. Salamon and M. Jaime, Rev. Mod. Phys. **73**, 583 (2001).
3. A. M. Kadomtseva, Yu. F. Popov, G. P. Vorob'ev, *et al.*, Fiz. Tverd. Tela (St. Petersburg) **42**, 1077 (2000) [Phys. Solid State **42**, 1110 (2000)].
4. A. I. Abramovich, L. I. Koroleva, A. V. Michurin, *et al.*, Fiz. Tverd. Tela (St. Petersburg) **42**, 1451 (2000) [Phys. Solid State **42**, 1494 (2000)].
5. P. Njvak, M. Marysko, M. M. Savosta, and A. N. Ulyanov, Phys. Rev. B **60**, 6655 (1999).
6. A. M. Aliev, Sh. B. Abdulvagidov, A. B. Batdalov, *et al.*, Pis'ma Zh. Éksp. Teor. Fiz. **72**, 668 (2000) [JETP Lett. **72**, 464 (2000)].
7. A. A. Bosak, O. Yu. Gorbenko, A. R. Kaul, *et al.*, J. Magn. Magn. Mater. **211**, 61 (2000).

8. Sh. B. Abdulvagidov, G. M. Shakhshaev, and I. K. Kamilov, Prib. Tekh. Éksp., No. 5, 134 (1996).
9. A. M. Aliev, Sh. B. Abdulvagidov, A. B. Batdalov, *et al.*, Fiz. Tverd. Tela (St. Petersburg) **45**, 124 (2003) [Phys. Solid State **45**, 130 (2003)].
10. J. M. De Teresa, M. R. Ibarra, P. Algarabel, *et al.*, Phys. Rev. B **65**, 100403(R) (2002).
11. N. D. Mathur and P. B. Littlewood, Solid State Commun. **119**, 271 (2001).
12. A. Asamitsu, Y. Moritomo, Y. Tomioka, *et al.*, Nature **373**, 407 (1995).
13. F. Millange, S. de Brion, and G. Chouteau, Phys. Rev. B **62**, 5619 (2000).
14. M. B. Salamon, P. Lin, and S. H. Chun, Phys. Rev. Lett. **88**, 197203 (2002).
15. L. D. Landau and E. M. Lifshitz, *Statistical Physics*, 3rd ed. (Nauka, Moscow, 1976; Pergamon Press, Oxford, 1980), Part 1.
16. A. Z. Patashinskiĭ and V. L. Pokrovskiĭ, *Fluctuation Theory of Phase Transitions*, 2nd ed. (Nauka, Moscow, 1982; Pergamon Press, Oxford, 1979).
17. A. P. Levanyuk, Zh. Éksp. Teor. Fiz. **36**, 810 (1959) [Sov. Phys. JETP **9**, 571 (1959)].
18. V. L. Ginzburg, Fiz. Tverd. Tela (Leningrad) **2**, 2031 (1960) [Sov. Phys. Solid State **2**, 1824 (1960)].
19. L. P. Gor'kov, Zh. Éksp. Teor. Fiz. **34**, 735 (1958) [Sov. Phys. JETP **7**, 505 (1958)].
20. L. G. Aslamazov and A. I. Larkin, Fiz. Tverd. Tela (Leningrad) **10**, 1104 (1968) [Sov. Phys. Solid State **10**, 875 (1968)].
21. Yu. A. Izyumov and Yu. N. Skryabin, Usp. Fiz. Nauk **171**, 121 (2001) [Phys. Usp. **44**, 109 (2001)].

Translated by P. Pozdeev

Journal of Experimental and Theoretical Physics, Vol. 96, No. 4, 2003, pp. 766–774.
Translated from Zhurnal Éksperimental'noĭ i Teoreticheskoĭ Fiziki, Vol. 123, No. 4, 2003, pp. 867–880.

SOLIDS
Electronic Properties

Electrons in a Curvilinear Quantum Wire

L. I. Magarill* and M. V. Éntin**

Institute of Semiconductor Physics, Siberian Division, Russian Academy of Sciences, pr. Akademika Lavrent'eva 13, Novosibirsk, 630090 Russia

**e-mail: levm@isp.nsc.ru*

***e-mail: entin@isp.nsc.ru*

Received October 28, 2002

Abstract—The adiabatic motion of electrons in curvilinear quantum wires was studied. It was assumed that the cross section of a wire was constant along its length. The potential that limited electron motion across a wire and the shape of the cross section of the wire were considered arbitrary, while the curvature and the torsion (defined as the derivative of the cross section rotation angle with respect to the length) were assumed to be small. An effective nonrelativistic Hamiltonian for the motion of electrons along a wire with the conservation of transverse quantum numbers was obtained. The spin-orbit coupling Hamiltonian related to the curvature and torsion of a wire was found. Particular cases of a rectilinear twisted quantum wire with a noncircular cross section and a curvilinear quantum wire on a plane were studied. Various transverse potential models limiting the motion of electrons were considered. In particular, the coefficients of the effective Hamiltonian for quantum wires with rectangular and circular cross sections and hard walls and for wires with a parabolic potential were found. © *2003 MAIK "Nauka/Interperiodica"*.

1. INTRODUCTION

In recent years, rapid technological progress has made it possible to create various low-dimensional systems of complex geometrical shapes [1–4] such as scrolls, rings, helices, and other structures. Electronic states in curved low-dimensional systems have been studied both theoretically and experimentally (e.g., see [5–10]). It is known that curvature causes the appearance of an effective geometrical potential [11–13], which augments the large energy levels of transverse quantization. Corrections to the longitudinal kinetic energy, which are inversely proportional to the thickness of the layer, arise in addition to the geometrical potential in systems with an asymmetric potential that limits transverse motion [14].

As distinct from mathematical lines, one-dimensional systems should be considered taking into account their transverse geometry. For this reason, one-dimensional systems are characterized not only by curvature but also by the shape of their cross sections. A local one-dimensional Hamiltonian of the system should be determined both by the local curvature of the system and by the torsion of the wire along its axis. The particular case of a rectilinear twisted wire was considered by us in [15].

The purpose of this work is to obtain an effective Hamiltonian for electrons in a curved and, possibly, twisted quantum wire. We consider quantum wires of a constant cross section with both circular and noncircular symmetry. We will construct an adiabatic Hamiltonian for the motion of electrons along wires. Next, some particular cases will be analyzed.

First, a Hamiltonian not taking into account the spin of electrons will be obtained. Further, an additional term for spin-orbit coupling will be found.

Various quantum wires to which the statement of our problem pertains are shown in Figs. 1 and 2. Some of them are particular cases of quantum wires obtained in [1].

2. PROBLEM STATEMENT

The motion of electrons with a quadratic and isotropic spectrum in a bent and/or twisted quantum wire is considered. Ignoring its thickness, we see that the wire satisfies the equation

$$\mathbf{r} = \mathbf{a}(q_3), \tag{1}$$

where q_3 is the length of the arc along the $\mathbf{a}(q_3)$ curve. This one-dimensional description is, however, insufficient for taking into account the shape of the wire. As the wire has a finite thickness, Eq. (1) only determines some medial line that passes along the wire.

The system under consideration can be treated as a uniform rectilinear wire whose cross section has an arbitrary shape; the wire is twisted along its axis and/or bent (see Fig. 2). The curvature and torsion of the wire will be considered to arbitrarily depend on q_3. Mathematically, a wire can be described as a locus formed by a plane figure that moves along a curve in such a way that its plane remains normal to the curve at points of intersection and that the curve everywhere intersects the figure at the same point of the figure.

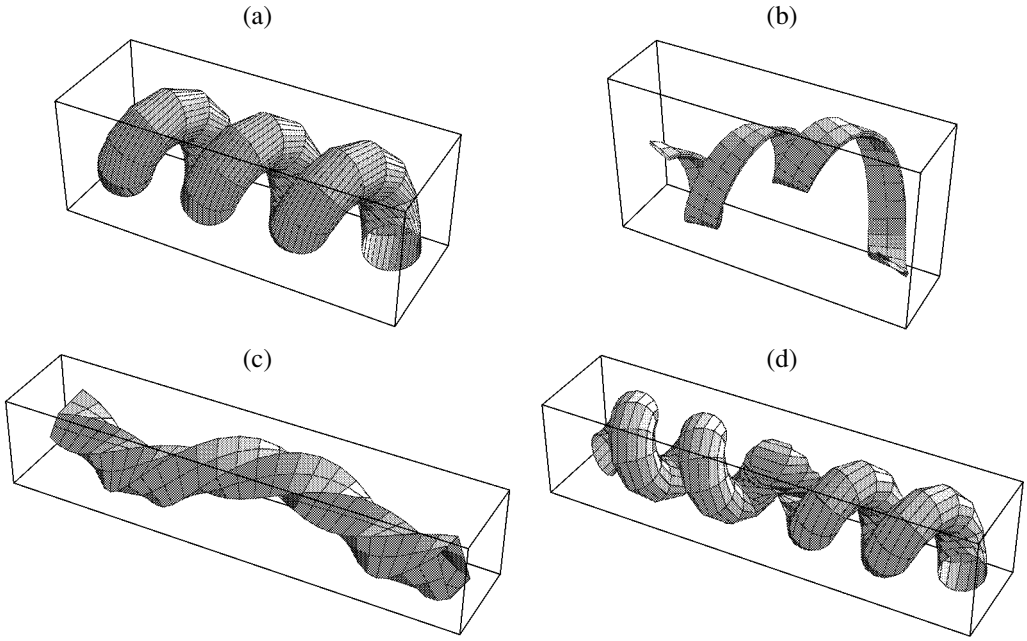


Fig. 1. Examples of curved quantum wires with constant cross sections: (a) helix with a constant pitch formed from a wire of a circular cross section, (b) curved thin band without internal torsion, (c) twisted curved wire with a square cross section, and (d) adjoined left and right helices with circular cross sections.

Generally, we will restrict the motion of electrons across a wire by an arbitrary (not necessarily rigid) potential $U(\mathbf{r})$. The potential that satisfies our assumptions should equally depend on local coordinates transverse with respect to the curve in an arbitrary cross section with accuracy to coordinate system rotations around the tangent to the curve. Let us select a system of (curvilinear) coordinates in which the potential does not depend on q_3. The moving trihedron of the curve will be constructed from the tangent $\mathbf{t}(q_3) = \partial_3\mathbf{a}$, normal $\mathbf{n}(q_3) = \partial_3\mathbf{t}/|\partial_3\mathbf{t}|$, and binormal $\mathbf{b}(q_3) = \mathbf{t} \times \mathbf{n}$. We also introduce the vectors related to it,

$$\mathbf{n}_1 = \mathbf{n}\cos\phi - \mathbf{b}\sin\phi, \quad \mathbf{n}_2 = \mathbf{n}\sin\phi + \mathbf{b}\cos\phi. \quad (2)$$

Here, $\phi(q_3)$ is the angle of rotation of the cross section around $\mathbf{t}$. Let us define new coordinates by the relation

$$\mathbf{r}(\mathbf{q}) = \mathbf{a}(q_3) + q_1\mathbf{n}_1 + q_2\mathbf{n}_2. \quad (3)$$

In q_i coordinates, the transverse potential has the form $U = U(\mathbf{q}_\perp)$, where $\mathbf{q}_\perp = (q_1, q_2)$.

We assume that the thickness of the wire is small compared with its radius of curvature; in addition, the product of wire torsion $\omega \equiv \partial_3\phi$ and its thickness will be considered small. These parameters determine the adiabatic character of electron movement along the wire; namely, electrons that move along the wire retain the number of their transverse quantization subband. Schematically, an effective Hamiltonian is constructed as follows. The initial Hamiltonian is expanded into a series in powers of thickness, and the wave function is sought in the form of its expansion in transverse adiabatic states. This gives a one-dimensional Hamiltonian for motion in a certain transverse quantization subband.

3. AN EFFECTIVE NONRELATIVISTIC HAMILTONIAN

The Schrödinger equation in the new coordinates takes the form

$$\left[-\frac{1}{2\mu}\frac{1}{\sqrt{G}}\sum_{i,j}\partial_i(\sqrt{G}G^{ij}\partial_j) + U(q_1, q_2)\right]\Psi = E\Psi. \quad (4)$$

Here, G^{ij} is the contravariant metric tensor inverse to the covariant metric tensor $G_{ij} = \partial_i\mathbf{r} \cdot \partial_j\mathbf{r}$ and $G = \det G_{ij}$.

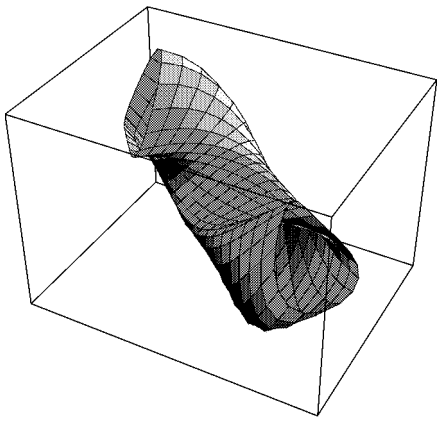


Fig. 2. Segment of a curved and twisted wire with an asymmetric cross section.

Based on (3) and (2), we obtain the $\partial_i \mathbf{r}$ derivatives in the form

$$\partial_3 \mathbf{r} = \mathbf{t}(1 - \kappa(\mathbf{q}_\perp \cdot \mathbf{l})) + (\tau - \omega)([\mathbf{q}_\perp \times \mathbf{l}]_3 \times \mathbf{n} + (\mathbf{q}_\perp \cdot \mathbf{l})\mathbf{b}), \quad (5)$$

$$\partial_{1,2} \mathbf{r} = \mathbf{n}_{1,2}. \quad (6)$$

In (5), we use the notation $\mathbf{l} = (\cos\phi, \sin\phi)$; $\kappa = |\partial_3 \mathbf{t}|$ is the curvature of the curve; and $\tau(q_3) = (\mathbf{t} \cdot \partial_3 \mathbf{t} \times \partial_3^2 \mathbf{t})/\kappa^2$ is the geometrical torsion of the curve. This equation follows from the Serret–Frenet formulas $\partial_3 \mathbf{n} = -\kappa \mathbf{t} + \tau \mathbf{b}$, $\partial_3 \mathbf{b} = -\tau \mathbf{n}$.

Using (5) and (6), we obtain

$$G_{ij} = \begin{pmatrix} 1 & 0 & -q_2\zeta \\ 0 & 1 & q_1\zeta \\ -q_2\zeta & q_1\zeta & G + \zeta^2 q_\perp^2 \end{pmatrix}, \quad (7)$$

$$G^{ij} = \frac{1}{G}\begin{pmatrix} G + \zeta^2 q_2^2 & -\zeta^2 q_1 q_2 & \zeta q_2 \\ -\zeta^2 q_1 q_2 & G + \zeta^2 q_1^2 & -\zeta q_1 \\ \zeta q_2 & -\zeta q_1 & 1 \end{pmatrix}, \quad (8)$$

$$G = (1 - \kappa(\mathbf{q}_\perp \cdot \mathbf{l}))^2 \equiv (1 - Q)^2, \quad (9)$$

where $\zeta = \tau - \omega$. The wave function should be normalized according to the condition

$$\int |\Psi|^2 \sqrt{G} d^3 q = 1.$$

Let us introduce the new function $\Psi = \Phi G^{-1/4}$ to eliminate $\sqrt{G}$ from the volume element in the normalization of the wave function. As we are interested in the limit of small wire thickness d, the coefficients of the Schrödinger equation will be expanded in $\zeta q_\perp$ and Q bearing in mind that $q_\perp \sim d$ and the action of $\partial_{1,2}$ on transverse wave functions gives values on the order of $1/d$.

After the corresponding substitutions, the Schrödinger equation takes the form

$$\left[\mathbf{p}_\perp^2 + (p - \zeta M)^2 - \frac{\kappa^2}{4}\right]\Phi = 2\mu(E - U)\Phi, \quad (10)$$

where $p_j = -i\partial_j$, $\mathbf{p} = (\mathbf{p}_\perp, p_3)$, and $M = q_1 p_2 - q_2 p_1$ is the operator of momentum projection onto the third axis.

Let us expand wave function Φ in transverse states $\psi_n(q_\perp)$ in the wire,

$$\Phi(\mathbf{q}) = \sum_n \psi_n(\mathbf{q}_\perp)\chi_n(q_3). \quad (11)$$

These states satisfy the equation

$$\left(\frac{\mathbf{p}_\perp^2}{2\mu} + U\right)\psi_n = E_n \psi_n, \quad (12)$$

where E_n is the transverse-state energy in the wire. As a result, we obtain the $\chi_n(q_3)$ longitudinal part of the wave function as a solution to the system of equations

$$\left[\frac{p^2}{2\mu} + E_n - \frac{\kappa^2}{8\mu}\right]\chi_n + \frac{1}{2\mu}\sum_{n'}[\zeta^2 (M^2)_{nn'} - 2\{\zeta, p\}(M)_{nn'} + 2\{\kappa \mathbf{l}, p^2\}(\mathbf{q}_\perp)_{nn'}]\chi_{n'} = E\chi_n. \quad (13)$$

Here and throughout, braces denote the symmetrization of operators,

$$\{A, B\} = \frac{1}{2}(AB + BA).$$

In one-dimensional equations, we will omit index 3 in coordinate and momentum denotations. In (13), the commutators of ζ, κ, and $\mathbf{l}$ with longitudinal momentum p are small. Nevertheless, we cannot ignore them, because the Hamiltonian should remain Hermitian.

First, consider a nondegenerate spectrum of transverse states. We will seek states formed from a certain transverse state n and assume that all other $\chi_{n'}$ ($n' \neq n$) values are small compared with χ_n. Accordingly, we will only leave the diagonal term in (13). The $(M)_{nn}$ diagonal matrix element vanishes, and (13) is simplified to

$$h_n \chi_n \equiv \left[\frac{p^2}{2\mu} - \overbrace{\frac{\kappa^2(q)}{8\mu}}^{a} + \overbrace{\frac{1}{2\mu}\beta_n \zeta^2(q)}^{b} + \overbrace{\frac{1}{\mu}\{\kappa \mathbf{l}, p^2\}(\mathbf{q}_\perp)_{nn}}^{c}\right]\chi_n = (E - E_n)\chi_n. \quad (14)$$

Here, $\beta_n = (M^2)_{nn}$. An additional criterion for the validity of the approximation that we use is fairly small values of the $\{\zeta, p\}(M)_{nn'}$ off-diagonal terms. We will show that this condition is satisfied if $pd \ll 1$, that is, in the vicinity of transverse quantization subband bottoms, but is violated as the longitudinal energy of electrons increases.

Term a in (14) is the well-known geometrical potential [11] caused exclusively by the curvature of the wire and independent of its internal structure.

Term b is determined by the internal and geometrical torsion of the wire. This term depends on the form of the transverse potential and transverse wave functions via the $(M^2)_{nn}$ matrix element. The value of this element is on the order of one if the cross section of the

wire is neither too prolate nor too symmetrical. This term is positive (a particle is repelled from the region where $|\zeta|$ is large), which distinguishes it from the geometrical potential, which attracts a particle to the region with maximum curvature. Term b vanishes if internal and geometrical torsion values are equal. In addition, it vanishes for a wire with a circular cross section for all nondegenerate states (states with zero momentum projections onto the axis), in particular, for the ground state.

Term c in (14) contains the diagonal matrix element that corresponds to the transverse coordinate $[(\mathbf{q}_\perp)_{nn}]$. If we are only interested in states in one (for instance, the lower) transverse quantization subband, $(\mathbf{q}_\perp)_{nn}$ can be reduced to zero by selecting the origin for $\mathbf{q}_\perp$. The $(\mathbf{q}_\perp)_{nn}$ matrix elements with all n are zero for $U(\mathbf{q}_\perp)$ potentials symmetrical with respect to rotations around axis 3 through $2\pi/N$ ($N = 2, 3, \ldots$) angles including axially symmetrical potentials. In particular, parabolic potentials and quantum wires with rectangular cross sections and hard walls have such symmetry properties.

In the more general case of a quantum wire with a nonsymmetrical potential, the c term does not vanish. This term is centrifugal in origin; that is, the centrifugal force drives an electron traveling along a curvilinear wire to the external side of a bent wire and thereby changes the potential energy of the electron in the nonsymmetrical well potential. We obtained a similar term for motion over a curvilinear surface [14].

This contribution depends on the longitudinal momentum of the electron and is a small addition to its kinetic energy. The c term is, however, significant if the longitudinal energy of the electron becomes fairly high. A comparison of terms a and c shows that c is larger than a if $p^2 \geq \kappa/d$. At low energies, the geometrical potential makes the major contribution. Note that term c is responsible for the dependence of the Hamiltonian on angle ϕ, that is, on the orientation of the cross section with respect to the normal to the curve.

Note that our result differs from that obtained in [11–13] by the presence of the b and c terms, which depend on the torsion of the wire, in addition to the geometrical term determined by curvature. Generally, these contributions cannot be ignored. Only if the wire has a circular cross section can the result be reduced to the geometrical potential (see below). The reason for the error is the unjustified neglect of the internal quantum wire structure in passing to the limit of an infinitely thin wire. The b and c terms, which depend on transverse wave functions, make the Hamiltonian not unambiguous in the limit of an infinitely thin wire. (However, the absence of a limit already follows from the divergence of the energy of transverse states.)

At a fairly high longitudinal energy, we must take into account terms linear in p in (13). Formally, these terms are largest in order of magnitude among all the terms caused by wire curvature and torsion. Nevertheless, the diagonal elements of this perturbation vanish, and the effective one-subband Schrödinger equation should be constructed in second-order perturbation theory.

Let us write the main terms of (13) in the form

$$(h_n + E_n - E)\chi_n = \sum_{n'} \frac{1}{\mu}\{\zeta, p\}(M)_{nn'}\chi_{n'}. \qquad (15)$$

Next, express small $\chi_{n'}$ ($n' \neq n$) components through χ_n, substitute them into the equation for χ_n, and then stop iterations,

$$\left[h_n - \frac{1}{\mu^2}\sum_{n'}\{\zeta, p\}\frac{|(M)_{nn'}|^2}{h_{n'} + E_{n'} - E}\{\zeta, p\}\right]\chi_n = (E - E_n)\chi_n. \qquad (16)$$

The $\zeta(q)$ function is a smooth function of coordinates. For this reason, the action of the h_n and $\{\zeta, p\}$ operators on the wave function only slightly changes the longitudinal momentum. The denominator in (16) can therefore be replaced by $E_n - E_{n'}$, which gives

$$H_n = \frac{1}{2\mu}\Big[p^2 - \frac{1}{4}\kappa^2 + \beta_n\zeta^2 - \overbrace{\gamma_n\{\zeta, p\}^2}^{d} + 2(\mathbf{q}_\perp)_{nn}\{\kappa\mathbf{l}, p^2\}\Big], \qquad (17)$$

$$H_n\chi_n = (E - E_n)\chi_n, \qquad (18)$$

where

$$\gamma_n = \frac{2}{\mu}\sum_{n' \neq n}\frac{|(M)_{nn'}|^2}{E_{n'} - E_n}. \qquad (19)$$

The β_n and γ_n constants are fully determined by the transverse potential. Generally, the β_n dimensionless value is on the order of 1, and the γ_n dimensional constant is of the order of the squared wire thickness. The term with γ_n gives a correction to the kinetic energy which is small in the $\zeta^2\gamma_n \sim (\zeta d)^2 \ll 1$ parameter. The ratio between this term and the contribution proportional to β_n is determined by the ratio between the $p^2/2\mu$ longitudinal electron energy and the distance between transverse levels.

In the particular case of a quantum wire with a symmetrical potential, the effective Hamiltonian can be written in the form

$$H_n = \frac{1}{2\mu}\Big[p^2 - \frac{1}{4}\kappa^2 + \beta_n\zeta^2 - \gamma_n\{\zeta, p\}^2\Big]. \qquad (20)$$

If the curvature and torsion ζ are constant, they shift subband bottoms by $(4\beta_n\zeta^2 - \kappa^2)/8\mu$ and change the

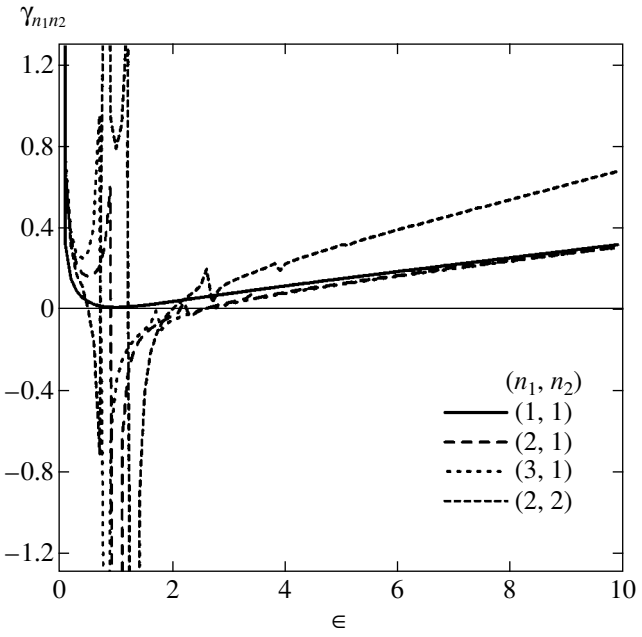


Fig. 3. Dependence of γ_n/ab on ϵ for various n.

effective mass of the electron, $1/\mu \longrightarrow 1/\mu(1 - \gamma_n\zeta^2)$. The sign of subband bottom shifts is determined by the ratio between the curvature and torsion values.

Another particular case is a twisted rectilinear wire. In such a wire, $\kappa = \tau = 0$. If the state in the vicinity of the band bottom is only considered, (20) can be used to obtain a one-dimensional Schrödinger equation with the effective potential $\beta_n\omega^2/2\mu$:

$$[p^2 + \beta_n\omega^2]\chi_n = 2\mu(E - E_n)\chi_n. \tag{21}$$

This result was obtained by us in [15].

The derived effective Hamiltonian allows the problem of electron movement in a quantum wire to be reduced to solving a purely one-dimensional Schrödinger equation with variable coefficients. The known, exactly solvable, one-dimensional problems can be used to consider special types of bent and/or twisted quantum wires.

4. THE β_n AND γ_n CONSTANTS FOR PARTICULAR QUANTUM WIRE POTENTIALS

Let us calculate the β_n and γ_n values for typical $U(\mathbf{q}_\perp)$ potentials. First, consider a potential of the rectangular box type $-a_i/2 < q_i < a_i/2$ with solid walls. The $n = (n_1, n_2)$ transverse states are characterized by two numbers $n_{1,2} = 1, 2, \ldots$, and

$$E_n = \frac{\pi^2}{2\mu}\left(\frac{n_1^2}{a_1^2} + \frac{n_2^2}{a_2^2}\right).$$

The β_n and γ_n coefficients are found in the form

$$\beta_n = \frac{1}{12}\left[\pi^2\left((\epsilon n_1)^2 + \left(\frac{n_2}{\epsilon}\right)^2\right) - 6\left(\left(\frac{\epsilon n_1}{n_2}\right)^2 + \left(\frac{n_2}{\epsilon n_1}\right)^2\right)\right], \tag{22}$$

$$\gamma_n = \frac{2^{12}ab}{\pi^6\epsilon}\sum_{n' \neq n} C(n_1, n_1')C(n_2, n_2')\frac{n_1^2 n_2^2 n_1'^2 n_2'^2}{(n_1^2 - n_1'^2)^4(n_2^2 - n_2'^2)^4} \times \frac{(n_2^2 - n_2'^2 - \epsilon^2(n_1^2 - n_1'^2))^2}{\epsilon^2(n_1^2 - n_1'^2) + n_2'^2 - n_2^2}, \tag{23}$$

where $\epsilon = a_2/a_1$ is the ratio between the sides of the rectangle and

$$C(n_1, n_1') = \frac{1 - (-1)^{(n_1 + n_1')}}{2}.$$

The dependence of $\gamma_{n_1n_2}/a_1a_2$ on parameter ϵ is shown in Fig. 3. The $\gamma_{n_1n_2}(\epsilon)$ value has the obvious symmetry property $\gamma_{n_1n_2}(\epsilon) = \gamma_{n_2n_1}(1/\epsilon)$. The $\gamma_{n_1n_2}(\epsilon)$ dependence for the $n = (1, 1)$ ground state does not have singularities. Resonances at higher levels correspond to the conditions of the arising of degenerate states,

$$\epsilon^2(n_1^2 + n_1'^2) + n_2^2 - n_2'^2 = 0.$$

Another example is the parabolic confining potential

$$U(\mathbf{q}_\perp) = \frac{\mu(\omega_1^2 q_1^2 + \omega_2^2 q_2^2)}{2}.$$

We then have

$$E_n = \omega_1\left(n_1 + \frac{1}{2}\right) + \omega_2\left(n_2 + \frac{1}{2}\right), \quad n_{1,2} = 0, 1, 2, \ldots$$

The β_n and γ_n values are given by

$$\beta_n = \frac{1}{4\omega_1\omega_2} \times [(2n_1 + 1)(2n_2 + 1)(\omega_1^2 + \omega_2^2) - 2\omega_1\omega_2], \tag{24}$$

$$\gamma_n = \frac{1}{2\mu\omega_1\omega_2} \times \left[(\omega_1 - \omega_2)^2\frac{n_1 + n_2 + 1}{\omega_1 + \omega_2} + (\omega_1 + \omega_2)^2\frac{n_2 - n_1}{\omega_1 - \omega_2}\right]. \tag{25}$$

The applicability of (23) and (25) is limited by the closeness of the denominators to zero.

5. A DEGENERATE SPECTRUM

If the spectrum of transverse states is degenerate, perturbation theory is inapplicable because the energy denominators then vanish. At the same time, the momentum projection M and $q_\perp$ operators give matrix elements between different states of this group that generally do not vanish. A degenerate spectrum arises, for instance, in quantum wires with a square section or in a parabolic potential with multiple frequencies.

In the case of degeneracy, terms that belong to the degenerate group only remain in Eq. (13), which can then be used as an effective Schrödinger equation.

A particular case of degeneracy is axial potential symmetry, $U(\mathbf{q}_\perp) = U(q_\perp)$. The selection of angle $\phi(q)$ is then arbitrary for symmetry reasons. For instance, ϕ can be set equal to zero. Clearly, the observed values should not depend on the selection of ϕ. Transverse motion states ψ_{nm} are classified according to the projection of momentum m values. All states except those with zero projections are doubly degenerate with respect to the sign of the momentum projection. The matrix element of $\mathbf{q}_\perp$ vanishes, and the Schrödinger equation for the states with a given momentum projection can then be diagonalized,

$$\left[\frac{(p-m\zeta)^2}{2\mu} - \frac{\kappa^2}{8\mu}\right]\chi_m = (E-E_n)\chi_m. \tag{26}$$

The $m\zeta$ value plays the role of a vector potential. Torsion can be excluded from this equation by passing to a new $u(q)$ wave function,

$$\chi_m = \exp\left(im\int\zeta dq\right)u(q),$$

$$\left[\frac{p^2}{2\mu} - \frac{\kappa^2}{8\mu}\right]u = (E-E_n)u.$$

The possibility of excluding torsion from (26) is directly related to the arbitrariness of selecting ϕ. Indeed, the geometrical and internal torsion terms enter into the equation as additive values, and if physical values do not depend on ϕ, they should also be independent of τ. It follows that the effective Schrödinger equation for a wire with a circular cross section reduces to the introduction of the geometrical potential. However, if the frame of reference is selected along geometrical directrices **n** and **b**, the wave function acquires an additional phase (Berry phase [16]), which is multiple to the angle of rotation of the geometrical directrices. Indeed, if a bent quantum wire has a circular cross section, this wire is nevertheless characterized by anisotropy imposed by the normal to the curve. Rotations of the normal along the wire are equivalent to rotations of wire directrices, which determine the Berry phase.

6. A PLANE QUANTUM WIRE

Most often, quantum wires formed on a plane surface have been studied. Such wires are described by plane curves, and one of the directrices of their cross section is the normal to the surface. This means that both τ and ω vanish, and Hamiltonian (20) only retains the geometrical potential,

$$H_n = \frac{1}{2\mu}\left[p^2 - \frac{1}{4}\kappa^2\right]. \tag{27}$$

By way of example, consider a plane snakelike curve $y = a\sin(bx)$. The curvature of the curve at point x is

$$\kappa = -\frac{ab^2\sin(bx)}{[1+a^2b^2\cos^2(bx)]^{3/2}}.$$

Let us return from q to Cartesian variable x. The Schrödinger equation is then obtained in the form

$$\frac{1}{\sqrt{1+a^2b^2\cos^2(bx)}}\frac{\partial}{\partial x}\frac{1}{\sqrt{1+a^2b^2\cos^2(bx)}}\frac{\partial}{\partial x}\chi + \frac{a^2b^4\sin^2(bx)}{4[1+a^2b^2\cos^2(bx)]^3}\chi = 2\mu(E_n-E)\chi. \tag{28}$$

On the assumption that the curve is not steep ($ab \ll 1$), we obtain the Mathieu equation in the principal order,

$$\chi'' + \left[2\mu(E-E_n) + \frac{a^2b^4}{2}(1-\cos(2bx))\right]\chi = 0. \tag{29}$$

7. A HELICAL QUANTUM WIRE

Next, consider a quantum wire that has the form of a helix wound on the surface of a cylinder with a constant pitch D. The $\mathbf{a}(q_3)$ function for such a helix has the form

$$\mathbf{a}(q_3) = \left(R\cos\frac{2\pi q_3}{S}, R\sin\frac{2\pi q_3}{S}, \frac{Dq_3}{S}\right), \tag{30}$$

where $S = \sqrt{(2\pi R)^2 + D^2}$. Using (30), we easily obtain equations for the curvature and geometrical torsion,

$$\kappa = R\left(\frac{2\pi}{S}\right)^2, \quad \tau = \frac{2\pi D}{S^2}. \tag{31}$$

In the nondegenerate case and in the absence of internal torsion ($\omega = 0$), the Hamiltonian takes the form

$$H_n = \frac{1}{2\mu}p^2[1 + 2|(\mathbf{q}_\perp)_{nn}|\kappa - \gamma_n\tau^2] + \frac{1}{2\mu}\left[\beta_n\tau^2 - \frac{1}{4}\kappa^2\right]. \tag{32}$$

Hamiltonian (32) is independent of coordinates, preserves the momentum, and immediately gives the

energy spectrum. The multiplier of p^2 determines the renormalized effective mass, but the corresponding corrections are small in adiabatic parameters. If the transverse potential has the symmetry properties specified above, the term with $|(\mathbf{q}_\perp)_{nn}|$ vanishes. The second term in the right-hand part of Eq. (32), which equals

$$\frac{1}{2\mu R^2}\frac{\beta_n(D/2\pi R)^2-1}{(1+(D/2\pi R)^2)^2},$$

is the shift of subband bottoms caused by curvature and torsion. This shift may be either positive or negative, and it depends on the number of the subband.

8. SPIN-ORBIT COUPLING

The initial spin-orbit coupling Hamiltonian for the conduction band has the form

$$\mathcal{H}^{SO} = -i\alpha[\boldsymbol{\sigma}\times\nabla U]\nabla. \tag{33}$$

Here, α is the effective spin-orbit coupling constant in a volume crystal. In semiconductors of the A_3B_5 type, this constant for the Γ_1 band valley is given by

$$\alpha = (2E_g m)^{-1}\frac{\Delta(2E_g+\Delta)}{(E_g+\Delta)(3E_g+2\Delta)}$$

(e.g., see [17]), where E_g is the forbidden bandwidth and Δ is the spin-orbit splitting of the valence band.

We will only consider nondegenerate transverse states. Let us transform Hamiltonian (33) and write it in the curvilinear system of coordinates introduced above. Averaging over transverse states then gives the effective spin-orbit coupling Hamiltonian.

After the transformation, Hamiltonian (33) in the space of Φ functions takes the form

$$\mathcal{H}^{SO} = \alpha\epsilon^{ijk}\frac{1}{G^{1/4}}\bar{\sigma}_j(\partial_k U)p_i\frac{1}{G^{1/4}}. \tag{34}$$

Here, the covariantly transformed Pauli matrices are determined by the equations

$$\bar{\sigma}_j = K_{ji}\sigma_i, \quad K_{ji} = \partial_j r_i. \tag{35}$$

Using (5) and (6), we obtain

$$\bar{\sigma}_j = \tilde{\sigma}_j + \zeta\delta_{j3}(q_1\tilde{\sigma}_2 - q_2\tilde{\sigma}_1), \quad \tilde{\sigma}_j = (n_j)_i\sigma_i. \tag{36}$$

For convenience, we introduced the $\mathbf{n}_3 \equiv \mathbf{t}$ vector in (36). The rotated $\tilde{\sigma}_j$ Pauli matrices are functions of only the longitudinal coordinate (q_3).

After expanding (34) in powers of $q_{1,2}$ and averaging over transverse states with number n, we obtain the effective spin-orbit coupling Hamiltonian in the principal order,

$$H_n^{SO} = \alpha[\zeta(A_1\tilde{\sigma}_1 + A_2\tilde{\sigma}_2) + \kappa\tilde{\sigma}_3(l_1A_2 - l_2A_1) + \{\kappa[\tilde{\sigma}_1(l_1A_{12} + l_2A_{22}) - \tilde{\sigma}_2(l_1A_{11} + l_2A_{21})], p\}]. \tag{37}$$

The A_i and A_{ij} values are determined by the values averaged over transverse wave functions ($\langle\ldots\rangle_n = (\ldots)_{nn}$),

$$A_1 = i\langle q_2((\partial_1 U)\partial_2 - (\partial_2 U)\partial_1)\rangle_n,$$
$$A_2 = -i\langle q_1((\partial_1 U)\partial_2 - (\partial_2 U)\partial_1)\rangle_n, \tag{38}$$

$$A_{ij} = \langle q_i\partial_j U\rangle_n. \tag{39}$$

The A_i values can be written differently without explicitly including potential U,

$$A_1 = -\frac{i}{2\mu}\langle\partial_2^2\partial_1\rangle_n, \quad A_2 = -\frac{i}{2\mu}\langle\partial_1^2\partial_2\rangle_n, \tag{40}$$

$$A_{11} = \frac{1}{2\mu}\Big\langle(1+2q_1\partial_1)(\partial_1^2+\partial_2^2)\Big\rangle_n,$$
$$A_{22} = \frac{1}{2\mu}\Big\langle(1+2q_2\partial_2)(\partial_1^2+\partial_2^2)\Big\rangle_n, \tag{41}$$

$$A_{12} = A_{21} = -\frac{1}{\mu}\Big\langle q_1\partial_2(\partial_1^2+\partial_2^2)\Big\rangle_n. \tag{42}$$

The order of magnitude of $\mathbf{A}$ and A_{ij} can be inferred from the last equations. Generally, $A_i \sim E_n^{1/2}\mu^{1/2}$ and $A_{ij} \sim E_n$.

Equation (37) can also be written in terms of the projections of the $\sigma_\mathbf{t}$, $\sigma_\mathbf{n}$, and $\sigma_\mathbf{b}$ Pauli matrices onto the unit vectors of the basis ($\mathbf{t}$, $\mathbf{n}$, $\mathbf{b}$),

$$H_n^{SO} = \alpha[\zeta(\mathbf{A}\cdot\mathbf{l}\sigma_\mathbf{n} - [\mathbf{A}\times\mathbf{l}]_3\sigma_\mathbf{b}) - \kappa[\mathbf{A}\times\mathbf{l}]_3\sigma_\mathbf{t} + \{\kappa[((l_1^2-l_2^2)A_{12} - l_1l_2(A_{11}-A_{22}))\sigma_\mathbf{n} - (l_1^2A_{11} + l_2^2A_{22} - 2l_1l_2A_{12})\sigma_\mathbf{b}], p\}]. \tag{43}$$

Spin-orbit coupling Hamiltonian (37) contains torsion only in combination with vector $\mathbf{A}$. Vector $\mathbf{A}$ vanishes if the $q_{1,2}$ variables in the $U(\mathbf{q}_\perp)$ potential are separable or if this potential possesses a center of inversion. Spin-orbit coupling is then determined exclusively by the curvature of the wire. The A_{12} tensor component vanishes if the potential has a symmetry plane ($q_1 \longleftrightarrow -q_1$ or $q_2 \longleftrightarrow -q_2$). The spin-orbit coupling Hamiltonian can then be written as

$$H_n^{SO} = -\frac{\alpha}{2}\{\kappa(A_{11}+A_{22})\sigma_\mathbf{b} + \kappa(A_{11}-A_{22})(\sigma_\mathbf{n}\sin 2\phi + \sigma_\mathbf{b}\cos 2\phi), p\}. \tag{44}$$

In particular, for nondegenerate subbands in an axially symmetrical potential (including the lowest subband), Hamiltonian (44) takes the simple form

$$H_n^{SO} = -\alpha A_{11}\{\kappa\sigma_{\mathbf{b}}, p\}. \quad (45)$$

9. SPIN-ORBIT COUPLING FOR PARTICULAR POTENTIALS

Let us calculate A_i and A_{ij} for the above-mentioned wire models; that is, for a harmonic potential, for a box with a rectangular cross section and hard walls, and for a cylindrical wire with hard walls. In these models, $A_1 = A_2 = A_{12} = 0$.

The A_{ii} values for a harmonic potential are easily calculated directly and are determined by the equation $A_{ii} = \omega_i(n_i + 1/2)$.

In order to determine matrix elements containing the derivatives of the potential for a well with infinite walls, we must accurately perform the passage to the limit of an infinite potential, because the wave function tends to zero as $U \longrightarrow \infty$ in the region of potential action. For this purpose, we can use the following technique.

Consider the generalized problem of a quantum wire with an arbitrary shape and hard walls. Let its boundary be described by the equation $f(\mathbf{q}_\perp) = 0$. For definiteness, we assume that $f(\mathbf{q}_\perp) < 0$ everywhere inside and $f(\mathbf{q}_\perp) > 0$ everywhere outside the wire. The potential of the walls will be selected in the form $U(\mathbf{q}_\perp) = U_0\theta(f(\mathbf{q}_\perp))$, where $U_0 \longrightarrow \infty$. We must find the matrix elements $(f_1(\mathbf{q}_\perp)\boldsymbol{\partial}U(\mathbf{q}_\perp))_{nn'}$, where $f_1(\mathbf{q}_\perp)$ is some function.

At large but finite U_0 values, the wave function behaves as $\exp(-k\xi)$, where $k = \sqrt{2\mu U_0}$ and ξ is the shortest distance from the boundary. It follows from the continuity of the wave function and its derivatives at the boundary that $\partial_\xi\psi/\psi|_{\xi=0} = -k$. When U_0 tends to infinity, $\psi(\xi = 0)$ tends to zero, whereas $\partial_\xi\psi|_{\xi=0}$ remains finite. Expressing the wave function through its derivative yields

$$(f_1(\mathbf{q}_\perp)\boldsymbol{\partial}U(\mathbf{q}_\perp))_{nn'} = -\frac{1}{2\mu}(\partial_\xi f_1(\mathbf{q}_\perp)(\boldsymbol{\partial} f(\mathbf{q}_\perp))\delta(f(\mathbf{q}_\perp))\partial_\xi)_{nn'}. \quad (46)$$

This formula can be treated as the replacement of the $f_1(\mathbf{q}_\perp)\boldsymbol{\partial}U(\mathbf{q}_\perp)$ perturbation by the operator

$$-\frac{1}{2\mu}(\mathbf{N}\cdot\boldsymbol{\partial})f_1(\boldsymbol{\partial}f)\delta(f)(\mathbf{N}\cdot\boldsymbol{\partial}),$$

where $\mathbf{N}(\mathbf{q}_\perp) = \boldsymbol{\partial}f/|\boldsymbol{\partial}f|$ is the normal to the boundary. The new form of perturbation is independent of the potential value at the boundary.

This technique can also be used to calculate the A_i and A_{ij} values according to (37). As a result, we obtain diagonal elements of the form

$$A_{ii} = \frac{\pi^2 n_i^2}{\mu a_i^2} \quad (47)$$

for a quantum wire with a rectangular cross section and hard walls.

The wave functions with $m = 0$ are nondegenerate for a cylindrical wire of radius R with hard walls. They are given by the Bessel functions $\psi_n(r) = CJ_0(\lambda_n r/R)$, where $r = |\mathbf{q}_\perp|$, λ_n are zeros of the zero-order Bessel function, and $E_n = \lambda_n^2/2\mu R^2$. For symmetry reasons, $A_{11} = A_{22}$. Using (39) and then (46), we find

$$A_{11} \equiv \frac{1}{2}(A_{11} + A_{22}) \equiv \frac{1}{2}\langle r(\partial_r U)\rangle_n \equiv \frac{1}{4\mu}R^2(\psi_n'(R))^2 = E_n. \quad (48)$$

Note in conclusion that, in this work, we presented adiabatic effective Hamiltonians for the motion of electrons in bent and twisted quantum wires. We considered the general case of a nonplanar quantum wire with a constant cross section, not necessarily circular. The problem was solved for a quadratic simple isotropic energy spectrum of electrons in the presence and absence of degeneracy of transverse states. In the derivation of the Hamiltonian, we did not restrict our consideration to states near transverse quantization subband bottoms. The analysis was performed taking into account spin-orbit coupling. Various examples of curved and twisted wires with different transverse potentials were considered.

ACKNOWLEDGMENTS

This work was financially supported by the Russian Foundation for Basic Research (project nos. 02-02-16377 and 00-02-17658).

REFERENCES

1. V. Ya. Prinz, V. A. Seleznev, and A. K. Gutakovsky, in *Proceedings of the 24th International Conference on Physics of Semiconductors*, Ed. by D. Gershoni (World Sci., Singapore, 1999); V. Ya. Prinz, V. A. Seleznev, A. K. Gutakovsky, *et al.*, Physica E (Amsterdam) **6**, 828 (2000).
2. V. Ya. Prinz *et al.*, Inst. Phys. Conf. Ser., No. 166, 203 (2000).
3. V. Ya. Prinz, D. Grützmacher, A. Bayer, *et al.*, Nanotechnology **12**, S1 (2001).
4. O. G. Schmidt and K. Eberl, Nature **410**, 168 (2001).
5. M. L. Leadbeater, C. L. Foden, T. M. Burke, *et al.*, J. Phys.: Condens. Matter **57**, L307 (1995); M. L. Lead-

beater, C. L. Foden, J. H. Burroughes, *et al.*, Phys. Rev. B **52**, 8629 (1995).
6. C. L. Foden, M. L. Leadbeater, and M. Pepper, Phys. Rev. B **52**, 8646 (1995).
7. A. A. Bykov, A. K. Bakarov, A. V. Goran, *et al.*, Pis'ma Zh. Éksp. Teor. Fiz. **74**, 182 (2001) [JETP Lett. **74**, 164 (2001)].
8. L. I. Magarill, D. A. Romanov, and A. V. Chaplik, Zh. Éksp. Teor. Fiz. **113**, 1411 (1998) [JETP **86**, 771 (1998)]; L. I. Magarill and A. V. Chaplik, Zh. Éksp. Teor. Fiz. **115**, 1478 (1999) [JETP **88**, 815 (1999)].
9. A. I. Vedernikov and A. V. Chaplik, Zh. Éksp. Teor. Fiz. **117**, 449 (2000) [JETP **90**, 397 (2000)].
10. V. M. Osadchiĭ and V. Ya. Prinz, Pis'ma Zh. Éksp. Teor. Fiz. **72**, 451 (2000) [JETP Lett. **72**, 312 (2000)].
11. R. C. T. da Costa, Phys. Rev. A **23**, 1982 (1981).
12. H. Jensen and H. Koppe, Ann. Phys. (N.Y.) **46**, 586 (1971).
13. N. Ogawa, K. Fujii, and A. Kobushukin, Prog. Theor. Phys. **83**, 894 (1990).
14. M. V. Entin and L. I. Magarill, Phys. Rev. B **64**, 085330 (2001).
15. M. V. Entin and L. I. Magarill, Phys. Rev. B **66**, 205308 (2002).
16. M. V. Berry, Proc. R. Soc. London, Ser. A **392**, 45 (1984).
17. V. F. Gantmakher and I. B. Levinson, *Scattering of Carriers in Metals and Semiconductors* (Nauka, Moscow, 1984).

Translated by V. Sipachev