

Fermion Zero Modes in Painlevé–Gullstrand Black Hole¹

G. E. Volovik

Low Temperature Laboratory, Helsinki University of Technology, FIN-02015 HUT, Finland
Landau Institute for Theoretical Physics, Russian Academy of Sciences, Moscow, 117940 Russia

Received May 14, 2001

The Painlevé–Gullstrand metric of a black hole allows one to discuss the fermion zero modes inside the hole. The statistical mechanics of the fermionic microstates can be responsible for the black hole thermodynamics. These fermion zero modes also lead to quantization of the horizon area. © 2001 MAIK “Nauka/Interperiodica”.

PACS numbers: 04.70.Dy; 05.30.Fk

1. INTRODUCTION

In general relativity, there are different nonequivalent metrics $g_{\mu\nu}$, which describe the same gravitational object. Although they can be obtained from each other by coordinate transformations, in the presence of an event horizon they produce nonequivalent quantum vacua. Among the other metrics used for description of a black hole, the Painlevé–Gullstrand metric [1] has many advantages [2, 3] and now has become more popularity (see, e.g., [4]; for the extension of the Painlevé–Gullstrand metric to a rotating black hole see [5]). Also, such a stationary but not static metric naturally arises in the condensed matter analogues of gravity [6–10]. Here, using the Painlevé–Gullstrand metric, we consider the structure of low-energy fermionic microstates in the interior of the black hole and their contribution to the black hole thermodynamics.

The interval in the Painlevé–Gullstrand space–time is

$$ds^2 = -(c^2 - v_s^2(r))dt^2 - 2v_s(r)drdt + dr^2 + r^2d\Omega^2. \quad (1)$$

In acoustic black and white holes, $v_s(r)$ is the radial velocity of fluid, which produces the effective metric for acoustic waves—phonons—propagating in liquid [6–8]. For the gravitational field produced by a point source of mass M , the function $v_s(r)$ has the form

$$v_s(r) = \pm c\sqrt{r_h/r}, \quad r_h = 2MG, \quad (2)$$

where r_h is the radius of horizon, G is the Newton gravitational constant, and $c = \hbar = 1$. The Painlevé–Gullstrand metric breaks the time reversal symmetry; the time reversal operation transforms a black hole to a

white hole (see also [3]). The minus sign in Eq. (2) gives the metric for the black hole. In the fluid analogue of gravity, this corresponds to the liquid flowing inward. The plus sign characterizes the white hole and, correspondingly, the flow outward in the fluid analogy. The time reversal operation reverses the direction of flow.

The Painlevé–Gullstrand metric describes the space–time both in exterior and in interior regions. This space–time, though not static, is stationary. That is why the energy in the interior region is well determined. Moreover, in contrast to the Schwarzschild metric, the particle energy spectrum $E(\mathbf{p})$ (the solution of equation $g^{\mu\nu}p_\mu p_\nu + m^2 = 0$, where $p_0 = -E$) is well defined for any value of the momentum \mathbf{p} . This allows us to determine the ground state (vacuum) of the Standard Model in the interior region and the thermal states—the black hole matter. We consider here only the fermionic vacuum of the Standard Model, assuming that the Standard Model is an effective theory and, thus, the bosonic fields are the collective modes of the fermionic vacuum.

2. FERMI SURFACE FOR STANDARD MODEL FERMIONS INSIDE HORIZON

We will see that the main contribution to the thermodynamics of a black hole comes from very short wavelengths on order of the Planck length. That is why the ultraviolet cutoff must be introduced. We introduce the cutoff using the nonlinear dispersion of a particle spectrum in the ultrarelativistic region, which violates the Lorentz symmetry at short distances. Such nonlinear dispersion of the particle spectrum is now frequently used both in black hole physics and in cosmology [11–12]. We will use the superluminal dispersion, for which the particle velocity becomes superluminal at very high momentum. For the simplest superluminal

¹ This article was submitted by the author in English.

dispersion, the energy spectrum of a fermionic particle in the Painlevé–Gullstrand metric becomes

$$E = p_r v_s \pm c \sqrt{p^2 + p^4/p_0^2}, \quad (3)$$

where p_r is the radial momentum of particle; p_0 plays the role of the cutoff momentum, which is somewhat less than the Planck momentum p_{Planck} ; and we neglected all masses of the Standard Model fermions, since they are much less than the characteristic energy scales.

Because of the possibility of superluminal propagation, the surface $r = r_h$ is not the true horizon. This surface marks the boundary of the ergoregion: at $r < r_h$, particles with positive square root in Eq. (3) can acquire negative energy. As a result, at $r < r_h$ the Fermi surface appears—the surface in the 3D momentum space, where the energy of particles is zero, $E(\mathbf{p}) = 0$. For the spectrum in Eq. (3), the surface is given by the equation which expresses the radial momentum in terms of the transverse momentum p_\perp :

$$\begin{aligned} p_r^2(p_\perp) &= \frac{1}{2} p_0^2 (v_s^2 - 1) \\ -p_\perp^2 &\pm \sqrt{\frac{1}{4} p_0^4 (v_s^2 - 1)^2 - p_0^2 v_s^2 p_\perp^2}. \end{aligned} \quad (4)$$

This surface exists at each point \mathbf{r} within the horizon (ergosurface), where $v_s^2 > 1$. It exists only in the restricted range of the transverse momenta, with the restriction provided by the cutoff parameter p_0 :

$$p_\perp < \frac{1}{2} p_0 \left| v_s - \frac{1}{v_s} \right|. \quad (5)$$

This means that the Fermi surface is a closed surface in the 3D momentum space \mathbf{p} .

The Fermi surface provides the finite density of fermionic states (DOS) at $E = 0$

$$\begin{aligned} N(E = 0) &= N_F \sum_{\mathbf{p}, \mathbf{r}} \delta(E(\mathbf{p})) \\ &= \frac{4\pi N_F}{(2\pi)^3} \int_0^{r_h} r^2 dr \int d^3 p \delta \left(p_r v_s + c \sqrt{p^2 + \frac{p^4}{p_0^2}} \right) \\ &= \frac{N_F}{\pi} \int_0^{r_h} r^2 dr \int_0^{p_\perp^2(r)} \frac{d(p_\perp^2)}{|v_G|}. \end{aligned} \quad (6)$$

Here, $N_F = 16N_g$ is the number of massless chiral fermionic species in the Standard Model with N_g generations, and v_G is the radial component of the group velocity of particles at the Fermi surface:

$$v_G(E = 0) = \frac{dE}{dp_r} = \mp \sqrt{\left(v_s - \frac{1}{v_s} \right)^2 - 4 \frac{p_\perp^2}{p_0^2}}. \quad (7)$$

Integration over p_\perp^2 in Eq. (6) gives for the density of states

$$N(E = 0) = \frac{N_F p_0^2}{\pi} \int_0^{r_h} r^2 dr \left| v_s - \frac{1}{v_s} \right| = \frac{4N_F}{35\pi} p_0^2 r_h^3. \quad (8)$$

The main contribution to DOS and, thus, to the thermodynamics comes from the momenta p comparable with the cutoff momentum p_0 . That is why all masses of fermions were neglected.

The DOS $N(E = 0)$ determines the thermodynamics of black hole matter at $T \neq 0$. The thermal energy $\mathcal{E}(T)$ and entropy $\mathcal{S}(T)$ carried by the Standard Model fermions in the interior of the black hole with nonzero temperature is

$$\begin{aligned} \mathcal{E}(T) &= N(0) \int dE E f(E/T) = \frac{\pi^2}{6} N(0) T^2, \\ \mathcal{S}(T) &= \frac{\pi^2}{3} N(0) T, \end{aligned} \quad (9)$$

where $f(x) = 1/(e^x + 1)$ is the Fermi distribution function.

We considered a large temperature, as compared to the Hawking temperature $T_H = \hbar c/4\pi r_h$ [13]. At lower energies, the discrete nature of the spectrum of Standard Model fermions bound to the black hole becomes important.

At $T \sim T_H$, the entropy becomes of the order of

$$\mathcal{S}(T \sim T_H) \sim N_F p_0^2 r_h^2. \quad (10)$$

The cutoff momentum p_0 can be expressed in terms of the effective gravitational constant G , which is determined by the same cutoff according to the Sakharov induced gravity [14]. The effective action for gravity is obtained by the integration over vacuum fermions, and thus all fermionic species must add to produce the effective Newton constant: $G^{-1} \sim p_0^2 N_F$. Thus, the thermal entropy in Eq. (10) is scaled as G^{-1} , i.e., $\mathcal{S}(T \sim T_H) \sim r_h^2/G$. The same occurs with the Bekenstein–Hawking entropy of a black hole, $S_{\text{BH}} = \pi r_h^2/G$. As was first shown by Jacobson, S_{BH} is renormalized by the same quantum fluctuations as G^{-1} and, thus, is proportional to G^{-1} [15]. Thus, the thermal entropy of the Standard Model fermionic microstates within the black hole at $T \sim T_H$ has the same behavior and the same order of magnitude as the Bekenstein–Hawking entropy of the black hole.

3. DISCRETE ENERGY LEVELS INSIDE HORIZON

We now turn to low energy, where quantization is important and gives discrete energy levels for the Stan-

dard Model fermions within the horizon. Since the momenta of particles are large compared to the size of the horizon, one can use the quasiclassical approximation for the radial motion and the Bohr–Sommerfeld quantization rule. We consider here low-energy states whose energy E is much less than the characteristic energy scale of a Fermi liquid: $E \ll p_0 c$. In this limit, the classical trajectories, which determine the Bohr–Sommerfeld quantization, can be obtained by perturbation theory. Let us start with the zero-order trajectories, i.e., trajectories with $E = 0$. After quantization of the azimuthal motion, one obtains the following dependence of the radial momentum p_r on r , which determines the classical trajectories along the radius at a given value of angular momentum L [compare this with Eq. (4)]:

$$\begin{aligned}
 p_r^2(r, E = 0, L) &= \frac{1}{2} p_0^2 (v_s^2 - 1) \\
 &- \frac{L^2}{r^2} \pm \sqrt{\frac{1}{4} p_0^4 (v_s^2 - 1)^2 - p_0^2 v_s^2 \frac{L^2}{r^2}}.
 \end{aligned} \quad (11)$$

Since for the typical bound states one has $L \gg 1$, the difference between expressions $L(L+1)$, $(L + \frac{1}{2})^2$, and L^2 is not important.

The trajectories in Eq. (11) are closed; there are two turning points on each trajectory. A particle moves back and forth between the zeroes r_1 and r_2 of the square root on the right-hand side of Eq. (11):

$$\left| v_s(r_{1,2}) - \frac{1}{v_s(r_{1,2})} \right| = \frac{2L}{p_0 r_{1,2}}. \quad (12)$$

At the turning points, the group velocity of the particle v_G in Eq. (7) becomes zero and changes to the opposite sign. This is very similar to the Andreev reflection [16]: the velocity changes sign after reflection, while the momentum p_r does not.

For nonzero but small energy, $E \ll p_0 c$, the trajectories are obtained by perturbation theory. The first-order correction gives

$$\begin{aligned}
 p_r(r, E, L) &= p_r(r, E = 0, L) + \frac{dp_r}{dE} E \\
 &= p_r(r, E = 0, L) + \frac{E}{v_G(E = 0)}.
 \end{aligned} \quad (13)$$

The Bohr–Sommerfeld quantization gives

$$2\pi(n_r + \gamma(L)) = \oint dr p_r(r, E = 0, L) + E \oint \frac{dr}{v_G}, \quad (14)$$

where n_r is the radial quantum number and $\gamma(L)$ is a parameter of order unity, which is not determined within this quantization scheme. Numerical integration

of $\oint dr p_r(r, E = 0, L)$ shows that it is very close to the following equation:¹

$$\begin{aligned}
 \oint dr p_r(r, E = 0, L) &= \pm \frac{3\sqrt{3}\pi}{2} (L_{\max} - L), \\
 L_{\max} &= \frac{p_0 r_h}{3\sqrt{3}}.
 \end{aligned} \quad (15)$$

As a result, one obtains the following equidistant energy levels for each L :

$$E = \omega_0(L) \left(n_r + \gamma(L) \mp \frac{3\sqrt{3}}{4} (L_{\max} - L) \right), \quad (16)$$

$$\frac{\pi}{\omega_0(L)} = \int_{r_1}^{r_2} dr \frac{1}{\sqrt{(v_s - 1/v_s)^2 - 4L^2/r^2 p_0^2}}. \quad (17)$$

In two limiting cases, the interlevel distance

$$\begin{aligned}
 \omega_0(L \ll p_0 r_h) &\approx \frac{\pi \hbar c}{r_h \ln(p_0 r_h / L)} = T_H \frac{4\pi^2}{\ln(p_0 r_h / L)}, \\
 \omega_0(L \rightarrow L_{\max}) &= \frac{3\hbar c}{r_h} = 12\pi T_H.
 \end{aligned} \quad (18)$$

Thus, for every spherical harmonic L , L_r , there are bound states in the black hole interior whose energy as function of the radial quantum number n_r crosses the zero-energy level at $n_r \approx \pm(3\sqrt{3}/4)(L_{\max} - L)$. These are the branches of the fermion zero modes. The total number of such branches is

$$\begin{aligned}
 N_{zm} &= 2N_F \sum_{L < L_{\max}} (2L + 1) \\
 &\approx 2N_F L_{\max}^2 = \frac{2}{27} N_F p_0^2 r_h^2 \sim \frac{A}{G},
 \end{aligned} \quad (19)$$

where $A = 4\pi r_h^2$ is the area of the black hole horizon. The estimation of the density of states remains the same as in Eq. (8), which was obtained within the Fermi-surface approach:

$$\begin{aligned}
 N(0) &= 2N_F \sum_L \frac{2L + 1}{\omega_0(L)} = 2N_F \int_0^{\infty} d(L^2) \frac{1}{\omega_0(L)} \\
 &= \frac{N_F p_0^2 r_h}{\pi} \int_0^{r_h} r^2 dr \left| v_s - \frac{1}{v_s} \right| = \frac{4N_F}{35\pi} p_0^2 r_h^3.
 \end{aligned} \quad (20)$$

¹ I am indebted to V.B. Eltsov for these calculations.

4. FERMION ZERO MODES: VORTEX VS. BLACK HOLE

The energy spectrum of the ultrarelativistic fermions within the black hole in Eq. (16) resembles the spectrum of fermionic bound state within the core of vortices in Fermi superfluids and superconductors. The energy levels are also equidistant there [17]. The energy spectrum of the fermion zero modes in the vortex core depends on two quantum numbers that are appropriate for the states within the linear object: linear and angular momenta along the vortex axis:

$$E(L_z, p_z) = \omega_0(p_z)(L_z + \gamma), \quad (21)$$

where parameter γ is either 0 or 1/2 depending on the type of vortex (see review paper [8]). For vortices with $\gamma = 0$, the energy levels with $L_z = 0$ have exactly zero energy. For such a vortex, the entropy is nonzero even at $T = 0$. Each of the states with $E = 0$ can be either free or occupied by a fermion. This gives the zero-temperature entropy $\ln 2$ per each $E = 0$ state with given p_z . Thus, the total entropy of the vortex at $T = 0$ is proportional to the length l of the vortex line:

$$\mathcal{S}(T = 0) = N_{zm} \ln 2, \quad N_{zm} \propto p_0 l. \quad (22)$$

Here, $p_0 = p_F$ is the Fermi momentum of the Fermi superfluid or superconductor; it plays the role of a cut-off parameter. It is interesting that, as in the case of the black hole, the wavelengths of fermions comprising the fermion zero modes in the vortex core are much shorter than the core size. This allows us to use the quasiclassical theory. However, even within the quasiclassical theory, one can, using the symmetry or other arguments, find the value of phase shift γ in the Bohr–Sommerfeld quantization scheme [8] and predict for which vortex the system of the equidistant levels of fermions contains the states with exactly zero energy.

For the fermionic states bound to the black hole, the parameter $\gamma(L)$ in Eq. (16) is still unknown. That is why one cannot say whether the system of equidistant levels contains a level with zero energy or not. If so, then each state with zero energy contributes to the entropy $\ln 2$; the total entropy provided by the fermion zero mode at $T = 0$ is

$$\mathcal{S}(T = 0) = N_{zm} \ln 2. \quad (23)$$

5. DISCUSSION

From Eq. (19), it follows that for the Painlevé–Gullstrand black hole the area of the black hole horizon is expressed in terms of the integer valued quantity

$$Ap_{\text{Planck}}^2 = \sigma \mathcal{N}, \quad (24)$$

where \mathcal{N} is the number of fermion zero modes within the black hole: $\mathcal{N} = N_{zm}$; and σ is of order unity. This formula with different values of parameter σ was discussed in many modern theories of black holes (see

[18, 19] and references therein). It was interpreted as a quantization of the horizon area, with \mathcal{N} being the quantum number which characterizes the black hole as an “atom” [18]. If one uses $\sigma = 4\ln 2$ as in [20], one obtains the Bekenstein–Hawking entropy in Eq. (23). \mathcal{N} was also interpreted as the number of “constituents” of the black hole interior—“gravitational atoms” [21]. In our case, both interpretations apply, though with some reservation.

The quantization of area in Eq. (24) usually suggests that the spacing between the levels is uniform and on the order of $dM/d\mathcal{N} \propto E_{\text{Planck}}^2/M$ [18]. This is in agreement with Eq. (18) for the interlevel distance ω_0 in the fermionic spectrum. However, the quantization in terms of the number of fermion zero modes suggests another possible interpretation:

The area A of the black hole is a continuous parameter. When it changes, the number of fermion zero modes $N_{zm}(A)$ as a function of A changes in a stepwise manner at some critical values of A . This is what happens, say, in the integer quantum Hall effect, where the integer topological charge \mathcal{N} of quantum vacuum as a function of external parameters has plateaus. If the external parameter is a magnetic field B , then $\mathcal{N}(B)$ and the Hall conductivity $\sigma_{xy}(B) = (e^2/h)\mathcal{N}(B)$ change abruptly when the critical values of magnetic field B are crossed. Similar behavior of the topological charge \mathcal{N} of quantum vacuum occurs in other quasi-2D fermionic systems as well; e.g., the momentum–space topological charge \mathcal{N} of a film of quantum liquid is a stepwise function of a continuous parameter—thickness of the film (see [22, Chapter 9]).

\mathcal{N} in Eq. (24) can also be related to the number of constituents, as suggested in [21]. According to the Fermi liquid description, the number of thermal fermions in a Fermi liquid at temperature T is $N_{\text{thermal}} \sim N(0)T \sim A/G$. According to Eq. (9), each fermion carries energy of order T . At $T \sim T_H$, their total thermal energy is on the order of the mass M of the black hole, and they carry thermal entropy of the order of the Hawking–Bekenstein entropy S_{BH} . Assuming that the whole mass M of the black hole comes from thermal fermions within the horizon, one has $M = E + 3pV$, where $V = (4\pi/3)r_h^3$ is the volume within the horizon and p is the pressure of the fermionic system. Using the equation of state of thermal fermions forming the Fermi surface, $E = ST/2 = pV$, which follows from Eq. (9), one obtains the correct relation between the mass, Hawking temperature, and Bekenstein entropy of the black hole:

$$M = E + 3pV = 4E = 2TS. \quad (25)$$

The Bekenstein–Hawking entropy $S = \pi r_h^2$ and the Hawking temperature $T_H = 1/4\pi r_h$ are reproduced if the relation between the cutoff p_0 and the Newton constant is $G^{-1} = N_F p_0^2/105\pi$. Thus, the black hole fermionic

matter at the Hawking temperature can provide the mass and the entropy of the black hole; those fermions thermally excited within the horizon can serve as constituents.

These constituents do not actually represent the gravitational atoms which form the quantum vacuum and give rise to the phenomenon of gravitation according to [23]. These are conventional elementary particles of the Standard Model (quarks and leptons) that are excited within the black hole. Their contribution is essential even at a temperature as low as T_H because of the huge DOS within the black hole.

On the other hand, these constituents have little to do with the matter absorbed by the black hole during its formation. The fermions, which form the matter within the black holes, all are fermions of the Standard Model, quarks and leptons, which are highly ultrarelativistic. The black hole metric emerging after collapse perturbs significantly the spectrum of Standard Model fermions, so that the Fermi surface appears which provides a huge DOS at zero energy. A substantial part of these fermions have momenta on the order of the Planck scale; to them, the effective gravitational theory probably does not apply. In this sense, these fermions are close to gravitational atoms of trans-Planckian physics.

We considered the vacuum of the Standard Model fermions and their thermal states as viewed in the Painlevé–Gullstrand metric. This vacuum is substantially different from the vacuum state as viewed by a comoving observer. The reconstruction of the vacuum within the black hole involves the Planck energy scale, and the results depend on the cutoff procedure. The cutoff procedure, on the other hand, depends on the coordinate system used, and it assumes the existence of a preferred coordinate frame at high energy. That is why the vacuum structure depends on the coordinate system.

Since the Planck energy scale is involved, it is not clear whether the traditional description of a black hole is applicable. Moreover, the stability of this new vacuum is not guaranteed. In most of those condensed matter systems, where an analogue of the event horizon is possible, the vacuum becomes unstable in the presence of a horizon; i.e., the quantum vacuum of the condensed matter resists the formation of a horizon [8, 24]. Also, the huge density of states may generate symmetry breaking in the black hole interior, as happens in the core of vortices [25] and cosmic strings [26].

Even if the black hole survives under such reconstruction of the Standard Model vacuum, there is another problem to be solved. When thermal states of the fermionic black hole matter are considered, their energy and pressure must serve as a source of gravitational field according to (maybe somewhat modified) Einstein equations. This will change the field v_s which enters the black hole metric, and, thus, the energy spectrum will be modified, but remain equidistant for each L .

In conclusion, we considered the statistical mechanics of fermionic microstates—the Standard Model fermions—in the interior of a black hole. Fermion zero

modes give a correct dependence of the entropy of the Painlevé–Gullstrand black hole on the area of the hole, on the number of fermionic species, and on the Planck cutoff parameter. They also lead to quantization of the horizon area. That is why fermion zero modes can be the true microstates that are responsible for the black hole thermodynamics.

I thank T. Jacobson, P. Mazur, R. Schützhold, A. Starobinsky, W. Unruh, and M. Visser for illuminating discussions and criticism. This work was supported in part by the Russian Foundation for Basic Research and the European Science Foundation.

REFERENCES

1. P. Painlevé, C. R. Hebd. Seances Acad. Sci. **173**, 677 (1921); A. Gullstrand, Ark. Mat. Astron. Fys. **16** (8), 1 (1922).
2. K. Martel and E. Poisson, Am. J. Phys. **69**, 476 (2001).
3. R. Schützhold, Phys. Rev. D **64**, 024029 (2001).
4. M. K. Parikh and F. Wilczek, Phys. Rev. Lett. **85**, 5042 (2000).
5. C. Doran, Phys. Rev. D **61**, 067503 (2000).
6. W. G. Unruh, Phys. Rev. Lett. **46**, 1351 (1981); Phys. Rev. D **51**, 2827 (1995).
7. M. Visser, Class. Quantum Grav. **15**, 1767 (1998).
8. G. E. Volovik, gr-qc/0005091.
9. T. Jacobson and G. E. Volovik, Phys. Rev. D **58**, 064021 (1998).
10. M. Stone, cond-mat/0012316.
11. S. Corley and T. Jacobson, Phys. Rev. D **54**, 1568 (1996); S. Corley, Phys. Rev. D **57**, 6280 (1998); S. Corley and T. Jacobson, Phys. Rev. D **59**, 124011 (1999).
12. A. Starobinsky, Pis'ma Zh. Éksp. Teor. Fiz. **73**, 415 (2001) [JETP Lett. **73**, 371 (2001)].
13. S. W. Hawking, Nature **248**, 30 (1974).
14. A. D. Sakharov, Dokl. Akad. Nauk SSSR **177**, 70 (1967) [Sov. Phys. Dokl. **12**, 1040 (1968)].
15. T. Jacobson, gr-qc/9404039.
16. A. F. Andreev, Zh. Éksp. Teor. Fiz. **46**, 1823 (1964) [Sov. Phys. JETP **19**, 1228 (1964)].
17. C. Caroli, P. G. de Gennes, and J. Matricon, Phys. Lett. **9**, 307 (1964).
18. J. D. Bekenstein, in *Proceedings of the Eighth Marcel Grossman Meeting*, Ed. by Tsvi Piran (World Scientific, Singapore, 1999); gr-qc/9710076.
19. H. A. Kastrup, Phys. Lett. B **413**, 267 (1997).
20. V. F. Mukhanov, Pis'ma Zh. Éksp. Teor. Fiz. **44**, 50 (1986) [JETP Lett. **44**, 63 (1986)].
21. P. O. Mazur, in *Proceedings of the Eighth Marcel Grossman Meeting*, Ed. by Tsvi Piran (World Scientific, Singapore, 1999); hep-th/9712208.
22. G. E. Volovik, *Exotic Properties of Superfluid ^3He* (World Scientific, Singapore, 1992).
23. P. O. Mazur, hep-th/9708133.
24. S. Liberati, S. Sonego, and M. Visser, Class. Quantum Grav. **17**, 2903 (2000).
25. Yu. G. Makhlin and G. E. Volovik, Pis'ma Zh. Éksp. Teor. Fiz. **62**, 719 (1995) [JETP Lett. **62**, 737 (1995)].
26. S. G. Naculich, Phys. Rev. Lett. **75**, 998 (1995).

Photon Damping Caused by Electron–Positron Pair Production in a Strong Magnetic Field

N. V. Mikheev and N. V. Chistyakov*

Yaroslavl State University, ul. Sovetskaya 14, Yaroslavl, 150000 Russia

* e-mail: mch@uniyar.ac.ru

Received May 8, 2001

Damping of an electromagnetic wave in a strong magnetic field is analyzed in the kinematic region near the threshold of electron–positron pair production. Damping of the electromagnetic field is shown to be noticeably nonexponential in this region. The resulting width of the photon $\gamma \rightarrow e^+e^-$ decay is considerably smaller than previously known results. © 2001 MAIK “Nauka/Interperiodica”.

PACS numbers: 13.40.Hq; 95.30.Cq

The problem of propagation of electromagnetic fields through an active medium is inherent in a variety of physical phenomena. The birth and evolution of supernova and neutron stars, where the matter density can be on the order of nuclear density $\rho \sim 10^{14}–10^{15}$ g/cm³ and the temperature can achieve several tens of megaelectronvolts, are the largest scale and the most interesting such phenomena. In addition to dense and hot matter, a strong magnetic field [1, 2], which can be several orders of magnitude higher than the so-called critical, or Schwinger, value $B_e = m_e^2/e \approx 4.41 \times 10^{13}$ G, can be generated in the above-mentioned objects.¹ This strong magnetic field can induce new phenomena which can considerably affect the evolution of these astrophysical objects. Electromagnetic-field damping caused by electron–positron pair production in an external magnetic field is one of these phenomena. Recall that the $\gamma \rightarrow e^+e^-$ process is kinematically forbidden in vacuum. The magnetic field changes the kinematics of charged particles, electrons and positrons, allowing the production of an electron–positron pair in the kinematic region $q_{\parallel}^2 = q_0^2 - q_3^2 \geq 4m_e^2$, where q_0 is the photon energy and q_3 is the momentum component along the magnetic field.²

In 1954, Klepikov [3] examined the production of an electron–positron pair by a photon in a magnetic field and obtained the amplitude and width of the $\gamma \rightarrow e^+e^-$ decay in the semiclassical approximation. Later, the authors of [4–9] considered this process in the context of its astrophysical applications. It was pointed out in [7, 8] that the use of the expression derived in [3] for the width considerably overestimates the result in the strong magnetic field limit. In this case, one should use

an exact expression for the width of one-photon production of a pair when electrons and positrons occupy only the ground Landau level. However, it was found that the expression for the decay width in the limit of a strong magnetic field has a root singularity at the point $q_{\parallel}^2 = 4m_e^2$. Shabad [9] emphasized that this behavior indicates that the decay width calculated in perturbation theory cannot be treated as a damping coefficient. In this case, the damping coefficient is primarily determined from the time evolution of the photon wave function in the presence of a magnetic field. Shabad [9] suggested that this dependence be obtained by solving the dispersion equation with account taken of the vacuum polarization in a magnetic field with complex values of photon energy. In our opinion, this method has several disadvantages. First, it is well known but rarely mentioned that the dispersion equations with complex energies have no solutions in the physical sheet. Solutions are in the nonphysical Riemannian sheets (analyticity region of the polarization operator), which are generally infinite in number. As a result, the dispersion equation has an infinite number of solutions with both positive and negative imaginary parts of energy. The physical status of these solutions requires a separate investigation.

Shabad [9] used the asymptotic form of the polarization operator near the pair production threshold and erroneously treated it as a two-sheet complex function. This circumstance led to the existence of two complex conjugate solutions, one of which is physically meaningless because it has a positive imaginary part and, therefore, provides exponentially increasing amplitude of electromagnetic wave. Therefore, to obtain a physically meaningful result, one should artificially discard the redundant solutions.

¹ We use the system of units where $\hbar = c = 1$.

² Hereafter, we consider the magnetic field directed along the third axis.

Second, this approach cannot correctly describe the substantially nonexponential damping in the near-threshold region in a strong field.

Thus, damped electromagnetic waves in a magnetic field cannot be completely described by solving the dispersion equation.

In this work, we use a method that is applied in the field theory at finite temperatures and in plasma physics (see, e.g., [10]). It consists of the determination of a retarded solution to the electromagnetic field equation that includes an external source and takes into account the vacuum polarization in a magnetic field. Time damping of the electromagnetic wave is analyzed in a uniform external magnetic field, whose intensity is the largest parameter of the problem, $B_e \gg q^2, m_e^2$.

To describe the time evolution of electromagnetic wave $\mathcal{A}_\alpha(x)$ [$x_\mu = (t, \mathbf{x})$], we consider a linear response of the system ($\mathcal{A}_\alpha(x)$ and a vacuum polarized in a magnetic field) to an external source, which is adiabatically turned on at $t = -\infty$ and turned off at $t = 0$. At $t > 0$, the electromagnetic wave evolves independently. Thus, the source is necessary for creating an initial state. For simplicity, we consider the evolution of a monochromatic wave. In this case, the source function should be taken in the form

$$\mathcal{J}_\alpha(x) = j_\alpha e^{i\mathbf{kx}} e^{\varepsilon t} \theta(-t), \quad \varepsilon \rightarrow 0^+. \quad (1)$$

The time dependence of $\mathcal{A}_\alpha(x)$ is determined by the equation

$$(g_{\alpha\beta} \partial_\mu^2 - \partial_\alpha \partial_\beta) \mathcal{A}_\beta(x) + \int d^4 x' \mathcal{P}_{\alpha\beta}(x-x') \mathcal{A}_\beta(x') = \mathcal{J}_\alpha(x), \quad (2)$$

where $\mathcal{P}_{\alpha\beta}(x-x')$ is the photon polarization operator in a magnetic field. We note that, for the source on the right-hand side of Eq. (2) to be conserved, $\partial_\alpha \mathcal{J}_\alpha = 0$, the current j_α must have the form $j_\alpha = (0, \mathbf{j})$, $\mathbf{j} \cdot \mathbf{k} = 0$. The evolution of $\mathcal{A}_\alpha(x)$ is described by the retarded solution

$$\mathcal{A}_\alpha^R(x) = \int d^4 x' G_{\alpha\beta}^R(x-x') \mathcal{J}_\beta(x'), \quad (3)$$

where $G_{\alpha\beta}^R(x-x')$ is the retarded Green's function, which is defined through the commutator of the Heisenberg operators $\hat{A}_\alpha(x)$ of the electromagnetic field as (see, e.g., [11]),

$$G_{\alpha\beta}^R(x-x') = -i \langle 0 | [\hat{A}_\alpha(x), \hat{A}_\beta(x')] | 0 \rangle \theta(t-t'). \quad (4)$$

It is instructive to express this function in terms of the causal Green's function

$$G_{\alpha\beta}^C(x-x') = -i \langle 0 | \mathbf{T} \hat{A}_\alpha(x) \hat{A}_\beta(x') | 0 \rangle, \quad (5)$$

by using the relationship

$$G_{\alpha\beta}^R(x-x') = 2 \operatorname{Re} G_{\alpha\beta}^C(x-x') \theta(t-t'). \quad (6)$$

In the presence of a magnetic field, it is convenient to decompose Green's function (5) in the eigenvectors of polarization operator [9]:

$$G_{\alpha\beta}^C(x) = \int \frac{d^4 q}{(2\pi)^4} G_{\alpha\beta}^C(q) e^{-iqx}, \quad (7)$$

$$G_{\alpha\beta}^C(q) = \sum_{\lambda=1}^3 \frac{b_\alpha^{(\lambda)} b_\beta^{(\lambda)}}{(b^{(\lambda)})^2 q^2 - \mathcal{P}^{(\lambda)}(q)}, \quad (8)$$

where $\mathcal{P}^{(\lambda)}(q)$ are the eigenvalues of polarization operator. The eigenvectors $b_\alpha^{(\lambda)}$

$$b_\alpha^{(1)} = (q\varphi)_\alpha,$$

$$b_\alpha^{(2)} = (q\tilde{\varphi})_\alpha, \quad (9)$$

$$b_\alpha^{(3)} = q^2 (q\varphi\varphi)_\alpha - (q\varphi\varphi)q_\alpha,$$

together with the 4-vector q_α , form a complete orthogonal basis in the Minkowski 4-space. In Eqs. (9), $\varphi_{\alpha\beta} = F_{\alpha\beta}/B$ and $\tilde{\varphi}_{\alpha\beta} = \frac{1}{2} \varepsilon_{\alpha\beta\mu\nu} \varphi_{\mu\nu}$ are a dimensionless magnetic-field tensor and dual tensor, respectively, $(q\varphi)_\alpha = q_\sigma \varphi_{\sigma\alpha}$, $(q\varphi\varphi) = q_\alpha \varphi_{\alpha\beta} \varphi_{\beta\sigma} q_\sigma$. Substituting Eqs. (1) and (6) into Eq. (3) and using Eqs. (7) and (8), we obtain after simple integration the following result at $t > 0$:

$$\begin{aligned} \mathcal{A}_\alpha^R(x) &= \sum_{\lambda=1}^3 V_\alpha^{(\lambda)}(x) = 2e^{i\mathbf{kx}} \\ &\times \operatorname{Re} \sum_{\lambda=1}^3 \int \frac{dq_0}{2\pi i} \frac{b_\alpha^{(\lambda)}(b^{(\lambda)}j)/(b^{(\lambda)})^2 e^{-iq_0 t}}{(q_0 - i\varepsilon)(q_0^2 - \mathbf{k}^2 - \mathcal{P}^{(\lambda)}(q))}, \end{aligned} \quad (10)$$

where $q_\alpha = (q_0, \mathbf{k})$. Note that the definition of the integral in Eq. (10) should be completed because the integrand can include singularities, which are due, on the one hand, to the zeros of its denominator and, on the other, to the domain of its definition. To analyze these singularities, it is necessary to know the explicit form and analytical properties of the eigenvalues $\mathcal{P}^{(\lambda)}(q)$ of the polarization operator, which was examined in detail in a number of studies. In the limit of a strong magnetic field, the functions $\mathcal{P}^{(\lambda)}(q)$, which we are interested in, can be borrowed, e.g., from [9, 12, 13] and represented as [with $O(1/B)$ accuracy]

$$\mathcal{P}^{(1)}(q) \approx -\frac{\alpha}{3\pi} q_\perp^2 - q^2 \Lambda(B, q^2), \quad (11)$$

$$\mathcal{P}^{(3)}(q) \approx -q^2 \Lambda(B, q^2), \quad (12)$$

$$\begin{aligned} \mathcal{P}^{(2)}(q) \approx &-\frac{2\alpha e B}{\pi} \left(\frac{1}{\sqrt{z}(1-z)} \arctan \sqrt{\frac{z}{1-z}} - 1 \right) \\ &- q^2 \Lambda(B, q^2), \end{aligned} \quad (13)$$

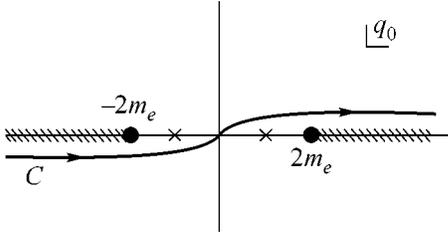


Fig. 1. The path of integration C in the complex q_0 plane. The crosses are the poles corresponding to the real solutions of dispersion Eq. (14). The shaded parts of the real axis are cuts.

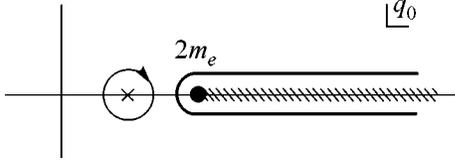


Fig. 2. The path of integration C after the transformation allowing one to separate the pole $F_{\text{pole}}(t)$ and cut $F_{\text{cut}}(t)$ contributions.

where

$$\Lambda(B, q^2) = \frac{\alpha}{3\pi} [1.792 - \ln(B/B_e)] + \pi(q^2),$$

$$z = q_{\parallel}^2/4m_e^2, \quad q_{\perp}^2 = q_1^2 + q_2^2,$$

and $q^2\pi(q^2)$ is the photon polarization operator in the absence of a magnetic field. Note that the contribution from the pole $q_{\parallel}^2 = 0$ that results from the normalization of the basis vectors $b_{\alpha}^{(2)}$ and $b_{\alpha}^{(3)}$ is nonphysical and, taking into account explicit form (11)–(13) of the polarization operator, can be removed by gauge transformation after summation over polarizations. Thus, the contribution to the solution can be made only by the poles corresponding to the dispersion equation

$$q^2 - \mathcal{P}^{(\lambda)}(q) = 0. \quad (14)$$

Using solution (10), one can demonstrate on the basis of Eqs. (11)–(13) that only two modes, $\lambda = 1$ and 2 with the polarization vectors

$$\epsilon_{\alpha}^{(1)} = \frac{b_{\alpha}^{(1)}}{\sqrt{(b^{(1)})^2}} = \frac{(q\varphi)_{\alpha}}{\sqrt{q_{\perp}^2}},$$

$$\epsilon_{\alpha}^{(2)} = \frac{b_{\alpha}^{(2)}}{\sqrt{(b^{(2)})^2}} = \frac{(q\tilde{\varphi})_{\alpha}}{\sqrt{q_{\parallel}^2}},$$
(15)

are physically meaningful for real photons.³ A photon of the third, $\lambda = 3$, mode is nonphysical [9]. Indeed,

³ Modes with the polarizations $\epsilon_{\alpha}^{(1)}$ and $\epsilon_{\alpha}^{(2)}$ correspond to so-called parallel (\parallel) and perpendicular (\perp) modes, respectively, in the Adler notation [14].

substitution of the expression for $\mathcal{P}^{(3)}(q)$ into Eq. (14) gives an equation that has the only solution $q^2 = 0$. Therefore, the contribution of the third mode to solution (10) is proportional to the total divergence and can be eliminated by the corresponding gauge transformation.

In the limit of a strong magnetic field, only the mode with the polarization vector $\epsilon_{\mu}^{(2)}$ can decay into an electron–positron pair, because only the eigenvalue of the polarization operator $\mathcal{P}^{(2)}(q)$ (13) has an imaginary part at $q_{\parallel}^2 \geq 4m_e^2$. Therefore, to analyze time damping of the electromagnetic field, it is sufficient to consider only the term with $V_{\alpha}^{(2)}(x)$ in Eq. (10).

The further calculations can be simplified by going over to the reference frame where $\mathbf{k} = (k_1, k_2, 0)$, which can always be done without disturbing the properties of the external magnetic field. In this frame, $q_{\parallel}^2 = q_0^2$ and the polarization vector of the second mode takes the form $\epsilon_{\alpha}^{(2)} = (0, 0, 0, -1)$. As a result, $V_{\alpha}^{(2)}(x)$ is expressed as

$$V_{\alpha}^{(2)}(x) = V_{\alpha}^{(2)}(0, \mathbf{x}) \frac{\text{Re}F(t)}{\text{Re}F(0)}, \quad (16)$$

where

$$F(t) = \int_C \frac{dq_0}{2\pi i} \frac{e^{-iq_0 t}}{(q_0 - i\epsilon)(q_0^2 - \mathbf{k}^2 - \mathcal{P}^{(2)}(q))}, \quad (17)$$

$$V_{\alpha}^{(2)}(0, \mathbf{x}) = 2\epsilon_{\alpha}^{(2)} j_3 e^{i\mathbf{k}\mathbf{x}} \text{Re}F(0).$$

The path of integration C in Eq. (17) is determined by the analytical properties of $\mathcal{P}^{(2)}(q)$ and is shown in Fig. 1. The function $\mathcal{P}^{(2)}(q)$ is analytical in the complex plane q_0 with cuts along the real axis (see Fig. 1). In the kinematic region $|q_0| < 2m_e$, the eigenvalue $\mathcal{P}^{(2)}(q)$ is real and Eq. (14) has real solutions which govern the photon dispersion in this region.

For further analysis, it is convenient to transform the path of integration to the path shown in Fig. 2. In this case, the integral in Eq. (17) is represented as

$$F(t) = F_{\text{pole}}(t) + F_{\text{cut}}(t), \quad (18)$$

where the first term is determined by the residue at the point $q_0 = \omega$, which is the solution to dispersion Eq. (14). This term corresponds to the undamped solution in the region $\omega < 2m_e$ [9]. The second term determines the time dependence of the electromagnetic field above the threshold of electron–positron pair production and has the form of the Fourier integral

$$F_{\text{cut}}(t) = \int_{-\infty}^{\infty} \frac{dq_0}{2\pi} F_{\text{cut}}(q_0) e^{-iq_0 t}, \quad (19)$$

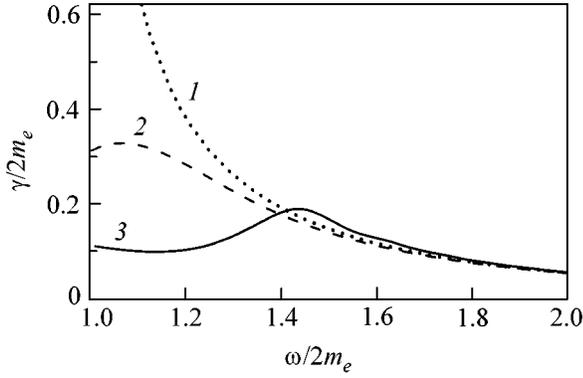


Fig. 3. The frequency dependence of the $\gamma \rightarrow e^+e^-$ decay width in the near-threshold region for the magnetic field $B = 200B_e$. Line 1 is the tree approximation including the root singularity; line 2 is obtained from the complex solution of the dispersion equation in the second Riemannian sheet [9]; and line 3 is γ_{eff} from approximation (23).

$$F_{\text{cut}}(q_0) = \frac{2\theta(q_0 - 2m_e)I}{q_0([q_0^2 - \mathbf{k}^2 - R]^2 + I^2)}, \quad (20)$$

$$R \equiv \text{Re}\mathcal{P}^{(2)}(q_0) = \frac{\alpha}{\pi}eB \left(\frac{1}{\sqrt{z(z-1)}} \ln \frac{\sqrt{z} + \sqrt{z-1}}{\sqrt{z} - \sqrt{z-1}} + 2 \right), \quad (21)$$

$$I \equiv -\text{Im}\mathcal{P}^{(2)}(q_0 + i\epsilon) = \frac{\alpha eB}{\sqrt{z(z-1)}}, \quad z = \frac{q_0^2}{4m_e^2}. \quad (22)$$

Expressions (19)–(22) together with Eq. (16) determine the time evolution of the photon wave function above the pair production threshold in a strong magnetic field.

Strictly speaking, because of the threshold behavior of the Fourier transform $F_{\text{cut}}(q_0)$, time damping of the function $F_{\text{cut}}(t)$ and, therefore, of the wave function $\mathcal{A}_\mu(t)$ is nonexponential. However, in some characteristic time interval (the inverse effective width of the $\gamma \rightarrow e^+e^-$ decay can naturally be chosen as such an interval), the time dependence of the wave function can approximately be represented as exponentially damped harmonic oscillations:

$$\mathcal{A}_\mu(t) \sim e^{-\gamma_{\text{eff}}t/2} \cos(\omega_{\text{eff}}t + \phi_0). \quad (23)$$

Here, ω_{eff} and γ_{eff} are, respectively, the effective frequency and width of photon decay, which should be found by using Eqs. (19)–(22) for each value of momentum \mathbf{k} to determine the effective photon dispersion law above the threshold of electron-positron pair production.

The quantity γ_{eff} , which governs the intensity of photon absorption due to e^+e^- pair production in a magnetic field, is important for astrophysical applications. The absorption coefficient obtained from the $\gamma \rightarrow e^+e^-$

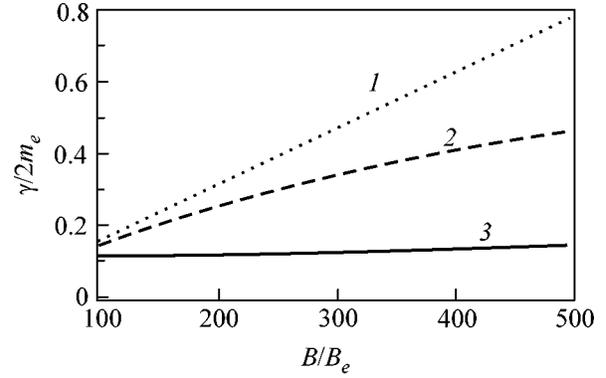


Fig. 4. The decay width vs. the magnetic field for the frequency $\omega = 2.5m_e$. The meaning of lines is the same as in Fig. 3.

decay probability and containing a root singularity is usually employed in astrophysics (see, e.g., [15]). Shabad [9] pointed out that this leads to the overestimation of the intensity of e^+e^- pair production. Our analysis demonstrates that the calculation of the absorption coefficient (decay width) by using the complex solution in the second Riemannian sheet [9] also leads to a considerable overestimation in the near-threshold region, as is seen from Figs. 3 and 4.

Nonexponential damping in the near-threshold region is known for the processes in vacuum and matter [16, 17]; however, as far as we know, it has not been considered in an external field so far. In contrast to vacuum or medium, the near-threshold effect in the magnetic field is kinematically enhanced due to the singular behavior of the polarization operator in this field. Therefore, this phenomenon is not only topical for astrophysical application but is of conceptual interest.

We are grateful to S.S. Gershtein, G.P. Pron'ko, V.V. Kiselev, and A.K. Likhoded for stimulating discussion and useful remarks.

This work was supported in part by the Ministry for Education of the Russian Federation (project no. E00-11.0-5) and the Russian Foundation for Basic Research (project no. 01-02-17334).

REFERENCES

1. G. S. Bisnovatyĭ-Kogan and S. G. Moiseenko, *Astron. Zh.* **69**, 563 (1992) [*Sov. Astron.* **36**, 285 (1992)]; G. S. Bisnovatyĭ-Kogan, *Astron. Astrophys. Trans.* **3**, 287 (1993).
2. R. C. Duncan and C. Thompson, *Astrophys. J. Lett.* **392**, L9 (1992); C. Thompson and R. C. Duncan, *Mon. Not. R. Astron. Soc.* **275**, 255 (1995).
3. N. P. Klepikov, *Zh. Ėksp. Teor. Fiz.* **26**, 19 (1954).
4. P. A. Sturrock, *Astrophys. J.* **164**, 529 (1971).
5. E. Tadermaru, *Astrophys. J.* **183**, 625 (1973).
6. M. A. Ruderman and P. S. Sutherland, *Astrophys. J.* **196**, 51 (1975).

7. V. S. Beskin, *Astrofizika* **18**, 439 (1982).
8. J. K. Daugherty and A. K. Harding, *Astrophys. J.* **273**, 761 (1983).
9. A. E. Shabad, *Tr. Fiz. Inst. Akad. Nauk SSSR* **192**, 5 (1988).
10. D. Boyanovsky, H. J. de Vega, Y. J. Ng, *et al.*, *Phys. Rev. D* **59**, 105001 (1999).
11. L. D. Landau and E. M. Lifshitz, *Course of Theoretical Physics*, Vol. 5: *Statistical Physics* (Nauka, Moscow, 1981; Pergamon, Oxford, 1980), Part 2.
12. W. Tsai and T. Erber, *Phys. Rev. D* **12**, 1132 (1975).
13. D. B. Melrose and R. J. Stoneham, *Nuovo Cimento A* **32**, 435 (1976).
14. S. L. Adler, *Ann. Phys. (N.Y.)* **67**, 599 (1971).
15. A. K. Harding, M. G. Baring, and P. L. Gonthier, *Astrophys. J.* **476**, 246 (1997).
16. L. A. Khal'fin, *Zh. Éksp. Teor. Fiz.* **33**, 1371 (1957) [*Sov. Phys. JETP* **6**, 1053 (1958)].
17. I. Joichi, Sh. Matsumoto, and M. Yoshimura, *Phys. Rev. A* **57**, 798 (1998).

Translated by R. Tyapaev

Second Optical Harmonic Generation in Nonlinear Crystals with a Disordered Domain Structure

E. Yu. Morozov¹, A. A. Kaminskiĭ², A. S. Chirkin^{1,*}, and D. B. Yusupov³

¹ Moscow State University, Moscow, 119899 Russia

*e-mail: chirkin@squeeze.phys.msu.su

² Shubnikov Institute of Crystallography, Russian Academy of Sciences, Moscow, 117333 Russia

³ Tashkent State Aircraft Institute, Tashkent, 702132 Uzbekistan

Received May 29, 2001

A theory describing the second harmonic generation in nonlinear-optical crystals with random domains is developed. The theory takes into account the fluctuations of both the phase mismatch and the nonlinear coupling coefficient of the interacting waves. It is shown that, in such crystals, the maximal efficiency of fundamental-to-second harmonic conversion is low, and the dependence of the average intensity of the second harmonic on the fundamental intensity differs considerably from a square law. © 2001 MAIK "Nauka/Interperiodica".

PACS numbers: 42.65.Ky; 42.25.Dd; 42.70.Mp

This paper presents the theory of second optical harmonic generation (SHG) in disordered polydomain structures and interprets on its basis the behavior of the second harmonic intensity observed in experiments on some ferroelectric crystals with random needlelike domains (microdomains), such as $\text{Sr}_x\text{Ba}_{1-x}(\text{NbO}_3)_2 : \text{Nd}^{3+}$ ($x \cong 0.6$) and $\text{Ba}_2\text{NaNb}_5\text{O}_{15} : \text{Nd}^{3+}$ [1–3]. In these crystals, the dependence of the second harmonic intensity on the intensity of the exciting laser radiation was found to differ considerably from a quadratic dependence and approach a linear one [1], while the conversion efficiency was fairly low. The cited publications [1–3] provided no explanation for this unexpected behavior of the second harmonic intensity. It should be noted that earlier experimental studies of the SHG in optically nonlinear inhomogeneous media (see, e.g., [4]) did not reveal any deviation of the dependence of the second harmonic intensity on the fundamental intensity from the square law that is characteristic of a low conversion efficiency in homogeneous crystals. The experimental data were explained in terms of the SHG theory that was developed in the given-field [4] or given-intensity [5] approximations.

Below, we develop the theory of the SHG in statistically inhomogeneous nonlinear media by going beyond the limits of the aforementioned approximations. We take into account the fluctuations of the phase difference and the fluctuations of the coefficient of nonlinear wave coupling. We establish that there is a region of the problem parameters for which the dependence of the second harmonic intensity on the fundamental intensity is virtually linear. Qualitatively, the essence of the effect can be explained as follows. In stochastically

inhomogeneous nonlinear media with a low efficiency of the fundamental-to-second harmonic conversion, the intensity of the second harmonic is proportional to the squared intensity of the fundamental radiation, the length of the medium, and the so-called coherent length. The latter proves to be dependent on the fundamental radiation intensity, and, therefore, the dependence of the second harmonic intensity on the pumping intensity differs from quadratic.

We describe the SHG process by the following system of reduced equations (compare with [4]):

$$\begin{aligned} dA_2/dz &= -ig(z)\beta A_1^2(z)\exp[-i\Phi(z)], \\ dA_1/dz &= -ig(z)\beta A_2(z)A_1^*(z)\exp[i\Phi(z)], \end{aligned} \quad (1)$$

where $A_1(z)$ and $A_2(z)$ are the complex amplitudes of the fundamental wave and the second harmonic, respectively, and β is the absolute value of the coefficient of nonlinear wave coupling. The function $g(z)$ describes the modulation of the nonlinear coefficient; we assume that this function has the form of a random telegraph process taking the values $+1$ and -1 with equal probability ($g^2(z) \equiv 1$). This model corresponds to a crystal with 180° random domains.

The function $\Phi(z)$ takes into account the phase shift due to the phase mismatch of the interacting waves:

$$\Phi(z) = \int_0^z (2k_1(x) - k_2(x)) dx = \Delta_0 z + \int_0^z \Delta(x) dx, \quad (2)$$

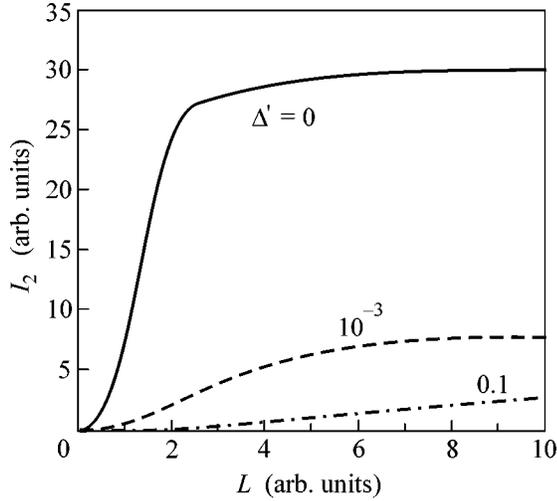


Fig. 1. Relative intensity of the second harmonic (normalized to the fundamental intensity $I_0 = 10 \text{ MW/cm}^2$) versus the reduced interaction length $L = z/L_{nl}$ for $\alpha' = 0.1$ and for different values of Δ' (the values are indicated in the plot).

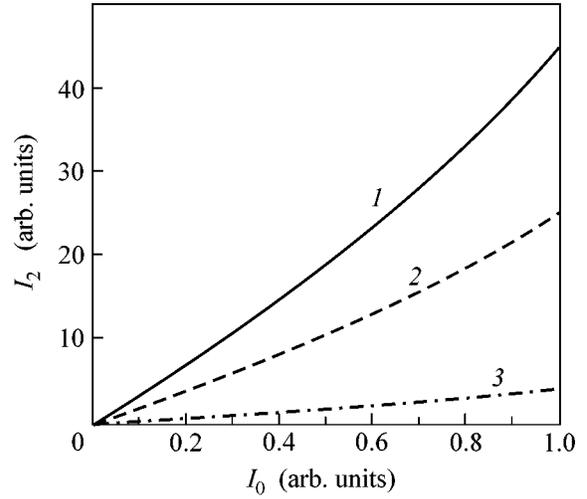


Fig. 2. Relative intensity of the second harmonic versus the fundamental intensity (both are normalized to the intensity $I_0 = 10 \text{ MW/cm}^2$) for $\alpha' = 0.1$ and for different values of Δ' : (1) 0, (2) 1×10^{-3} , and (3) 0.1.

where $k_j(z)$ is the wave number at the corresponding frequency and Δ_0 and $\Delta(z)$ are the regular and fluctuating parts of the phase mismatch. We assume that $\Delta(z)$ is a δ -correlated Gaussian process satisfying the conditions

$$\begin{aligned} \langle \Delta(z) \rangle &= 0, \\ \langle \Delta(z)\Delta(z') \rangle &= B(z, z') = 2K\delta(z-z'). \end{aligned} \quad (3)$$

The system of Eqs. (1) can be represented in the form

$$\begin{aligned} \frac{dI_2}{dz} &= ig(z)\beta U_-, \quad \frac{dI_1}{dz} = -ig(z)\beta U_-, \\ \frac{dU_-}{dz} &= i\frac{d\Phi(z)}{dz}U_+(z) + i4\beta g(z)I_1I_2 - i2\beta g(z)I_1^2, \end{aligned} \quad (4)$$

$$\frac{dU_+}{dz} = i\frac{d\Phi(z)}{dz}U_-(z),$$

where I_1 and I_2 are the intensities of the fundamental wave and the second harmonic, respectively; $I_j = |A_j|^2$. In Eqs. (4), we used the notation

$$U_-(z) = A_2 A_1^{*2} \exp[i\Phi(z)] - \text{complex conjugate},$$

$$U_+(z) = A_2 A_1^{*2} \exp[i\Phi(z)] + \text{complex conjugate}.$$

Let us average Eqs. (4) over the realizations of the random processes $g(z)$ and $\Delta(z)$, which are assumed to be uncorrelated. For this purpose, it is necessary to determine the correlators of the type of $\langle g(z)U_{\mp} \rangle$ and

$\langle \Delta(z)U_{\mp} \rangle$. For the random telegraph process $g(z)$, the following relation is valid [6]:

$$\begin{aligned} \frac{d}{dz} \langle g(z)U_{\mp}(z) \rangle &= -2\nu \langle g(z)U_{\mp}(z) \rangle + \left\langle g(z) \frac{dU_{\mp}(z)}{dz} \right\rangle, \end{aligned} \quad (5)$$

where the quantity ν^{-1} has the meaning of a ‘‘mean modulation scale,’’ i.e., a mean domain size.

For the random Gaussian process $\Delta(z)$, the Furutsu-Novikov formula is appropriate [6]:

$$\langle \Delta(z)F(z) \rangle = \int_0^z B(z, z') \left\langle \frac{\delta F(z')}{\delta \Delta(z')} \right\rangle dz'. \quad (6)$$

Taking into account Eqs. (3), one obtains

$$\langle \Delta(z)F(z) \rangle = K \left\langle \frac{\delta F(z)}{\delta \Delta(z)} \right\rangle. \quad (7)$$

Using Eqs. (4) and (5), one has

$$\begin{aligned} \frac{d}{dz} \langle g(z)U_- \rangle &= -2\nu \langle g(z)U_- \rangle \\ &+ \langle ig(z)[\Delta_0 + \Delta(z)]U_+ + i4\beta I_1I_2 - i2\beta I_1^2 \rangle, \\ \frac{d}{dz} \langle g(z)U_+ \rangle &= -2\nu \langle g(z)U_+ \rangle \\ &+ \langle ig(z)[\Delta_0 + \Delta(z)]U_- \rangle. \end{aligned} \quad (8)$$

According to Eqs. (7) and (8), one has for the correlator $\langle \Delta(z)g(z)U_{\pm} \rangle$

$$\begin{aligned} & \langle \Delta(z)g(z)U_{\pm}(z) \rangle \\ &= K \left\langle \frac{\delta(g(z)U_{\pm}(z))}{\delta\Delta(z)} \right\rangle = iK \langle g(z)U_{\mp}(z) \rangle. \end{aligned} \quad (9)$$

Introducing the notation $\langle g(z)U_{-}(z) \rangle = i\psi_1$, $\langle g(z)U_{+}(z) \rangle = \psi_2$, and $2\nu + K = \alpha$, one finally obtains the averaged system of equations in the form (compare with [5])

$$\begin{aligned} \frac{d\bar{I}_2}{dz} &= -\beta\psi_1, & \frac{d\bar{I}_1}{dz} &= \beta\psi_1, \\ \frac{d\psi_1}{dz} &= -\alpha\psi_1 + \Delta_0\psi_2 + 4\beta_1\langle I_1I_2 \rangle - 2\beta_2\langle I_1^2 \rangle, & (10) \\ \frac{d\psi_2}{dz} &= -\alpha\psi_2 - \Delta_0\psi_1. \end{aligned}$$

It is reasonable to assume that the field statistics for the second harmonic excited in a disordered nonlinear medium is Gaussian. Then, taking into account the relation $I_1(z) + I_2(z) = I_{10}$ (I_{10} is the intensity of the fundamental wave at the medium input), one can write

$$\begin{aligned} \langle I_1I_2 \rangle &= I_{10}\bar{I}_2 - 2\bar{I}_2^2, \\ \langle I_1^2 \rangle &= I_{10}^2 - 2I_{10}\bar{I}_2 + 2\bar{I}_2^2. \end{aligned} \quad (11)$$

Substituting Eqs. (11) in Eqs. (10), one obtains a closed system of equations.

For a numerical solution of Eqs. (10), we change to dimensionless quantities:

$$\begin{aligned} x &= \frac{\bar{I}_2}{I_{10}}, & y_i &= \frac{\Psi_i}{I_{10}^{3/2}}, & \zeta &= \frac{z}{L_{nl}}, \\ \alpha' &= \alpha L_{nl}, & \Delta'_0 &= \Delta_0 L_{nl}, \end{aligned}$$

where $L_{nl} = 1/\beta\sqrt{I_{10}}$ is the so-called characteristic nonlinear length. As a result, the system of Eqs. (10) can be reduced to the form

$$\begin{aligned} \frac{dx}{d\zeta} &= -y_1, \\ \frac{dy_1}{d\zeta} &= -\alpha'y_1 + \Delta'_0 y_2 + (8x - 12x^2 - 2), & (12) \\ \frac{dy_2}{d\zeta} &= -\alpha'y_2 - \Delta'_0 y_1. \end{aligned}$$

Note that the parameter α' characterizes the degree of disorder of the crystal structure within the nonlinear length, and the parameter Δ'_0 characterizes the phase mismatch within this length.

The system of Eqs. (12) was solved numerically for different values of parameters α' and Δ'_0 and for differ-

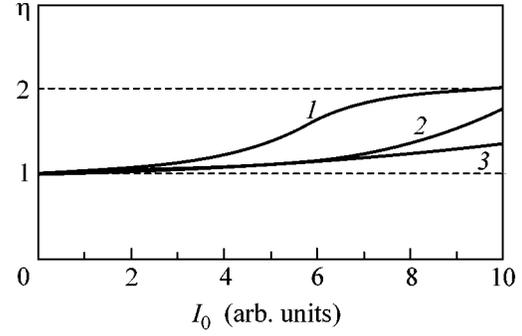


Fig. 3. The exponent characterizing the dependence of the second harmonic intensity on the fundamental intensity as a function of the fundamental intensity. The curves are calculated for the intensities normalized to $I_0 = 100 \text{ MW/cm}^2$ with the parameter α' and the values of Δ' corresponding to the curves shown in Fig. 2.

ent fundamental intensities I_{10} . The results of calculations are presented in Figs. 1–3. From Fig. 1, it follows that the maximal efficiency of the fundamental-to-second harmonic conversion in crystals with a disordered domain structure is small and saturated within several nonlinear lengths of the wave interaction. Figure 2 shows the dependences of the second harmonic intensity on the intensity of fundamental radiation. Such dependences were experimentally studied in [1–3]. The curves shown in Fig. 2 are plotted for those values of parameter α' for which the dependence of the mean intensity of the second harmonic on the pumping intensity is close to linear. Thus, Fig. 3 allows one to explain the linear dependence of the second harmonic intensity on the fundamental intensity that was experimentally observed for crystals with a disordered domain structure [1].

The differential exponent of the nonlinear dependence of \bar{I}_2 on I_{10} ($\eta = d\ln I_2/d\ln I_0$) is represented in Fig. 3. One can see that, in the region of high pumping intensities, this dependence is close to quadratic, as in the case of homogeneous crystals with a low efficiency of conversion to the second harmonic. At the same time, there is an interval of fundamental intensities within which the dependence of the mean intensity of the second harmonic on the fundamental intensity is close to linear. Thus, Fig. 3 allows one to explain the linear dependence of the second harmonic intensity on the fundamental intensity that was experimentally observed for crystals with a disordered domain structure [1].

In closing, we note that the main result of this study is the development of the stochastic theory of the SHG in statistically inhomogeneous media. In particular, this study reveals that the dependence of the second harmonic intensity on the intensity of the laser radiation differs considerably from a square law in the case of a low efficiency of the fundamental-to-second harmonic conversion, which agrees well with the experimental results [1–3]. The approach described above can be generalized to the case of a quasiphase-mismatch conversion of optical frequencies in an active nonlinear medium or in a superlattice.

The results and the approach described in this paper can also be of interest in application to such objects as an inhomogeneous acoustically nonlinear medium or an inhomogeneous plasma.

We are grateful to G.D. Laptev and A.A. Novikov for useful discussions. This work was supported in part by the Russian Foundation for Basic Research and the programs “Basic Metrology” and “Basic Spectroscopy” of the Ministry of Industry and Science of the Russian Federation.

REFERENCES

1. S. Kawai, T. Ogawa, H. S. Lee, *et al.*, Appl. Phys. Lett. **73**, 768 (1998).
2. A. A. Kaminskiĭ, J. Garcia-Sole, S. N. Bagaev, *et al.*, Kvantovaya Élektron. (Moscow) **25**, 1059 (1998) [Quantum Electron. **28**, 1031 (1998)].
3. A. A. Kaminskiĭ, D. Jaque, G. N. Bagaev, *et al.*, Kvantovaya Élektron. (Moscow) **26**, 95 (1999) [Quantum Electron. **29**, 95 (1999)].
4. S. A. Akhmanov and A. S. Chirkin, *Statistical Phenomena in Nonlinear Optics* (Mosk. Gos. Univ., Moscow, 1971).
5. A. S. Chirkin and D. B. Yusupov, Radiotekh. Élektron. (Moscow) **23**, 636 (1978).
6. V. I. Klyatskin, *Stochastic Equations and Waves in Randomly-Inhomogeneous Media* (Nauka, Moscow, 1980).

Translated by E. Golyamina

Generation of High-Energy Secondary Pulsed Molecular Beams

V. M. Apatin, G. N. Makarov*, and V. V. Nesterov

Institute of Spectroscopy, Russian Academy of Sciences, Troitsk, Moscow region, 142190 Russia

* e-mail: g.makarov@isan.troitsk.ru

Received May 3, 2001

A method is suggested for generating high-intensity secondary pulsed molecular beams in which the kinetic energy of molecules can be controlled by an intense laser IR radiation through the vibrational excitation of molecules in the source. High-intensity [$\geq 10^{20}$ molecule/(sr s)] SF₆ molecular beams with a kinetic energy of $\cong 1.0$ eV without carrier gas and of $\cong 1.9$ and $\cong 2.4$ eV with carrier He (SF₆/He = 1/10) and H₂ (SF₆/H₂ = 1/10) gases, respectively, were obtained. © 2001 MAIK "Nauka/Interperiodica".

PACS numbers: 07.77.Gx; 33.80.-b

1. High-intensity beams [$\geq 10^{20}$ molecule/(sr s)] of accelerated molecules with kinetic energy ranging approximately from one to several electron-volts are used in various fields of fundamental and applied studies (investigation of chemical reactions with energy barriers, elastic and inelastic collisions, interaction of molecules with surface, etc.) [1].

The use of skimmers for separating molecules from gas-dynamically cooled jets formed by pulsed nozzles is the most widespread method of obtaining high-intensity molecular beams [2]. The main characteristics of pulsed beams are as follows: intensity, duration, velocity, and the scatter of molecular velocities in the beam (degree of gas cooling).

The above-mentioned energy range is difficult to utilize. No universal methods of generating molecular (atomic) beams in this range exist at present. Such beams are formed by different methods [1] (see also [3, 4] and references cited therein). However, the majority of these methods are rather complicated (e.g., initiation of optical breakdown or rf or arc discharge inside a nozzle) and apply only to atoms and not to molecules.

The aerodynamic acceleration [5, 6] of molecules diluted in a lighter carrier gas (e.g., He or H₂) and the heating of a gas in a nozzle to high temperatures ($T_0 \cong 3000$ K), as well as the combination of these two methods, are the methods that are most frequently used for this purpose. The aerodynamic acceleration is not too efficient if the gas-to-carrier mass ratio is small.

In the case of nozzle heating, the energy of beam molecules is determined by the gas temperature before the expansion through the nozzle:

$$\frac{1}{2} m v^2 = \frac{\gamma}{\gamma - 1} k(T_0 - T), \quad (1)$$

where v is the steady-state flow velocity, m is the molecular mass, $\gamma = c_p/c_v$ is the ratio of specific heats, k is the Boltzmann constant, and T is the steady-state temperature.

In [7], argon atoms diluted in helium were accelerated to several electron-volts in a continuous beam by the combination of nozzle heating and aerodynamic acceleration, and in [8] this method was used to obtain continuous beams of xenon atoms diluted in hydrogen (Xe/H₂ = 0.23/100) and having kinetic energy as high as $\cong 30$ eV.

Heating of pulsed nozzles to high temperatures is a challenge because they are made from materials (in particular, elastomers and plastics) that fail at $T \geq 200^\circ\text{C}$ [2]. In [9, 10], to obtain accelerated molecular beams, it was suggested to use the vibrational excitation of molecules by an IR laser pulse in the gas-dynamic expansion zone at the nozzle outlet. The accelerated beams of the SF₆ and CF₃I molecules with kinetic energies of $\cong 0.5$ and $\cong 0.74$ eV, respectively, were obtained in [10, 11]. However, this approach cannot provide high excitation energy densities because of the optical breakdown at the nozzle outlet. Moreover, only a small fraction of molecules can be efficiently accelerated in a beam by this method.

It follows from the aforesaid that the excitation of molecules by an intense IR laser radiation inside the source of a pulsed beam, i.e., before the gas outflow from the nozzle, would be promising for obtaining high-energy molecular beams. This program was implemented in this work by the method of formation of a secondary pulsed molecular beam, as a result of which the intense beams of accelerated molecules were obtained.

2. In the method proposed, the pressure shock (shock wave) was used as a source of the secondary

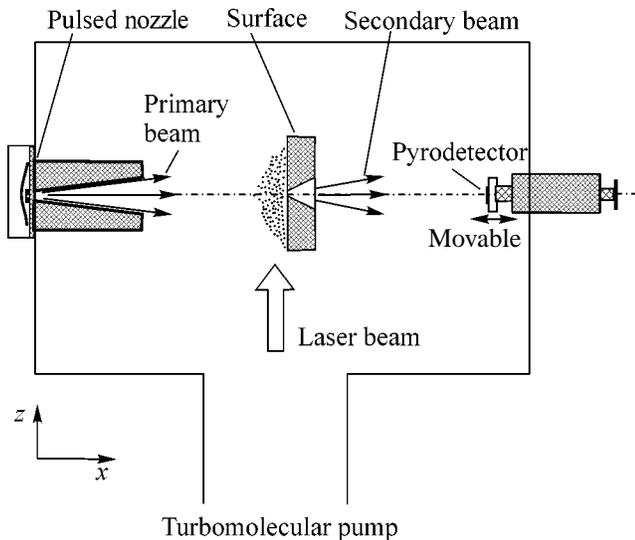


Fig. 1. Scheme of the experiment. Formation of the secondary molecular beam using a plate with a hole shaped like a divergent cone.

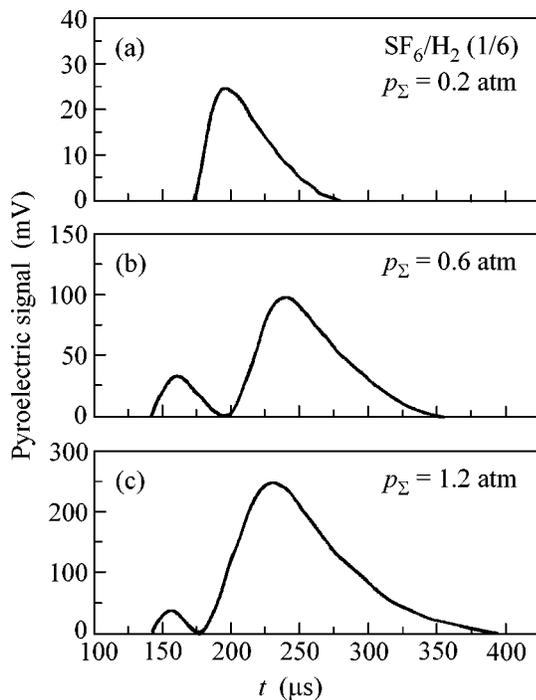


Fig. 2. Time evolution of a pulse of molecular beam ($\text{SF}_6/\text{H}_2 = 1/6$) passing through a hollow convergent truncated cone (inlet diameter 11 mm, outlet diameter 2.8 mm, and length 32 mm). The distance from the nozzle to the cone waist is 83 mm and from the nozzle to the detector is 143 mm.

molecular beam [12–14]. The shock was formed due to the interaction of a gas-dynamically cooled pulsed molecular flow with a solid surface [15, 16]. The essence of the experiment is illustrated in Fig. 1. An

intense [$\geq 10^{20}$ molecule/(sr s)] wide-aperture (with divergence $\omega \cong 0.05$ sr) molecular beam (or flow) was incident on the surface of a solid plate (a polished duralumin 7.5-mm-thick plate was placed at a distance $x \cong 60$ mm from the nozzle) having a hole in its center. The hole represented a divergent cone with the inlet diameter $d_{\text{in}} \cong 2$ mm and the outlet diameter $d_{\text{out}} \cong 5$ mm. The hole walls were polished.

As the primary beam was incident on the surface, a pressure shock was formed ahead of it, where the gas density, pressure, and temperature were appreciably higher than in the incident beam [17, 18]. According to the estimates made in [17], the concentration of SF_6 molecules in the shock varied from $\cong 10^{16}$ to $\cong 5 \times 10^{17}$ cm^{-3} , depending on the intensity of the primary beam. While the pressure shock persisted ahead of the surface, the gas flew out from the shock through the hole in the plate into a high-vacuum section of the chamber to form a new (secondary) pulsed molecular beam, whose characteristics were, generally, different from those of the primary beam.

The primary beam was formed by a pulsed nozzle of the current loop type [19]. The hole diameter was 0.75 mm. The time of opening the nozzle was equal to $\cong 40$ μs (FWHM). Gas pressure in the nozzle was varied from $\cong 0.1$ to 7 atm. The nozzle cross section was shaped like a cone with an angle of 15° . The cone length was 35 mm. A vacuum chamber where the molecular beam was formed was evacuated to $\cong 1 \times 10^{-6}$ torr by a turbomolecular pump. The number of molecules flowing out of the nozzle during a pulse depended on the pressure in the nozzle and was varied from $\cong 3 \times 10^{15}$ to $\cong 1.1 \times 10^{17}$ molecule/pulse.

The intensity of the secondary beam depended on the intensity of the primary beam and on the hole diameter and cone angle in the plate. It increased appreciably (by 5–6 times) when, instead of the flat surface, a hollow truncated cone was used to form the shock (or when the secondary beam was formed by a convergent–divergent cone of the Laval nozzle type).

One can see from Fig. 1 that the excitation of molecules in the shock wave presents no special problems. When the high-energy beams were obtained in the scheme with a cone, its convergent part was replaced by a hollow tetrahedral truncated pyramid made from thin NaCl plates, which are transparent to the radiation of a CO_2 laser. This allowed the molecules to be excited inside the pyramid immediately before the exit from the secondary nozzle. The internal (mainly vibrational) energy of molecules increased due to the multiphoton absorption in a strong IR field [20]. The energy transfer from the vibrational to the translational degrees of freedom occurred due to the subsequent vibrational–translational V – T relaxation in the vacuum-expanded gas, resulting in the acceleration of molecules.

In our experiments, the characteristics of the secondary molecular beam and the possibility of obtaining

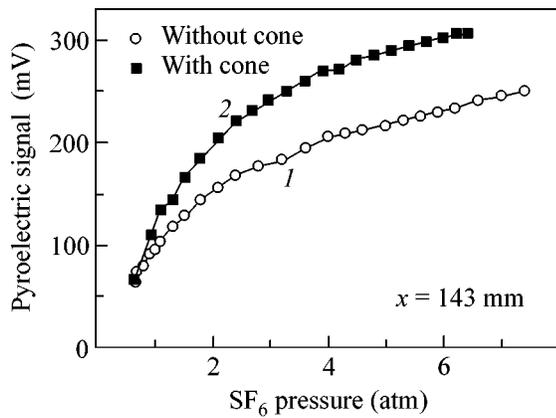


Fig. 3. Pyroelectric signals induced at the detector by the unperturbed primary molecular beam (curve 1) and secondary molecular beam (curve 2) as functions of the SF₆ pressure in the nozzle. The secondary beam was formed using a convergent-divergent cone of the Laval nozzle type. The cone parameters and the conditions of the experiment are given in the text.

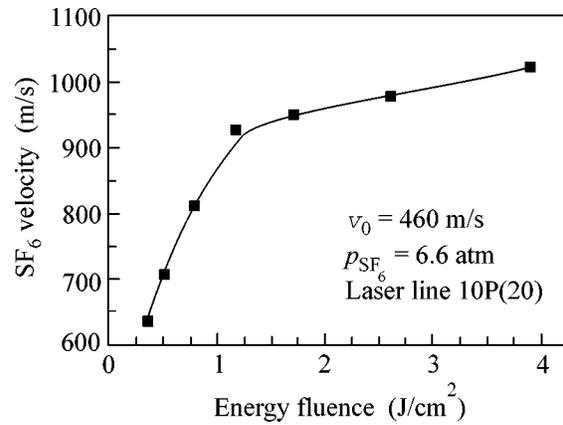


Fig. 4. The velocity of the secondary SF₆ beam as a function of laser fluence. The molecules were excited at a frequency of 944.2 cm⁻¹ [laser 10P(20) line]. The SF₆ pressure in the nozzle is equal to 6.6 atm.

high-energy molecules were examined. Measurements were carried out by the time-of-flight method using a noncooled pyrodetector with time resolution $\approx 3\text{--}5\ \mu\text{s}$ as a molecular beam detector [21, 22]. The time-of-flight spectra of molecules were measured at different distances from the sources of the primary and secondary beams and used to determine the beam velocities and the scatter of molecular velocities in the beams.

3. The time evolution of a pulsed molecular beam (time-of-flight spectrum of molecules) passing through the hollow cone is shown in Fig. 2 for different gas pressures (SF₆/H₂ = 1/6) in the nozzle. The distance from the nozzle to the cone waist was 83 mm and to the detector was 143 mm. At a low gas pressure in the nozzle ($p_{\Sigma} \leq 0.2$ atm), the intensity of the primary beam is low and the shock does not form ahead of the surface [17, 18], so that only the primary beam passes through the cone (Fig. 2a). As the pressure in the nozzle increases, the pressure shock is formed in the cone (in Fig. 1 ahead of the surface) and the primary beam pulse shortens. Simultaneously, a pulse of the secondary molecular beam appears and starts to rapidly increase in amplitude (Figs. 2b, 2c). At a relatively high intensity of the primary molecular beam [$\geq 10^{20}$ molecule/(sr s)],

the intensity of the secondary beam was comparable to the intensity of the unperturbed primary beam (in the absence of the plate).

Figure 3 shows the intensities of the primary (curve 1) and secondary (curve 2) molecular beams as functions of the SF₆ pressure in the nozzle. The secondary beam was formed using a hollow convergent-divergent cone of the Laval nozzle type ($d_{\text{in}} = 14$ mm, $d_0 = 2$ mm, $d_{\text{out}} = 7$ mm, overall length 40 mm, and the length of the convergent part 30 mm). The distance from the nozzle to the detector was 143 mm and from the cone waist to the detector was 79 mm. One can see that, at a SF₆ pressure $p \geq 1.0$ atm in the nozzle, the pyroelectric signal induced in the detector by the secondary molecular beam is stronger than the signal from the primary beam. For this reason, the intensity of the secondary beam in Fig. 3 is comparable with the intensity of the primary beam even if one takes into account an approximately quadratic dependence of the pyroelectric signal on the distance between the beam source and the detector.

It was established experimentally that the duration of the secondary molecular beam, its velocity, and the scatter of molecular velocities do not differ appreciably

Acceleration of SF₆ in the secondary molecular beam

Gas composition	Pressure in the nozzle, atm	CO ₂ laser line	Energy fluence, J/cm ²	Mean molecular velocity in the beam, m/s		Kinetic energy of molecules, eV	
				v_0	V_L	E_{kin}^0	E_{kin}^L
SF ₆	6.6	10P(26)	3.7	460	1150	0.163	1.0
SF ₆ + H ₂ (1 : 10)	3.2	10P(16)	3.5	1020	1765	0.8	2.40
SF ₆ + He (1 : 10)	4.0	10P(20)	3.7	1050	1580	0.85	1.92

from those in the primary beam. As expected, the difference in the scatter of velocities is the greatest. It was 20–30% higher in the secondary beam than in the primary one. For instance, when the secondary beam was formed using the cone whose parameters are given above and the SF₆ pressure in the nozzle was equal to 6.6 atm, the velocities of molecular beams and the scatter of velocities in the primary and secondary molecular beams were, respectively, $v_1 \cong 480$ m/s, $\Delta v_1 \cong 77$ m/s and $v_2 \cong 447$ m/s, $\Delta v_2 \cong 87$ m/s. It follows from this example that the gas in the secondary beam is cooled rather strongly (the Mach number was $M_2 \cong v_2/\Delta v_2 \cong 5$).

The velocity of the secondary SF₆ beam as a function of the CO₂ laser energy fluence is shown in Fig. 4. The molecules were excited in the source of a secondary beam immediately before the exit from the nozzle [inside a hollow truncated pyramid made from the NaCl plates and attached to the front of the plate with a conic hole (Fig. 1)]. The laser was tuned to a frequency of 944.2 cm^{-1} [10P(20) line], which is close to the frequency of the SF₆ v_3 mode ($\cong 948 \text{ cm}^{-1}$ [23]). The mean molecular velocity in the absence of laser excitation was $v_0 = 460$ m/s and achieved $v_L > 1000$ m/s in the presence of laser excitation.

The results on the acceleration of SF₆ in the secondary pulsed beam are presented in the table. We obtained molecular beams of SF₆ with kinetic energies $E_{\text{kin}}^L \cong 1$ eV ($v_L \cong 1150$ m/s) in the absence of a carrier gas and $E_{\text{kin}}^L \cong 1.9$ and 2.4 eV in the presence of, respectively, He and H₂ as carrier gases. These values are appreciably higher than those obtained in [4, 10].

Thus, the method suggested in this work allows one to obtain intense molecular beams in which the kinetic energy of molecules can be controlled through the vibrational excitation by a strong IR laser pulse in the source. Note in conclusion that the dissociation of molecules by the IR or UV laser radiation in the secondary source or in the beam can also be used to obtain accelerated radicals.

We are grateful to V.N. Lokhman and A.N. Petin for assistance. This work was supported in part by the Russian Foundation for Basic Research (project no. 00-03-33003a) and the American Foundation for Civil Research and Development of the Independent States of the Former Soviet Union (grant no. RC1-2206).

REFERENCES

1. H. Pauly, in *Atomic and Molecular Beam Methods*, Ed. by G. Scoles (Oxford Univ. Press, New York, 1988).
2. W. R. Gentry, in *Atomic and Molecular Beam Methods*, Ed. by G. Scoles (Oxford Univ. Press, New York, 1988).
3. J. M. Girard, A. Lebehot, and R. Compargue, *J. Phys. D* **26**, 1382 (1993).
4. G. N. Makarov, *Zh. Éksp. Teor. Fiz.* **108**, 404 (1995) [*JETP* **81**, 218 (1995)].
5. E. W. Becker, K. Bier, and H. Burghoff, *Z. Naturforsch. A* **10**, 7 (1955).
6. E. Kolodney and A. Amirav, *Chem. Phys.* **82**, 269 (1983).
7. R. Comparque, A. Lebehot, J. C. Lemonnier, and D. Murette, in *Rarefied Gas Dynamics: Technical Papers Selected from the 12th International Symposium on Rarefied Gas Dynamics, 1980*, Ed. by S. S. Fisher (American Inst. of Aeronautics and Astronautics, New York, 1980), p. 823.
8. U. Buck, E. Lessner, and D. Pust, *J. Phys. B* **13**, L125 (1980).
9. V. N. Lokhman and G. N. Makarov, *Pis'ma Zh. Éksp. Teor. Fiz.* **61**, 172 (1995) [*JETP Lett.* **61**, 175 (1995)].
10. G. N. Makarov, *Chem. Phys. Lett.* **237**, 361 (1995).
11. V. N. Lokhman, G. N. Makarov, D. D. Ogurok, *et al.*, *Khim. Fiz.* **17**, 35 (1998).
12. Ya. B. Zel'dovich and Yu. P. Raizer, *Physics of Shock Waves and High-Temperature Hydrodynamic Phenomena* (Nauka, Moscow, 1966; Academic, New York, 1966).
13. L. D. Landau and E. M. Lifshitz, *Course of Theoretical Physics, Vol. 6: Fluid Mechanics* (Nauka, Moscow, 1986; Pergamon, New York, 1987).
14. G. N. Abramovich, in *Applied Gas Dynamics* (Nauka, Moscow, 1991), Part 1.
15. G. N. Makarov and A. N. Petin, *Pis'ma Zh. Éksp. Teor. Fiz.* **71**, 583 (2000) [*JETP Lett.* **71**, 399 (2000)].
16. G. N. Makarov and A. N. Petin, *Chem. Phys. Lett.* **323**, 345 (2000).
17. G. N. Makarov and A. N. Petin, *Zh. Éksp. Teor. Fiz.* **119**, 5 (2001) [*JETP* **92**, 1 (2001)].
18. G. N. Makarov and A. N. Petin, *Chem. Phys.* **266**, 125 (2001).
19. W. R. Gentry and C. F. Giese, *Rev. Sci. Instrum.* **49**, 595 (1978).
20. V. N. Bagratashvili, V. S. Letokhov, A. A. Makarov, and E. A. Ryabov, *Multiple Photon Infrared Laser Photo-physics and Photochemistry* (Harwood Academic, New York, 1985).
21. V. M. Apatin, L. M. Dorozhkin, G. N. Makarov, and L. M. Pleshkov, *Appl. Phys. B: Photophys. Laser Chem.* **B29**, 273 (1982).
22. V. M. Apatin and G. N. Makarov, *Zh. Éksp. Teor. Fiz.* **84**, 15 (1983) [*Sov. Phys. JETP* **57**, 8 (1983)].
23. R. S. McDowell, J. B. Krohn, H. Flicker, and M. C. Vasquez, *Spectrochim. Acta A* **42**, 351 (1986).

Translated by V. Sakun

Influence of Angles of Incidence of Laser Radiation on the Generation of Fast Ions

S. Yu. Gus'kov^{1,*}, N. N. Demchenko¹, K. N. Makarov², S. G. Nishchuk²,
T. A. Pikuz³, Yu. A. Satov², I. Yu. Skobelev³, Yu. B. Smakovskii², A. Ya. Faenov³,
S. V. Khomenko², and B. Yu. Sharkov⁴

¹ Lebedev Physical Institute, Russian Academy of Sciences, Leninskii pr. 53, Moscow, 117924 Russia

*e-mail: guskov@sci.lebedev.ru

² Troitsk Institute for Innovation and Thermonuclear Research, Troitsk, Moscow region, 142190 Russia

³ Center of Spectral Data for Multiply Charged Atoms,
State Scientific Center All-Russia Research Institute of Physicotechnical and Radio Engineering Measurements,
Mendeleevo, Moscow region, 141570 Russia

⁴ Institute of Theoretical and Experimental Physics, Bol'shaya Cheremushkinskaya ul. 25, Moscow, 117218 Russia

Received May 16, 2001

It was established experimentally that the number and energy of fast ions in laser plasma increased with increasing angle of focusing laser radiation onto a flat target. Numerical calculations showed that the increase in angle of focusing brought the mean angle of incidence of laser radiation closer to the optimal angle corresponding to the maximal efficiency of the resonance absorption mechanism and, as a result, increased the fraction of absorbed laser energy in the energy of fast electrons and increased the number of fast electrons. In turn, the increase in the energy and number of fast electrons resulted in an increase in the number of fast electrons involved in the formation of a self-consistent electric field at the target edge and led to the growth of the field strength, which, eventually, was the reason for the increase in the number and energy of fast ions. © 2001 MAIK "Nauka/Interperiodica".

PACS numbers: 52.38.Kd; 52.50.Jm

1. The studies of fast particle formation in laser plasma is a topical problem not only in the light of the development of our knowledge about the fundamental physical processes occurring in high-temperature plasma but also for solving a number of important applied problems such as design of injectors of multiply charged ions for heavy-ion accelerators or laser fusion, where the fast particles may play both a negative role (premature heating of a spherical target and deterioration of its compression) and a positive role (fast-ion fusion reaction). In the last 30 years, experimental and theoretical investigations have shown that the product $q_{\text{las}} \lambda_{\text{las}}^2$, where q_{las} and λ_{las} are the laser radiant flux and laser wavelength, respectively, is the key parameter for the physics of fast-particle generation in laser plasma (see, e.g., review [1]). Clearly, this parameter alone cannot describe a great diversity of physical processes resulting in the electron and ion acceleration. Under different experimental conditions, the efficiency of acceleration processes is different even at a fixed value of parameter $q_{\text{las}} \lambda_{\text{las}}^2$. For instance, it is evident that the efficiency of acceleration processes can depend on the duration of laser pulse (especially for ultrashort pulses) or on the aggregate state of the target substance (solid state, gas, or clusters). The experimen-

tal studies carried out in this work showed that there is one more and not quite apparent parameter that determines the number of generated fast particles. This parameter proved to be the aperture of a lens used for focusing the laser pulse, or, more precisely, the ratio of the beam diameter D_{las} to the lens focal length F . Our experiments showed that, all things being the same, the number of fast ions increased with increasing parameter D_{las}/F . This effect was explained by the theoretical calculations, and it was shown that it is caused by the resonance absorption of that part of laser radiation which is incident on the target not normally to the surface.

2. Experiments were performed at the Troitsk Institute for Innovation and Thermonuclear Research on a TIR setup [2]. This setup consisted of a CO₂ laser with an output of up to 100 J and allowed the generation of pulses with a smooth temporal shape and a duration of 15–80 ns. The pulse duration was varied by changing the level of pumping active mixture and the operating mode of a master oscillator and by varying the composition of cells containing a saturated absorber (SF₆ in a mixture with air). The experiments described in this work were carried out with pulse durations of 20 and 14 ns at half-widths. The radiation was focused onto a target using lenses with focal lengths $F = 60$ and

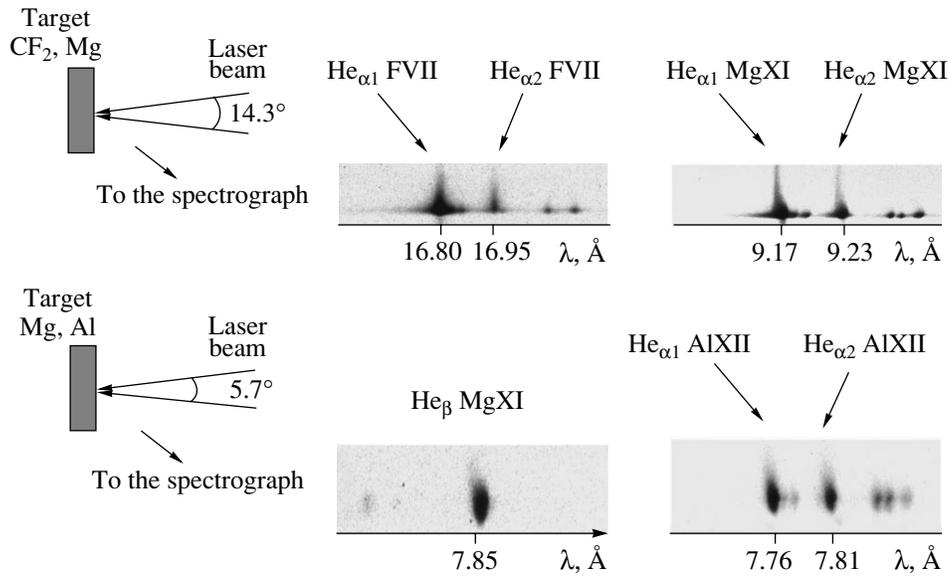


Fig. 1. Scheme of the experiment and the spectrograms of the He-like FVII, MgXI, and AIXII ions, as observed for laser radiation focused by lenses with $D_{\text{las}}/F =$ (up) 0.25 and (down) 0.1.

150 cm, and the beam diameter D_{las} was 15 cm. The laser power density on the target was $\sim(3-5) \times 10^{12}$ W/cm².

Soft X-rays from plasma were recorded by spectrographs with spherically curved mica crystals. The mica crystals, the target, and a photographic film were arranged following the FSSR-1D scheme [3–5]. The spectra were recorded with high spectral ($\lambda/\Delta\lambda \approx 2000-10000$, depending on the spectral range) and spatial ($\Delta x \approx 20$ μm) resolutions. Flat massive plates made from Teflon, magnesium, and aluminum were used as targets. The observation angle was $\approx 55^\circ$ to the target surface (Fig. 1). With such an arrangement of the spectrograph, the ions flying from the target had a velocity component along the observation direction and, hence, the emitted photons were short-wavelength-shifted because of the Doppler effect. The appearance of the fast ions could be judged from the presence of broad short-wavelength wings in the observed spectral lines, with the wing shape approximately reproducing the ion velocity distribution.

The experiments were performed with two apertures of focusing lenses: (1) $D_{\text{las}}/F = 0.25$ and (2) $D_{\text{las}}/F = 0.1$. In the first case, broad short-wavelength wings were clearly observed for all recorded spectral lines, regardless of the target material; in the second case, the wing intensities were no higher than the noise level (Fig. 1). It is worth noting that the laser power density in the second case was even slightly higher than in the first one (in the first case, a pulse with $\tau_{\text{las}} = 20$ ns and $E_{\text{las}} = 60$ J was focused to a spot of diameter 300 μm , while, in the second case, a higher power pulse with $\tau_{\text{las}} = 14$ ns and $E_{\text{las}} = 100$ J was focused to a smaller spot with a diameter of 180 μm).

Figure 2 shows the ion velocity distribution reproduced from the shape of the short-wavelength wing of the MgXI He $_{\alpha}$ line for both focusing situations. One can see that, in the first case (high-aperture lens), there were many plasma ions with a velocity of up to $(1.5-2) \times 10^8$ cm/s, whereas in the second case (lower aperture lens) only ions with appreciably lower velocities of $(5-6) \times 10^7$ cm/s were recorded with certainty (Fig. 2b).

3. The ion velocity as a function of the angle of focusing laser radiation suggests that the accelerating field is created by fast electrons produced upon resonance absorption at the critical surface. Indeed, some photons of the laser radiation focused onto the target are incident on the surface not along its normal, and the number of such photons increases with increasing lens aperture. A half of radiation flow is *p*-polarized and can excite resonance at the critical surface. The efficiency of resonance absorption depends on the parameter $\tau = (k_0 L)^{1/3} \sin \theta_0$ [6–8], where $k_0 = \omega/c$ is the wave number, ω is the laser frequency, L is the plasma characteristic inhomogeneity scale at the critical surface, and θ_0 is the angle of incidence on the target. The maximal efficiency of resonance absorption corresponds to $\tau_0 = 0.7$. The absorption efficiency, the fraction of resonance absorption, and the plasma hydrodynamic parameters were calculated using the physical model developed in [7, 8] (the RAPID program). This model is based on the equations of two-temperature plasma hydrodynamics and on a combination of the Maxwell equations with the geometrical optics approximation in describing the propagation of laser radiation in plasma. Strictly speaking, if a flat target is irradiated by a single beam, the calculation of plasma dynamics in the plasma torch geom-

etry necessitates a two-dimensional description of plasma expansion. However, it was pointed out in [9, 10] that the expansion in a transverse direction may qualitatively be taken into account using spherical geometry, provided that the initial radius of the target is equal to the focal spot diameter. (In this case, the laser power density should be fixed). Note also that, if the focal point in the case of a spherical target is shifted from the target center along the beam direction at a distance equal to the radius of critical surface, then the angular distribution of incident radiation will exactly coincide with the angular distribution of the radiation focused onto the flat target. This model was used to carry out calculations for aluminum targets with initial radii 300 and 500 μm (typical size of a focal point).

In compliance with the experimental conditions, two variants of laser pulse and lens aperture were considered. In both cases, the time dependence for the flow was taken in the form of an isosceles triangle and the lens focal point was positioned behind the target center in such a way that the distance from it to the center was equal to the radius of critical surface at the moment of pulse maximum. The table reports the computational results for some of the parameters characterizing the light absorption process (δ_{abs} is the total fraction of absorption due to inverse bremsstrahlung and resonance mechanisms, and $\delta_{\text{abs}}^{\text{res}}$ is the fraction of resonance absorption) and for the following main plasma hydrodynamic parameters at the moment of maximal pulse power: T_e^{max} is the maximal electron temperature in corona, $L = [1/\rho(\partial\rho/\partial r)_c]^{-1}$ is the characteristic scale of plasma inhomogeneity at the critical point, r_c is the radius of the critical surface, $u(\rho_c)$ is the plasma velocity at the point with critical density ρ_c , and $u(\rho_c/10)$ is the plasma velocity at the point with density $\rho_c/10$. The calculated density ρ , velocity u , and electron-temperature T_e and ion-temperature T_i profiles are shown in Fig. 3 for the first variant with high-aperture lens.

It is seen from the table and Fig. 3 that the hydrodynamic velocities are not high enough to account for the observed Doppler shift at relatively small (~ 400 – $800 \mu\text{m}$) distances from the target. Moreover, because of a higher flux density, the velocity in the variant with $F = 150 \text{ cm}$ is slightly higher than for $F = 60 \text{ cm}$. According to the table, the hydrodynamic parameters of plasma depend weakly on the lens aperture. However, the fractions of resonance absorption differ appreciably. In the case of a target with $R_0 = 300 \mu\text{m}$, it is 3.8 times higher for $F = 60 \text{ cm}$ than for $F = 150 \text{ cm}$. Therefore, one should take into account a mechanism of additional acceleration of plasma ions by the electric field created by fast electrons.

4. Let us consider a simple model allowing the estimation of the characteristic energy and density of fast electrons produced through the resonance absorption of laser radiation. At the ray turning point, plasma can be

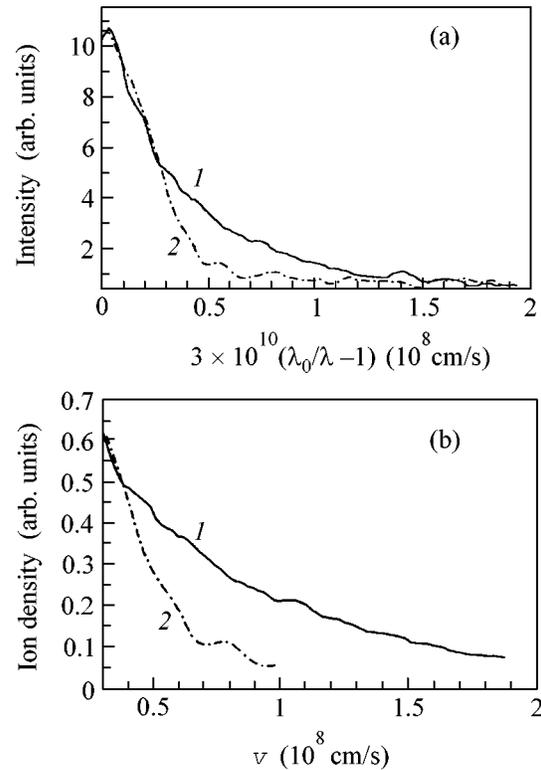


Fig. 2. (a) The short-wavelength wing of the MgXI He $_{\alpha}$ line and (b) the ion velocity distribution, as observed for two variants of focusing laser radiation: $F = (1)$ 60 and (2) 150 cm.

considered as a flat layer with density changing along its gradient [7, 8] (the x axis is chosen in this direction). It follows from the Maxwell equations that the longitudinal field at the critical surface is limited by the dissipative processes [6]

$$|E_{xc}| = \alpha_0 |H_c| \omega / \nu, \quad (1)$$

where ν is the effective frequency that is related to the high-frequency conductivity σ by formula $\sigma = \nu/4\pi$, H_c is the magnetic field at the critical point, and $\alpha_0 = \sin\theta_0$ (θ_0 is the angle of incidence). In the spherical case, the angle of incidence is determined from the relation $\sin^2\theta_0 = \epsilon_1(r_t)$, where $\epsilon_1(r_t)$ is the real part of the dielectric constant at the ray turning point [7, 8].

In weak fields, where the electron oscillation energy is lower than its thermal energy, the field dissipation is due to the electron–ion collisions, so that $\nu = \nu_{ei}$. In strong fields, where the electron oscillation amplitude becomes comparable with the characteristic width of plasma resonance, the longitudinal field is mainly limited by the generation of superthermal electrons, so that $\nu = \nu_h$ [in the general case, $\nu = \nu_{ei} + \nu_h$ in Eq. (1)]. During the oscillation period, the field can expend work on an electron flying through the resonance region. The resulting electron characteristic energy is equal to its oscillation energy in the resonant field. The superther-

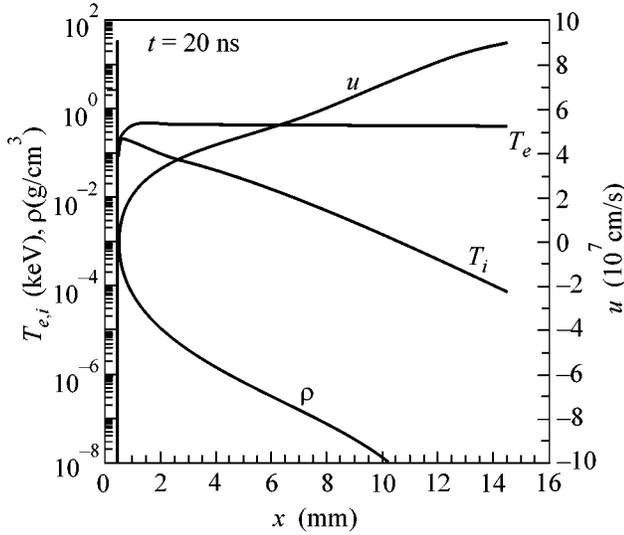


Fig. 3. Density ρ , velocity u , and electron-temperature T_e and ion-temperature T_i profiles at the moment $t = 20$ ns of maximal pulse power for the variant with $F = 60$ cm and $R_0 = 500$ μm .

mal electrons carry away energy from the resonance region:

$$n_h(m_e v_h^3/2) = (v_h/v)q_{\text{abs}}^{\text{res}}, \quad (2)$$

where n_h and v_h are, respectively, the density and the velocity of superthermal electrons and $q_{\text{abs}}^{\text{res}}$ is the energy flux density absorbed in the resonance region. One has for the velocity v_h

$$v_h = e|E_{xc}|/m_e\omega. \quad (3)$$

The current of fast electrons is fully dissipative,

$$en_h v_h = \sigma_h|E_{xc}|, \quad (4)$$

where $\sigma_h = v_h/4\pi$. Replacing the density profile in the vicinity of the critical point by a linear curve with characteristic inhomogeneity scale L , one can obtain [6]

$$q_{\text{abs}}^{\text{res}} = \int_{\Delta x_{\text{res}}} \sigma|E_x|^2 dx = \frac{1}{8}\alpha_0^2|H_{yc}|^2\omega L. \quad (5)$$

From Eqs. (2)–(5) one gets

$$\frac{v_h}{\omega} + \frac{v_{ei}}{\omega} = \left[\frac{e\alpha_0|H_{yc}|}{\pi m_e \omega^2 L} \right]^{1/2}. \quad (6)$$

The relativistic case can be considered in a similar manner (the result is presented in [11]). Interestingly, Eq. (6) for the effective frequency can be derived by another method through setting the electron oscillation amplitude in resonant field $a_{os} = e|E_{xc}|/m_e\omega^2$ equal to the resonance characteristic width $\Delta x = 2Lv_h/\omega$. To estimate the energy of fast electrons, the relationship between the magnetic field $|H_{yc}|$ at the critical point and its vacuum value $|H_0|$ (in the incident wave) should be known. For the plasma with a linear density profile, this relationship is written as [6]

$$\alpha_0|H_{yc}| = \frac{\Phi(\tau)|H_0|}{\sqrt{2\pi k_0 L}}, \quad (7)$$

$$\Phi(\tau) = 4\tau V(\tau^2) \sqrt{\frac{V(\tau^2)}{-V'(\tau^2)}},$$

where V and V' are, respectively, the Airy function and its derivative. The function $\Phi(\tau)$ has a maximum at $\tau = 0.7$: $\Phi(0.7) = 1.2$; at $\tau \rightarrow 0$, $\Phi(\tau) \approx 2\tau$. Estimates show that the electron oscillation energy in the resonance is far above the thermal energy, so that the Coulomb frequency v_{ei} is determined not by the thermal but by the oscillation velocity, allowing v_{ei} to be neglected compared to v_h . One finally obtains for the longitudinal field in the resonance

$$|E_{xc}| = |H_0|^{1/2}[\Phi(\tau)]^{1/2}(\pi k_0 L/2)^{1/4}(m_e c\omega/e)^{1/2}.$$

The energy of fast electrons is $\epsilon_h = e^2|E_{xc}|^2/m_e\omega^2$. From Eqs. (3) and (4), one has for the density of fast electrons $n_h = n_c(v_h/\omega)$, where n_c is the critical density (for a CO₂ laser, $n_c = 10^{19}$ cm⁻³).

Let us estimate ϵ_h and n_h using the plasma parameters obtained from the hydrodynamic calculations. For $F = 60$ cm and $t = 20$ ns, one has $L = 4 \times 10^{-2}$ cm, $k_0 L = 240$, $\alpha_0 = 0.124$ (for the rays with the maximal angle of incidence), $|H_0| = |E_0| = 1.58 \times 10^5$ CGS units for $q_{\text{las}} = 3 \times 10^{12}$ W/cm², and $\tau = 0.77$. With these values, one

Table

$R_0, \mu\text{m}$	δ_{abs}	$\delta_{\text{abs}}^{\text{res}}$	$T_e^{\text{max}}, \text{keV}$	$L, \mu\text{m}$	$r_c, \mu\text{m}$	$u(\rho_c), 10^7 \text{ cm/s}$	$u(\rho_c/10), 10^7 \text{ cm/s}$
$E_{\text{las}} = 60 \text{ J}, \tau_{\text{las}} = 20 \text{ ns}, F = 60 \text{ cm}$							
300	0.642	0.106	0.518	362	900	2.73	4.24
500	0.752	0.094	0.462	442	1200	2.57	3.87
$E_{\text{las}} = 100 \text{ J}, \tau_{\text{las}} = 14 \text{ ns}, F = 150 \text{ cm}$							
300	0.517	0.0279	0.654	369	1040	3.5	4.91
500	0.620	0.0297	0.593	441	1340	3.06	4.62

obtains $v_h/\omega = 7.8 \times 10^{-4}$, $|E_{xc}| = 6.15 \times 10^6$ CGS units, $\epsilon_h = 180$ keV, and $n_h = 7.8 \times 10^{15} \text{ cm}^{-3}$. Clearly, the Coulomb collision frequency can be ignored for the electron energy of 180 keV ($v_{ei}/\omega \sim 10^{-7}$). For $F = 150$ cm and $t = 14$ ns one has $\epsilon_h = 160$ keV and $n_h = 7.6 \times 10^{15} \text{ cm}^{-3}$. Consequently, the maximal energy and density of fast electrons change only slightly. A considerable difference is observed for the fraction of resonance absorption (by 3–4 times), because it is an integral over the angles of incidence and, hence, is sensitive to the distinctions in the angular distributions of incident radiation. The total number of fast electrons is determined by the fraction of resonance absorption and is equal to 2.21×10^{14} for $F = 60$ cm ($R_0 = 300 \mu\text{m}$) and to 1.09×10^{14} for $F = 150$ cm if the maximal energies of fast electrons are taken as mean values.

5. Thus, the use of a high-aperture focusing lens enhances the role of resonance absorption and, therefore, must increase the number of fast electrons generated in plasma, which, in turn, must increase the fraction of fast ions. Based on the above-mentioned results concerning the generation of fast electrons, one can obtain simple estimates for the energy and fraction of fast ions produced under the conditions of our experiments. As is well known, the ion acceleration inside a plasma torch is of little importance at laser power densities below $q_{\text{las}} = 10^{18} - 10^{19} \text{ W/cm}^2$ because the current of fast electrons is compensated for by the conduction current of plasma thermal electrons. In this case, fast ions can be accelerated efficiently only under the action of a self-consistent electric field arising at the plasma boundary with vacuum because of the charge separation due to the escape of fast electrons from the plasma at a distance equal to the Debye radius.

The self-consistent field at the plasma boundary with vacuum can be estimated as

$$E_S = 4\pi J_h e t_h, \quad (8)$$

where J_h is the flow of fast electrons at the plasma boundary and t_h is the residence time of a fast electron beyond the plasma (the inverse of the Debye frequency of a fast electron).

Assuming that the energy of fast electrons corresponds to the energy of laser radiation absorbed through the resonance mechanism, the flow of fast electrons at the plasma boundary, with allowance made for plasma expansion, is

$$J_h = \frac{\delta_{\text{abs}}^{\text{res}} \beta q_{\text{las}}}{\epsilon_h} \left(\frac{R_{\text{las}}}{R_{\text{las}} + u_b \tau_{\text{las}}} \right)^2, \quad (9)$$

where β is the fraction of fast electrons flying in the direction opposite to the beam and u_b is the characteristic velocity of the plasma boundary.

By solving the equation of motion for a fast electron in field (8), one gets for the time t_h

$$t_h = 2^{-1/4} \left(\frac{m_e^{1/2} \epsilon_h^{3/2}}{\pi \delta_{\text{abs}}^{\text{res}} \beta q_{\text{las}}} \right)^{1/2} \left(1 + \frac{u_b \tau_{\text{las}}}{R_{\text{las}}} \right). \quad (10)$$

Substituting Eqs. (9) and (10) into Eq. (8), one finally obtains for the self-consistent field

$$E_S = 2^{1/4} \left(\frac{\pi \delta_{\text{abs}}^{\text{res}} \beta q_{\text{las}} m_e^{1/2}}{\epsilon_h^{1/2}} \right) \left(\frac{R_{\text{las}}}{R_{\text{las}} + u_b \tau_{\text{las}}} \right). \quad (11)$$

The mean number of accelerated ions is equal to the mean number of fast electrons beyond the target. These create a self-consistent field

$$N_i \equiv \pi (R_{\text{las}} + u_b \tau_{\text{las}})^2 J_h t_h. \quad (12)$$

During the laser pulse, the ions are accelerated to the velocity

$$V_i \equiv \frac{ze}{m_i} E_S \tau_{\text{las}}. \quad (13)$$

Under the conditions of our problem, $u_b \tau_{\text{las}} \gg R_{\text{las}}$ and one has from Eqs. (12), (13), and (9)–(11) for the mean number of fast ions and their velocity

$$N_i \approx 2^{-3/4} \left(\frac{\delta_{\text{abs}}^{\text{res}} \beta q_{\text{las}} m_e^{1/2}}{\pi e^2 \epsilon_h^{1/2}} \right)^{1/2} R_{\text{las}} u_b \tau_{\text{las}}, \quad (14)$$

$$V_i \approx 2^{1/4} \frac{ze}{m_i} \left(\frac{\pi \delta_{\text{abs}}^{\text{res}} \beta q_{\text{las}} m_e^{1/2}}{\epsilon_h^{1/2}} \right)^{1/2} \frac{R_{\text{las}}}{u_b}. \quad (15)$$

It is noteworthy that the results obtained for the acceleration of fast ions in the self-consistent field of fast electrons predict an increase both in the ion energy and in the number of ions with increasing fraction of resonantly absorbed energy; i.e., $V_i \propto N_i \propto (\delta_{\text{abs}}^{\text{res}})^{1/2}$.

Substituting the experimental parameters and the computational results given in the table and assuming additionally that only half of the fast electrons fly toward the target boundary with vacuum, one can readily obtain for the mean energy and the number of fast aluminum ions that $V_i \approx 4 \times 10^8 \text{ cm/s}$ and $N_i \approx 2.5 \times 10^{11}$ in the experiment with a short-focus lens and $V_i \approx 1.5 \times 10^8 \text{ cm/s}$ and $N_i \approx 10^{11}$ in the experiment with a long-focus optics, in good agreement with the experimental data.

Thus, it has been observed experimentally in this work that the use of a high-aperture focusing lens increases the energy and the number of fast ions in laser plasma. This phenomenon is explained by the fact that the use of short-focus optics increases the fraction of resonantly absorbed laser energy and, hence, the number of fast electrons. In our opinion, the use of this

effect in the design of injectors of multiply charged ions for heavy-ion accelerators holds much promise.

This work was supported in part by the INTAS (grant no. 97-2090) and the MNTTs (project no. 856).

REFERENCES

1. S. J. Gitomer, R. D. Jones, F. Begay, *et al.*, Phys. Fluids **29**, 2679 (1986).
2. V. Yu. Baranov, K. N. Makarov, V. C. Roerich, *et al.*, Laser Part. Beams **14**, 347 (1996).
3. I. Yu. Skobelev, A. Ya. Faenov, B. A. Bryunetkin, *et al.*, Zh. Éksp. Teor. Fiz. **108**, 1263 (1995) [JETP **81**, 692 (1995)].
4. T. A. Pikuz, A. Ya. Faenov, S. A. Pikuz, *et al.*, J. X-Ray Sci. Technol. **5**, 323 (1995).
5. B. K. F. Young, A. L. Osterheld, D. F. Price, *et al.*, Rev. Sci. Instrum. **69**, 4049 (1998).
6. V. L. Ginzburg, *The Propagation of Electromagnetic Waves in Plasmas* (Nauka, Moscow, 1967; Pergamon, Oxford, 1970).
7. Yu. V. Afanas'ev, E. G. Gamaliĭ, N. N. Demchenko, *et al.*, Zh. Éksp. Teor. Fiz. **106**, 1553 (1980) [Sov. Phys. JETP **79**, 837 (1980)].
8. Yu. V. Afanas'ev, E. G. Gamaliĭ, N. N. Demchenko, *et al.*, Tr. Fiz. Inst. Akad. Nauk SSSR **134**, 3 (1980).
9. N. N. Demchenko and V. B. Rozanov, J. Sov. Laser Res. **12**, 250 (1991).
10. N. N. Demchenko and V. B. Rozanov, Preprint No. 209, FIAN (Lebedev Institute of Physics, USSR Academy of Sciences, Moscow, 1989).
11. M. P. Kalashnikov, P. Nickles, Th. Schlegel, *et al.*, Phys. Rev. Lett. **73**, 260 (1994).

Translated by V. Sakun

Remanent States of Small Ferromagnetic Cylinder¹

V. Ternovsky, B. Luk'yanchuk, and J. P. Wang

Data Storage Institute of Singapore, 117608 Singapore

e-mail: vladimir@dsi.nus.edu.sg

Received May 7, 2001; in final form, May 21, 2001

The internal magnetic structures of a cylinder are computed without prior assumption of axial symmetry. A three-dimensional magnetostatic solution is presented. A detailed distribution of the magnetization is obtained through numerical integration of the Landau–Lifshitz equation with the demagnetization field. Various entry conditions are used. Materials with small uniaxial anisotropy typically demonstrate vortex structures, while a higher anisotropy leads to the formation of concentric cylindrical magnetic bubbles inside the cylinder and a monodomain flower state. © 2001 MAIK “Nauka/Interperiodica”.

PACS numbers: 75.60.Lr; 75.60.Ej; 75.60.Ch

The problem of the shaped body magnetization under the action of some initial magnetic field refers to a well-known problem in magnetism [1, 2]; however, until now it has been insufficiently analyzed. Solution of this problem is important for the design of high-density magnetic recording media or magnetic random access memory. The modern line in the fabrication of these devices is the usage of patterned magnetic recording media with nanoscale sizes, where each element stores one bit of data (see, e.g., [3]). This patterning permits one to enlarge the signal-to-noise ratio. The elements in arrays are designed so that an element's magnetostatic interactions with its neighbors are small in comparison to its coercivity [4]. Such patterned media have been proposed for high-density data storage (above 500 Gbt/in²); see [4] and references therein.

Particles with dimensions of about 100 nm (and below) are usually too small to show a well-developed domain structure, but show a variety of nonuniform magnetization configurations. These remanent states depend on the properties of the magnetic body and its geometry, as well as the magnetization prehistory.

The first problem of this kind, which was explored in detail, refers to the magnetic states of small cubic particles [1, 2, 5, 6]. An analysis shows that magnetic states in a small ferromagnetic cube can be classified in terms of some elementary structures, namely, flower state (single domain), two domains, and vortex and twisted vortex states. The phase diagram of these states at zero magnetic fields was recently calculated in coordinates relative to exchange interaction and uniaxial anisotropy [6]. These remanent states were identified experimentally using magnetic force microscopy [7], Kerr microscopy [8], and other high-resolution mag-

netic imaging methods [9]. The reason for a variety of magnetization configurations in a cubic particle is related to its high symmetry.

Although a high-density (100 Gbt/in²) writing and reading on perpendicular recording media patterned by square islands was recently demonstrated [10], many researchers do not consider the cubic (or thin square) particles as optimal for magnetic recording. Thus, particles with lower symmetry are intensively discussed. Among the shapes experimentally analyzed are elliptical [11], conical [12], cylindrical [13], and spherical [14] particles. Analysis of 3D magnetic structures for these particle shapes is more complex than for cubic particles (see, e.g., comments in [1]).

Below, we suggest an effective iterative solver for cylindrical particles. A realistic representation of a ferromagnetic cylinder assumes its finite height, $z = h$, and radius, $r = a$; i.e., the 3D case is considered. The parameters chosen in numerical calculations correspond to those commonly used for patterned magnetic media [15]. The easy axis is the cylinder axis (z).

The magnetic structure is calculated by solving the Landau–Lifshitz–Gilbert (LLG) equation (see, e.g., in [1, 2, 16]) for natural movement of magnetization. For convenience in the calculations, we consider \mathbf{M} as a normalized (to M_s , which is the saturated value) magnetization vector. Thus, the normalized LLG equation is represented by [1]

$$(1 + \alpha^2) \frac{d\mathbf{M}}{dt} = \mathbf{H} \times \mathbf{M} - \alpha \mathbf{M} \times (\mathbf{M} \times \mathbf{H}), \quad (1)$$
$$\mathbf{M}|_{t=0} = \mathbf{M}_0,$$

where α is the dimensionless damping factor introduced to specify quasi-local dissipative phenomena. This factor depends nonlinearly on the magnetization [17], but in the search for the magnetization structures

¹ This article was submitted by the authors in English.

this dependence can be omitted. The “classical” form of LLG Eq. (1) does not include gradient dissipation terms [18]. These terms play an important role for the line broadening in ferromagnetic resonance; however, they are unimportant for the analysis of magnetization structures.

The quantity \mathbf{H} in Eq. (1) is a local effective magnetic field. This field is given by the sum of anisotropy, exchange and demagnetization fields:

$$\mathbf{H} = -4\pi Q\mathbf{M}_\perp + \mathbf{H}^{(ex)} + \mathbf{H}^{(m)}. \quad (2)$$

Here, $Q = K_a/2\pi M_s^2$ and K_a is the anisotropy constant. The Laplacian in the expression for exchange field $\mathbf{H}^{(ex)} = 2\Delta\mathbf{M}$ is also written in dimensionless form, and the characteristic exchange length $\lambda = \sqrt{A}/M_s$ (A is the exchange stiffness constant) is used to normalize the derivations.

The main complexity refers to the demagnetizing field $\mathbf{H}^{(m)}$ created by \mathbf{M} . This field can be defined from the magnetostatic problem [16]:

$$\text{rot } \mathbf{H}^{(m)} = 0, \quad \text{div}(\mathbf{H}^{(m)} + 4\pi\mathbf{M}) = 0 \quad (3)$$

with proper jump (the absence of magnetization outside the particle) and boundary conditions (continuity of the tangential component of \mathbf{H} and normal component of magnetic induction \mathbf{B}). It is convenient to rewrite Maxwell Eqs. (3) using potential U , $\mathbf{H}^{(m)} = -\text{grad } U$. This potential can be represented in the form $U = F + V$, where the functions F and V satisfy the Poisson and Laplace equations:

$$\Delta F = 4\pi \text{div} \mathbf{M}, \quad \Delta V = 0. \quad (4)$$

The Poisson equation (4) is solved with zero boundary condition on the surface. The boundary condition for the Laplace equation provides continuity of the magnetic induction:

$$F|_S = 0, \quad V|_S = \oiint_S \left[-\frac{1}{4\pi} \frac{\partial F(\mathbf{r}_1)}{\partial \mathbf{n}} + \mathbf{M}(r_1)\mathbf{n} \right] \frac{d\mathbf{r}_1}{|\mathbf{r} - \mathbf{r}_1|}. \quad (5)$$

Unit vector \mathbf{n} in Eq. (5) characterizes the normal to the surface. Micromagnetic equations are typically solved by the finite difference technique using some equidistant 3D grid in Cartesian coordinates [19]. This technique is more suitable for parallelepiped particle shapes, although the problem of the corner singularities [20, 21] is open with this technique.

We suggest the iterative technique, which permits one to solve Eqs. (1)–(5) for a cylinder. We used the analytical solution of Maxwell equations (4), (5) for a fixed distribution of magnetization \mathbf{M} . The solution is written in cylindrical coordinates, $\{r, \varphi, z\}$, where z is the cylinder axis. Keeping in mind the periodic condi-

tions, $U(r, \varphi + 2\pi, z) = U(r, \varphi, z)$, one can use the Fourier series:

$$U(r, \varphi, z) = \frac{1}{2}[V_0(r, z) + W_0(r, z)] + \sum_{k=1}^{\infty} [(V_k + W_k)\cos k\varphi + (S_k + E_k)\sin k\varphi], \quad (6)$$

$$\mathbf{M}(r, \varphi, z) = \frac{1}{2}\mathbf{A}_0(r, z) + \sum_{k=1}^{\infty} [\mathbf{A}_k(r, z)\cos k\varphi + \mathbf{B}_k(r, z)\sin k\varphi]. \quad (7)$$

Here, we do not assume axial symmetry of the solution in advance. The relation of the coefficients in Eq. (6) to the Fourier components of magnetization vector (7) can be found with the Hankel and Fourier transforms (with respect to r and z coordinates):

$$\begin{aligned} V_k(r, z) &= \int_0^{\infty} ds s J_k(sr) \tilde{V}_k(s, z), \\ S_k(r, z) &= \int_0^{\infty} ds s J_k(sr) \tilde{S}_k(s, z), \\ \tilde{V}_k(s, z) &= \frac{1}{2} \int_0^z dz_1 e^{-s(z-z_1)} \Pi_k(s, z_1) \\ &\quad - \frac{1}{2} \int_z^h dz_1 e^{s(z-z_1)} \Pi_k(s, z_1), \\ \tilde{S}_k(s, z) &= \frac{1}{2} \int_0^z dz_1 e^{-s(z-z_1)} T_k(s, z_1) \\ &\quad - \frac{1}{2} \int_z^h dz_1 e^{s(z-z_1)} T_k(s, z_1), \\ \Pi_k(s, z) &= 4\pi \int_0^{\infty} dr_1 r_1 J_k(sr_1) A_{kz}(r_1, z), \\ T_k(s, z) &= 4\pi \int_0^{\infty} dr_1 r_1 J_k(sr_1) B_{kz}(r_1, z), \end{aligned} \quad (8)$$

where A_{kz} and B_{kz} are the z components of the corresponding vector coefficients in Eq. (7). In a similar way,

the coefficients W_k and E_k are expressed as

$$\begin{aligned}
 W_k(r, z) &= \frac{1}{\pi} \int_0^\infty \int_0^\infty d\omega d\xi \cos \omega(\xi - z) \tilde{W}_k(\omega, r, \xi), \\
 E_k(r, z) &= \frac{1}{\pi} \int_0^\infty \int_0^\infty d\omega d\xi \cos \omega(\xi - z) \tilde{E}_k(\omega, r, \xi), \\
 \tilde{W}_k(\omega, r, \xi) &= -4\pi K_k(\omega r) \int_0^r dr_1 \{ k [B_{k\phi}(r_1, \xi) \\
 &\quad - A_{kr}(r_1, \xi)] I_k(\omega r_1) - r\omega A_{kr}(r_1, \xi) I_{k+1}(\omega r_1) \} \\
 &\quad - 4\pi I_k(\omega r) \int_r^a dr_1 \{ k [B_{k\phi}(r_1, \xi) \\
 &\quad - A_{kr}(r_1, \xi)] K_k(\omega r_1) + r\omega A_{kr}(r_1, \xi) K_{k+1}(\omega r_1) \}, \\
 \tilde{E}_k &= 4\pi K_k(\omega r) \int_0^r dr_1 \{ k [A_{k\phi}(r_1, \xi) \\
 &\quad + B_{kr}(r_1, \xi)] I_k(\omega r_1) + r\omega B_{kr}(r_1, \xi) I_{k+1}(\omega r_1) \} \\
 &\quad + 4\pi I_k(\omega r) \int_r^a dr_1 \{ k [A_{k\phi}(r_1, \xi) \\
 &\quad + B_{kr}(r_1, \xi)] K_k(\omega r_1) - r\omega B_{kr}(r_1, \xi) K_{k+1}(\omega r_1) \},
 \end{aligned} \tag{9}$$

where J_k , K_k , and I_k are Bessel functions.

To combine this result with the LLG equation, we use the solution of Eq. (1) with a fixed effective field and perform iteratively one time step ($n, n+1$) at a fixed grid point $\{r_i, \phi_j, z_k\}$:

$$\begin{aligned}
 \mathbf{M}(t) &= \frac{1}{C(t_n, t)} [\mathbf{H}_1(t_n, t) + \mathbf{H}_2(t_n, t)], \\
 \mathbf{H}_1 &= \left\{ |\mathbf{H}| + \mathbf{H}\mathbf{M}_n \right. \\
 &\quad \left. + (\mathbf{H}\mathbf{M}_n - |\mathbf{H}|) \exp \left[-\frac{2\alpha}{1+\alpha^2} F(t_n, t) \right] \right\} \mathbf{H}, \\
 \mathbf{H}_2 &= 2 \exp \left[-\frac{\alpha}{1+\alpha^2} F(t_n, t) \right] \\
 &\quad \times \left\{ |\mathbf{H}| \sin \left[\frac{F(t_n, t)}{1+\alpha^2} \right] \mathbf{H} \times \mathbf{M}_n \right.
 \end{aligned} \tag{10}$$

$$\left. - \cos \left[\frac{F(t_n, t)}{1+\alpha^2} \right] \mathbf{H} \times (\mathbf{H} \times \mathbf{M}_n) \right\},$$

$$C = \left\{ |\mathbf{H}| + \mathbf{H}\mathbf{M}_n - (\mathbf{H}\mathbf{M}_n - |\mathbf{H}|) \right.$$

$$\left. \times \exp \left[-\frac{2\alpha}{1+\alpha^2} F(t_n, t) \right] \right\} |\mathbf{H}|,$$

$$F(t_n, t) = \int_{t_n}^t dt_1 |\mathbf{H}(t_n, t_1)|,$$

where $\mathbf{M}_n = \mathbf{M}(t_n)$. Solution (10) is obtained by writing Eq. (1) in cylindrical coordinates with the z axes directed along the vector \mathbf{H} .

The magnetic vector field $\mathbf{H}(t_n, t)$ in Eq. (10) is an effective magnetic field from the previous step, i.e., $\mathbf{H}(t_n)$. Therefore, the additional time-dependent external magnetic field $\mathbf{H}^{(\text{ext})}(t)$ can be included as $\mathbf{H}(t_n, t) = \mathbf{H}(t_n) + \mathbf{H}^{(\text{ext})}(t)$. Field $\mathbf{H}^{(\text{ext})}(t)$ can be used for the simulation of the magnetic writing process or for the simulation of the influence of neighboring magnetic elements within the array of patterned media.

We should emphasize that the computer calculation of a sequence of onefold integrals in Eqs. (7)–(9) is very fast. For a small cylinder, one needs about 1 sec with a computer of 1.2 GHz to calculate the field at all grid points $\{r_i, \phi_j, z_k\}$ for one iteration step. Naturally, the computation time depends on the total grid number, and big particles need a longer time. We used a combination of calculation with software written in the FORTRAN and Mathematica languages.

The final magnetization (in zero field at $t \rightarrow \infty$) depends on the magnetization history, i.e., on the initial distribution \mathbf{M}_0 and precise external magnetic pulse shape, $\mathbf{H}^{(\text{ext})}(t)$. In this paper, we do not discuss the dynamic processes and set $\mathbf{H}^{(\text{ext})}(t) = 0$. The following initial magnetizations are typically considered. (1) A random initial state is used for finding the lowest minima of Gibbs energy. It is also used for modeling the “erasing head” work, e.g., with fast laser heating of a material above the Curie temperature [22]. (2) $\mathbf{M}_0 = \mathbf{M}_z = \{0, 0, 1\}$ and/or (3) $\mathbf{M}_0 = \mathbf{M}_r = \{1, 0, 0\}$ are used for modeling the transverse recording [23, 24]. (4) $\mathbf{M}_0 = \mathbf{M}_x = \{\cos \phi, -\sin \phi, 0\}$ is used for modeling the longitudinal recording [24].

The iterative method of solving the LLG equation was first proposed for a cubic particle [25]. In this case, complicated Eqs. (6)–(9) are replaced by simpler formulas for the Fourier expansion in Cartesian coordinates. For small-sized cubic particles, the remanent state has a flowerlike or vortex magnetization configu-

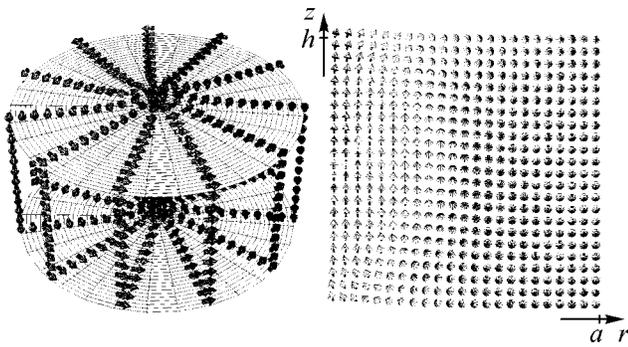


Fig. 1. Side faces of the vortex configuration for a cylinder with $a = 5\lambda$, $h = 10\lambda$, $Q = 10^{-3}$. Cross section shows magnetization in the $\{r, z\}$ plane. Random initial state.

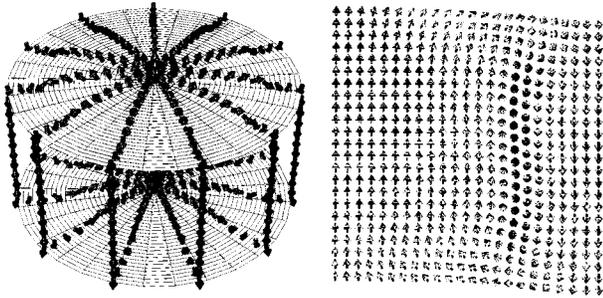


Fig. 2. Single-domain state with concentric Bloch wall along the z direction. This wall is clearly seen in the $\{r, z\}$ cross section. Input parameters are: $a = 5\lambda$, $h = 10\lambda$, $Q = 0.2$. Random initial state.

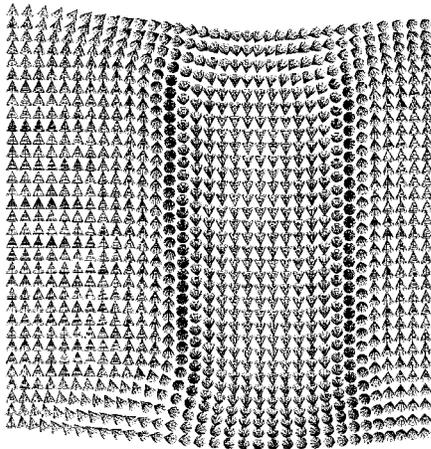


Fig. 3. The $\{r, z\}$ cross-section images of a two-domain structure with concentric Bloch domain walls. Parameters: $a = 10\lambda$, $h = 20\lambda$, $Q = 0.3$. Vortex initial state.

ration, which is in good agreement with many previous calculations and experimental results.

In contrast to cubic particles, there are just a few results for the magnetic states in a small cylinder [1, 2, 26, 27]. Experimental results [12, 13] demonstrate that

for cylindrical particles (with a low-anisotropy material) the flowerlike configurations are not favorable. It can be understood along the line that the magnetization tends to avoid magnetic charges except for two topological “singularities” [1, 26, 27].

The typical lattice size in calculations was changed from $20 \times 20 \times 20$ to $120 \times 120 \times 120$. The parameters of the material were varied in the range $Q \in [10^{-3}, 1]$, and the cylinder geometry was varied within $\{a, h\} \in [\lambda, 30\lambda]$. The magnetic behavior of the cylinder was studied as a function of size, aspect ratio a/h , and the prehistory. Below, we present the main results of calculations.

1. Low anisotropy, $Q \leq 10^{-2}$. With random initial conditions (and/or with $\mathbf{M}_0 = \mathbf{M}_r$), the cylinder with $a, h \leq 30\lambda$ has a stable vortex configuration (Fig. 1). The stability of a single magnetic vortex for a small isotropic ferromagnetic cylinder was recently discussed in [28]. The vortex in Fig. 1 has a rather spiral structure; i.e., $M_r \neq 0$. The relative volume of homogeneous state with $M_z \approx 1$ increases with diminishing cylinder radius. As a result, the flowerlike structure coexists with the vortex solution for a small cylinder at $a, h \leq 2\lambda$. With $\mathbf{M}_0 = \mathbf{M}_x$ for sufficiently big cylinders ($a > 10\lambda$), one can find the axially asymmetric solutions with spiral branches, similar to those which were discussed in [29]. These solutions become more distinct with increasing anisotropy.

2. Intermediate anisotropy, $10^{-2} < Q < 0.5$. With $a, h \geq 2\lambda$, magnetization tends to create a magnetic bubble, with the Bloch-type domain wall; see in Fig. 2. This magnetic bubble develops monotonically from the vortex state. With increasing radius, $a \geq 8\lambda$, two concentric domain walls appear; see Fig. 3. For tall cylinder with $a/h \ll 1$ and $a \geq 10\lambda$, the internal magnetic domains are formed around the points of “singularity” ($r = 0, z = 0, h$) and do not grow inside the center of the cylinder. With longitudinal initial magnetization $\mathbf{M}_0 = \mathbf{M}_x$ and sufficiently large particles, $a \geq 30\lambda$, the complex magnetization configurations with broken axial symmetry can develop. The other initial conditions with nonzero radial component also yield complex multidomain structures for big particles. The domains can be formed in the z directions. A vortex structure with different chirality on the top and the bottom can be obtained for the initial state $\mathbf{M}_0 = \mathbf{M}_z$. A change in chirality occurs through the nonsymmetrical Bloch domain wall, typical of weak anisotropy materials [1, 2]. This state seems to be metastable.

3. High anisotropy, $Q \geq 0.5$. The “flower” state, Fig. 4, is favorable for a small cylinder at the $\mathbf{M}_0 = \mathbf{M}_z$ initial condition. This configuration is stable with respect to small magnetic perturbations. A single-domain cylindrical particle has “critical size,” $a \approx 2\lambda$, when it remains “uniformly” magnetized. It is near the limit of micromagnetic model validity. Thus, the application of the micromagnetic model to such sizes is questionable;

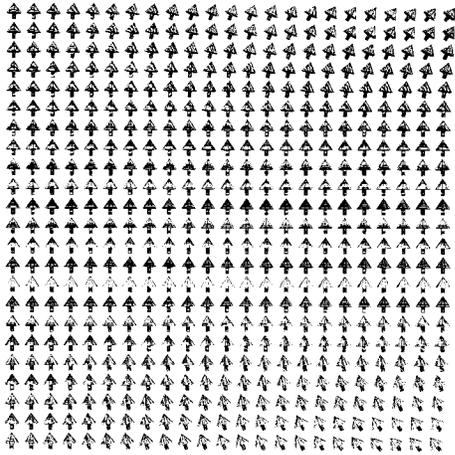


Fig. 4. The “flower” state. Parameters: $a = 5\lambda$, $h = 10\lambda$, $Q = 0.5$. Initial state $\mathbf{M}_0 = \mathbf{M}_z$.

it should be proved additionally by quantum mechanics calculations.

In conclusion, we state that the energetic consideration, which presents the “phase diagram” of different states in a particle [1, 2], is sometimes insufficient for the identification of states in the magnetic recording problem. For example, authors [26] consider magnetic bubble improbable for media with uniaxial anisotropy. Meanwhile, the integration of the LLG equation demonstrates that this state (Fig. 2), on the contrary, is very probable for a wide class of initial conditions. Metastable states in magnetization can be separated by sufficiently large energy barriers. These states represent attractors with some “trapping region” with respect to the initial conditions. The dynamic theory based on the solution of Landau–Lifshitz equation has no problem with the identification of the corresponding states, while the pure energetic consideration that ignores the magnetization prehistory may lead to some confusion.

The authors are very grateful to Profs. S.I. Anisimov and E.I. Katz for useful discussions.

REFERENCES

1. A. Hubert and R. Schäfer, *Magnetic Domains. The Analysis of Magnetic Microstructures* (Springer-Verlag, Berlin, 1998).
2. A. Aharoni, *Introduction to the Theory of Ferromagnetism* (Oxford Univ. Press, Oxford, 2001).
3. J. F. Smyth, S. Schultz, D. R. Fredkin, *et al.*, *J. Appl. Phys.* **69**, 5262 (1991).
4. K. J. Kirk, J. N. Chapman, S. McVitie, *et al.*, *J. Appl. Phys.* **87**, 5105 (2000).
5. M. E. Schabes and H. M. Bertram, *J. Appl. Phys.* **64**, 1347 (1988).
6. W. Rave, K. Fabian, and A. Hubert, *J. Magn. Magn. Mater.* **190**, 332 (1998).
7. L. Folks and C. R. Woodward, *J. Magn. Magn. Mater.* **190**, 28 (1998).
8. E. Zueco, W. Rave, R. Shafer, *et al.*, *J. Magn. Magn. Mater.* **190**, 42 (1998).
9. W. Wulfhekel, H. F. Ding, W. Lutzke, *et al.*, *Appl. Phys. A: Mater. Sci. Process.* **A72**, 463 (2001).
10. J. Lohau, A. Moser, C. D. Rettner, *et al.*, *Appl. Phys. Lett.* **78**, 990 (2001).
11. A. Fernandez, M. R. Gibbons, M. A. Wall, and C. J. Cerjan, *J. Magn. Magn. Mater.* **190**, 71 (1998).
12. C. A. Ross, M. Farhoud, M. Hwang, *et al.*, *J. Appl. Phys.* **89**, 1310 (2001).
13. A. Lebib, S. P. Li, M. Natali, and Y. Chen, *J. Appl. Phys.* **89**, 3892 (2001).
14. V. P. Puntès, K. V. Krishnan, and P. Alivisatos, *Appl. Phys. Lett.* **78**, 2187 (2001).
15. *Magneto-optical Recording Materials*, Ed. by R. J. Gambino and T. Suzuki (IEEE Press, New York, 2000).
16. E. M. Lifshitz and L. P. Pitaevsky, *Course of Theoretical Physics, Vol. 5: Statistical Physics* (Nauka, Moscow, 1978; Pergamon, New York, 1980), Part 2.
17. J. Fidler and T. Schrell, *J. Phys. D* **33**, R135 (2000).
18. V. G. Baryakhtar, B. A. Ivanov, A. L. Sukstanskii, and E. Y. Melikhov, *Phys. Rev. B* **56**, 619 (1997).
19. W. Rave and A. Hubert, *IEEE Trans. Magn.* **36**, 3886 (2000).
20. W. Rave, K. Ramstöck, and A. Hubert, *J. Magn. Magn. Mater.* **183**, 329 (1998).
21. A. Thiaville, D. Tomas, and J. Miltat, *Phys. Status Solidi A* **170**, 125 (1998).
22. J. Gudde, U. Conrad, V. Jahuke, *et al.*, *Phys. Rev. B* **59**, R6608 (1999).
23. K. Okuda, K. Sueoka, and K. G. Ashar, *IEEE Trans. Magn.* **24**, 2479 (1988).
24. H. N. Bertram, *Theory of Magnetic Recording* (Cambridge Univ. Press, Cambridge, 1994).
25. V. Ternovsky, J. P. Wang, and T. C. Chong, in *Proceedings of the 8th Magnetism and Magnetic Materials, Intermag Conference, San Antonio, 2001*, p. 340.
26. A. S. Arrott and T. L. Templeton, *Physica B (Amsterdam)* **233**, 259 (1997).
27. A. Hubert and W. Rave, *J. Magn. Magn. Mater.* **184**, 67 (1998).
28. K. Yu. Guslenko and K. L. Metlov, *Phys. Rev. B* **63**, 100403 (2001).
29. A. B. Borisov, *Pis'ma Zh. Éksp. Teor. Fiz.* **73**, 279 (2001) [*JETP Lett.* **73**, 242 (2001)].

Electron Gas Compression and Coulomb Explosion in the Surface Layer of a Conductor Heated by Femtosecond Laser Pulse

S. I. Kudryashov* and V. I. Emel'yanov

International Laser Center, Faculty of Physics, Moscow State University, Moscow, 119899 Russia

* e-mail: sergeikudryashov@chat.ru

Received May 14, 2001

The hypothesis is put forward on the basis of experimental data that strong inhomogeneous heating of the skin layer of conducting materials by a femtosecond pulse gives rise to a double electrical layer that is formed of a “surface” layer of positive ions and a thin (about 1 nm) “subsurface” layer of a superdense (10^{23} – 10^{25} cm $^{-3}$) degenerate electron gas. The double layer breaks within one picosecond through the Coulomb explosion.
© 2001 MAIK “Nauka/Interperiodica”.

PACS numbers: 79.20.Ds; 78.47.+p

The interaction of ultrashort laser pulses with conductors is traditionally considered using a two-temperature model (TTM) [1], which treats the electron and phonon subsystems separately on the electron–phonon interaction time scale (on the order of picoseconds for most metals). The vast majority of experimental studies devoted to the characteristic parameters of this model (electron–electron relaxation time, electron–gas thermalization time, and electron–phonon coupling constant [2–4]) and to the energy transport in a conductor through the ballistic flight of hot electrons and through electron heat conductivity [2, 5] are carried out at low fluences F of heating beam, no higher than 10 mJ cm $^{-2}$, which corresponds to peak electron temperatures T_e on the order of 10^4 K, i.e., considerably lower than the typical Fermi temperature $T_F \sim 10^5$ K. Meanwhile, the experimental results obtained in [6, 7] at higher values $F \geq 0.1$ J cm $^{-2}$ (peak $T_e \sim T_F$) indicate that a different type of lattice–electron interaction may take place upon the subpicosecond disordering of a metal crystal lattice (“electron melting”). This calls for further analysis of the conditions for the validity of the TTM. Moreover, a sharp pressure gradient ∇P_e occurring at peak temperatures $T_e \sim T_F \sim 10^5$ K in the electron gas within the skin layer of a conductor may induce the redistribution of electron–gas density in the skin layer and the formation of a double electrical layer (DEL). The electric field $E = QV T_e$ arising in a DEL with characteristic width $\alpha^{-1} \sim 10^{-8}$ m owing to the thermoelectric effect is determined by the differential thermal emf $Q \approx 10^{-8} T \sim 10^{-3}$ V K $^{-1}$ (at $T_e \sim 10^5$ K) and is comparable with the intraatomic values $E_{at} \sim 10^{11}$ V m $^{-1}$. To our knowledge, the physics of these phenomena has not yet been explored, neither experimentally nor theoret-

ically, under the conditions of ultrafast inhomogeneous heating of conductors to peak values $T_e \sim T_F$.

In this work, the transient optical characteristics and the dynamics of mass transfer in the skin layers of metals (Al and Cu), as well as the dynamics of the laser-induced metallic layers in melts of semiconductors (Si and GaAs) heated by 100-fs laser pump pulses, were studied by ellipsometry (self-reflection of s - and p -polarized pump pulses), time-resolved optical microscopy, and stationary interference microscopy.

A standard femtosecond laser setup of the Institute of Laser and Plasma Physics (University of Essen, Germany) was used, whose components (oscillator and the regenerative and multipass sapphire amplifiers) are described in [8]. The parameters of laser radiation that was led into the pump and probe channels were as follows: first-harmonic (FH) wavelength ($\hbar\omega = 1.56$ eV) $\lambda \approx 800$ nm, pulse (Gaussian) duration $\tau \approx 100$ fs, pulse energy (TEM $_{00}$) 1.5 mJ, repetition rate 10 Hz, and relative amplitude of the pre- and postpulses no higher than 5–7%. In the pump channel, the focused linearly s - or p -polarized FH beam was directed at an angle of 45° onto a target that was moved from pulse to pulse. The energy of the mirror-reflected s - or p -polarized FH beam was measured by a pyroelectric detector at different energies of the single incident pulses. Al and Cu films (30–40 nm thick) on glass and optically “thick” metallic films of the l -Si and l -GaAs melts (~ 20 nm thick) prepared at the surfaces of bulk undoped samples of these materials ([100] orientation) under the action of a single pump pulse with effective fluence F_{eff} far above the nonthermal melting threshold ($F_{\text{eff}}^M \approx 0.25$ J cm $^{-2}$ [9] and 0.15 J cm $^{-2}$, respectively) were used as targets. In the probe channel, the FH beam was led through an

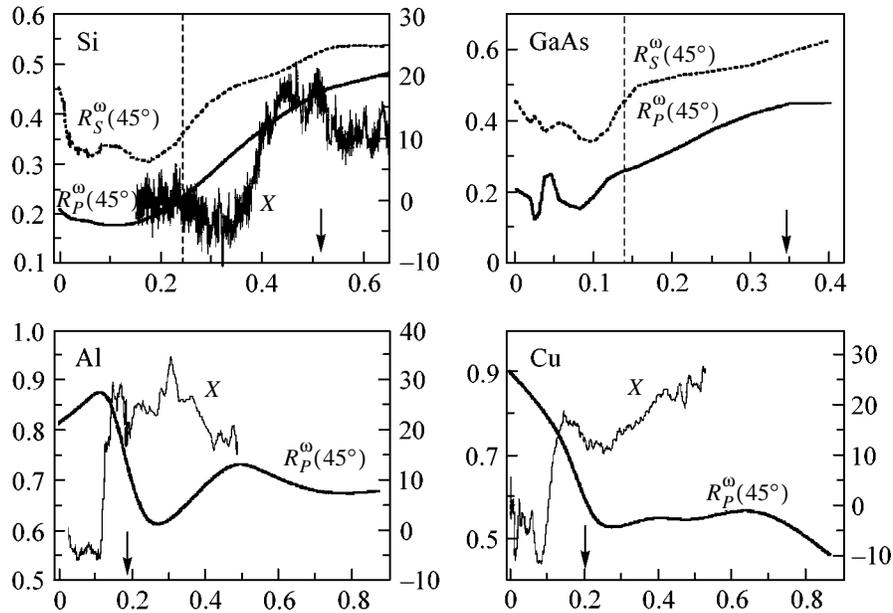


Fig. 1. (Left ordinate axes) self-reflectivities R_S^0 and R_P^0 as functions of instantaneous effective (absorbed) fluence $F_{\text{eff}}(t)$ of pump beam for the Cu, Al, Si, and GaAs samples and (right ordinate axes) crater depths (profiles) X (in nm) as functions of integrated F_{eff} for the Cu, Al, and Si samples. The vertical dashed lines indicate the positions of the melting thresholds during the pump pulse for Si and GaAs, and the arrows indicate the onset of anomalous optical properties (R_S^0 and R_P^0 plateaus) for all samples.

optical delay line whereupon the frequency doubled, and the second harmonic (SH, $\hbar 2\omega = 3.12$ eV) was used at normal incidence for probing, through the objective of an optical microscope, the target area excited by the FH. The mirror-reflected probe SH images of the target were photographed by a synchronized CCD camera for several time delays of the probe pulse relative to the pump pulse and, after normalization to the image of the unexcited sample and calibration, represented the two-dimensional distribution of sample reflectivity $R^{2\omega}$ symmetric about the center of the pump beam spot. The characteristics of the probe channel allowed the target reflectivity transients to be studied with a spatial resolution of $2 \mu\text{m}$ and a time resolution of 100 fs. The morphology and the depth profiles X of the craters formed at the target surface by the pump beam were studied by stationary interference microscopy with a depth resolution up to 2 nm.

The self-reflectivities of all samples were experimentally studied as functions of pump fluence for the p -polarized beam (and also s -polarized beam for Si and GaAs). The resulting R_P^0 functions (R_P^0 and R_S^0 for Si and GaAs) were processed to eliminate their spatial averaging caused by the inhomogeneous distribution of pump fluence F over the TEM₀₀ beam spot on the target. In addition, to eliminate the time averaging of R_P^0 and R_S^0 during the laser pulse, the R_P^0 and R_S^0 depen-

dences on $F_{\text{eff}} = (1 - R_{S,P}^0)F$ were processed graphically as in [9]. Recall that this procedure is based on the assumption that R_S^0 and R_P^0 depend on the instantaneous effective fluence $F_{\text{eff}}(t)$ during the pulse rather than on the intensity $I_{\text{eff}}(t)$. This assumption may be considered to be valid for ultrashort pulses with moderate $F_{\text{eff}} \leq 1\text{--}10 \text{ J cm}^{-2}$ acting on metals. The resulting $R_S^0(F_{\text{eff}}(t))$ and $R_P^0(F_{\text{eff}}(t))$ functions are presented in Fig. 1.

The $R_P^0(F_{\text{eff}}(t))$ curves in Fig. 1 show plateaus at 0.5–0.6 for all types of metal films exposed to the pulses with characteristic $F_{\text{eff}}^1(t) \approx 0.2 \text{ J cm}^{-2}$ (hereafter F_{eff}^1 ; for l -Si and l -GaAs, counted from the melting threshold). An analogous effect was earlier observed at close F_{eff} values for the laser-induced melts of Al and Ag films and explained, within the framework of TTM, by the fact that the characteristic time τ_{ee} of electron–electron scattering decreased drastically (to 10^{-16} s) upon heating the electron subsystem to 10^5 K [6]. At the same time, the $X(F_{\text{eff}})$ curves in Fig. 1 show a sharp drop in the crater depth down to the skin depth near F_{eff}^1 . This is evidence for a more complicated character of the processes occurring in the skin layer of such a strongly heated conductor.

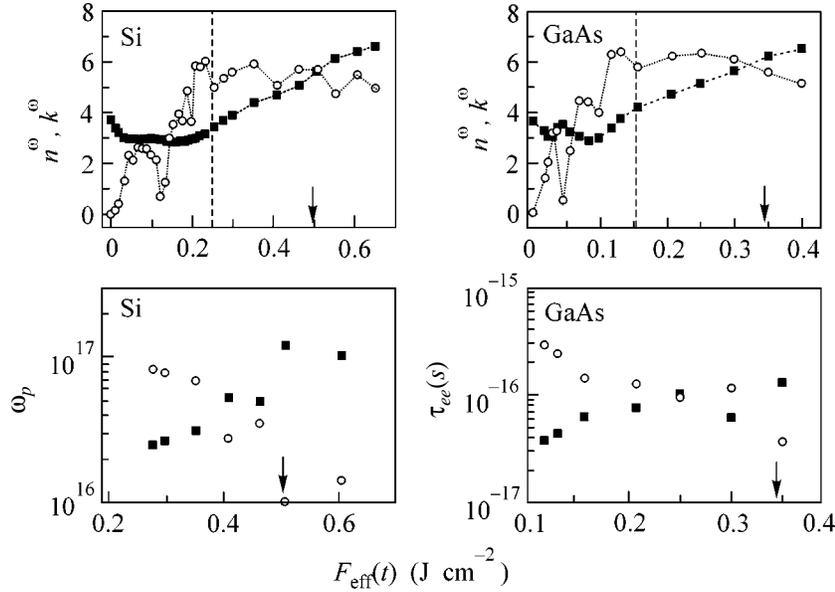


Fig. 2. Optical constants n^ω (dark squares), k^ω (light circles), ω_p (dark squares), and τ_{ee} (light circles) as functions of $F_{\text{eff}}(t)$ for the Si and GaAs samples (the ω_p and τ_{ee} axes for both samples coincide). The vertical dashed lines indicate the melting thresholds during the pump pulse for Si and GaAs, and the arrows indicate the onset of anomalous optical properties ($n^\omega > k^\omega$ and a sharp change in ω_p and τ_{ee}) for these samples.

The $R_S^\omega(F_{\text{eff}}(t))$ and $R_P^\omega(F_{\text{eff}}(t))$ dependences obtained in this work for the conducting l -Si and l -GaAs films were used to study their optical characteristics n^ω and k^ω near F_{eff}^1 (Fig. 2). The calculations were carried out using the Fresnel formulas by choosing pairs of n^ω and k^ω values minimizing the discrepancy between the experimental and calculated values of R_S^ω and R_P^ω . The fact that $n^\omega(F_{\text{eff}}(t))$ continuously increases while k^ω is constant above the melting threshold of the samples studied indicates that the optical properties of the substances are nonuniform (have effective character) within the skin layer. At the same time, the fact that $n^\omega > k^\omega$ at $F_{\text{eff}}(t) \geq F_{\text{eff}}^1$ indicates that a dielectric phase with positive real part of the dielectric constant appears within the skin layer. With the aim of elucidating the nature of the observed anomalies, the effective values of n^ω and k^ω were used to obtain the corresponding dependences of the effective plasma frequency ω_p and τ_{ee} on $F_{\text{eff}}(t)$ (Fig. 2). It is seen from these curves that ω_p increases sharply and τ_{ee} decreases at $F_{\text{eff}}(t) \geq F_{\text{eff}}^1$. Since the measured electron optical mass is unity for l -Si and l -GaAs, the observed substantial monotonic increase in $\omega_p(F_{\text{eff}}(t))$ can be associated with the respective increase in the effective electron density N_e in the skin layer and, hence, increase in the effective Fermi energy E_{F0} (Fig. 3). In the Fermi liquid theory [10], a decrease in $\tau_{ee}(F_{\text{eff}}(t))$ at $F_{\text{eff}}(t) \geq F_{\text{eff}}^1$ can be caused by the increase in ω_p and T_e (Fig. 3). The lat-

ter, like the corresponding electron–electron scattering constants K_{ee} (2.5 and 1.5 fs $^{-1}$ eV $^{-2}$ for l -Si and l -GaAs, respectively), were calculated by formulas

$$\tau_{ee} = K_{ee}^{-1} \frac{1 + \exp[-(\hbar\omega/kT_e)]}{(\pi kT_e)^2 + (\hbar\omega)^2}, \quad (1)$$

$$K_{ee} = C\omega_p/E_{F0}^2 \quad (2)$$

using the functions $\omega_p(F_{\text{eff}}(t))$ and $E_{F0}(F_{\text{eff}}(t))$ obtained in this work (Figs. 2, 3) and the fitting parameter C estimated at $F_{\text{eff}}(t) \approx F_{\text{eff}}^M$ (i.e., at $T_e \approx 0$).

The dependences of N_e , E_{F0} , and T_e on $F_{\text{eff}}(t)$ (Fig. 3) suggest that, despite a substantial rise in T_e during the pump pulse, electron gas in the skin layers of the metallic l -Si and l -GaAs films remains degenerate because of an increase in its effective density N_e , likely, owing to the thermoelectric effect [11]. Indeed, an increase in $T_e(F_{\text{eff}}(t))$ gives rise to the temperature gradient ∇T_e and the gradient of electron-gas pressure ∇P_e in the skin layers of the l -Si and l -GaAs films, where

$$P_e(T_e) = 2/5 N_e E_{F0} [1 + 5\pi^2/12 (kT_e/E_{F0})^2 + \dots]. \quad (3)$$

Inasmuch as the electron gas in the skin layer drifts along the ∇P_e gradient, the N_e distribution shifts from the sample surface so that the “effective surface” of skin layer, i.e., the surface of maximal energy release, shifts deeper into the skin layer and ∇P_e sharpens. As a result, the effective skin surface shifts deeper into the

skin layer even more rapidly than the deeper lying layers of the moving N_e distribution, ∇P_e sharpens further, and the actual skin depth decreases further; i.e., electron gas becomes compressed, and the corresponding local absorption coefficient increases. Evidently, the trailing edge of the N_e distribution driven by ∇P_e into the bulk propagates faster than its leading edge and, hence, gradually forms a nonstationary “shock” wave P_e with compression of electron gas at its leading edge. The exponential character of the N_e , E_{F0} , and T_e functions of instantaneous $F_{\text{eff}}(t)$ (Fig. 3), which are represented as sets of instantaneous values of these parameters during the laser pulse, confirms the nonstationary interpretation given above for this phenomenon. At the same time, it is evident that an increase in the degree of ion screening by electron gas with increasing N_e in the propagating wave will result in the saturation of absorption at the effective skin surface with a certain N_e^{max} , and the further shock-wave evolution under the action of the laser pulse will be due to an increase in the spatial width of a compressed electron-gas layers with density N_e^{max} and, likely, to the formation of a universal “plasma mirror” that was observed for a wide range of materials at $I \approx 10^{15} \text{ W cm}^{-2}$ [12]. Note that electron gas in the *l*-Si film was compressed 20-fold in our experiments, without any indications of saturation of the $N_e(F_{\text{eff}}(t))$ dependence (Fig. 3).

Clearly, as the shock wave P_e propagates, a DEL formed by a “surface” layer of positively charged ions and a “subsurface” layer of a superdense and “cold” degenerate electron gas may arise within the skin layer (Fig. 3). The thickness of the ion layer with density N_i and charge state z_i and the thickness of the electron layer can be estimated from the electroneutrality requirement $z_i N_i l_i \approx (N_e - N_e^0) l_e$, where $l_i + l_e \approx \alpha^{-1}$ and $N_e^0 \approx 2 \times 10^{23} \text{ cm}^{-3}$ is the equilibrium density of electron gas in the *l*-Si and *l*-GaAs films. For the average bulk energy density of about 10 eV atom⁻¹ deposited in the skin layer at $F_{\text{eff}}(t) \approx F_{\text{eff}}^1$, which is close to the first ionization potential of the Si, Ga, As, Al, and Cu atoms (8.2, 6.0, 9.8, 6.0, and 7.7 eV, respectively [13]), one has $l_e \approx \alpha^{-1} N_e^0 / N_e$ and $l_i \approx \alpha^{-1} (1 - N_e^0 / N_e)$, where N_e / N_e^0 is the compression ratio for the electron layer. For the compression ratios of 10–20 that were achieved for electron gas in this work, l_e equals ~ 1 nm and decreases with increasing N_e , whereas l_i increases to α^{-1} . For $F_{\text{eff}}(t) \approx F_{\text{eff}}^1$, the internal electric field in the DEL, according to the Poisson equation, is equal to $E \approx e N_i l_i / \epsilon_0 \approx 10^{13} \text{ V m}^{-1}$ immediately after the pump pulse (the corresponding potential difference $\Delta\phi \approx 10^5 \text{ V}$), which qualitatively agrees with the experimental data [14] on the carbon ions with energies up to 10 keV at close pump fluences, if it is taken into account that $\Delta\phi$

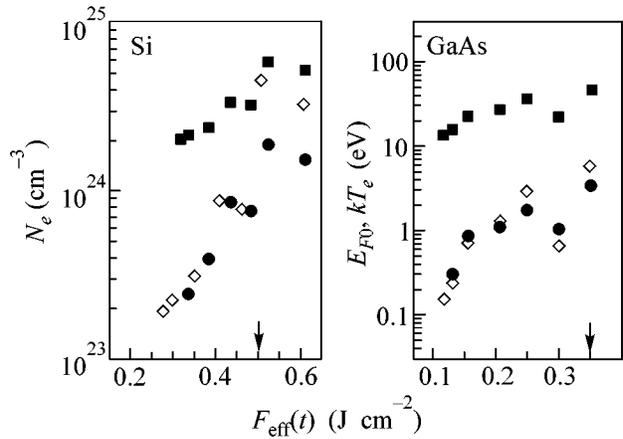


Fig. 3. N_e (light rhombi), E_{F0} (dark squares), and kT_e (dark circles) as functions of $F_{\text{eff}}(t)$ for the Si and GaAs samples (the ordinate axes for both samples coincide). The arrows indicate the onset of the N_e , E_{F0} , and kT_e anomalies during the pump pulse for the Si and GaAs samples.

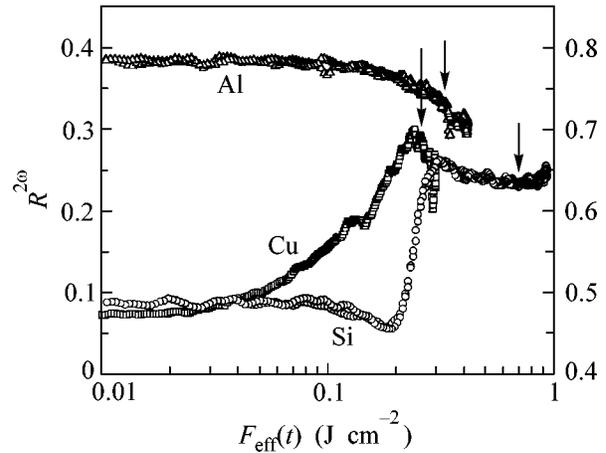


Fig. 4. Reflectivity $R^{2\omega}$ ($\hbar 2\omega = 3.12 \text{ eV}$; normal incidence) as a function of effective pump fluence F_{eff} for the (right axis) Al, Si, and (left axis) Cu samples with time delays of, respectively, 0.3, 0.5, and 0.5 ps from the beginning of pump pulse. The arrow indicates the subpicosecond DEL expansion threshold for the indicated samples.

decreases rapidly after the heating pulse because of the relaxation processes in the DEL.

During the first picosecond after the pump pulse, $R^{2\omega}$ decreases at $F_{\text{eff}}(t) \geq F_{\text{eff}}^1$ for the Cu and Al films and increases for the *l*-Si film (Fig. 4), likely, because of the removal of the heated skin layers of metals and of the melt from the surface of excited silicon. Specifically, due to the internal electric field, the DEL breaks through the “relief” (Coulomb explosion) of the unstable surface layer of positive ions, while the electron-gas compression vanishes due to the electron–electron repulsion and electron heat conductivity. A comparison

of the X values at $F_{\text{eff}}(t) \approx F_{\text{eff}}^1$ (Fig. 1) with the skin width α^{-1} (8–12 nm) for the *l*-Si, Cu, and Al films shows that the DEL breaking time, which restricts the duration of energy transfer from the skin layer to the bulk of the sample, is too short to retain the whole absorbed laser energy in the DEL plasma.

In summary, it has been assumed on the basis of the experimental data obtained in this work that the strong inhomogeneous heating of the skin layer of conducting materials during the laser pulse gives rise to a shock electron-pressure wave that creates a thin layer of a superdense (10^{23} – 10^{25} cm $^{-3}$) degenerate electron gas in the region of the skin boundary under the “surface” layer of positive ions. The resulting surface double electrical layer breaks within one picosecond through the Coulomb explosion and electron heat conductivity.

We are grateful to D. von der Linde, K. Sokolowski-Tinten, and V.V. Temnov for their assistance in the organization of the experiments and to the German Academic Exchanges Service for partial support (fellowship for S.I.K. in 2000).

REFERENCES

1. S. I. Anisimov, B. L. Kapeliovich, and T. L. Perel'man, *Zh. Éksp. Teor. Fiz.* **66**, 776 (1974) [*Sov. Phys. JETP* **39**, 375 (1974)].
2. R. H. M. Groeneveld, R. Sprik, and A. Legendijk, *Phys. Rev. B* **51**, 11433 (1995).
3. V. E. Gusev and O. B. Wright, *Phys. Rev. B* **57**, 2878 (1998).
4. M. Bonn, D. N. Denzler, S. Funk, *et al.*, *Phys. Rev. B* **61**, 1101 (2000).
5. C. Suarez, W. E. Bron, and T. Juhasz, *Phys. Rev. Lett.* **75**, 4536 (1995).
6. X. Y. Wang and M. C. Downer, *Opt. Lett.* **17**, 1450 (1992).
7. C. Guo and A. J. Taylor, *Phys. Rev. B* **62**, 5382 (2000).
8. *Femtosecond Laser Pulses: Principles and Experiments*, Ed. by C. Rulliere (Springer-Verlag, Berlin, 1998).
9. S. I. Kudryashov and V. I. Emel'yanov, *Pis'ma Zh. Éksp. Teor. Fiz.* **73**, 263 (2001) [*JETP Lett.* **73**, 228 (2001)].
10. D. Pines, *The Theory of Quantum Liquids* (Benjamin, New York, 1966).
11. A. A. Abrikosov, *An Introduction to the Theory of Normal Metals* (Nauka, Moscow, 1972), Chap. 6; in *Solid State Physics* (Academic, New York, 1972), Suppl. 12.
12. D. F. Price, R. M. More, R. S. Walling, *et al.*, *Phys. Rev. Lett.* **75**, 252 (1995).
13. *Physical Quantities: A Handbook*, Ed. by I. S. Grigor'ev and E. Z. Meilikhov (Énergoizdat, Moscow, 1991), Chap. 19.
14. H. Köster and K. Mann, *Appl. Surf. Sci.* **109/110**, 428 (1997).

Translated by V. Sakun

Direct Experimental Study of the Equilibrium Diffusion of Carbon Atoms between the (100)Mo Surface and the Bulk

N. R. Gall', E. V. Rut'kov, and A. Ya. Tontegode

Ioffe Physicotechnical Institute, Russian Academy of Sciences, Politekhnikeskaya ul. 26, St. Petersburg, 194021 Russia

e-mail: gall@ms.ioffe.rssi.ru

Received May 17, 2001

A direct experimental study of the diffusion of carbon atoms between the (100)Mo surface and the bulk has been carried out at process temperatures in the range 1400–2000 K, and the total balance of carbon atoms in the system has been determined. The difference in the activation energies of carbon dissolution and precipitation $\Delta E = E_{S1} - E_{1S}$ has been found under conditions of a dynamic equilibrium between both processes. This difference determines the temperature dependence of the degree of surface enrichment with carbon in reference to the bulk. The activation energy of the dissolution of carbon atoms has been determined in special experiments ($E_{S1} = 3.9$ eV), and the activation energy of the precipitation of carbon atoms E_{1S} has been calculated ($E_{1S} = 1.9$ eV), which turns out to be close to the energy of carbon bulk diffusion in molybdenum. © 2001 MAIK "Nauka/Interperiodica".

PACS numbers: 66.10.-x; 68.35.Fx

Numerous studies have been devoted to the diffusion of particles in solids [1–4]. Nevertheless, the regularities of the diffusion of atoms between the surface and the bulk have not been adequately studied [4, 5], though these regularities are important in all physical and chemical processes at the surface of heated bodies. The goal of this work was to study the regularities of the diffusion of carbon atoms between the surface and the bulk in thin molybdenum ribbons. Preliminarily, we comprehensively studied processes that occur during the interaction of carbon atoms with molybdenum over the wide temperature range 300–2000 K and a wide range of surface and bulk carbon concentrations [6, 7].

Experiments were carried out in an ultrahigh-vacuum (UHV, $P \sim 1 \times 10^{-10}$ torr) high-resolution Auger electron ($\Delta E/E \sim 0.1\%$) spectrometer with a prism energy analyzer [8]. The Auger spectra were measured directly in strongly heated (up to 2000 K) samples, which was of principal importance in studying equilibrium processes at medium and high temperatures. Thin molybdenum ribbons ($1 \times 0.02 \times 40$ mm³), which were carefully purified from possible impurities by high-temperature heating in an oxygen atmosphere and under UHV conditions, were used as samples. The texture of ribbons was expressed by the (100) face by no more than 99.5%; the surface was uniform in work function and possessed $e\phi = 4.45$ eV. Carbon was deposited on the surface from the source developed by the authors. The source created an atomic C flux free from carbon clusters [9]. The absolute flux density was determined by the procedure described in [7].

Consider experiments under diffusion equilibrium conditions, when the carbon atom flux from the surface

to the metal bulk v_{S1} is equal to the flux from the first near-surface layer onto the surface v_{1S} . For this purpose, carbon was deposited onto the ribbon at 300 K with $N_{\text{dep}} = 1.5 \times 10^{16}$ cm⁻², and, next, it was spread over the entire molybdenum bulk by heating to 2000 K. Using the known data on the bulk diffusion of carbon in molybdenum [10, 11], it can be easily shown that, in the range 1500–2000 K, carbon atoms will pass through the whole width $h = 20$ μm of the ribbon rather rapidly. Actually, in the case of unidirectional diffusion, the length of the diffusion front λ is related to the diffusion time by the equation [12]

$$\lambda = \sqrt{2Dt}, \quad (1)$$

where

$$D[\text{cm}^2/\text{s}] = 34.8 \times 10^{-2} \exp[1.78 \times 11600/T] \quad (2)$$

with allowance made for the diffusion rate constants of C atoms in Mo proposed in [10]. Setting $\lambda = h$, we find the rated traveling time of carbon from one side of the ribbon to the other as a function of temperature (see table).

It is evident that the uniform filling of the molybdenum bulk with carbon at $T > 1500$ K will be attained rapidly. If now the temperature is decreased stepwise from 2000 K, and the Auger signal of carbon is recorded directly at the heated sample rather than at room temperature, we will obtain curve I presented in Fig. 1. This figure displays the temperature dependence of the surface carbon concentration normalized to the carbon concentration in the surface carbide with $N_{SC} = 1 \times 10^{15}$ cm⁻², which we take as $\theta = 1$.

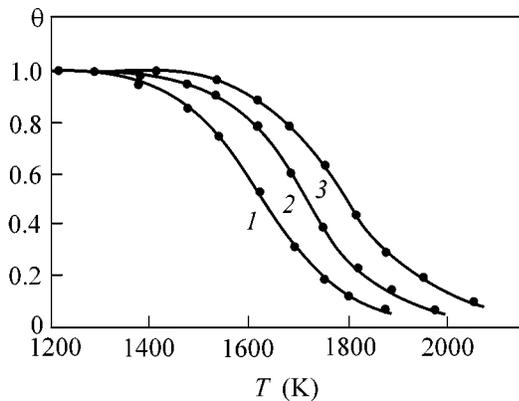


Fig. 1. Equilibrium coverage of carbon on (100)Mo as a function of the substrate temperature. The number of carbon atoms per 1 cm^2 of the interstitial plane equals: (1) 1.0×10^{11} , (2) 1.5×10^{11} , and (3) 3×10^{11} . The carbon concentration in the surface carbide, which equals $N_S = 1 \times 10^{15} \text{ cm}^{-2}$, is taken as $\theta = 1$.

It turned out that the dependence is reproduced well at a decrease or increase in temperature. It is evident that equilibrium conditions in our experiments are attained when the diffusion fluxes between the surface and the bulk of the ribbon become equal and the mass transfer of carbon is terminated. Because carbon is uniformly spread over the entire bulk, its concentration can be estimated in one interstitial layer. Let us assume that $T = 1400 \text{ K}$. Then, the surface carbon concentration on each side of the ribbon will be equal to $N_S = 1 \times 10^{15} \text{ cm}^{-2}$. The amount of carbon that passes into the ribbon bulk will be equal to

$$N_b = N_{\text{dep}} - 2N_S. \quad (3)$$

Hence, the carbon concentration in 1 cm^2 of an interstitial layer will be equal to

$$N_1 = N_S/m = 1 \times 10^{11} \text{ cm}^{-2}, \quad (4)$$

where $m = h/2d \sim 1.4 \times 10^5$ is the number of interstitial planes in the ribbon with the allowance made for the occurrence of two equivalent interstitial planes per one unit cell, and d is the lattice constant in the [100] direction, $d \approx 3.04 \text{ \AA}$ [13]. The huge difference in the carbon concentrations at the surface ($\sim 10^{15} \text{ cm}^{-2}$) and in the near-surface layer ($\sim 10^{11} \text{ cm}^{-2}$) can exist at equilibrium if the precipitation of carbon is a very fast process compared to its dissolution; that is, it should be expected that $E_{1S} \ll E_{S1}$.

If the amount of carbon dissolved in the bulk is increased by its sputtering at 300 K followed by heating

Rated traveling time of carbon through a sample

$T, \text{ K}$	1500	1600	1700	1800	1900	2000
$t, \text{ s}$	56	24	11	6	3	1.8

the ribbon again to 2000 K, we will obtain new equilibrium curves 2 and 3 in Fig. 1. An increase in the bulk concentration of carbon shifts the entire curve toward the region of higher temperatures because of the growth of the precipitation flux. It is important to note here that the direct proportionality of N_S to the concentration of the dissolved carbon is observed in the region of high T for the case of $N_S < 5 \times 10^{14} \text{ cm}^{-2}$.

When the bulk carbon concentrations are far from tolerable ones, which is the case in this work at least at $T > 1500 \text{ K}$, the dissolution v_{S1} and precipitation v_{1S} fluxes with particle exchange can be described by the following equations [4]:

$$v_{S1} = N_S(1 - N_1/N_{1m})C_{S1} \exp[-E_{S1}/kT], \quad (5)$$

$$v_{1S} = N_1C_{1S} \exp[-E_{1S}/kT], \quad (6)$$

where k is the Boltzmann constant, N_{1m} is the carbon concentration in the first subsurface layer corresponding to the limiting solubility, and C_{S1} and C_{1S} are preexponential factors in the equations for the dissolution and precipitation fluxes.

The mass transfer of particles between the surface and the bulk of the metal determines the activation energies of dissolution E_{S1} and precipitation E_{1S} , which can differ in the general case from the activation energy of bulk diffusion $E_0 = 1.78 \text{ eV}$ [14]. If we make a section of curves $\theta = f(T)$ for $N_S = \text{const}$, the following equation can be obtained from Eqs. (5) and (6):

$$\ln N_{1S}(T) = \ln[N_{S1}C_{S1}/C_{1S}] - \Delta E/kT. \quad (7)$$

The experimental data actually fell on a straight line, and ΔE was found from its slope. It turned out that ΔE was constant and equal to $(2.0 \pm 0.1) \text{ eV}$ in the coverage range $0 < \theta < 0.5$. At $\theta > 0.5$, ΔE decreased slightly and reached 1.8 eV at $\theta = 0.95$ (Fig. 2). Given ΔE , the ratio $C_{S1}/C_{1S} = 370 = \text{const}$ was determined from Eq. (7).

Not only the energy difference, but the activation energy of carbon dissolution E_{S1} for the case $N_S < 5 \times 10^{14} \text{ cm}^{-2}$ was determined by setting up special experiments. For this purpose, a carbon dose ($N_{\text{dep}} = 5 \times 10^{14} \text{ cm}^{-2}$) was sputtered at 300 K onto a sample carefully purified from carbon. Then, the ribbon temperature was raised stepwise, and the surface carbon was tracked by Auger electron spectroscopy. It was found that the dissolution of this carbon started only at $T > 1350 \text{ K}$. Using the procedure proposed previously [15], we estimated the activation energy of carbon dissolution $E_{S1} = 3.9 \pm 0.3 \text{ eV}$ from the starting portions of the dependence for 1350 K. This value is in complete correspondence with the temperature of the onset of notable carbon dissolution. A similar significant excess of the activation energy of carbon dissolution ($E_{S1} = 5.0 \text{ eV}$) over the activation energy of bulk diffusion ($E_0 = 2.6 \text{ eV}$) was observed for the (100)W-C system in [5] and then for the (100)Ta-C system in [16]. The effect observed is evidently characteristic for many

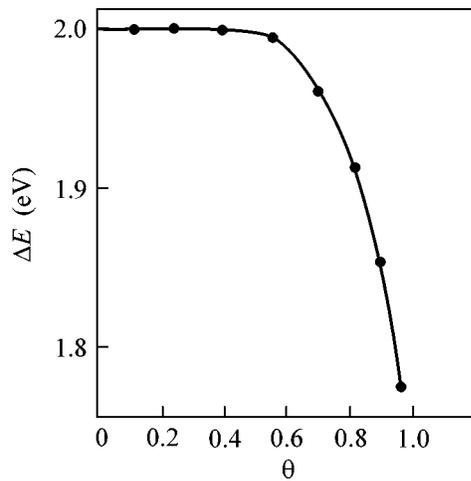


Fig. 2. Dependence $\Delta E = E_{S1} - E_{1S} = f(\theta)$ for the C/(100)Mo adsorption system under conditions of equilibrium between the diffusion fluxes from the surface to the bulk and from the bulk to the surface.

metal–carbon systems. Given E_{S1} , let us also estimate E_{1S} in the (100)Mo–C system: $E_{1S} = E_{S1} - \Delta E \sim 3.9 - 2.0 = 1.9$ eV (for $\theta < 0.5$); that is, E_{1S} is close to the activation energy of carbon bulk diffusion in molybdenum.

It is interesting to note that the dependence $\theta - f(T)$ reaches a plateau $\theta = 1$ at $T < 1400$ K, which corresponds to the formation of the surface MoC carbide. The reached value $\theta = 1$ depends on neither the further decrease in temperature nor the carbon concentration in the molybdenum bulk. It may be suggested that the energy barriers for carbon dissolution and precipitation at the surface become equal ($\Delta E \rightarrow 0$).

We would like to emphasize that obtaining reliable quantitative data on the processes of dissolution and precipitation required solving two important problems. It was necessary to know exactly the concentration of carbon deposited on the metal and vary it. It was also necessary to carry out all the measurements directly at high temperatures. We managed to overcome these difficulties, first, by using an absolutely calibrated flux of carbon atoms and, second, by using thin metal ribbons. Therefore, an equilibrium between the surface carbon and carbon in the bulk metal was attained rapidly. Finally, the unique prism Auger spectrometer allowed Auger spectra to be measured directly for a sample heated up to 2000 K and higher. This allowed errors associated with the rapid precipitation of carbon at the surface upon sample cooling to be avoided. Thus, the physical picture of processes was properly revealed, and some important kinetic parameters describing car-

bon transport between the metal bulk and surface were obtained.

The rate parameters for surface–bulk and bulk–surface diffusion transitions in the solid state were apparently obtained for the first time. The great difference between the activation energies of bulk diffusion (~ 1.8 eV) and dissolution (~ 3.9 eV) may lead to differences of many orders of magnitude in the rates of diffusion processes.

This work was supported by the Ministry of Science of Russian Federation, “Surface Atomic Structures” program, project no. 4.6.99.

REFERENCES

1. R. M. Barrer, *Diffusion in and through Solids* (Cambridge Univ. Press, Cambridge, 1941).
2. P. G. Shewmon, *Diffusion in Solids* (McGraw-Hill, New York, 1963; Metallurgiya, Moscow, 1966).
3. M. P. Seach, in *Practical Surface Analysis by Auger and X-ray Photoelectron Spectroscopy* (Wiley, London, 1983), p. 273.
4. A. Ya. Tontegode, *Zh. Tekh. Fiz.* **43**, 1042 (1973) [*Sov. Phys. Tech. Phys.* **18**, 657 (1973)].
5. K. J. Rawlings, S. D. Foulis, and B. J. Hopkins, *Surf. Sci.* **109**, 513 (1981).
6. A. Ya. Tontegode, M. M. Usufov, E. V. Rut'kov, and N. R. Gall', *Zh. Tekh. Fiz.* **62** (10), 148 (1992) [*Sov. Phys. Tech. Phys.* **37**, 1038 (1992)].
7. E. V. Rut'kov, A. Ya. Tontegode, M. M. Usufov, and N. R. Gall, *Appl. Surf. Sci.* **78**, 179 (1994).
8. V. N. Ageev, E. V. Rut'kov, A. Ya. Tontegode, and N. A. Kholin, *Fiz. Tverd. Tela (Leningrad)* **23**, 2248 (1981) [*Sov. Phys. Solid State* **23**, 1315 (1981)].
9. N. R. Gall, E. V. Rut'kov, A. Ya. Tontegode, *et al.*, *Chem. Vap. Deposition* **5**, 72 (1997).
10. E. Fromm and E. Gebhardt, *Gase und Kohlestoff in Metallen* (Springer-Verlag, Berlin, 1976).
11. V. G. Samsonov and I. M. Vinit'skii, *Handbook of Refractory Compounds* (Metallurgiya, Moscow, 1976; Plenum, New York, 1980).
12. G. Schulze, *Metallphysik* (Akademie-Verlag, Berlin, 1967; Mir, Moscow, 1971).
13. C. Kittel, *Introduction to Solid State Physics* (Wiley, New York, 1975; Nauka, Moscow, 1978).
14. P. S. Rudman, *Trans. Metall. Soc. AIME* **239**, 1949 (1967).
15. A. Ya. Tontegode and F. K. Yusifov, *Poverkhnost*, No. 4, 20 (1994).
16. N. R. Gall, E. V. Rut'kov, and A. Ya. Tontegode, *Surf. Sci.* **472**, 187 (2001).

Translated by A. Bagatur'yants

Magnetic Moment Relaxation of a Shallow Acceptor Center in Heavily Doped Silicon

T. N. Mamedov¹, D. G. Andrianov², D. Herlach³, V. N. Gorelkin⁴, K. I. Gritsaï¹, V. N. Duginov¹, O. Kormann⁵, J. Major⁵, A. V. Stoïkov¹, and U. Zimmermann³

¹ Joint Institute for Nuclear Research, Dubna, Moscow region, 141980 Russia

² State Research and Design Institute of Rare Metals “Giredmet,” Moscow, 109017 Russia

³ Paul Scherrer Institut, CH-5232 Villigen PSI, Switzerland

⁴ Moscow Institute of Physics and Technology, Dolgoprudnyĭ, Moscow region, 141700 Russia

⁵ Max-Planck-Institut für Metallforschung, D-70569 Stuttgart, Germany

e-mail: tmamedov@nu.jinr.ru

Received May 30, 2001

Results of studying the temperature dependence of the residual polarization of negative muons in crystalline silicon with germanium ($9 \times 10^{19} \text{ cm}^{-3}$) and boron (4.1×10^{18} , 1.34×10^{19} , and $4.9 \times 10^{19} \text{ cm}^{-3}$) impurities are presented. It is found that, similarly to *n*- and *p*-type silicon samples with impurity concentrations up to $\sim 10^{17} \text{ cm}^{-3}$, the relaxation rate ν of the magnetic moment of a μAl acceptor in silicon with a high impurity concentration of germanium ($9 \times 10^{19} \text{ cm}^{-3}$) depends on temperature as $\nu \sim T^q$, $q \approx 3$ at $T = (5\text{--}30) \text{ K}$. An increase in the absolute value of the relaxation rate and a weakening of its temperature dependence are observed in samples of degenerate silicon in the given temperature range. Based on the experimental data obtained, the conclusion is made that the spin-exchange scattering of free charge carriers makes a significant contribution to the magnetic moment relaxation of a shallow acceptor center in degenerate silicon at $T \leq 30 \text{ K}$. Estimates are obtained for the effective cross section of the spin-exchange scattering of holes (σ_h) and electrons (σ_e) from an Al acceptor center in Si: $\sigma_h \sim 10^{-13} \text{ cm}^2$ and $\sigma_e \sim 8 \times 10^{-15} \text{ cm}^2$ at the acceptor (donor) impurity concentration $n_a(n_d) \sim 4 \times 10^{18} \text{ cm}^{-3}$. © 2001 MAIK “Nauka/Interperiodica”.

PACS numbers: 71.55.Cn; 76.75.+i

In contrast to shallow donors, shallow acceptor centers in semiconductors with a diamond-type crystal lattice have not been adequately investigated [1]. The use of conventional methods (EPR, ENDOR) for studying shallow acceptors is essentially restricted because of the high spin–lattice relaxation rate of the magnetic moment of the acceptor and the broadening of the resonance line due to random crystal strains [2].

The use of negative muons [3–6] significantly extends the possibilities of experimental studies of shallow acceptor centers in silicon.

When a negative muon is implanted in silicon, the resulting muonic atom models the behavior of an acceptor aluminum impurity (μAl). The behavior of the residual polarization of a negative muon that occupies the $1S$ level of the muonic atom is determined by the hyperfine interaction (A_{hf} is the hyperfine coupling constant) in the muonic atom and by interactions of the electron shell of the muonic atom with the medium. In an external magnetic field transverse to the muon spin, the magnetic moment relaxation rate of the μAl acceptor is high ($\nu \gg A_{hf}$). Therefore, the polarization vector precession frequency is close to the precession frequency of the free muon spin, and the muon spin exhib-

its relaxation (with the relaxation rate λ) and a paramagnetic shift of the precession frequency ($\Delta\omega$). The relation between the parameters of muon spin polarization observed experimentally (λ , $\Delta\omega$) and the parameters (A_{hf} , ν) characterizing interactions of the paramagnetic acceptor center within the model of isotropic hyperfine interaction is determined by the equations [7]

$$\lambda = \frac{j(j+1)}{3} \left(\frac{A_{hf}^2}{\nu} + \frac{A_{hf}^2 \nu}{\nu^2 + \omega_e^2} \right), \quad (1)$$

$$\frac{\Delta\omega}{\omega_0} = -\frac{g\mu_B}{2\mu_B^\mu} \frac{j(j+1)\hbar A_{hf}}{3k_B T}, \quad (2)$$

where $\Delta\omega = \omega(T) - \omega_0$ (ω_0 is the angular muon spin precession frequency in the diamagnetic state of the muonic atom); $\hbar = h/2\pi$ (h is Planck's constant); k_B is the Boltzmann constant; μ_B and μ_B^μ are the electronic and muonic Bohr magnetons, respectively; g is the g factor of the acceptor center; $\omega_e = g\mu_B B/\hbar$ is the angular magnetic moment precession frequency of the electron shell of the acceptor center in the external mag-

netic field B ; and T is temperature in kelvin. For a shallow acceptor center in silicon, $j = 3/2$ [8] and $g = -1.07$ [9].

We found previously [4–6] that the magnetic moment relaxation of a μAl acceptor in nondegenerate silicon is due to spin–lattice interaction, and the relaxation rate depends on temperature as $\nu \sim T^q$, $q \approx 3$. An increase in the absolute value of the relaxation rate and a weakening of its temperature dependence were observed [6] in n - and p -type silicon at impurity concentrations of higher than $\sim 10^{18} \text{ cm}^{-3}$. It was suggested that the increase in the magnetic moment relaxation rate of an acceptor center in a degenerate semiconductor at low temperatures is due to the spin-exchange scattering of free charge carriers. The goal of this work was to elucidate the mechanisms of magnetic moment relaxation for an acceptor in a heavily doped semiconductor.

The temperature dependence of parameters of the residual polarization of negative muons was investigated in samples of silicon with germanium ($9 \times 10^{19} \text{ cm}^{-3}$, $\rho = 60 \text{ } \Omega \text{ cm}$) and boron (4.1×10^{18} , 1.34×10^{19} , and $4.9 \times 10^{19} \text{ cm}^{-3}$) impurities. Germanium, as well as silicon, belongs to the IV group of elements of the Periodic Table; the outer electron shell of these elements contains four valence electrons. In contrast to group III and V element impurities, a germanium impurity does not increase the concentration of free charge carriers in silicon as compared to “pure” samples, in which the concentration of free charge carriers at low temperatures is insignificant. Therefore, the contributions to the magnetic moment relaxation of an acceptor due to the interactions with free charge carriers and with the crystal lattice can be separated by comparing the results of investigations for silicon with a germanium impurity and silicon with concentrations of group III and V atoms of order n_c (n_c is the critical concentration corresponding to the dielectric–metal transition). When the concentration of group III and V atoms is equal to or higher than n_c , the concentration of free charge carriers in the semiconductor is approximately equal to the concentration of impurity atoms. Depending on the impurity atom, n_c for silicon varies from $\sim 3 \times 10^{18}$ to $\sim 7 \times 10^{18} \text{ cm}^{-3}$ [10].

The measurements were made on a Stuttgart LFQ spectrometer [11] and GPD [12] spectrometers with muon beams μE4 and μE1 of the proton accelerator at the Paul Scherrer Institute, PSI, Switzerland. The samples to be studied were cut from silicon single crystals in the shape of disks ($\sim 30 \text{ mm}$ in diameter and up to 15 mm in height) and were set so that the disk axis coincided with the muon beam axis. The magnetic field at the sample transverse to the muon spin was created by Helmholtz coils and equaled 0.1 T . The long-term stability of the magnetic field was no worse than 10^{-4} . The temperature of samples was stabilized with an accuracy of 0.1 K over the range $4.2\text{--}300 \text{ K}$.

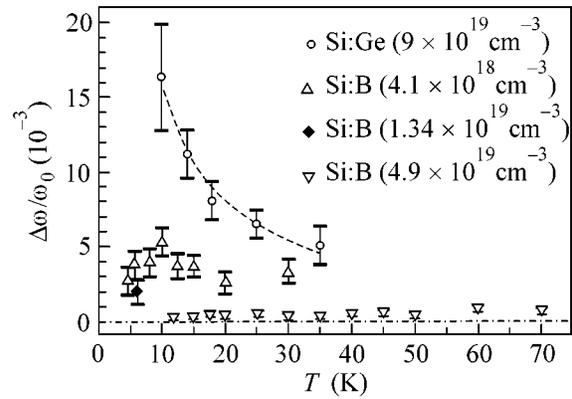


Fig. 1. Temperature dependence of the precession frequency shift of the muon spin for silicon samples with germanium ($9 \times 10^{19} \text{ cm}^{-3}$) and boron (4.1×10^{18} , 1.34×10^{19} , and $4.9 \times 10^{19} \text{ cm}^{-3}$) impurities; the dashed line corresponds to the dependence $\Delta\omega/\omega_0 = C/T$.

The procedure of recovering the polarization function of a negative muon from the experimental $\mu\text{-SR}$ spectra was described in detail in [4]. Recall that (see [4]) the polarization function of the muon spin in the paramagnetic state of the acceptor center takes the form (we consider the case when the sample is in an external magnetic field transverse to the muon spin)

$$p(t) = p_0 e^{-\lambda t} \cos(\omega t + \phi), \quad (3)$$

where p_0 is polarization at the zero instant of time (the residual polarization of the negative muon at the $1S$ level of the muonic atom), and λ , ω , and ϕ are, respectively, the relaxation rate, frequency, and starting precession phase of the muon spin.

Experimental data on the temperature dependence of the precession frequency shift of the muon spin for the silicon samples studied in this work are shown in Fig. 1. (The precession frequency of the muon spin at room temperature was taken as ω_0 .) It is evident in the figure that the frequency shift increases with decreasing temperature for the sample with a germanium impurity. The hyperfine coupling constant in the acceptor center formed by the muonic atom was determined by fitting the data obtained for the sample with a germanium impurity using Eq. (2). It was found that $A_{hf}/2\pi$ equals $23.6 \pm 1.5 \text{ MHz}$, which is in agreement with the average value ($26.5 \pm 2.3 \text{ MHz}$) of the hyperfine coupling constant obtained for more than ten samples of n - and p -type silicon with impurity concentrations up to $2 \times 10^{17} \text{ cm}^{-3}$ studied previously [4–6]. (In averaging the experimental data [4–6], the error in the $A_{hf}/2\pi$ value was calculated as the weighted average standard deviation of the obtained values relative to their average value.)

On the average, the precession frequency shift of the muon spin at $T \approx 30 \text{ K}$ for a silicon sample with a boron impurity ($4.1 \times 10^{18} \text{ cm}^{-3}$) is $\sim 3 \times 10^{-3}$. Within the limits

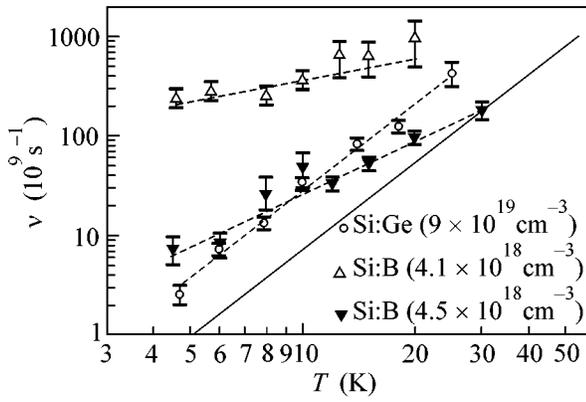


Fig. 2. Temperature dependence of the magnetic moment relaxation rate ν of an acceptor center for samples of silicon with germanium ($9 \times 10^{19} \text{ cm}^{-3}$), boron ($4.1 \times 10^{18} \text{ cm}^{-3}$), and phosphorus ($4.5 \times 10^{18} \text{ cm}^{-3}$) impurities; straight lines describe relationships of the form $\nu = C\tilde{T}^q$; the solid straight line is the averaged dependence $\nu(T)$ for samples of n - and p -type silicon with impurity concentrations up to $2 \times 10^{17} \text{ cm}^{-3}$.

of experimental errors, it hardly depends on the temperature. The precession frequency shift at a boron concentration of more than $\sim 10^{19} \text{ cm}^{-3}$ is close to zero within the limits of experimental errors.

The decrease in the absolute value of the precession frequency shift of the muon spin in silicon with a boron impurity observed in this work correlates with the behavior [13, 14] of the static paramagnetic susceptibility of acceptor (donor) impurities in silicon when their concentration $n_a(n_d)$ is on the order of higher than n_c .

Muon spin relaxation (λ) was observed only for samples of silicon with a germanium impurity ($9 \times 10^{19} \text{ cm}^{-3}$) and boron ($4.1 \times 10^{18} \text{ cm}^{-3}$). The experimental data $\lambda(T)$ were approximated by Eq. (1) under the suggestion that ν depends on temperature as $\nu = C\tilde{T}^q$, $\tilde{T} = T/1 \text{ K}$. The values of q were found to be equal to 2.87 ± 0.16 and 0.7 ± 0.2 for samples with germanium and boron impurities, respectively. The index of power dependence q for the sample of silicon with a germanium impurity, in spite of the very high concentration of Ge, coincides within the limits of errors with values $q \approx 3$ obtained previously for more than ten samples of silicon with various (P, B, As, Al, and Ga) impurities whose concentrations were no more than $2 \times 10^{17} \text{ cm}^{-3}$. In the case of silicon with a boron impurity ($4.1 \times 10^{18} \text{ cm}^{-3}$), the temperature dependence of ν is significantly weaker than a T^3 dependence. A similar weakening of the temperature dependence of the magnetic moment relaxation rate was observed previously in [5, 6] for an acceptor impurity in n - and p -type silicon with impurity concentrations of $\sim 10^{18} \text{ cm}^{-3}$.

The temperature dependence of the magnetic moment relaxation rate ν of an acceptor center is illustrated in Fig. 2 for samples of silicon with germanium ($9 \times 10^{19} \text{ cm}^{-3}$) and boron ($4.1 \times 10^{18} \text{ cm}^{-3}$) impurities. The value of ν was calculated from the experimental data on the muon spin relaxation rate by Eq. (1) at $A_{hf}/2\pi = 26.5 \text{ MHz}$. The values of ν calculated for silicon with a phosphorus impurity ($4.5 \times 10^{18} \text{ cm}^{-3}$) from the experimental data reported in [5] are also presented in Fig. 2. The straight lines in the figure correspond to relationships of the form $\nu = C\tilde{T}^q$, and the solid line corresponds to the averaged (by the data from [4–6]) temperature dependence $\nu(T)$ for silicon samples with n - and p -type impurity concentrations of less than $2 \times 10^{17} \text{ cm}^{-3}$: $C \approx 9 \times 10^6 \text{ s}^{-1}$, $q = 2.86 \pm 0.26$. The weighted average standard deviation of the results of separate measurements of q relative to their average value is given as the error of parameter q .

The analogous (close to T^3) temperature dependence of ν in silicon samples with $n_a(n_d) \leq 10^{17} \text{ cm}^{-3}$ and in silicon with a high ($\sim 10^{20} \text{ cm}^{-3}$) concentration of an isoelectronic impurity and, simultaneously, a significant deviation of ν from a T^3 dependence at $n_a(n_d) \sim n_c$ indicate that the magnetic moment relaxation of an acceptor center in the first case is due to spin–lattice interaction, whereas the spin-exchange scattering of charge carriers by an acceptor makes a significant contribution to relaxation in the second case. The rate of the spin-exchange scattering of free charge carriers from a paramagnetic center in a degenerate semiconductor, as well as in a metal, is proportional to the first degree of temperature

$$\nu = \beta \hbar^{-1} n^{2/3} \sigma k_B T, \quad (4)$$

where σ is the effective cross section of spin-exchange scattering, n is the impurity concentration, and β is a numerical coefficient (for silicon, $\beta \approx 1$ [15]).

Correspondingly, the temperature dependence of the magnetic moment relaxation rate of an acceptor center for degenerate silicon samples can be represented in the form

$$\nu(T) = C_1 \tilde{T} + C_2 \tilde{T}^3, \quad \tilde{T} = T/1 \text{ K}. \quad (5)$$

The values of parameters C_1 and C_2 obtained when the experimental data for degenerate silicon samples shown in Fig. 2 were fitted to Eq. (5) are given in the table. The effective cross sections of the spin-exchange scattering of holes (σ_h) and electrons (σ_e) from an Al acceptor center in Si calculated by Eq. (4) using the values of C_1 are given in the same table.

It is evident from the table that the effective cross section for the spin-exchange scattering of holes from an acceptor in a boron-doped silicon sample ([B] $\sim 4.1 \times 10^{18} \text{ cm}^{-3}$) is $\sigma_h \sim 10^{-13} \text{ cm}^2$ and exceeds by more than an order of magnitude the cross section for electron scattering in a sample with approximately the same

Values of C_1 and C_2 parameters and estimates of the effective cross section of the spin-exchange scattering of holes σ_h and electrons σ_e by a paramagnetic Al acceptor center in Si

Sample	$C_1, 10^9 \text{ s}^{-1}$	$C_2, 10^6 \text{ s}^{-1}$	$\sigma_h, \text{ cm}^2$	$\sigma_e, \text{ cm}^2$
Si : B ($4.1 \times 10^{18} \text{ cm}^{-3}$)	42 ± 4	9*	$\sim 10^{-13}$	–
Si : P ($4.5 \times 10^{18} \text{ cm}^{-3}$)	2.4 ± 0.4	4.9 ± 1.2	–	$\sim 8 \times 10^{-15}$

Note: When processing, the parameter C_2 was fixed according to its mean value in n - and p -type silicon samples with impurity concentrations less than $2 \times 10^{17} \text{ cm}^{-3}$.

concentration of phosphorus impurity. We know no experimental work in which the effective cross sections of the spin-exchange scattering of charge carriers by a shallow impurity center in silicon would be determined. For comparison, we cite the value of σ obtained in [16] for the scattering of conduction electrons by a paramagnetic Fe center in Si, which is a deep donor: $\sigma_e(\text{Fe}) \sim 10^{-14} \text{ cm}^2$. The effective $\sigma_h(\text{Al})$ and $\sigma_e(\text{Al})$ cross sections obtained in this work are not in contradiction with the data reported in [16] for $\sigma_e(\text{Fe})$, because the effective cross section of the spin-exchange scattering of charge carriers can be larger in the case of a shallow acceptor center due to a larger size of the hole localization region.

Thus, in this work, it is shown experimentally that the magnetic moment relaxation of a shallow acceptor center in silicon at $T \lesssim 30 \text{ K}$ in the case of a low concentration of free charge carriers (nondegenerate semiconductor, isoelectronic impurity) is due to spin–lattice interaction, and the relaxation rate depends on temperature as $\nu \sim T^q$, $q \approx 3$. The magnetic moment relaxation of an acceptor in degenerate silicon at $T \lesssim 30 \text{ K}$ is virtually determined by the spin-exchange scattering of free charge carriers. Estimates are obtained for the effective cross section of the spin-exchange scattering of holes (σ_h) and electrons (σ_e) by an Al acceptor center in Si.

The authors are grateful to the administration of the Paul Scherrer Institute (Switzerland) for the possibility of performing the measurements reported in this work.

REFERENCES

1. G. D. Watkins, *Fiz. Tverd. Tela (St. Petersburg)* **41**, 826 (1999) [*Phys. Solid State* **41**, 746 (1999)].

2. G. W. Ludwig and H. H. Woodbury, *Solid State Phys.* **13**, 223 (1962).
 3. V. N. Gorelkin and V. P. Smilga, *Zh. Éksp. Teor. Fiz.* **66**, 1201 (1974) [*Sov. Phys. JETP* **39**, 586 (1974)].
 4. T. N. Mamedov, I. L. Chaplygin, V. N. Duginov, *et al.*, *J. Phys.: Condens. Matter* **11**, 2849 (1999).
 5. T. N. Mamedov, K. I. Gritsaj, A. V. Stoykov, *et al.*, *Physica B (Amsterdam)* **289–290**, 574 (2000).
 6. T. N. Mamedov, D. G. Andrianov, D. Gerlakh, *et al.*, Preprint R14-2000-151, OIYaI (Joint Inst. for Nuclear Research, Dubna, 2000).
 7. V. N. Gorelkin, T. N. Mamedov, and A. S. Baturin, *Physica B (Amsterdam)* **289–290**, 585 (2000).
 8. G. L. Bir and G. E. Pikus, *Symmetry and Strain-Induced Effects in Semiconductors* (Nauka, Moscow, 1972; Wiley, New York, 1975).
 9. H. Neubrand, *Phys. Status Solidi B* **86**, 269 (1978).
 10. P. Ph. Edwards and M. J. Sienko, *Phys. Rev. B* **17**, 2575 (1978).
 11. R. Scheuermann, J. Schmidl, A. Seeger, *et al.*, *Hyperfine Interact.* **106**, 295 (1997).
 12. R. Abela, C. Baines, X. Donath, *et al.*, *Hyperfine Interact.* **87**, 1105 (1994).
 13. M. P. Sarachik, D. R. He, W. Li, and M. Levy, *Phys. Rev. B* **31**, 1469 (1985).
 14. A. Roy, M. Turner, and M. P. Sarachik, *Phys. Rev. B* **37**, 5522 (1988); A. Roy and M. P. Sarachik, *Phys. Rev. B* **37**, 5531 (1988).
 15. Yu. G. Semenov, *Fiz. Tverd. Tela (Leningrad)* **22**, 3190 (1980) [*Sov. Phys. Solid State* **22**, 1868 (1980)].
 16. M. F. Deĭgen, V. Ya. Bratus', B. E. Vugmeister, *et al.*, *Zh. Éksp. Teor. Fiz.* **69**, 2110 (1975) [*Sov. Phys. JETP* **42**, 1073 (1975)].

Translated by A. Bagatur'yants

Apparent Anomalous Behavior of the Power Spectrum of $1/f$ Noise and Its Explanation

A. G. Budarin

Altair State Scientific and Production Association, Moscow, 111024 Russia

Received May 14, 2001

The explanation is given for the apparent anomalous behavior of the power spectrum of $1/f$ noise as if it corresponded to the total infinite power of noise sources. Physical mechanisms eliminating the apparent anomalies are described. With these mechanisms, the finite and integrable $1/f$ -noise power spectra fitting the known physical concepts of noise processes are obtained in the low- and high-frequency limits. © 2001 MAIK "Nauka/Interperiodica".

PACS numbers: 05.40.Ca

A low-frequency random signal with power spectrum $F(\omega) \sim \omega^{-\alpha}$ ($\alpha \sim 1$, $\omega = 2\pi f$), called $1/f$ noise, is observed in a wide diversity of transport phenomena, not necessarily physical, occurring in many-particle systems. A remarkable property of the $1/f$ -noise spectrum of the form $\omega^{-\alpha}$ is that it displays an anomalous behavior; namely, it tends to infinity in the vicinity of zero frequency, and the corresponding integral diverges at one of the edges or, for $\alpha = 1$, at both edges of the frequency scale. This property of the $1/f$ noise implies that it is nonstationary and that the total power of its sources is infinite, which is in contradiction with the properties of real noise signals that are generated and measured under stationary conditions. To account qualitatively for this paradoxical situation, it was repeatedly assumed that the $1/f$ spectrum is so modified at high and low frequencies that it becomes finite and integrable [1].

It is the purpose of this work to prove this assumption and establish the mechanism that is responsible for the deviation of the $1/f$ spectrum from the standard form $\omega^{-\alpha}$ and to estimate these distinctions on the basis of the mathematical theory of $1/f$ noise developed in [2, 3]. This theory states that the $1/f$ noise is generated due to the diffusional relaxation of the fluctuations of a physical parameter in a shock-type system, in which this parameter obeys a power-law density distribution [2, 3], while the fluctuations form a random Poisson sequence. Accordingly, the $1/f$ noise is a random signal formed as a result of measuring the diffusion flows $j_i(x - x_i, t - t_i)$, where x_i and t_i are, respectively, the coordinate and the moment at which the i th fluctuation occurs; $r(x - x_i, t - t_i)$ is the current value of the parameter at $t > t_i$; and $i \in (0, K)$, where K is the number of fluctuations in the Poisson sequence. The spectrum of $1/f$ noise has the required $\omega^{-\alpha}$ form if the relationships $\beta = 3/2 - \alpha$, $\beta \in (0, 1)$, $\alpha \in (0.5, 1.5)$ hold and the in-

qualities $h = (\lambda_1/\omega)^{1/2} \ll 1$ and $\delta^2 = (v^2/4\kappa\omega)^2 \ll 1$ are fulfilled, where λ_1 is the lowest eigenvalue of the diffusion equation $r_t = vr_x + \kappa r_{xx}$ with homogeneous boundary conditions, which describes the fluctuation relaxation process; v is the drift velocity of particles; and κ is the diffusion coefficient.

Let us first consider the low-frequency region. It is shown in [2, 3] that the frequency dependence $F(\omega)$ is modified at $v \neq 0$ because of replacing ω by $\omega' = \omega(1 + \delta^2)^{1/2}$, as a result of which the $1/f$ spectrum becomes different from $\omega^{-\alpha}$ at low frequencies. However, the mechanism of forming fluctuations in the general model [2–4] necessarily presumes a nonzero velocity of the jump in the physical parameter and, hence, the nonzero velocity v of the convective transport of particles involved in the fluctuation. For this reason, the $F(0)$ value in this theory is finite; i.e., the low-frequency anomaly of the $1/f$ noise vanishes.

The experimentally observed fact that $F(\omega)$ is close to $\omega^{-\alpha}$ is qualitatively explained in this theory as follows. Let us write δ^2 as $(v/v_d)^4$, where $v_d = 2(\kappa\omega)^{1/2}$ characterizes the velocity of the diffusion front moving from the fluctuation after the singularity appears in its density. Due to this singularity and to the rapidity of its formation because of the fluctuation front breaking, v_d is larger than v and, hence, the condition $\delta^2 \ll 1$ for the validity of the $F(\omega) \sim \omega^{-\alpha}$ dependence is fulfilled. For instance, it follows from [4, 5] that v in the electric conductors with high charge-carrier density is on the order of the drift velocity V of conduction electrons, so that for typical values $\kappa \sim 10^4$ cm²/s, $V \sim 10^{-2}$ cm/s, and $f = (\omega/2\pi) \sim 1$ s⁻¹ the value of δ^2 is $\sim 10^{-19}$, i.e., negligibly small. The values $\delta \sim 1$ are achieved only at $\omega \sim \Omega = v^2/4\kappa \sim 2.5 \times 10^{-9}$ Hz. For this reason, to distinguish between the $\omega^{-\alpha}$ spectrum and the $1/f$ spectrum pre-

dicted by this theory, an experimentally inaccessible time of $\Omega^{-1} \sim 10^2$ years is required.

Let us now consider the high-frequency domain. It is assumed that the $1/f$ spectrum decreases in this domain faster than $\omega^{-\alpha}$ because of a smoothing-out effect of the time correlation between different points of the particle trajectory in the diffusion flow. This effect was not taken into account in the basic model. The inclusion of this effect will lead to the suppression of the most rapid fluctuations in the diffusion flow and, as a consequence, to the narrowing of the $1/f$ spectrum up to its integrable form.

Let us prove this supposition by calculating the asymptotic form of $F(\omega)$ in the high-frequency limit with account taken of the time correlation. For this purpose, we generalize the model [2, 3] and assume that κ is a function of time. The resulting diffusion equation allows for the correlative distinctions in the fluctuation decay rates at different instants of time. To establish the form of $\kappa(t)$, let us introduce the coordinate $\xi = x + vt$ and transform the diffusion equation to the canonical form $w_t = \kappa(t)w_{\xi\xi}$, where $w(\xi, t) = r(\xi - vt, t)$. We next use two methods, statistical and phenomenological, to calculate the variance $D_\xi = d_\xi^2$ for the random displacement $d_\xi = \xi(t) - \xi(0)$ of a particle in the flow and obtain the desired function $\kappa(t)$ by equating the results of both calculations.

In the first method, one should use the relationship $d_\xi = \int_0^t u(t') dt'$ between d_ξ and the random velocity component $u(t)$ of a particle ($\langle u \rangle = 0$) to find

$$D_\xi = \int_0^t \langle u(t')u(t'') \rangle dt' dt'' = 2 \int_0^t (t-s)R(s)ds, \quad (1)$$

where $R(|t' - t''|) = \langle u(t')u(t'') \rangle$ is the particle velocity correlation function.

In the second method, one starts with the definition of the $\langle \xi^2(t) \rangle$ quantity

$$\langle \xi^2(t) \rangle = \int_{-\infty}^{\infty} \xi^2(t)w(\xi, t)d\xi / \int_{-\infty}^{\infty} w(\xi, t)d\xi. \quad (2)$$

Let us neglect the boundary effects (small at $h \ll 1$) in the object under observation and substitute the solution of the diffusion equation with $\kappa = \kappa(t)$ into Eq. (2):

$$w(\xi, t) = \int_{-\infty}^{\infty} w(\xi, 0)G(\xi, t; \xi', 0)d\xi', \quad (3)$$

where the Green's function of the diffusion equation is

$$G(\xi, t; \xi', 0) = (\pi g(t))^{-1/2} \exp[-(\xi - \xi')^2/g(t)],$$

$$g(t) = 4 \int_0^t \kappa(t')dt'. \quad (4)$$

This gives the following phenomenological expression for $D_\xi(t)$:

$$D_\xi(t) = 2 \int_0^t \kappa(t')dt'. \quad (5)$$

By setting Eq. (5) equal to its estimate (1), one obtains the desired quantity

$$\kappa(t) = \int_0^t R(s)ds. \quad (6)$$

For the correlation function

$$R(s) = \kappa_0 \tau^{-1} \exp(-s/\tau)$$

[6], one obtains

$$\kappa(t) = \kappa_0 [1 - \exp(-t/\tau)],$$

$$D_\xi = 2\tau\kappa_0 [t/\tau - (1 - \exp(-t/\tau))]. \quad (7)$$

At $t/\tau \rightarrow 0$, $\kappa(t) \rightarrow \kappa_0 t/\tau$ and $D_\xi(t) \rightarrow \kappa_0 t^2/\tau$ and at $t/\tau \rightarrow \infty$, $\kappa(t) \rightarrow \kappa_0$ and $D_\xi(t) \rightarrow 2\kappa_0 t$. One can show [6] that this limiting behavior of the functions $\kappa(t)$ and $D_\xi(t)$ does not depend on the form of $R(s)$ and is quite general.

By reproducing the calculations performed in [2, 3] and replacing $\kappa = \text{const}$ by $\kappa(t)$ from Eq. (6), one can recast $F(\omega)$ as a series

$$G^{(0)}(\omega) = \sum_{n=1}^{\infty} \langle b_n^2 \rangle k_n^2 |g_{\omega n 0}|^2 \quad (8)$$

with asymptotic form $\langle b_n^2 \rangle \sim (A^2/L)k_n^{-2(1-\beta)}$ ($\beta = 3/2 - \alpha$) differing from $G^{(0)}(\omega)$ in [2, 3] only by the frequency multiplier $|g_{\omega n 0}|^2$, for which one now has

$$g_{\omega n 0} = \int_0^{\infty} \kappa(t) \exp(-k_n^2 D(t)/2 + i\omega t) dt. \quad (9)$$

Thus, taking account of time correlations in this problem amounts to the estimation of $g_{\omega n 0}$. It is hard to apply directly the formulas of asymptotic analysis to this quantity because Eq. (9) involves two large parameters ω and n^2 that enter in $k_n^2 = (\pi n/L)^2$. Nevertheless, one can show [7] that, at large enough ω , the estimate of a given accuracy is expressed by the integral over small vicinity $[0, \Delta'(\omega, n^2)]$, which decreases indefinitely as ω increases, with $\Delta'(\omega, n_2^2) < \Delta'(\omega, n_1^2)$ at $n_1 < n_2$, as follows from Eq. (9). This allows the ω_0 and

n^2 -independent $\Delta = \Delta(\omega_0)$ quantities to be chosen in such a way that the integral estimate of $g_{\omega n}$ for $\omega \geq \omega_0$

$$g_{\omega n} = \int_0^{\Delta} \kappa_0(t/\tau) \exp(-\kappa_0 k_n^2 t^2 / 2\tau + i\omega t) dt, \quad (10)$$

obtained by replacing the upper limit ∞ in Eq. (9) by $\Delta = \Delta(\omega_0)$ and $\kappa(t)$ and $D_\xi(t)$ by their limiting values, will have the predetermined accuracy.

The asymptotic estimate of integral (10) can be done by its transformation to the special functions and the standard power series expansions. Let us introduce the parameters $\eta = (\omega\tau/2)^{1/2}$, $x = nh$, $\lambda = (\eta x)^2$, $\psi = \eta/x$, $\gamma = \Delta/\tau$, $\zeta = \gamma\lambda^{1/2} - i\psi$, and $Q = \pi^{1/2}\kappa_0\tau/2\lambda$ and substitute $\kappa_0 k_n^2 = (nh)^2\omega$ and $t = \tau\lambda^{-1/2}(y - i\psi)$ in the exponent, whereupon we transform Eq. (10) to

$$g_{\omega n} = Q \exp(-\psi^2) [\mathbf{i} \operatorname{erfc}(-i\psi) - \pi^{1/2} \exp(-\zeta^2) - \operatorname{erfc}(\zeta)]. \quad (11)$$

We now take into account that $|\zeta| = |\zeta(x)| \geq (\omega\Delta)^{1/2}$ and $\arg \zeta \in (0, -\pi/2)$ for any x and assume that $(\omega\Delta)^{1/2} \gg 1$. The replacement of $\operatorname{erfc}(\zeta)$ in Eq. (11) by its asymptotic estimate $\pi^{-1/2}\zeta^{-1}\operatorname{erfc}(-\zeta^2)$ and $\mathbf{i}\operatorname{erfc}(-i\psi)$ by its series expansion [Eq. (7.2.4) in [8, ch. 7]] gives the following estimate at $(\omega\Delta)^{1/2} \gg 1$:

$$g_{\omega n} \sim Q \exp(-\psi^2) \left[(1/2) \sum_{k=0}^{\infty} a_k (i\psi)^k - W \exp(-\zeta^2) \right], \quad (12)$$

where $a_k = 2^k [k! \Gamma((3-k)/2)]^{-1}$ at $k=1$ and $k=2m$ ($m \geq 0$) and $a_k = 0$ at $k=2m+1$ ($m > 0$) and $W = W(x, \gamma) = |W| \exp(i \arg W)$, where $|W| = \pi^{-1/2} \gamma x^2 (\gamma x^2 + 1)^{-1/2}$ and $\arg W = \arctan(\gamma x^2)^{-1}$.

The resulting general expression for $|g_{\omega n}|^2$ at $(\omega\Delta)^{1/2} \gg 1$ is

$$\begin{aligned} |g_{\omega n}|^2 &= I_0 = I_1 - I_2 + I_3, \\ I_1 &= (Q^2/4) \exp(-2(\eta/x)^2) \\ &\times \sum_{k=0}^{\infty} \sum_{j=0}^k a_k a_j (2 - \delta_{kj}) (\eta/x)^{k+j} \cos((k-j)\pi/2), \\ I_2 &= Q^2 |W| \exp[-(\gamma\eta x)^2 - (\eta/x)^2] \\ &\times \sum_{k=0}^{\infty} a_k (\eta/x)^k \cos(\omega\Delta + \arg W + k\pi/2), \\ I_3 &= Q^2 |W|^2 \exp(-2(\gamma\eta x)^2). \end{aligned} \quad (13)$$

Inserting $|g_{\omega n}|^2$ into Eq. (8) and passing from summation to integration in the same manner as in [2, 3], one

arrives at the correct (because of the condition $h \ll 1$) estimate $G_0(\omega)$ for the spectrum $G^{(0)}(\omega)$:

$$\begin{aligned} G_0(\omega) &= G_1(\omega) - G_2(\omega) + G_3(\omega), \\ G_i(\omega) &= (A^2/\pi L) (\omega/\kappa_0)^{\beta+1/2} \int_0^{\infty} x^{2\beta} I_i(x) dx. \end{aligned} \quad (14)$$

After the substitution $x = \eta(y/2)^{-1/2}$, the integrals in $G_1(\omega)$ are reduced to Euler integrals of the first kind [8], to give

$$\begin{aligned} G_1(\omega) &= A^2 (8L)^{-1} P(\alpha) (\kappa_0\tau)^\alpha (\omega\tau)^{-2\alpha}, \\ P(\alpha) &= \sum_{k=0}^{\infty} \sum_{j=0}^k a_k a_j (2 - \delta_{kj}) \\ &\times 2^{-(k+j)/2} \Gamma((k+j)/2 + \alpha) \cos((k-j)\pi/2). \end{aligned} \quad (15)$$

By the substitution $x = (\gamma\eta)^{-1}(y/2)^{1/2}$, the integral in $G_3(\omega)$ is approximated by an Euler integral of the first kind, to give

$$G_3(\omega) = (A^2/2\pi L) \Gamma(2 - \alpha) (\kappa_0\tau)^\alpha \gamma^{2(\alpha-1)} (\omega\tau)^{-2}. \quad (16)$$

The term $G_2(\omega)$ is transformed as

$$\begin{aligned} G_2(\omega) &= A^2 (\kappa_0\tau)^2 (\omega/\kappa_0)^{\beta+1/2} |W| (4L\eta^4)^{-1} \sum_{k=0}^{\infty} a_k f_k \eta^k, \\ f_k &= \int_0^{\infty} x^{2\beta-k-4} \exp[-(\gamma\eta x)^2 - (\eta/x)^2] \\ &\times \cos(\omega\Delta + \arg W - k\pi/2) dx. \end{aligned} \quad (17)$$

The asymptotic estimate of the integral f_k using the Laplace method [7] yields

$$\begin{aligned} f_k &\sim (\sqrt{\pi}/4) \gamma^{1-\beta+k/2} \\ &\times \exp(-\omega\Delta) \cos(\omega\Delta - (2k-1)\pi/4). \end{aligned} \quad (18)$$

Substituting Eq. (18) into Eq. (17), one obtains the estimate for $G_2(\omega)$ in the form

$$\begin{aligned} G_2(\omega) &\sim q \exp(-\omega\Delta) \\ &\times [S_s \sin(\omega\Delta + \pi/4) + S_c \cos(\omega\Delta + \pi/4)], \\ q &= (A\kappa_0\tau/4)^2 \gamma^{1-\beta} (\sqrt{2}\eta^5 L)^{-1} (\omega/\kappa_0)^{\beta+1/2}, \\ S_s &= \sum_{k=0}^{\infty} a_k (\omega\Delta/2)^{k/2} \sin(k\pi/2), \\ S_c &= \sum_{k=0}^{\infty} a_k (\omega\Delta/2)^{k/2} \cos(k\pi/2). \end{aligned}$$

The use of identities (6.1.17) and (6.1.18) in [8, ch. 6] for estimating the sums S_s and S_c gives

$$S_s = (2\omega\Delta/\pi)^{1/2},$$

$$S_c = \pi^{-1/2} \sum_{m=0}^{\infty} (\omega\Delta/2)^m (m!(m-1/2)^{-1})$$

$$\sim O(\exp(\omega\Delta/2)).$$

For this reason, the term $G_2(\omega)$ is on the order of $\exp(-\omega\Delta/2)$ and, being exponentially small, should be omitted from the sum in Eq. (14).

It follows from these formulas that, at $(\omega\Delta)^{1/2} \gg 1$, the high-frequency asymptotic expression for the $1/f$ spectrum has the form $\omega^{-2\alpha}$ for $\alpha \leq 1$ or ω^{-2} for $\alpha > 1$ and is integrable. Since the limiting forms of $\kappa(t)$ and $D_\xi(t)$ are independent of the correlation function $R(s)$, the result obtained is also independent of it and, in this respect, is general. It is noteworthy that the correlation times of the particle velocities in the diffusion processes are usually short, so that the above high-frequency asymptotic form of the $1/f$ spectrum is practically inaccessible to the observation.

In summary, the mathematic theory of $1/f$ noise suggested in [2, 3] was used to describe the physical mechanisms for the regularization of the $1/f$ spectrum at the edges of the frequency scale. The corresponding spectrum is finite and integrable over the entire frequency axis and does not show anomalies that are contradictory

to the spectrum of a stationary random process and to the physical concepts of noises. The reason for the apparent presence of these anomalies is that the edges of the $1/f$ spectrum, where it deviates from the known anomalous form $\omega^{-\alpha}$, are practically unattainable in the experiment.

REFERENCES

1. M. J. Buckingham, *Noise in Electronic Devices and Systems* (Halsted Press, New York, 1983; Mir, Moscow, 1986).
2. A. G. Budarin, Dokl. Akad. Nauk **359**, 615 (1998) [Dokl. Phys. **43**, 218 (1998)].
3. A. G. Budarin, Radiotekhnika (Moscow), No. 11, 8 (1997).
4. A. G. Budarin, Dokl. Akad. Nauk **372**, 326 (2000) [Dokl. Phys. **45**, 208 (2000)].
5. A. A. Abrikosov, *Fundamentals of the Theory of Metals* (Nauka, Moscow, 1987; North-Holland, Amsterdam, 1988).
6. G. I. Taylor, Proc. London Math. Soc. **20**, 196 (1921–1922).
7. M. V. Fedoryuk, *Saddle-Point Technique* (Nauka, Moscow, 1977).
8. *Handbook of Mathematical Functions*, Ed. by M. Abramowitz and I. A. Stegun (Dover, New York, 1971; Nauka, Moscow, 1979).

Translated by V. Sakun

The Role of Causality in Ensuring the Ultimate Security of Relativistic Quantum Cryptography

S. N. Molotkov and S. S. Nazin

Institute of Solid State Physics, Russian Academy of Sciences, Chernogolovka, Moscow region, 142432 Russia

Received May 15, 2001

One of the central points of quantum information theory is the problem of ultimate security of quantum cryptography; i.e., the security that is due only to the fundamental laws of nature rather than to technical restrictions. It is shown that a relativistic quantum cryptosystem proposed in this paper is ultimately secure against any eavesdropping attempts. The application of relativistic causality makes it possible to simply prove the security of the cryptosystem. Moreover, because the scheme does not involve collective measurements and quantum codes, it can be experimentally implemented even at the current level of optical fiber technologies. © 2001 MAIK "Nauka/Interperiodica".

PACS numbers: 03.67.Dd

Weisner has come up in [1] with a quantum cryptography concept, which became readily available after publication [2] by Bennett and Brassard. A considerable advance was achieved by Ekert [3] and Bennett *et al.* [4]. Ekert [3] proposed a cryptosystem based on the Einstein–Podolsky–Rosen (EPR) effect [5]. Bennett *et al.* [4] demonstrated that any eavesdropping attempts can be detected by using any pair of nonorthogonal states. Later, a variety of quantum cryptosystems and their implementations were proposed [6]. At present, there are three variants of proving the ultimate security. The proof by Mayers and Yao [7], as well as that by Biham *et al.* [8], concerns the so-called BB84 protocol [2]. The proof by Hoi-Kwong Lo and Chau [9] concerns the EPR-effect protocol [3] and, in contrast to [7], implies that the legal users have a quantum computer. Shor and Preskill [10] attempted to simplify the above-mentioned proofs by explicitly introducing quantum codes to a cryptoscheme. Goldenberg and Vaidman [11] proposed the first relativistic quantum cryptoscheme, whose ultimate security was briefly proven in [12]. As early as 1931, Landau and Peierls [13] first discussed the restrictions imposed on the measurability of quantum states in the relativistic region. Further investigation was carried out by Bohr and Rosenfeld [14].

A protocol is secure if, for any $N \geq 1$, $\varepsilon_1 > 0$, and $\varepsilon_2 > 0$, its parameters (states under use, measurements, etc.) can be chosen so that

(i) the probability that the rows $s_A(N)$ and $s_B(N)$ differ from each other in at least one bit is below ε_1 ; i.e.,

$$\Pr\{s_A(N) \neq s_B(N)\} \leq \varepsilon_1; \quad (1)$$

in other words (in terms of mutual information between

participants A and B), the inequality

$$I(A; B) \geq N - \varepsilon'_1 \quad (2)$$

can be satisfied for any $\varepsilon'_1 > 0$;

(ii) the probability that eavesdropper E knows the row $s_A(N)$ is higher than the probability 2^{-N} of simple guessing by no more than ε_2 ,

$$\Pr\{s_A(N) = s_E(N)\} \leq 2^{-N} + \varepsilon_2, \quad (3)$$

which means that he has only negligibly small information about the rows $s_A(N)$ and $s_B(N)$, which are taken as a key of the length N by the legal participants; i.e.,

$$I(A; E) \leq \varepsilon_2, \quad I(B; E) \leq \varepsilon_2. \quad (4)$$

If it is necessary to send one bit of classical information, participant A attributes the classical bits 0 and 1 to two quantum states with the density matrices ρ_0 and ρ_1 , respectively, which are chosen with the *a priori* probabilities π_0 and π_1 , $\pi_0 + \pi_1 = 1$. Measurements are described as decompositions of unity in the space of states, $\sum_i E_i = I$. As a result of measurements, the information available for participant B about the bit of participant A is determined as maximal mutual information over all possible measurements:

$$\begin{aligned} & I(A; B, \rho_0; \rho_1) \\ &= \max_{\{E_i\}} \sum_i \left\{ \pi_0 \text{Tr}\{\rho_0 E_i\} \log_2 \left(\frac{\text{Tr}\{\rho_0 E_i\}}{\text{Tr}\{\rho E_i\}} \right) \right. \\ & \quad \left. + \pi_1 \text{Tr}\{\rho_1 E_i\} \log_2 \left(\frac{\text{Tr}\{\rho_1 E_i\}}{\text{Tr}\{\rho E_i\}} \right) \right\}. \end{aligned} \quad (5)$$

A fundamental upper bound on the available information is represented by the following inequality, which was proven by Holevo [15] (see also [16]):

$$I(A; B, \rho_0; \rho_1) \leq S_{\text{vN}}(\rho) - \sum_{i=0,1} \pi_i S_{\text{vN}}(\rho_i), \quad (6)$$

$$S_{\text{vN}}(\rho) = -\text{Tr}\{\rho \log(\rho)\},$$

where $S_{\text{vN}}(\rho)$ is the von Neumann entropy [17] and the equality is attained if and only if the density matrices ρ_0 and ρ_1 commute with each other. For pure states, the latter statement means that the equality in Eq. (4) is achieved only for orthogonal states, i.e., when $\rho_{0,1} = |\psi_{0,1}\rangle\langle\psi_{0,1}|$ and $\langle\psi_0|\psi_1\rangle = 0$. In this case, available information attains the maximum

$$I^{\text{max}}(A; B, \rho_0; \rho_1) = 1, \quad (7)$$

$$E_0 = \mathcal{P}_0 = |\psi_0\rangle\langle\psi_0|, \quad E_1 = \mathcal{P}_1 = |\psi_1\rangle\langle\psi_1|.$$

For this reason, we will consider a simple one-dimensional model, which involves necessary restrictions dictated by the relativistic causality [18].

The legal participants check spatially remote regions Ω_A and Ω_B of the dimension L . At the beginning of the protocol, participant A equiprobably prepares one of the following two orthogonal states corresponding to 0 or 1:

$$|\psi_{0,1}\rangle = \int_0^\infty dk \mathcal{F}(k) a_{0,1}^\dagger(k) |0\rangle$$

$$= \int_0^\infty dk \mathcal{F}(k) |k, e_{0,1}\rangle = |\mathcal{F}, e_{0,1}\rangle, \quad (8)$$

$$|k, e_{0,1}\rangle = a_{0,1}^\dagger(k) |0\rangle,$$

$$\langle k, e_i | k', e_j \rangle = \delta(k - k') \delta_{ij},$$

where $a_{0,1}^\dagger(k)$ is the creation operator for a photon with momentum (energy) $k > 0$ and one of the orthogonal polarization states e_0 and e_1 ; $\mathcal{F}(k)$ is the state amplitude in the k representation; $i, j = 0, 1$; and $k \in (0, \infty)$. In the coordinate representation, the states have the form

$$|\psi_{0,1}\rangle = \int_{-\infty}^\infty \mathcal{F}(x-t) |x, t\rangle \otimes |e_{0,1}\rangle, \quad (9)$$

$$\mathcal{F}(x-t) = \int_0^\infty dk \mathcal{F}(k) e^{ik(x-t)}, \quad (10)$$

$$\langle k|x, t\rangle = \frac{1}{\sqrt{2\pi}} e^{ik(x-t)}, \quad x, t \in (-\infty, \infty).$$

The normalization condition in the τ representation, with allowance made for [19], has the form

$$\int_{-\infty}^\infty dx e^{ik(x-t)} \frac{1}{x-t+a} = i\pi \text{sgn}(k) e^{-ika}, \quad (11)$$

$$\langle\psi_{0,1}|\psi_{0,1}\rangle = \langle\mathcal{F}|\mathcal{F}\rangle$$

$$= \int_{-\infty}^\infty \int_{-\infty}^\infty dx dx' \mathcal{F}(x-t) \mathcal{F}^*(x'-t) \left[\frac{1}{2} \delta(x-x') \right. \quad (12)$$

$$\left. + \frac{i}{\pi x-x'} \right] = \int_{-\infty}^\infty |\mathcal{F}(x-t)|^2 dx.$$

The states are chosen to be almost monochromatic, so that the amplitude is a ‘‘step’’ $\mathcal{F}(\tau) \approx \text{const} \approx 1/\sqrt{L}$ with tails at the ends and

$$\int_{\{L\}} dx |\mathcal{F}(x-t)|^2 = 1 - \delta, \quad \delta \rightarrow 0. \quad (13)$$

The falloff at the ends can be chosen as sharp as is wished so that δ is as small as is wished. The latter is considered as being the case, and the parameter δ is the smallest parameter of the problem with any accuracy (for details, see [20–22]).

It is substantial for the protocol that the length L_{ch} of the quantum channel should be shorter than the effective state length, $L_{\text{ch}} < L$.

Later (at moment t_B), when the state entirely arrives at the region Ω_B of dimension L , participant B , who checks this region, carries out measurements that are described by the decomposition of unity

$$I = \int_{-\infty}^\infty dx |x, t_B\rangle\langle x, t_B| \otimes I_{\text{C}^2} \quad (14)$$

$$= \mathcal{P}_0(t_B) + \mathcal{P}_1(t_B) + \mathcal{P}_\perp(t_B),$$

$$\mathcal{P}_{0,1}(t_B) = |\mathcal{F}_{t_B}, e_{0,1}\rangle\langle\mathcal{F}_{t_B}, e_{0,1}|,$$

$$|\mathcal{F}_{t_B}, e_{0,1}\rangle = \int_{-\infty}^\infty dx \mathcal{F}(x-t_B) |x, t_B\rangle \otimes |e_{0,1}\rangle, \quad (15)$$

$$x \in \Omega_B, \quad \mathcal{P}_\perp(t_B) = I - \mathcal{P}_0(t_B) - \mathcal{P}_1(t_B).$$

These measurements mean that participant B takes the projection onto a state whose amplitude is entirely within the region Ω_B . The probabilities of outcomes at time t_B are

$$\text{Pr}\{i, t_B; j\} = \text{Tr}\{|\psi_i\rangle\langle\psi_i| \mathcal{P}_j(t_B)\} \quad (16)$$

$$= \delta_{ij} \int_{\{L\}} dx |\mathcal{F}(x-t_B)|^2 = \delta_{ij}, \quad x \in \Omega_B.$$

The total probability of error is the sum of two terms. The first term corresponds to the case when the eavesdropper's instrument fails to operate at all. In this case, the probability of error is equal to $1/2$ (this is the probability of error in simple guessing). The second term is the error when the eavesdropper's instrument comes into action in the region accessible to him. Because of the local orthogonality, the probability of this error is zero. Formally, the total probability of error is represented as

$$P_e(t_E) = P_e(\Omega_E, t_E) + P_e(\bar{\Omega}_E, t_E), \quad (17)$$

where Ω_E is the region accessible to the eavesdropper (correspondingly, $\bar{\Omega}_E$ is the inaccessible region; i.e., it complements Ω_E to the whole coordinate space) and t_E is the time of measurement in the region Ω_E . A complete unity decomposition is

$$I = I(\Omega_E, t_E) + I(\bar{\Omega}_E, t_E),$$

$$I(\bar{\Omega}_E, t_E) = \sum_{i=0,1} \int_{\bar{\Omega}_E} dx |x, t_E, e_i\rangle \langle x, t_E, e_i|, \quad (18)$$

$$I(\Omega_E, t_E) = \sum_{i=0,1} \int_{\Omega_E} dx |x, t_E, e_i\rangle \langle x, t_E, e_i|.$$

The minimal probability $P_e(\Omega_E, t_E)$ of error is determined by minimizing over all possible decompositions of unity $I(\Omega_E)$ [16]:

$$P_e(\Omega_E, t_E) = \min_{E_0, E_1} \left\{ \frac{1}{2} \text{Tr}\{|\psi_0\rangle \langle \psi_0| E_1\} + \frac{1}{2} \text{Tr}\{|\psi_1\rangle \langle \psi_1| E_0\} \right\}. \quad (19)$$

The quantities $E_{0,1}$ are easily obtained and the total error $P_e(\Omega_E, t_E) \equiv 0$ of discrimination is

$$E_{0,1} = \int_{\Omega_E} dx |x, t_E; e_{0,1}\rangle \langle x, t_E; e_{0,1}|, \quad (20)$$

$$P_e(\bar{\Omega}_E, t_E) = \frac{1}{2} N(\bar{\Omega}_E, t_E) = \frac{1}{2} \int_{\bar{\Omega}_E} dx |\mathcal{F}(x - t_E)|^2.$$

Therefore, the probability that eavesdropper E correctly identifies a bit of participant A at fixed size of the accessible region Ω_E is

$$P_{OK}(t_E) = 1 - P_e(\bar{\Omega}_E, t_E) - P_e(\Omega_E, t_E) = \frac{1}{2} \left(1 + \int_{\Omega_E} dx |\mathcal{F}(x - t_E)|^2 \right). \quad (21)$$

The information available to the eavesdropper about the bit of participant A is calculated by Eq. (5) taking into

account that measurement is described by the decomposition of unity $\{E_i\} = \{I_{\bar{\Omega}_E}, E_0, E_1\}$. The available information is the following sum of those parts of mutual information which are provided by the outcomes in the inaccessible and accessible Ω_E regions, respectively:

$$I(A; E, \Omega_E, t_E) = I(A; E, \rho_0, \rho_1, \Omega_E, t_E) + I(A; E, \rho_0, \rho_1, \bar{\Omega}_E, t_E). \quad (22)$$

Calculation by Eq. (5), with allowance made for $\{E_i\} = \{I_{\bar{\Omega}_E}, E_0, E_1\}$ and $\pi_0 = \pi_1 = 1/2$, yields

$$I(A; E, \rho_0, \rho_1, \bar{\Omega}_E, t_E) = 0, \quad (23)$$

$$\text{Tr}\{\rho_{0,1} I_{\bar{\Omega}_E}\} = \frac{1}{2} \text{Tr}\{\rho I_{\bar{\Omega}_E}\}, \quad \rho = \frac{1}{2}(\rho_0 + \rho_1),$$

$$I(A; E, \rho_0, \rho_1, \Omega_E, t_E) = \int_{\Omega_E} dx |\mathcal{F}(x - t_E)|^2, \quad (24)$$

$$\text{Tr}\{\rho_0 E_0\} = \text{Tr}\{\rho_1 E_1\} = \text{Tr}\{\rho E_{0,1}\}.$$

Let χ be the effective extension of the eavesdropper region (compared to the communication channel length accessible to him). The probability of the eavesdropper's error in distinguishing the states is

$$\text{Pr}_E\{\chi\} = \frac{1}{2} \left(1 + \int_{\{L_{\text{ch}} + \chi\}} dx |\mathcal{F}(x - t_E)|^2 \right) = \frac{1}{2} \left(1 + \frac{L_{\text{ch}} + \chi}{L} \right). \quad (25)$$

The probability that any state $|\tilde{\mathcal{F}}\rangle$ delayed for time (distance) χ passes the delay test by legal participant B when he carries out a measurement described by Eqs. (14), (15) is given by

$$\text{Pr}_B\{\chi\} = \text{Tr}\{|\tilde{\mathcal{F}}\rangle \langle \tilde{\mathcal{F}}| (\mathcal{P}_0(t_B) + \mathcal{P}_1(t_B))\} = \left| \int_{\{L - \chi\}} dx \mathcal{F}(x - t_B) \tilde{\mathcal{F}}^*(x - t_B) \right|^2 \leq \left(\int_{\{L - \chi\}} dx |\mathcal{F}(x - t_B)|^2 \right) \left(\int_{\{L - \chi\}} dx |\tilde{\mathcal{F}}(x - t_B)|^2 \right) \leq \left(1 - \frac{\chi}{L} \right). \quad (26)$$

It is sufficient to restrict our consideration to the pure delayed states $\tilde{\mathcal{F}}$. The restriction of the integration lim-

its to the region $L - \chi$ is caused by the existence of the limiting propagation speed and, therefore, by the fact that none of states delayed for time χ can arrive at the extreme right part of the region L by the measurement time t_B .

Thus, the probability for the eavesdropper to know the transmitted bit and pass the test by participant B is

$$\Pr\{\text{bit}_E = \text{bit}_A \wedge \text{pass test}, \chi\} = \Pr_E\{\chi\} \Pr_B\{\chi\} \\ = \frac{1}{2} \left(1 + \frac{L_{\text{ch}} + \chi}{L}\right) \left(1 - \frac{\chi}{L}\right), \quad \Pr = \frac{1}{2} \left(1 + \frac{L_{\text{ch}}}{L}\right). \quad (27)$$

The maximum of probability (27) is attained at the end of the interval at $\chi = 0$.

We demonstrate now that probability (27) in a channel with noise cannot exceed the corresponding value in an ideal channel. A noise-induced change in the state can be described by an instrument including the relativistic restrictions on it. The instrument has the general form [24–27]

$$\mathcal{T}[\dots] = \sum_k \mathcal{G}_k[\dots] \mathcal{G}_k^+, \quad \mathcal{G}_k = \sqrt{\lambda_k} |\phi_k\rangle \langle \phi_k|, \quad (28)$$

$$\sum_k \lambda_k \mathcal{G}_k \mathcal{G}_k^+ \leq 1, \quad \lambda_k \geq 0,$$

$$\begin{aligned} & \text{Tr}\{\mathcal{T}[|\psi_{0,1}\rangle \langle \psi_{0,1}|] I(\Omega_E, t_e)\} \\ &= \sum_k \text{Tr}\{|\psi_{0,1}\rangle \langle \psi_{0,1}| (\mathcal{G}_k I(\Omega_E, t_e) \mathcal{G}_k^+)\} \\ &\leq \sum_k \lambda_k \text{Tr}\{|\psi_{0,1}\rangle \langle \psi_{0,1}| (|\phi_k\rangle \langle \phi_k|)\} \\ &\leq \sum_k \lambda_k |\langle \phi_k | \psi_{0,1} \rangle|^2 \end{aligned} \quad (29)$$

$$\begin{aligned} &\leq \sum_k \lambda_k \langle \phi_k | \phi_k \rangle \langle \psi_{0,1} | \psi_{0,1} \rangle \leq \langle \psi_{0,1} | \psi_{0,1} \rangle \\ &= \int_{\Omega_A} dx |\mathcal{F}(x - t_A)|^2 = \int_{\Omega_E} dx |\mathcal{F}(x - t_E)|^2. \end{aligned}$$

The latter equality in Eq. (29) means that the amplitude $\mathcal{F}(x - t_A)$ of the state

$$\begin{aligned} |\psi_{0,1}\rangle &= \int_{\Omega_A} dx \mathcal{F}(x - t_A) |x, t_A\rangle \otimes |e_{0,1}\rangle \\ &= \int_{\Omega_E} dx \mathcal{F}(x - t_E) |x, t_E\rangle \otimes |e_{0,1}\rangle \end{aligned} \quad (30)$$

is entirely localized in the region $x \in \Omega_A$ at time t_A and will be entirely localized in the region $x \in \Omega_E$ at a time

that is later than or equal to $t_E = t_A + \text{dist}(\Omega_E, \Omega_A)$. Thus, the probability for the eavesdropper to know an individual transmitted bit and pass the test by legal participant B is no higher than $\Pr = 1/2(1 + L_{\text{ch}}/L)$ in an ideal channel.

- At *a priori* specified time instants, participant A prepares and sequentially sends the states $|\psi_{0,1}\rangle$ to the communication channel, while participant B carries out measurements described by the decomposition of unity [Eqs. (14), (15)]. Only those messages remain which pass the test.

- Participants A and B open part of messages, count the number of discrepancies, and estimate the probability of error p_{err} .

- Participant A specifies the numbers of those messages in the remaining sequence which carries only 0 or only 1. These message numbers merge into groups, each of k pieces. Participant B corrects the errors in each block by the majority principle [28]. The number k is taken so as to reduce the effective error in each block

bit $\tilde{\text{bit}}(i)$ ($\tilde{0} = \overbrace{\{0, 0, \dots, 0\}}^k$ and $\tilde{1} = \overbrace{\{1, 1, \dots, 1\}}^k$) down to $\approx p_{\text{err}}^k \ll p_{\text{err}}$. After that, the block bits are numbered.

- Using block bits, the participants form $N + M$ bits of parity $\text{Bit} = \sum_{i=1}^n \oplus \tilde{\text{bit}}(i)$. With this aim, participant

A specifies the numbers (in new number system of block bits) which will be included in each parity bit.

- For $N + M$ bits of parity $\text{Bit}(j)$, $j = 1, \dots, N + M$, the procedure including M steps of hashing is carried out. With this aim, participant A chooses at each step a random row s_l of length $N + M - l$ ($l = 1, \dots, M$) and openly reports it to participant B . Then, participants A and B check the parities of the subsets of bits in their rows (Bit_A and Bit_B) by comparing the parities with the row s_l because $s_l \text{Bit}_A = s_l \text{Bit}_B = (s_l \text{Bit}_A) \oplus (s_l \text{Bit}_B) = s_l (\text{Bit}_A \oplus \text{Bit}_B)$. If the parities of subrows coincide with each other, the participants discard one bit from the specified position in the sequences Bit_A and Bit_B . And all M steps are carried out in this way. As a result, the probability that tests for parity were passed and the remaining rows of N bits of parities Bit_A and Bit_B do not coincide with each other is [29]

$$\Pr\{s_A(N) \neq s_B(N)\} = 2^{-M}. \quad (31)$$

By choosing appropriate M , the probability that the N -bit rows are different can be made as small as is wished.

• Each parity bit can be formed from block bits in the following number of ways [30]:

$$\frac{1}{2} \sum_{i=0}^n C_{nk}^{ik} = \frac{2^{nk}}{2k} \sum_{l=1}^k \cos^{nk} \left(\frac{l\pi}{k} \right) \cos(nl\pi) \approx \frac{1}{2k} 2^{nk}. \quad (32)$$

The Hartley information I of the set of block rows is (within the accuracy of rounding off) the number of binary symbols necessary for identifying the row parity and is virtually equal to the total number of binary rows of length nk :

$$I = \log_2 \left(\frac{2^{nk}}{2k} \sum_{l=1}^k \cos^{nk} \left(\frac{l\pi}{k} \right) \cos(nl\pi) \right) \approx \eta nk, \quad (33)$$

$$\eta \approx 1;$$

i.e., almost all bits in the row must be known. The probability of knowing each bit and passing the test is less than in Eq. (27). Therefore, the conditional probability that the eavesdropper knows N resulting bits, which are sent by participant A and taken as a key, is [recall that $(1 + L_{\text{ch}}/L)/2 < 1$]

$$\begin{aligned} & \Pr\{s_A(N) = s_E(N)\} \\ &= 2^{-N} \{1 + 2 \times 2^{-\eta nk} [(1 + L_{\text{ch}}/L)]^{\eta nk}\}^N \\ &= 2^{-N} (1 + 2\zeta)^N, \\ & \zeta = 2^{-\eta nk} [(1 + L_{\text{ch}}/L)]^{\eta nk}. \end{aligned} \quad (34)$$

Mutual information acquired by the eavesdropper about the row of the resulting bits of participant A is

$$\begin{aligned} I(A; E) &= I(A) - I(A|E), \quad I(A) = -\log_2 2^{-N}, \\ I(A|E) &= -\log_2 \Pr\{s_A(N) = s_E(N)\}, \end{aligned} \quad (35)$$

where $I(A)$ is the information contained in the row of the resulting bits of length N and $I(A|E)$ is conditional information acquired by eavesdropper E about the row of the bits of participant A . Taking into account Eqs. (34) and (35), we obtain mutual information [31] between participant A and eavesdropper E in the form

$$\begin{aligned} I(A; E) &= N - N + N \log_2 (1 + 2\zeta) \approx \frac{2N\zeta}{\ln 2} \\ &= 2N \times 2^{-\eta nk} [(1 + L_{\text{ch}}/L)]^{\eta nk} / \ln 2 \ll 1. \end{aligned} \quad (36)$$

At given N and L_{ch} , L ($L_{\text{ch}} < L$) [see Eq. (36)] can be made exponentially small with respect to the parameter nk .

• We demonstrate now that the mutual information acquired by eavesdropper E about the row of resulting

bits of participant B is also exponentially small. The mutual information between participants A and B is

$$\begin{aligned} I(A; B) &= I(A) - I(A|B) \\ &= -\log_2 2^{-N} + \log_2 \Pr\{s_A(N) = s_B(N)\} \\ &= N + \log_2 (1 - 2^{-M}) \approx N - 2^{-M} / \ln 2. \end{aligned} \quad (37)$$

Using the triangle inequality for the conditional informations, we finally obtain

$$\begin{aligned} I(A|E) &\leq I(A|B) + I(B|E), \\ I(B|E) &\geq N - (2N\zeta - 2^{-M}) / \ln 2, \\ I(B; E) &\leq (2N\zeta - 2^{-M}) / \ln 2 \ll 1. \end{aligned} \quad (38)$$

Thus, the second part of the security criterion, Eqs. (3) and (4), is proved.

This study was supported by the Russian Foundation for Basic Research, project no. 99-02-18127.

REFERENCES

1. S. Wiesner, SIGACT News **15**, 78 (1983).
2. C. H. Bennett and G. Brassard, in *Proceedings of IEEE International Conference on Computer Systems and Signal Processing, Bangalore, India, 1984*, p. 175.
3. A. K. Ekert, Phys. Rev. Lett. **67**, 661 (1991).
4. C. H. Bennett, Phys. Rev. Lett. **68**, 3121 (1992); C. H. Bennett, G. Brassard, and N. D. Mermin, Phys. Rev. Lett. **68**, 557 (1992).
5. A. Einstein, B. Podolsky, and N. Rosen, Phys. Rev. **47**, 777 (1935).
6. N. Gisin, G. Ribordy, W. Tittel, and H. Zbinden, quant-ph/0101098.
7. D. Mayers and A. Yao, quant-ph/9802025.
8. E. Biham, M. Boyer, P. O. Boykin, *et al.*, quant-ph/9912053.
9. Hoi-Kwong Lo and H. F. Chau, quant-ph/9803006.
10. P. W. Shor and J. Preskill, quant-ph/0003004.
11. L. Goldenberg and L. Vaidman, Phys. Rev. Lett. **75**, 1239 (1995); quant-ph/9506030.
12. S. N. Molotkov and S. S. Nazin, quant-ph/0008008.
13. L. D. Landau and R. Peierls, Z. Phys. **69**, 56 (1931); *Collection of Works* (Nauka, Moscow, 1969), Vol. 1, p. 56.
14. N. Bohr and L. Rosenfeld, Math.-Fys. Medd. **12**, 3 (1933); *Collection of Scientific Works* (Nauka, Moscow, 1971).
15. A. S. Holevo, Probl. Peredachi Inf. **9**, 3 (1973) [Probl. Inf. Transm. **9**, 177 (1973)].
16. C. A. Fuchs, quant-ph/9601020.
17. J. von Neumann, *Mathematical Foundations of Quantum Mechanics* (Princeton Univ. Press, Princeton, 1955).
18. N. N. Bogolyubov, Izv. Akad. Nauk SSSR, Ser. Fiz. **19**, 237 (1955); N. N. Bogoliubov and D. V. Shirkov, *Introduction to the Theory of Quantized Fields* (Nauka, Moscow, 1973; Wiley, New York, 1980).
19. Yu. A. Brychkov and A. P. Prudnikov, *Integral Transforms of Generalized Functions* (Nauka, Moscow, 1977; Gordon and Breach, New York, 1989).

20. D. A. Kirzhnits, Usp. Fiz. Nauk **90**, 129 (1966) [Sov. Phys. Usp. **9**, 692 (1967)].
21. A. M. Jaffe, Phys. Rev. **158**, 1454 (1967).
22. I. Bialynicki-Birula, Phys. Rev. Lett. **80**, 5247 (1998).
23. N. Wiener and R. Paley, *Fourier Transform in the Complex Domain* (American Mathematical Society, New York, 1934; Nauka, Moscow, 1964).
24. E. B. Devis, *Quantum Theory of Open Systems* (Academic, London, 1976).
25. K. Kraus, *States, Effects and Operations* (Springer-Verlag, Berlin, 1983).
26. A. S. Holevo, in *Lectures on Statistical Structure of Quantum Theory* (1999), p. 1.
27. P. Busch, M. Grabowski, and P. J. Lahti, *Operational Quantum Physics* (Springer-Verlag, Berlin, 1995), Lect. Notes Phys. **31** (1995).
28. E. J. Mac Williams and N. J. A. Sloane, *The Theory of Error-Correcting Codes* (North-Holland, Amsterdam, 1977).
29. C. H. Bennett, Tal Mor, and J. Smolin, Phys. Rev. A **54**, 2675 (1996); Tal Mor, quant-ph/9906073.
30. A. P. Prudnikov, Yu. A. Brychkov, and O. I. Marichev, *Integrals and Series* (Nauka, Moscow, 1981; Gordon and Breach, New York, 1986).
31. C. E. Shannon, Bell Syst. Tech. J. **27**, 397 (1948); **27**, 623 (1948).

Translated by R. Tyapaev