Technical Physics, Vol. 45, No. 9, 2000, pp. 1099–1104. Translated from Zhurnal Tekhnicheskoĭ Fiziki, Vol. 70, No. 9, 2000, pp. 1–6.

On the Relationship between the Eulerian and Lagrangian Statistics of a Brownian Particle

E. Z. Gribova and A. I. Saichev

Nizhni Novgorod State University, pr. Gagarina 23, Nizhni Novgorod, 603600 Russia

E-mail: gribova@rf.um.rumet.ru

Received May 11, 1999; in final form, November 1, 1999

Abstract—Statistics of particles (smoke aerosol in the atmosphere) reaching a prescribed spatial region are analyzed. The effects of both regular and random forces acting on the particles are taken into account. A method is proposed for calculating the probability distributions of various quantities (time required to reach a prescribed location, particle velocity in its vicinity, etc.), and its scope is determined as depending on the parameters of the problem. © *2000 MAIK "Nauka/Interperiodica".*

INTRODUCTION

The increasingly heavy atmospheric and water pollution motivates studies of the physics of smoke aerosol precipitation in the ground layer. In particular, the content of several chemical elements in soil have been measured in the vicinity of an industrial facility [1]. The analysis was focused on gravitational precipitation of aerosol affected by the geoelectric field at the ground. However, diffusion should also be taken into account as a mechanism contributing to the motion of a particulate contaminant in an air flow. Diffusive mixing leads to stochastic behavior of particle trajectories, so that one may consider the statistics of both the domain that can be reached by a particle and the time required to reach it. This paper provides a description of aerosol particle motion allowing for both regular (considered in [1]) and stochastic forces.

STATEMENT OF THE PROBLEM

To analyze the diffusion of a passive contaminant in the atmosphere or ocean, it is frequently required to know the probabilistic characteristics of the contaminant particles contained in a certain spatial domain (e.g., the mean or most probable time required to reach a certain surface domain or the particle velocity statistics near the surface). In fluid dynamics, analysis of flow characteristics at fixed points is based on the Eulerian approach. Accordingly, the problem in question should naturally be described in terms of Eulerian statistics. Note that direct calculation of Eulerian statistics is a difficult task, because the particle velocity at a given point in space is a stochastic function of a random argument (the time required to reach the point).

At the same time, the Lagrangian approach is commonly used in studies of contaminant motion to obtain a probabilistic characterization of a certain particle at a current moment. The corresponding characteristics should naturally be referred to as Lagrangian. Generally, a Lagrangian statistical description is simpler, since it can be reduced to an analysis of the statistical properties of well-known solutions to Langevin stochastic equations [2]. In particular, one can readily calculate the Lagrangian statistics for passive contaminant particles. For this reason, we use known Lagrangian probability distributions to construct Eulerian ones in this paper.

Consider a contaminant particle driven by a stochastic force $\boldsymbol{\xi}(t)$ (e.g., by collisions with ambient particles) and a regular force $\boldsymbol{\xi}_0(t)$. Henceforth, both $\boldsymbol{\xi}(t)$ and $\boldsymbol{\xi}_0(t)$ are interpreted as forces per unit mass. The force $\boldsymbol{\xi}_0(t)$ may be of a different nature, depending on the particular problem under analysis. For example, the force considered in [1] represents the effects of gravity, the Coulomb force exerted by the geoelectric field on a charged particle, and advection by the wind. When a statistical characterization of the motion of charged aerosol particles includes effects of the geomagnetic field, $\boldsymbol{\xi}_0(t)$ is interpreted as an electromagnetic force. It is well known (see [2]) that the motion of a particle is described by the Langevin equations

$$\frac{d\mathbf{R}}{dt} = \mathbf{V}, \quad \frac{d\mathbf{V}}{dt} = \boldsymbol{\xi}_0(\mathbf{R}, \mathbf{V}, t) + \boldsymbol{\xi}(t) \qquad (1)$$

supplemented by the initial conditions

$$\mathbf{R}(t = 0) = \mathbf{R}_0, \quad \mathbf{V}(t = 0) = \mathbf{V}_0,$$

where $\mathbf{V}(t)$ and $\mathbf{R}(t)$ are the instantaneous velocity and location of the particle at a moment t.

The variables $\mathbf{V}(t)$ and $\mathbf{R}(t)$ are Lagrangian characteristics of diffusion, since they are associated with a specific moving particle. In seeking a statistical description, for example, of the velocities of the particles that reach a certain surface (referred to as a detector here), we establish the relationship between Lagrangian and Eulerian statistics. To provide a

1063-7842/00/4509-1099$20.00 © 2000 MAIK "Nauka/Interperiodica"

graphic illustration of the basic principles that determine this relationship, we simplify the physical setting and geometry of the problem as much as possible.

Consider a point source in space. At the initial moment, a particle having zero velocity ($\mathbf{V}_0 = 0$) is issued by the source, which is placed at the origin of the coordinate system ($\mathbf{R}_0 = 0$).

We introduce the following two assumptions concerning the interaction of the particle with the environment. First, we assume that the regular force is constant: $\boldsymbol{\xi}_0(t) = \boldsymbol{\xi}_0$. (Note that the case when $\boldsymbol{\xi}_0(t)$ is a friction force of the form $\boldsymbol{\xi}_0(t) = k(\mathbf{U} - \mathbf{V})$, where k is the effective friction coefficient and $\mathbf{U}$ is the wind velocity, was discussed in detail in [3, 4].) Second, $\boldsymbol{\xi}(t)$ is treated as a zero-mean Gaussian stochastic process with the correlation tensor

$$\langle \xi_i(t)\xi_j(t+t')\rangle = 2D\delta_{ij}\delta(t') \quad (i, j = 1, 2, 3), \quad (2)$$

where D is diffusivity.

The latter assumption means that, following [5], we consider a Brownian motion here.

We define a coordinate system $\mathbf{r} = \{\mathbf{r}_\perp, z\}$, where the z-axis is aligned with $\boldsymbol{\xi}_0$ and $\mathbf{r}_\perp = \{x, y\}$ denotes the transverse coordinates. As a detector, we choose the plane located at $z = L$. We seek the probability density functions for the moment t^* when the detector is reached by the particle and for the coordinates and velocity of the particle reaching the detector.

The simplifications concerning $\boldsymbol{\xi}_0$ and $\boldsymbol{\xi}(t)$, combined with the chosen geometry (with an infinite plane perpendicular to the direction of the regular force used as a detector), make it possible to avoid the introduction of parameters that are not essential for the desired relationship between the statistics [3, 4].

Since displacements in different directions are statistically independent by virtue of (2), the joint probability density function of the particle coordinates and velocity can be represented as

$$f(\mathbf{r}, \mathbf{v}; t) = f_\perp(\mathbf{r}_\perp, \mathbf{v}_\perp; t) f_z(z, v_z; t). \quad (3)$$

Here, $f_\perp$ and f_z are functions defined as

$$\begin{aligned} f_\perp(\mathbf{r}_\perp, \mathbf{v}_\perp; t) &= \langle \delta[\mathbf{r}_\perp - \mathbf{R}_\perp(t)]\delta[\mathbf{v}_\perp - \mathbf{V}_\perp(t)]\rangle, \\ f_z(z, v_z; t) &= \langle \delta[z - Z(t)]\delta[v_z - V_z(t)]\rangle \end{aligned} \quad (4)$$

(with averaging over realizations of the stochastic force $\boldsymbol{\xi}(t)$) and satisfying the Fokker–Planck equations [2, 6]

$$\begin{aligned} \frac{\partial f_\perp}{\partial t} + \mathbf{v}_\perp \frac{\partial f_\perp}{\partial \mathbf{r}_\perp} &= D\frac{\partial^2 f_\perp}{\partial \mathbf{v}_\perp^2}, \\ \frac{\partial f_z}{\partial t} + v_z \frac{\partial f_z}{\partial z} + \xi_0 \frac{\partial f_z}{\partial v_z} &= D\frac{\partial^2 f_z}{\partial v_z^2} \end{aligned} \quad (5)$$

subject to the initial condition

$$f(\mathbf{r}, \mathbf{v}; t = 0) = \delta(\mathbf{r} - \mathbf{R}_0)\delta(\mathbf{v} - \mathbf{V}_0).$$

The solution to (5) for $\mathbf{R}_0 = 0$ and $\mathbf{V}_0 = 0$ is expressed as

$$f_\perp(x, y, v_x, v_y; t) = \frac{3}{(2\pi D t^2)^2} \times \exp\left\{-\frac{3}{Dt}\left[\left(\frac{x^2}{t^2} - \frac{x v_x}{t} + \frac{v_x^2}{3}\right) + \left(\frac{y^2}{t^2} - \frac{y v_y}{t} + \frac{v_y^2}{3}\right)\right]\right\},$$

$$f_z(z, v_z; t) = \frac{\sqrt{3}}{2\pi D t^2}\exp\left\{-\frac{3}{Dt}\left[\frac{(z - \xi_0 t^2/2)^2}{t^2} - \frac{(z - \xi_0 t^2/2)(v_z - \xi_0 t)}{t} + \frac{(v_z - \xi_0 t)^2}{3}\right]\right\}.$$

Note that the probability density function $f(\mathbf{r}, \mathbf{v}; t)$ provides a Lagrangian statistical characterization of the coordinates and velocities of the particle at an arbitrary moment, whereas we seek Eulerian probability density functions for the coordinates $w_{\mathbf{r}_\perp}(\mathbf{r}_\perp; L)$ and velocities $w_{\mathbf{v}_\perp}(\mathbf{v}_\perp; L)$ and $w_{v_z}(v_z; L)$ of a particle at the detector. (It is obvious that the time required to reach the detector is a random variable.) In what follows, we use the known Lagrangian statistical characterization to find the desired Eulerian probability densities.

RELATIONSHIP BETWEEN LAGRANGIAN AND EULERIAN STATISTICS FOR A BROWNIAN PARTICLE

We define an auxiliary function $F(t, \mathbf{r}_\perp, \mathbf{v}; L)$ as

$$F(t, \mathbf{r}_\perp, \mathbf{v}; L) = |v_z| f(\mathbf{r}_\perp, L, \mathbf{v}; t). \quad (6)$$

Using (3) and (4) and invoking the representation of the delta function (e.g., see [7])

$$\delta[L - Z(t)] = \sum_i \delta(t - t_i)/|\dot{Z}|$$

where t_i is a root of the equation $Z(t) = L$ and the sum over all roots is taken, we rewrite (6) as

$$F(t, \mathbf{r}_\perp, \mathbf{v}; L) = \left\langle \sum_i \delta(t - t_i)\delta[\mathbf{r}_\perp - \mathbf{R}_\perp(t)]\delta[\mathbf{v} - \mathbf{V}(t)]\right\rangle. \quad (7)$$

Now, let us consider the Eulerian probabilistic characteristics of detected particles that can be found by using $F(t, \mathbf{r}_\perp, \mathbf{v}; L)$.

As a separate case, we consider the situation when the equation $Z(t) = L$ has a unique root t^*. Then, the sum to be averaged in (7) contains a single term and the

average, by definition, is the joint probability density function

$$w(t, \mathbf{r}_\perp, \mathbf{v}; L) = \langle \delta(t - t^*)\delta[\mathbf{r}_\perp - \mathbf{R}_\perp(t^*)]\delta[\mathbf{v} - \mathbf{V}(t^*)]\rangle \quad (8)$$

of the time when a particle reaches the detector and the particle's transverse coordinate and velocity at that moment. Averaging (8) over every pair of variables, we obtain the desired Eulerian probability density functions for the detected particle. In particular,

$$w_t(t; L) = \langle \delta[t - t^*(L)]\rangle$$

for the moment when the detector plane is crossed,

$$w_{\mathbf{r}_\perp}(\mathbf{r}_\perp; L) = \langle \delta[\mathbf{r}_\perp - \mathbf{R}_\perp(t^*)]\rangle$$

for the coordinates of the particle reaching the detector, and

$$w_{\mathbf{v}_\perp}(\mathbf{v}_\perp; L) = \langle \delta[\mathbf{v}_\perp - \mathbf{V}_\perp(t^*)]\rangle,$$

$$w_{v_z}(v_z; L) = \langle \delta[v_z - V_z(t^*)]\rangle$$

for the velocity components.

To find out when the sum in (7) effectively contains a single term, we apply the expression for the total probability

$$F(t, \mathbf{r}_\perp, \mathbf{v}; L) = \sum_{N=1}^{\infty} P(N; L) \sum_{i=1}^{N} g_i(t, \mathbf{r}_\perp, \mathbf{v}; L|N). \quad (9)$$

Here, $g_i(t, \mathbf{r}_\perp, \mathbf{v}; L|N)$ is the conditional joint probability density function for the moment t_i of the ith crossing of the detector plane and the coordinates and velocity of the detected particle subject to the condition that N crossings have occurred by the moment t, and $P(N; L)$ is the probability that the number of crossings is N. It is obvious that (9) is identical with (8) when $P(N; L) = 0$ for $N > 1$. In this case, (9) yields $w(t, \mathbf{r}_\perp, \mathbf{v}; L) = g_1(t, \mathbf{r}_\perp, \mathbf{v}; L|1)$ and the Eulerian probability density functions enumerated above are expressed as

$$w_t(t; L) = \int_{-\infty}^{\infty} d\mathbf{v}|v_z| \int_{-\infty}^{\infty} d\mathbf{r}_\perp f(\mathbf{r}_\perp, L, \mathbf{v}; t), \quad (10)$$

$$w_{\mathbf{r}_\perp}(\mathbf{r}_\perp; L) = \int_0^{\infty} dt \int_{-\infty}^{\infty} d\mathbf{v}|v_z| f(\mathbf{r}_\perp, L, \mathbf{v}; t), \quad (11)$$

$$w_{\mathbf{v}_\perp}(\mathbf{v}_\perp; L) = \int_0^{\infty} dt \int_{-\infty}^{\infty} d\mathbf{r}_\perp \int_{-\infty}^{\infty} dv_z|v_z| f(\mathbf{r}_\perp, L, \mathbf{v}; t), \quad (12)$$

$$w_{v_z}(v_z; L) = \int_0^{\infty} dt \int_{-\infty}^{\infty} d\mathbf{r}_\perp \int_{-\infty}^{\infty} d\mathbf{v}_\perp|v_z| f(\mathrm{r}_\perp, L, \mathbf{v}; t). \quad (13)$$

The conditions under which the assumption ensuring the validity of (10)–(13) holds with sufficient accuracy is discussed below.

PARAMETRIC DEPENDENCE OF THE SOLUTION

The derivation of (10)–(13) was based on the assumption that the particle can cross the detector plane only once. To check if this assumption actually holds, one must know how to calculate $P(N; L)$ for any N.

Unfortunately, we cannot calculate the probability $P(N; L)$ in the general case. The approximate method for evaluating $P(1; L)$ and $P(2; L)$ proposed in [3, 4] is actually based on the assumption that any probability corresponding to $N > 2$ can be neglected. According to the numerical simulation of particle motion reported in [4], this is the case when advection in the positive direction of the z-axis is so strong that the particle cannot return to the detector from the half-plane $z \geq L$. Then, the probability $P(N; L)$ of N crossings of the detector plane satisfies the inequality $P(M; L) \ll P(1; L)$, and the corresponding expressions in (10)–(13) are approximately valid. However, the problem remains open to discussion in the case when the diffusivity D and advection ξ_0 have arbitrary values. Here, we formulate a somewhat different condition (perhaps more amenable to analysis) under which the formulas obtained above are valid.

Let us use the fact that the problem involves the characteristic time $t_0 = \sqrt{2L/\xi_0}$ required for a particle to reach the detector in the absence of random forcing, the velocity scale $v_0 = \xi_0 t_0 = \sqrt{2L\xi_0}$, and the single dimensionless parameter $\alpha = \xi_0 v_0/D$. This parameter characterizes the relative contributions of regular and random displacements; in particular, the case of advection playing a dominant role considered in [3] corresponds to $\alpha \gg 1$. We must determine the values of α for which the average number $\langle N\rangle$ of crossings of the detector plane is close to unity. The condition $\langle N\rangle \approx 1$ (interpreted in a probabilistic sense) means precisely that the particle reaches the detector only once.

To calculate $\langle N\rangle$, we return to expression (9). Calculating the integral over all variables and using the fact that the conditional probability density functions are normalized to unity, we find the average number $\langle N\rangle$ of crossings of the detector by the particle:

$$\langle N\rangle = \int_0^{\infty} dt \int_{-\infty}^{\infty} d\mathbf{v} \int_{-\infty}^{\infty} d\mathbf{r}_\perp F(t, \mathbf{r}_\perp, \mathbf{v}; L) = \int_0^{\infty} dt \int_{-\infty}^{\infty} d\mathbf{v}|v_z| \int_{-\infty}^{\infty} d\mathbf{r}_\perp f(\mathrm{r}_\perp, L, \mathbf{v}; t). \quad (14)$$

Incidentally, we note that the right-hand side of (14) can be interpreted as a normalization condition for the

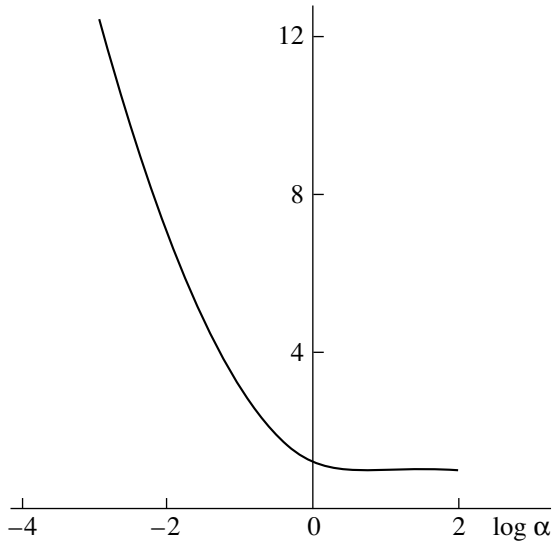


Fig. 1. Mean number of detector crossings versus the parameter α representing the relative contributions of advection and diffusion.

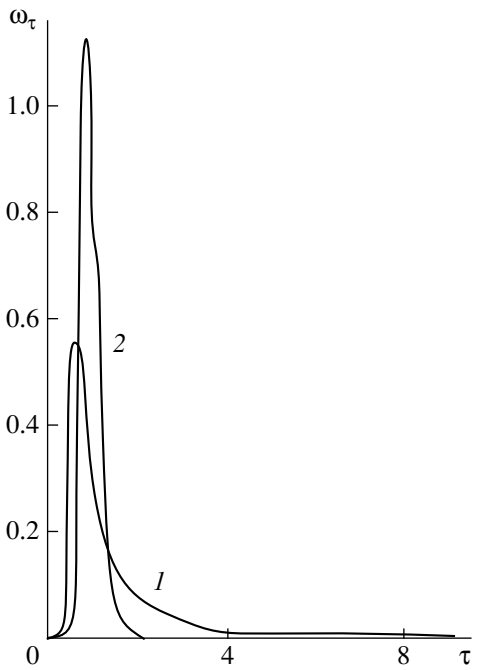


Fig. 2. Probability density function of the time required for a particle to reach the detector for α = (*1*) 1 and (*2*) 10.

probability density functions given by (10)–(13). In particular, if the particle actually crosses the detector plane only once, then $\langle N \rangle = 1$ and (10)–(13) are exact expressions for conventional (normalized to unity) probability distributions.

Figure 1 shows the dependence of $\langle N \rangle$ on the logarithm of α derived from (14). This result is quite obvious: in the case of strong advection ($\alpha \gg 1$), a particle driven by the constant force $\boldsymbol{\xi}_0$ crosses the detector plane and "escapes" into the domain $z > L$; in the opposite case ($\alpha \ll 1$, diffusion is much stronger than advection), the particle reaches the detector and crosses it many times. The results obtained here are in complete agreement with a qualitative understanding of the particle motion (trajectories corresponding to various α were presented in [4]). Thus, the single-crossing condition is fulfilled a fortiori when $\alpha \geq 1$.

Note that the relation derived here differs from that obtained in [3], where the probability of two crossings was found to be low, irrespective of the value of α. This is explained by the fact that the detector standby time is not limited in the present analysis, and the particle eventually reaches the detector even if advection is very weak (which may take infinite time). In [3], the detector standby time was assumed to be finite. Therefore, the particle was "late" in arriving at the detector: with $\alpha \ll 1$, even the probability of a single crossing was much less than unity, and the probability of two crossings was negligible (a numerical simulation of the particle motion predicts $P(2; L) \cong 10^{-1}P(1; L)$ in this case for any detector standby time consistent with the approximate normalization $P(1; L) + P(2; L) \approx 1$.

In the case of precipitation of soot emitted by a chimney stack (with an average density of 1800 kg/m^3) under real atmospheric conditions, the parameter α can be estimated as follows. Using the data presented in [1] (geoelectric field strength $E \approx 100$ V/m) and [8] (diffusivity $D \sim (0.3\ldots30)$ m^2/s^3 for a wind velocity varying from 1 to 10 m/s) for precipitating particles of radius ~100 μm emitted by a chimney stack 50 m high, we obtain $\alpha \geq 10$; i.e., relations (10)–(13) actually do hold.

EXAMPLES OF CALCULATED EULERIAN DISTRIBUTIONS

We now give some examples of the application of relations (10)–(13). For convenience, we change to the dimensionless variables $\tau = t/t_0$, $\mathbf{u} = \mathbf{v}/v_0$, and $\boldsymbol{\rho} = \mathbf{r}_\perp/L$ and consider the dimensionless probability density functions

$$\omega_\tau(\tau; \alpha) = t_0 w_t(t; L), \quad \omega_\rho(\rho; \alpha) = L^2 w_{\mathbf{r}_\perp}(\mathbf{r}_\perp; L),$$

$$\omega_{u_z}(u_z; \alpha) = v_0 w_{v_z}(v_z; L),$$

$$\omega_{u_\chi}(u_\chi; \alpha) = v_0 w_{v_x}(v_z; L)$$

(the probability densities are identical as functions of transverse coordinates $\chi = x/L$ and $\zeta = y/L$ and transverse velocity components u_χ and u_ζ). Integrating (10), (12), and (13) with respect to the coordinates and (10), (11), and (13) with respect to the transverse velocity components, we obtain

$$\omega_\tau(\tau; \alpha) = \int_{-\infty}^{\infty} du_z |u_z| \varphi_z(u_z; \tau; \alpha), \qquad (15)$$

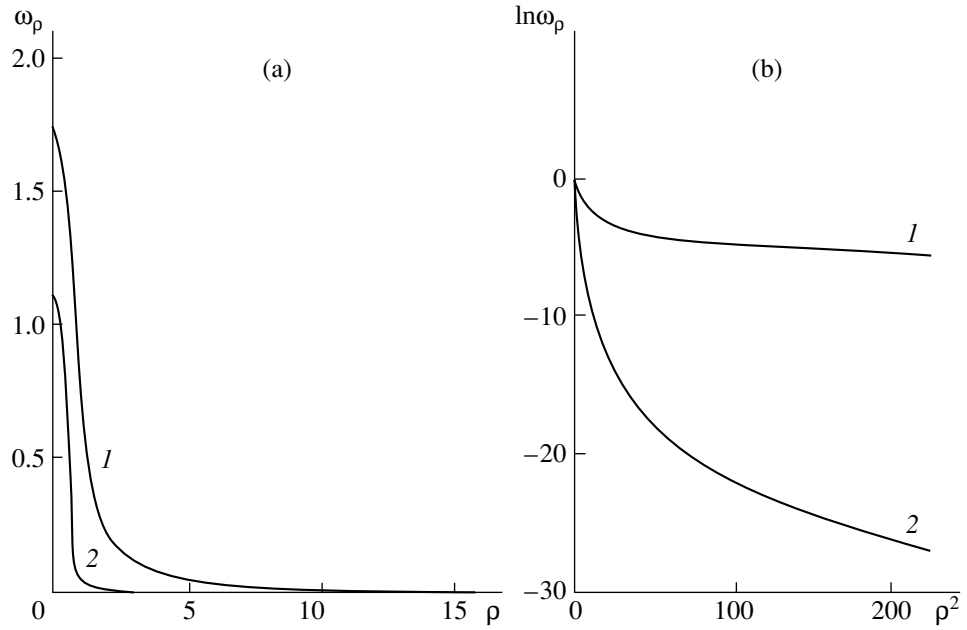


Fig. 3. (a) Probability density function of the transverse coordinate of a particle reaching the detector and (b) the logarithm of the probability for α = (*1*) 1 and (*2*) 10.

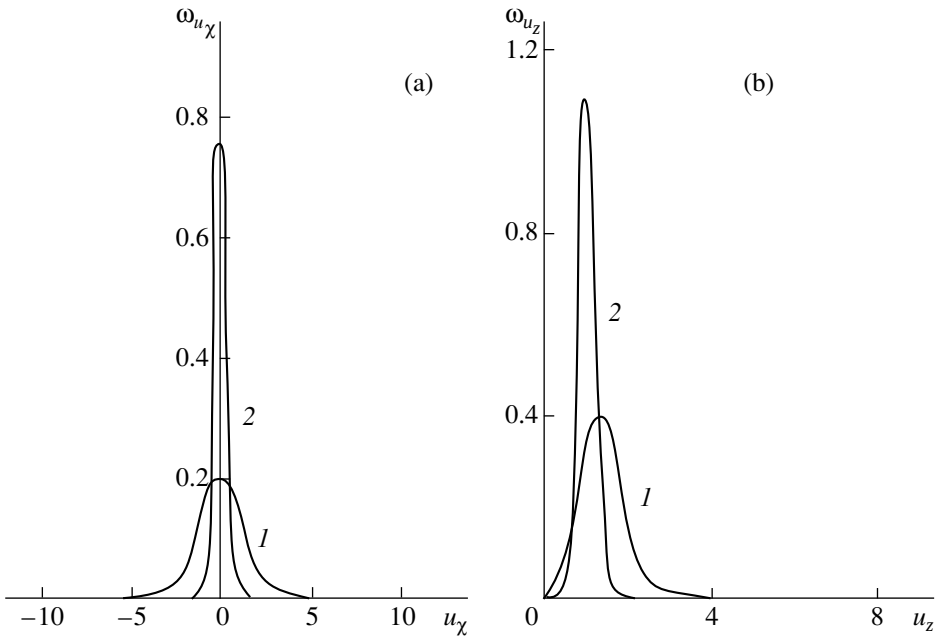


Fig. 4. (a) Probability density functions of (a) the transverse and (b) longitudinal components of the velocity of a detected particle for α = (*1*) 1 and (*2*) 10.

$$\omega_\rho(\rho;\alpha) = \frac{3\alpha}{8\pi}\int_0^\infty \frac{d\tau}{\tau^3}\exp\left(-\frac{3\alpha\rho^2}{8\tau^3}\right) \times \int_{-\infty}^{\infty} du_z |u_z| \varphi_z(u_z;\tau;\alpha), \quad (16)$$

$$\omega_{u_\chi}(u_\chi;\alpha) = \sqrt{\frac{\alpha}{2\pi}}\int_0^\infty d\tau \exp\left(-\frac{\alpha u_\chi^2}{2\tau}\right) \times \int_{-\infty}^{\infty} du_z |u_z| \varphi_z(u_z;\tau;\alpha), \quad (17)$$

$$\omega_{u_z}(u_z;\,\alpha) = |u_z|\int_0^\infty d\tau\varphi_z(u_z;\,\tau;\,\alpha), \qquad (18)$$

where

$$\varphi_z(u_z;\,\tau;\,\alpha) = v_0 f_z(L,\,v_z;\,t) = \frac{\sqrt{3\alpha}}{2\pi\tau^2}\exp\left\{-\frac{3\alpha}{8\tau}\right.$$

$$\left.\times\left[\frac{(1-\tau^2)^2}{2\tau^2} - \frac{(1-\tau^2)(u_z-\tau)}{\tau} + \frac{2(u_z-\tau)^2}{3}\right]\right\}.$$

The results of calculations of (15)–(18) performed for two values of α are shown in Figs. 2–4, where curves *1* and *2* correspond to $\alpha = 1$ and $\alpha = 10$, respectively. The figures demonstrate that an increase in α (with decreasing diffusivity D) results in decreasing variances of the transverse coordinate and velocity components at the detector (see Figs. 3a and 4), which could be expected for a Brownian particle. An increasing velocity variance is associated with a long "tail" appearing in the probability distribution of the detector-crossing time (Fig. 2) when $\alpha = 1$: when α is large, the particles are characterized by a narrower velocity scatter, reaching the detector "simultaneously." In addition, Figure 3b demonstrates the deviation of the distribution of the transverse coordinate from a Gaussian shape, which increases with α.

CONCLUSIONS

We have found the solution to a problem that is important for applications: we have obtained the probability distributions for a Brownian particle reaching a detector. In doing so, we used a well-known solution to a classical problem: the joint probability density function for particle coordinates and velocity at an arbitrary moment. The simplest possible case considered here provides a suitable framework for formulating a condition under which the relations between the probability density functions are valid. Numerical results are in complete agreement with a qualitative understanding of Brownian motion.

Note that formulas relating the Eulerian and Lagrangian statistics of a Brownian particle can readily be established for multivariate probability densities. For example, when the joint distribution of coordinates is known, probability density functions analogous to (10)–(13) can be found by invoking some properties of the delta function

$$f_2(L,\,\mathbf{r}'_\perp,\,\mathbf{r}''_\perp,\,\mathbf{v}',\,\mathbf{v}'',\,t',\,t'')$$
$$= \langle\delta[L - Z(t')]\delta[L - Z(t'')]\delta[\mathbf{r}'_\perp - \mathbf{R}_\perp(t')]$$
$$\times\delta[\mathbf{r}''_\perp - \mathbf{R}_\perp(t'')]\delta[\mathbf{v}' - \mathbf{V}(t')]\delta[\mathbf{v}'' - \mathbf{V}(t'')]\rangle,$$

Moreover, if a particle crosses the detector plane many times, the function f_2 can be used to express a quantity that is important for practical applications: the mean square number of detector crossings

$$\langle N^2\rangle = \int_0^\infty\int_0^\infty dt'dt''\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} d\mathbf{v}'d\mathbf{v}''\left|v'_z v''_z\right|$$
$$\times\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} d\mathbf{r}'_\perp d\mathbf{r}''_\perp f_2(L,\,\mathbf{r}'_\perp,\,\mathbf{r}''_\perp,\,\mathbf{v}',\,\mathbf{v}'',\,t',\,t'').$$

It is obvious that a change in the law of particle–environment interaction (e.g., allowance for nonzero correlation time of the stochastic process $\boldsymbol{\xi}(t)$ or particle-velocity relaxation time when $\boldsymbol{\xi}_0(t)$ is a viscous friction force) or in the geometry of the problem (finite size or curvature radius of the detector) will lead to a dependence on additional parameters. However, since the relationship established above relies only on general properties of the delta function, it should be expected that the relations obtained here remain valid under more stringent conditions.

ACKNOWLEDGMENTS

This work was supported by the Russian Foundation for Basic Research, project nos. 97-02-16521 and 95-IN-RU-723; by the Ministry of Education of the Russian Federation, project no. 3878; and by the Nizhni Novgorod International Center for Basic and Applied Research, project no. 99-2-09.

REFERENCES

1. A. I. Grigor'ev and T. I. Sidorova, Zh. Tekh. Fiz. **68** (3), 20 (1998) [Tech. Phys. **43**, 283 (1998)].
2. S. Chandrasekhar, *Stochastic Problems in Physics and Astronomy* (AIP, New York, 1943; Inostrannaya Literatura, Moscow, 1947).
3. E. Z. Gribova and A. I. Saichev, Izv. Akad. Nauk, Fiz. Atmos. Okeana **33**, 654 (1997).
4. E. Z. Gribova and A. I. Saichev, Izv. Vyssh. Uchebn. Zaved., Radiofiz. **41**, 1301 (1998).
5. R. H. Kraichnan, J. Fluid Mech. **64**, 737 (1974).
6. V. I. Klyatskin, *Stochastic Equations and Waves in Randomly Inhomogeneous Media* (Nauka, Moscow, 1980).
7. A. I. Saichev and W. A. Wojczynski, *Distributions in the Physical and Engineering Sciences* (Birkhäuser, Boston, 1997).
8. G. T. Csanady, *Turbulent Diffusion in the Environment* (Reidel, Dordrecht, 1980).

Translated by A. Betev

Technical Physics, Vol. 45, No. 9, 2000, pp. 1105–1109. Translated from Zhurnal Tekhnicheskoĭ Fiziki, Vol. 70, No. 9, 2000, pp. 7–12.

Depth Distribution of Light Ions under Grazing Incidence on a Target

V. V. Marinyuk and V. S. Remizovich

Moscow State Engineering Physics Institute (Technical University), Kashirskoe sh. 31, Moscow, 115409 Russia

E-mail: remix@glasnet.ru

Received June 28, 1999

Abstract—The depth distribution of light ions under grazing incidence on the surface of a semi-infinite layer was analytically derived. It was assumed that the interaction between the ions and atoms of the medium is described by potentials in the form of an inverse power function: ($V(r) \sim r^{-1/\nu}$). Calculations showed that the ion distribution (ion density) peaks at some depth, rather than being a monotonic function. The more slowly the potential decreases (the larger the value of ν), the more distinct the ion density peak and the deeper its position. At large depths, the ion density drops according to a power law. *© 2000 MAIK "Nauka/Interperiodica".*

INTRODUCTION

When a target is irradiated by medium-energy ions at grazing angles, the reflected flux consists largely of ions that have undergone multiple small-angle scattering. The problem of small-angle scattering has been comprehensively treated over the last few decades. Using the Fokker–Planck approximation, Firsov derived the angular spectrum of reflected ions (regardless of azimuth) for purely elastic scattering [1, 2]. Without using the diffusion approximation, he also obtained the angular spectrum of reflected radiation for the case when the interaction potential has the form of an inverse quadratic function ($V(r) \sim r^{-2}$) [3]. Subsequently, Firsov's results were repeatedly extended. Within the Fokker–Planck approximation, the distributions of reflected ions over both (polar and azimuth) exit angles [4] and range [5] were deduced. In [6], the angular spectrum of reflected radiation was found for purely elastic scattering of ions (regardless of azimuth) for inverse power interaction potentials without using the Fokker–Planck approximation.

In treating solid-state-related phenomena, such as sputtering and defect formation, it is, however, insufficient to know the distribution of ions only at the target surface. Since recoil atoms are produced inside the target, the depth distribution of ions, which specifies the density of recoil atoms knocked out by an ion flux, is of great importance. To this point, the stopped ion distributions over ranges and penetration depths have been studied, most extensively [7]. However, not only stopped but also moving particles generate defects. Therefore, new information on depth distributions of ions is needed. Of particular interest is grazing incidence of an ion beam on the target. In this case, most of today's theories of sputtering [8, 9] fail, so that analysis of sputtering spectra for grazing incidence of ions still remains topical [10].

While particle reflection at grazing incidence has been extensively discussed in the literature, the depth distribution of the particle density has remained practically untouched. In this work, an analytical depth dependence of the volumetric ion density is derived for the case when ion–medium interaction is described by potentials in the form of an inverse power function ($V(r) \sim r^{-1/\nu}$). In what follows, by particles in the medium, we mean moving ions. It is worth noting that the ratio of ion fluxes moving into and out of the medium does not depend on the depth and is specified by the parameter exponent ν. We found the characteristic depth where the reflected ion flux forms. For small glancing angles, this depth turns out to be much smaller than the final ion range. Thus, the purely elastic ion scattering approximation can be used in calculations.

STATEMENT OF THE PROBLEM: ION FLUX DENSITY VS. DEPTH

Let a broad ion beam with an initial energy T_0 strike the surface of a semi-infinite homogeneous target $z \geq 0$ (the z-axis is directed normally to the surface and inward into the medium) at an angle ζ_0. The grazing angle ζ_0 is assumed to be small: $\zeta_0 \ll 1$. Ion movement is defined by the angles ζ and φ, where ζ is the angle between the ion velocity vector and the surface and φ is the azimuth angle (the initial azimuth is $\varphi_0 = 0$). The outward- and inward-directed fluxes correspond to angles $\zeta > 0$ and $\zeta = -|\zeta| < 0$, respectively. Medium-energy ions are scattered at small angles $\vartheta_{\text{eff}} \ll 1$ (ϑ_{eff} is the effective angle of single scattering). Therefore, under grazing incidence of the ions ($\zeta_0 \ll 1$), the angles ζ and φ will also be small [1–6, 11]. For the motion of light ions ($M_1 \ll M_2$, where M_1 and M_2 are the masses of a light ion and an atom of the medium, respectively), energy losses due to elastic collisions can be neglected.

1063-7842/00/4509-1105$20.00

Assume that ionization losses of ions in the medium can also be neglected (this assumption will be justified below). Then, the transport equation for ion flux density $N(\tilde{z}, \zeta, \varphi)$ takes the form [6]

$$\zeta\frac{\partial N(\tilde{z}, \zeta, \varphi)}{\partial\tilde{z}} = \int_{-\infty}^{\infty} d\varphi' \int_{-\infty}^{\infty} d\zeta' I_\nu(((\zeta-\zeta')^2 + (\varphi-\varphi')^2) \times \{N(\tilde{z}, \zeta', \varphi') - N(\tilde{z}, \zeta, \varphi)\}, \quad (1)$$

$$N(\tilde{z}=0, \zeta>0, \varphi) = N_0\delta(\zeta-\zeta_0)(\delta(\varphi)); \quad N(\tilde{z}\longrightarrow\infty, \zeta, \varphi)\longrightarrow 0. \quad (2)$$

Here, N_0 is the density of an incident ion flux; $\tilde{z} = w_{el}(T_0)z$, a depth in free path units; $w_{el}(T_0)$, the probability of an ion with an energy T_0 being scattered per unit path length; $I_\nu(\gamma^2)$, the probability per collision that an ion with the state (ζ', φ') will pass to the state (ζ, φ); and $\gamma^2 \approx (\zeta-\zeta')^2 + (\varphi+\varphi')^2$. Within the small-angle approximation, ζ and φ are usually allowed to vary infinitely. For $I_\nu(\gamma^2)$, we will use the two-parameter expression [6]

$$I_\nu(\cos\gamma) = \frac{\nu\vartheta_{\text{eff}}^{2\nu}(4+\vartheta_{\text{eff}}^2)^\nu}{\pi[(4+\vartheta_{\text{eff}}^2)^\nu - \vartheta_{\text{eff}}^{2\nu}][\vartheta_{\text{eff}}^2 + 2(1-\cos\gamma)]^{1+\nu}}, \quad (3a)$$

$$2\pi\int_0^\pi \sin\gamma I_\nu(\cos\gamma)d\gamma = 1.$$

For highly anisotropic scattering (ϑ_{eff}, $\gamma^2 \ll 1$), expression (3a) yields

$$I_\nu(\gamma^2) = \frac{1}{\pi}\frac{\nu\vartheta_{\text{eff}}^{2\nu}}{[\vartheta_{\text{eff}}^2+\gamma^2]^{1+\nu}}(\nu>0), \quad 2\pi\int_0^\infty \gamma I_\nu(\gamma^2)d\gamma = 1. \quad (3b)$$

The parameter ν specifies the rate of decrease of the scattering probability with increasing γ. For Rutherford scattering, $\nu = 1$. It was shown [6] that, if $\nu > 1$ [that is, $I_\nu(\gamma^2)$ falls faster than γ^{-4}], computations agree with results obtained using the Fokker–Planck approximation for the angular variables. If $0 < \nu < 1$ [$I_\nu(\gamma^2)$ falls more slowly than γ^{-4}], the expression obtained for the angular spectrum of backscattered ions essentially differs from Firsov's results and the Fokker–Planck approximation becomes invalid. The case $\nu = 0$ is singular. Formally, the normalizing integral for small-angle scattering indicatrix (3b) diverges at its upper limit in this case. This means that, for such ν, the small-angle approximation has virtually no domain of applicability and deserves special consideration [12].

The parameter ϑ_{eff} specifies the scattering anisotropy. If $\nu < 1$, ϑ_{eff} is related to the mean square of the single-scattering angle as

$$\langle\gamma^2\rangle_\nu \approx 2\langle 1-\cos\gamma\rangle_\nu \approx \frac{4\nu}{1-\nu}\left(\frac{\vartheta_{\text{eff}}}{2}\right)^{2\nu} \quad (0<\nu<1). \quad (4)$$

For small ϑ_{eff}, indicatrix (3b) describes the small-angle scattering probability for an inverse power potential $V(r) \sim r^{-1/\nu}$ [6, 11].

Using the method of eigenfunctions, one of the authors [6] derived a simple analytical expression for the distribution of reflected ions over polar exit angles (regardless of azimuth) in a quasi-diffusion approximation (ζ_0, $|\zeta| \gg \vartheta_{\text{eff}}$):

$$S_\nu(|\psi|) = |\zeta|\int_{-\infty}^{\infty} d\varphi N(\tilde{z}=0, -|\zeta|, \varphi) = \frac{1+2\nu}{\pi(1+\nu)}\frac{\sin(\pi\nu/(1+\nu))}{2\cos(\pi\nu/(1+\nu)) + |\psi|^{\frac{1+2\nu}{1+\nu}} + |\psi|^{-\frac{1+2\nu}{1+\nu}}}. \quad (5)$$

Here, $S_\nu(|\psi|)$ is the integral of the ion reflection function over azimuth angles and $|\psi| = |\zeta|/\zeta_0$ is the reduced reflection angle.

The quantity $S_\nu(|\psi|)d|\psi|$ is the number of ions leaving a unit surface area in a unit time at angles between $|\psi|$ and $|\psi| + d|\psi|$. The integrated-over-azimuth ion flux density at an arbitrary depth is given by

$$N(\tilde{z}, \zeta) = \int_0^\infty \lambda C_\nu(\lambda)\exp(-\lambda^3\tilde{z})\Phi_\lambda(\zeta)d\lambda. \quad (6)$$

Expression (6) is the expansion of the solution of transport equation (1) over a complete set of orthonormal eigenfunctions $\Phi_\lambda(\zeta)$ at grazing angles. In the posed problem, the eigenfunctions $\Phi_\lambda(\zeta)$ have the form [6]

$$\Phi_\lambda(\zeta) = (\sqrt{3}/\pi\lambda)\int_0^\infty d\omega\cos\left(\omega\zeta - \frac{a_\nu}{\lambda^3}\omega^{2\nu+1}\right),$$

$$a_\nu = \frac{1}{2\nu+1}\frac{\Gamma(1-\nu)}{\Gamma(1+\nu)}\left(\frac{\vartheta_{\text{eff}}}{2}\right)^{2\nu}, \quad (7a)$$

$$\int_{-\infty}^{\infty}\Phi_\lambda(\zeta)\Phi_{\lambda'}(\zeta)\zeta d\zeta = \frac{1}{\lambda}\delta(\lambda-\lambda'). \quad (7b)$$

Here, $\Gamma(x)$ is the Euler gamma function [13].

Using orthogonality condition (7b) for the functions $\Phi_\lambda(\zeta)$, one can find the expansion coefficients $C_\nu(\lambda)$:

$$C_\nu(\lambda) = \zeta_0\left\{\Phi_\lambda(\zeta_0) - \int_0^\infty d|\psi|S_\nu(|\psi|)\Phi_\lambda(-\zeta_0|\psi|)\right\}, \quad (8)$$

where the reflection function $S_\nu(|\psi|)$ is defined by (5).

To explicitly find the expansion coefficients $C_\nu(\lambda)$, it is convenient to use the Mellin transformation [14]. Eventually, we have

$$C_\nu(\lambda) = \frac{\pi\sqrt{3}}{\lambda}\int_C \frac{ds}{2\pi i}\left(\frac{a_\nu}{\zeta_0^{2\nu+1}\lambda^3}\right)^{-s} \times \frac{1}{\Gamma(1-s)\Gamma((2\nu+1)s)\sin(\pi(\nu+1)s)}. \quad (9)$$

The integral in (9) is taken along any straight line lying in the domain of analyticity of the integrand, $-(1+\nu)^{-1} < \mathrm{Res} < (1+\nu)^{-1}$. Thus, using formulas (6), (7b), and (9), we can find the integrated-over-azimuth ion flux density $N(\tilde{z}, \zeta)$ at any depth inside the medium.

DISTRIBUTION OF IONS WITH DEPTH

The calculation of the space–angular ion distribution $N(\tilde{z}, \zeta)$ at an arbitrary depth is a challenge. It was noted, however, that sometimes we need only know the ion distribution with depth, that is, their volumetric density $\rho(\tilde{z})$ (hereafter, the ion density), regardless of the direction of their movement:

$$\rho(\tilde{z}) = \int_{-\infty}^{\infty} d\zeta f(\tilde{z}, \zeta). \quad (10)$$

The ion distribution function $f(\tilde{z}, \zeta)$ is related to the flux density $N(\tilde{z}, \zeta)$ as $N(\tilde{z}, \zeta) = \nu_0 f(\tilde{z}, \zeta)$, where ν_0 is the ion velocity [15].

Let us distinguish ions moving inward into the medium (inward-directed flux) and those moving in the opposite direction, i.e., to the boundary (outward-directed flux). Then, the ion density at a depth $\tilde{z}$ can be represented in the form

$$\rho(\tilde{z}) = \rho_+(\tilde{z}) + \rho_-(\tilde{z}). \quad (11)$$

The parameters $\rho_+(\tilde{z})$ and $\rho_-(\tilde{z})$ stand for the depth ion distributions in the inward- and outward-directed fluxes, respectively:

$$\rho_\pm(\tilde{z}) = \nu_0^{-1}\int_0^\infty d\zeta N(\tilde{z}, \pm\zeta) = N_0\nu_0^{-1}\int_0^\infty \lambda C_\nu(\lambda)\exp(-\lambda^3\tilde{z})d\lambda\int_0^\infty \Phi_\lambda(\pm\zeta)d\zeta. \quad (12)$$

When substituting expression (7a) for the eigenfunctions $\Phi_\lambda(\pm\zeta)$ into (12), one should bear in mind that the order of integration with respect to ω and ζ cannot be changed. Therefore, we will proceed as follows. Multiply $\Phi_\lambda(\pm\zeta)$, introduced by (7a), by $\exp(-\zeta\delta)$ and take the integral with respect to ζ from zero to infinity, changing the order of integration with respect to ω and ζ. Then, proceeding to the limit $\delta \longrightarrow +0$ under the integral with respect to ω, we will eventually obtain

$$\int_0^\infty d\zeta\Phi_\lambda(\pm\zeta) = \frac{\sqrt{3}}{2\lambda}\left(1 \pm \frac{1}{1\nu+1}\right). \quad (13)$$

From (13), (12), and (11), we have

$$\rho(\tilde{z}) = \sqrt{3}N_0\nu_0^{-1}\int_0^\infty d\lambda C_\nu(\lambda)\exp(-\lambda^3\tilde{z}), \quad (14)$$

$$\rho_+(\tilde{z}) = \frac{\nu+1}{2\nu+1}\rho(\tilde{z}),$$
$$\rho_-(\tilde{z}) = \frac{\nu}{2\nu+1}\rho(\tilde{z}) \quad (0 < \nu \le 1). \quad (15)$$

Thus, the density of outward-moving ions is $(\nu + 1)/\nu$ times less than that of inward-moving ions at any depth. The ratio $\rho_-(\tilde{z})/\rho_+(\tilde{z})$ does not depend on depth and monotonically decreases from 1/2 at $\nu = 1$ to zero at $\nu \longrightarrow 0$.

Using the explicit form (9) of the expansion coefficients $C_\nu(\lambda)$, one easily finds the depth distribution of the ion density from (14):

$$\rho(\tau) = \pi\rho_0\int_C \frac{ds}{2\pi i}\tau^{-s} \times \frac{\Gamma(s)}{\Gamma(1-s)\Gamma((2\nu+1)s)\sin(\pi(\nu+1)s)}, \quad (16)$$

where

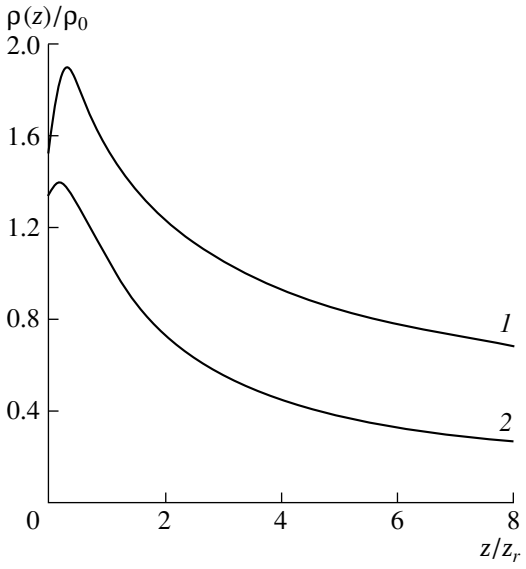
$$\tau = \frac{a_\nu\tilde{z}}{\zeta_0^{2\nu+1}} = \frac{\Gamma(1-\nu)}{(2\nu+1)\Gamma(1+\nu)}\left(\frac{\vartheta_{\mathrm{eff}}}{2}\right)^{2\nu}\frac{w_{el}(T_0)z}{\zeta_0^{2\nu+1}}, \quad \rho_0 = \frac{\nu_0}{v_0} \quad (17)$$

(ρ_0 is the incident ion density).

In (16), the integral is taken along any straight line parallel to the imaginary axis in the range $0 < \mathrm{Res} < (\nu+1)^{-1}$ in the complex plane. As follows from (16), the problem has the property of self-similarity: all the problem parameters (except ν) are combined into the single reduced depth τ, which governs the ion density. In view of (4), the reduced depth [see (17)] can be written as

$$\tau = \frac{\Gamma(2-\nu)}{\nu(2\nu+1)\Gamma(1+\nu)}\frac{\langle\Theta_s^2(T_0)\rangle z}{4\zeta_0^{2\nu+1}}, \quad (18)$$

where $\langle\Theta_s^2(T_0)\rangle = w_{el}(T_0)\langle\gamma^2\rangle$ is the mean square of angle of scattering for an ion with an energy T_0 over a unit length.



Ion density vs. depth (in terms of $z_r = \zeta_0^{2\nu+1}/\langle\Theta_s^2(T_0)\rangle$ for ν = (*1*) 1 and (*2*) 1/2.

Expressions (16)–(18) show that the quantity

$$z_r \sim \zeta_0^{2\nu+1}/\langle\Theta_s^2(T_0)\rangle \sim \zeta_0^{2\nu+1} l_{\text{tr}} \tag{19}$$

is the only problem parameter that has the dimension of length. It is the characteristic depth of formation of the reflected ion flux ($l_{\text{tr}} \approx 2/\langle\Theta_s^2(T_0)\rangle$ is the transport length of elastic scattering [15]).

From (16), one easily obtains the asymptotic expression for ion density at small and large depths. The poles of the integrand in (16) all lie on the real axis. Therefore, the contour of integration can be deformed in such a way that it does not intersect the poles. Bending the contour to the left and using the residue theorem, one can expand $\rho(\tau)$ in a generalized power series in terms of positive powers of τ. Bending the contour of integration to the right yields the expansion of $\rho(\tau)$ in a series in terms of negative powers of τ. From (16), the ions density at small depths ($\tau \ll 1$) depends on the residues of the integrand at the poles $s = 0$ and $s = -(1+\nu)^{-1}$; at large depths ($\tau \gg 1$), the density depends on the residue at the pole $s = (1+\nu)^{-1}$. Hence,

$$\rho(\tau) \approx \rho_0 \begin{cases} \dfrac{2\nu+1}{\nu+1} + (2\nu+1)\Gamma\left(\dfrac{2\nu+1}{\nu+1}\right) \\ \times\left[\Gamma\left(\dfrac{1}{\nu+1}\right)\right]^{-2}\tau^{\frac{1}{\nu+1}}, \quad \tau \ll 1 \\ \dfrac{1}{\nu}\Gamma\left(\dfrac{1}{\nu+1}\right)\left[\Gamma\left(\dfrac{1}{\nu+1}\right)\right]^{-2}\tau^{-\frac{1}{\nu+1}}, \quad \tau \gg 1. \end{cases} \tag{20}$$

Thus, if $0 < \nu \le 1$, the near-surface ion density grows with depth. The larger the value of ν, the faster the growth. At large depths, the smaller the value of ν, the faster the decrease in the ion density. Such behavior can be explained as follows. For steady-state target irradiation and purely elastic scattering, the number of ions striking a unit area per unit time equals that of ions leaving the target (the total reflection coefficient is unity [6, 15]). Before leaving the target, the ions must undergo multiple scattering in order to "face" the boundary. To do this, they must travel some distance, i.e., penetrate to some depth. Therefore, the majority of the ions will be concentrated at some depth rather than on the surface.

The rate with which $\rho(\tau)$ grows (at small depths) or drops (at large depths) depends on ν. The larger the value of ν, the more anisotropic scattering. Consequently, the smaller the value of ν, the more probable large-angle scattering becomes and the faster the ions turn about to leave the medium. That is why, as ν decreases, the ion density grows more slowly at small depths and drops faster at large depths.

To conclude this section, we note the following.

Energy loss of ions due to ionization can be neglected until the ion path in the medium is less than the maximum ion range R_0. Therefore, it should be required that the characteristic depth of reflection z_r [see (19)] be much smaller than the depth $\zeta_0 R_0$ at which the ions lose a considerable part of their energy; that is, the inequality

$$\frac{\langle\Theta_s^2(T_0)\rangle R_0}{\zeta_0^{2\nu}} \gg 1. \tag{21}$$

must be met.

With condition (21) fulfilled, the ion density $\rho(\tau)$ has time to become asymptotic ($\tau \gg 1$, or $z \gg z_r$) at depths where ion deceleration can still be neglected. Neglecting inelastic loss of ion energy in calculating reflection spectra within the Fokker–Planck approximation ($\nu = 1$) was discussed in [5]. It was shown that, if $\langle\Theta_s^2(T_0)\rangle R_0/\zeta_0^2 \gg 1$, the reflected flux consists mainly of purely elastically scattered ions. Thus, inequality (21) is the generalization of this condition for inverse power potentials that decay faster than the Coulomb potential.

ION DISTRIBUTION WITH DEPTH FOR SEVERAL INVERSE POWER INTERACTION POTENTIALS

Consider specific values of the parameter ν that simplify analytic representations for $\rho(\tau)$. The value $\nu = 1/2$ corresponds to the Firsov inverse quadratic potential ($V(r) \sim r^{-2}$). In this case, expression (16) is reduced to

$$\rho_{\nu=1/2}(\tau) = 2\sqrt{\pi}\rho_0 \int_C \frac{ds}{2\pi i}(4\tau)^{-s} \times \frac{\Gamma(s)}{\Gamma(s+1/2)}\frac{\sin\pi s}{\sin(3\pi s/2)}. \tag{22}$$

The integrand here is the product of two functions: $[\Gamma(s)/\Gamma(s + 1/2)]$ and $[\sin\pi s/\sin(3\pi s/2)]$, the Mellin inverse being known for each [16]. Then, applying the convolution theorem to the Mellin transformation, one can write the integral in (22) in the form

$$\rho_{\nu = 1/2}(\tau) = \frac{2\rho_0}{\pi\sqrt{3}}\int_0^1 \frac{dx}{x\sqrt{1-x}} \times \frac{1}{(x/4\tau)^{2/3} + (4\tau/x)^{2/3} - 1}. \quad (23)$$

At $\nu = 1/2$, the reduced depth [see (18)] is $\tau = \langle\Theta_s^2(T_0)\rangle z/4\zeta_0^2$.

The integral in (16) can also be simplified at $\nu = 1$. In this case, indicatrix (3b) stands for the probability of single scattering for the Coulomb potential ($V(r) \sim r^{-1}$). Ion density (16) can then be written in the easy-to-use form

$$\rho_{\nu = 1}(\tau) = \sqrt{3}\rho_0\int_C \frac{ds}{2\pi i}(27\tau)^{-s} \times \Gamma(s)\frac{\Gamma(s + 1/2)\Gamma(1/2 - s)}{\Gamma(s + 1/3)\Gamma(s + 2/3)}. \quad (24)$$

Since the Mellin inverses of the functions $\Gamma(s + 1/2)\Gamma(1/2 - s)[\Gamma(s + 1/3)\Gamma(s + 2/3)]^{-1}$ and $\Gamma(s)$ are tabulated [14], the application of the convolution theorem to $\rho_{\nu = 1}(\tau)$ yields

$$\rho_{\nu = 1}(\tau) = \rho_0\int_0^\infty \frac{dx}{x}\exp\left(-\frac{27\tau}{x}\right)\frac{1}{\sqrt{x}} \times \left\{J_{1,3}\left(\frac{2}{\sqrt{x}}\right) + J_{-1/3}\left(\frac{2}{\sqrt{x}}\right)\right\}, \quad (25)$$

where

$$J_\alpha(t) = \pi^{-1}\int_0^\pi d\Theta\cos(\alpha\Theta - t\sin\Theta)$$

is the Anger function [13].

Substituting the integral representation for the Anger function into (25) and changing the order of integration, we obtain

$$\rho_{\nu = 1}(\tau) = \frac{2\rho_0}{3\sqrt{3\pi\tau}}\int_0^\pi d\Theta\cos(\Theta/3)\exp\left\{-\frac{\sin^2\Theta}{27\tau}\right\}. \quad (26)$$

For $\nu = 1$, reduced depth (18) is $\tau = \langle\Theta_s^2(T_0)\rangle z/12\zeta_0^3$.

The figure plots the ion densities given by (23) and (26) with depth [in units of $\zeta_0^{2\nu+1}/\langle\Theta_s^2(T_0)\rangle$ for $\nu = 1$ and 1/2. As ν decreases, the peak of the distribution is seen to become less pronounced and shift toward smaller depths. The explanation is, as before, that at smaller values of ν the ions turn about and leave the medium faster.

Some concluding remarks. At small depths ($z \ll z_r$), the number of ions inside the medium grows, peaks at $z \sim z_r$, and then ($z \gg z_r$) drops with depth according to a power law. At $z \sim \zeta_0 R_0$, expression (16) for ion density fails, because it does not take into account the energy loss of the ions. For $z > \zeta_0 R_0$, the ion density vs. depth dependence is essentially defined by ionization deceleration.

REFERENCES

1. O. B. Firsov, Dokl. Akad. Nauk SSSR **169**, 1311 (1966) [Sov. Phys. Dokl. **11**, 732 (1967)].
2. O. B. Firsov, Fiz. Tverd. Tela (Leningrad) **9**, 2145 (1967) [Sov. Phys. Solid State **9**, 1687 (1967)].
3. O. B. Firsov, Zh. Tekh. Fiz. **40** (1), 83 (1970) [Sov. Phys. Tech. Phys. **15**, 57 (1970)].
4. V. S. Remizovich and I. S. Tilinin, Zh. Tekh. Fiz. **50**, 1524 (1980) [Sov. Phys. Tech. Phys. **25**, 883 (1980)].
5. V. S. Remizovich, I. S. Tilinin, and M. I. Ryazanov, Zh. Éksp. Teor. Fiz. **79**, 448 (1980) [Sov. Phys. JETP **52**, 225 (1980)].
6. V. S. Remizovich, Zh. Éksp. Teor. Fiz. **87**, 506 (1984) [Sov. Phys. JETP **60**, 290 (1984)].
7. W. Eckstein, *Computer Simulation of Ion–Solid Interactions* (Springer-Verlag, Berlin, 1991; Mir, Moscow, 1995).
8. *Sputtering by Particle Bombardment*, Ed. by R. Behrisch (Springer-Verlag, New York, 1981; Mir, Moscow, 1984).
9. *Sputtering of Solids: Principles and Applications*, Ed. by E. S. Mashkova, (Mir, Moscow, 1989).
10. E. S. Mashkova and V. A. Molchanov, Poverkhnost, No. 3, 5 (1995).
11. V. A. Kurnaev, E. S. Mashkova, and V. A. Molchanov, *Light Ion Reflection from Solid Surface* (Énergoatomizdat, Moscow, 1985).
12. P. Sigmund, Phys. Rev. **184**, 383 (1969).
13. *Higher Transcendental Functions (Bateman Manuscript Project)*, Ed. by A. Erdelyi (McGraw-Hill, New York, 1953; Nauka, Moscow, 1965, 1966), Vols. 1, 2.
14. O. I. Marichev, *Handbook of Integral Transforms of Higher Transcendental Functions: Theory and Algorithmic Tables* (Nauka i Tekhnika, Minsk, 1978; Horwood, Chichester, England, 1983).
15. N. P. Kalashnikov, V. S. Remizovich, and M. I. Ryazanov, *Collisions of Fast Charged Particles in Solids* (Atomizdat, Moscow, 1980).
16. *Tables of Integral Transforms (Bateman Manuscript Project)*, Ed. by A. Erdelyi (McGraw-Hill, New York, 1954; Nauka, Moscow, 1968).

Translated by V. Isaakyan

Technical Physics, Vol. 45, No. 9, 2000, pp. 1110–1114. Translated from Zhurnal Tekhnicheskoĭ Fiziki, Vol. 70, No. 9, 2000, pp. 13–16.

Classical Physics and Electron Spin

A. G. Chirkov and I. V. Kazinets

St. Petersburg State Technical University,
ul. Politekhnicheskaya 29, St. Petersburg, 195251 Russia

Received December 27, 1999

Abstract—In the framework of classical physics, the formation of the intrinsic angular momentum of an electron is described and the correct value of the gyromagnetic ratio is obtained. The Maslov–Leray quasi-classical quantization rules result in the exact values of Landau levels obtained from the Pauli equation. © *2000 MAIK "Nauka/Interperiodica".*

Goudsmit and Uhlenbeck [1], in putting forward the hypothesis for electron spin, thought of an electron as a rigid body rotating about its axis. Afterwards, this model was rejected, since the rotational velocity turned out to exceed the velocity of light. In the course of time, physicists began to regard spin as a fundamental quantum feature of an electron that still cannot be explained physically. However, the expression for the gyromagnetic ratio does not include Planck's constant; therefore, the correct result can be obtained in terms of classical physics. The aim of this study is to elucidate a formation mechanism of electron spin in classical physics.

In this paper, the term "electron" refers to a point particle that has a charge but does not have higher order moments (electric and magnetic dipole moments, quadrupole moments, etc.). We will also show that the Maslov–Leray quasi-classical quantization rules [2, 3] allow one to obtain the exact Landau spectrum that follows from the Pauli equation.

1. OBSERVABLES AS TRANSFORMATION GENERATORS

Let us consider a canonical transformation due to a close-to-identical generating function S [4]:

$$S(q, P) = q^i P_i + \varepsilon f(q, P), \quad 0 < \varepsilon \ll 1. \tag{1}$$

Hereafter, we assume summation over i from 1 to n. Under transformation (1), an arbitrary observable $g(P, q)$ turns into $g'(P, Q)$, so that

$$\delta g = \varepsilon\{f, g\}, \tag{2}$$

where $\{f, g\}$ is the Poisson bracket of the functions f and g:

$$\{f, g\} = \frac{\partial f}{\partial p_i}\frac{\partial g}{\partial q^i} - \frac{\partial f}{\partial q^i}\frac{\partial g}{\partial p_i}. \tag{3}$$

Formulas (1)–(3) show that observables play a twofold part. First, they are smooth real-valued functions of the system state; second, they are canonical transformation generators.

Indeed, putting $f(p, q) = P_i$, we find that $Q^i = q^i + \varepsilon$; i.e., the ith momentum component generates translation along the ith coordinate axis, etc. Such a correspondence between observables and canonical transforms allows for the determination of observables in classical mechanics through their associated canonical transforms. Moreover, one should take into account the Noether theorem, which relates classical integrals of motion (energy, momentum, and angular momentum) of a mechanical system to fundamental properties of space and time [4].

Thus, we define momentum as a conserved quantity that is the generator of translation in space; angular momentum, as the generator of rotation; and energy, as the generator of translation in time. Note that these definitions are absolutely general. They apply not only to discrete mechanics but to arbitrary (classical or quantum) field theories as well.

2. ELECTRON MOMENTUM IN A MAGNETIC FIELD

Consider a nonrelativistic point particle of mass m and charge e that moves in a homogeneous static magnetic field $\mathbf{B}$. The vector potential may be chosen in the form $\mathbf{A} = \mathbf{B} \times \mathbf{r}/2$. Directing the third Cartesian coordinate axis along $\mathbf{B}$, we obtain the Lagrangian function

$$L = \frac{m}{2}(\dot{x}_1^2 + \dot{x}_2^2 + \dot{x}_3^2) + \frac{eB}{2c}(x_1\dot{x}_2 - \dot{x}_1 x_2) \tag{4}$$

and the Hamiltonian function

$$H = \frac{1}{2m}\left(\pi_1 + \frac{eB}{2c}x_2\right)^2 + \frac{1}{2m}\left(\pi_2 - \frac{eB}{2c}x_1\right)^2 + \frac{1}{2m}\pi_3^2, \tag{5}$$

1063-7842/00/4509-1110$20.00 © 2000 MAIK "Nauka/Interperiodica"

where $\pi_1 = m\dot{x}_1 - (eB/2c)x_2$, $\pi_2 = m\dot{x}_2 + (eB/2c)x_1$, and $\pi_3 = m\dot{x}_3$ are the components of the generalized momentum.

The quantities $m\dot{\mathbf{r}}$, $\mathbf{p} = m\dot{\mathbf{r}}$, and $\boldsymbol{\pi}$, which coincide for a free particle, differ in the presence of the field. Hence, we face the problem of determining a quantity that for a charged particle plays the same role as momentum for an uncharged one.

According to the definition in Sect. 1, neither $m\dot{\mathbf{r}}$ nor $\boldsymbol{\pi}$ can be the electron momentum. However, using the second Noether theorem [4], one can easily find the integral of motion corresponding to the transformation of translation in space:

$$\mathbf{p} = m\dot{\mathbf{r}} + \frac{e}{c}\mathbf{B}\times\mathbf{r}. \tag{6}$$

Note that (6) is an evident integral of motion due to the Lorentz force. In this case, the Poisson bracket $\{H, \mathbf{p}\} = 0$, which corresponds to the momentum conservation law. In contrast to $\boldsymbol{\pi}$, quantity (6) is gauge-invariant.

Thus, the quantity $\mathbf{p}$ meets all of the requirements in Sect. 1 and should be considered as the momentum of a charged particle in a static homogeneous magnetic field. However, the transition from $\boldsymbol{\pi}$ to $\mathbf{p}$ is not canonical, and the Hamiltonian function cannot be expressed in terms of $\mathbf{p}$. The complexity of describing magnetic moment (spin) through the Lagrangian and Hamiltonian functions was noted as early as in 1934 by Frenkel' [5] and has still not been overcome.

Conventional analysis of the interaction between a charged particle and a magnetic field that uses the Hamiltonian function expressed in terms of the generalized momentum $\boldsymbol{\pi}$ is incorrect. Indeed, the energy of a charged particle moving in a magnetic field (Hamiltonian function) expressed in terms of velocity coincides with that of a free particle (the particle energy is conserved in a magnetic field) [5].

The incorrect interpretation results from the uncertain definition of dynamic systems in physics literature. A Hamiltonian mechanical system is defined by an even-dimensional manifold (phase space); a closed nondegenerate 2-form Ω on it, which takes the canonical form $\Omega = \sum_i dp_i\Lambda dx_i$ in simplectic coordinates x and p; and a smooth function (the Hamiltonian function) on this manifold [6]. Of the three required elements that specify Hamiltonian systems, physicists commonly pay attention to the last one alone, i.e., the Hamiltonian function. However, according to the definition, a Hamiltonian system changes when the phase space M and/or differential form Ω defined on this space change, even if the Hamiltonian function is invariable.

Consider an additional closed 2-form F on a configurational manifold N: $F = \sum_{i,j} F_{ij}\,dx_i\Lambda dx_j$ ($dF = 0$). This form is referred to as the form of gyroscopic forces. The sum $\Omega + F$ introduces a new simplectic structure on the space of the cotangent bundle of the configurational manifold N. If $H(q, p)$ is some Hamiltonian function on M, the pair $(\Omega + F, H)$ defines a new Hamiltonian system (with the same Hamiltonian function) on M. The form $\Omega + F$ may be represented canonically by using the Darboux theorem [4]. To do this, we write locally $F = dA$ and $A = \sum_k A_k(x)dx_k$, since F is the closed form. Finally, we obtain $Q + F = \sum d\,p_i\Lambda dx_i + \sum_i d\,A_i\Lambda dx_i = \sum d\,(p_i + A_i)\Lambda dx_i$ in terms of the x and p variables. Consequently, the variables $x'_k = x_k$ and $\pi_k = p_k = A_k(x)$ are new canonical coordinates on M. The Hamiltonian equations in these new coordinates have the canonical form with the Hamiltonian function $H(x', \pi' - A) = H(x, p)$.

The Lagrangian description of a point particle with a magnetic moment does not exist, since the state space of such a system $R^6 \times S^2$ is not the cotangent bundle of any configurational manifold.

In electrodynamics, F is the Faraday 2-form, F_{ik} is the electromagnetic field tensor, and A_k are the components of the vector potential. The closure of F ($dF = 0$) is the consequence of the homogeneous Maxwell equations.

3. ANGULAR ELECTRON MOMENTUM IN A MAGNETIC FIELD

Let us apply the Noether theorem to system (4) of the Lagrangian functions to find the integrals of motion that correspond to rotations about the coordinate axes. The Lagrangian function is found to vary under rotation about the first and second coordinate axes. The additional term in the Lagrangian function is not reduced to the total time derivative; i.e., there are no integrals of motion corresponding to these transformations.

The integral of motion for rotation about the third axis is defined and has the form

$$J_3 = m(x_1\dot{x}_2 - x_2\dot{x}_1) + \frac{eB}{2c}(x_1^2 + x_2^2). \tag{7}$$

Thus, the situation is entirely similar to the quantum mechanical case. The components of the angular momentum cannot simultaneously be canonical momenta. Only the projection of the angular momentum onto a specified direction (quantization axis) is conserved (defined). On the other hand, one can easily verify that $\{J^2, H\} = 0$ and $\{J^2, J_i\} = 0$, so that the magnitude of the total angular momentum is defined and can be the canonical momentum simultaneously with one of its components.

The augend in (7) is not an orbital angular momentum, since $\mathbf{p} \neq m\dot{\mathbf{r}}$. Having expressed the total angular momentum J_3 through orbital angular momentum $L_3 = x_1p_2 - x_2p_1$, we obtain

$$J_3 = L_3 - \frac{eB}{2c}(x_1^2 + x_2^2) = L_3 + S_3; \tag{8}$$

i.e., only the total angular momentum is conserved just as in the Dirac theory. The addend S_3 in (8) should be identified with the projection of the intrinsic angular momentum (spin) on the third axis. In the vector form, the corresponding quantity is given by

$$\mathbf{S} = \frac{e}{2c}\mathbf{r} \times (\mathbf{B} \times \mathbf{r}), \tag{9}$$

but only the quantity $\mathbf{Sn}$ ($\mathbf{n} = \mathbf{B}/B$) is defined (conserved).

4. ELECTRON GYROMAGNETIC RATIO

We will derive the expression for the electron gyromagnetic ratio using the relationships obtained in Sections 1–3. The third projection of the electron magnetic moment is

$$(\mathbf{M})_3 = \frac{e}{2mc}(\mathbf{r} \times m\dot{\mathbf{r}})_3 = \frac{e}{2mc}J_3 - \frac{e^2B}{4mc^2}(x_1^2 + x_2^2). \tag{10}$$

Having expressed M_3 in terms of the projection of orbital angular momentum L_3, we obtain

$$\begin{aligned} M_3 &= \frac{e}{2mc}L_3 - \frac{e^2B}{2mc^2}(x_1^2 + x_2^2) \\ &= \frac{e}{2mc}L_3 + \frac{e}{mc}S_3 = M_{\text{orb}} + M_{\text{spin}}. \end{aligned} \tag{11}$$

Thus, the orbital and spin gyromagnetic ratios are equal to $e/2mc$ and e/mc, respectively.

It is worth noting that, as in the Dirac theory, the obtained results are valid only for elementary (not composite) particles which have no internal charged structure. The term added to the momentum $m\mathbf{v}$ of an uncharged particle should be considered as an attribute of a charged particle just as the spin term in the expression for angular momentum. The origin of the additive terms is evident. In a nonrelativistic approximation, the field near an electron can be considered as Coulomb if one disregards radiation processes. Then a circulating energy flux appears around the particle in the plane orthogonal to the **B** direction. This flux is responsible for field terms added to the observables of a "bare" (having no eigenfield) particle.

The field additive to the momentum $m\mathbf{v}$ of a bare particle is calculated by the formula

$$\begin{aligned} \mathbf{p}_f &= \frac{1}{4\pi c}\int_V \mathbf{E}_0 \times \mathbf{B}\,dV \\ &= \frac{1}{2\pi c}\int_V \mathbf{A}\,\mathrm{div}\mathbf{E}\,dV = \frac{2e}{c}\mathbf{A}(\mathbf{r}_0(t)), \end{aligned} \tag{12}$$

where $\mathbf{E}_0 = e(\mathbf{r} - \mathbf{r}_0(t))/|\mathbf{r} - \mathbf{r}_0(t)|^2$ is the eigenfield of the particle, $\mathbf{r}_0(t)$ is the electron radius vector, $\mathbf{B}$ is a homogeneous external field, and V is the volume of a sphere centered at the electron position.

The field term added to the momentum gives, in its turn, the field contribution to the orbital angular momentum:

$$\mathbf{L}_f = \mathbf{r}_0 \times \mathbf{p}_f = \frac{e}{c}\mathbf{r}_0 \times (\mathbf{B} \times \mathbf{r}_0),$$

which exactly coincides with field term (9).

5. QUASI-CLASSICAL QUANTIZATION OF ELECTRON ENERGY

In cylindrical coordinates ρ, α, and z that is directed along $\mathbf{B}$, the Hamilton–Jacobi equation for the characteristic action function S has the form

$$\begin{aligned} \left(\frac{\partial S}{\partial \rho}\right)^2 + \left(\frac{1}{\rho}\frac{\partial S}{\partial \alpha} - \frac{1}{2}m\omega\rho\right)^2 + \left(\frac{\partial S}{\partial z}\right)^2 &= 2mE, \\ \omega &= \frac{eB}{mc}. \end{aligned} \tag{13}$$

The coordinates α and z are cyclic, and hence, the solution of equation (13) can be represented in the form

$$S = R(\rho) + b_\alpha\alpha + b_z z, \tag{14}$$

where b_α and b_z are arbitrary constants.

Substituting (14) into (13), we obtain

$$R(\rho) = \pm\int [2mE - b_z^2 - (b_\alpha/\rho - m\omega\rho/2)^2]^{1/2} d\rho. \tag{15}$$

The principal Hamiltonian function $V(\rho, \alpha, z; E, \pi_\alpha, \pi_z; t) = -Et + S$ is a function of canonical transformation from the variables ρ, α, z; π_α, π_z to the variables t_0, a_α, a_z, E, b_a, and b_z, which are constants. The new and old variables are related as

$$\begin{aligned} t - t_0 &= \partial S/\partial E \\ &= \pm m\int [2mE - b_z^2 - (b_\alpha/\rho - m\omega\rho/2)^2]^{-1/2} d\rho \end{aligned}$$

$$\begin{aligned} a_\alpha &= \partial S/\partial b_k = \pm\int (m\omega/2 - b_\alpha/\rho^2) \\ &\times [2mE - b_z^2 - (b_\alpha/\rho - m\omega\rho/2)^2]^{-1/2} d\rho + \alpha, \end{aligned}$$

$$\begin{aligned} a_z &= \partial S/\partial z \\ &= \mp\int[2mE - b_z^2 - (b_\alpha/\rho - m\omega\rho/2)^2]^{-1/2}d\rho + z, \\ \pi_\rho &= \partial S/\partial\rho \\ &= \pm[2mE - b_z^2 - (b_\alpha/\rho - m\omega\rho/2)^2]^{1/2}, \\ \pi_\alpha &= \partial S/\partial\alpha = b_\alpha, \\ \pi_z &= \partial S/\partial s = b_z. \end{aligned} \tag{16}$$

According to [4], instead of the generalized momenta π_ρ and π_α, we introduce new constant momenta

$$I_\rho = \frac{1}{2\pi}\oint\pi_\rho d\rho, \tag{17}$$

$$I_\alpha = \frac{1}{2\pi}\oint\pi_\alpha d\alpha, \tag{18}$$

where integration is over the period of the associated coordinate.

Expression (17) then takes the form

$$I_\rho = \frac{1}{\pi}\int_{\rho_1^*}^{\rho_2^*}[2mE - b_z^2 - (b_\alpha/\rho - m\omega\rho/2)^2]^{1/2}d\rho,$$

where $\rho_{1,2}^*$ are the zeros of the integrand.

Taking the integral with the theory of residues, we come to

$$I_\rho = \frac{E}{\omega} - \frac{b_z^2}{2m\omega} + \frac{b_\alpha}{2} - \frac{|b_\alpha|}{2}. \tag{19}$$

Note that this expression differs from those used in electrodynamics earlier [7].

Projecting the particle trajectory onto the plane (ρ, α), we obtain a circle that either encloses, or does not enclose, the origin, depending on the initial conditions. Here, a positively charged particle moves clockwise in this plane. The expression for b_α is given in the following form:

$$b_\alpha = \pi_\alpha = m\rho^2\dot\alpha + m\omega\rho^2/2 = m\rho^2(\dot\alpha + \omega/2). \tag{20}$$

If the circle does not enclose the origin, there exist points A and B on this circle where $\dot\alpha = 0$. Then, as follows from (20), $b_\alpha > 0$ in these points.

In the second case, there exist points C and D such that $v_\rho = \dot\rho = 0$. At the point D, $Dv_\alpha = -\omega R_L$, where R_L is the circle radius, and $v_\alpha = \dot\alpha|OD|$, where $|OD| < 2R_L$ is the distance of the point D to the origin. In this case, $\dot\alpha < -\omega/2$ and $b_\alpha < 0$.

We will consider the case where the circle encloses the origin more carefully. In this case, $I_\alpha = -b_\alpha$, since $b_\alpha = \pi_\alpha = \text{const}$ and the coordinate α changes by 2π over the period. Then, the Hamiltonian function has the form

$$H = E = \omega(I_\rho + I_\alpha) + \frac{b_z^2}{2m}. \tag{21}$$

The rates of variation of ρ and α coincide (degeneracy), since I_ρ and I_α add up. It is common practice in this case to introduce the momenta $I = I_\rho + I_\alpha$ and $P = I_\rho$ instead of the nonexistent momenta I_ρ and I_α. Then, the Hamiltonian depends only on I and b_z:

$$H = \omega I + \frac{b_z^2}{2m}. \tag{22}$$

The canonically conjugate variable Ψ that corresponds to the action I is a linear function of time: $\Psi = \omega t + \Psi_0$. The canonically conjugate variable $Q = \alpha_L$ corresponding to the new momentum P is a constant and has the meaning of the angular coordinate of the Larmor circle center. The momentum P is given by $P = m\omega\rho_L^2/2$, where ρ_L is the coordinate of the Larmor circle center.

If the circular projection of the particle trajectory onto the (ρ, α) plane does not enclose the origin, we have, instead of (21),

$$H = \omega I_\rho + \frac{b_z^2}{2m}. \tag{23}$$

Considering $I = I_\rho$, $P = I + b_\alpha$, and b_z as new momenta, one arrives at the same expression for Hamiltonian function (22). However, one should take into account that other variables are used. Using the Maslov–Leray quantization rules [2, 3], we obtain

$$\begin{aligned} I_\rho &= \hbar\left(n_\rho + \frac{1}{2}\right), \quad n_\rho = 0, 1, \ldots, \\ I_\alpha &= n_\alpha\hbar, \quad n_\alpha = 0, \pm1, \pm2, \ldots \end{aligned} \tag{24}$$

and using (19),

$$E = \frac{eB\hbar}{mc}\left(n_\rho + \frac{1}{2} + \frac{|n_\alpha| - n_\alpha}{2}\right) + \frac{b_z^2}{2m}. \tag{25}$$

The latter exactly coincides with the Landau spectrum [8] resulting from the Pauli Hamiltonian.

The coincidence of the Maslov quasi-classical quantization rules [3] with the exact quantum results was first mentioned by J. Leray [2]. Maslov himself did not point clearly to such a way of applying his quantization method. He studied only the case of infinitely

large quantum numbers, i.e., the "correspondence principle" of quantum mechanics.

REFERENCES

1. S. A. Goudsmit, Phys. Today **29**, 40 (1976).
2. J. Leray, *Analyse lagrangienne et mecanique quantique* (Strasbourg, 1978).
3. M. V. Karasev and V. P. Maslov, *Nonlinear Poisson Brackets* (Nauka, Moscow, 1991).
4. V. I. Arnold, *Mathematical Methods of Classical Mechanics* (Nauka, Moscow, 1979; Springer-Verlag, New York, 1989).
5. Ya. I. Frenkel', *Electrodynamics* (ONTI, Moscow, 1934).
6. V. V. Kozlov, *Symmetry, Topology, and Resonance in Hamiltonian Mechanics* (Izhevsk, 1995).
7. L. D. Landau and E. M. Lifshitz, *The Classical Theory of Fields* (Nauka, Moscow, 1967; Pergamon, Oxford, 1975).
8. L. D. Landau and E. M. Lifshitz, *The Course of Theoretical Physics*, Vol. 3: *Quantum Mechanics: Non-Relativistic Theory* (Nauka, Moscow, 1967; Pergamon, New York, 1987).

Translated by M. Fofanov

Technical Physics, Vol. 45, No. 9, 2000, pp. 1115–1121. Translated from Zhurnal Tekhnicheskoĭ Fiziki, Vol. 70, No. 9, 2000, pp. 17–23.

Stripping of Fast Oxygen Ions Colliding with Atoms of Light Elements

A. V. Bakaldin*, S. A. Voronov*, S. V. Koldashov*, and V. P. Shevel'ko**

** Moscow State Engineering Physics Institute (Technical University), Kashirskoe sh. 31, Moscow, 115409 Russia*

*** Lebedev Institute of Physics, Russian Academy of Sciences, Leninskiĭ pr. 53, Moscow, 117924 Russia*

E-mail: shev@sci.lebedev.ru

Received January 28, 1999

Abstract—Based on the available experimental data and the results of calculations carried out in this paper, cross sections are estimated for single-electron stripping of O^{Z+} (Z = 1–7) ions with energies in the range E = 0.5–200 MeV/u due to collisions with H, He, N, O, or Ar atoms. Analytical approximations of the cross sections are presented. *© 2000 MAIK "Nauka/Interperiodica".*

INTRODUCTION

Processes in which positive ions colliding with atoms or molecules change their charge (ionization or charge transfer) play an important role in many problems of atomic-collision physics [1–3], plasma physics [4], developing methods for accumulating heavy ions in accelerators [5], the physics of cosmic rays and the earth's magnetosphere [6], etc. For example, such processes occur in the interaction of ions from an anomalous component of cosmic rays (ACCR) (H, He, C, N, O, Ne, and Ar ions) with ions and molecules of the earth's upper atmosphere and determine the formation and dynamics of the nuclear component of the earth's radiation belt. In spite of extensive experimental and theoretical studies of the mechanisms for capturing ACCR ions by the geomagnetic field and the dynamics of ion fluxes in the earth's radiation belt, these processes have not yet been investigated quantitatively. Information on the charge-state distribution of captured ions, the spatial distribution of ion fluxes along the geomagnetic field lines, etc. is still lacking. This information can be obtained by numerically simulating the propagation of ACCR ions in the earth's magnetic field, subsequent capture of these ions by a geomagnetic trap, and their motion in the radiation belt. In order to perform such calculations, one needs to know the cross sections for stripping and charge transfer of ACCR ions with energies in the range of 1–200 MeV/u in their interaction with atoms and molecules of the upper atmosphere.

In this paper, we determine the cross sections for stripping of oxygen ions (the most abundant ions in the ACCR) with energies in the range E = 1–200 MeV/u due to collisions with atoms of light elements:

$$O^{Z+} + A \longrightarrow O^{(Z+1)+} + A + e^-; \quad Z = 1\text{–}7; \quad A = \text{H, He, N, O, Ar}. \tag{1}$$

The stripping cross sections (1) are calculated numerically in the Bohr–Born approximation using the Atom code. On the basis of the calculated cross sections, available experimental data, and scalings for the stripping cross sections, we estimate the stripping cross sections for processes (1) at energies in the range E = 0.5–200 MeV/u, accurate to a factor of 2. These cross sections can be used to numerically simulate the capture of ACCR ions by the earth's magnetosphere; the results of the simulations will be published in a separate paper.

BASIC PROCESSES RESULTING IN A CHANGE IN THE ION CHARGE STATE: CHARGE TRANSFER AND IONIZATION

Charge transfer and ionization are the basic processes resulting in a change in the charge state of an incident ion colliding with an atom or molecule. Such processes were considered in [1–8] for a wide energy range; a bibliography on the effective cross sections (up to 1997) is presented in [9].

The charge of an ion colliding with a neutral atom or molecule can either increase due to *stripping* (ionization),

$$X^{Z+} + A \longrightarrow X^{(Z+a)+} + \ldots, \quad a \geq 1, \tag{2}$$

or decrease due to *charge transfer* (the capture of the target electrons),

$$X^{Z+} + A \longrightarrow X^{(Z-b)+} + \ldots, \quad b \geq 1. \tag{3}$$

Here, X^{Z+} is the incident ion and A is the target atom or molecule. The dots in processes (2) and (3) mean that processes resulting in the excitation and ionization of the target A can also occur. The total cross section for the change in the charge of the incident ion is deter-

1063-7842/00/4509-1115$20.00

mined by the sum of the cross sections for the single-electron and multielectron ionization (ion) and electron capture (ec) processes, respectively:

$$\sigma_{\text{tot}} = \sum_{a \geq 1} \sigma_{\text{ion}}^{(a)} + \sum_{b \geq 1} \sigma_{\text{ec}}^{(b)}. \quad (4)$$

As a rule, single-electron processes ($a = b = 1$) make the major contribution to sum (4). Nevertheless, the contribution from multielectron processes to the total cross section may reach 30–50% under certain conditions (see, e.g., [10, 11]).

In general, the cross sections for ionization and charge transfer of an incident ion have different dependences on such atomic parameters as the relative velocity v, the charge of the incident ion Z, and the charge of the target nucleus Z_T. Therefore, the relative contribution from each process is determined by the energies and charges of the colliding particles. For high collision energies $v/Z > 1$ au (1 au of velocity = 2.2×10^8 cm/s), the cross section for the ionization of an incident ion colliding with a neutral atom has the asymptotic form

$$\sigma_{\text{ion}} \sim Z_T^2/(Z_P^4 v^2), \quad (5)$$

and the charge-transfer cross section is determined by the sum of the cross sections for radiative electron capture (REC) and nonradiative capture (NRC),

$$\sigma_{\text{ec}} = \sigma_{\text{NRC}} + \sigma_{\text{REC}} \sim C_1 Z_T^5 Z_P^5/v^{11} + C_2 Z_T Z_P^5/v^{5/2}, \quad (6)$$

where C_1 and C_2 are constants. The charge-transfer processes (3) are of great importance when a bare nucleus is used as an incident ion and the capture of the target electrons is a single process that changes the ion charge. The s_{NRC} cross sections are generally calculated in the Brinckman–Cramers approximation, or its modifications [2], or approximated by semiempirical formulas [12], whereas the s_{REC} cross sections are calculated by the Stobbe formula [13–15]. The processes of radiative and nonradiative charge transfer and their relative contribution to the total cross section are considered, e.g., in [15]. For fast oxygen ions with energies E = 1–200 MeV/u, the charge-transfer cross sections (3) are small compared to the stripping cross sections (2). Therefore, the former are not considered in this paper.

METHODS FOR CALCULATING THE STRIPPING CROSS SECTIONS AT HIGH COLLISION ENERGIES

At high collision energies, the Born approximation [16, 17] is used to calculate the cross section for the ionization of incident ions (stripping cross section). In this method, the scattering amplitude is given by the product of the form factors of the incident ion and target atom in the representation of the transferred momentum (q-representation):

$$\sigma_{\text{ion}} = \frac{8\pi a_0^2}{v^2} \sum_P \int_0^\infty d\varepsilon \sum_{i(T)} \int_{q_{\min}}^\infty \frac{dq}{q^3} |F_P(q)|^2 |F_T(q)|^2, \quad (7)$$

where a_0 is the Bohr radius, q is the transferred momentum, $F_{P,T}(q)$ are the form factors of the incident ion and target atom, and ε is the energy of the knocked-out electron.

The sum over P denotes summation over all electron shells of the incident ion; and the sum over $i(T)$ denotes summation over all the final target states i, including excitation and ionization. The minimum transferred momentum is defined by the expression

$$q_{\min} = (I_P + \varepsilon + \Delta E_T)/v, \quad (8)$$

where I_P is the binding energy of an escaped electron and ΔE is the excitation electron energy in the target atom. The form factors $F(q)$ have the form

$$|F_P(q)|^2 = |\langle \varepsilon | e^{i\mathbf{qr}} | 0 \rangle_P|^2, \quad (9)$$

$$|F_T(q)|^2 = \left| Z_T \delta_{i0} - \sum_{j=1}^{N_r} \langle i | e^{i\mathbf{qr}} | 0 \rangle_j \right|^2, \quad (10)$$

where $|0\rangle_P$ and $|0\rangle_j$ denote the electron wave functions in the incident ion and the target atom, respectively, and N_T is the total number of target electrons.

For neutral atoms, we have $N_T = Z_T$.

The form factor $|F_T(q)|^2 = Z_T^2$ for bare nuclei and the form factor for targets with 1s electrons (H or He) are [17]

$$F_{1s,1s}(q) = Z_T - N_T \frac{1}{[1 + (q/2d)^2]^2}, \quad (11)$$

where d is a constant determining the degree of screening of a nucleus by 1s electrons.

For hydrogen atoms, we have $N_T = Z_T = 1$ and $d = 1$, and for helium atoms, we have $N_T = Z_T = 2$ and $d = 1.69$. The form factors $F_T(q)$ for other atomic shells are tabulated in [18].

The calculation of the cross sections for stripping an incident ion by formula (7) is rather complicated even for single-electron targets because of the intricate dependence of the lower limit of integration in (7) on the energy of the knocked-out electron and the excitation energy of the target-atom electron. Formulas (7)–(10) are substantially reduced when the kinetic energy of the incident ion is much higher than the excitation energy of the target atom, i.e., $v^2 \gg \Delta E$. In this case, one can take

$$q_{\min} \approx (I_P + \varepsilon)/v, \quad (12)$$

i.e., the target electrons may be treated as motionless. Under condition (12), formulas (7)–(10) are simplified and the stripping cross section takes the form [19, 20]

$$\sigma_{\text{ion}} = \frac{8\pi a_0^2}{v^2}\sum_P \int_0^\infty d\varepsilon \sum_T \int_{q_{\min}}^\infty \frac{dq}{q^3}|F_P(q)|^2 S_T(q), \quad (13)$$

$$S_T(q) = \left[Z_T - \sum_j^{N_T} F_{jT}(q)\right]^2 + \left[N_T - \sum_j^{N_T}|F_{jT}(q)|^2\right], \quad (14)$$

$$F_{jT}(q) = \langle j|e^{i\mathbf{qr}}|j\rangle, \quad q_{\min} = (I_P + \varepsilon)/v. \quad (15)$$

From the physical point of view, the two terms in sum (14) correspond to ionization by the screened target nucleus Z_T and screened electrons N_T. For $S_T(q) \equiv 1$, Eqs. (13) and (15) correspond to the Born approximation for ionization of an atom or ion by electron (or proton) impact (see, e.g., [16]).

Interaction between the electrons of the incident ion and the target electrons may increase or decrease the total stripping cross section, depending on the transferred momentum q. As $q \longrightarrow 0$ (large impact parameters), the form factor is $F_{jT} \longrightarrow 1$, whereas for $q \longrightarrow \infty$ (small impact parameters), we have $F_{jT} \longrightarrow 0$ [cf. (11)]. Accordingly, the effective screening is minimum ($S_T(q) \longrightarrow 0$) as $q \longrightarrow 0$ and is maximum for large transfers of momentum q:

$$S_{\max}(q) \longrightarrow Z_T^2 + N_T, \quad q \longrightarrow \infty. \quad (16)$$

Therefore, the maximum cross section for stripping of an incident ion by an atom is approximately proportional to $Z_T^2 + N_T$ for high collision energies, where Z_T^2 corresponds to the contribution from the interaction with the nucleus, whereas N_T stems from the interaction with the target electrons. Thus, at high collision energies, the target electrons do not screen the nucleus and even make an additional contribution to the stripping cross section that is proportional to N_T. The heavier the target atom, the larger the cross section for stripping the incident ion. For example, $S_{\max}(\text{H}) = 2$ for ionization by hydrogen atoms, whereas $S_{\max}(\text{Ar}) = 342$ for argon atoms. For intermediate and low collision energies, scaling law (16) is no longer valid due to considerable screening effects and the stripping cross sections are scaled according to the semiempirical law $\sigma_{\text{ion}} \sim Z_T^{3/2}$, which was obtained in [7].

Note that, for $S_T(q) = Z_T^2 + N_T$, the stripping cross section (13) can be represented as a linear combination of cross sections $\sigma_e(v)$ and $\sigma_p(v)$ for ionization by electron and proton impacts, respectively:

$$\sigma_{\text{ion}}(v) = Z_T^2\sigma_p(v) + N_T\sigma_e(v), \quad \sigma_p(v) \approx \sigma_e(v), \quad v \gg I^{1/2}, \quad (17)$$

where I is the binding energy of the incident-ion electron.

Equation (17) corresponds to the Bohr approximation [21], where the velocity of the incident ion is so high that the nucleus and electrons of the target can be considered to be immobile [cf. (12)]. The Bohr approximation (17) is applicable at sufficiently high collision energies, when the velocity of an incident particle exceeds the velocity of the K-electron of a target atom.

Another (ad hoc) approximation was proposed in [20]. In this approximation, the dependence of the disturbing charge $S_T(q)$ on the transferred momentum q and the threshold dependence on the cross section for electron-impact ionization $\sigma_e(v)$ is presented in the form

$$S_T(q) = \left[Z_T - \sum_j^{N_T} F_{jT}(q)\right]^2 + \left[N_T - \sum_j^{N_T}|F_{jT}(q)|^2(\sigma_e(v)/\sigma_p(v))\right]. \quad (18)$$

Thus, the calculation of the cross section for stripping an incident ion colliding with a structural particle is a rather laborious problem even in the Born approximation.

CALCULATION OF THE EFFECTIVE CROSS SECTIONS FOR STRIPPING OF OXYGEN IONS BY LIGHT ATOMS

Experimental data on the cross sections for stripping of fast ions by atomic and molecular targets are reported in a number of papers (see, e.g., [22–34]), and the results of numerical calculations are presented in [29, 31, 35, 36]. It is worth noting that experimental data on stripping of light ions by light atoms ($Z_T < 10$) for energies $E > 5$ MeV/u are still lacking. However, it is the energy range $E = 0.5$–30 MeV/u that is of practical interest for various applications, because the stripping cross sections are maximum in this case. The cross sections for stripping of oxygen ions by gas targets (H_2, He, N_2, or Ar) were measured only for the energy range $0.1 < E < 5$ MeV/u [22, 27, 28, 31], and the calculations of the cross sections for $O^{7+} + H_2$ and O^{7+} + He collisions in the same energy range were performed in [29, 31].

In this paper, we present the cross sections for stripping of oxygen ions by light atoms for $E = 1$–200 MeV/u. For this purpose, we used the experimental data for $E < 5$ MeV/u and performed numerical calculations for higher energies. The cross sections were calculated in the Bohr–Born approximation (17), in which the experimental cross sections $\sigma_e(v)$ for electron-impact ionization of oxygen ions were taken from [37] and the cross sections $\sigma_p(v)$ for proton-impact ioniza-

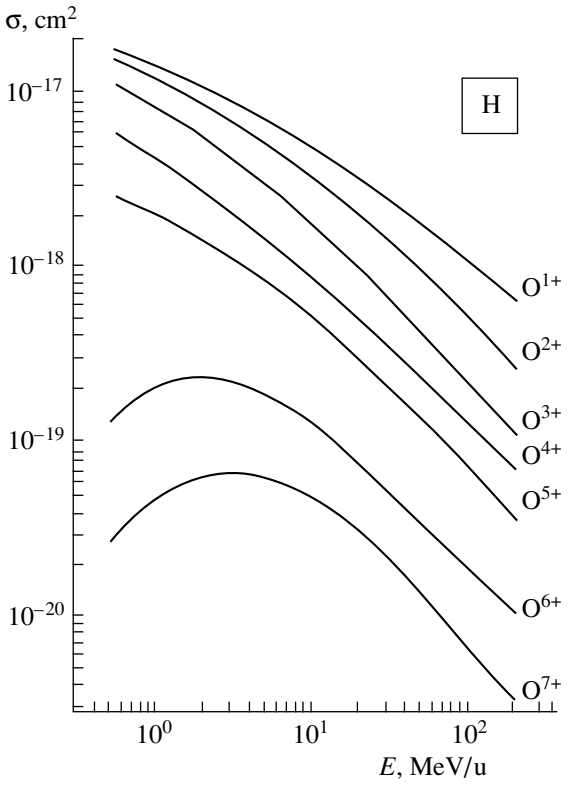


Fig. 1. Cross sections for stripping of O^{Z+} ions by H atoms.

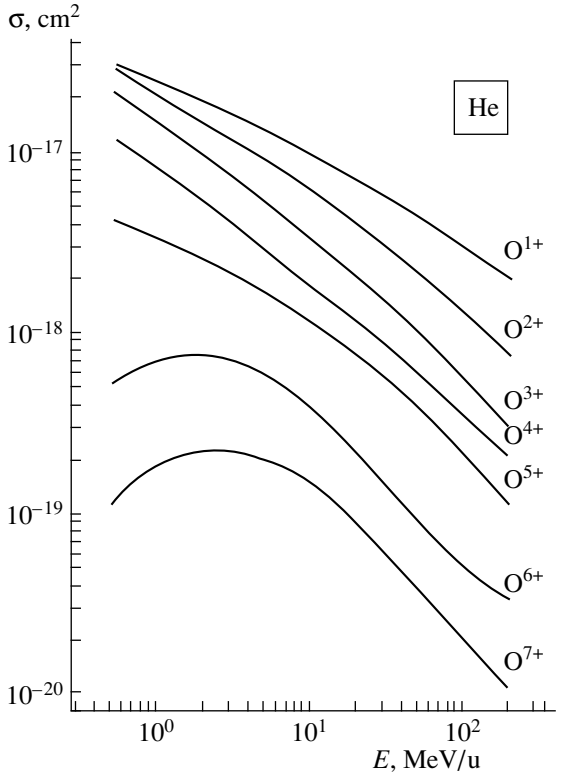


Fig. 2. Cross sections for stripping of O^{Z+} ions by He atoms.

tion were calculated in the Born approximation by the Atom code (see [38]). The total cross sections for ionization by protons were calculated taking into account ionization involving electrons from all the shells of an oxygen ion and the contribution from the processes of excitation of inner electrons to autoionization states. The wave functions of the outer and inner electrons of an oxygen ion in both discrete and continuous spectra were calculated by numerically solving the radial Shrödinger equation with the effective potential of the atomic residual [38]. The total cross sections $\sigma_p(v)$ for proton-impact ionization of oxygen ions for E = 1–200 MeV/u are listed in Table 1.

The calculated cross sections for stripping of oxygen ions (17) by atoms of light elements in the energy range $E < 5$ MeV/u exceeds the experimental data by a factor of 5. This is not surprising, because the Born approximation is no longer valid in this energy range. Therefore, we used the experimental data [22, 27–31] to estimate the cross sections for low energies, whereas for high energies, the stripping cross sections were calculated by formula (17). In the intermediate energy range and in cases where experimental data were lacking (e.g., for O^{Z+} + O collisions), the cross sections were estimated by the semiempirical scaling $Z_T^{3/2}$ proposed in [7] for the intermediate energy range. The

Table 1. Total cross sections (cm^2) for ionization of oxygen ions by protons as a function of proton energy

E, MeV	O^{1+}	O^{2+}	O^{3+}	O^{4+}	O^{5+}	O^{6+}	O^{7+}
1.56	1.8×10^{-17}	9.5×10^{-18}	4.5×10^{-18}	2.3×10^{-18}	1.1×10^{-18}	1.5×10^{-19}	5.5×10^{-20}
3.12	1.0×10^{-17}	5.0×10^{-18}	2.6×10^{-18}	1.3×10^{-18}	6.9×10^{-19}	1.3×10^{-19}	4.7×10^{-20}
6.24	5.9×10^{-18}	2.8×10^{-18}	1.5×10^{-18}	7.8×10^{-19}	4.1×10^{-19}	8.4×10^{-20}	3.2×10^{-20}
12.5	3.2×10^{-18}	1.6×10^{-18}	8.0×10^{-19}	4.5×10^{-19}	2.3×10^{-19}	5.1×10^{-20}	1.9×10^{-20}
25.	1.8×10^{-18}	8.6×10^{-19}	4.4×10^{-19}	2.5×10^{-19}	1.3×10^{-19}	2.9×10^{-20}	1.1×10^{-20}
50.	9.5×10^{-19}	4.7×10^{-19}	2.3×10^{-19}	1.3×10^{-19}	7.4×10^{-20}	1.6×10^{-20}	6.2×10^{-21}
100.	5.2×10^{-19}	2.5×10^{-19}	1.3×10^{-19}	7.8×10^{-20}	3.8×10^{-20}	8.9×10^{-21}	3.4×10^{-21}
200.	2.8×10^{-19}	1.3×10^{-19}	6.9×10^{-19}	3.8×10^{-20}	2.1×10^{-20}	4.8×10^{-21}	1.8×10^{-21}

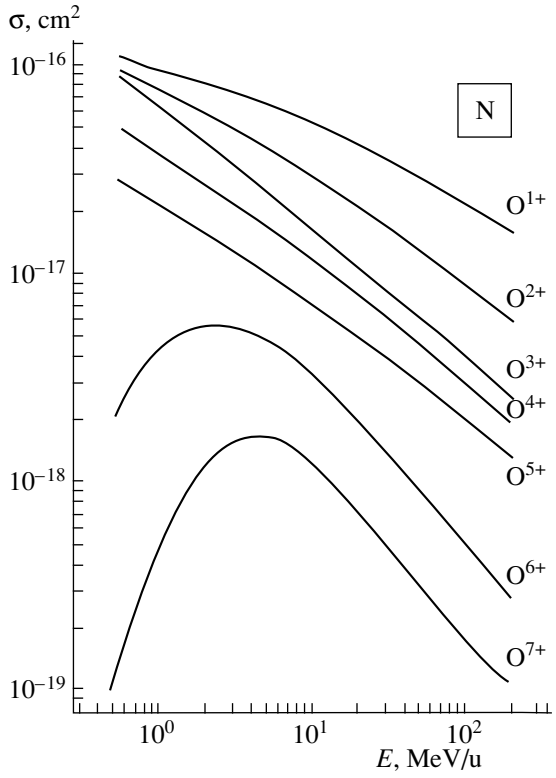


Fig. 3. Cross sections for stripping of O^{Z+} ions by N atoms.

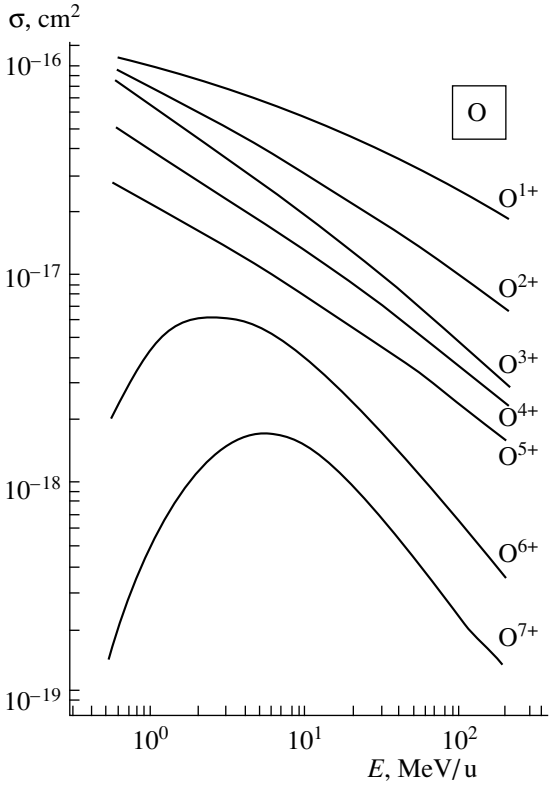


Fig. 4. Cross sections for stripping of O^{Z+} ions by O atoms.

cross sections for stripping the oxygen ions colliding with H, He, N, O, or Ar atoms are presented in Figs. 1–5 for the energy range E = 0.5–200 MeV/u. The estimated error in the cross sections is about a factor of 2.

Similar estimations of the cross sections for stripping of O^{Z+} ions (Z = 1–7) by H, N, or Ar atoms in the energy range E = 0.1–200 MeV/u were obtained in [39]. In this case, the cross sections at high energies were normalized using asymptotic semiempirical formulas [40], which were obtained for the K-electron ionization of H- and He-like ions. In general, the stripping cross sections obtained in [39] for ion energies $E >$ 50 MeV/u are from three to five times greater than those calculated in this paper. Apparently, this stems from the fact that, in this paper, the stripping cross sections are calculated taking into account the shell structure of oxygen ions, i.e., ionization of all the electrons of an incident ion, including direct ionization and the contribution from the excitation of inner electrons to autoionization states.

To numerically simulate the dynamics of the elementary processes occurring in laboratory and astrophysical plasmas, self-contained analytical formulas for the cross sections are preferable. In this paper, the effective cross sections for single-collision stripping of oxygen ions in gases for energies of 0.5–200 MeV/u

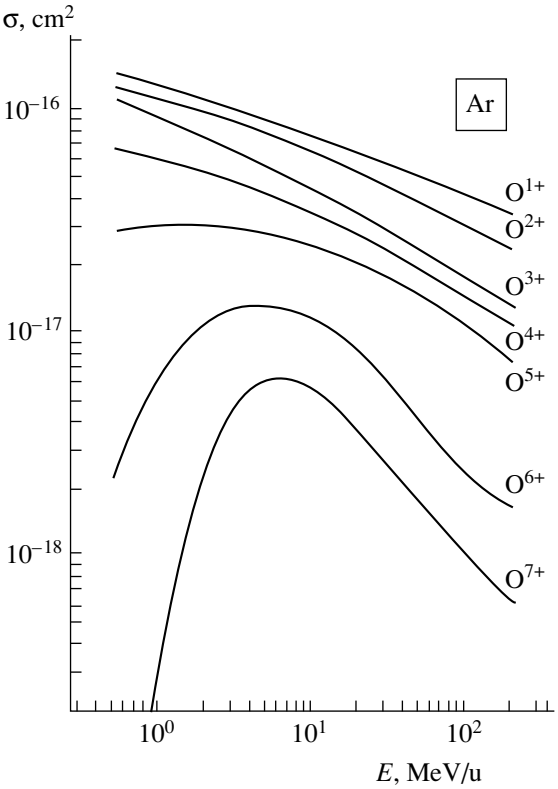


Fig. 5. Cross sections for stripping of O^{Z+} ions by Ar atoms.

Table 2. Approximation parameters of cross sections for stripping of oxygen ions (19) by hydrogen atoms (O^{Z+} + H) for energies E = 0.5–200 MeV/u

Ion	a_1	a_2	a_3	a_4	a_5
O^{1+}	–16.88	–0.4012	–0.07731	0	0
O^{2+}	–16.94	–0.4730	–0.1080	0	0
O^{3+}	–17.11	–0.5510	–0.1138	0	0
O^{4+}	–17.40	–0.6154	–0.06562	0	0
O^{5+}	–17.72	–0.4767	–0.1184	0	0
O^{6+}	–18.69	0.3595	–0.7577	0.1542	0
O^{7+}	–19.32	0.6130	–0.7524	0.1146	0

Table 3. Same as in Table 2 for stripping of O^{Z+} ions by He atoms

Ion	a_1	a_2	a_3	a_4	a_5
O^{1+}	–16.63	–0.3736	–0.04589	0	0
O^{2+}	–16.70	–0.5054	–0.05561	0	0
O^{3+}	–16.85	–0.6032	–0.05757	0	0
O^{4+}	–17.12	–0.6255	–0.02963	0	0
O^{5+}	–17.48	–0.3852	–0.1135	0	0
O^{6+}	–18.12	0.3076	–0.8090	0.1814	0
O^{7+}	–18.73	0.4927	–0.7582	0.1336	0

Table 4. Same as in Table 2 for stripping of O^{Z+} ions by N atoms

Ion	a_1	a_2	a_3	a_4	a_5
O^{1+}	–16.02	–0.2178	–0.0561	0	0
O^{2+}	–16.13	–0.3785	–0.0462	0	0
O^{3+}	–16.22	–0.5522	–0.02178	0	0
O^{4+}	–16.44	–0.4742	–0.04067	0	0
O^{5+}	–16.68	–0.4403	–0.03870	0	0
O^{6+}	–17.36	0.6971	–1.250	0.4670	–0.0670
O^{7+}	–18.30	1.723	–1.673	0.3476	0

Table 5. Same as in Table 2 for stripping of O^{Z+} ions by O atoms

Ion	a_1	a_2	a_3	a_4	a_5
O^{1+}	–15.99	–0.2146	–0.04480	0	0
O^{2+}	–16.10	–0.3791	–0.03430	0	0
O^{3+}	–16.20	–0.4777	–0.04310	0	0
O^{4+}	–16.41	–0.4526	–0.03131	0	0
O^{5+}	–16.66	–0.4057	–0.03462	0	0
O^{6+}	–17.30	0.7231	–1.282	0.5093	–0.07984
O^{7+}	–18.24	1.557	–1.409	0.2659	0

Table 6. Same as in Table 2 for stripping of O^{Z+} ions by Ar atoms

Ion	a_1	a_2	a_3	a_4	a_5
O^{1+}	–15.91	–0.2102	–0.01757	0	0
O^{2+}	–15.96	–0.2132	–0.03263	0	0
O^{3+}	–16.05	–0.2849	–0.03587	0	0
O^{4+}	–16.23	–0.1885	–0.05909	0	0
O^{5+}	–16.52	0.04542	–0.1375	0	0
O^{6+}	–17.21	1.085	–1.032	0.1971	0
O^{7+}	–18.54	4.342	–4.627	1.824	–0.2635

were approximated by the expression

$$\sigma(E) = 10^{a_1 + a_2\log(E) + a_3\log^2(E) + a_4\log^3(E) + a_5\log^4(E)}, \quad (19)$$

where $\sigma(E)$ is the stripping cross section in cm^2/atom, E is the kinetic energy of oxygen ions in MeV/u, and a_i are the approximation coefficients.

The approximation coefficients a_i obtained for stripping of oxygen ions by light elements are listed in Tables 2–6.

CONCLUSION

In this paper, based on the available experimental data, numerical calculations, and scalings, we have obtained the cross sections for the single-collision stripping of O^{Z+} oxygen ions (Z = 1–7) colliding with H, He, O, Ne, or Ar atoms in the energy range E = 0.5–200 MeV/u. These data may be used for a number of physical applications, e.g., for numerical simulations of ion capture from the ACCR by a geomagnetic trap and the determination of the lifetime of oxygen-ion beams

in accelerators in which the atoms of light elements are the basic components of the residual gas in a vacuum system.

ACKNOWLEDGMENTS

This work was supported in part by the Russian Foundation for Basic Research, project nos. 96-02-16090 and 97-02-16919.

REFERENCES

1. H.-D. Betz, Rev. Mod. Phys. **44**, 465 (1972).
2. L. P. Presnyakov, V. P. Shevel'ko, and R. K. Yanev, *Elementary Processes with the Participation of Multiply Charged Ions* (Énergoatomizdat, Moscow, 1986).
3. C. L. Cocke and R. E. Olson, Phys. Rep. **205**, 153 (1991).
4. *Atomic and Molecular Processes in Fusion Edge Plasmas*, Ed. by R. R. Janev (Plenum, New York, 1995).
5. D. Habs, Nucl. Instrum. Methods Phys. Res. B **43**, 390 (1989).
6. W. N. Spjeladvik, Space Sci. Rev. **23**, 499 (1979).
7. V. S. Nikolaev, Usp. Fiz. Nauk **85** (4), 6 (1965).
8. P. H. Mokler and Th. Stöhlker, Adv. At. Mol. Opt. Phys. **37**, 297 (1996).
9. H. Tawara, Report NIFS-DATA-42/Nagoya (Japan, 1997).
10. J. H. McGuire, Adv. At. Mol. Opt. Phys. **29**, 217 (1992).
11. V. P. Shevelko and H. Tawara, in *Atomic Multielectron Processes* (Springer-Verlag, Berlin, 1998).
12. A. S. Schlachter, J. W. Stearns, W. G. Graham, *et al.*, Phys. Rev. A **27**, 3372 (1983).
13. M. Stobbe, Ann. Phys. (Leipzig) **7**, 661 (1930).
14. Th. Stöhlker, C. Kozhuharov, P. H. Mokler, *et al.*, Phys. Rev. A **51**, 2098 (1995).
15. J. Eichler and Th. Stöhlker, in *Atomic Physics with Heavy Ions* (Springer-Verlag, Berlin, 1999), p. 251.
16. D. R. Bates, *Atomic and Molecular Processes* (Academic, New York, 1962; Mir, Moscow, 1964).
17. J. S. Briggs and K. Taulbjerg, in *Topics in Current Physics*, Vol. 5: *Structure and Collisions of Ions and Atoms* (Springer-Verlag, Berlin, 1978).
18. J. H. Hubbel, W. J. Veigele, E. A. Briggs, *et al.*, J. Phys. Chem. Ref. Data **4**, 471 (1975).
19. J. H. McGuire, N. Stolterfoht, and P. R. Simony, Phys. Rev. A **24**, 97 (1981).
20. R. Anholt, Phys. Lett. A **114**, 126 (1988).
21. N. Bohr, K. Dan. Vidensk. Selsk. Mat.-Fys. Medd. **1** (8), Sect. 3.3, 4.2 (1948).
22. J. R. Macdonald and F. W. Martin, Phys. Rev. A **4**, 1965 (1971); **4**, 1974 (1971).
23. W. Erb, Report GSI-P-7-78 (Darmstadt, Germany, 1978).
24. B. Franzke, IEEE Trans. Nucl. Sci. **NS-28**, 2116 (1981).
25. W. G. Graham, K. H. Berkner, E. V. Bernstein, *et al.*, J. Phys. B **18**, 2503 (1985).
26. R. Hippler, S. Datz, P. D. Miller, *et al.*, Phys. Rev. A **35**, 585 (1987).
27. R. Anholt, X.-Y. Xu, Ch. Stoller, *et al.*, Phys. Rev. A **37**, 1105 (1988).
28. T. N. Tipping, J. M. Sanders, J. Hall, *et al.*, Phys. Rev. A **37**, 2906 (1988).
29. H.-P. Hülskötter, W. E. Meyerhof, E. Dillard, *et al.*, Phys. Rev. Lett. **63**, 1938 (1989).
30. S. A. Boman, E. M. Bernstein, and J. A. Tanis, Phys. Rev. A **39**, 4423 (1989).
31. H.-P. Hülskötter, B. Feinberg, W. E. Meyerhof, and A. Belkacem, Phys. Rev. A **44**, 1712 (1991).
32. P. Rymuza, Th. Stöhlker, and C. L. Cocke, J. Phys. B **26**, L169 (1993).
33. O. Heber, G. Sampoll, and B. B. Bandong, Phys. Rev. A **52**, 4578 (1995).
34. W. Wu, K. L. Wong, E. C. Montenegro, *et al.*, Phys. Rev. A **55**, 2771 (1997).
35. V. P. Shevelko, D. Böhne, and Th. Stöhlker, Nucl. Instrum. Methods Phys. Res. A **415**, 609 (1998).
36. V. P. Shevelko, O. Brinzanescu, W. Jacoby, *et al.*, Hyperfine Interact. **114**, 289 (1998).
37. T. Kato and H. Tawara, At. Data Nucl. Data Tables **36**, 167 (1987).
38. L. A. Vaĭnshteĭn and V. P. Shevel'ko, in *Structure and Characteristics of Ions in a Hot Plasma* (Nauka, Moscow, 1986); Preprint No. 43, FIAN (Lebedev Institute of Physics, Russian Academy of Sciences, Moscow, 1996).
39. I. S. Dmitriev, Vestn. Mosk. Univ., Ser. 3: Fiz., Astron., No. 2, 22 (1999).
40. I. S. Dmitriev, V. P. Zaikov, and Yu. A. Tashaev, Nucl. Instrum. Methods **164**, 329 (1979).

Translated by M. Fofanov

Technical Physics, Vol. 45, No. 9, 2000, pp. 1122–1127. Translated from Zhurnal Tekhnicheskoĭ Fiziki, Vol. 70, No. 9, 2000, pp. 24–29.

GASES AND LIQUIDS

A Perfect Centrifuge

O. E. Aleksandrov

Ural State Technical University, ul. Mira 19, Yekaterinburg, 620002 Russia

E-mail: aleks@dpt.ustu.ru

Received April 13, 1999; in final form, September 27, 1999

Abstract—The theory of separation in a Zippe gas centrifuge based on the radial averaging method has been developed. A more general way of deriving the one-dimensional diffusion equation is proposed. The applicability of the averaging method for an arbitrary distribution of the circulation flow along the centrifuge axis has been demonstrated. The general solution has been obtained, and the optimal internal circulation, that is, a perfect centrifuge, has been defined. The maximum separative power of the perfect centrifuge is expressed as a function of the external parameters. © *2000 MAIK "Nauka/Interperiodica".*

To design new effective gas centrifuges (GCs) for isotope separation, it is necessary to have an accurate evaluation of the reserves of the centrifuge output. At present, the main efforts to describe GCs are directed at developing calculation and simulation programs (see, for example [1, 5]), although the GC efficiency limit remains unknown. The GC efficiency is usually expressed in units of the separative work done in a unit time, that is, in units of the separative power of a GC:

$$\delta U = PV(N_P) + WV(N_W) - FV(N_F), \qquad (1)$$

where δU is the separative power of a GC; F, P, and W are the feed, product, and waste flows in a GC, respectively; N_F, N_P, and N_W are the concentration in the flows of feed, product, and waste of a GC, respectively; and V is the value function

$$V(N) = (2N-1)\ln\frac{N}{1-N}.$$

The separative power of GCs is governed by many factors. One important factor is an internal gas flow in the separation chamber of the GC rotor (Fig. 1). The limiting estimate of the separative power of a GC according to Dirac [6] is used rather widely:

$$\delta U_{\max} = \frac{\pi}{2} H\rho D\left(\frac{\Delta m(R\Omega)^2}{2kT}\right)^2, \qquad (2)$$

where H is the GC rotor length, D is the diffusion coefficient, ρ is the density of the gas being separated, Δm is the difference between the molecular masses of the isotopes ($\Delta m = |m_1 - m_2|$)), R is the inner radius of the GC rotor, Ω is the angular rotation speed of the GC rotor, k is Boltzmann's constant, and T is the gas temperature.

It follows from (2) that the separative power of modern centrifuges is several times less than the maximum possible one. Attempts to achieve a GC efficiency comparable with the estimate in (2) have so far turned out to be futile. There are also other disagreements between experiment and the results of (2). For example, the experimental dependence of δU on the GC rotor length is nonlinear. On the other hand, modern centrifuges demonstrate an efficiency which is 10–20% higher than predicted by the more detailed analytical Cohen model. This gives cause to consider this model unusable.

The estimate in (2) has one more major drawback; namely, it does not give any connection between the separative power of a GC and the external (feed, product, and waste) and internal (circulation) flows. The latter flow makes this estimate applicable only to a limiting operation regime of a GC, namely, that of infinite feed and circulation flows in the rotor. Centrifuges are unable to operate in such a regime, one of the reasons being that the separation factor in this limiting regime tends to zero. Working parameters of the flows in GCs are chosen as a trade-off between the requirements of stability of the circulation in the rotor and a finite value of the separation factor. Thus, the problem of defining a set of parameters that ensure maximum efficiency of GCs is of great importance.

In this article, we attempt a new approach to an analytical description of the separation process in a GC using the radial averaging method in order to (1) prove its applicability to an arbitrary distribution of the circulation flow in the GC rotor and (2) find an optimal relation between the internal circulation in the rotor and the external parameters of a GC, that is, to design a perfect centrifuge (PC).

In Cohen's method [7] the problem of two-dimensional distribution of the concentration in the rotor is reduced to a one-dimensional problem of the concentration distribution along the GC axis. An advantage of the method is its simplicity and the possibility of deriving analytical relations between the gas flow in the rotor and the GC efficiency.

In Cohen's original derivation of the one-dimensional diffusion equation, the gas flow pattern in the

1063-7842/00/4509-1122$20.00

rotor has been significantly simplified: (i) It was assumed that the circulation flow in the rotor did not vary along the GC and that flow closure occurred in negligibly thin layers near the rotor ends. (ii) The feed flow was assumed to be negligible in comparison with the circulation flow in the rotor. It was therefore erroneously believed that Cohen's one-dimensional diffusion equation was inapplicable to a quantitative description of the separation in GCs under the conditions of nonuniform circulation flow along the rotor and finite feed flow. For example, it was affirmed that a flow in the radial direction would reduce the GC efficiency because of convective mixing [6].

We will show that the one-dimensional diffusion equation adequately describes an arbitrary circulation flow in the GC rotor without sacrificing description accuracy. In Fig. 1, a cylindrical coordinate system is shown. The GC rotation axis corresponds to $r = 0$, and the point $z = 0$ coincides with the feeding point. It is assumed that the problem is two-dimensional and that an isotopic approximation can be used; that is, the isotope concentration does not affect the gas flow as a whole.

Let us divide the flow inside the rotor into two components: (1) the circulation (closed stream lines) and (2) the transit flow (open stream lines). Mathematically, this is equivalent to decomposition of the mass flux vector field into vortex and potential components. Such a decomposition exists and is unique for any continuous vector field [8]:

$$\rho\mathbf{V} = \rho\mathbf{V}_\Psi + \rho\mathbf{V}_\varphi = \rho D\,\mathrm{curl}(\mathbf{\Psi}) + \mathrm{grad}(\varphi), \quad (3)$$

where $\mathbf{\Psi} = \begin{pmatrix} 0 \\ \Psi \\ 0 \end{pmatrix}$ is the vector potential of the circulation flow; φ is the scalar potential of the transit flow; $\mathbf{V}_\Psi$ and $\mathbf{V}_\varphi$ are the corresponding flow velocities; the factor ρD is introduced for convenience, and it is assumed that ρD does not depend on the coordinates; and the single nonzero component of $\mathbf{\Psi}$ is due to the two-dimensionality of the flow.

In a regime with no product withdrawal ($F = 0$) and zero radial mass flux ($\rho V_r = 0$), $\rho D\Psi r = \psi$, where ψ is the Cohen stream function [6]. In other cases, the circulation potential does not coincide with the stream function. The main distinction is that the feed flow does not make a direct contribution to ψ. This does not mean that the feed flow does not affect the circulation. The feed can still be considered as one of the mechanisms inducing the vortex gas flow in the GC rotor (along with the temperature gradient and mechanical agitation).

Cohen ignored the distinction between the two kinds of flows in the rotor, on the assumption that the feed flow is small. In real GCs, the feed and circulation flows are comparable, which is a necessary condition for optimal operation, as shown below.

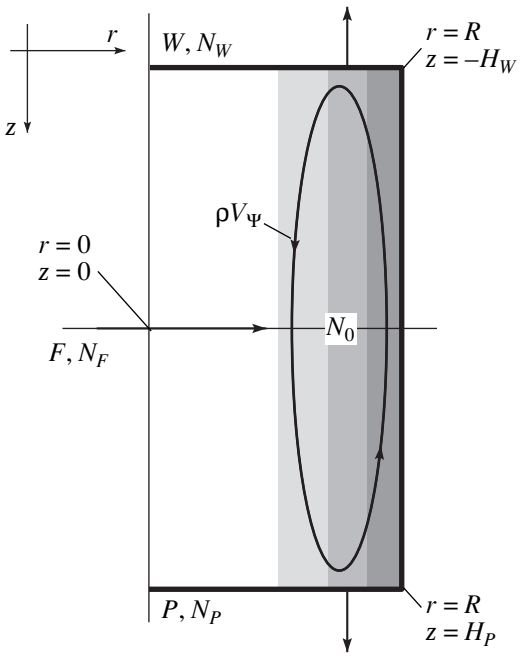


Fig. 1. Schematic diagram of a centrifuge.

Similar to the expression for the mass flux, the isotope mass flux $\mathbf{\Phi}$ can be represented as

$$\mathbf{\Phi} = \mathrm{curl}(\mathbf{\Psi}_D) + \mathrm{grad}(\varphi_D). \quad (4)$$

The vector potential $\mathbf{\Psi}_D$ defines the vortex isotope mass flux, and the scalar potential φ_D defines the transit isotope mass flux.

On the other hand, by definition, the isotope mass flux is

$$\Phi_r = -\rho D\left[\frac{\partial N}{\partial r} + frN(1-N)\right] + \rho V_r N,$$
$$\Phi_z = -\rho D\frac{\partial N}{\partial z} + \rho V_z N, \quad (5)$$

where N is the molar concentration of the isotope and $f = (\Delta m\Omega^2)/(kT)$.

It is possible to represent (4) in the form

$$\mathbf{\Phi}_i = \mathrm{curl}(\mathbf{\Psi}_{Di}) + N_i\rho\mathbf{V}_{\varphi i}, \quad (6)$$

where $\rho\mathbf{V}_\varphi$ is the vector of the potential component of the mass flux in a GC and i is an index equal to P for the enriching section of a GC and to W for the extracting one.

With expression (6), the diffusion equation

$$\mathrm{div}(\mathbf{\Phi}_i) = 0 \quad (7)$$

is reduced to an identity. Taking $\mathbf{\Psi}_D$ in the form $\mathbf{\Psi}_D = (\rho D\mathbf{\Psi}N + \delta\boldsymbol{\phi})$, where $\delta\boldsymbol{\phi}$ is an unknown function, gives by implication the deviation of the vortex of the isotope mass flux from the vortex of the average mass flux. Equating expressions (5) and (6) to each other allows one to transform Eq. (7) into the set of equations

$$-\left[\frac{\partial N}{\partial r} + frN(1-N)\right] + \frac{\rho V_{\varphi ri}}{\rho D}(N - N_i)$$

$$-\Psi\frac{\partial N}{\partial z} - \frac{1}{\rho D}\left(\frac{\partial \delta\phi_i}{\partial z}\right) = 0, \qquad (8)$$

$$-\frac{\partial N}{\partial z} + \frac{\rho V_{\varphi zi}}{\rho D}(N - N_i) + \Psi\frac{\partial N}{\partial r} - \frac{1}{\rho Dr}\frac{\partial(r\delta\phi_i)}{\partial r} = 0.$$

The following substitutions were made in (5): $N\rho\mathbf{V} = N\rho\mathbf{V}_\Psi + N\rho\mathbf{V}_\phi$ and

$$\operatorname{curl}(\mathbf{\Psi}_D) = N\operatorname{curl}(\rho D\Psi) + \rho D\mathbf{\Psi}\operatorname{grad}(N) + \operatorname{curl}(\delta\phi)$$
$$= N\rho\mathbf{V}_\Psi + \rho D\mathbf{\Psi}\operatorname{grad}(N) + \operatorname{curl}(\delta\phi).$$

Thus, the terms containing $N\rho\mathbf{V}_\Psi$ cancel out and no derivatives of Ψ remain in the equations, whereas the terms involving $\rho\mathbf{V}_\varphi$ remain. Eliminating the terms containing $\partial N/\partial z$ from the first equation of (8) and the terms containing $\partial N/\partial r$ from the second equation gives

$$-(1+\Psi^2)\frac{\partial N}{\partial r} - frN(1-N) + \vartheta_r(N - N_i)$$

$$-\left(\frac{\Psi}{r}\frac{\partial(r\delta\phi)}{\partial r} + \frac{\partial\delta\phi}{\partial z}\right) = 0,$$

$$-(1+\Psi^2)\frac{\partial N}{\partial z} - fr\Psi N(1-N) + (\vartheta_z + \Psi\vartheta_r) \qquad (9)$$

$$\times(N - N_i) + \left(\frac{1}{r}\frac{\partial(r\delta\phi)}{\partial r} - \Psi\frac{\partial\delta\phi}{\partial z}\right) = 0,$$

where $\boldsymbol{\vartheta} = (\mathbf{V}_\varphi/D)$.

Up to now, no simplifications have been made; therefore, set (9) is equivalent to Eq. (7). The second equation in (9) can be subjected to radial averaging under the conventional assumption that the concentration depends weakly on r, that is, by the substitution $\langle N(r)X(r)\rangle \approx \langle N(r)\rangle\langle X(r)\rangle$. At the same time,

$$\int_0^R \frac{2\pi r}{r}\frac{\partial(r\delta\phi)}{\partial r}dr = 2\pi r\delta\phi\big|_0^R = 2\pi R\delta\phi(R) = 0$$

due to imperviousness of the boundaries to the isotope flow ($\delta\varphi(R) = 0$), and the term

$$\int_0^R 2\pi r\Psi\frac{\partial\delta\phi}{\partial z}dr \approx 0,$$

because the deviation of the vortex of the isotope mass flux from the vortex of the average mass flux is small. Now we obtain an equation similar to that of Cohen:

$$-\frac{\partial N}{\partial z} - \frac{f\langle r\Psi\rangle}{1+\langle\Psi^2\rangle}N(1-N) + \frac{\vartheta_z}{1+\langle\Psi^2\rangle}$$
$$\times\left(1 + \frac{\langle\Psi\vartheta_r\rangle}{\vartheta_z}\right)(N - N_i) = 0. \qquad (10)$$

Here,

$$\langle X\rangle = \frac{1}{\pi R^2}\int_0^R 2\pi r X(r)dr.$$

Equation (10) differs from Cohen's original equation in that it is applicable to an arbitrary distribution of the circulation flow, and in the absence of circulation in the rotor ($\Psi = 0$), the concentration gradient along the 0z-axis becomes equal to zero, although the feed flow may be nonzero ($\vartheta_z \neq 0$).

The radial velocity component $\vartheta_r = \vartheta_r(z)$ related to the feed flow depends on the feeding method. If feeding is implemented near the centrifuge axis, then $\vartheta_r > 0$ and the GC efficiency may be higher. Here, we will not elaborate on this point. Let us suppose that the feed flow does not give rise to the radial velocity component ($\vartheta_r = 0$). This will not affect the radial velocity related to circulation. In this case, the term containing ϑ_r vanishes from Eq. (10) and it becomes entirely identical in form to Cohen's equation.

To integrate Eq. (10), it is necessary either to know the circulation potential or to make some assumption concerning its form. Let us introduce the following assumption: $\Psi = \Psi(r)\Psi(z)$. Then one can introduce the circulation amplitude in the form $\Psi_0(z) = \sqrt{\langle\Psi^2(z)\rangle}$, as well as the profile coefficient

$$a = AR = \frac{\langle\Psi_r\rangle}{\sqrt{\langle\Psi^2\rangle}}, \qquad (11)$$

where A is a constant characterizing the flow profile along the 0r-axis; its maximum value is $A = \sqrt{1/2}$ when $\Psi = \Psi_0(z)r$.

Taking into account assumption (11), Eq. (10) takes the form

$$-\frac{d}{dz}\left(\ln\left(\frac{N}{1-N}\right)\right) + \frac{fa\Psi_0(z)}{1+\Psi_0(z)^2}$$
$$-\frac{\vartheta_z}{1+\Psi_0(z)^2}\frac{N - N_i}{N(1-N)} = 0. \qquad (12)$$

To solve equation (12), it should be linearized. Cohen linearized it on the assumption that the isotope concentration is small ($N \ll 1$). A more general linearized form can be obtained by requiring only that the

concentration difference be small ($\Delta N \ll 1$) and taking into account that

$$(N - N_i)\frac{d}{dN}\left(\ln\left(\frac{N(1-N_i)}{N_i(1-N)}\right)\right) = \frac{N - N_i}{N(1-N)}; \quad (13)$$

that is,

$$\frac{N - N_i}{N(1-N)}$$

is the first term of the expansion of the separation factor α_i into a series in N in the vicinity of N_i, where

$$\alpha_i = \frac{N(1-N_i)}{N_i(1-N)}.$$

To within the next series term $\sim(N - N_i)^2$ one can obtain[1]

$$-\frac{d}{dz}(\ln\alpha_i) + \frac{fa\Psi_0}{1+\Psi_0^2} - \frac{\vartheta_z}{1+\Psi_0^2}\ln\alpha_i = 0. \quad (14)$$

The general integral of Eq. (14) is

$$\ln\alpha_i(z) = fa\exp\left(-\vartheta_z\int_{H_i}^{z}\frac{dz}{1+\Psi_0^2}\right) \times\int_{H_i}^{z}\left[\frac{\Psi_0}{1+\Psi_0^2}\exp\left(\vartheta_z\int_{H_i}^{z}\frac{dz}{1+\Psi_0^2}\right)\right]dz, \quad (15)$$

and the total separation factor of a GC is

$$\ln\chi = \ln\alpha_P(0) - \ln\alpha_W(0).$$

Integral (15) can be expressed analytically in two important cases: (1) when the circulation does not depend on z and (2) for the case of optimal circulation. For circulation flow Ψ_0 constant along the z-axis we have

$$\ln\chi = \frac{\pi a\Psi_0\rho D\Delta m(\Omega R)^2}{kTF\Theta(1-\Theta)} \times(1-(1-\Theta)e^{\Theta KH_P} - \Theta e^{(1-\Theta)KH_W}), \quad (16)$$

where

$$K = -\frac{F}{\pi R^2\rho D(1+\Psi_0^2)},$$

and $\Theta = P/F$ is the ratio of the product/feed flows in a GC (flow division factor).

A detailed analysis of the solution for constant circulation is outside the scope of this article.

In addition, Eq. (12) can be compared with the equation for a symmetrical cascade

$$\frac{dN}{ds} = \varepsilon N(1-N) - \frac{P}{L''(s)}(N_P - N), \quad (17)$$

where the cascade stage number s corresponds to the z coordinate, the enrichment factor ε is given by

$$\varepsilon = \frac{fa\Psi_0}{1+\Psi_0^2}, \quad (18)$$

the feed flow is ϑ_z, and the stage waste flow L'' is given by

$$L''(s) = 1 + \Psi_0(z)^2. \quad (19)$$

The main difference between a centrifuge and a cascade is that in the former a connection exists between the waste flow (19) and the enrichment factor (18) via the circulation potential Ψ_0, making possible a variation of the enrichment factor in a GC along its inner cascade, whereas in a conventional cascade the enrichment factor is considered to be independent of the stage number.

An optimization task can be set for the inner cascade of a centrifuge. For an ideal cascade, the goal is a cascade with minimum power consumption for a given separation regime (for given F, P, N_F, and N_P), that is, a cascade consisting of the minimum number of centrifuges. For an isolated centrifuge, minimization of the power expended on maintaining the circulation flow inside the rotor appears senseless. However, it is possible to state the problem of defining the circulation in the rotor in such a way that the separative power is maximum for the given external parameters (the separation regime). Below, this centrifuge will be called a perfect centrifuge (PC).

There are four external parameters in both a centrifuge and a cascade: (1) concentration N_F in the feed flow, (2) concentration N_0 at the feeding point (a more convenient parameter for a GC than the product concentration), (3) the feed flow F, and (4) the flow division factor Θ. A relation between two of these for the ideal case can be indicated at once: $N_0 = N_F$. Relations between the others should be established from the condition of maximum separative power. The only internal parameter in the proposed centrifuge model is the circulation potential $\Psi_0(z)$. It is necessary to define $\Psi_0(z)$ in such a way as to obtain the maximum separative power for the given external parameters.

[1] For linearization one can also use the function

$$Y(N) = \frac{N - N_i}{N(1-N)}$$

and derive an analogous equation

$$-\frac{dY}{dz} + \frac{fa\Psi_0}{1+\Psi_0^2} - \frac{\vartheta_z}{1+\Psi_0^2}Y = 0.$$

The solution of the initial nonlinear equation is in between these two approximations.

For a given feed flow F and the flow division factor Θ, the separative work [9] is

$$\delta U = \Theta(1-\Theta)F(\ln\chi)^2 \tag{20}$$

and will be a maximum along with the enrichment factor χ. A maximum of $\ln\chi$ is attained if $d\ln\chi/dz$ is a maximum at every point on the z-axis. The only remaining arbitrary parameter in Eq. (14) is ψ_0. By differentiating[2] (14) with respect to Ψ_0 (this step is analogous to searching for the distribution of the flows in the ideal cascade [6]), we obtain

$$\frac{d}{d\Psi_0}\left(\frac{fa\Psi_0 - \vartheta_z\ln\alpha_i}{1+\Psi_0^2}\right) = \frac{fa(1-\Psi_0^2) + 2\vartheta_z\Psi_0\ln\alpha_i}{(1+\Psi_0^2)^2} = 0, \tag{21}$$

from which it follows that

$$\Psi_{0\ \mathrm{opt}} = \frac{\vartheta_z\ln\alpha_i}{fa} + \sqrt{\left(\frac{\vartheta_z\ln\alpha_i}{fa}\right)^2 + 1}. \tag{22}$$

Let us note that $\ln\alpha_i(z=0) = \ln\alpha_P$ and $\ln\alpha_i(z=H_P) = 0$. Consequently, the optimal $\Psi_{0\mathrm{opt}}(z)$ has a maximum at the feeding point and decreases down to

$$\Psi_{0\ \mathrm{opt}}(z = H_p, H_W) = 1 \tag{23}$$

at the rotor ends.

Since the circulation flow is common to the enriching and extracting centrifuge sections, the centrifuge optimization with respect to the circulation flow imposes one more relation on the parameters:

$$\frac{\Theta}{1-\Theta} = \frac{\ln\alpha_W}{\ln\alpha_P}. \tag{24}$$

Thus, at $H_P = H_W$ it is necessary that $\Theta = 0.5$.

Substituting expression (22) into Eq. (14) gives the following relation for the PC:

$$\frac{d\ln\alpha_i}{dz} = \frac{fa}{2\Psi_{0\ \mathrm{opt}}}. \tag{25}$$

Equation (25) leads to the conclusion that long centrifuges are rather inefficient. This follows from the observation that the central part of a centrifuge in the optimal regime makes a smaller contribution to the separation factor than the end parts (Ψ_0 is maximal and $d(\ln\alpha)/dz$ minimal).

Equation (25) can be integrated in order to derive an expression for the maximum attainable separation factor in a GC.

By introducing the denotations

$$\beta_i = \frac{\vartheta_{zi}\ln\alpha_i}{fa}; \quad l = z\vartheta_{zi}; \quad l_i = H_i\vartheta_{zi}, \tag{26}$$

the solution of Eq. (25) can be represented in the following form:

$$l_i = \beta_i(\sqrt{1+\beta_i^2} + \beta_i) + \ln(\sqrt{1+\beta_i^2} + \beta_i). \tag{27}$$

It is impossible to derive an analytical expression for $\beta_i(l_i)$ from function (27).

The optimal circulation potential (22) can be written in the same notation:

$$\Psi_{0\ \mathrm{opt}} = \sqrt{1+\beta_i^2} + \beta_i. \tag{28}$$

The relation

$$\frac{\beta_i}{l_i} = \frac{\ln(\alpha_i)}{faH_i}$$

does not explicitly depend on the feed flow and can be considered as the normalized separation factor per unit length of the centrifuge, and quantity l_i can be considered as the effective length of a GC. An ideal dependence of the separation factor on the feed flow and the centrifuge length is plotted in Fig. 2. A data point of a real centrifuge cannot occur above the curve presented in the figure. For comparison, in the same plot, the maximum separation factor is given for an optimized GC with a constant circulation flow along the 0z-axis. The points for real centrifuges presumably fall between these two curves. Figure 2 shows that an increase in l, that is, in either the rotor length (H) or the feed flow (ϑ_z), reduces the separation factor per unit length of the GC.

The maximum separative power of a GC is plotted in Fig. 3. The present theory allows one to more accu-

[2] From (14) it follows that

$$\frac{d}{dz}(\ln\alpha_i) = \frac{fa\Psi_0 - \vartheta_z\ln\alpha_i}{1+\Psi_0^2}.$$

The condition for the extremum is

$$\frac{d}{d\Psi_0}\left(\frac{d}{dz}(\ln\alpha_i)\right) = 0$$

or

$$\frac{d}{d\Psi_0}\left(\frac{fa\Psi_0 - \vartheta_z\ln\alpha_i}{1+\Psi_0^2}\right) = 0.$$

In turn,

$$\frac{d}{d\Psi_0}\left(\frac{fa\Psi_0 - \vartheta_z\ln\alpha_i}{1+\Psi_0^2}\right) = \frac{\partial}{\partial\Psi_0}\left(\frac{fa\Psi_0 - \vartheta_z\ln\alpha_i}{1+\Psi_0^2}\right) + \frac{\partial}{\partial\ln\alpha_i}\left(\frac{fa\Psi_0 - \vartheta_z\ln\alpha_i}{1+\Psi_0^2}\right)\frac{d\ln\alpha_i}{d\Psi_0}.$$

However, we should have obtained $(d\ln\alpha_i)/(d\Psi_0) = 0$, because $\ln\alpha_i$ is the extremal separation factor.

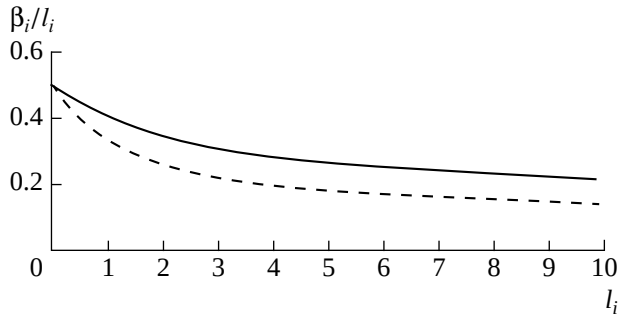


Fig. 2. The normalized separation factor per unit length of the rotor of a perfect centrifuge (see Eq. (27) and definition (26)). The solid line represents the perfect centrifuge; the dashed one represents an optimized centrifuge with constant circulation flow along the 0z-axis (see Eq. (16)).

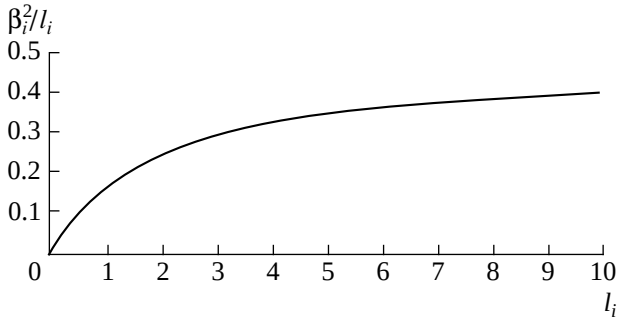


Fig. 3. The normalized separative power per unit length of the rotor for the perfect centrifuge (see Eqs. (27) and (20), as well as definition (26)).

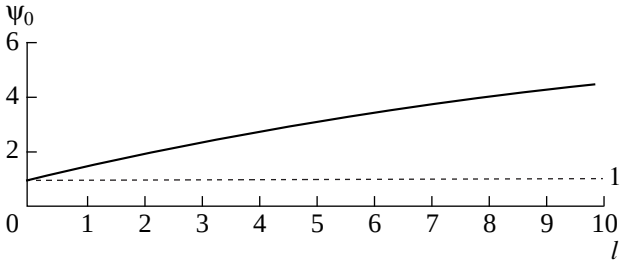


Fig. 4. The ideal profile of the intrarotor circulation potential (see Eqs. (27) and (28), as well as definition (26)). The reference point for l values is the rotor cover.

rately estimate the maximum separative power of a GC by taking into account the influence of the feed flow.

An ideal $\Psi_{0\text{opt}}(l)$ dependence is given in Fig. 4. According to expression (11) for Ψ_0, the optimal dependence of the flux ρV_z on l coincides with $\Psi_0(l)$ everywhere except for thin layers near the rotor covers. To obtain a finite magnitude of Ψ_0 at $z = 0$, it is necessary to have a deltalike flow profile ($\rho V_z(l) \sim \delta(l - l_i)$) near the cover.

CONCLUSIONS

(1) The radial averaging method is applicable, without sacrificing accuracy, for describing the separation in GCs for an essentially nonuniform distribution of the circulation flow along the rotor length.

(2) The radial mass fluxes in GCs do not inadvertently cause a decrease in the separative power.

(3) For the internal circulation in a GC there exists an ideal distribution of the circulation potential and a corresponding ideal distribution of the mass flux along the rotor axis.

ACKNOWLEDGMENTS

The present work was supported by the Electrochemical Plant PO (city of Zelenogorsk) and the Ministry of General and Vocational Education of the Russian Federation.

REFERENCES

1. V. D. Borisevich, E. V. Levin, and V. V. Naumochkin, At. Énerg. **70** (1), 28 (1991).
2. V. D. Borisevich, V. V. Naumochkin, and B. M. Smakov, Izv. Akad. Nauk SSSR, Mekh. Zhidk. Gaza, No. 4, 24 (1987).
3. V. D. Borisevich, E. V. Levin, and V. V. Naumochkin, Izv. Akad. Nauk SSSR, Mekh. Zhidk. Gaza, No. 4, 39 (1989).
4. É. M. Aĭsen, V. D. Borisevich, and E. V. Levin, At. Énerg. **72** (1), 44 (1992).
5. E. V. Levin, At. Énerg. **72** (1), 39 (1992).
6. *Uranium Enrichment*, Ed. by S. Villani (Springer-Verlag, Berlin, 1979; Énergoatomizdat, Moscow, 1983).
7. K. Cohen, *The Theory of Isotope Separation as Applied to the Large-Scale Production of* ^{235}U (McGraw-Hill, New York, 1951).
8. G. A. Korn and T. M. Korn, *Mathematical Handbook for Scientists and Engineers* (McGraw-Hill, New York, 1968; Nauka, Moscow, 1978).
9. V. A. Palkin, At. Énerg. **84** (3), 253 (1998).

Translated by N. Mende

Technical Physics, Vol. 45, No. 9, 2000, pp. 1128–1134. Translated from Zhurnal Tekhnicheskoĭ Fiziki, Vol. 70, No. 9, 2000, pp. 30–36.

On the Influence of Viscosity on the Characteristic Time of the Instability of a Charged Drop

S. O. Shiryaeva

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

E-mail: shir@uniyar.ac.ru

Received November 19, 1999.

Abstract—A nonlinear integral equation describing the evolution of spheroidal deformation of a drop that is unstable with respect to its intrinsic charge is derived and solved for arbitrary values of viscosity. It was shown that, due to an essentially nonlinear character of the phenomenon, the characteristic time of instability development equals the time of tenfold increase in the amplitude of an initial, physically infinitesimal spheroidal deformation of an unstable drop. The dependence of the instability characteristic time on the drop viscosity is described by an increasing linear function. *© 2000 MAIK "Nauka/Interperiodica".*

INTRODUCTION

Theoretical estimation of the characteristic time of the instability of a highly charged drop is of interest for various applications of the phenomenon in geophysics and technical physics. It is also important, because the instability development is difficult to observe directly in experiments. In the numerous experiments conducted to verify Rayleigh's stability criterion for a highly charged drop, observers have detected only the initial and final states of the drops (see [1, 2] and references therein). Theoretical estimates of the instability development time for a highly charged drop were obtained in [3–6] for an inviscid liquid, in which case the characteristic time scale of the process depends only on the relation between the unstable drop deformation and the extent of drop charge supercriticality. Using the same approach as in [3–5] and taking into account the decrease in the instability growth rate due to the viscous damping of liquid capillary motion in spherical and spheroidal drops [7, 8], one can estimate the influence of viscosity.

1. The spectrum of capillary oscillation of an isolated drop of a conducting inviscid liquid of radius R carrying a charge Q and characterized by a surface tension σ has the form

$$\omega_n^2 = \frac{\sigma}{\rho R^3} n(n-1)[(n+2) - W], \quad W = \frac{Q^2}{4\pi\sigma R^3}, \quad (1)$$

where n is the number of a capillary oscillation eigenmode and ρ is the liquid density [9].

It follows from (1) that, when $W > 4$, the principal mode ($n = 2$) becomes unstable and its amplitude ζ grows with time as $\zeta \sim \exp(\gamma_0 t)$, where

$$\gamma_0 = \left\{\frac{2\sigma}{\rho R^3}(W - W_{2*})\right\}^{1/2},$$

and $W_{2*} = 4$ is the critical value of parameter W for the onset of the principal-mode instability.

If the liquid is characterized by a kinematic viscosity ν_*, then the growth rate of the principal-mode instability is lower by the value of the viscous damping decrement (see Appendix):

$$\eta_g \equiv \gamma_0 - \gamma_g \equiv \left\{\frac{2\sigma}{\rho R^3}(W - W_{2*})\right\}^{1/2}\left\{1 - \frac{1}{1 + \nu k(W)}\right\},$$

$$\gamma_g \approx \gamma_0[1 + \nu k(W)]^{-1}, \quad (2)$$

$$k(W) = \left[1 + \frac{16}{W+4} + \frac{1}{2(W - W_{2*})}\right],$$

where γ_g is the instability growth rate for a viscous spherical drop and ν is the dimensionless viscosity normalized to its characteristic value $(R\sigma/\rho)^{1/2}$. The denominator of the last term in the expression for $k(W)$ characterizes the degree of drop charge subcriticality.

2. When $W = 4$, the drop is unstable with respect to an infinitesimal surface deformation of the form $\zeta = \zeta_0 P_2(\cos\Theta)$, which corresponds to excitation of the principal mode of capillary oscillation. Such a capillary oscillation can be excited, for example, by thermal motion of liquid molecules. In this case, the amplitude ζ_0 is expressed as $\zeta_0 = (\sigma/kT)^{1/2}$, where k is the Boltzmann constant and T is the absolute temperature of the liquid. The excitation of a capillary wave $\sim P_2(\cos\Theta)$ corresponds to the elongation of the drop into a spher-

1063-7842/00/4509-1128$20.00

oid of eccentricity $e_0 = (3\zeta_0/R)^{1/2}$ [3–5]. For a spheroid, the critical value of W for the onset of the principal-mode instability is known to be a decreasing function of eccentricity [10]. In the linear approximation with respect to the squared eccentricity e^2, this function has the form

$$W_{2*} = 4(1 - \alpha e^2), \tag{3}$$

where $\alpha = 1/3$ [10].

Therefore, when $W = 4$, the amplitude $\zeta = \zeta_0 P_2(\cos\Theta)$ of a thermal perturbation of an initially spherical drop grows with time as an exponential with the growth rate

$$\gamma_d \equiv \gamma_0 - \eta_g = \left\{\frac{8\sigma}{\rho R^3}\alpha e_0^2\right\}^{1/2}\frac{1}{1 + \nu k(W)},$$

where $k(W)$ for $W = 4$ is rewritten by using (3) as follows:

$$k(W) \equiv k(e_0^2) = \left[3 + \frac{1}{8\alpha e_0^2}\right].$$

An increase in the amplitude of a perturbation $\sim P_2(\cos\Theta)$ corresponds to a further elongation of the drop. Its eccentricity grows, the critical value of W decreases according to (3) (i.e., the degree of charge supercriticality increases), and the instability growth rate increases. As the drop geometry changes from spherical to spheroidal, the growth rate of the principal-mode instability becomes a function of the extent of spheroidal deformation, i.e., the squared eccentricity e^2. As shown in the Appendix, this function is

$$\chi \equiv \gamma_{gd} = \gamma_g\left[1 - \frac{2}{1 + 0.8\nu}e^2\right], \tag{4}$$

where χ denotes the instability growth rate for a charged viscous spheroidal drop.

Thus, the amplitude ζ of a spherical surface perturbation $\sim P_2(\cos\Theta)$ grows with time as

$$\zeta = \zeta_0 e^{\chi t}$$
$$= \zeta_0 \exp\left(\sqrt{\frac{8\sigma}{\rho R^3}\alpha e^2}\frac{1}{1 + \nu k(e^2)}\left[1 - \frac{2}{1 + 0.8\nu}e^2\right]t\right). \tag{5}$$

3. Note that (5) is valid only when the relation between the perturbation growth rate $d\zeta/dt$ and the perturbation amplitude ζ is linear. To find the value of ζ at any moment, we consider the sequence of values of ζ taken after time intervals Δt_i ($i = 1, 2, 3, 4, \ldots$) during each of which expression (5) can be assumed to be valid. The growth rate χ is a function of the spheroidal deformation e^2 increasing with time, where χ depends on time. The interval Δt_i is determined by the condition that change in the growth rate over the interval $\chi_i - \chi_{i-1}$ is much less than the growth rate at the end of the preceding time interval:

$$\Delta\chi \equiv \chi_i - \chi_{i-1} \ll \chi_{i-1}.$$

Then, the value of χ_i can be treated as constant and equal to χ_{i-1} within the interval Δt, and expression (5) can be used to calculate the increase in amplitude $\Delta\zeta_i$.

Suppose that an initially spherical drop is unstable and begins to elongate with growth rate χ_0. The perturbation amplitude will increase over an interval Δt_1 from ζ_0 to

$$\zeta_1 = \zeta_0 \exp\chi_0\Delta t_1.$$

This will cause an increase in e^2 and, according to (4), raise the growth rate to χ_1. Over the next interval Δt_2, the perturbation amplitude will vary as

$$\zeta_2 = \zeta_1 \exp\chi_1\Delta t_2.$$

By the end of the interval Δt_2, the perturbation amplitude, and hence e^2, will have grown; accordingly, the growth rate χ will become χ_2 at the end of Δt_2. During the next time interval Δt_3, the perturbation amplitude will increase as

$$\zeta_3 = \zeta_2 \exp\chi_2\Delta t_3.$$

Thus, for the ith interval, we write

$$\zeta_i = \zeta_{i-1} \exp\chi_{i-1}\Delta t_i. \tag{6}$$

In (6), ζ_{i-1} can be expressed in terms of ζ_{i-2}, ζ_{i-3}, and all the preceeding values of the amplitude up to ζ_0. As a result, expression (6) becomes

$$\zeta_i = \zeta_0 \exp\left\{\sum_{k=1}^{i}\chi_{k-1}\Delta t_k\right\}.$$

In the limit of $\Delta t_k \longrightarrow 0$ and $i \longrightarrow \infty$, we have

$$\zeta(t) = \zeta_0 \exp\left\{\int_0^t \chi(e^2(t))dt\right\}.$$

Substituting (4) for $\chi(e^2(t))$ in the last expression, we obtain

$$\zeta(t) = \zeta_0 \exp\left\{\int_0^t\left(\sqrt{\frac{8\sigma}{\rho R^3}\alpha e^2(t)}\frac{1}{1 + \nu k(e^2(t))}\right.\right.$$
$$\left.\left.\times\left[1 - \frac{2}{1 + 0.8\nu}e^2(t)\right]\right)dt\right\}.$$

In the linear approximation with respect to e^2, valid when $e^2 \ll 1$, we can write

$$e^2 \approx \frac{3\zeta}{R},$$

This means that the analysis that follows is valid provided that $\zeta(t)/R \ll 0.3$. The equation then takes the form

$$\zeta(t) = \zeta_0 \exp\left\{\int_0^t \left(\sqrt{\frac{24\sigma}{\rho R^3}}\alpha\frac{\zeta(t)}{R}\frac{1}{1+\nu k(\zeta(t)/R)}\right.\right.$$
$$\left.\left.\times\left[1-\frac{6}{1+0.8\nu}\frac{\zeta(t)}{R}\right]\right)dt\right\}, \quad (7)$$
$$k\left(\frac{\zeta(t)}{R}\right) \equiv \left[3+\frac{1}{24\alpha}\frac{R}{\zeta(t)}\right].$$

Thus, we obtain a nonlinear integral equation for the amplitude $\zeta(t)$ of the unstable principal mode of surface oscillation for a viscous charged drop.

4. Let us rewrite the equation in the following dimensionless form:

$$X(t) = X_0 \exp\left\{\int_0^t \left(X^{1/2}\frac{1}{1+\nu k(X)}\left[1-\frac{6}{1+0.8\nu}X\right]\right)\kappa dt\right\},$$
$$X \equiv \frac{\zeta}{R}, \quad \kappa = \sqrt{\frac{24\sigma}{\rho R^3}}\alpha, \quad k(X) \equiv \left[3+\frac{1}{24\alpha X}\right].$$

To find a solution, we take the logarithm of the equation:

$$\ln\frac{X}{X_0} = \int_0^t \left(X^{1/2}\frac{1}{1+\nu k(X)}\left[1-\frac{6}{1+0.8\nu}X\right]\right)\kappa dt.$$

Then, we differentiate both sides of the resulting expression:

$$\frac{1}{X}\frac{dX}{d(\kappa t)} = \frac{X^{1/2}}{1+\nu k(X)}\left[1-\frac{6}{1+0.8\nu}X\right].$$

Separating the variables, we integrate the result to find

$$\frac{2a}{3}X^{-3/2} + 2(ac+b)X^{-1/2} - 2c^{1/2}(ac+b)\operatorname{arctanh}(cX)^{1/2}$$
$$= \frac{2a}{3}X_0^{-3/2} + 2(ac+b)X_0^{-1/2} - 2c^{1/2}(ac+b) \quad (8)$$
$$\times \operatorname{arctanh}(cX_0)^{1/2} - \kappa t,$$
$$X_0 \equiv \frac{\zeta_0}{R}, \quad a \equiv \frac{\nu}{24\alpha}, \quad b \equiv 1+3\nu, \quad c \equiv \frac{6}{1+0.8\nu}.$$

The terms containing the function arctanh in (8) are much smaller than the remaining terms when $X \ll 0.3$. They affect the evolution of a spheroidal deformation of the drop only at the final stage of the instability development, when $X \sim 1$. In experiments on the Rayleigh breakup of drops, we had $R \sim 100$ μm [1, 2] and the amplitude of thermal capillary waves was $\zeta_0 \sim 0.1$ nm; therefore, $X_0 \sim 10^{-6}$. The breakup of a drop that is unstable with respect to its intrinsic charge (the loss of excessive charge) starts when $X \approx 1$ [2]. Hence, the influence of the terms containing arctanh in (8) on the drop evolution can be neglected. This is confirmed by a direct calculation based on (8).

In the general case, equation (8) is too cumbersome to yield an explicit form of the amplitude $\zeta = \zeta(t)$ as a function of time. Therefore, we consider asymptotic situations. When $\nu = 0$ (i.e., $a = 0$ and $b = 1$), by omitting the terms containing arctanh, expression (8) is transformed into

$$X^{-1/2} = X_0^{-1/2} - 0.5\kappa t,$$

which yields

$$\zeta = \frac{\zeta_0}{(1-t/T_0)^2},$$
$$T_0 = \frac{2}{X_0^{1/2}\kappa} = R^2\left(\frac{\rho}{6\alpha\zeta_0\sigma}\right)^{1/2}. \quad (9)$$

Thus, we obtain the time dependence of the spheroidal deformation amplitude for an inviscid liquid drop that is unstable with respect to its intrinsic charge known from [3]. In the low-viscosity approximation, only the terms containing either $\sim bX^{-1/2}$ or $bX_0^{-1/2}$ can be retained in (8), and the spheroidal deformation amplitude for a low-viscosity liquid drop unstable with respect to its intrinsic charge has the same time dependence as in (9) but with T_0 multiplied by b; i.e., viscosity increases the characteristic time of instability development.

When the kinematic viscosity ν_*, drop radius R, liquid density ρ, and surface tension σ are such that $\nu_*(\rho\sigma^{-1}R^{-1})^{1/2} \geq 1$, the drop can be treated as highly viscous [2]. When $X \ll 1$, the terms containing $\sim X^{-3/2}$ in (8) are dominant and (8) transforms into

$$\frac{2a}{3}X^{-3/2} = \frac{2a}{3}X_0^{-3/2} - \kappa t$$

yielding

$$\zeta = \frac{\zeta_0}{(1-t/T_\nu)^{2/3}},$$
$$T_\nu = \frac{2}{3}\frac{a}{X_0^{3/2}\kappa} = \nu R^3\left(\frac{\rho}{3^5\times 2^7\alpha^3\zeta_0^3\sigma}\right)^{1/2}. \quad (10)$$

It is clear that the time dependence of a spheroidal perturbation amplitude for a highly viscous drop with $\nu \geq 1$ is weaker than for an inviscid drop. The characteristic time T_ν of instability development for a viscous drop is $aX_0^{-1}/3$ times longer than T_0 for an inviscid liq-

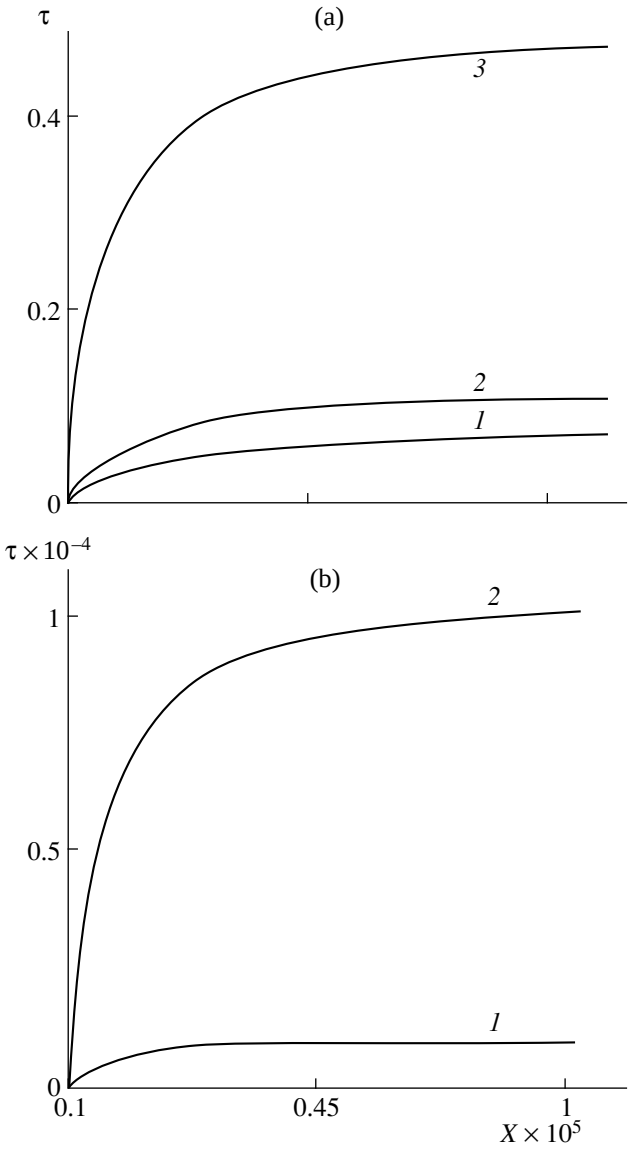


Fig. 1. Dimensionless time versus dimensionless amplitude of spheroidal deformation of a drop that is unstable with respect to intrinsic charge for $\alpha = 1/3$ and $X_0 = 10^{-6}$ in the case of (a) an inviscid liquid (*1*) and a viscous liquid with $\nu = 10^{-5}$ (*2*) and 10^{-4} (*3*) and (b) a viscous liquid with $\nu = 0.25$ (*1*) and 2.5 (*2*).

uid drop. Since X_0 is small, the effect of viscosity turns out to be significant.

The results of numerical calculations based on the complete version of (8) and performed for various values of viscosity are shown in Figs. 1 and 2. They demonstrate that the characteristic time of spheroidal deformation of an unstable viscous drop is substantially longer compared to the case of an inviscid liquid drop even for $\nu \ll 1$. Figure 1 shows that, due to the strongly nonlinear character of the process, the characteristic time of instability development for a drop that is unstable with respect to its intrinsic charge is the time

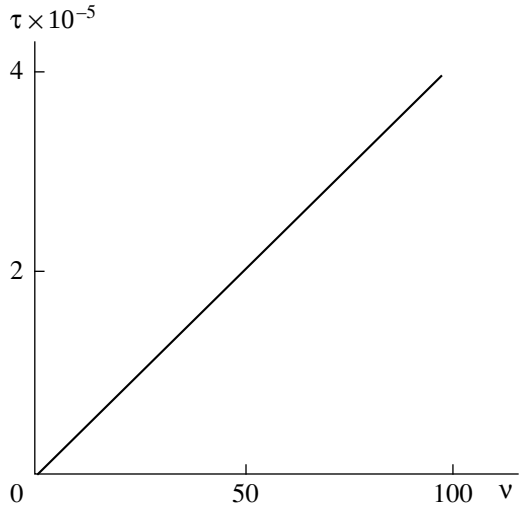


Fig. 2. Dimensionless time of tenfold increase in the initial amplitude of spheroidal deformation of a drop versus the drop's dimensionless viscosity, calculated numerically by using (8) with $\alpha = 1/3$ and $X_0 = 10^{-6}$.

required for an initial deformation to increase tenfold. After this time period, the curve $t(X)$ levels out.

The dependence of the characteristic growth time T_ν of a spheroidal deformation of a drop that is unstable with respect to intrinsic charge on the dimensionless viscosity is represented by the simple linear graph in Fig. 2, valid for both low and high values of viscosity. The dependence T_ν on ν shown in Fig. 2 can be analytically approximated by the expression

$$T_\nu = T_0 + g\nu, \tag{11}$$

where T_0 is characteristic dimensionless time of the instability development for an inviscid liquid drop and g is a numerical factor.

For $\alpha = 1/3$, $\rho = 10^{-3}$ kg/m^3, $X_0 = 10^{-6}$, and $\sigma = 0.07N/m$, numerical calculations predict $T_0 \approx 0.0684$ and $g \approx 4035$. The linear behavior of $T_\nu(\nu)$ is in agreement with the asymptotic formula (10), which is valid only for large values of viscosity.

CONCLUSIONS

The predicted effect of viscosity on the growth rate of the spheroidal deformation amplitude for a drop that is unstable with respect to its intrinsic charge is quite obvious: it is a priori clear that the viscosity must increase the characteristic time of instability development. However, we provided the first quantitative estimate for the viscosity effect on the essentially nonlinear growth of a very small thermal deformation of the drop geometry at the stability threshold. It is interesting that, despite the nonlinearity, the characteristic time of insta-

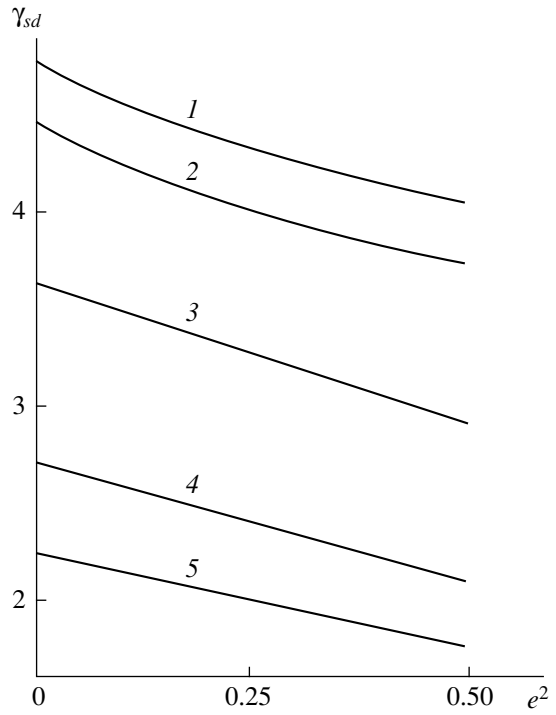


Fig. 3. Principal-mode ($n = 2$) instability growth rate $\gamma_{sd} = \gamma_{sd}(e^2)$ calculated for a charged spheroidal drop versus squared eccentricity by using (1A) for $W = 4$ and $\nu = (1)$ 0.03, (2) 0.1, (3) 0.36, (4) 0.8, and (5) 1.2.

bility development turns out to be a linear increasing function of viscosity.

APPENDIX

It is known from [9, 10] that, when a spherical conducting drop carries a charge slightly greater than the critical one, the drop is unstable and the amplitude of the principal mode of capillary oscillation grows exponentially, which corresponds to the drop's elongation into a spheroid. This process is accompanied by charge redistribution over the drop surface, with its surface density increasing at the apexes and higher modes of the drop's capillary oscillation becoming unstable. The instability growth rate γ_n of any mode of a perfectly conducting drop depends on the Rayleigh parameter W, dimensionless viscosity ν, and eccentricity of the spheroidal drop (the extent of its elongation) [11]. This is illustrated by Fig. 3, where growth rate γ_2 of the principal mode of a spheroidal drop is shown as a function of e^2 calculated for constant W and various ν by using the following dispersion equation obtained in [11]:

$$
\begin{aligned}
&s[s^2 + n(n-1)(n+2)\alpha_n] + 2\nu\Big[s^2(n-1)(2n+1) \\
&\quad - \sqrt{\frac{s}{\nu}}f_n\Big(\sqrt{\frac{s}{\nu}}\Big)(s^2 + n(n-1)(n+2)\alpha_n)\Big] \\
&\quad - 4\nu^2 n(n-1)(n+2)s\sqrt{\frac{s}{\nu}}f_n\Big(\sqrt{\frac{s}{\nu}}\Big) + e^2\kappa_n \\
&\quad \times\Bigg\{s\Big[(s^2 + n(n-1)(n+2)\alpha_n)\Big(2(n-1) \\
&\quad + \sqrt{\frac{s}{\nu}}f_n\Big(\sqrt{\frac{s}{\nu}}\Big)\Big) - 3((2n-1)(n+2)\alpha_n + n^3)\Big] \\
&\quad + 2\nu\Big[s^2\Big(2n^3 - 8n^2 + 4n - 10 + \frac{9}{n}\Big) \\
&\quad + \sqrt{\frac{s}{\nu}}f_n\Big(\sqrt{\frac{s}{\nu}}\Big)\Big(s^2\Big(2n^2 - n + 2 + \frac{9}{n(n+1)}\Big) \\
&\quad + 3\Big(2n^2 - n + 3 - \frac{6}{(n+1)}\Big)(n+2)\alpha_n + 3n^3\Big)\Big] \\
&\quad + 4\nu^2 s\sqrt{\frac{s}{\nu}}f_n\Big(\sqrt{\frac{s}{\nu}}\Big)\Big[5n^3 + 5n^2 + 2n - \frac{18}{(n+1)}\Big]\Bigg\} = 0.
\end{aligned}
\tag{1A}
$$

$$
\kappa_n \equiv \frac{n(n+1)}{3(2n-1)(2n+3)}, \quad W = \frac{Q^2}{16\pi\sigma R^3},
$$

$$
\alpha_n = 1 - \frac{W}{W_{n*}}, \quad W_{n*} \equiv \frac{n+2}{4},
$$

$$
f_n(x) \equiv i_{n+1}(x)/i_n(x), \quad x \equiv (s/\nu)^{1/2}.
$$

Here, the following notation is used: $s = \mathrm{Re}s + i\mathrm{Im}s$ is the complex frequency; i is the imaginary unit; n is the number of a capillary oscillation mode; σ is surface tension; $i_n(\zeta)$ is the modified spherical Bessel function of the first kind with a complex argument, defined as

$$
\begin{aligned}
i_n(\zeta) &= \sqrt{\frac{\pi}{2\zeta}}I_{n+1/2}(\zeta) \\
&= \begin{cases}
\exp(-0.5n\pi i)j_n(\zeta\exp(0.5\pi i)) \\
\text{at} \quad (-\pi < \arg\zeta \le 0.5\pi), \\
\exp(1.5n\pi i)j_n(\zeta\exp(-1.5\pi i)) \\
\text{at} \quad (0.5\pi < \arg\zeta \le \pi),
\end{cases} \\
n &= 0, \pm1, \pm2, \ldots,
\end{aligned}
$$

where the expression

$$
f_n(\zeta) = \sqrt{\frac{\pi}{2\zeta}}J_{n+1/2}(\zeta)
$$

represents the spherical Bessel function of the first kind; $I_{n+1/2}(\zeta)$ and $J_{n+1/2}(\zeta)$ are, respectively, the modified and usual Bessel functions of half-integer order; and W is the Rayleigh parameter characterizing the drop's stability with respect to its intrinsic charge [2].

If the drop charge is greater than the critical one, i.e., $W > W_{n*}$, then the nth-mode frequency s_n is imaginary and the nth-mode amplitude grows exponentially.

Equation (1A) is presented here in dimensionless form (with $R = 1$, $\rho = 1$, and $\sigma = 1$). When $e^2 = 0$, the dispersion equation (1A) of capillary motion in a charged viscous spheroidal drop reduces to the equation of capillary motion in a charged viscous spherical drop derived in [8].

When the charge is slightly supercritical (i.e., the value of $|\alpha_n|$ is small), the growth rate increases with the eccentricity. This is the case when a drop exhibiting self-charge instability carries Rayleigh's limit charge, as in the experiments on Rayleigh's criterion verification (see [1, 2] and references therein). Then, the exponential elongation of the drop is initiated by a thermal fluctuation of its geometry proportional to the Legendre polynomial $P_2(\cos\Theta)$ [3, 4].

Let us analyze the effects of viscosity, charge, and eccentricity on the instability growth rate using the dispersion equation above. These effects cannot be described in analytical form by using the dispersion equation (1). For this reason, we derive the desired analytical results by approximating numerical results.

Figure 4 shows the real part of the complex frequency s of the principal mode ($n = 2$) of capillary oscillation of a spherical drop (i.e., the instability growth rate γ_g) as a function of dimensionless viscosity ν plotted for several supercritical values of the Rayleigh parameter W ($W > W_{2*}$). In the domain of W values that are supercritical for drop instability, the curves shown in this figure can be approximated by the analytical formula

$$\gamma_g = \frac{\gamma_0}{1 + \nu k(W)}, \qquad k(W) = \left[1 + \frac{4}{W+1} + \frac{1}{2(W-4)}\right], \tag{2A}$$

where the expression

$$\gamma_0 = \left(\frac{2\sigma}{\rho R^3}\right)^{1/2} (W - W_{2*})^{1/2}$$

represents the instability growth rate for an inviscid liquid drop. Thus, viscosity reduces the instability growth rate γ_s compared to the growth rate γ_0 for an inviscid liquid drop by the viscous damping rate n_g:

$$\gamma_g = \gamma_0 - \eta_g, \qquad \eta_h = \gamma_0\left\{1 - \frac{1}{1 + \nu k(W)}\right\}. \tag{3A}$$

In the asymptotic limit of a low-viscosity drop, when $\nu k(W) < 1$, (2A) can be rewritten as

$$\gamma_g \approx \gamma_0[1 - \nu k(W)],$$

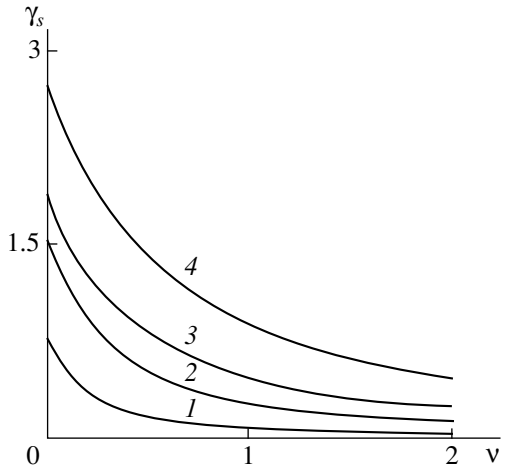


Fig. 4. Principal-mode ($n = 2$) instability growth rate γ_s calculated numerically for a charged spherical drop as a function of its dimensionless viscosity ν for supercritical $W =$ (*1*) 1.1, (*2*) 1.3, (*3*) 1.5, and (*4*) 2.

and (3A) takes the form

$$\eta_g = \gamma_0 \nu k(W). \tag{4A}$$

Figure 3 shows the principal-mode growth rate γ_{gd} for a charged spheroidal drop of a viscous conducting liquid as a function of the squared eccentricity e^2, calculated for several values of viscosity and constant supercritical charge ($W = 4$). In the linear approximation with respect to e^2, these curves can be approximated by the expression

$$\gamma_{gd} = \gamma_g\left[1 - \frac{2}{1 + 0.8\nu}e^2\right], \tag{5A}$$

where γ_g is given by (2A). For low viscosity, (5A) takes the form

$$\gamma_{gd} \approx \gamma_0[1 - \nu k(W)][1 - 2e^2(1 - 0.8\nu)].$$

Using (5A), we express the viscous damping rate (3A) of the principal-mode oscillation of a spherical conducting drop as

$$\eta_{gd} \equiv \gamma_0 - \gamma_{gd} = \gamma_0 - \frac{\gamma_0}{1 + \nu k(W)}\left[1 - \frac{2}{1 + 0.8\nu}e^2\right].$$

For a low-viscosity drop, this expression reduces to

$$\eta_{gd} \approx \eta_g[1 - 2e^2(1 - 0.8\nu)],$$

where η_g is from (4A).

REFERENCES

1. A. I. Grigor'ev and S. O. Shiryaeva, Zh. Tekh. Fiz. **61** (3), 19 (1991) [Sov. Phys. Tech. Phys. **36**, 258 (1991)].
2. A. I. Grigor'ev and S. O. Shiryaeva, Izv. Akad. Nauk, Mekh. Zhidk. Gaza, No. 3, 3 (1994).

3. S. O. Shiryaeva and I. D. Grigor'eva, Pis'ma Zh. Tekh. Fiz. **20** (6), 1 (1994) [Tech. Phys. Lett. **20**, 214 (1994)].
4. S. O. Shiryaeva, A. I. Grigor'ev, and I. D. Grigor'eva, Zh. Tekh. Fiz. **65** (9), 39 (1995) [Tech. Phys. **40**, 885 (1995)].
5. A. I. Grigor'ev, Pis'ma Zh. Tekh. Fiz. **24** (24), 36 (1998) [Tech. Phys. Lett. **24**, 962 (1998)].
6. D. F. Belonozhko and A. I. Grigor'ev, Pis'ma Zh. Tekh. Fiz. **25** (15), 41 (1999) [Tech. Phys. Lett. **25**, 610 (1999)].
7. S. O. Shiryaeva and A. I. Grigor'ev, Pis'ma Zh. Tekh. Fiz. **21** (16), 17 (1995) [Tech. Phys. Lett. **21**, 639 (1995)].
8. S. O. Shiryaeva, M. I. Munichev, and A. I. Grigor'ev, Zh. Tekh. Fiz. **66** (7), 1 (1996) [Tech. Phys. **41**, 635 (1996)].
9. Lord Rayleigh (J. W. Strutt), Philos. Mag. **14**, 184 (1882).
10. A. I. Grigor'ev, Zh. Tekh. Fiz. **55**, 1272 (1985) [Sov. Phys. Tech. Phys. **30**, 736 (1985)].
11. S. O. Shiryaeva, Zh. Tekh. Fiz. **68** (4), 20 (1998) [Tech. Phys. **43**, 366 (1998)].

Translated by V. Gursky

Technical Physics, Vol. 45, No. 9, 2000, pp. 1135–1140. Translated from Zhurnal Tekhnicheskoĭ Fiziki, Vol. 70, No. 9, 2000, pp. 37–43.

Study of Directed Ion Velocities in a Vacuum Arc by an Emission Method

A. S. Bugaev, V. I. Gushenets, A. G. Nikolaev, E. M. Oks, and G. Yu. Yushkov

High Current Electronics Institute, Siberian Division, Russian Academy of Sciences, Akademicheskiĭ pr. 4, Tomsk, 634055 Russia

E-mail: gyushkov@hcei.tomsk.ru

Received June 22, 1999

Abstract—Directed ion velocities in a vacuum arc discharge plasma are measured on the basis of a study of the ion emission current response to a rapid change of arc current. It is shown that these velocities are about 10^6 cm/s, are determined by the cathode material, and are almost independent of the ion charge number. Applying a magnetic field results in an increase in the directed ion velocity. As the gas pressure increases, the directed ion velocity decreases; this is the only case where the directed velocities are observed to depend on the ion charge number. © *2000 MAIK "Nauka/Interperiodica".*

INTRODUCTION

Cathode spots occurring in a vacuum arc discharge emit a plasma that, being a good conductor, lets the discharge current flow through the gap between the cathode and anode. The main parameters and applications of this type of discharge are described in detail in [1–5]. Although the vacuum arc discharge has long been actively studied and is commonly used in high-current switches and plasma–ion technology, there is no generally accepted detailed physical concept of the mechanism for such a discharge. As is known [1], the electron energy distribution in a vacuum arc plasma is close to Maxwellian, with a temperature of T_e = 3–6 eV. The ion plasma component is determined by the cathode material; the ion charge number is Q = 1–6 [2]. The ions move from the cathode spots with directed velocities of $v_i \sim 10^6$ cm/s. A number of papers were devoted to studying the dependence of the ion velocity v_i on the ion charge number Q. In [6–8], it was shown that v_i increases with increasing Q as $v_i \propto Q^\alpha$, where $0.5 < \alpha < 1$. In [8], this behavior was explained by the concept of ion acceleration due to both gas-dynamic effects and the passage of ions through the so-called "potential hump," which presumably exists in the cathode region of a vacuum arc. According to the former mechanism, the directed ion velocity does not depend on the ion charge number, whereas the electrostatic acceleration in the potential hump should result in an increase in the directed velocity v_i with increasing Q. According to [9], in some cases, the velocities of ions emitted from cathode spots depend on Q. The authors explain this fact by ion acceleration in the potential hump, whereas, in other cases, the ion velocities are independent of Q, which is explained by acceleration due solely to gas-dynamic effects. Therefore, at present, the available experimental data concerning the influence of the ion charge number on the directed ion velocity in a vacuum arc are rather contradictory. As a result, the same physical process, namely, ion acceleration in the cathode region of a vacuum arc discharge, has been explained on the basis of quite different (although not mutually exclusive) hypotheses or even their combinations.

This study is devoted to the experimental investigation of directed ion velocities in the plasma of a vacuum arc discharge. A distinctive feature of the study is the use of the emission technique to diagnose a vacuum arc.

EXPERIMENTAL TECHNIQUE

A schematic of the experimental setup is presented in Fig. 1. The design and operating principles of the discharge circuit are the same as in [10, 11]. A trigger pulse from a power supply *7* with a duration of 40 μs and current of 20 A was applied to the gap between the cathode *1* and the enveloping trigger electrode *2*. Applying the trigger pulse led to the breakdown between electrodes *1* and *2* over the end surface of an insulator *3* separating the electrodes. The discharge plasma on the dielectric surface initiated a pulse vacuum arc between the cathode *1* and anode *4*. The duration (about 500 μs) and current (100–300 A) of the arc discharge were determined by the power supply *9*. During the vacuum arc discharge, a plasma of the cathode material emitted from cathode spots occupied the anode cavity. We investigated cathodes made from C, Mg, Al, Ni, Ti, Zr, Nb, Cu, Pb, or Bi. On the end surface of the anode *4*, at a distance of 12.6 cm from the cathode, there were holes for extracting the ions. The ions were extracted and accelerated by a multiaperture accelerating–decelerating electrode system *6* fed with a dc voltage of 10–25 kV from the power supply *10*. Next, the accelerated ions entered a time-of-flight mass spectrometer with a 1.2-m base. The operating princi-

1063-7842/00/4509-1135$20.00

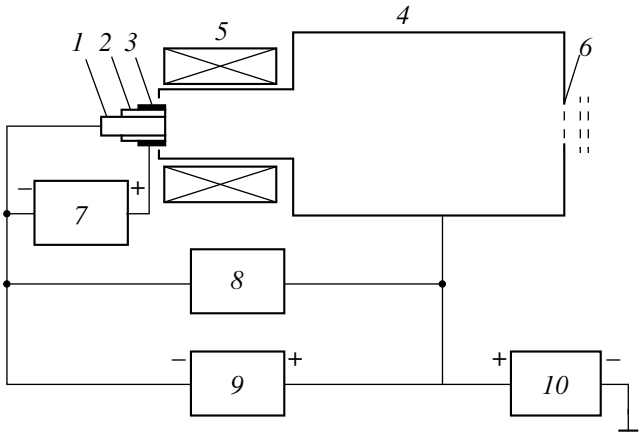


Fig. 1. Schematic of the experimental setup: (*1*) cathode, (*2*) trigger electrode, (*3*) ceramic insulator, (*4*) anode, (*5*) magnetic coil, (*6*) multiaperture acceleration–deceleration electrode system, (*7*) trigger-pulse power supply, (*8*) switch or jump-current power supply, (*9*) vacuum-arc power supply, and (*10*) accelerating voltage source.

ples, design, and parameters of the mass spectrometer are described in detail in [12, 13]. The discharge chamber and mass spectrometer were pumped out to $P \sim (0.8–1.5) \times 10^{-5}$ torr with a 500-l/s turbopump. To test the mass spectrometer, the measured ion charge composition was compared with the results from other experiments. Thus, our results are in excellent agreement with the well-known and reliable data of Brown [14] (see Table 1).

The emission methods for investigating a vacuum arc discharge that are used in this study are based on examining the response of the ion emission current from the plasma boundary to a short-time influence of external factors on the vacuum arc. Such a factor can be a rapid change in the arc current, which leads to either the disappearance or the additional creation of cathode spots. In this study, we either short-circuited the vacuum-arc power supply with a high-speed switch (short-current mode) or applied an additional current pulse to the discharge gap (jump-current mode). In the first case, the vacuum arc current dropped from 100–300 A to zero in 1.4 μs and the discharge voltage dropped from U_{arc} = 25–50 V, which is typical of vacuum arcs, to less than 10 V, which is insufficient to maintain the discharge. In the second case, the current increased by 100–800 A in 2–4 μs. The power supply used to form the additional current pulse included a 0.1-μF capacitor, which discharged through the gap between the cathode *1* and anode *4* when a control pulse was applied to the high-voltage switch. Applying an additional current pulse resulted in an increase in the discharge voltage to 100 V or even higher. Then, in about 5 μs, the discharge voltage decayed exponentially to the usual U_{arc} value. Both jump- and short-current modes were switched on within 150–200 μs after the ignition of a vacuum arc, when the basic parameters of the arc had already reached their steady-state values. The response of both the ion emission current and ion charge composition to this switching was somewhat delayed. The time delay was determined mainly by the time required for the ions to cover the distance from the cathode to the emissive plasma boundary. Thus, the directed ion velocity was found from the time required for ions to cross the discharge gap.

ANALYSIS OF THE EXPERIMENTAL RESULTS

Figure 2 presents the time evolution of the ion current I_i through the vacuum arc after switching off the power supply (short-current mode) for Al and Bi cathodes. It is worth noting that, for ions with different charge numbers, the currents vary in proportion to each other and the time evolutions are almost the same. Similar results were obtained for all the cathode materials we investigated. When a magnetic field from a short magnetic coil *5* was applied to the discharge gap (Fig. 1), the ion current decayed more rapidly (Fig. 3). It is seen from Fig. 3 that, in the presence of a magnetic

Table 1. Comparison of the measured charge composition of a vacuum arc plasma with the data from [14]

Cathode material	Data from [14]						Our measurements						Comparison
	charge number composition, %					average charge number $\langle Q \rangle$	charge number composition, %					average charge number $\langle Q_m \rangle$	$\frac{\langle Q_m \rangle - \langle Q \rangle}{\langle Q \rangle}$, %
	1^+	2^+	3^+	4^+	5^+		1^+	2^+	3^+	4^+	5^+		
C	100	–	–	–	–	1.00	98	2	–	–	–	1.04	+4.0
Mg	37	63	–	–	–	1.63	30	70	–	–	–	1.70	+4.3
Ti	6	82	12	–	–	2.06	14	58	27	1	–	2.15	+6.4
Zr	9	55	30	6	–	2.33	8	48	34	10	–	2.46	+5.6
Nb	5	46	37	12	–	2.56	7	40	35	15	3	2.67	+4.3
W	8	34	36	19	3	2.75	7	29	33	25	6	2.94	+6.9
Ta	13	39	28	18	2	2.57	9	40	30	16	5	2.68	+4.2

field, the time evolutions for ions with different charge numbers are almost identical. As in [15], the magnetic field resulted in an increase in both the discharge voltage and the fraction of multiply charged ions in the vacuum arc plasma.

Variations in the distance between the cathode and emissive plasma boundary resulted in proportional variations in the characteristic decay time of the ion current. Thus, reduction of the spacing L_{ca} between the Pb cathode and the emitting electrode from 12.6 to 7.8 cm led to a proportional drop in the time in which the ion current fell by one-half (from 39 to 24 μs for Pb^+ and from 35 to 22 μs for Pb^{2+} ions).

In the jump-current mode, the delay of the ion emission current response was generally determined by the cathode material and the spacing between the cathode and emitting electrode. The time evolution of the ion current through the vacuum arc for a Mg cathode after triggering the jump-current mode is presented in Fig. 4. Note that the currents of ions with different charge numbers reach their maxima simultaneously. This was true for all the cathode materials included in the study.

In both the short-current and jump-current modes, the ion emission current response to a rapid change in the arc current took place after a certain time interval t. Since ionization of the cathode material in a vacuum arc occurs mainly near cathode spots (no more than 1 mm from the cathode surface, which is much less than the spacing L_{ca}) and then the resulting plasma only expands, the time t can be represented as

$$t = t_p + t_{acc} + t_g, \quad (1)$$

where t_p, t_{acc}, and t_g are the time intervals during which ions pass from the cathode to the emissive boundary, cross the accelerating gap, and pass from the accelerating gap to the spectrometer shutter, respectively.

The times t_{acc} and t_g for ions with different charges and masses can be easily calculated from the known length of the accelerating gap, the spacing between the emissive plasma boundary and the mass-spectrometer shutter L_g, and the accelerating voltage U_{acc}. Simple estimates of the time $t_{acc} + t_g$ carried out for different ions and a typical value of U_{acc} (21 kV in most of the experiments) show that the time t is one order of magnitude longer than $t_{acc} + t_g$. Thus, the delay in the ion current response is determined mainly by the time required for ions to cross the discharge gap rather than either the accelerating gap or the distance from the emissive boundary to the spectrometer shutter. Nevertheless, the minor discrepancy between the time evolutions in Fig. 2 may be related to a scatter in $t_{acc} + t_g$ values for ions with different charge numbers.

Thus, from the measured cathode–emitting electrode spacing L_{ac}, for the directed plasma ion velocity v_i, we obtain

$$v_i = L_{ac}/(t - (t_{acc} + t_g)). \quad (2)$$

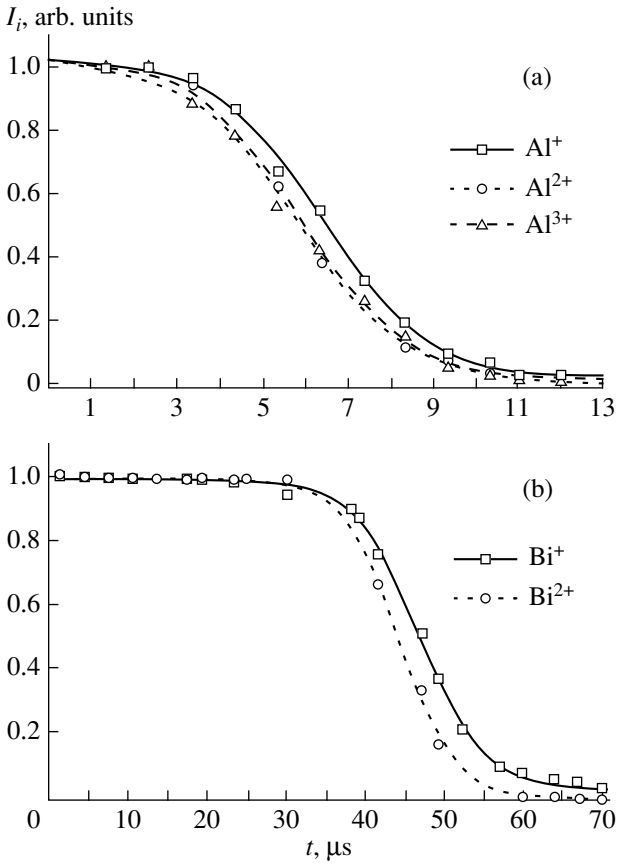


Fig. 2. Time dependences of the ion emission currents in the short-current mode for (a) Al and (b) Bi cathode

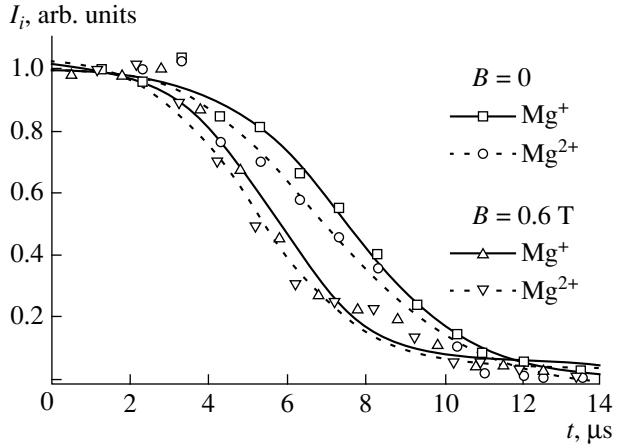


Fig. 3. Same as in Fig. 2 for a Mg cathode with and without magnetic field.

In the jump-current mode, the time t was defined as the time interval between the maxima of the discharge current and ion current I_i. In the short-current mode, the time t was defined as the time interval between the maxima of the time derivatives of those currents. The ion velocities determined by both these methods are

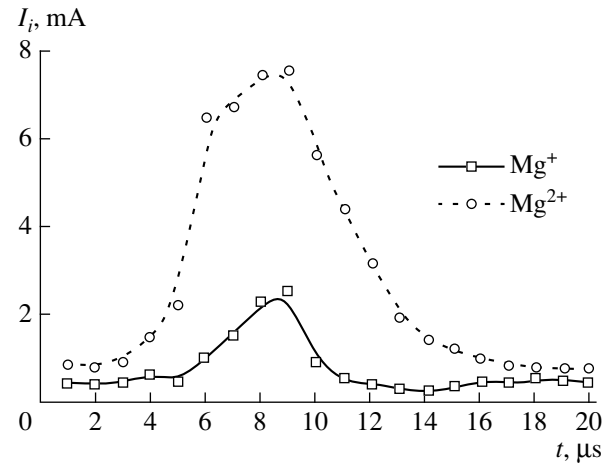


Fig. 4. Time dependences of the ion emission currents in the jump-current mode (600 A, 2 μs).

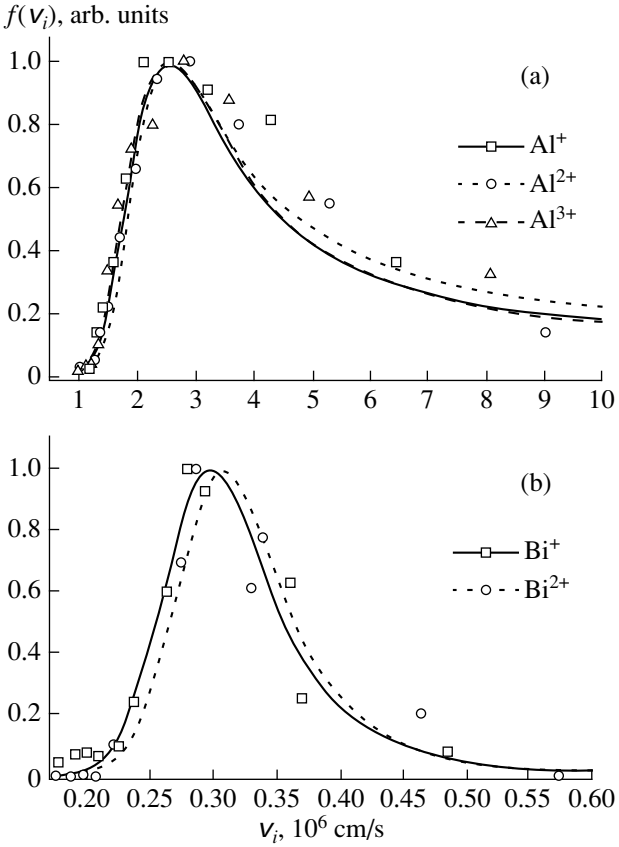


Fig. 5. Ion velocity distribution functions for a vacuum arc with (a) Al and (b) Bi cathodes.

presented in Table 2 for different cathode materials. It is seen that the difference between the directed ion velocities determined by the two methods is relatively small (within ±35%). It is also seen from the table that the light (e.g., Mg) ions are faster than the heavy (e.g., Bi) ions. We point out again the important result that the velocity, and hence the kinetic energy, of the directed motion of the ions of a given element is almost independent of the ion charge number.

A decrease in the distance between the cathode and the emissive boundary results in an almost proportional drop in the ion current decay time. Indeed, taking into account the value of $t_{acc} + t_g$, which is included in the measured t value, we find that a decrease in L_{ca} by a factor of 1.6 (from 12.6 to 7.8 cm) leads to exactly the same decrease in the time in which the Pb^+ ion current falls by one-half. For Pb^{2+} ions, the decay time decreases by a factor of 1.5. This experimental fact is additional evidence that the ion velocity in the expanding plasma of a vacuum arc is constant and that the size of the ionization zone is much less than the characteristic length of the discharge gap. The close values of t for Pb^+ and Pb^{2+} ions show that, for ions of a given element, the ion velocity as a function of the ion charge number ($v_i = f(Q)$) is almost constant.

By differentiating the ion current with respect to time and taking into account the value of $t_{acc} + t_g$ for the ions with different charge numbers, we can obtain the distribution function over the directed ion velocities in the short-current mode. The results of applying this procedure to the dependences from Fig. 2 are shown in Fig. 5. It is seen that, for Al and Bi ions, not only the maxima of the distribution functions, but also the distribution functions themselves match each other fairly well. On the assumption that the ion velocity distribution is isotropic, we find that the kinetic energy is equally shared among ions with different charge numbers. Note that the distribution function obtained in this way decreases with increasing ion velocity much more slowly than the Maxwellian distribution.

Applying a magnetic field to the discharge gap results in an increase in the arc voltage U_{arc} and the appearance of ions with a higher charge number in the plasma. In this case, however, the decay time of the ion current in the short-current mode and the time in which the ion emission current reaches its maximum in the jump-current mode both decrease. Figure 6 shows the dependence of the measured ion velocity v_i for different materials on the arc voltage U_{arc}. For each specific material, the leftmost point in Fig. 6 corresponds to the usual arc voltage U_{arc} without a magnetic field, whereas the rightmost point corresponds to the increased value of U_{arc} when a magnetic field is applied. It is seen that the directed ion velocity increases with increasing arc voltage. At the same time, the fact that, in the presence of a magnetic field, the values of t for ions with different charge numbers are almost the same indicates that the increase in the ion velocities is independent of Q. Thus, the additional energy supplied to the cathode region of a vacuum arc due to increasing arc voltage is equally shared among the ions with different charge numbers; i.e. the velocity, and hence the kinetic energy,

of each ion increases by approximately the same quantity independently of the ion charge.

The above results do not fully agree with the results of [7], where the directed ion velocity v_i was observed to depend on the ion charge number Q. It was suggested in [7] that the cathode drop in the potential of a vacuum arc plasma was followed by a "potential hump" that was higher than the arc voltage. The ions moving from the cathode region of the vacuum arc acquired directed velocities due to electrostatic acceleration when sliding down this hump. Clearly, the kinetic energy of an ion in this case is proportional to its charge number Q, and, on the assumption that the ion motion is collisionless, the ion velocity is proportional to $Q^{0.5}$.

In experiments [7], the duration of the vacuum arc discharge was 0.5 s and a 20-l/s magnetic-discharge pump was used to evacuate the discharge gap. In this case, because of the long duration of the discharge and the relatively low pumping rate (which decreases even more with decreasing gas pressure), the actual pressure in the system during the arc discharge might have been much higher than that reported in [7]. In order to verify this hypothesis, we carried out measurements of the ion velocities at an elevated gas pressure in the discharge gap. The experiments showed that the increase in the gas pressure in the discharge gap during forced gas puffing led to an increase in the time in which the ion current decayed (see Fig. 7). In this case (as in [16]), the fraction of multiply charged ions decreased. The effect of a decrease in the fraction of multiply charged ions with increasing pressure was thoroughly studied in [17]. At the same time, the influence of an increase in pressure on the above dependences was different for ions with different charge numbers. Thus, for the ions with lesser charge numbers, the increase in the time in which the ion current decayed was more pronounced than for the ions with greater charge numbers. As an example of the influence of an elevated gas pressure on the ion velocities, we present the time dependence of the ion current in the short-current mode for a copper cathode during forced argon puffing into the discharge system (Fig. 8).

The shift of the dependences in Fig. 7 toward longer times is clear evidence of a decrease in the directed ion velocities. The dependences presented in Fig. 8 show that the time in which the ion current falls by one-half is 9 μs for doubly charged copper ions, whereas for singly charged ions, this time is 7 μs. The figure also presents the dependences calculated using the distribution function over directed velocities of copper ions obtained in [4], based on an analysis of data from [7]. One can see the tendency toward a disproportional change in the velocities of singly and doubly charged ions with increasing gas pressure. As the pressure increases, the time dependences of the ion current approach those described in [4]. The above influence of an increase in the pressure on the velocities of ions with different charge numbers was confirmed by measure-

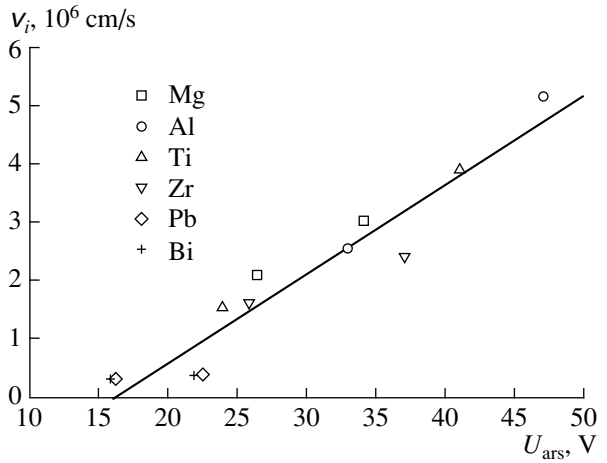


Fig. 6. Ion velocity vs. the vacuum arc voltage.

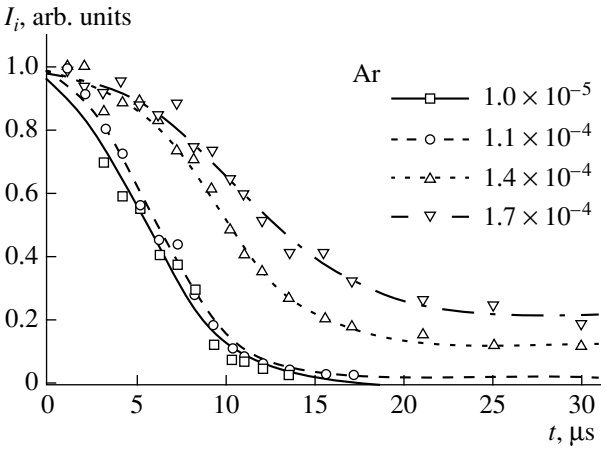


Fig. 7. Time dependences of the current of Al^+ ions in the short-current mode for an Al cathode and different pressures of the puffed gas (Ar).

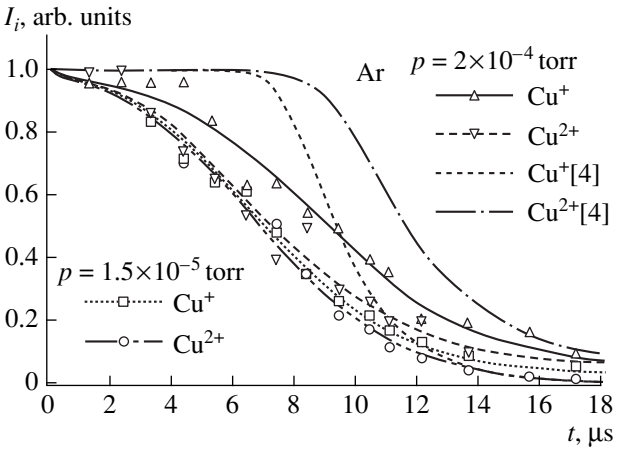


Fig. 8. Comparison of the experimental time dependences of the ion current in the short-current mode with those calculated using the data from [4].

Table 2. Directed ion velocities v_i in a vacuum arc measured in (1) the jump-current and (2) short-current modes

Ions	Q	v_i, 10^6 cm/s (1)	v_i, 10^6 cm/s (2)	Velocity ratio
Mg	1+	1.8	2.0	0.90
	2+	1.7	2.1	0.81
Al	1+	1.8	2.5	0.72
	2+	1.8	2.6	0.69
	3+	1.6	2.5	0.64
Ti	1+	1.6	1.5	1.07
	2+	1.5	1.4	1.07
	3+	1.4	1.5	0.93
Pb	1+	0.5	0.4	1.25
	2+	0.5	0.4	1.25
Bi	1+	0.4	0.3	1.33
	2+	0.4	0.3	1.33

ments of v_i in the jump-current mode. Unfortunately, in our case, because of the requirements related to the electric strength of the accelerating gap, the gas pressure was limited to a value of $(4–7) \times 10^{-4}$ torr. However, the fact of a disproportional change in the velocities of ions with different charge numbers was established for different cathode materials and different gases puffed into the discharge gap. As in [17], in order to observe this effect, the pressure of heavier gases (xenon or argon) can be lower than for lighter gases (nitrogen or air). Therefore, the difference between our results and the results of [7] can be explained by the different gas pressures in the discharge gap of the vacuum arc. At the same time, for a pressure of 10^{-5} torr, the directed ion velocities are observed to depend on the ion charge number.

CONCLUSIONS

(1) The directed ion velocity in a vacuum arc discharge plasma can be deduced from the time evolution of the parameters of the ion current extracted from the discharge plasma after a rapid change in the arc current.

(2) The directed ion velocity is on the order of 10^6 cm/s, is determined by the cathode material, and is almost the same for ions with different charge numbers.

(3) When a magnetic field is applied to the vacuum arc, the directed ion velocity increases, which correlates with an increase in the arc voltage.

(4) An increase in the pressure of the residual gas in the discharge gap of a vacuum arc results in a decrease in the directed ion velocity; this decrease is more pronounced for ions with lower charge numbers.

ACKNOWLEDGMENTS

This work was supported in part by the Russian Foundation for Basic Research (project no. 99-02-18163) and a grant for young scientists from the Siberian Division of the Russian Academy of Sciences.

REFERENCES

1. I. G. Kesaev, *Cathode Processes in Electric Arc* (Nauka, Moscow, 1968).
2. G. A. Mesyats, *Ectons* (Nauka, Yekaterinburg, 1994).
3. *Handbook of Vacuum Arc Science and Technology*, Ed. by R. L. Boxman, D. M. Sanders, and P. I. Martin (Noyes, Park Ridge, New York, 1995).
4. *Vacuum Arcs: Theory and Application*, Ed. by J. M. Lafferty (Wiley, New York, 1980; Mir, Moscow, 1982).
5. G. A. Mesyats and D. I. Proskurovskiĭ, *Pulsed Electric Discharge in Vacuum* (Nauka, Novosibirsk, 1984).
6. V. M. Lunev, V. G. Padalka, and V. M. Khoroshikh, Zh. Tekh. Fiz. **47**, 1491 (1977) [Sov. Phys. Tech. Phys. **22**, 858 (1977)].
7. W. D. Davis and H. C. Miller, J. Appl. Phys. **40**, 2212 (1969).
8. H. C. Miller, J. Appl. Phys. **52**, 4523 (1981).
9. K. Tsuruta, K. Sekiya, O. Tan, and G. Watanabe, in *Proceedings of the XVII International Symposium on Discharges and Electrical Insulation in Vacuum, Berkeley, 1996*, Vol. 1, p. 181.
10. E. M. Oks, I. G. Brown, M. R. Dickinson, and R. A. MacGill, Rev. Sci. Instrum. **67**, 959 (1996).
11. E. M. Oks, I. G. Brown, M. R. Dickinson, *et al.*, Appl. Phys. Lett. **67**, 200 (1995).
12. I. G. Brown, J. E. Galvin, R. A. MacGill, and R. T. Wright, Rev. Sci. Instrum. **58**, 1589 (1987).
13. I. G. Brown and X. Godechot, IEEE Trans. Plasma Sci. **19**, 713 (1991).
14. I. G. Brown and J. E. Galvin, IEEE Trans. Plasma Sci. **17**, 679 (1989).
15. E. M. Oks, A. Anders, I. G. Brown, *et al.*, IEEE Trans. Plasma Sci. **24**, 1174 (1996).
16. E. M. Oks, B. H. Wolf, P. Spadtke, and H. Emig, in *Proceedings of the Beijing Workshop on Mevva Ion Sources and Application, Beijing, 1993*, p. 24.
17. P. Spadtke, H. Emig, B. H. Wolf, and E. M. Oks, Rev. Sci. Instrum. **65**, 3113 (1994).

Translated by N. Ustinovskiĭ

Technical Physics, Vol. 45, No. 9, 2000, pp. 1141–1146. Translated from Zhurnal Tekhnicheskoĭ Fiziki, Vol. 70, No. 9, 2000, pp. 51–56.

SOLIDS

Radiation-Induced Conductivity of Alkali Halide Crystals in Strong Electric Fields under X-ray and Photoexcitation

V. D. Kulikov and Yu. V. Lisyuk

Tomsk Polytechnical Institute, pr. Lenina 30, Tomsk, 634004 Russia

Received January 8, 1998; in final form, December 1, 1999

Abstract—A study is undertaken into radiation-induced conductivity of alkali halide crystals under X-ray excitation and sequential excitation with X-ray and laser pulses within the absorption band of *F*- and F^--centers. The basic conduction parameters (concentration and lifetime of carriers upon X-ray and photoexcitation) are estimated. The possible processes responsible for the nonlinearity of the current–voltage characteristics are discussed. It is shown that an increase in the conductivity in strong electric fields may be due to a decrease in the spatial localization of electrons in the conduction band of the insulator. *© 2000 MAIK "Nauka/Interperiodica".*

INTRODUCTION

Investigations into radiation-induced conductivity provide important information on the mechanisms of generation, transfer, and recombination of carriers. This information is needed for a deeper insight into the behavior of insulator materials under ionizing radiation. Of particular interest is the study of radiation-induced conductivity in strong electric fields, which, as was shown experimentally, exhibits substantial deviations from Ohm's law [1, 2]. The origin of the nonlinear behavior of the current–voltage characteristics still remains unclear. Some authors [1] suppose that the nonlinearity is due to the presence of nonohmic contacts. According to [2, 3], the superlinear behavior of the current–voltage characteristics of crystalline quartz, glasses, and alkali halide crystals under X-ray irradiation is associated with the Onsager effect. Moreover, the above nonlinearity can stem from phenomena such as charge carrier injection from metallic electrodes, electrostatic and collisional ionization of color centers in ionic crystals, and variations in the carrier mobility in electric fields.

The purpose of the present work was to investigate the nonlinearity of current–voltage characteristics in alkali halide crystals at 10^4–10^5 V/cm. The alkali halide crystals were chosen for our study owing to the known mechanisms of radiation-induced carrier generation and defect formation. We examined X-ray induced conductivity and the photoconductivity induced by laser excitation of *F* and F^- color centers (the *F*- and F^--centers are the positively charged anionic vacancies with one and two trapped electrons, respectively). Two excitation methods were appropriate for use in our experiments by the following reasoning. First, X-ray excitation produces oppositely charged carrier pairs coupled with each other through the Coulomb interaction. The Onsager effect reduces the probability of their recombination due to the spatial separation of carriers in the external electric field. Upon photoexcitation, a similar situation can arise when the ionization of an *F*-center creates a photoelectron and a positively charged anionic vacancy. A different situation is observed upon excitation of an F^--center, when the formation of a photoelectron and a neutral *F*-center prevents the Onsager effect. Thus, application of two excitation methods makes it possible to determine how the Onsager mechanism affects the behavior of current–voltage characteristics. Second, it should be expected that the mechanism of electron–hole recombination upon X-ray excitation differs from the mechanism of photoelectron recombination at lattice defects under light excitation of color centers. The difference in carrier recombination can be used as a guideline in distinguishing electrons created through band-to-band generation and electrons released from traps. Identification of the electron origin is of prime interest for the determination of the contribution from carriers generated by collisional ionization of color centers to the conductivity.

In this work, we evaluated the lifetime and concentration of carriers under X-ray and photoexcitation. The possible approaches to the estimation of carrier lifetime are of considerable interest, because, in most cases, the lifetime measurements from conductivity decay are limited by instrumental temporal resolution [4].

EXPERIMENTAL TECHNIQUE

The schematic diagram of the experiment is shown in the inset of Fig. 1. We studied single crystals of KBr, KCl, NaCl, KI, and CsI grown from the salt melts (high-purity grade). Thin plates with a cross section of 15×15 mm and a thickness of ≈150–350 μm were used as samples. Platinum electrodes and a guard ring were evaporated onto the samples to exclude the surface current.

1063-7842/00/4509-1141$20.00

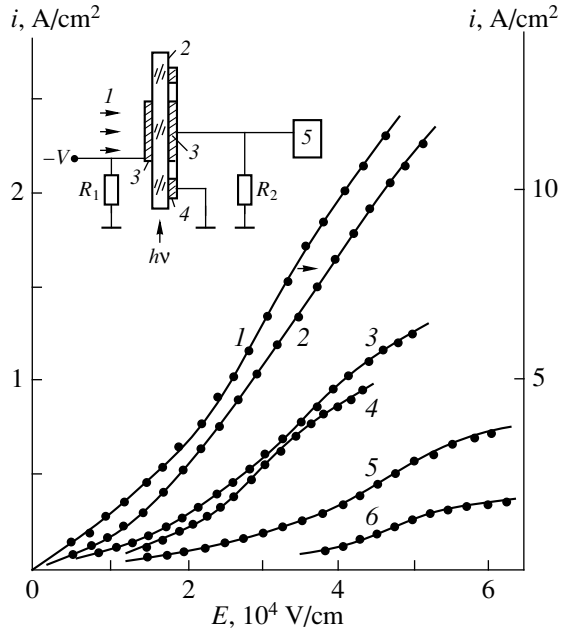


Fig. 1. Current–voltage characteristics of crystals (*1*) CsI, (*2*) KBr, (*3*) KCl, (*4*) KI, (*5*) NaCl, and (*6*) CsI–Tl (X-ray excitation). The inset shows the schematic diagram of the experiment: (*1*) X-ray radiation, (*2*) sample, (*3*) electrodes, (*4*) guard ring, and (*5*) oscillograph.

X radiation was produced by exposure of an Al target about 300 μm thick to a high-current electron beam. The parameters of the electron beam were as follows: maximum electron energy ~0.3 MeV, pulse duration (at half-height) $\tau_i \approx 200$ ns, and current density ≈400 A/cm^2.

Photoexcitation was achieved with a single mode Q-switched YAG : Nd laser. The interelectrode space of the sample was irradiated through a lateral face. The energy fluence of laser radiation at a fundamental wavelength of 1.06 μm at the sample location was equal to 0.2 J/cm^2 (half-height duration $\tau_l \approx 30$ ns). The F^--centers were excited at the fundamental wavelength. For excitation of the *F*-centers, the laser radiation was converted with the use of a lithium niobate crystal into the second harmonic with a wavelength of 0.53 μm. The conversion efficiency was about 10%.

In order to compare the results of X-ray and photoconductivity measurements, it was expedient to perform cascade excitation. For this purpose, the crystal subjected to X-ray coloration with the first pulse was further excited under laser radiation. The effect of cascade excitation on the KBr crystal is illustrated in Fig. 2a by the oscillogram of optical transmission of the sample at a He–Ne laser wavelength (λ = 0.63 μm) (upper inset) and conduction currents. Note that light pulse excitation within the absorption *F*-band results in complete decay of the *F*-centers. The concentrations of *F*- and F^--centers under X-ray irradiation were determined from the optical transmission data (Fig. 2a) according to the Smakula formula.

The current–voltage characteristics were measured in air at room temperature. Negative pulsed voltage *V* with a duration of 20 μs was applied across the sample electrodes (Fig. 1). An electron accelerator was switched on within 8 μs after the voltage application.

The time constant τ_{RC}, which is related to recharging of the sample capacitance with a change in the sample resistance, has been determined as $\tau_{RC} = (R_1 + R_2)C_1$, where $C_1 = 10^{-11}F$ is the interelectrode capacitance of the sample, $R_1 = 100\ \Omega$, and $R_2 = 50\ \Omega$ (Fig. 1a). In this case, τ_{RC} (~1 ns) is less than the pulse rise time of the X and laser radiation (~5 ns) and the signal is not distorted. The current–voltage curves were plotted as the current pulse amplitudes vs. the corresponding voltages.

The absorbed X-ray energy of the samples was determined from the dependence of the luminescence of the CsI crystal on the excitation energy density. The CsI sample was irradiated with electron beams of different powers, and the luminescence intensity was measured. The electron beam energy incident on the sample was measured with an IMO-2N calorimeter. Then, upon X-ray excitation of CsI, the energy *W* absorbed by the sample was determined from the luminescence intensity.

RESULTS

The dependences of the current density *i* and conductivity σ on the electric field strength *E* upon excitation of alkali halide crystals with X and laser radiation are shown in Figs. 1 and 2. It is seen from Fig. 1 that, under irradiation with equal doses, the current density decreases in the series CsI, KBr, KCl, KI, and NaCl. In CsI, the current is one order of magnitude larger than that in the other crystals. However, the conductivity in the CsI–Tl crystal is rather small, which is likely due to the capture of carriers by the activator ions [5]. For KBr and KCl crystals (Fig. 2), the current density upon excitation with green light (λ = 0.53 μm), which ionizes *F*-centers, exceeds the current density upon excitation of the F^--centers with the first harmonic of laser radiation (λ = 1.06 μm) by a factor of three or four.

For all the studied samples, the current–voltage characteristics are significantly nonlinear upon X-ray and photoexcitation. In the range of electric field strength $E \sim 10^4$–10^5 V/cm, four portions could be distinguished in the current–voltage curves. In the first portion (at $E < 10^4$ V/cm), the current–voltage characteristics are almost linear, whereas in the second portion ($E \sim 10^4$–2×10^4 V/cm), a superlinear increase in the current is observed with an increase in the electric field strength ($i \sim E^2$). In the third portion, the dependence of *i* on *E* becomes more gently sloping, and in the fourth portion (at large *E*), the current–voltage charac-

teristics are close to linear again. The threshold field strengths E_m, which correspond to the transition of current–voltage characteristics to the linear current regime, differ slightly for different crystals and fall in the range $(4–6) \times 10$ V/cm. The dependences of the current density on E upon X-ray and photoexcitation (Fig. 2a) are similar in identical materials. Figure 2b demonstrates the conductivity as a function of the field strength. As can be seen, these dependences reach saturation in the range $(1–6) \times 10^4$ V/cm.

The radiation-induced conductivity of KBr and CsI crystals was studied at different absorbed energy densities upon X-ray excitation and photoexcitation within the absorption band of F- and F^--centers in KBr and KCl crystals. The X-ray-induced conductivity is proportional to the excitation energy density; i.e., $\sigma \sim \sqrt{W}$. The photoconductivity excited within the absorption band of F- and F^--centers changes linearly with a change in the laser fluence. The results obtained confirm a fundamental difference in carrier recombination: it follows the quadratic law under X-ray excitation and exhibits a linear behavior under photoexcitation. This is in reasonable agreement with the available data [6, 7].

DISCUSSION

According to [5, 8, 9], holes in alkali halide crystals exposed to ionizing irradiation undergo a fast autolocalization (10^{-12} s). These holes and the neighboring atoms form an X_2^--type configuration (X is a halogen atom). This is the so-called V_k-center. The hole mobility is close to zero, and only free electrons contribute to the conductivity, either from their recombination with V_k-centers or by their trapping. The main mechanism of free electron capture in alkali halide crystals is capture in F-centers with the formation of F^--centers.

The concentration and lifetime of carriers can be determined from the experimental data in the following way. The electron concentration in the conduction band of alkali halide crystals under homogeneous excitation with X and laser radiation in the bulk is given by the equation

$$\partial n_{1,2}/\partial t = G_{1,2} - n_{1,2}/\tau_{1,2}, \qquad (1)$$

where n_1 and n_2 are the electron concentrations; G_1 and G_2 are the generation rates of carrier pairs; τ_1 and τ_2 are the lifetimes of electrons under X-ray and photoexcitation, respectively; and t is time.

In the case of X-ray excitation, Eq. (1) holds without regard for electron capture in trapping centers (i.e., at a low concentration of color centers).

In alkali halide crystals at different excitation energy densities, $\tau_1 \sim 10^{-10}–10^{-11}$ s [8], which is considerably less than the X-ray and laser pulse lengths ($\sim 10^{-8}$ s). Therefore, as a first approximation, we can

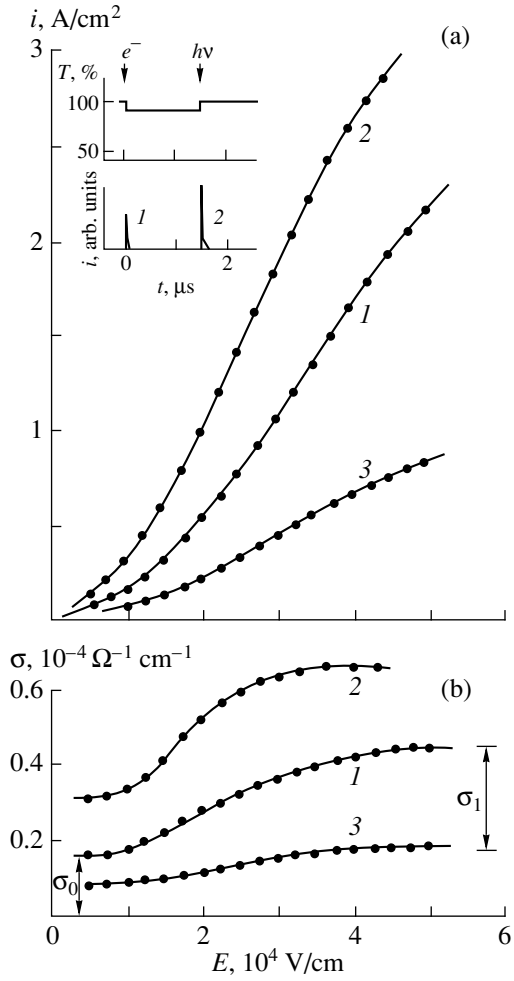


Fig. 2. Dependences of (a) current density and (b) conductivity on the electric field strength upon (*1*) X-ray and laser excitation within the absorption bands of (*2*) F- and (*3*) F^--centers in a KBr crystal. The insets show cascade excitation of a KBr crystal with X-ray and laser (λ = 0.53 μm) pulses and the oscillogram of optical transmission of the F-centers: (*1*) X-ray current and (*2*) photocurrent pulses.

consider the quasi-stationary case and rewrite Eq. (1) in the form

$$n_{1,2} = G_{1,2}\tau_{1,2}. \qquad (2)$$

On the other hand, $n_{1,2}$ can be determined from the experimental current–voltage characteristics by using the relationship

$$i = en_{1,2}\mu E, \qquad (3)$$

where e is electron charge and μ is the mobility.

For X-ray excitation, the generation rate of electron–hole pairs per unit volume can be derived from the condition $G_1 = W/A\tau_i$, where $A = 1.5E_g$ is the energy spent in generating one electron–hole pair in the material (E_g is the band gap). In the case when the KBr sample absorbs the X-ray energy $W = 0.6 \times 10^{-2}$ J/cm^3, μ =

10 $cm^2/(V\ s)$ [5], and $E = 10^4$ V/cm, from relationships (2) and (3), we obtain $n_1 = 1.1 \times 10^{13}\ cm^{-3}$, $G_1 = 3.4 \times 10^{23}\ cm^{-3}\ s^{-1}$, and $\tau_1 = 3 \times 10^{-11}$ s.

Upon photoexcitation in the absorption *F*-band of the KBr sample ($E = 10^4$ V/cm), from formula (3), we have $n_2 = 2.25 \times 10^{13}\ cm^{-3}$. According to the experimental data, X-ray excitation results in the concentration of *F*-centers $N_F = 3 \times 10^{14}\ cm^{-3}$. In the limiting case when all the *F*-centers are assumed to be excited by light, we obtain $G_2 = N_F/\tau_l = 10^{22}\ cm^{-3}\ s^{-1}$ and $\tau_2 = 2 \times 10^{-9}$ s. The values of τ_1 and τ_2 agree well with carrier recombination, which follows the quadratic law upon X-ray excitation and exhibits a linear behavior upon photoexcitation.

The nonlinearity of the current–voltage characteristics can be explained by the mechanism known as the space-charge-limited current [10]. According to this mechanism, a potential difference applied across a thin layer of a high-resistivity sample induces a current in the material due to charge carrier injection from a metallic electrode. In this case, $i \sim E^2$.

In our case (Fig. 1), the metal–insulator contact is nonohmic. The second insulator–metal contact is not electron-blocking. The absence of contact-limited conduction leads to the fact that, at the instant of irradiation, a positively charged layer of holes is formed at the metal–insulator interface due to electrons traveling into the bulk of the sample. Although the external field is screened in the bulk of the insulator, tunnel injection of electrons from the metal can occur in the transition region in large electric fields $\geq 10^6$ V/cm, which, in turn, brings about an increase in the current in the sample.

The degree of spatial inhomogeneity of the electric field in the bulk of the sample can be estimated from the concentration and lifetime of carriers. The thickness of the positively charged layer *d* is limited by the low mobility and short lifetime of carriers $\tau \sim 10^{-10}$ s and is estimated as $d = \mu E\tau = 0.1$ μm ($E = 10^4$ V/cm). At the electron–hole pair concentration $n = 10^{13}\ cm^{-3}$, this layer corresponds to the surface density of holes $N = 10^8\ cm^{-2}$. For a KBr sample 350 μm thick at $E = 10^4$ V/cm, the negative charge density in a metallic electrode is $N_0 = 10^{11}\ cm^{-2}$. Therefore, in our case, the layer of holes compensates for the external field by less than 1%. This suggests that no strong electric fields occur in the sample and the spatial distribution of *E* is nearly uniform.

Electron injection from the contact is disproved by the fact that photoionization of F^--centers should not result in the formation of positive charge layers, because the *F*-centers are electrically neutral. However, the current–voltage characteristics remain nonlinear, even though the *F*-centers are induced without applied voltage. It should be noted that, in the case of the space-charge-limited current mechanism, the voltage is mainly applied across the region of carrier drift. In our case, when carrier injection occurs from the contact and the penetration depth of carriers is considerably less than the sample thickness, the current will be limited by the bulk resistance of the sample.

The experimental dependences observed allow us to exclude the contribution of the Onsager effect to the nonlinearity of the current–voltage characteristics by the following reasoning. First, as noted above, the current–voltage characteristics are similar for all three excitation methods (excitations of *F*-centers, F^--centers, and interband excitation). However, it could be expected that neutrality of the *F*-centers excludes the Onsager effect and results in linearity of the photocurrent upon excitation of F^--centers. Second, according to the estimates [11], the field strength that corresponds to the complete separation of charges due to the Onsager effect is approximately equal to 10^6 V/cm. This value significantly exceeds our experimental estimate when the linear dependence is observed in the field range $4–5 \times 10^4$ V/cm.

Under electric fields higher than 10^4 V/cm, the probability of electrostatic ionization is rather high [12], i.e., a decrease in the energy of ionizing electrons at *F*- and F^--centers due to a decrease in the potential barrier of the center. However, the electrostatic effect should increase the conductivity with an increase in the electric field $\sim \exp(\sqrt{E})$, which disagrees with the behavior of radiation-induced conductivity in electric fields (Fig. 2).

For the current–voltage characteristics $i \sim E^2$, the conductivity increment in the electric field $\Delta\sigma$ can be written as

$$\Delta\sigma \sim E. \tag{4}$$

Relationship (4) for radiation-induced conductivity is similar to that for carrier scattering by polar optical lattice phonons or dipoles [13]. For this type of scattering, the mean free time of carriers is $\tau_p \sim (W_0 + W(E))^{1/2}$, where W_0 is the thermal energy of an electron, $W(E) = el_0$, and *E* is the energy increment of carriers in the electric field (l_0 is the mean free path of an electron between collisions with phonons). At $W_0 \ll W(E)$, we can assume that $\tau_p \sim \sqrt{W} + W(E)/2\sqrt{W_0}$. This result agrees well with the inference that, in ionic crystals, electron scattering by optical lattice phonons predominates if the energies of the carriers are less than 2–3 eV [14, 15].

In the third portion of the current–voltage characteristics, $\sigma(E)$ can be represented by the empirical formula

$$\sigma(E) = \sigma_0 + \Sigma\sigma_j \exp(-b_j/E), \tag{5}$$

where σ_0 is the conductivity in a field less than $E \approx 10^4$ V/cm; b_1 and b_2 are the parameters specifying the slopes of flattened and steep portions of the $\sigma(E)$ curve;

and σ_1 and σ_2 are the conductivity increments in these portions, respectively.

The experimental curves $\sigma(E)$ in $\ln(\sigma(E) - \sigma_0)$ vs. $1/E$ coordinates are depicted in Fig. 3. For KBr, KCl, CsI, and KI crystals, the $\sigma(E)$ dependence is quite well linearized by using the b_1 parameter. For NaCl, CsI–Tl, crystalline quartz, and K208 glass [2], the electric field dependence of the conductivity is more complex and contains two (flattened and steep) portions. Since the slopes of the linear portions for the former group of materials are close to those of the flattened portions for the latter group, these slopes can be characterized by the same parameter b_1.

The b_1 and b_2 parameters for the studied materials are presented in the table. The maximum difference in b_1 for alkali halide crystals is no more than 30%. The b_1 parameter for SiO_2 and the glass deviates from the mean value for alkali halide crystals by ~50%. For X-ray and photostimulated conductivity in KBr and KCl crystals, the b_1 values are equal to within 10%. The dispersion of b_2 for materials of the second group is rather large (approximately three or four times). There is a tendency for b_2 and E_m to increase in going from NaCl and CsI–Tl crystals to oxide compounds such as crystalline quartz and glass. The conductivity increments are $\sigma_1 \approx (0.7–1.5)\sigma_0$ and $\sigma_2 \approx \sigma_0$.

It can be assumed that dependence (5) corresponds to activation-type conductivity. Indeed, in electric fields of 10^4–10^5 V/cm, the field-accelerated carriers, whose mean free path l is more than l_0, likely overcome the energy barrier. The fraction of these electrons $\beta \sim (-l/l_0) \approx (W_b/el_0E)$ [16], where $W_b = elE$ is the activation energy of over-the-barrier motion.

For ionic crystals, according to the data of [5], $l_0 \approx 10a$, where a is the lattice constant. By comparing Eq. (5) and the dependence for β, we obtain $W_{b1,2} = el_0b_{1,2}$. For all the materials, the W_{b1} values are almost identical and are of the order $\sim kT$, even though a and b_1 are somewhat different. The ionization energy W_{b2}, which corresponds to the steep portion of the experimental dependence $\sigma(E)$, considerably exceeds kT in the electric fields $E \sim (4–8) \times 10^4$ V/cm and falls in the range from 0.05 eV for NaCl to 0.19 eV for K208 glass (see table).

The dependence of σ on E can be very similar to that described by formula (5) in the case of carrier release from traps under collisional ionization. According to [12], the collisional ionization coefficient α is proportional to $\exp(-W_i/el_0E)$, where W_i is the ionization energy.

A nonlinear increase in the conductivity in the KBr and KCl crystals can be attributed to ionization of F- and F^--centers. However, the energy, $W_{i1} = 0.025$ eV, is considerably less than the ionization energies of F- and F^--centers in the ground (~1.5 eV) and excited (0.1 eV [8, 9]) states. The ionization mechanism in these crystals is also disproved by the fact that, upon X-ray excitation in strong fields $E \sim 5 \times 10^4$ V/cm, the quadratic character of recombination does not change to the linear recombination which is observed upon light excitation of F^-- and F-centers.

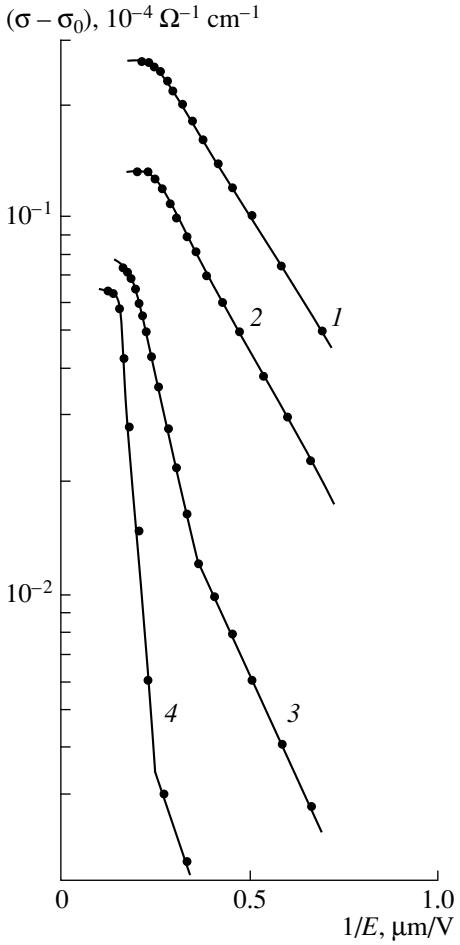


Fig. 3. Dependences of radiation-induced conductivity on the electric field strength: (*1*) KBr, (*2*) KCl, (*3*) NaCl, and (*4*) crystalline quartz.

In our opinion, the above regularities can be interpreted in terms of density-of-states function tails observed both for crystals and (even more probably) for glasses. These tails arise from the modulation of allowed-band edges by fluctuations of the density or internal electric field strengths [12, 17, 18].

As follows from the table, the W_{b1} values for all the materials are close to one another and equal to ~0.025 eV. The energy barrier height is of the same order of magnitude as kT and, most likely, is determined by either the optical phonon energy (0.03 eV) or density fluctuations caused by lattice vibrations [12].

Among materials, such as NaCl, CsI–Tl, crystalline quartz, and K208 glass, in electric fields of $(4–8) \times 10^4$ V/cm, the glass exhibits the highest activation energy ~0.19 eV. This value agrees well with an estimate of 0.2–0.4 eV [18], which was obtained for the

Parameters $b_{1,2}$, $W_{b1,2}$, and E_m characterizing radiation-induced conductivity in strong electric fields

Material	b_1	b_2	W_{b1}	W_{b2}	E_m
	10^4 V/cm		eV		10^4 V/cm
KBr	3.7		0.024		4.5
KCl	3.9		0.025		4.5
CsI	5.1		0.023		4.5
KI	6.1		0.043		4.0
NaCl	5.1	9.2	0.029	0.052	5.5
CsI–Tl	5.3	21.0	0.024	0.096	5.7
Crystalline quartz	7.6	23.0	0.037	0.11	7.5
Glass K208	7.8	38.4	0.038	0.19	8.0

modulation depth of the conduction band in glass. Most probably, the ionization of *F*-centers occurs in the NaCl crystal, whereas in CsI–Tl, the electric field of a charged impurity locally changes the energy distribution of the density of states [17]. In oxide compounds, the density fluctuations caused by variations in chemical bonds can make a certain contribution, which results in the formation of conduction band tails [18]. Consequently, an increase in the conductivity at $E \sim (1–5) \times 10^4$ V/cm can be due to an increase in the carrier mobility, whereas at higher fields ($E \geq 5 \times 10^4$ V/cm), the release of localized electrons occurs through the electron–electron interaction according to the mechanism of collisional ionization.

Thus, we can conclude that, in alkali halide crystals (KBr, KCl, NaCl, and CsI), crystalline SiO_2, and K208 glass in electric fields of 10^4–10^5 V/cm, the nonlinearity of the current–voltage characteristics is most likely associated with the decrease in spatial localization of charge carriers in the conduction band of the insulator. An increase in the conductivity can be caused by both a decrease in the scattering by lattice vibrations due to over-the-barrier motion of carriers and an increase in the number of free carriers due to electron transport from the band tail to higher energy levels. This conclusion is confirmed by two characteristic activation energies: the lower energy ($\sim kT$), whose values are close for all the studied materials, and the higher energy ($> kT$), which is higher in disordered materials compared to crystals.

REFERENCES

1. *High-Energy Electronics of Solids*, Ed. by D. I. Vaĭsburd (Nauka, Novosibirsk, 1982).
2. V. D. Kulikov, Zh. Tekh. Fiz. **66** (8), 181 (1996) [Tech. Phys. **41**, 841 (1996)].
3. V. D. Kulikov, in *Proceedings of the 9th International Conference on Radiation Physics and Chemistry of Inorganic Materials, Tomsk, 1996*, pp. 234–235.
4. B. P. Aduev, G. M. Belokurov, and V. P. Shvaĭko, Fiz. Tverd. Tela (S.-Peterburg) **37**, 2537 (1995) [Phys. Solid State **37**, 1392 (1995)].
5. É. D. Aluker, D. Yu. Lusis, and S. A. Chernov, *Electron Excitations and Radioluminescence in Alkali Halide Crystals* (Zinatne, Riga, 1979).
6. V. V. Butkov and D. I. Vaĭsburd, Dokl. Akad. Nauk SSSR **293** (3), 598 (1987) [Sov. Phys. Dokl. **32**, 223 (1987)].
7. B. G. Gorshkov, A. S. Epifanov, A. A. Manenkov, and A. A. Panov, Tr. Inst. Obshch. Fiz. Akad. Nauk SSSR **4**, 99 (1986).
8. R. T. Williams, J. N. Bredford, and W. L. Faust, Phys. Rev. B **18** (12), 7038 (1978).
9. V. M. Lisitsyn, V. I. Korepanov, and V. Yu. Yakovlev, Izv. Vyssh. Uchebn. Zaved., Fiz., No. 11, 5 (1996).
10. M. A. Lampert and P. Mark, *Current Injection in Solids* (Academic, New York, 1970; Mir, Moscow, 1973).
11. A. V. Vannikov, V. K. Matveev, V. P. Sichkar', and A. P. Tyutnev, *Radiation Effects in Polymers: Electrical Properties* (Nauka, Moscow, 1983).
12. K. V. Shalimova, *The Physics of Semiconductors* (Énergiya, Moscow, 1976).
13. Yu. K. Pozhela, *Plasma and Current Instabilities in Semiconductors* (Nauka, Moscow, 1997).
14. A. Ausmees, M. Elango, A. Kikas, and J. Pruulmann, Phys. Status Solidi B **137** (2), 495 (1986).
15. M. Sparks, D. L. Mils, W. Warren, *et al.*, Phys. Rev. B **24** (6), 3519 (1981).
16. *Applied Electroluminescence*, Ed. by M. V. Fok (Sov. Radio, Moscow, 1974).
17. J. I. Pankove, *Optical Processes in Semiconductors* (Prentice-Hall, Englewood Cliffs, 1971; Mir, Moscow, 1973).
18. A. R. Silin' and Yu. R. Zakis, Izv. Akad. Nauk Latv. SSR, Ser. Fiz. Tekh. Nauk, No. 5, 68 (1987).

Translated by S. Egorov

Technical Physics, Vol. 45, No. 9, 2000, pp. 1147–1153. Translated from Zhurnal Tekhnicheskoĭ Fiziki, Vol. 70, No. 9, 2000, pp. 57–63.

SOLIDS

Adsorption of Water Molecules on Yttrium Barium Cuprate Superconductors

B. M. Gorelov, D. V. Morozovskaya, V. M. Pashkov, and V. A. Sidorchuk

Institute of Surface Chemistry, National Academy of Sciences of Ukraine, Kiev, 252022 Ukraine

Received May 18, 1999; in final form, September 10, 1999

Abstract—Gravimetry and thermogravimetric analysis were used to study the adsorption of water molecules on the high temperature superconductor $YBa_2Cu_3O_7$ at room temperature. It was found that water adsorption subdivides into surface adsorption and bulk adsorption, which starts after the formation at the surface of a physically bound water layer no less than 65–100 Å thick. During bulk adsorption, H_2O molecules diffuse from this surface layer to the lattice, where they form four bound states with desorption temperatures of ~208, 330, 370, and 775°C and heats of formation of 38, 99, 72, and 68 kJ/mol, respectively, and mainly occupy interstitial sites of the intermediate layers. The presence of molecules in the lattice does not affect either the superconducting transition temperature or resistance to direct current; however, it results in an increase in the surface resistance. The resistance to direct current increases due to the formation of dielectric inclusions of other phases. © *2000 MAIK "Nauka/Interperiodica".*

INTRODUCTION

It is known [1–5] that high-temperature superconductors degenerate when exposed to water. Their degradation depends on the activity of atoms at the crystal surface and interface (superconductor–liquid, superconductor–vapor), ceramic density, and impurities in crystallites or the medium. Thus, under standard conditions, monocrystals and ceramics of $YBa_2Cu_3O_7$ with a density exceeding 4.5 g/cm^3 hardly interact at all with water vapor, while less dense samples can decompose and their degradation is enhanced in the presence of carbon dioxide or halogen impurities. This process includes fast and slow stages [6] and results in a decrease in the superconducting phase content.

In addition, water molecules can enter the crystal lattice of $YBa_2Cu_3O_7$ without causing decomposition of the compound [7, 8]. H_2O molecules can occupy more than ten nonequivalent interstitial sites in the lattice, and because of the pronounced layered structure of the lattice, the barriers to diffusion along and at the normal to the **c**-axis are fundamentally different; therefore, the distribution of molecules in the layers can be characterized by nonuniform filling of interstitial sites. The incorporation of H_2O molecules into the $YBa_2Cu_3O_7$ crystal lattice can result in an increase in the critical temperature (T_c) [8], although the influence of the molecules occupying different lattice sites on the behavior of T_c appears to be different. Since the roles of the cuprate and intermediate layers in high-temperature superconductivity are different, preferential filling of interstitial sites in CuO_2 or Ba–O and Cu1–O layers by H_2O molecules can be used to determine their role in superconductivity.

It should be noted that penetration and diffusion of H_2O molecules into the lattice and degradation during adsorption occur after chemical bonding of the molecules with atoms at the $YBa_2Cu_3O_7$ surface and apparently exhibit different rates, since decomposition is determined by the reactive diffusion rate, whereas penetration is determined by the rate of diffusion into the lattice. A study of water adsorption can identify the conditions for initial incorporation and diffusion of the molecules into the lattice, localization sites, the numbers and evolution of the bound states, their influence on T_c, and the emergence of other phases.

The objective of this work is to study the adsorption of water molecules onto $YBa_2Cu_3O_7$, the conditions for initial incorporation into the crystal lattice, the kinetics and localization sites of the bound states of molecules in the lattice, and their influence on the critical temperature and resistance. To this end, we investigated surface and bulk adsorption of H_2O molecules onto $YBa_2Cu_3O_7$ at different pressures and constant temperature, and onto CuO, Y_2O_3, BaO_2, and BaO compounds, which were used to simulate the interaction of H_2O molecules with the superconductor atoms. The conditions for bulk adsorption of water molecules and their penetration into the crystal lattice, bound states of the molecules in the lattice, and their effect on T_c and the resistance were also studied. The investigations were carried out using gravimetry and thermogravimetry methods, as well as measurements of the resistance to direct and alternating current.

1063-7842/00/4509-1147$20.00

SAMPLES AND EXPERIMENTAL PROCEDURE

Samples of dispersed $YBa_2Cu_3O_7$ and single-phase ceramics (checked by X-ray studies) with a density of 5.5 g/cm^3 and lattice parameters a = 3.821 Å, b = 3.889 Å, and c = 11.667 Å were studied. The dispersed samples were obtained by dispersing ceramics and subsequently annealing the powder in oxygen for 6–8 h at 420°C. The scale-shaped powder particles had a diameter of 17–20 μm, a thickness of ~10 μm, and a specific surface of 1 m^2/g. The nonstoichiometric compounds $YBa_2Cu_3O_{7-\delta}$ with $\delta > 0$, which were obtained by vacuum annealing of samples with $\delta = 0$, were used as well.

For the simulation experiments, high-purity powders of copper, yttrium, and barium oxides, as well as of barium peroxide, were used. The specific surface of CuO and Y_2O_3 was 1 m^2/g, while that of BaO and BaO_2 was 5–10 m^2/g. Adsorption was carried out at room temperature on samples that were preliminarily annealed in vacuum (10^{-3} torr) at a temperature of ~150°C for 2.5–3 h. Double-distilled water was used as an adsorbate. The adsorption measurements were carried out using McBain scales with a sensitivity of 2×10^{-5} g. Thermogravimetric (TG) and differential thermal analysis (DTA) were performed using a Q-1500 drift gauge at a heating rate of 5°C/min. The dc resistance was measured by the four-contact technique, and the resistance at a frequency of 10 GHz, by the resonant cavity technique with the H_{011} oscillation mode.

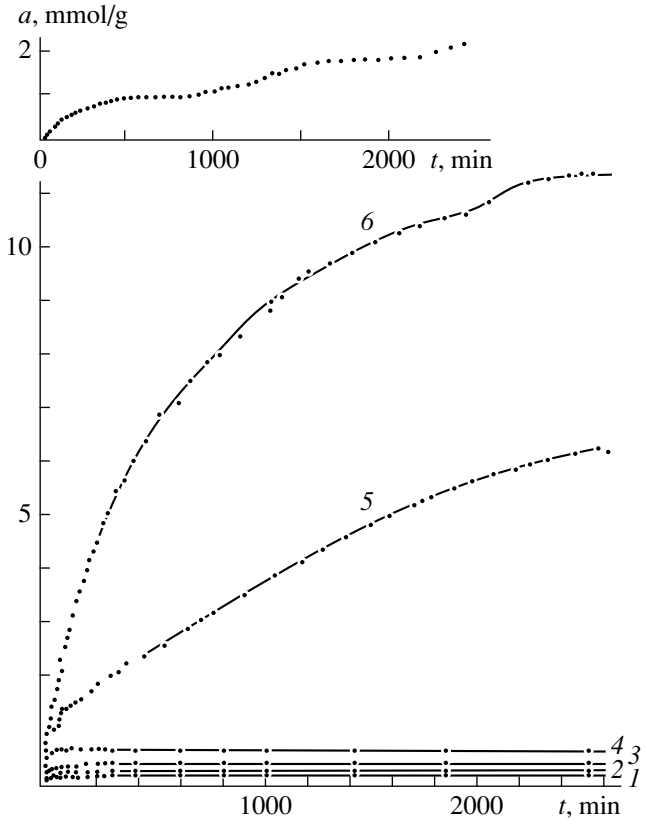


Fig. 1. Kinetics of water adsorption onto $YBa_2Cu_3O_7$ at vapor pressures β = (*1*) 3, (*2*) 6, (*3*) 10, (*4*) 13, (*5*) 16, and (*6*) 18.7 torr. The insert shows the adsorption kinetics of water for samples of $YBa_2Cu_3O_7$ ceramics at 18.7 torr.

EXPERIMENTAL RESULTS AND DISCUSSION

The adsorption kinetics a at different water vapor pressures is presented in Fig. 1. At pressures $p \le 16$ torr, a shows fast growth to the saturation value a_∞ in a time $t \le 18$–20 min. The kinetics can be described by the expression

$$a = a_\infty[1 - \exp(-Kt)], \quad (1)$$

where a_∞ is the adsorption limit at $t \longrightarrow \infty$.

The magnitude of a_∞ is proportional to pressure, and at $p \le 16$ torr, $a_\infty \le 1$ mmol/g and $K \le 4.7 \times 10^{-4}$ s^{-1} (Fig. 1, curves *1–5*). In the pressure range $p \ge 16$ torr after fast growth of the adsorbed water quantity for a time $t \ge 20$ min, bulk adsorption is observed; it is characterized by weak growth of a with a tendency toward saturation at $t > 1800$ min and shows a threshold behavior beginning at a pressure $p \ge 16$ torr and an adsorption time $t > 60$ min (Fig. 1, curves *5*, *6*).

In the ceramic samples, the adsorption behavior is similar but a_∞ is considerably smaller and bulk adsorption starts at p = 18.7 torr (Fig. 1, insert).

The results of thermogravimetric and differential thermal analysis of the $YBa_2Cu_3O_7$ samples before and after adsorption of H_2O molecules are presented in Fig. 2. In the starting samples, the weight decrease at $T \ge 420$°C is due to desorption of O1 oxygen (Fig. 2a). After adsorption at $p < 16$ torr, the loss of weight seen in the TG and DTA curves and the endothermic signal at temperatures 85–105°C indicate evaporation of water adsorbed at the surface and absorption of the evaporation heat $Q_a \simeq 26$ kJ/mol (Fig. 2b). After adsorption at $p \ge 16$ torr, with a greater number of H_2O molecules adsorbed at the surface, heating of the samples reveals a narrow exothermic DTA peak at ~82°C (Fig. 2c). This signal is related to energy release by the H_2O molecules and indicates a first-order phase transition in the water layer at the surface on heating. The amount of released heat is ~3–5 kJ/mol, and the amount of absorbed heat increases to 32–36 kJ/mol. The resulting Q_a values correspond to the heat of physical adsorption [9].

As bulk adsorption begins, first three, and then four, bound states of water molecules become apparent in the TG and DTA curves, corresponding to a decrease in m and absorption of desorption heat $Q_d \simeq 38$, 99, 72, and 68 kJ/mol with temperatures of the DTA signal maxima of $T_d \simeq 208$, 330, 370, and 775°C, respectively (Fig. 2d). The Q_d and T_d values give an indication of the coordination hydrogen bond of H_2O molecules with the

lattice atoms. Note that the states with $T_d \simeq 208$ and 330°C were observed after adsorption at temperatures $T \geq 127$°C [10, 11].

Thus, water molecules penetrate and diffuse into the $YBa_2Cu_3O_7$ lattice after formation of the polylayer of physically bound water at the particle surface. The thickness (h) of the polylayer of physically bound H_2O molecules, at which diffusion starts, can be evaluated using the relationship

$$h = \mu_a (ms)^{-1} \omega_{H_2O} N_A^{2/3} M^{-2/3} \rho^{-1/3}, \quad (2)$$

where μ_a, m, and M are the weight of adsorbed water, the weight of the sample, and the gram-molecular weight of water, respectively; s is the specific surface of the sample; ρ is the density of water; ω_{H_2O} is the area occupied by an H_2O molecule at the surface; and N_A is Avogadro's number.

At $\omega_{H_2O} = 10.2$ A^2 [12] and $\rho = 1$ g/cm^3, the polylayer thickness is $h \simeq 100$ Å and the minimum h value at which diffusion can be observed is equal to ~65 Å.

The phase transition in the polylayer of adsorbed water is similar to the order–disorder transition in liquid crystals and the orientation transitions at the surface [13, 14]; therefore, it can be related to the transition from the ordered state of dipoles at $T < 82$°C to the disordered state at $T > 82$°C. The ordering of dipoles sets in at a polylayer thickness of $h > 40$ Å and is possible through interaction with $YBa_2Cu_3O_7$ cations, which binds the dipoles at the surface, and through the dipole interaction in the polylayer, which aligns the dipoles [9].

It should be noted that for $YBa_2Cu_3O_{7-\delta}$ compounds with $\delta \geq 0.3$ the exothermic DTA signal is not observed when the water polylayer is heated. This shows that dipoles in the adsorbed polylayer are not ordered, probably due to reduction of the orientation interaction energy with increasing δ as the distance between the lattice cations and bound dipoles in the polylayer increases.

The beginning of penetration and diffusion of the H_2O molecules into the $YBa_2Cu_3O_7$ lattice after the formation of a polylayer with a thickness $h \geq 65$ Å can be related to lowering of the barrier to molecular chemisorption as h increases and the monolayer of physically adsorbed water transforms into a polylayer. The possibility of such a relation follows from the lowering of the interaction energy of atoms of the high-temperature semiconductor of species α in the charge state q_α^0 with atoms of H_2O molecules:

$$U = \frac{1}{8\pi\varepsilon_0\varepsilon} \sum_{i,\alpha} q_\alpha^0 V_\alpha^i, \quad (3)$$

where ε_0 is the dielectric constant of vacuum,

$$V_\alpha^i = \sum_{j,\beta} \frac{q_\beta^i}{|r_{i\alpha} - r_{j\beta}|}$$

is the electrostatic potential created by the jth atoms of β-H_2O molecules at the location of an ith atom as the polylayer is formed, and ε is the dielectric constant of the polylayer.

The diffusion coefficient D of water molecules into the particle bulk can be determined from the relationship [15]

$$D = \pi V^2 \Gamma (4 S_1^2 t)^{-1} [a(t)/a_\infty]^2, \quad (4)$$

where V is the volume of particles, $S_1 = \rho V s$ is the outer surface area of the particles (ρ is the density of the sam-

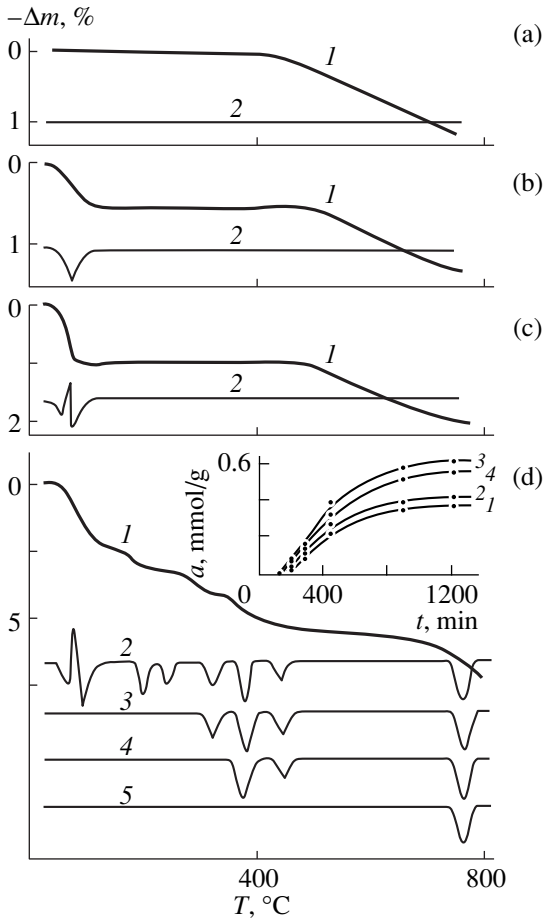


Fig. 2. Temperature dependence of (*1*) the weight loss and (*2*) DTA signal of $YBa_2Cu_3O_7$ samples prior to water adsorption (a) and after surface adsorption of H_2O molecules at 18.7 torr for 10 (b), 50 min (c), and after bulk adsorption (d) with subsequent annealing of the states at 250 (DTA, *3*), 400 (DTA, *4*), 450 (DTA, *5*). The insert shows the occupation kinetics of the state by H_2O molecules for desorption temperatures of ~208 (*1*), 330 (*2*), 370 (*3*), and 775°C (*4*).

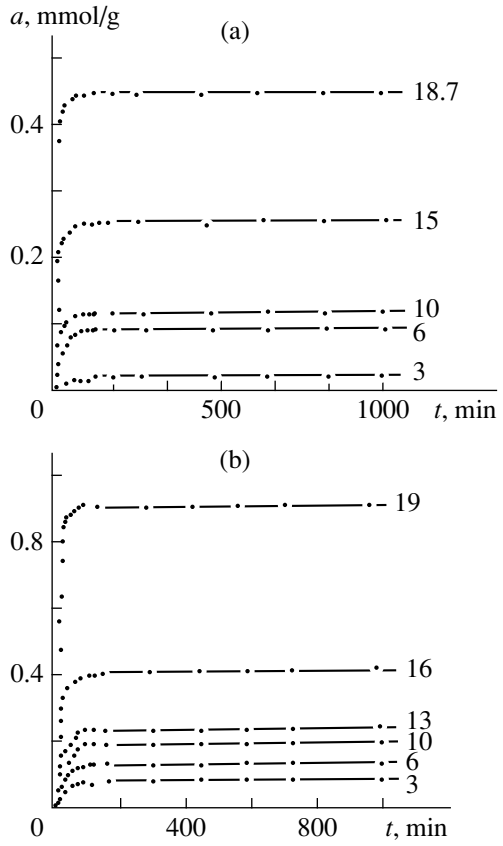


Fig. 3. Kinetics of adsorption of water molecules (a) on Y_2O_3 at pressures of 3, 6, 10, 15, and 18.7 torr and (b) on CuO at p = 3, 6, 10, 13, 16, and 19 torr. Numbers near the curves denote p values.

ples), and $\Gamma = a_\infty/c_0$ (c_0 is the amount of adsorbate at the boundary of the samples).

Assuming that the $YBa_2Cu_3O_7$ particles are of uniform size, $a_\infty = a = 10.2$ mmol/g at $t \longrightarrow 3000$ min and the surface reaction rate, that is, the transition rate of H_2O molecules from free to bound state, is considerably higher than the diffusion rate, one can obtain $D = 2.6 \times 10^{-12}$ cm^2/s.

With increasing bulk adsorption time, the number of molecules in bound states increases and the occupation of different states is different (insert in Fig. 2d). The molecules localized in the lattice do not interact with each other, since in successive annealings at ~250, 340, 400, and 800°C, each type of state is annealed without affecting Q_d of the other types of states. In addition, DTA curves of the samples held in an atmosphere of water vapor for $t > 300$ min have two peaks at temperatures $T_{max} \simeq 240$ and 445°C, which are not accompanied by any change of weight (Fig. 2d). This can imply that, in $YBa_2Cu_3O_7$, compounds of different phase composition form, which absorb heat in the process of melting or decomposition.

Thus, the water molecules adsorbed in the $YBa_2Cu_3O_7$ lattice are found in four bound states.

To determine the positions where water molecules are localized in the $YBa_2Cu_3O_7$ lattice, adsorption on yttrium and copper oxides, barium peroxide, and nonstoichiometric cuprates $YBa_2Cu_3O_{7-\delta}$ with $\delta = 0.27$ and 0.7 was studied. In Y_2O_3, CuO, and BaO_2, the charge states of Y^{3+}, Cu^{2-}, Ba^{2+}, O^{2-}, and O^- ions are comparable with the atomic charges in $YBa_2Cu_3O_7$, namely, $Y^{3.05+}$, $Cu2^{2+}$, $Cu1^{1.46+}$, $O2^{2.01-}$, $O3^{2.16-}$, $Ba^{1.45+}$, $O4^{0.67-}$, and $O1^{1.76-}$ [16], and the peroxide bond O=O in BaO_2 is similar to the bond between O4 and O1 atoms. Therefore, the charge states of atoms of such compounds can simulate the barrier to chemisorption and diffusion of H_2O molecules in the $YBa_2Cu_3O_{7-\delta}$ lattice. In addition, by varying δ, the atomic charge states in Ba–O and Cu1–O layers can be changed, while keeping unchanged those in the Y and Cu2–O layers [16, 17], which allows the Q_d values of H_2O molecules in the intermediate layers to be varied and the positions of the molecules in the $YBa_2Cu_3O_7$ lattice to be determined.

The adsorption kinetics of water molecules on oxides of yttrium and copper are similar, with a increasing until saturation at $t \le 20$ min and a constant value of a at $t > 20$ min (Fig. 3). In the pressure range $p \le 19$ torr, the behavior of a can be described by expression (1) with parameters $a_\infty \le 0.97$ mmol/g, $K \le 3.5 \times 10^{-4}$ s^{-1} and $a_\infty \le 0.47$, $K \le 4.6 \times 10^{-4}$ s^{-1} for adsorption on Y_2O_3 and CuO, respectively. The TG and DTA curves of Y_2O_3 and CuO samples after water adsorption at p = 18.7 torr are identical to those presented in Fig. 2b and indicate the formation at the oxides' surface of a layer of physically bound water with $h \simeq 100$ Å and the absence of penetration of the H_2O molecules into the lattice.

The adsorption behavior on BaO_2 and $YBa_2Cu_3O_7$ is essentially similar (Fig. 4). At pressures $p \le 16$ torr, a reaches saturation in the time interval $t \le 20$ min and stays constant for $t > 20$ min, with parameters $a_\infty \le 1$ mmol/g and $K \le 4.1 \times 10^{-4}$ s^{-1}. At $p \ge 16$ torr, bulk adsorption is observed, which features a slow increase in a with time and a tendency to saturation for $t \longrightarrow 1800$–2000 min. The fast increase in a corresponds to the formation on the BaO_2 surface of a physically bound water layer, whose evaporation heat Q_a = 18 kJ/mol (Fig. 5b); the slow increase corresponds to diffusion of the H_2O molecules into the lattice, where in a short period of time t they form one state (Fig. 5c) and, at a later time t, form a second state with parameters Q_d = 23 and 21 kJ/mol and T_d = 95 and 125°C, respectively (Fig. 5d). Molecules diffuse into the lattice after a polylayer of thickness $h \simeq 10$–20 Å has been formed on the surface. In addition, in the DTA curves, two signals are present at $T_{max} \simeq 240$ and 380°C (Fig. 5d), which are not accompanied by a change of m and are possibly related to the formation of compounds

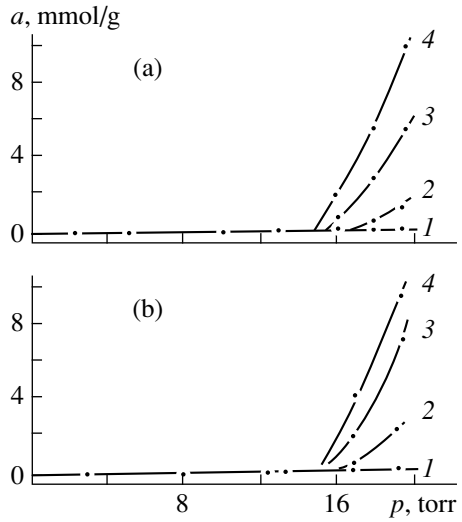


Fig. 4. Isotherms of adsorption of water molecules on (a) BaO_2 and (b) $YBa_2Cu_3O_7$ for exposures of (*1*) 15, (*2*) 60, (*3*) 1100, and (*4*) 1400 min.

of different phase compositions. The values of T_{max} of the first signal are identical in BaO_2 and $YBa_2Cu_3O_7$; and those of the second are close, which is evidence of identical or close compositions of the phases formed in the two compounds on contact with water.

Thus, the identical behavior of bulk adsorption of H_2O molecules in BaO_2 and $YBa_2Cu_3O_7$ and its absence in Y_2O_3 and CuO may be evidence that the formation of bound states in $YBa_2Cu_3O_7$ is due to interaction of the H_2O dipoles with barium and oxygen atoms in the Ba–O layer.

In $YBa_2Cu_3O_{7-\delta}$, with a decreasing number of O1 oxygen atoms, the values of the heat of formation Q_d of the four bound states of the H_2O molecules change (see table). Thus, as δ is increased, Q_d of the states with $T_d \simeq 200$ and 770°C varies nonmonotonically, while Q_d of the states with $T_d \simeq 310$ and 360°C smoothly increases and decreases, respectively. Since the heat of formation Q_d is released on breakage of the bond between atoms of the water molecule in charge states q_β^j and lattice atoms in charge states q_α^0 with energy U (equation (3)), the behavior of $Q_d(\delta)$ is presumably determined by variation of q_α^0 with increasing δ. If at $\delta \longrightarrow 1$ the charge state of Cu1 smoothly decreases from +1.5 to +0.8 and that of Ba increases from +1.5 to +2 [16, 18], the variation of Q_d of the states with $T_d \sim 310$ and 360°C is related to variation of q_{Ba} and q_{Cu1}, respectively, and can be an indication that the H_2O molecules in these states are located near Ba and Cu1 atoms (Fig. 6). The desorption heat of the state with $T_d \simeq 310$°C exceeds the Q_d of the other states by a factor of 1.4–3.5, which points to stronger binding of the molecules with the surrounding atoms. In this state, the molecules are presumably located in the interstitial

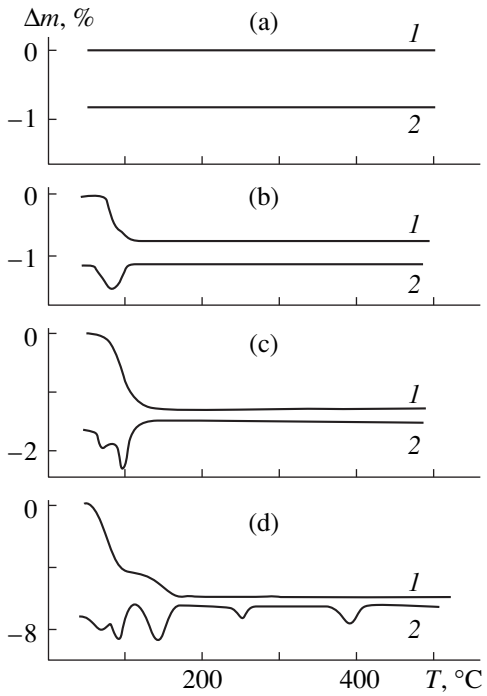


Fig. 5. Temperature dependences of (*1*) the weight loss and (*2*) DTA signal for BaO_2 samples before (a) and after adsorption of water molecules at 18 torr for (b) 25, (c) 70, and (d) 360 min.

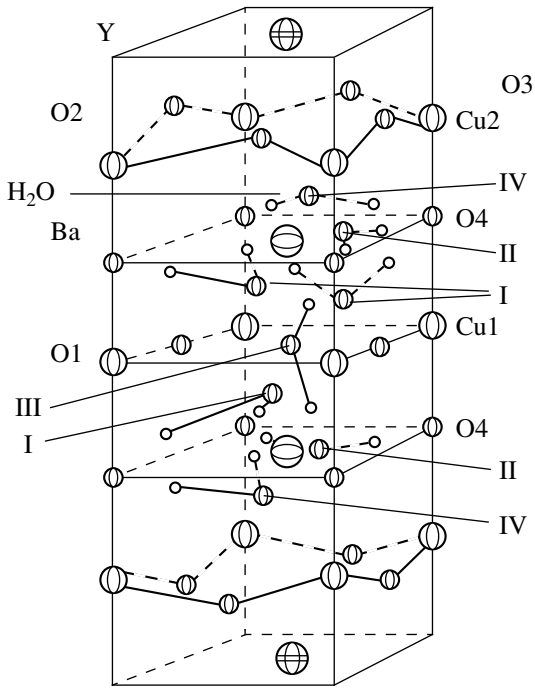


Fig. 6. Positions of H_2O molecules in the $YBa_2Cu_3O_7$ lattice: states with $T_d \simeq$ (I) 208, (II) 330, (III) 370, and (IV) 775°C.

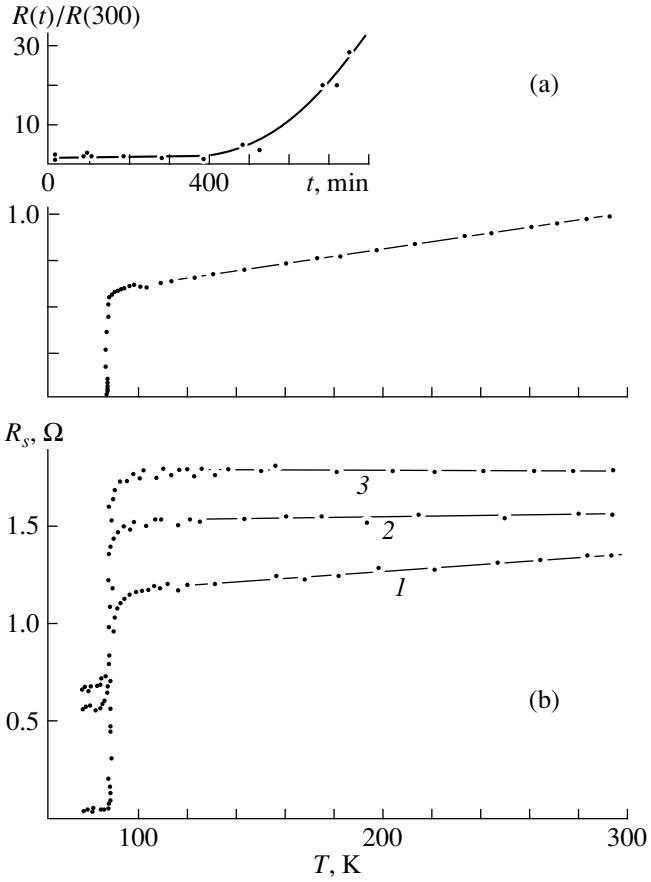


Fig. 7. Temperature dependences of the direct current resistance (a) of $YBa_2Cu_3O_7$ ceramics before and after adsorption at 18.7 torr for $t \le 240$ min (insert shows the dependence of the resistance at room temperature on adsorption time) and of the surface resistance (b) before (*1*) and after adsorption of H_2O molecules for (*2*) 240 and (*3*) 480 min.

sites of the Ba–O layer and have semipolar bonds with Ba and O4 atoms (state II in Fig. 6), and the increase in Q_d at $\delta \longrightarrow 1$ is due to the increase in q_{Ba}, q_{O4}, and energy U. H_2O molecules in the states with $T_d \simeq 360$°C are bound with Cu1, O4, and O1 atoms by semipolar bonds (state III in Fig. 6); and the decrease in Q_d at $\delta \longrightarrow 1$ is caused by the reduction in U, because of the decrease in both q_{Cu1} and the number of O1 atoms. The H_2O molecules in the other two states are probably located in the interstitial sites of Ba–O, Cu1–O and Ba–O, Cu2–O layers and interact with O4, Cu2, Cu1, and Ba atoms. The variation of $Q_d(\delta)$ of these states as $\delta \longrightarrow 1$ is probably determined by the nonmonotonic change of q_{O4} from –0.67 to –2.07 and the electron density distribution between Cu2 orbitals as carriers from the intermediate layers enter cuprate layers [16, 19] (state IV in Fig. 6), as well as by the decrease in the Cu1–O4 distance [20] and the weakening of repulsion as O1 atoms leave the lattice (state I in Fig. 6).

Desorption heats of the bound states of water in the $YBa_2Cu_3O_{7-\delta}$ lattice

Compound	Desorption heat (kJ/mol) of the states with T_d (°C)			
	208 ± 10	330 + 20	370 + 10	775 ± 5
$YBa_2Cu_3O_7$	38	79	72	68
$YBa_2Cu_3O_{6.73}$	23	81	51	51
$YBa_2Cu_3O_{6.3}$	56	125	42	75

It should be noted that the dependence of Q_d of the four states on the oxygen content in Cu1–O layers indicates that water molecules are not localized in the interstitial sites of the Y and Cu2–O layers. This can be explained by the fact that the repulsion between Y atoms and the H atoms in the H_2O molecule exceeds the attraction between Y and O atoms hindering the formation of stable states. The lack of correlation between the number of O1 atoms and H_2O molecules in the lattice at $\delta \longrightarrow 1$ suggests that the localization site of the molecules is determined by the interaction of cations and oxygen of the molecule, while the orientation is determined by the interaction of O4 anions with hydrogen atoms.

Thus, in the $YBa_2Cu_3O_7$ lattice, the water molecules are mainly found at the interstitial sites of the intermediate layers and are bound with Ba, Cu1, and O4 atoms by semipolar bonds.

For adsorption times $t \le 240$ min, penetration of the H_2O molecules into the $YBa_2Cu_3O_7$ lattice does not change the temperature dependence of direct current resistance or the critical temperature (Fig. 7a). However, with increasing t, R is rising (Fig. 7a, insert). The increase in R is presumably due to inclusions of dielectric phases emerging and growing in number at $t >$ 300 min, which give rise to DTA signals at $T_{max} \simeq 240$ and 445°C as they decompose.

In contrast to R, after penetration of H_2O molecules (up to $\sim 2.6 \times 10^{21}$ cm^{-3}) into the crystal lattice of $YBa_2Cu_3O_7$, its surface resistance R_s increases at both $T > T_c$ and $T < T_c$. The temperature of the superconducting transition is not changed by water adsorption (Fig. 7b). Since at $T > T_c$, $R_s = (\mu_0\mu\omega\rho/2)^{1/2}$, where μ_0 and μ are the permeabilities of vacuum and the sample, respectively; ω is the circular frequency; and ρ is the specific resistance

$$\rho = \rho_i + \rho_d + \rho_{ph} + \rho_e. \qquad (5)$$

Here ρ_i, ρ_d, ρ_{ph}, and ρ_e are contributions to the specific resistance due to scattering by impurities, defects, phonons, and electrons, the increase in R_s following penetration of H_2O molecules into the lattice can be explained by the greater contribution from ρ_d because of additional scattering of carriers on atoms of H_2O molecules. At $T < T_c$, when $R_s = -\mu_0\omega \mathrm{Im}\bar{\lambda}$ ($\bar{\lambda}$ is the complex penetration depth [21]) is determined by car-

rier energy dissipation in the layer of depth $\bar{\lambda}$; the increase in R_s is presumably related to the increase in dissipated energy due to scattering by atoms of the H_2O molecules.

In conclusion, the water molecules adsorbed on the surface of $YBa_2Cu_3O_7$ form a polylayer of physically bound H_2O molecules. When a layer thickness of $h \geq$ 65–100 Å is reached, the molecules diffuse into the lattice, localizing mainly at the interstitial sites of the intermediate layers and forming four bound states. Such molecules do not affect the temperature of the superconducting transition and direct current resistance; however, they cause an increase in the surface resistance. The increase in the surface resistance to direct current is presumably related to the formation of dielectric compounds of different phase compositions.

REFERENCES

1. V. I. Nefedov, A. N. Sokolov, M. A. Tyzikhov, *et al.*, Poverkhnost, No. 9, 22 (1989).
2. V. I. Nefedov and A. N. Sokolov, Zh. Neorg. Khim. **36**, 2723 (1989).
3. I. V. Pulyaeva, N. S. Granova, Yu. N. Velikhov, and V. V. Usenkova, Preprint No. IMK-92-13 (Kharkov, 1992).
4. A. Barkatt, H. Hujaji, R. W. Vasentha, *et al.*, MRS Bull., September, 45 (1993).
5. M. U. Kalanov, M. Karimov, M. S. Paĭzulakhanov, *et al.*, Uzb. Fiz. Zh., No. 2, 56 (1993).
6. E. A. Eremina, N. N. Oleĭnikov, V. I. Nefedov, and A. N. Sokolov, Zh. Vses. Khim. O–va. im. D. I. Mendeleeva **34**, 528 (1989).
7. A. I. Cheranev, S. P. Rozov, V. S. Mironov, and L. I. Barsova, Sverkhprovodimost: Fiz., Khim., Tekh. **3** (6), 37 (1990).
8. L. B. Harris and F. K. Nuang, Solid State Commun. **67**, 359 (1988).
9. V. Ya. Antonchenko, A. S. Davydov, and V. V. Il'in, *Foundations of Water Physics* (Naukova Dumka, Kiev, 1991).
10. Yu. M. Baĭkov, E. A. Egorov, V. V. Zhizhenkov, *et al.*, Sverkhprovodimost: Fiz., Khim., Tekh. **3**, 99 (1990).
11. E. A. Eremina, A. B. Yaroslavtsev, N. N. Oleĭnikov, *et al.*, Sverkhprovodimost: Fiz., Khim., Tekh. **2**, 617 (1989).
12. R. N. Meals and F. M. Lemis, *Silicones* (Reihold, New York, 1963).
13. P. de Gennes, *The Physics of Liquid Crystals* (Clarendon, Oxford, 1974; Mir, Moscow, 1977).
14. V. M. Ogenko, V. M. Rozenbaum, and A. A. Chuĭko, *Theory of Vibration and Reorientation of Surface Atomic Groups* (Naukova Dumka, Kiev, 1991).
15. D. P. Timofeev, (Akad. Nauk SSSR, Moscow, 1962).
16. V. E. Gusakov, Fiz. Nizk. Temp. **21**, 805 (1995) [Low Temp. Phys. **21**, 621 (1995)].
17. A. V. Vishnyakov, in *High-Temperature Superconductivity* (Mashinostroenie, Leningrad, 1990), pp. 377–404.
18. Y. Fukuda, N. Nagoshi, T. Suzuki, *et al.*, Phys. Rev. B: Condens. Matter **39**, 11494 (1989).
19. A. Bianconi, M. De Santis, A. Di Cicco, *et al.*, Phys. Rev. B: Condens. Matter **38**, 7196 (1988).
20. A. Manthiram, X. X. Tang, and J. B. Goodenough, Phys. Rev. B: Condens. Matter **42** (1A), 138 (1990).
21. N. H. Tea, M. B. Salamon, T. Datta, *et al.*, Phys. Rev. B: Condens. Matter **45**, 5628 (1992).

Translated by V. Lebedev

Technical Physics, Vol. 45, No. 9, 2000, pp. 1154–1158. Translated from Zhurnal Tekhnicheskoĭ Fiziki, Vol. 70, No. 9, 2000, pp. 64–69.

SOLIDS

An Adsorption Mechanism Underlying Phase Transformation in Stabilized Zirconia

V. I. Alekseenko and G. K. Volkova

Galkin Physicotechnical Institute, National Academy of Sciences of Ukraine, Donetsk, 340114 Ukraine

Received September 14, 1999

Abstract—Tetragonal-to-monoclinic martensite transformation due to water adsorption was found on the surface of both dispersed and ceramic oxide-stabilized zirconia. The transformation is due to local stretching strains arising in the lattice when the charge state of a stabilizer changes because of water molecules that are chemisorbed on the powder or ceramic surface. © *2000 MAIK "Nauka/Interperiodica".*

INTRODUCTION

The structure and properties of powder (tetragonal β phase) and ceramic zirconia stabilized by metal oxides (Y_2O_3, CaO, MgO, etc.) are sensitive to temperature, pressure, and various reactive gaseous and liquid media. When exposed to external effects, the high-temperature β phase may transform to the monoclinic α phase [1–3], which changes the properties of the material. Specifically, the physical, mechanical, and service properties of β-ZrO_2, used in the production of cutting tools, machine parts, medical prostheses, refractory materials, and ceramic materials for electronics, are severely degraded. Note that the low-temperature monoclinic α phase of crystalline unstabilized zirconia is stable to 1170°C; its high-temperature β phase, between 1170 and 2277°C; and the high-temperature cubic γ phase, between 2277 and 2690°C [2]. The stabilized β and γ phases may exist at near-room temperatures.

Today, the stabilization of ZrO_2 powder and ceramic is a key issue, since their properties are highly unstable under various environmental conditions. The solution to the problem is difficult, because stabilization mechanisms for the high-temperature phases are poorly understood. As a result, the physics of instability of dispersed and sintered materials is still unknown. In other words, microscopic mechanisms that relate external effects to the phase transformations responsible for the material structure and properties have not been found.

In this work, we studied the β-to-α transformation mechanism in Y_2O_3-stabilized zirconia. This transformation occurs when the surface of the β material (powder or ceramic) adsorbs water molecules.

Before discussing the mechanisms of high-temperature phase modification, we will outline the current concepts of structure stabilization. This information should be useful for studying structure instability and phase transformation. In [4–6], the stabilization of cubic ZrO_2 due to nonstoichiometry of oxygen vacancies was suggested and experimentally proved. In this case, the material is stabilized through generation of the necessary concentration of anion vacancies in the lattice. Resulting local stresses may "hold" the high-temperature cubic phase down to room temperature. Such an explanation obviously applies to β-ZrO_2 as well. Thus, in [4–6], nonstoichiometry appears as the decisive factor in high-temperature phase stability.

Stabilization of high-temperature ZrO_2 phases may be treated in terms of the energy theory [7, 8]. The internal energy U of a crystal is the sum of the volume, U_v, and surface, σS, energies (σ is the specific surface energy): $U = U_v + \sigma S$. Since $U_{v\alpha} < U_{v\beta} < U_{v\gamma}$ and $\sigma_\alpha > \sigma_\beta > \sigma_\gamma$, the high-temperature phase may become energetically favorable at room temperature if $S < S_{cr}$; that is, the smaller the grain size of the high-temperature phase, the more stable the phase.

A number of authors believe that the high-temperature ZrO_2 phases may exist under normal conditions owing to the stabilizing action of impurities. It was noted [9] that the interaction of Zr^{4+}, Hf^{4+}, Th^{4+}, and U^{4+} ions with water or hydroxyl ions promotes the formation and existence of the high-temperature phases of the associated oxides. This assumption was also supported in [10]. Metal oxides dissolved in ZrO_2 exert a similar stabilizing effect. Their presence in the lattice also generates oxygen vacancies [1]. It is believed that the structure is stabilized by vacancy-induced stresses, as in the case of nonstoichiometric stabilization.

RESULTS AND DISCUSSION

We investigated the structure and properties of powder and ceramic Y_2O_3-stabilized β-ZrO_2 samples. They were subjected to hydrothermal treatment by dipping them into distilled water and boiling water for a certain time. The grain size of the β modification was 120 Å, and the specific surface area measured by the Brunauer–Emmett–Teller method was found to be

1063-7842/00/4509-1154$20.00 © 2000 MAIK "Nauka/Interperiodica"

100 m^2/g. β-ZrO_2 ceramic samples sintered at 1500°C have accessible porosities from 5% to 0, depending on the hydrostatic pressure applied to the extrusion billets. The grain size was about 0.4 μm. After such treatment, the powder and ceramic samples undergo β-to-α transformation. Of special importance is the fact that the transformation proceeds on the surface, as evidenced by layer-by-layer polishing of the ceramic samples. The thickness (hence, amount) of the surface α phase is obviously independent of the time of hydrothermal treatment. For example, in the above range of accessible porosity, the amount of the α phase did not change when the time of hydrothermal treatment was varied from 60 min to several days. The α phase amount, however, depends on the porosity, which is in direct proportion to the specific surface area of that part of the sample absorbing X-ray radiation during phase analysis. The amount of α phase vs. porosity dependence is given in Fig. 1.

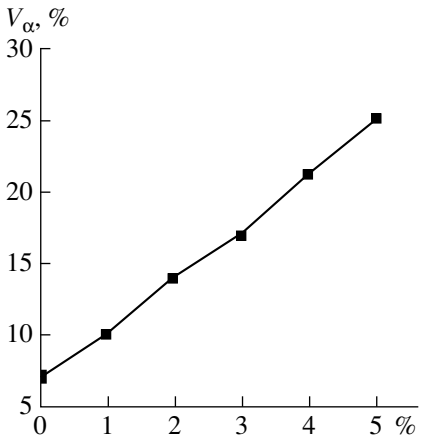


Fig. 1. Amount of nonoriented α phase in hydrothermally treated β-ZrO_2 vs. accessible porosity. The treatment time was 60 min.

The same surface transformation is observed in the powders. This follows from X-ray diffraction studies (a DRON-3 diffractometer, CuK_α radiation) of powders with different α phase contents (to 16%). It was shown that, for the α phase, the region of coherent X-ray scattering, which characterizes the grain size, is always less than for the β phase, although the energy theory implies the reverse relationship. In our opinion, this finding is of paramount importance and calls for special investigation, since it is believed today that phase transformation in dispersed systems proceeds in the entire volume, not only in part of it, as follows from our work. This is a possible reason why the energy theory disagrees with experiments [12].

The results for hydrothermally treated samples are in conflict with ideas put forward in [9, 10], where it is argued that water stabilizes the high-temperature ZrO_2 phases. Neither can they be explained by the concept [3], according to which the diffusion of surface-adsorbed water quasi-molecules (hydroxyl ions) into the yttria-stabilized β-ZrO_2 lattice is responsible for the phase transformation. The above ideas rely on the following experimental data: (1) ceramic disordering takes place in the surface layers. (2) Disordering is observed in air at temperatures between 200 and 300°C. (3) Water favors the β-to-α transformation. We nevertheless guess that the existing ideas of transformation mechanism are insufficiently advanced. They cannot answer two questions: (1) why does the diffusion of water quasi-molecules proceed only at the surface of powder grains or a ceramic sample? (2) Why does this diffusion have an optimum temperature range? Within the model in [3], the transformation rate must be proportional to the time and temperature of hydrothermal treatment. However, this is not confirmed by results obtained in [3], much less by our results. The list of questions can be extended. The lack of answers initiates a search for alternative models for β-to-α transformation due to adsorption.

We will discuss an electrophysical adsorption mechanism of β-to-α transformation in stabilized ZrO_2. In this mechanism, β-ZrO_2 is stabilized by Y_2O_3, whose dissolution in the zirconia lattice, having ionic–covalent bonds, generates substitutional ions Y^{3+} and oxygen vacancies V [11, 13].

Using the band diagram, we will consider the electronic structure of β-ZrO_2 when Y^{3+} ions are substituted for Zr^{4+} ions to generate oxygen vacancies (vacancies replace oxygen atoms). Trivalent Y^{3+} ions present in the forbidden gap of the β-ZrO_2 wide-gap dielectric produce acceptors with an ionization energy E_a. Then, the Fermi energy E_F would have to lie closer to the valence band (Fig. 2a). Oxygen vacancies in β-ZrO_2 produce donors with an ionization energy E_d, since the concentration of electrons exceeds that necessary to provide ionic bonds. In this case, the Fermi level tends to the conduction-band bottom (Fig. 2b). The simultaneous presence of acceptors (Y^{3+}) and donors (oxygen vacancies) in β-ZrO_2 was verified in [14]. In this case, vacancies donate electrons to Y^{3+} acceptors, turning into $V^{(+)}$ and converting Y^{3+} into Y^{2+} (or $Y^{3+(-)}$). The resulting unlike charges neutralize each other due to ionic bonding, which, in our opinion, serves as the lattice stabilizer. In addition, the concentrations of the donors and acceptors are counterbalanced, so that there is no deficit (or excess) of electrons needed to produce ionic bonds. That is why stabilized β-ZrO_2 is a nonconducting dielectric. In the last case, when β-ZrO_2 has $V^{(+)}$ and $Y^{3+(-)}$ (ionized donors and acceptors), the Fermi level is near the midgap (Fig. 2c).

Thus, Fig. 2c is a qualitative picture of the initial electronic structure of stabilized β-ZrO_2. Our next step is to consider β-to-α transformation due to water adsorption.

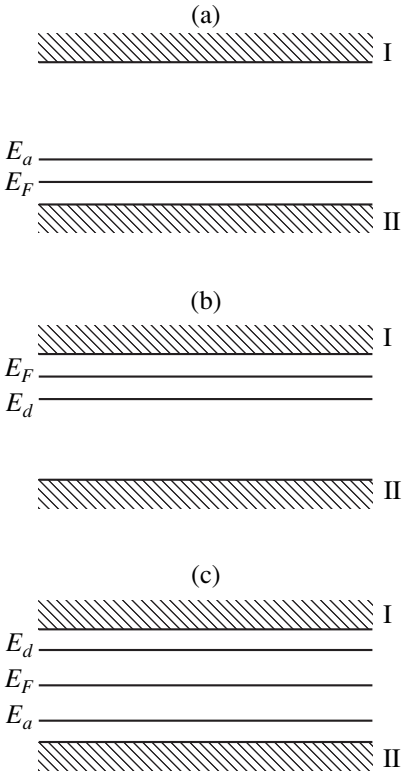


Fig. 2. Qualitative band diagram of β-ZrO_2 with (a) yttrium, (b) oxygen vacancies, and (c) yttrium and oxygen vacancies. I, conduction band; II, valence band.

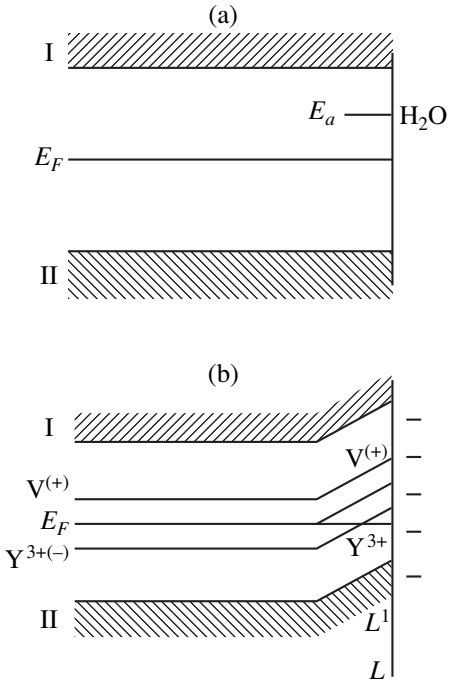


Fig. 3. Qualitative band diagram of the β-ZrO_2 surface with a chemisorbed water quasi-molecule: (a) electroneutral surface and (b) negatively charged surface. I and II, the same as in Fig. 2.

A real solid surface has a high density of surface states. The energy spectrum of nonadsorption states and that of local levels of chemisorbed species superimpose on one another [15]. At equilibrium, the occupation of the levels is uniquely defined by the position of the Fermi level on the surface.

Let water molecules and quasi-molecules physically and chemically adsorbing on the β-ZrO_2 powder and ceramic surfaces act as sorbents [16]. At chemisorption on the surface, the wave functions of the lattice and quasi-molecules overlap so much that the latter can be thought of as impurities locally disturbing the lattice potential periodicity [17]. Such a periodicity disturbance causes discrete donor or acceptor levels to arise in the β-ZrO_2 forbidden gap. Acceptor levels attract electrons, and the surface gains a local negative charge. When examined in a JEM-200A electron microscope, dispersed β-ZrO_2 powder with adsorbed quasi-molecules shows up as negatively charged. In terms of the elaborated concept, this allows us to argue that the surface layer of the adsorbed water is of an acceptor nature. Figure 3a qualitatively shows the band diagram of the electroneutral β-ZrO_2 surface with a chemisorbed water molecule. The position of the surface acceptor with an energy E_a in the forbidden gap depends on the type of lattice and sorbed molecule, as well as the distance of the molecule to the surface [17]. Note that the occupation of a surface level by a carrier is of a probabilistic nature and is described by Fermi–Dirac statistics. The probability of a local surface level being occupied is given by [17]

$$f = \{1 + \exp[(E_F - E_a)/kT]\}^{-1}. \quad (1)$$

Let N_0 be the number of adsorbed molecules per unit surface area; then, the number of charged molecules on it is

$$N^- = N_0\{1 + \exp[(E_F - E_a)/kT]\}^{-1}. \quad (2)$$

Assuming that each of the molecules has a charge equal to the electron charge e, one can determine the surface charge density:

$$\sigma^- = eN^- = eN_0\{1 + \exp[(E_F - E_a)/kT]\}^{-1}. \quad (3)$$

As follows from (3), the larger the value of $(E_F - E_a)$ or the farther E_a is from E_F, the smaller the surface charge density. Such a situation occurs when the adsorbed molecule is far away from the surface, thus causing the local level to tend to the forbidden gap.

According to the electroneutrality condition, charging of the surface must lead to the formation of a near-surface space charge region (SCR) to neutralize the surface charge. Neutralization implies that charges of sign opposite to that of the surface are attracted to the surface and like charges are repulsed. The SCR thickness L, depending on the material permittivity ε and carrier

concentration n, is expressed as [17]

$$L = [\varepsilon\varepsilon_0 kT(e^2 n)]^{1/2}. \quad (4)$$

The SCR results in a potential difference between the surface and interior and, eventually, in band (level) bending. If the surface is negative, the energy levels bend up, since the energy of an electron moving from the bulk to the surface grows (Fig. 3b). The bend extends a distance L into the material. Of interest is the near-surface region of thickness L^1, where, as seen in Fig. 3b, the yttrium acceptor lies above the Fermi level and, hence, releases an electron captured from the donor level of the vacancy. Its charge state thus changes from $Y^{3+(-)}$ to Y^{3+}, which changes the strength of the vacancy–yttrium ionic bond. As a result, the distance between zirconium and yttrium atoms in an Y–V–Zr chain will also change. Thus, the material will experience local stretching strains (Δl), causing the β-to-α martensite transformation to occur and the lattice volume to increase (Fig. 4). This strain is not a mere lattice distortion; instead, it is locally directed along the specific crystallographic direction. Thus, we can suppose that the α phase resulting from the β-to-α transformation in β-ZrO_2 will have a certain crystallographic orientation. This point will be touched upon below. It should be noted that the occurrence of local strains due to a change in the charge state of ions is also predicted by the theory of lattice symmetry [18].

Thus, the near-surface region of size L^1 is just that region where the β-to-α transformation in β-ZrO_2 takes place owing to the chemisorption of water quasi-molecules. Accordingly, the amount of the α phase is proportional to its volume. Note that the α phase volume increases with the charge surface density (or surface density of adsorbed species), since the number of ions changing their charge state and, hence, causing local strains also grows with the charge density on the surface. It becomes obvious from the aforesaid that the amount of the α phase on the surface can be analytically estimated. Its volume V_α in a powder grain is

$$V_\alpha = L^1 S(N^-/N_0), \quad (5)$$

where S is the grain surface area and N^-/N_0 is the relative surface density of charged adsorbed quasi-molecules.

Passing from the grain surface area to the total surface area S_c of a powder or ceramic sample of mass m_c, we rearrange (5) to the more convenient form

$$V_\alpha/V_\beta = L^1 S_0 \rho_\beta \{1 + \exp[(E_F - E_a)/kT]\}^{-1}. \quad (6)$$

Here, we used the classical expressions $S_c = S_0 m_c$ (S_0 is the specific surface area) and $m_c = \rho_\beta V_\beta$ (ρ_β is the β phase density, and V_β is the β phase volume where the transformation takes place).

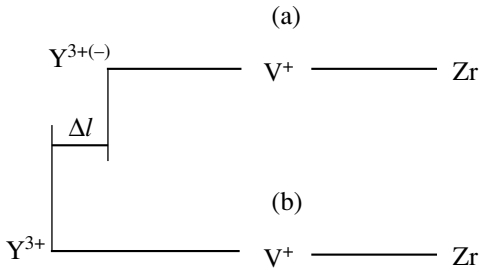


Fig. 4. Crystallographically oriented local stretching strain in the β-ZrO_2 lattice. The strain arises when an yttrium ion changes its charge state because of water adsorption: (a) neutral surface and (b) charged region.

Since the surface charge is distributed statistically, the parameter L^1 in (6) can be replaced by L from (4). Expression (6) implies the following.

(1) The amount of the resulting α phase depends largely on the difference between the Fermi level and the ionization energy of a surface level. Hence, by varying the Fermi level position (for example, by doping), we can precisely control the stability of high-temperature β-ZrO_2 under adsorption. The ionization energy of a surface level can also be changed, for example, by varying the molecule–adsorbent distance.

(2) At physical adsorption of molecules, when E_a tends to zero, the high-temperature phase is expected to be more stable than at chemisorption.

(3) If $kT \ll (E_F - E_a)$, the possibility of the transformation is low but grows with temperature.

Also noteworthy is the fact that L generally does not depend on the dispersity of the material, i.e., on S_0. It is related to the degree of charge carrier compensation. For high ionization densities (deep yttrium and vacancy levels), the compensation is the highest and the thickness of the unstable surface α layer, the smallest.

The validity of expression (6) can be judged from Fig. 1. The amount of the α phase is seen to grow with accessible porosity, the latter being proportional to S_0. We were also interested in how the degree of hydration (the amount of water bonded to grains) of β-ZrO_2 powder affects the β-to-α transformation. The experiments were aimed at verifying expression (6). In particular, it was of interest to see whether such a V_α vs. $(E_F - E_a)$ dependence does exist. For this purpose, variously hydrated powders were subjected to a hydrostatic pressure of 1 GPa (10 kbar), and the α amount was measured with X-ray diffraction. Different degrees of hydration were needed to discover the role of water in the transformation. Pressure served as the parameter that controls the $(E_F - E_a)$ value, changing the distance between an adsorbed molecule and the grain surface and thus changing E_a.

Figure 5 plots V_α against hydrostatic pressure P in variously hydrated β-ZrO_2 powders. It is seen that the

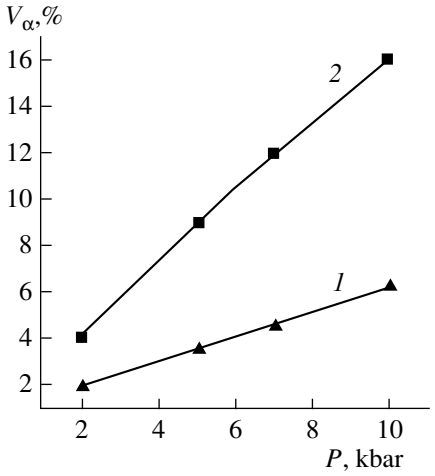


Fig. 5. Amount of nonoriented α phase in variously hydrated β-ZrO_2 powders vs. hydrostatic pressure: (*1*) 5 and (*2*) 15 wt% water.

α amount rises with the degree of hydration. In addition, for more highly hydrated samples, V_α depends more strongly on *P*. The effect of water on the transformation is thus obvious. An increase in V_α and (dV_α/dP) with degree of hydration can be associated with a decrease in $E_F - E_a$ under pressure. Recall that E_a rises under pressure, because adsorbed water molecules are brought close to the grain surface, which strengthens surface–molecule bonds.

Expression (6) also offers a satisfactory explanation of the weak dependence of V_α on temperature (room temperature and boiling point at hydrothermal treatments). This dependence may be insignificant if the condition $kT > (E_F - E_a)$ is fulfilled.

The most important argument in favor of the discussed transformation mechanism seems to be the near-surface crystallographically oriented α phase, which was found in the hydrothermally treated powder and ceramic β-ZrO_2 samples. It was revealed with X-ray topography (URS-2.0 equipment, CoK_α radiation). Morphologically, this phase, unlike the ordinary (nonoriented) α phase, exists as thin lamellas orientationally related to the (111) plane of the β phase. Its volume is no more than 5% of the total β volume, while that of the nonoriented α phase may be much larger. It should be emphasized that both α phases arise simultaneously under hydrothermal treatments. This is not surprising, since the oriented α phase forms on the surface, where the bond density is relatively low. In the bulk, where the bond density is higher, elastic strain energy is scattered over many more crystallographic directions, which prevents the oriented transformation. In other words, the transformation due to adsorption is directly related to the formation of the oriented α phase. In the grains, the thickness of the oriented α layer on the surface was estimated at 3–5 Å. The experimentally discovered presence of the orientated α phase is of paramount value, since it was predicted within the suggested mechanism.

CONCLUSIONS

(1) It was found experimentally that the β-to-α martensite transformation due to water adsorption takes place on the surface of powder and ceramic stabilized β-ZrO_2 samples.

(2) A possible mechanism of the transformation is advanced.

(3) An analytic expression for predicting the stability of the dispersed and ceramic systems and controlling their properties was obtained.

(4) If the suggested electrophysical mechanism is valid, the structure and phase composition of the stabilized materials can be controlled in wide limits by applying, for example, an electric or electromagnetic field. This problem will be highlighted in forthcoming papers.

REFERENCES

1. N. Claussen, Mater. Sci. Eng. **71**, 23 (1985).
2. G. Ya. Akimov, V. M. Timchenko, and I. V. Gorelik, Fiz. Tverd. Tela (St. Petersburg) **36**, 3582 (1994) [Phys. Solid State **36**, 1906 (1994)].
3. M. Yoshimura, Ceram. Bull. **67**, 1950 (1988).
4. I. I. Vishnevskiĭ, A. M. Gavrish, and B. Ya. Sukharevskiĭ, Tr. Ukr. Nauchno-Issled. Inst. Ogneuporov **6** (53), 74 (1962).
5. J. Livage and Ch. Mazieres, C. R. Hebd. Seances Acad. Sci. **261**, 4433 (1965).
6. M. Denoux, C. R. Hebd. Seances Acad. Sci. **260**, 5003 (1965).
7. A. Krauth and H. Meyer, Ber. Dtsch. Keram. Ges. **42** (3), 61 (1965).
8. Yu. M. Polezhaev, Zh. Fiz. Khim. **41**, 2958 (1967).
9. F. Mumpton and R. Roy, J. Am. Ceram. Soc. **43** (5), 234 (1960).
10. V. B. Glushkova, in *Rare-Earth Metals, Alloys, and Compounds* (Nauka, Moscow, 1973), pp. 216–217.
11. F. Hund, J. Electrochem. Soc. **55**, 363 (1951).
12. T. Mitsuhasi, M. Ichihara, and U. Tatsuke, J. Electrochem. Soc. **57** (2), 97 (1974).
13. C. Waguer, Naturwissenschaften **31**, 265 (1943).
14. I. I. Naumov, G. A. Ol'khovik, O. I. Velikokhatnyĭ, and N. N. Aparov, Neorg. Mater. **28**, 805 (1992).
15. F. F. Vol'kenshteĭn, *The Electronic Theory of Catalysis on Semiconductors* (Fizmatgiz, Moscow, 1960; Pergamon, Oxford, 1963).
16. V. N. Strekalovskiĭ, Yu. M. Polezhaev, and S. F. Pal'guev, *Oxides with Impurity Disordering* (Nauka, Moscow, 1987).
17. G. I. Epifanov, *Physical Principles of Microelectronics* (Sov. Radio, Moscow, 1971).
18. M. Lanno and J. Bourgoin, *Point Defects in Semiconductors* (Springer-Verlag, New York, 1981; Mir, Moscow, 1984).

Translated by V. Isaakyan

Technical Physics, Vol. 45, No. 9, 2000, pp. 1159–1163. Translated from Zhurnal Tekhnicheskoĭ Fiziki, Vol. 70, No. 9, 2000, pp. 70–74.

SOLIDS

Nonlocal Josephson Electrodynamics of Electromagnetic Excitations in a Superconducting Magnetic (Three-Dimensional) Thin Film

A. I. Lomtev

Donetsk Physicotechnical Institute, National Academy of Sciences of Ukraine, Donetsk, 340114 Ukraine
E-mail: lomtev@kinetic.ac.donetsk.ua

Received November 17, 1999

Abstract—The integro-differential equation for the dynamics of the phase difference with temporal and spatial nonlocality for a Josephson junction in a superconducting three-dimensional magnetic thin film is derived. The equation is valid for any type of magnetic ordering. The magnetic subsystem is shown to significantly renormalize the spectrum of small-amplitude electromagnetic excitations, which, as a result, are damped. © *2000 MAIK "Nauka/Interperiodica".*

(1) By now, equations of the nonlocal Josephson electrodynamics have been derived and investigated in the following six cases: 1) the butted tunneling junction formed by two superconducting ultrathin films with thicknesses much less than the London penetration depth; 2) the tunneling junction formed by bulk superconductors with thicknesses much greater than the London penetration depth; 3) the tunneling junction formed by superconducting layers of finite thickness in the direction orthogonal to a magnetic field; 4) the butted tunneling junction formed by superconducting plates of finite thickness along a magnetic field; 5) the inclined (sloped) Josephson junction formed by superconducting plates of finite thickness along a magnetic field; 6) the Josephson junction in a superconducting magnetic (two-dimensional) thin film with a thickness much less than the London penetration depth.

As is shown in [1–8], the nonlocality effects may be significant even in contacts with large thickness d ($d \gg \lambda$, where λ is the London penetration depth) along the magnetic field (in the vortex direction), i.e., in the situations previously considered in the local approximation. In the opposite extreme case of thin-film contacts, when $d \ll \lambda$, nonlocality is very significant and becomes a dominant factor. The corresponding equations were derived and studied in [9–12]. A Josephson junction formed by two superconducting layers of finite thickness in the orthogonal direction to a magnetic field of vortices was studied in [13]. The butted contact and the inclined (sloped) junction of finite thickness along the magnetic field of vortices for an arbitrary ratio d/λ were considered in [14, 15], respectively. In a recent study [16], an integro-differential equation for the phase difference dynamics with spatial and temporal nonlocality was derived for the Josephson junction in a superconducting magnetic (two-dimensional) thin film.

At the present time, there are a large number of magnetic superconductors exhibiting new unique properties [17–19]. In addition to ternary compounds [20], superconductivity and magnetism coexist in high-temperature superconductors (HTS), such as REBaCuO and RECuO, where RE is a rare-earth ion. The strong antiferromagnetic correlation of copper spins in CuO_2-planes is one of the most important features of HTS materials [21].

(2) The investigation of the nonlocal electrodynamics of Josephson junctions in ultrathin magnetic superconducting films (for which $d \ll \lambda$) has become an urgent problem. In this case, the problem may be reduced to that of an infinite thin two-dimensional superconducting plane where the superconducting physical characteristics of a contact are independent of the coordinate perpendicular to the film plane; i.e., the current and magnetic field are assumed to alter slightly along the plate thickness [22]. Suppose, however, that the magnetic properties of the system under consideration are nevertheless three-dimensional; i.e., the permeability of the film depends on three spatial and temporal coordinates $\mu(\mathbf{r} - \mathbf{r}', t - t')$, where $\mathbf{r} = (x, y, z)$.

Let the geometry of the problem be the following: the film plane coincides with the $X0Y$ plane, the current flows along the y-axis, and a line of weak links lies along the x-axis. Following [23, 24], we consider the Maxwell equations for the magnetic induction $\mathbf{B}(\mathbf{r}, t) = \mathrm{curl}\mathbf{A}(\mathbf{r}, t)$ ($\mathbf{A}(\mathbf{r}, t)$ is the vector potential) defined by the sum of the magnetic field $\mathbf{H}(\mathbf{r}, t)$ produced by the superconducting current $\mathbf{j}(\mathbf{r}, t)$ and the magnetization $\mathbf{M}(\mathbf{r}, t)$:

$$\mathrm{curl}\mathbf{B}(\mathbf{r}, t) = \frac{4\pi}{c}\mathbf{j}(\mathbf{r}, t) + 4\pi\mathrm{curl}\mathbf{M}(\mathbf{r}, t). \qquad (1)$$

The relation between the current, the potential, and the phase of the order parameter $\Theta(\mathbf{r}, t)$ (for the London

calibration of the potential $\mathrm{div}\mathbf{A}(\mathbf{r}, t) = 0$) is determined by the expression

$$\mathbf{j}(\mathbf{r}, t) = (c/4\pi\lambda^2)[\mathbf{S}(\boldsymbol{\rho}, t) - \mathbf{A}(\mathbf{r}, t)], \quad (2)$$

where $\boldsymbol{\rho} = (x, y)$ and the vector $\mathbf{S}(\boldsymbol{\rho}, t)$ is defined by the phase gradient

$$\mathbf{S}(\boldsymbol{\rho}, t) = \frac{\Phi_0}{2\pi}\nabla\Theta(\boldsymbol{\rho}, t). \quad (3)$$

Here, Φ_0 is the flux quantum. The phase of the order parameter satisfies the condition

$$\mathrm{curl}\mathbf{S}(\boldsymbol{\rho}, t) = -\hat{\mathbf{z}}\frac{\Phi_0}{2\pi}\frac{\partial\varphi(x, t)}{\partial x}\delta(y), \quad (4)$$

where $\hat{\mathbf{z}}$ is a unit vector along the z-axis, $\delta(y)$ is the Dirac delta function, and $\varphi(x, t)$ is determined by the phase difference of the order parameter across a junction:

$$\varphi(x, t) = \Theta(x, +0, t) - \Theta(x, -0, t). \quad (5)$$

The system of Eqs. (1) and (2) leads to the relationship

$$\mathrm{curl}\mathbf{H}(\mathbf{r}, t) = \lambda^{-2}[\mathbf{S}(\boldsymbol{\rho}, t) - \mathbf{A}(\mathbf{r}, t)], \quad (6)$$

which can be written for an ultrathin film $d \ll \lambda$ in the form

$$\mathrm{curl}\mathbf{H}(\mathbf{r}, t) = \lambda_{\mathrm{eff}}^{-1}[\mathbf{S}(\boldsymbol{\rho}, t) - \mathbf{A}(\mathbf{r}, t)]\delta(z), \quad (7)$$

where $\lambda_{\mathrm{eff}} = \lambda^2/d$ is the effective Pearl penetration depth.

The magnetic field $\mathbf{H}(\mathbf{r}, t)$ is related to the magnetic induction $\mathbf{B}(\mathbf{r}, t)$ by the integral relationship

$$\mathbf{H}(\mathbf{r}, t) = \int_{-\infty}^{t} dt' \int_{-\infty}^{\infty} d\mathbf{r}'\mu^{-1}(\mathbf{r} - \mathbf{r}', t - t')\mathbf{B}(\mathbf{r}', t'). \quad (8)$$

From the system of relationships (7) and (8), we obtain the equation for the vector potential

$$\mathrm{curl}\left\{\int_{-\infty}^{t} dt' \int_{-\infty}^{\infty} d\mathbf{r}'\mu^{-1}(\mathbf{r} - \mathbf{r}', t - t')\mathrm{curl}\mathbf{A}(\mathbf{r}', t')\right\} = \lambda_{\mathrm{eff}}^{-1}[\mathbf{S}(\boldsymbol{\rho}, t) - \mathbf{A}(\boldsymbol{\rho}, z, t)]\delta(z), \quad (9)$$

which is expressed by means of the source field $\mathbf{S}(\boldsymbol{\rho}, t)$.

From the continuity equation $\mathrm{div}\mathbf{j}(\mathbf{r}, t) = 0$ and Eq. (2), it follows that $\mathrm{div}\mathbf{S}(\boldsymbol{\rho}, t) = 0$. Consequently, the vector field $\mathbf{S}(\boldsymbol{\rho}, t)$ may be represented as the curl of a solenoidal vector field $\mathbf{F}(\boldsymbol{\rho}, t)$ in the form

$$\boldsymbol{S}(\boldsymbol{\rho}, t) = \mathrm{curl}\mathbf{F}(\boldsymbol{\rho}, t), \quad (10)$$

where

$$\mathbf{F}(\boldsymbol{\rho}, t) = \hat{\mathbf{z}}F(\boldsymbol{\rho}, r), \quad \mathrm{div}\mathbf{F}(\boldsymbol{\rho}, t) = 0, \quad (11)$$

Substituting expression (10) with allowance for (11) into relationship (4), we obtain an equation describing the function $F(\boldsymbol{\rho}, t)$:

$$\Delta F(\boldsymbol{\rho}, t) = \frac{\Phi_0}{2\pi}\frac{\partial\varphi(x, t)}{\partial t}\delta(y). \quad (12)$$

On the one hand, the current density across the Josephson junction $j_y(x, 0, t)$ is equal to the sum, for example, of three terms: the Josephson supercurrent density, the normal (quasiparticle) current, and the capacitive displacement current

$$j_y(x, 0, t) = j_c\left[\sin\varphi(x, t) + \frac{\beta}{\omega_J^2}\frac{\partial\varphi(x, t)}{\partial t} + \frac{1}{\omega_J^2}\frac{\partial^2\varphi(x, t)}{\partial t^2}\right]d\delta(z), \quad (13)$$

where β is the dissipative parameter and j_c and ω_J are the critical current and the Josephson frequency, respectively.

On the other hand, it follows from Eq. (2) that the same current density $j_y(x, 0, t)$ may be represented in the form

$$j_y(x, 0, t) = \frac{c}{4\pi\lambda_{\mathrm{eff}}}[S_y(x, 0, t) - A_y(x, 0, 0, t)]\delta(z). \quad (14)$$

Equating the expressions for the current density (13) and (14), we see that

$$j_c\left[\sin\varphi(x, t) + \frac{\beta}{\omega_J^2}\frac{\partial\varphi(x, t)}{\partial t} + \frac{1}{\omega_J^2}\frac{\partial^2\varphi(x, t)}{\partial t^2}\right] = \frac{c}{4\pi\lambda^2}[S_y(x, 0, t) - A_y(x, 0, 0, t)]. \quad (15)$$

In order to derive a closed equation for the phase difference dynamics $\varphi(x, t)$, it is necessary to find a functional relationship between

$$\Delta_y(x, t) = S_y(x, 0, t) - A_y(x, 0, 0, t) \quad (16)$$

and $\varphi(x, t)$. To solve the system of equations (9), we introduce the Fourier transforms for the vectors $\mathbf{A}(\mathbf{r}, t)$ and $\mathbf{S}(\boldsymbol{\rho}, t)$ and the scalar function $\mu^{-1}(\mathbf{r}, t)$ in the form

$$\mathbf{A}(\boldsymbol{\rho}, z, t) = \int\frac{d^2qdq\omega}{(2\pi)^4}\exp(i\mathbf{q}\boldsymbol{\rho} + ipz - i\omega t)\mathbf{A}(\mathbf{q}, p, \omega), \quad (17)$$

$$\mathbf{S}(\boldsymbol{\rho}, t) = \int\frac{d^2qd\omega}{(2\pi)^3}\exp(i\mathbf{q}\boldsymbol{\rho} - i\omega t)\mathbf{S}(\mathbf{q}, \omega) \quad (18)$$

and

$$\mu^{-1}(\mathbf{r}, t) = \int\frac{d^2qdpq\omega}{(2\pi)^4}\exp(i\mathbf{q}\boldsymbol{\rho} + ipz - i\omega t)\mu^{-1}(\mathbf{q}, p, \omega). \quad (19)$$

Using relationships (17) and (18), we represent the quantity $\Delta_y(x, t)$ in the integral form

$$\Delta_y(x, t) = \frac{1}{(2\pi)^2}\int_{-\infty}^{\infty} d\omega \int_0^{\infty} q\,dq \int_{-\pi}^{\pi} d\vartheta [S_y(q, \omega) - A_y(q, \omega)] \times \exp(iqx\cos\vartheta - i\omega t), \quad (20)$$

where $q = (q_x^2 + q_y^2)^{1/2}$, ϑ is the polar angle on the plane (q_x, q_y), and

$$\mathbf{A}(\mathbf{q}, \omega) = \int_{-\infty}^{\infty} \frac{dp}{2\pi} \mathbf{A}(\mathbf{q}, p, \omega). \quad (21)$$

We now represent the solution of Eq. (9) in the form

$$\mathbf{A}(\mathbf{q}, p, \omega) = \lambda_{\text{eff}}^{-1} \frac{\mu(\mathbf{q}, p, \omega)}{(p^2 + q^2)} [\mathbf{S}(\mathbf{q}, \omega) - \mathbf{A}(\mathbf{q}, \omega)]. \quad (22)$$

As follows from definition (21) and Eq. (22), the relation between $\mathbf{S}(\mathbf{q}, \omega)$ and $\mathbf{A}(\mathbf{q}, \omega)$ can be written in the form

$$\mathbf{A}(\mathbf{q}, \omega) = \left(1 + \frac{\lambda_{\text{eff}}}{R(\mathbf{q}, \omega)}\right)^{-1} \mathbf{S}(\mathbf{q}, \omega), \quad (23)$$

where $R(\mathbf{q}, \omega)$ is determined by the integral

$$R(\mathbf{q}, \omega) = \int_{-\infty}^{\infty} \frac{dp}{2\pi} \frac{\mu(\mathbf{q}, p, \omega)}{p^2 + q^2}; \quad (24)$$

therefore,

$$S_y(\mathbf{q}, \omega) - A_y(\mathbf{q}, \omega) = \frac{\lambda_{\text{eff}}}{R(\mathbf{q}, \omega) + \lambda_{\text{eff}}} S_y(\mathbf{q}, \omega). \quad (25)$$

To find $S_y(\mathbf{q}, \omega)$, it is necessary to differentiate Eq. (12) with respect to the coordinate x, since $S_y(\boldsymbol{\rho}, t)$ is equal to $\partial F(\boldsymbol{\rho}, t)/\partial x$. As a result, we have

$$\Delta S_y(\boldsymbol{\rho}, t) = -\frac{\Phi_0}{2\pi} \frac{\partial^2 \varphi(x, t)}{\partial x^2} \delta(y). \quad (26)$$

The relationship written below is the corollary of this equation:

$$S_y(\mathbf{q}, \omega) = \frac{\Phi_0}{2\pi q^2} \int_{-\infty}^{\infty} dx \int_{-\infty}^{\infty} dt \times \exp(-iq\cos\vartheta x + i\omega t) \frac{\partial^2 \varphi(x, t)}{\partial x^2}. \quad (27)$$

Combining expressions (27), (25), (20), and (15), we obtain the integro-differential equation for the phase difference dynamics across the Josephson junction, valid for any type of magnetic ordering in a superconducting film:

$$\sin\varphi(x, t) + \frac{\beta}{\omega_J^2} \frac{\partial \varphi(x, t)}{\partial t} + \frac{1}{\omega_J^2} \frac{\partial^2 \varphi(x, t)}{\partial t^2} = I_J \int_{-\infty}^{\infty} dx' \int_{-\infty}^{\infty} dt' K\left(\frac{x - x'}{2\lambda_{\text{eff}}}, t - t'\right) \frac{\partial^2 \varphi(x', t')}{\partial x'^2}, \quad (28)$$

where $I_J = \lambda_J^2/\lambda$, λ_J is the Josephson penetration depth and the integral kernel

$$K\left(\frac{x - x'}{2\lambda_{\text{eff}}}, t - t'\right),$$

which is nonlocal over spatial and temporal variables, takes the form

$$K\left(\frac{x - x'}{2\lambda_{\text{eff}}}, t - t'\right) = \int_0^{\infty} \frac{dq}{\pi} \int_{-\infty}^{\infty} \frac{d\omega}{2\pi} \times \frac{\lambda_{\text{eff}} J_0[q(x - x')] \exp[-i\omega(t - t')]}{q[R(\mathbf{q}, \boldsymbol{\omega}) + \lambda_{\text{eff}}]}. \quad (29)$$

Here J_0 is the zeroth-order Bessel function. In view of (29), the temporal nonlocality of Eq. (28) is governed by the frequency dispersion of the permeability $\mu(\mathbf{q}, p, \omega)$ (through the function $R(\mathbf{q}, \omega)$).

(3). Since $\lambda \gg a$ (a is the crystal lattice constant), it is natural to use a hydrodynamic description of the magnetic subsystem. Considering only the paramagnetic temperature range for the permeability, we have the expression [24, 25]

$$\mu(\mathbf{q}, p, \omega) = 1 + i4\pi \frac{\chi_0 D(q^2 + p^2)}{\omega + iD(q^2 + p^2)}, \quad (30)$$

where χ_0 is the static permeability and D is the coefficient of the spin diffusion for three-dimensional Heisenberg magnetic substances.

In a strict sense, the superconducting currents shield the long-wavelength part of the exchange and electromagnetic interactions, renormalizing the parameters of the magnetic subsystem [26]. However, this circumstance is not taken into account, since we are dealing with the paramagnetic temperature range and are interested in evaluating only the order of magnitude.

According to formula (24) and in view of (30), the function $R(\mathbf{q}, \omega)$ takes the form

$$R(\mathbf{q}, \omega) = \frac{f_0(q, \omega) + i4\pi\chi_0 q}{2q f_0(q, \omega)}, \quad (31)$$

where, in its turn, the function $f_0(q, \omega)$ is equal to

$$f_0(q, \omega) = [(q^4 + \omega^2/D^2)^{1/2}/2 - q^2/2]^{1/2} + i[(q^4 + \omega^2/D^2)^{1/2}/2 + q^2/2]^{1/2}. \quad (32)$$

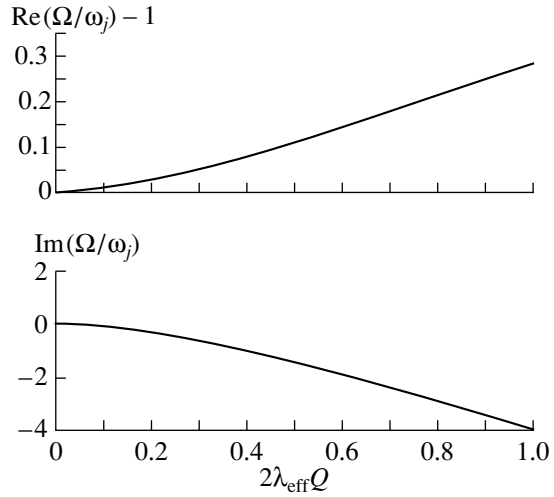


Fig. 1.

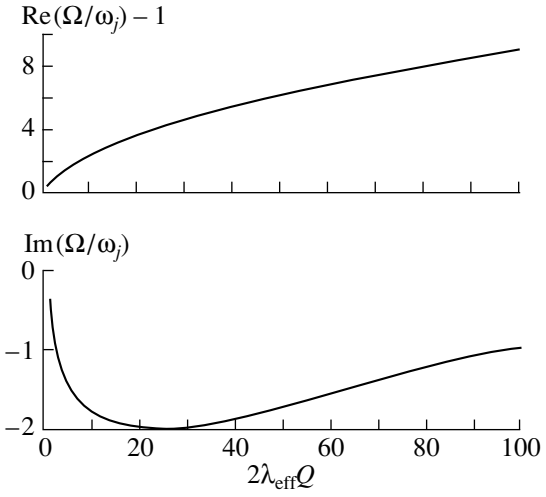


Fig. 2.

We consider the spectrum of small-amplitude electromagnetic excitations with the wave vector Q and the frequency Ω,

$$\varphi(x, t) = \varphi_0 \exp[i(Qx - \Omega t)], \quad |\varphi_0| \ll 1, \qquad (33)$$

propagating along the Josephson junction.

From Eq. (28), with allowance for (29), (31), and (32) for solutions (33) in the nondissipative limit (at $\beta = 0$) and in the absence of the transport current and an external magnetic field, we obtain the dispersion equation $\hat{\Omega} = \hat{\Omega}(\hat{Q})$:

$$\hat{\Omega}^2 = 1 + \frac{2}{\pi} I \hat{Q}^2 I(\hat{Q}, \hat{\Omega}), \qquad (34)$$

where

$$I(\hat{Q}, \hat{\Omega}) = \int_0^\infty \frac{dx F_0(\hat{Q}\cosh x, \hat{\Omega})}{F_0(\hat{Q}\cosh x, \hat{\Omega})(1 + \hat{Q}\cosh x) + i4\pi\chi_0 \hat{Q}\cosh x} \qquad (35)$$

and

$$F_0(x, y) = [(x^4 + \eta^2 y^2)^{1/2}/2 - x^2/2]^{1/2} + i[(x^4 + \eta^2 y^2)^{1/2}/2 + x^2/2]^{1/2}. \qquad (36)$$

Here, the dimensionless quantities $\hat{Q} = 2\lambda_{\text{eff}} Q$ and $\hat{\Omega} = \Omega/\omega_J$ are introduced, and the following designations $I = I_J/2\lambda_{\text{eff}}$, $\eta = \omega_J/\Omega_{\text{eff}}$, and $\Omega_{\text{eff}} = D/(2\lambda_{\text{eff}})^2$ are used.

At $\chi_0 = 0$ (for a nonmagnetic superconductor), Eq. (34), with allowance for (35) and (36), leads to the well-known spectrum of small-amplitude electromagnetic waves derived in [12] for the long-wave and short-wave ranges. The frequency $\hat{\Omega}$ is a real function of the wave vector $\hat{Q}$, and the damping of modes in the nondissipative limit (at $\beta = 0$) is absent. The other situation is observed for a magnetic superconductor at $\chi_0 = 0$. For antiferromagnetics, a typical value of χ_0 lies between 10^{-5} and 10^{-3}. Numerical simulations have shown that, for $\chi_0 = 10^{-3}$ (compared with the susceptibility of the copper subsystem in HTS materials), $I = 1$ and $\eta = 10^3$ from Eq. (34) (implicit relative to $\hat{\Omega}$), and with allowance for (35) and (36), we obtain the complex solution $\hat{\Omega} = \hat{\Omega}(\hat{Q}) = \text{Re}\hat{\Omega}(\hat{Q}) + i\text{Im}\hat{\Omega}(\hat{Q})$ with a small negative imaginary part satisfying the condition

$$\text{Im}\hat{\Omega}(\hat{Q})/\text{Re}\hat{\Omega}(\hat{Q}) \ll 1. \qquad (37)$$

This indicates that, due to the effect of the magnetic subsystem, damping of small-amplitude electromagnetic excitations exists. However, no rising solutions were observed.

It is clear from the above that the nonlocality over spatial variables for the Josephson electrodynamics of magnetic superconducting thin films is very significant and becomes the main factor determining its specific features. The temporal nonlocality of the Josephson electrodynamics of magnetic superconducting thin films is associated with the frequency dispersion of the material permeability, which is finally produced by the delay processes.

According to (30), damping of the electromagnetic excitations in a Josephson junction in magnetic superconducting thin films is related to the complexity of the material permeability functions. The physical mechanism of such damping is a process of spin diffusion with a finite relaxation time, during which a small part of the electromagnetic field energy is irreversibly dissipated. The presence itself of the damping of linear electromagnetic waves in a Josephson junction (in the nondissipative limit) may provide information about whether or not the superconductor is magnetic.

For the wave vector values $0 \le \hat{Q} \le 1$ and $1 \le \hat{Q} \le 100$, the spectrum renormalized by the magnetic subsystem and the damping decrement of electromagnetic waves in a Josephson junction are shown in Figs. 1 and 2, respectively.

ACKNOWLEDGMENTS

I am grateful to Yu.V. Medvedev for critical remarks; A.S. Zel'tser for help in performing the numerical simulations; and V.N. Krivoruchko, Yu.E. Kuzovleva, Yu.G. Pashkevich, and A.É. Fillipov for fruitful discussions, attention, and support.

REFERENCES

1. Yu. M. Aliev, V. P. Silin, and S. A. Uryupin, Sverkhprovodimost: Fiz. Khim. Tekh. **5** (2), 228 (1992).
2. A. Gurevich, Phys. Rev. B **46** (5), 3187 (1992).
3. Yu. M. Aliev, V. P. Silin, and S. A. Uryupin, Pis'ma Zh. Éksp. Teor. Fiz. **57** (3), 187 (1993) [JETP Lett. **57**, 193 (1993)].
4. Yu. M. Aliev and V. P. Silin, Zh. Éksp. Teor. Fiz. **104**, 2526 (1993) [JETP **77**, 142 (1993)].
5. Yu. M. Aliev and V. P. Silin, Phys. Lett. A **177** (3), 259 (1993).
6. V. P. Silin, Pis'ma Zh. Éksp. Teor. Fiz. **58** (9), 726 (1993) [JETP Lett. **58**, 701 (1993)].
7. G. L. Alfimov and V. P. Silin, Zh. Éksp. Teor. Fiz. **106**, 671 (1994) [JETP **79**, 369 (1994)].
8. V. P. Silin, Pis'ma Zh. Éksp. Teor. Fiz. **60** (6), 442 (1994) [JETP Lett. **60**, 460 (1994)].
9. Yu. M. Ivanchenko and T. K. Soboleva, Pis'ma Zh. Éksp. Teor. Fiz. **51** (2), 100 (1990) [JETP Lett. **51**, 114 (1990)].
10. Yu. M. Ivanchenko and T. K. Soboleva, Phys. Lett. A **147** (1), 65 (1990).
11. Yu. M. Ivanchenko and T. K. Soboleva, Fiz. Tverd. Tela (Leningrad) **32** (7), 2029 (1990) [Sov. Phys. Solid State **32**, 1181 (1990)].
12. R. G. Mints and I. B. Snapiro, Phys. Rev. B **51** (5), 3054 (1995).
13. I. O. Kulik and I. K. Yanson, in *Josephson Effect in Superconducting Tunnel Structures* (Nauka, Moscow, 1970), Chap. 1.
14. Yu. E. Kuzovlev and A. I. Lomtev, Zh. Éksp. Teor. Fiz. **111** (5), 1803 (1997) [JETP **84**, 986 (1997)].
15. A. I. Lomtev, Zh. Éksp. Teor. Fiz. **113** (6), 2256 (1998) [JETP **86**, 1234 (1998)].
16. A. I. Lomtev, Pis'ma Zh. Éksp. Teor. Fiz. **69** (2), 132 (1999) [JETP Lett. **69**, 148 (1999)].
17. A. I. Buzdin, L. N. Bulaevskiĭ, M. L. Kulich, and S. V. Panyukov, Usp. Fiz. Nauk **144** (4), 597 (1984) [Sov. Phys. Usp. **27**, 927 (1984)].
18. A. I. Buzdin and L. N. Bulaevskiĭ, Usp. Fiz. Nauk **149** (1), 45 (1986) [Sov. Phys. Usp. **29**, 412 (1986)].
19. Yu. A. Izyumov, N. M. Plakida, and Yu. N. Skryabin, Usp. Fiz. Nauk **159** (4), 621 (1989) [Sov. Phys. Usp. **32**, 1060 (1989)].
20. *Superconductivity in Ternary Compounds*, Vol. 2: *Superconductivity and Magnetism*, Ed. by E. Fisher and M. Maple (Springer, Heidelberg, 1982; Mir, Moscow, 1985),
21. *Physical Properties of High-Temperature Superconductors*, Ed. by D. M. Ginsberg (World Scientific, Singapore, 1989; Mir, Moscow, 1990), Chaps. 4, 6.
22. A. A. Abrikosov, in *Fundamentals of the Theory of Metals* (Nauka, Moscow, 1987; North-Holland, Amsterdam, 1988), Chap. 18.
23. E. M. Lifshitz and L. P. Pitaevskiĭ, in *Course of Theoretical Physics*, Vol. 5: *Statistical Physics* (Nauka, Moscow, 1978; Pergamon, New York, 1980), Part 2, Chap. 5.
24. H. Umezawa, H. Matsumoto, and M. Tachiki, in *Thermo-Field Dynamics and Condensed States* (North-Holland, Amsterdam, 1982; Mir, Moscow, 1985), Chap. 11.
25. B. I. Halperin and P. C. Honenberg, Phys. Rev. **188** (2), 898 (1969).
26. A. I. Buzdin, Pis'ma Zh. Éksp. Teor. Fiz. **40** (5), 193 (1984) [JETP Lett. **40**, 956 (1984)].

Translated by Yu. Vishnyakov

Technical Physics, Vol. 45, No. 9, 2000, pp. 1164–1169. Translated from Zhurnal Tekhnicheskoĭ Fiziki, Vol. 70, No. 9, 2000, pp. 75–81.

OPTICS, QUANTUM ELECTRONICS

Light Propagation in a Directional Coupler with Saturable Nonlinearity

P. I. Khadzhi and O. K. Orlov

Pridnestrovsk State University, Tiraspol, 278000 Moldova

Received January 11, 1999; in final form, December 10, 1999

Abstract—The propagation of laser radiation in a nonlinear directional coupler with identical optical waveguides is studied theoretically. Account is taken of refractive index saturation. Bifurcations of the intensity distribution and transmission characteristic are revealed, and respective regions in the parameter plane are delineated. Two points of input intensity self-switching are found. A coupling length vs. normalized input intensity is investigated for different values of the nonlinearity parameter. Distinctive features of the coupler with saturated nonlinearity compared with Kerr-type couplers are highlighted. *© 2000 MAIK "Nauka/Interperiodica".*

INTRODUCTION

Research on nonlinear directional couplers (NLDCs) with distributed coupling of copropagating light waves sets the stage for the progress toward ultrafast all-optical switches [1–3]. Under nonlinear conditions, when the refractive index depends on light intensity, the transmission loss of each wave depends on input intensity. Furthermore, the self-switching of the waves was predicted [1]. The phenomenon shows up as sharp changes in the intensities of both waves at the output in response to a slight variation in the intensity of one of them at the input. A satisfactory theory has been developed for the propagation in NLDCs where the intensity-dependent refractive index or propagation constant obeys the Kerr model. Exact analytical solutions to the nonlinear equations for wave intensities have been obtained in terms of elliptic functions, and key features of switching processes have been investigated. However, Kerr-type NLDCs are of limited practical value, because their switching power has turned out to be rather high. Much lower switching powers are offered by semiconductor NLDCs, which are remarkable for their giant resonant-type nonlinearities. Semiconductors are basically non-Kerr-type media, although they can be regarded as Kerr-type materials as long as the susceptibility is expanded in powers of intensity up to a linear term. Such an approximation may, however, fail at high excitation levels. For example, a medium modeled by two-level atoms exhibits saturation, which considerably changes the refractive index and absorption coefficient. In [4–12], the effect of refractive-index saturation was studied numerically. It was shown that the performance of such NLDCs qualitatively differs from that of Kerr-type NLDCs. A non-Kerr-type NLDC with a power-law index–intensity characteristic was considered in [13]. Yet, works [4–13] give no exact analytical solutions of the nonlinear equations in the non-Kerr-type case.

Even more promising seem to be materials with excitonic and biexcitonic optical nonlinearities [14–16]. They offer a wealth of coherent nonlinear effects. On the other hand, their index–intensity characteristic, which can self-consistently be derived from Bloch-like equations, may be rather complicated.

In view of the foregoing, it seems worthwhile to seek analytical solutions of the NLDC equations for non-Kerr-type media. This paper presents such solutions, the equations being solved in quadrature. Exponential saturation of the refractive index (propagation constant) is taken into account.

BASIC EQUATIONS

Consider an NLDC with two identical optical waveguides. Following [5, 9], the dependence of the propagation constant β on the wave intensity J is defined as

$$\beta = \beta_0 + \alpha \exp\left(-\frac{J}{J_S}\right), \qquad (1)$$

where β_0 and α are constants and J_S is the saturation intensity.

Formula (1) will be used for modeling the nonlinear medium with a saturable refractive index (propagation constant). As J increases from zero to $J \gg J_S$, $\beta(J)$ decreases monotonically from $\beta_0 + \alpha$, approaching β_0. Expanding (1) in powers of J/J_S up to a linear term, we obtain an expression for β in the Kerr limit for a self-defocusing medium. As is usually the case, the coupling constant γ of the waveguides is independent of J [1–13]. Optical absorption in the medium is assumed to be negligible.

Then, the amplitudes E_1 and E_2 of coupled waves propagating through the respective waveguides along their axes obey the following nonlinear differential

1063-7842/00/4509-1164$20.00

equations [1–13]:

$$\frac{dE_1}{dx} = -i\left(\beta_0 + \alpha\exp\left(-\frac{J_1}{J_S}\right)\right)E_1 + i\gamma E_2, \quad (2)$$

$$\frac{dE_2}{dx} = -i\left(\beta_0 + \alpha\exp\left(-\frac{J_2}{J_S}\right)\right)E_2 + i\gamma E_1. \quad (3)$$

It is usually assumed that the fields differ in amplitude and phase. System (2)–(3) is then replaced by nonlinear equations for amplitudes and phase difference [1]. We suggest another approach that seems to be more convenient and enables one to solve system (2)–(3) by quadrature. Let us define new functions

$$J_1 = \frac{c}{8\pi}|E_1|^2, \quad J_2 = \frac{c}{8\pi}|E_2|^2,$$
$$Q = \frac{c}{8\pi}(E_1^*E_2 - E_2^*E_1), \quad R = \frac{c}{8\pi}(E_1^*E_2 + E_2^*E_1). \quad (4)$$

Using (2), (3), and the conjugate equations, we obtain the following system of coupled nonlinear differential equations:

$$\frac{dJ_1}{dx} = i\gamma Q, \quad \frac{dJ_2}{dx} = -i\gamma Q, \quad (5)$$

$$\frac{dQ}{dx} = i\alpha\left(\exp\left(-\frac{J_1}{J_S}\right) - \exp\left(-\frac{J_2}{J_S}\right)\right)R + 2i\gamma(J_1 - J_2), \quad (6)$$

$$\frac{dR}{dx} = i\alpha\left(\exp\left(-\frac{J_1}{J_S}\right) - \exp\left(-\frac{J_2}{J_S}\right)\right)Q. \quad (7)$$

Let us solve the system for the case where one of the waveguides (waveguide 1) is fed by laser radiation of amplitude E_0 and intensity J_0. Then the boundary conditions for (5)–(7) are

$$J_{1|x=0} = J_0, \ J_{2|x=0} = 0, \ Q_{|x=0} = R_{|x=0} = 0. \quad (8)$$

From equations (5), one easily obtains the integral of motion

$$J_1 + J_2 = J_0, \quad (9)$$

according to the energy conservation law. As suggested by (5), (7), and (8), the second integral of motion is

$$R = \frac{\alpha}{\gamma}J_S\left(1 - \exp\left(-\frac{J_1}{J_S}\right)\right)\left(1 - \exp\left(-\frac{J_2}{J_S}\right)\right). \quad (10)$$

Substituting (10) into (6) and using (5), we come to the third integral:

$$Q^2 = -4J_1(J_0 - J_1) + \frac{\alpha^2}{\gamma^2}J_S^2 \times\left(1 - \exp\left(-\frac{J_1}{J_S}\right)\right)^2\left(1 - \exp\left(-\frac{J_2}{J_S}\right)\right)^2. \quad (11)$$

From (9)–(11), one easily finds the fourth integral, which involves all of the functions:

$$R^2 = Q^2 + 4J_1J_2. \quad (12)$$

Finally, substituting (11) into (5) immediately gives the nonlinear differential equation for intensity distribution over the length of waveguide 1:

$$\frac{dJ_1}{dx} = -2\gamma\left\{J_1(J_0 - J_1) - \frac{\alpha^2}{4\gamma^2}J_S^2 \times\left(1 - \exp\left(-\frac{J_1}{J_S}\right)\right)^2\left(1 - \exp\left(-\frac{J_2}{J_S}\right)\right)^2\right\}^{1/2}. \quad (13)$$

Having found $J_1(x)$ from (13), one determines $J_2(x)$ with the help of (9). Let us define normalized quantities

$$Y_1 = \frac{J_1}{J_S}, \quad Y_2 = \frac{J_2}{J_S}, \quad Y_0 = \frac{J_0}{J_S},$$
$$z = 2\gamma x, \quad a = \frac{\alpha^2}{4\gamma^2}. \quad (14)$$

Then (13) is solved by quadrature as follows:

$$\int_{Y_1}^{Y_0}[Y(Y_0 - Y) - a(1 - \exp(-Y))^2 \times(1 - \exp(Y - Y_0))^2]^{-1/2}dY = 2\gamma x. \quad (15)$$

To facilitate the qualitative analysis of solutions, we rewrite (13) in the form

$$\left(\frac{dY_1}{dz}\right)^2 + W(Y_1) = 0, \quad (16)$$

where

$$W(Y_1) = -Y_1(Y_0 - Y_1) + a(1 - \exp(Y_1))^2 \times(1 - \exp(Y_1 - Y_0))^2. \quad (17)$$

We thus arrive at the equation of nonlinear conservative oscillator with the potential energy $W(Y_1)$. The motion of the oscillator is confined to the region where $W(Y_1) < 0$. Formulas (15) and (17) indicate that the behavior of the solutions is governed by two system parameters, Y_0 and a. The former is the normalized intensity at the input of waveguide 1, and the latter characterizes nonlinearity.

It follows from (15) and (17) that the intensity in waveguide 1 varies periodically from the upper limit Y_0 to the minimum Y_m. The latter can be found from the equation

$$Y_m(Y_0 - Y_m) = a(1 - \exp(Y_m))^2(1 - \exp(Y_m - Y_0))^2 \quad (18)$$

by taking the positive root closest to Y_0.

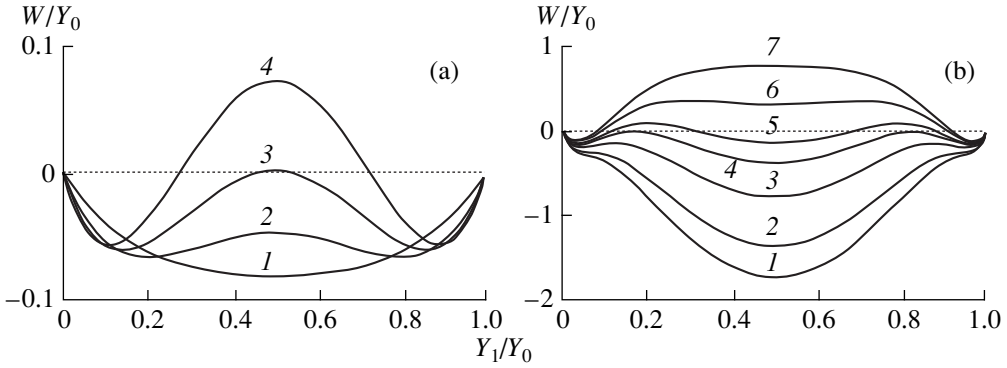


Fig. 1. Potential energy W of the equivalent nonlinear oscillator vs. Y_1 for different values of a. In panel (a), a = (*1*) 0.3, (*2*) 0.5, (*3*) 0.5988, and (*4*) 0.7. In panel (b), a = (*1*) 13, (*2*) 12, (*3*) 10.5, (*4*) 9.5537, (*5*) 9, (*6*) 8, and (*7*) 7.

BIFURCATIONS OF THE OSCILLATOR POTENTIAL ENERGY

The properties of intensity distributions in NLDC waveguides can be explained by investigating the behavior of the oscillator potential energy $W(Y_1)$ at different values of a and Y_0 (Fig. 1) and by plotting bifurcation curves in the (Y_0, a) plane (Fig. 2). These curves separate the plane into regions where the behavior of $W(Y_1)$ qualitatively differs. It follows from (17) that $W(Y_1)$ has the zeros $Y_1 = 0$ and $Y_1 = Y_0$. Based on the requirement that $W(Y_1)$ and its derivative be zero, we infer that there exists a critical value of Y_0, Y_{0C} at which the behavior of $W(Y_1)$ changes qualitatively. This value can be determined from the equation

$$\cosh\left(\frac{Y_{0C}}{2}\right) = 1 + \frac{1}{2}Y_{0C}. \qquad (19)$$

Equation (19) has a single positive real root: $Y_{0C} \approx 5.97$. The corresponding value of a,

$$a_{0C} = Y_0^{-2}\exp(Y_0), \qquad (20)$$

equals $a_{0C} = 10.98$. In the (Y_0, a) plane, Y_{0C} and a_{0C} are represented by point D (see Fig. 2).

Equation (16) implies that, with $0 \le Y_0 \le Y_{0C}$, increasing a produces a smoothly growing peak in the curve $W(Y_1)$ at $Y_1 = Y_0/2$ (Fig. 1a). If a is below a certain critical value a_C, max $W(Y_1) < 0$; then, the permissible (allowed) region of Y_1 lies between the turning points $Y_1 = 0$ and $Y_1 = Y_0$, at which $W(Y_1) = 0$. This means that the intensity Y_1 in waveguide 1 varies periodically between Y_0 and zero. If $a = a_C$, then the peak of $W(Y_1)$ is at the point $Y_1 = Y_0/2$, for which $W(Y_0/2) = 0$. Consequently, when $a \ge a_C$, a new turning point arises in the region $Y_0/2 < Y_1 < Y_0 (< Y_C)$ and the periodically varying intensity Y_1 is above zero at any z. The critical (or bifurcation) value a_C is determined by Y_0 as follows:

$$a_C = \frac{Y_0^2}{4(1 - \exp(-Y_0/2))^4}. \qquad (21)$$

In Fig. 2, equation (21) is represented by the concave curve that separates region B from the rest of the plane. The bifurcation curve $a_C(Y_0)$ has a valley at the point $Y_{0\min} \approx 2.514$, which is found from the equation $\exp(Y_0/2) = 1 + Y_0$. This indicates that switching occurs at progressively decreasing values of a as Y_0 increases within the range $0 < Y_0 < Y_{0\min}$. Since switching is a nonlinear phenomenon, a decrease in the nonlinearity parameter a must be compensated for by an increase in Y_0. Qualitatively, this agrees with the behavior of switching intensity in a system of two coupled identical waveguides with Kerr-type nonlinearity [1]. Under saturation, such behavior occurs only for $a_C < a_{C\min}$. Also note that switching is impossible at any Y_0 if $a < a_{C\min} = Y_{0\,\min}^2(1 + Y_{0\min})^4/4 \approx 6.029$. Even more intriguing is that, if $Y_0 > Y_{0\min}$ in region B, switching occurs at progressively increasing values of a as Y_0 grows, which is in sharp contrast to the properties of Kerr-type media [1]. The point is that, if $Y_0 \ll 1$, the NLDC obviously tends to a linear regime. If $Y_0 \gg 1$, the effect of the exponential term in (1) becomes negligible, so that the NLDC goes to a linear regime in an unconventional fashion as Y_0 increases. The nonlinearity at high Y_0's can be enhanced by raising a.

Thus, only a fraction of power is transferred from waveguide 1 to waveguide 2 if the NLDC operates in region B (Fig. 2). At $Y_0 \ge Y_{0C}$, another bifurcation arises. When a is small and lies in region A (Fig. 2), the curve $W(Y_1)$ has two peaks symmetric about $Y_1 = Y_0/2$. Their height grows with a (Fig. 1b). If a takes on a new critical value, which can be found from the system of equations

$$\frac{Y_0^2}{4} = z^2 + 4a\exp(-Y_0)\left(\cosh\left(\frac{Y_0}{2}\right) - \cosh(z)\right)^2, \qquad (22)$$

$$1 = 4a\exp(-Y_0)\frac{\sinh(z)}{z}\left(\cosh\left(\frac{Y_0}{2}\right) - \cosh(z)\right) \quad (23)$$

by eliminating z at $Y_0 > Y_{0C}$, where Y_{0C} is determined by (19), $W(Y_1)$ vanishes at two points symmetric about $Y_1 = Y_0/2$ (curve *4* in Fig. 1b). A new turning point Y_1 thus appears in the potential well $W(Y_1)$, where $Y_0/2 < Y_1 < Y_0$. Look at the narrow region *C* in the (Y_0, a) plane (Fig. 2). It forms when the initial curve [formula (21)] splits at the point *D*. This region is analogous to region *B* in the sense that only a fraction of the power is transferred from waveguide 1 to waveguide 2. However, regions *B* and *C* are of differing origins. In gradually passing from region *A* to region *C*, the NLDC stepwise switches from the total to partial power transfer mode, so that less than half of the input power is transferred to waveguide 2. A similar effect is observed for the *A*-to-*B* transition except that precisely half of the power is transferred. Remarkably, while a representative point moves from region *C* to region *B*, the transferred fraction of power rises gradually, reaches $Y_0/2$ at the dividing line, and then falls.

As Y_0 increases at a fixed a, the point goes from *A* to *B* and then enters *A* directly or by crossing *C*. Accordingly, the NLDC goes from the total to partial power transfer mode and then returns to the former (Fig. 2). As a increases at a fixed Y_0, the NLDC goes from the total to partial power transfer mode. For comparison, Fig. 2 includes the bifurcation curve for the Kerr-type nonlinearity $\beta = \beta_0 + \alpha J$ controlled by a (see the dashed curve).

The properties presented above also follow from phase trajectories in the $(Y_1, dY_1/dz)$ plane for variable a and Y_0.

INTENSITY DISTRIBUTION ALONG WAVEGUIDE 1: COUPLING LENGTH

Figure 3 shows the minimum intensity Y_m in waveguide 1 as a function of Y_0 at various a's. The graphs were computed from (18). If $a < a_{C\min}$, $Y_m = 0$ for any Y_0. If $a_{C\min} < a < a_{0C}$, Y_m equals zero while Y_0 is sufficiently small. At a certain Y_0 such that the representative point lies on the bifurcation curve between regions *A* and *B* (Fig. 2), Y_m jumps to $Y_0/2$ and then grows with Y_0. When the right-hand segment of the bifurcation curve is reached, Y_m equals $Y_0/2$ again. As soon as the point enters the region *A*, Y_m drops to zero and stays at that level while Y_0 grows further. If $a > a_{0C}$, Y_m in regions *A* and *B* varies in the same fashion with increasing Y_0. In region *C*, the growth of Y_m continues. Note that $Y_m > Y_0/2$ at the boundary between *C* and *A*. Looking at Fig. 3, one can see a short falling segment on each of the curves for large Y_0. When the point enters region *A* (Fig. 2), Y_m again turns to zero. Thus, $Y_m = 0$ everywhere in region *A*: the power can be totally

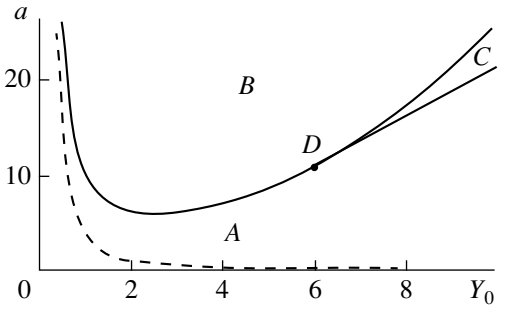


Fig. 2. Bifurcation curve $a(Y_0)$.

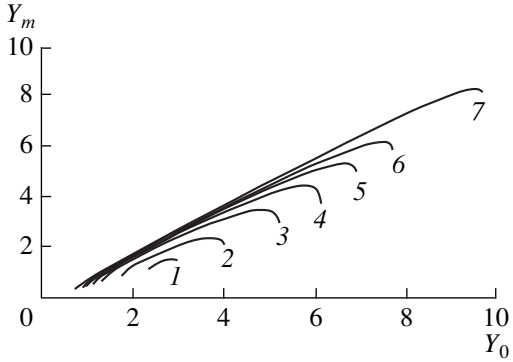


Fig. 3. Minimum intensity $Y_m(Y_0)$ for a = (*1*) 6.1, (*2*) 7, (*3*) 9, (*4*) 11, (*5*) 13, (*6*) 15, and (*7*) 20.

transferred to waveguide 2. In *B* and *C*, the power transfer is partial: $Y_0/2 < Y_m < Y_0$.

Now, consider the intensity distribution along waveguide 1, the intensity in waveguide 2 being determined by the relation $Y_2 = Y_0 - Y_1$ [see (9)]. Figure 4 shows distribution curves for different values of a and Y_0. It is seen that the power transfer to waveguide 2 may be total if the representative point lies in region *A* (Fig. 2). The coupling length L, i.e., the length along which Y_1 goes from Y_0 to zero, can be found from

$$L = \frac{1}{2\gamma}\int_0^{Y_0}[Y(Y_0 - Y) - a(1 - \exp(-Y))^2 \times (1 - \exp(Y - Y_0))^2]^{-1/2}dY. \quad (24)$$

As Y_0 increases, $L(Y_0)$ first rises to a maximum and then falls (Fig. 5a). In the linear-medium limit ($a = 0$), formula (24) gives $L = L_0 = \pi(2\gamma)$.

For regions *B* and *C* (Fig. 2), the distribution of Y_1 is much more complicated (Fig. 4b). When the point crosses the bifurcation curve $a(Y_0)$, the distribution changes radically. If Y_0 is so small that the representative point moves in region *A* (total transfer), L rises monotonically to infinity as the point approaches the bifurcation curve in Fig. 2, formula (24) remaining valid. In regions *B* and *C*, the function $Y_1(z)$ is also peri-

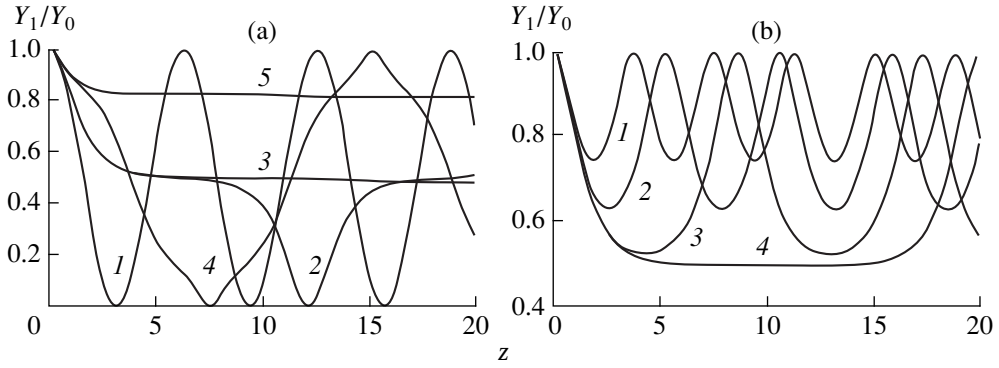


Fig. 4. Intensity distribution $Y_1/Y_0(z)$ along waveguide 1 at $a = 20$ and different values of Y_0. In panel (a), $Y_0 =$ (*1*) 0.05, (*2*) 0.59, (*3*) 0.5988, (*4*) 10, and (*5*) 9.55. In panel (b), $Y_0 =$ (*1*) 5.9, (*2*) 6.0, (*3*) 6.3, and (*4*) 7.3.

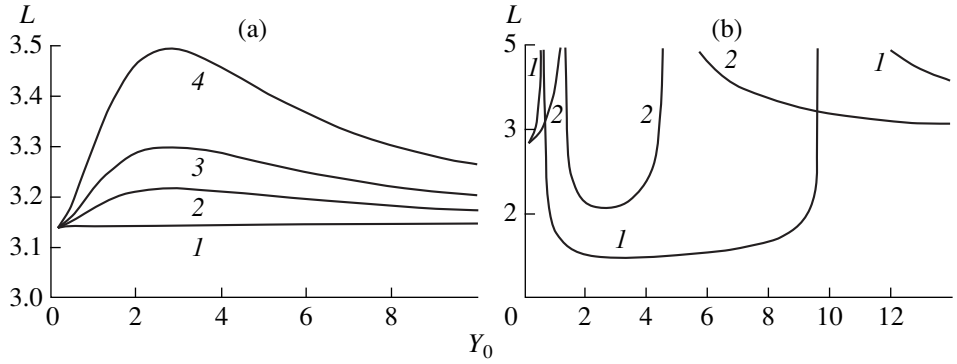


Fig. 5. Coupling length vs. Y_0 for waveguide 1 at different values of a. In panel (a), $a =$ (*1*) 0, (*2*) 0.5, (*3*) 1, and (*4*) 2. In panel (b), $a =$ (*1*) 8 and (*2*) 20.

odic but the transfer cannot be total (Fig. 4b) and the coupling length is expressed as

$$L = \frac{1}{2\gamma}\int_{Y_m}^{Y_0} [Y(Y_0 - Y) - a(1 - \exp(-Y))^2 \times (1 - \exp(Y - Y_0))^2]^{-1/2} dY, \quad (25)$$

where Y_m is determined from (18).

In B and C, L first rapidly goes down to a minimum with increasing Y_0 and then tends to infinity as the representative point approaches the right-hand segment of the bifurcation curve $a(Y_0)$. The higher a, the lower the minimum value of $L(Y_0)$. When the point enters region A, the conditions for total transfer are reestablished. The coupling length obeys (24) again and decreases monotonically as Y_0 goes up (Fig. 5b).

TRANSMISSION FUNCTION

Figure 6 depicts the transmission function $T_1 = Y_1/Y_0$ of waveguide 1 as a function of Y_0 at various a's. The NLDC length l is 5 and exceeds the coupling length of a linear coupler, L_0. It is seen that NLDCs with saturable nonlinearity offer a much wider variety of transmission functions than Kerr-type NLDCs [1]. If $a < a_{C\min}$, the curve $T_1(Y_0)$ has smoothly falling and smoothly rising segments for small Y_0 (Fig. 6a). The fall results from an increase in L with increasing Y_0. When Y_0 is such that $L = l$, the entire power is transferred to waveguide 2, so that $T_1 = 0$. It is important to note that, unlike Kerr-type NLDCs [1], the segment oscillating near a level of 1.0 for large Y_0 here is absent.

If $a > a_{C\min}$, the curve $T_1(Y_0)$ has two zeros and oscillates near $T_1 \approx 1$ in between (Fig. 6b). The number of oscillation periods grows with a. The curve includes two steep segments. They respectively correspond to the left- and right-hand segments of the bifurcation curve $a(Y_0)$ (see Fig. 2). At such values of Y_0 and a, one has $L = l$. The steepness of the two segments of $T_1(Y_0)$ (Fig. 6b) corresponds to that of $L(Y_0)$ (Fig. 5b). If $l < L_0$, $T_1(Y_0)$ has a single zero at larger Y_0's. Thus, the transmission function of a waveguide in a saturable NLDC is essentially governed by the normalized input intensity, nonlinearity parameter, and NLDC length.

Comparing the above results with the properties of Kerr-type NLDCs [1] suggests that self-switching is possible in saturable NLDCs as well. In contrast to the

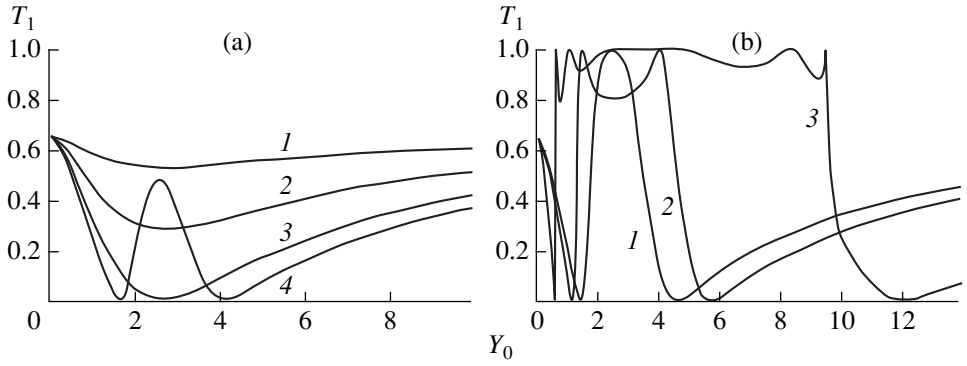


Fig. 6. Transmission function $T_1(Y_0)$ of waveguide 1 for $l = 5$ and different values of a. In panel (a), $a = (1)$ 1, (2) 2, (3) 5, and (4) 6. In panel (b), $a = (1)$ 6.5, (2) 8.0, and (3) 20.

no-saturation case [1], the transmission characteristic of a saturable device has two points of self-switching (Fig. 6b). Either lies in the middle of the respective steep and almost linear segment. If the input intensity is slightly higher or lower than the switching (critical) level, the input power is almost entirely applied to one or the other waveguide. Thus, a small variation in the input intensity around its critical value causes a sharp change in the output intensity. The effect may find application in all-optical switches, small-signal amplifiers, optical transistors, and other integrated-optics devices [1].

CONCLUSION

This study revealed high functionality of a simple symmetric saturable NLDC that is built around two optical waveguides and is fed at one of the inputs. Obviously, an asymmetric NLDC would exhibit more useful properties. Although NLDC performance is affected by optical absorption as well [5], our results seem to provide a fairly deep understanding of the subject. More comprehensive analysis will be reported at a later time.

REFERENCES

1. A. A. Maĭer, Usp. Fiz. Nauk **165**, 1037 (1995) [Phys. Usp. **38**, 991 (1995)].
2. A. A. Maĭer, Usp. Fiz. Nauk **166**, 1171 (1996) [Phys. Usp. **39**, 1109 (1996)].
3. S. M. Jensen, IEEE J. Quantum Electron. **18**, 1580 (1982).
4. F. Kh. Abdullaev and R. Gulyamov, Pis'ma Zh. Tekh. Fiz. **18** (20), 10 (1992) [Sov. Tech. Phys. Lett. **18**, 653 (1992)].
5. G. I. Stegeman, C. T. Seaton, C. N. Ironside, *et al.*, Appl. Phys. Lett. **50**, 1035 (1987).
6. G. I. Stegeman, E. L. Gaglioti, S. Trillo, *et al.*, Opt. Commun. **63** (5), 281 (1987).
7. E. L. Gaglioti, S. Trillo, S. Wabnitz, *et al.*, Appl. Phys. Lett. **51**, 293 (1987).
8. E. L. Gaglioti, S. Trillo, S. Wabnitz, *et al.*, J. Opt. Soc. Am. B **5**, 472 (1988).
9. Y. Chen, J. Opt. Soc. Am. B **8**, 986 (1991).
10. Y. Chen and A. W. Snyder, J. Appl. Phys. **67**, 1614 (1990).
11. J. D. Begin and M. Cada, IEEE J. Quantum Electron. **30**, 3006 (1994).
12. A. Guzman, M. Romagnoli, and S. Wabnitz, Appl. Phys. Lett. **56**, 614 (1990).
13. D. J. Mitchell, Y. Chen, and A. W. Snyder, Opt. Lett. **15**, 535 (1990).
14. S. A. Moskalenko, P. I. Khadzhi, and A. Kh. Rotaru, *Solitons and Nutation in the Exciton Range of the Spectrum* (Shtiintsa, Kishinev, 1980).
15. P. I. Khadzhi, *Nonlinear Optical Processes in Exciton and Biexciton Systems in Semiconductors* (Shtiintsa, Kishinev, 1985).
16. P. I. Khadzhi, G. D. Shibarshina, and A. Kh. Rotaru, *Optical Bistability in a System of Coherent Excitons and Biexcitons in Semiconductors* (Shtiintsa, Kishinev, 1988).

Translated by A. Sharshakov

Technical Physics, Vol. 45, No. 9, 2000, pp. 1170–1176. Translated from Zhurnal Tekhnicheskoĭ Fiziki, Vol. 70, No. 9, 2000, pp. 82–89.

OPTICS, QUANTUM ELECTRONICS

Reflection and Transmission of Light in Medium/Cholesteric/Substrate and Glass (1)/Cholesteric/Glass (2) Systems

A. A. Gevorgyan

Yerevan State University, ul. Manukyana 1, Yerevan, 375049 Armenia
E-mail: yndantiz@sun.ysu.am
Received February 2, 1999

Abstract—Reflection and transmission of light in medium/cholesteric/substrate and glass (1)/cholesteric/glass (2) systems are discussed. These systems can be viewed as a cholesteric-filled Fabry–Perot resonator. A modified layer-addition method is used. Rigorous explicit expressions for the elements of the Jones matrices for these systems are derived. Reflection, transmission, Faraday-rotation, and polarization-ellipticity spectra were studied. The effect of quasi-monochromatism and degree of polarization of light on the reflection properties was treated in terms of Müller matrices. The influence of the dielectric boundaries on intrinsic cholesteric polarizations was investigated. It is shown that initially quasi-circular intrinsic polarizations turn to quasi-linear as the difference between the cholesteric mean permittivity and the permittivity of the surrounding medium increases. Features of the Borrmann effect in the presence of dielectric boundaries are discussed. © *2000 MAIK "Nauka/Interperiodica".*

INTRODUCTION

The optical properties of structures based on liquid crystals, in particular, cholesteric liquid crystals, are attracting much attention. Most often used in practice are medium/cholesteric/substrate (MCS) and glass (1)/cholesteric/glass (2) (GCG) systems. In the former, cholesteric is applied on an isotropic substrate, while, in the latter, it is sandwiched in isotropic plates. This work is concerned with the optical properties of these systems under normal incidence of light (the cholesteric axis is normal to the plate surfaces).

The properties of cholesterics can be simulated by artificial helical media with desired parameters and also ferromagnetic helical media, which simulate their properties at high frequencies. This greatly extends investigations and applications of the systems. Since a GCG system can be considered as a cholesteric-filled Fabry–Perot resonator, the derivation of explicit analytical expressions for Jones matrix elements seems to be topical. In cholesteric-filled Fabry–Perot resonators, the number of control parameters grows; in addition, they perform not only frequency but also polarization selection [1, 2].

The propagation of plane monochromatic waves in homogeneous helical media has been the subject of investigation in many reports (see, e.g., [3–12]). The boundary-value problem of light transmission through a layer of a finite thickness was solved both analytically and numerically. However, the obtained solutions were approximate. In [13], this problem was rigorously solved for light transmission through a cholesteric layer. This allows cholesteric-based multilayer structures to be treated analytically. Yet, it is known that a problem of interaction of electromagnetic waves with multilayer structures with three or more interfaces cannot adequately be solved by immediately applying the boundary conditions method. An example is works [1, 2], where the optical properties of a GCG structure were numerically analyzed by immediately applying boundary conditions. Thirty-two linear equations with 32 unknowns were eventually derived. Various layer-addition methods are available in the literature (see [7, 14–18], and references therein). Here, an elegant and efficient method developed in [14, 16, 17] will be used. In [17], it was applied in solving the problems of light reflection and transmission in cholesteric/glass/cholesteric and cholesteric(1)/cholesteric(2) systems. In this work, the application of this method to solve the same problems in MCS and GCG systems is discussed.

REFLECTION AND TRANSMISSION OF LIGHT IN MCS SYSTEM

The situation where polarized light is reflected from or transmitted through a substrate covered by a homogeneous cholesteric layer is of considerable interest for ellipsometry. The substrate is assumed to be of a finite thickness. Outer media (0 and 3) and substrate (2) are homogeneous, optically isotropic, and have complex refractive indexes n_0, $n_3 = n_0$, and n_2, respectively. The cholesteric layer will be joint together with an isotropic plate. The solution method also leaves room for a substrate to be an anisotropic (or gyrotropic) plate. In this case, the Jones matrices for an isotropic substrate are

1063-7842/00/4509-1170$20.00

replaced by those for an anisotropic (or gyrotropic) plate.

Let a wave with an amplitude $\mathbf{E}_i$ be incident on the system. Designating the reflected and transmitted field amplitudes as $\mathbf{E}_r$ and $\mathbf{E}_t$ and expanding them in basis circular polarizations

$$\mathbf{E}_{i,r,t} = E^+_{i,r,t}\mathbf{n}_+ + E^-_{i,r,t}\mathbf{n}_- = \begin{bmatrix} E^+_{i,r,t} \\ E^-_{i,r,t} \end{bmatrix}$$

($\mathbf{n}_+$ and $\mathbf{n}_-$ are the unit vectors of circular polarizations), we will represent the solution in the form

$$\mathbf{E}_r = \hat{R}\mathbf{E}_i, \quad \mathbf{E}_t = \hat{T}\mathbf{E}_i, \tag{1}$$

where $\hat{R}$ and $\hat{T}$ are the Jones matrices for the system.

If there exists a system consisting of two layers A and B applied to each other "from left to right," then, according to [17], the Jones matrices of the system $A + B$ $\hat{R}_{A+B}$ and $\hat{T}_{A+B}$ are represented through the matrices of the component layers with the matrix equations [17]

$$\begin{aligned} \hat{R}_{A+B} &= \hat{R}_A + \tilde{\hat{T}}_A\hat{R}_B[\hat{I} - \tilde{\hat{R}}_A\hat{R}_B]^{-1}\hat{T}_A, \\ \hat{T}_{A+B} &= \hat{T}_B[\hat{I} - \tilde{\hat{R}}_A\hat{R}_B]^{-1}\hat{T}_A, \end{aligned} \tag{2}$$

where $\hat{I}$ is the unity matrix. The Jones matrices marked with a tilde refer to backward light propagation.

For example, if a cholesteric layer is between the same media, the Jones matrices for light incident "from the right" and "from the left" are related as

$$\hat{\tilde{T}} = \hat{F}^{-1}\hat{T}\hat{F}, \quad \hat{\tilde{R}} = \hat{F}^{-1}\hat{R}\hat{F}, \tag{3}$$

where

$$\hat{F} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$

for circular basis polarizations.

Let A be a cholesteric layer and B be a substrate. Then, using the Jones matrices for a glass plate [18] and a cholesteric [13], we obtain from (1)–(3) the expressions for the elements of the Jones matrices for the cholesteric/substrate system:

$$\begin{aligned} T^a_{ii} &= t_2[T^c_{ii} + r_2(T^c_{ji}R^c_{ji} - T^c_{ii}R^c_{ii})]/\Delta_1, \\ T^a_{ij} &= t_2[T^c_{ij} + r_2(T^c_{jj}R^c_{ji} - T^c_{ij}R^c_{ii})]/\Delta_1, \\ R^a_{ii} &= R^c_{ii} + r_2[(1 - r_2R^c_{ii})(T^c_{ii}T^c_{jj} + T^c_{ji}T^c_{ji}) \\ &\quad + r_2T^c_{ji}(T^c_{22}R^c_{21} + T^c_{11}R^c_{12})]/\Delta_1, \\ R^a_{ij} &= R^c_{ij} + r_2[T^c_{jj}(T^c_{ij} + T^c_{ji})(1 - r_2R^c_{ii}) \\ &\quad + r_2(T^c_{jj}T^c_{jj}R^c_{ji} + T^c_{ji}T^c_{ij}R^c_{ij})]/\Delta_1, \end{aligned} \tag{4}$$

$$i, j = 1, 2, \quad i \neq j,$$

where

$$\Delta_1 = (1 - r_2R^c_{11})^2 - r_2^2R^c_{12}R^c_{21}, \quad R^c_{11} = R^c_{22} = H,$$

$$R^c_{12} = Q + iF, \quad R^c_{21} = Q - iF,$$

$$T^c_{11} = (S - iN)\exp(i\Theta), \quad T^c_{12} = V\exp(i\Theta),$$

$$T^c_{21} = V\exp(-i\Theta), \quad T^c_{22} = (S + iN)\exp(-i\Theta),$$

$$H = \{\chi^2 f_2 f_1(c_1c_2 - 1) + 2u^2[f_2f_1(2\chi^2m_1 - \gamma^2) - \alpha^2\delta^2\gamma^2]s_1s_2 - iu\sqrt{\alpha}\gamma(p_1s_1c_2 + p_2s_2c_1)\}/\Delta,$$

$$F = \delta\chi\sqrt{\alpha}\{-2u\gamma\sqrt{\alpha}(s_1c_2 - s_2c_1) - i[f_2(c_1c_2 - 1) + 4u^2(m_1f_2 + \alpha\gamma^2)]s_1s_2\}/\Delta,$$

$$Q = u\gamma\delta\sqrt{\alpha}[4u\sqrt{\alpha}\gamma s_1s_2 + i(g_1s_2c_1 - g_2s_1c_2)]/\Delta,$$

$$S = \gamma\sqrt{\alpha}[\gamma\sqrt{\alpha}(c_1 + c_2) - iu(b_1s_1 + b_2s_2)]/\Delta,$$

$$V = \gamma\sqrt{\alpha}\delta[\sqrt{\alpha}(c_2 - c_1) - iu(q_2s_2 - q_1s_1)]/\Delta,$$

$$N = \gamma\sqrt{\alpha}\chi[if_1(c_2 - c_1) - 2u\sqrt{\alpha}(l_1s_1 + l_2s_2)]/\Delta,$$

$$\Delta = -\chi^2f_2^2 + (\chi^2f_2^2 + 2\alpha\gamma^2)c_1c_2 + 2u^2[\alpha^2\delta^2\gamma^2 - f_1^2(2\chi^2m_2 + \delta^2) + 4\alpha\chi^2(\delta^2 - 2m_2)] \times s_1s_2 - 2ui\gamma\sqrt{\alpha}(b_1s_1c_2 + b_2s_2c_1),$$

$$b_{1,2} = f_1w_{1,2} \pm \alpha\delta^2, \quad p_{1,2} = f_2w_{1,2} \mp \alpha\delta^2, \quad q_{1,2} = f_1 \pm \alpha\gamma,$$

$$g_{1,2} = f_2 \pm \alpha\gamma, \quad w_{1,2} = \gamma \pm 2\chi^2, \quad l_{1,2} = \gamma \pm 2, \quad f_{1,2} = 1 \pm \alpha,$$

$$s_{1,2} = \sin(k_{1,2}d)/(k_{1,2}d), \quad c_{1,2} = \cos(k_{1,2}d),$$

$$k_{1,2} = 2u\sqrt{m_1 \pm \gamma}/d, \quad m_{1,2} = 1 \pm \chi^2,$$

$$\alpha = \varepsilon_m/\varepsilon_v, \quad \varepsilon_v = n_0^2, \quad \chi = \lambda/(\sigma\sqrt{\varepsilon_m}),$$

$$\varepsilon_m = (\varepsilon_1 + \varepsilon_2)/2, \quad \varepsilon_a = (\varepsilon_1 - \varepsilon_2)/2, \quad \delta = \varepsilon_a/\varepsilon_m,$$

$$\Theta = 2\pi d/\sigma, \quad u = \pi d\sqrt{\varepsilon_m}/\lambda, \quad \gamma = \sqrt{\delta^2 + 4\chi^2},$$

$$r_2 = \tau_2[1 - \exp(-i2\beta_2)]/[1 - \tau_2^2\exp(-i2\beta_2)],$$

$$t_2 = 4n_2n_0\exp(-i\beta_2)/\{[1 + \tau_2^2\exp(-i2\beta_2)](n_2 + n_0)^2\},$$

$$\tau_2 = (n_0 - n_2)/(n_0 + n_2),$$

$\beta_2 = 2\pi d_2n_2/\lambda$, λ is the wavelength in a vacuum, σ is the helix pitch, ε_1 and ε_2 are the principal values of the cholesteric permittivity tensor in the plane normal to the

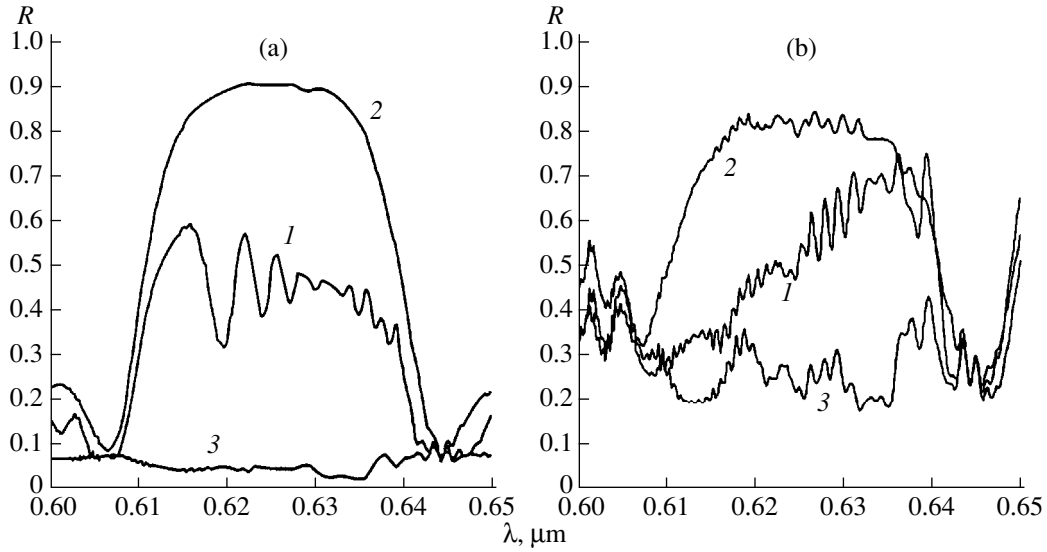


Fig. 1. Reflection coefficient R vs. wavelength for light incidence on the glass (1)/cholesteric/glass (2) system. (a) $n_1 = 2.5$ and $n_2 = 1.5$; (b) $n_1 = 1.5$ and $n_2 = 2.5$. Cholesteric layer parameters are $\varepsilon_1 = 2.29$, $\varepsilon_2 = 2.143$, $\sigma = 0.42$ μm, and $d = 20\sigma$. Other parameters: $d_1 = d_2 = 1000$ μm and $n_0 = 1$.

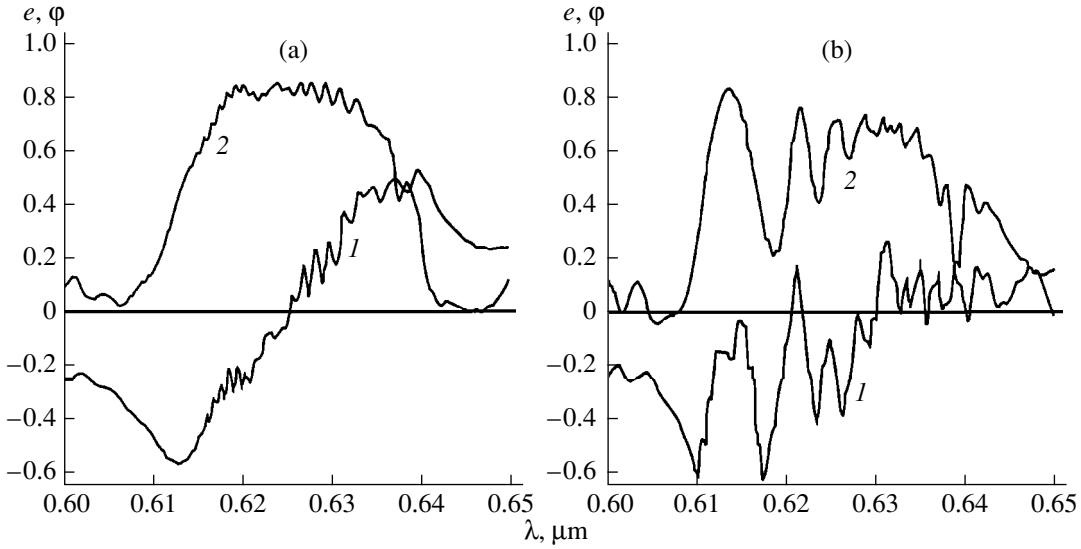


Fig. 2. Wavelength dependence of (*1*) Faraday rotation φ and (*2*) polarization ellipticity e for incidence of linearly polarized light on the glass (1)/cholesteric/glass (2) system. (a) $n_1 = 2.5$ and $n_2 = 1.5$; (b) $n_1 = 1.5$ and $n_2 = 2.5$. Other parameters as in Fig. 1.

axis of the medium, d is the cholesteric thickness, and d_2 is the substrate thickness.

When calculating the transmission coefficient T (from the intensity values), one should bear in mind that, when the refractive indexes n_3 and n_0 diverge, $T = n_3|E_t|^2/(n_0|E_i|^2)$ [18].

Consider the important specific case $\alpha = 1$; that is, the permittivity of medium 0 equals the mean permittivity of the cholesteric. Then, the explicit expressions for the matrices $\hat{R}$ and $\hat{T}$ are greatly simplified:

$$T^a_{ii} = t_2[h_2a_i + h_1a_j + iur_2\delta(s_2 - s_1)]/(\Delta\Delta_1),$$

$$T^a_{ij} = t_2\delta[a_2 - a_1 + iur_2(h_1s_i + h_2s_j)]/(\Delta\Delta_1),$$

$$R^a_{ii} = \{iu\delta^2(a_2s_1 - a_1s_2) + 2r_2[\delta^2(a_2 - a_1)^2 + \gamma\Delta + iur_2\delta^2\gamma(a_2s_2 - a_1s_1)]/(\Delta\Delta_1)\}/\Delta,$$

$$R^a_{ij} = \{iu\delta(h_1a_is_j + h_2a_js_i) + 2r_2\delta[(a_2 - a_1)(h_2a_j + h_1a_i) + iur_2\gamma(h_ja_2s_2 + h_ia_1s_1)]/(\Delta\Delta_1)\}\Delta, \quad (5)$$

$$i, j = 1, 2; \quad i \neq j,$$

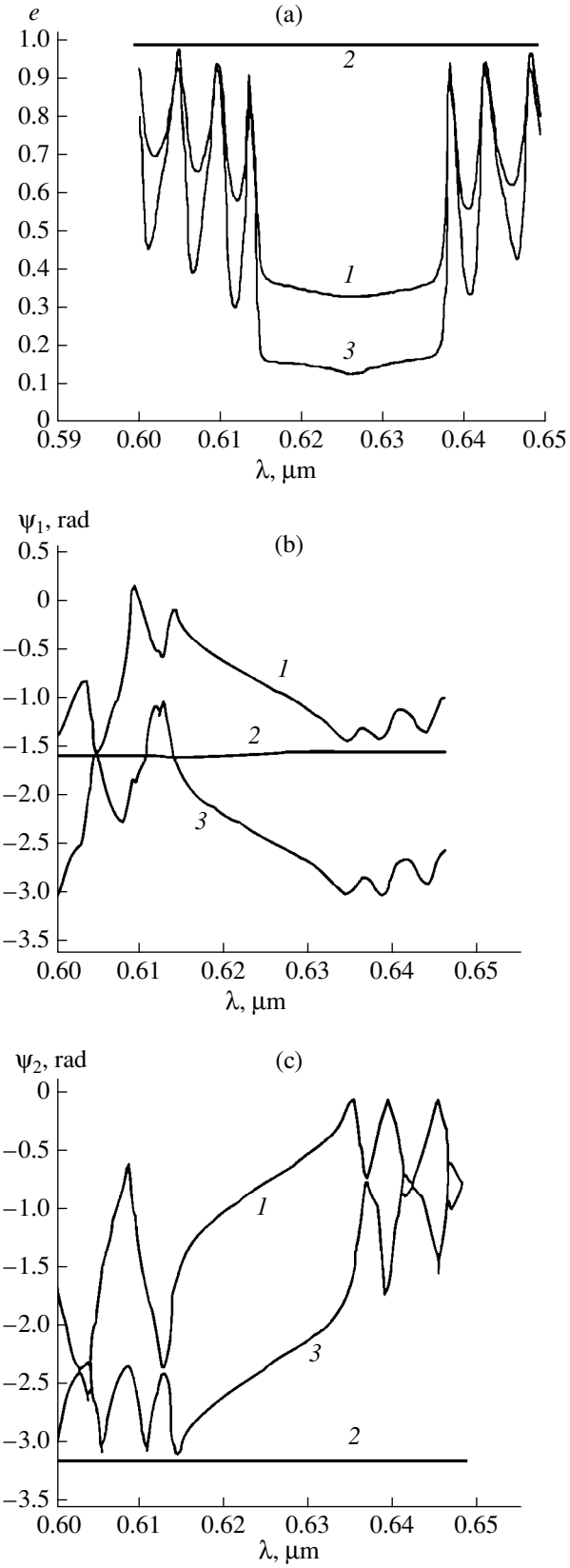


Fig. 3. Wavelength dependence of (a) IP ellipticity e and azimuths (b) Ψ_1 and (c) Ψ_2 at α = (*1*) 8.87, (*2*) 1, and (*3*) 0.0222. $d = 50\sigma$. Other parameters as in Fig. 1.

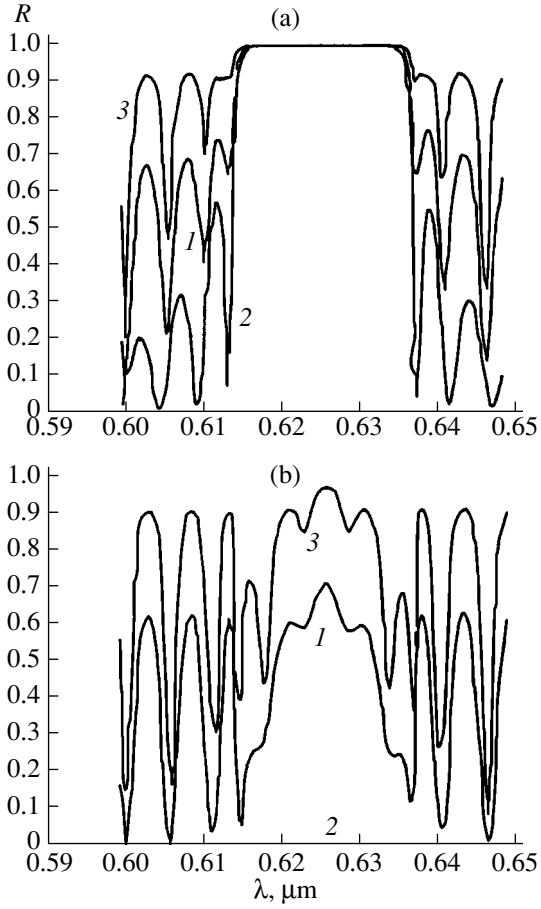


Fig. 4. Reflection coefficient R vs. wavelength. Curves *1–3* refer to the values of α in Figs. 3a, 3b. Other parameters as in Fig. 1. (a) First IP and (b) second IP.

where

$$\Delta_1 = 1 - r_2 u\delta[2r_2 u s_2 s_1 - i(s_1 a_2 - s_2 a_1)/(4\gamma)]/(a_1 a_2),$$

$$\Delta = 2\gamma a_1 a_2, \quad a_{1,2} = c_{1,2} = uil_{1,2}s_{1,2},$$

$$h_{1,2} = \gamma \pm 2\chi.$$

REFLECTION AND TRANSMISSION OF LIGHT IN GCG SYSTEM

Consider normal incidence of light on a glass (1)/cholesteric/glass (2) system. The glasses can obviously be dissimilar with different refractive indexes. The problem is to mathematically join an isotropic glass plate to the left side of an MCS system. Let us designate the first glass (the plate to be joined) by A and the MCS system (where the substrate has the properties of the second glass) by B. Then, from (1)–(3), we obtain for the glass (1)–cholesteric–glass (2) system

$$T_{ii} = t_1[T^a_{ii} + r_1(T^a_{ij}R^a_{ji} - T^a_{ii}R^a_{ii})]/\Delta_2,$$

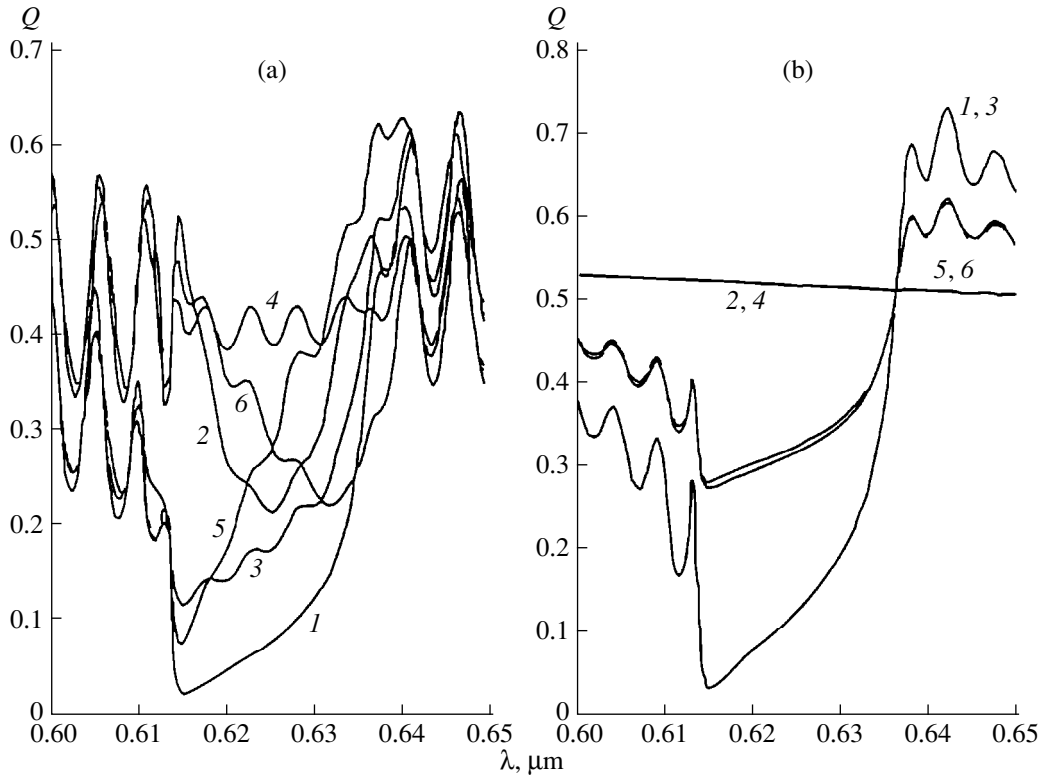


Fig. 5. Absorption vs. wavelength at α = (a) 0.1128 and (b) 1. Polarizations of light incident on the cholesteric layer are (*1*) diffraction, (*2*) nondiffraction, (*3*) counterclockwise circular, (*4*) clockwise circular, (*5*) linear along the *x*-axis, and (*6*) linear along the *y*-axis. $\mathrm{Im}\varepsilon_1 = 0.01$ and $\mathrm{Im}\varepsilon_2 = 0$. Other parameters as in Fig. 1.

$$T_{ij} = t_1[T^a_{ij} + r_1(T^a_{ii}R^a_{ij} - T^a_{ij}R^a_{ii})]/\Delta_2,$$
$$R_{ii} = r_1 + t_1^2[R^a_{11} + r_1(R^a_{12}R^a_{21} - R^a_{11}R^a_{11})]/\Delta_2, \quad (6)$$
$$R_{ij} = t_1^2 R^a_{ij}/\Delta_2,$$

where

$$\Delta_2 = (1 - r_1 R^a_{11})^2 - r_1^2 R^a_{12}R^a_{21},$$

$$r_1 = \tau_1[1 - \exp(-i2\beta_1)]/[1 - \tau_1^2\exp(-i2\beta_1)],$$

$$t_1 = 4n_1n_0\exp(-i\beta_1)/\{[1 + \tau_1^2\exp(-i2\beta_1)](n_1 + n_0)^2\},$$

$$\tau_1 = (n_0 - n_1)/(n_0 + n_1), \quad \beta_1 = 2\pi d_1 n_1/\lambda,$$

and d_1 is the thickness of glass 1.

The reflection, transmission, and polarization characteristics of glass/cholesteric/glass systems were given a detailed numerical analysis in [1, 2], where the isotropic glasses were assumed to be identical. Of interest is also the case when they are dissimilar. Figure 1 shows the wavelength dependence of the refractive index for incident light with linear polarization (curve *1*), as well as diffraction (*2*) and counterclockwise circular polarizations (*3*). The wavelength dependences of the Faraday rotation γ and polarization ellipticity *e* for linearly polarized light are depicted in Fig. 2 (curves *1* and *2*, respectively).

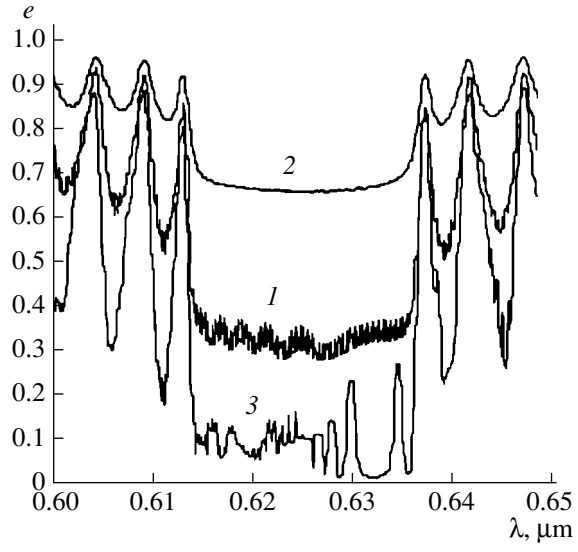


Fig. 6. IP ellipticity *e* vs. wavelength for n_1 = (*1*) 0.5, (*2*) 1.5, and (*3*) 10; $n_2 = n_1$, $d_1 = d_2 = 1000$ μm, and $n_0 = 1$. Other parameters as in Fig. 1.

INTRINSIC POLARIZATIONS

As is known, intrinsic polarizations (IPs) are two polarization remaining unchanged when light passes through a system. Intrinsic polarizations and their values are convenient for the classification of optical devices [15]. Therefore, the study of IPs in these systems is of certain interest. The IPs of cholesterics were considered, e.g., in [19], where the associated values were calculated using exact boundary conditions. It was found that, if the anisotropy is weak ($\delta \ll 1$) and absorption is absent, IPs are two orthogonal circular polarizations. Light with one of them undergoes diffraction reflection, while light with the other does not. If a cholesteric is highly anisotropic ($\delta \gg 1$), IPs are orthogonal quasi-linear polarizations. In what follows, we will first describe the effect dielectric boundaries have on the IPs of the cholesteric and then discuss the reflection of optical waves having polarizations coincident with the IPs. Subsequently, IPs of the glass (1)/cholesteric/glass (2) system will be considered.

In the presence of dielectric boundaries ($\alpha \neq 1$), IPs cease to be orthogonal, as follows from calculations. The same is also observed in the presence of absorption. It is known [20] that the IPs of normal gyrotropic media are also nonorthogonal. For $\alpha \neq 1$, the IPs are nonorthogonal but their ellipticities differ only by sign. At $\alpha \approx 1$, the IPs of weakly anisotropic cholesterics are two, clock- and counterclockwise, quasi-circular polarizations. As α increases (decreases), the IP ellipticities decrease (in magnitude) and tend to zero in the limits $\alpha \gg 1$ and $\alpha \ll 1$. Thus, for $\alpha \gg 1$ and $\alpha \ll 1$, IPs are quasi-linear polarizations, as in the case of high anisotropy.

Figure 3 shows the wavelength dependences of the ellipticity (the ellipticity of the other polarization differs only by sign). Also given are the wavelength dependences of the IP azimuths at different α's. The high dispersion of the IP ellipticity at $\alpha \neq 1$ is noteworthy. It is due to multiple reflections from the dielectric boundaries and increases as α deviates more and more from unity. The wavelength dependence of the reflection coefficient for the same range of α as in Fig. 3 and both IPs is shown in Fig. 4.

Since the IP changes with α, so does the nature of absorption. For example, at $\alpha \approx 1$, absorption is totally suppressed (Borrmann effect) when incident light has quasi-circular polarization. At $\alpha \gg 1$ or $\alpha \ll 1$, the Borrmann effect is observed at quasi-linear polarization. This is explained by the fact that absorption is completely suppressed when light incident on the cholesteric layer has diffraction polarization rather than circular or linear polarization. Figure 5 plots the quantity $Q = 1 - (R + T)$, which characterizes absorbed optical energy, against wavelength for various incident polarizations at $\alpha = 0.1128$ and 1. Under inclined incidence of light on a highly anisotropic cholesteric layer, absorption features were similar [21]. The observed large-amplitude oscillations of Q are associated with multiple reflections from the dielectric boundaries. These oscillations demonstrate the interference character of optical absorption in the medium. Unlike the case when the dielectric boundaries are absent, an oscillatory wavelength dependence of Q was also observed for light with nondiffraction IP.

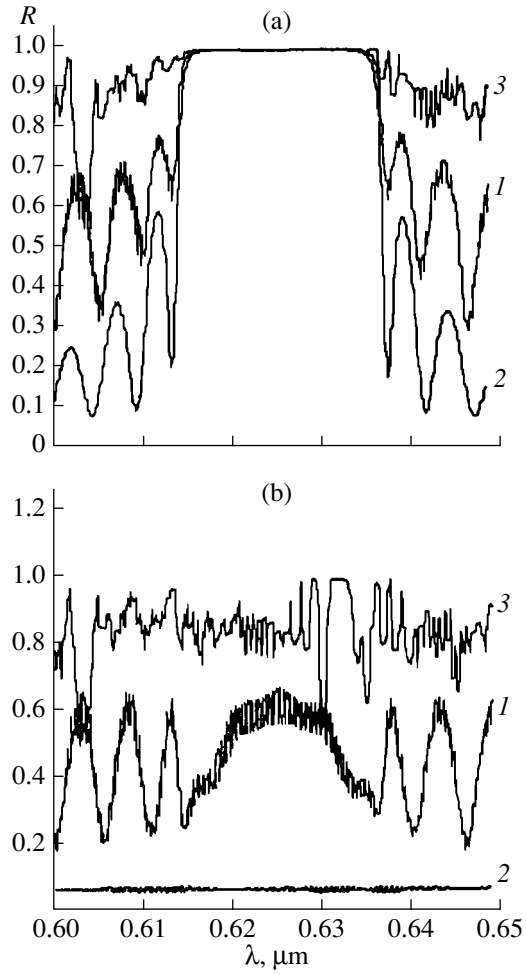


Fig. 7. R vs. wavelength. The refractive indexes and other glass parameters as in Fig. 6. (a) First and (b) second IP.

Now turn to IPs of the glass (1)/cholesteric/glass (2) system. Figure 6 plots the IP ellipticity (the ellipticity of the other IP differs only in sign) against wavelength for glasses with various refractive indexes. The wavelength dependence of the reflection coefficient for the same refractive indexes as in Fig. 6 and both IPs is depicted in Fig. 7.

MÜLLER MATRICES

The interaction of quasi-monochromatic partially polarized light with optical systems is usually described by Müller matrices. In this case, the solution of the problem is represented as

$$\mathbf{S}_t = \hat{M}_t \mathbf{S}_i, \quad \mathbf{S}_r = \hat{M}_r \mathbf{S}_i, \tag{7}$$

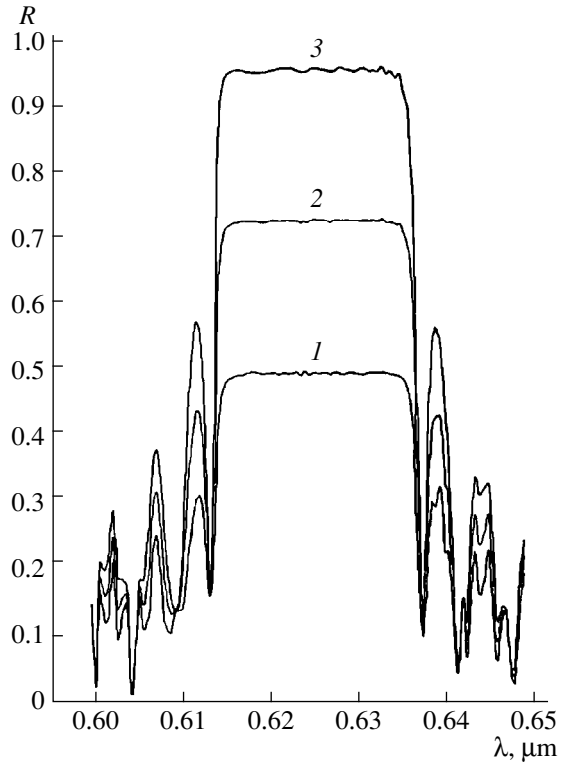


Fig. 8. R vs. wavelength at the degree of polarization of incident wave P = (*1*) 0, (*2*) 0.5, and (*3*) 1. Other parameters as in Fig. 6.

where $\mathbf{S}_i$, $\mathbf{S}_t$, and $\mathbf{S}_r$ are 4×1 Stokes column vectors for incident, reflected, and transmitted waves, respectively; $\mathbf{S}_i = I\{1,\ P\cos(2\mathbf{\Phi}_i)\cos(2\mathbf{\Psi}_i),\ P\cos(2\mathbf{\Phi}_i)\sin(2\mathbf{\Psi}_i),\ P\sin(2\mathbf{\Phi}_i)\}$; I is the total intensity of the incidence wave; $\mathbf{\Psi}_i$ is azimuth; $\mathbf{\Phi}_i$ is the ellipticity angle of the polarization ellipse for a completely polarized component in the incident wave; P is the degree of polarization of the incident wave; and $\hat{M}_t$ and $\hat{M}_r$ are 4×4 Müller matrices for the transmitted and reflected waves, respectively.

Müller matrices can conventionally be obtained from Jones matrices [15]. Explicit expressions for the Müller matrix elements are very tedious and are omitted here.

The wavelength dependence of the reflection coefficient R for various degree of polarization P of the incident wave is given in Fig. 8. Light passes through the glass (1)/cholesteric/glass (2) system. The completely polarized component in the incident wave has counterclockwise circular polarization.

REFERENCES

1. S. M. Arakelyan, O. S. Eritsyan, A. S. Karayan, and Yu. S. Chilingaryan, Kvantovaya Élektron. (Moscow) **7**, 959 (1980).
2. S. M. Arakelyan, O. S. Eritsyan, A. S. Karayan, and Yu. S. Chilingaryan, Opt. Spektrosk. **50**, 550 (1981).
3. Hl. De Vries, Acta Crystallogr. **4**, 219 (1951).
4. E. I. Kats, Zh. Éksp. Teor. Fiz. **59**, 1854 (1970) [Sov. Phys. JETP **32**, 1004 (1970)].
5. V. N. Belyĭ and A. N. Serdyukov, Dokl. Akad. Nauk BSSR **18**, 402 (1974).
6. V. A. Belyakov, *Diffraction Optics of Complex Periodic Media* (Nauka, Moscow, 1988).
7. S. Chandrasekhar, *Liquid Crystals* (Cambridge Univ. Press, Cambridge, 1977; Mir, Moscow, 1980).
8. V. N. Kapshaĭ, V. A. Kienya, and I. V. Semchenko, Kristallografiya **36**, 822 (1991).
9. W. S. Weiglhofer and A. Lakhtakia, Microwave Opt. Technol. Lett. **6**, 804 (1993).
10. W. S. Weiglhofer and A. Lakhtakia, Optik (Stuttgart) **96** (4), 179 (1994).
11. R. H. Good, Jr. and A. Karali, J. Opt. Soc. Am. A **11**, 2145 (1994).
12. W. S. Weiglhofer and A. Lakhtakia, Int. J. Appl. Electromagn. Mech. **6**, 221 (1995).
13. G. A. Vardanyan and A. A. Gevorgyan, Kristallografiya **42**, 710 (1997) [Crystallogr. Rep. **42**, 663 (1997)].
14. O. V. Pikichyan, Soobshch. Byurak. Obs., Akad. Nauk Arm. SSR, No. LV, 5 (1984).
15. R. M. Azzam and N. M. Bashara, *Ellipsometry and Polarized Light* (North-Holland, Amsterdam, 1977; Mir, Moscow, 1981).
16. V. A. Ambartsumyan, Izv. Akad. Nauk Arm. SSR, Estestv. Nauki, Nos. 1–2, 31 (1944); *Scientific Works*, Ed. by V. V. Sobolev (Akad. Nauk Arm. SSR, Yerevan, 1960), Vol. 1, pp. 263–268.
17. A. A. Gevorgyan, K. V. Papoyan, and O. V. Pikichyan, Opt. Spektrosk. **88** (4) (2000) (in press) [Opt. Spectrosc. **88**, 586 (2000)].
18. M. Born and E. Wolf, *Principles of Optics* (Pergamon, Oxford, 1969; Nauka, Moscow, 1970).
19. G. A. Vardanyan, A. A. Gevorgyan, and O. S. Eritsyan, Opt. Spektrosk. **85**, 640 (1998) [Opt. Spectrosc. **85**, 585 (1998)].
20. A. I. Okorochkov and A. F. Konstantinova, Kristallografiya **30**, 105 (1985) [Sov. Phys. Crystallogr. **30**, 57 (1985)].
21. Sah. Yuvaraj and K. A. Suresh, J. Opt. Soc. Am. A **11**, 740 (1994).

Translated by V. Isaakyan

Technical Physics, Vol. 45, No. 9, 2000, pp. 1177–1181. Translated from Zhurnal Tekhnicheskoĭ Fiziki, Vol. 70, No. 9, 2000, pp. 90–95.

OPTICS, QUANTUM ELECTRONICS

Estimation of the Petroleum Product Knock Rating by Regression Analysis of Near-Infrared Absorption Spectra

V. N. Korolev, A. V. Marugin, and V. B. Tsaregradskiĭ

Lobachevsky State University, pr. Gagarina 23, Nizhni Novgorod, 630600 Russia

E-mail: quant@rf-nngu.nnov.su

Received June 30, 1999

Abstract—Spectroscopic measurements in the near-infrared region are suggested to determine the octane number of petroleum products. Statistical regression analysis of the absorption spectra of hydrocarbons is used for calculating the gasoline octane number and several other physicotechnical parameters of fuel. The knock rating of commercial gasoline was determined with a specially designed analyzer. Its working parameters and limiting capabilities in determining the octane number are discussed. *© 2000 MAIK "Nauka/Interperiodica".*

INTRODUCTION AND STATEMENT OF THE PROBLEM

The development of analytical spectroscopic methods for studying petrochemical compounds is nowadays a topical problem. Of special importance are near-infrared spectroscopy methods for optical control of the knock rating of organic fuel [1, 2]. This property (i.e., the stability of fuel against spontaneous combustion) is conventionally characterized by octane number. Under laboratory conditions, the octane number is usually determined with special motor setups, where the knock rating of tested fuel (gasoline) is compared with that of calibrating mixtures of isooctane and heptane. The octane number is, by definition, the percentage (volume fraction) of isooctane in the calibrating mixture where knock is initiated at the same compression ratio as in the gasoline. The method has significant drawbacks, which are expensive and bulky setups, long-term test procedure, consumption of expensive chemicals, and the need for a special-purpose room. An alternative approach based on analytical spectroscopy holds much promise for developing express petroleum product testing, which will lead to fundamentally new opportunities for knock rating control [2, 3].

A physicochemical parameter of gasoline (octane number, in our case) can be calculated from its spectral characteristics by relating the parameter value to the absorption factors at a number of wavelengths [1, 4]. This optical method is one of the promising techniques recently developed for the investigation of petroleum products. However, its experimental implementation requires a number of problems to be solved. They are the determination of the applicability of the method to petroleum products with various compositions, search for optimal spectral ranges, the design of optical devices suited to gasoline analysis, etc.

Applying spectral analysis to gasoline, one should take into account that it may contain up to 100 different chemical components (hydrocarbons), each influencing the knock rating of the fuel. Therefore, traditional analytical methods, which determine the physicochemical characteristics of a fuel (in particular, octane number) from its exact composition, fail in this case because of the complex composition of petroleum products and the necessity to determine the knock rating of all constituents with allowance for their interaction. Consequently, spectral methods that do not involve the intermediate determination of the exact composition of the fuel are required.

In this paper, we apply regression analysis to the absorption spectra of the fuel to determine its octane number and check the accuracy of the determined parameters depending on the measuring techniques. The prospects of such analytical systems will be briefly outlined.

ANALYSIS OF ABSORPTION SPECTRA BY LINEAR REGRESSION METHOD

We calculated the octane number of gasoline from the absorption spectra by means of linear regression analysis [5, 6], which, as was noted, makes it possible to relate the octane number of gasoline to its spectral characteristics without the intermediate determination of its composition. In the simplest form, this relation can be represented as a linear combination of the gasoline absorption factors at various wavelengths. Physically, this relation stems from the fact that the absorption spectrum of gasoline is uniquely determined by its composition, which also governs the knock rating of the fuel. Consequently, to the absorption factor of gasoline at a particular wavelength, there corresponds a numerical coefficient that reflects the integrated contribution of gasoline components to the octane number. The linear regression analysis of spectral data is based

1063-7842/00/4509-1177$20.00

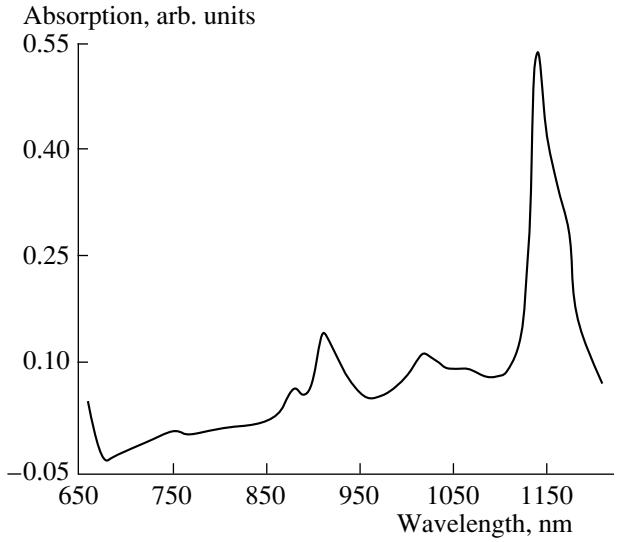


Fig. 1. Absorption spectrum of gasoline in the near-infrared region.

on the assumption that the octane number of a gasoline sample can be written, with a certain accuracy, as

$$y_i = a_0 + \sum_{j=1}^{k} a_j x_{ij}, \qquad (1)$$

where i is the number of a gasoline sample, y_i is the calculated (predicted) octane number of the ith sample, a_j are linear regression coefficients, and x_{ij} are the absorption factors of the ith sample at a wavelength λ_j.

The regression coefficient a_j accounts for the influence of the components responsible for absorption at the wavelength λ_j on the octane number. We assume that a set of the coefficients a_i is the same for all of the gasoline samples. This assumption is valid, since gasolines are chemically similar petroleum products and, hence, exhibit similar relations between the absorption spectrum and knock rating. Therefore, in order to find the gasoline octane number from the absorption spectrum, the set of the regression coefficients a_j should be determined.

The regression coefficients are calculated in several steps. At the initial stage, one makes up a set of different gasoline samples with known (laboratory-found) octane numbers y_i^{lab}, where i is the sample number. Next, the absorption factors of the samples a_i are measured at wavelengths λ_j to obtain the coefficients x_{ij}. Finally, the coefficients a_j are determined by the least-squares method: to do this, one performs a computer search for those a_j's minimizing the root-mean-square deviation of the regression-predicted octane numbers from the laboratory values, i.e., minimizing the quantity W given by

$$W^2 = \frac{1}{n-1}\sum_{i=1}^{n} (y_i^{\text{lab}} - y_i)^2, \qquad (2)$$

where n is the number of gasoline samples used in experiment.

Clearly, the unambiguous determination of the regression coefficients a_j requires the number of samples to exceed the number of wavelengths λ_j used for the absorption measurements. In this work, the standard linear regression analysis was performed to calculate a_j by the least-squares method [6]. Once the initial stage of calibration is complete and a_j's are calculated, the octane number of a gasoline sample under test can be determined from its absorption spectra.

EXPERIMENTAL SETUP

To control the gasoline characteristics, its absorption spectrum is best taken in the near-infrared region at about 1 μm, where several absorption peaks are observed (Fig. 1) [2]. As will be shown later, for the correct calculation of the octane number, the optical absorption (density) of a gasoline should be measured with an accuracy of about 10^{-4} [4]. In order to attain such an accuracy with thermal sources of radiation, one should use special spectral devices, since conventional grating spectrophotometers (for example, SP-6) cannot ensure the required accuracy.

To remedy the situation, we developed a spectral octanometer, which is actually an optical spectrum analyzer adapted to measuring the gasoline parameters in the near-infrared range. The experimental setup is shown in Fig. 2. It is a variant of the two-beam configuration, where one channel is measuring and the other provides referencing to the zero absorption level. The radiation from thermal source *1* is shaped into a narrow parallel beam and passes through a time-variable set of 15 narrow (spectral width $\Delta\lambda \sim 10$ nm) interference filters *2*, which select the spectral range of the probing radiation. Next, passing through an optical power divider, the beam bifurcates, and one of its part falls on chopper–commutator *3*, which eliminates drift in the optical arm at every measuring cycle. The optical absorption of gasoline in each spectral range is varied with an optical cell of length $L = 5$ cm that contains a gasoline sample. Then, the parts of the initial beam combine, and the integral beam is sensed by a silicon photodetector (PD). The PD signal is applied to an electronic processing unit consisting of an input amplifier, analog-to-digital converter, and digital signal processor realizing an algorithm of linear multicomponent analysis.

The described optical arrangement, combined with digital processing of the signals, enables one to measure the optical density accurate to 3×10^{-5}. Then, with the use of the 5-cm-long cell for gasoline samples, the

absorption factor can be measured with an accuracy of approximately 6×10^{-6} cm^{-1}. This value can be considered as ultimate for optical systems based on thermal radiation sources. A further improvement of the accuracy needs laser sources of radiation.

RESULTS AND DISCUSSION

Our system was applied to study commercial gasolines A-92 and A-76 with different octane numbers. The results for A-92 and A-76 are shown in Figs. 3 and 4, respectively. Here, the abscissa is the octane number measured with a UIT-85 test bench, and the ordinate is the octane number calculated by the regression algorithm from the spectral data obtained on the developed setup. The regression coefficients a_j obtained for the corresponding wavelengths are given in Table 1.

It should be pointed out that a more sophisticated treatment of the obtained regression coefficients a_j is impossible because of the complex chemical composition of the fuel. They represent the integral effect of the gasoline components (which absorb at the corresponding wavelengths) on fuel detonation.

The analysis shows that, with the optical analyzer developed to study the knock rating of fuels, the octane number can be measured with a standard deviation of $W \approx 0.1$ for A-92 and $W \approx 0.3$ for A-76 gasolines. This difference can be attributed to the various chemical compositions of the gasolines. A-76 gasoline usually has more components than A-92; hence, the octane number of A-76 is less accurate than that of chemically more homogeneous A-92.

In choosing the spectral range for measuring the octane number, one should take into account the required accuracy of measurements, the hardware complexity of an optical setup, the chemical composition of a fuel, etc. In practice, the number of wavelengths used for absorption measurements with a spectral octanometer is limited primarily by the need to use simple measuring systems. On the other hand, the use of few wavelengths limits the spectral range or spectral resolution of the system. In this connection, it is of interest to determine a probing radiation spectrum that is optimal for obtaining the absorption spectra, i.e., the spectrum that provides sufficiently accurate octane numbers at a small number of wavelengths employed for the spectral measurements. Note that the maximum permissible number of wavelengths needed for regression analysis is limited by special features of the gasoline absorption spectra. As is seen from a typical absorption spectrum of gasoline (Fig. 1), peaks in the 880–1050 nm range are several times lower than the one between 1100 and 1200 nm. Consequently, the measurements of the gasoline absorption will be effective if the optical system is tuned to only one of the two ranges. In this study, the 880–1050 nm range was chosen as it is wider and more suitable for photorecording.

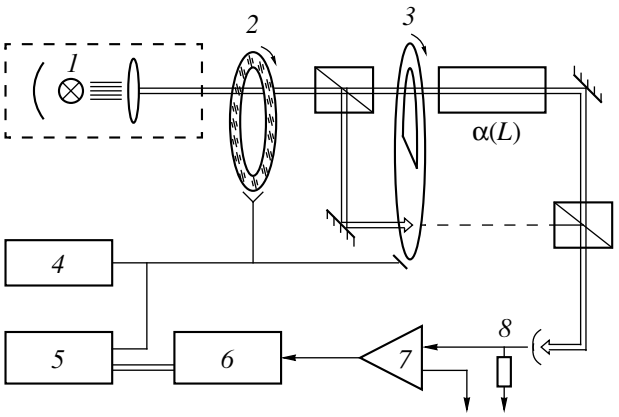


Fig. 2. Spectrum analyzer operating in the near-infrared region. (*1*) radiation source, (*2*) interference filters, (*3*) chopper–commutator, (*4*) power supply, (*5*) computer, (*6*) ADC, (*7*) amplifier, and (*8*) photodetector.

The greatest possible number of wavelengths involved in regression analysis is defined by a relationship between the characteristic spectral width $\Delta\lambda_p$ of absorption peaks and the selected spectral range. The absorption spectrum will obviously be measured correctly if $\Delta\lambda_p/\Delta\lambda_f > 1$, where $\Delta\lambda_f$ is the spectral gap between wavelengths of measurement. For the gasoline spectra, $\Delta\lambda_p \approx 20$–$30$ nm; hence, $\Delta\lambda_f$ can be estimated as ~10–15 nm with some margin. For our octanometer, $\Delta\lambda_f \approx 12$ nm, and the absorption is measured at 15 wavelengths. It was found, however, that an increase in the number of wavelengths does not provide additional information but considerably complicates the

Table 1. Regression coefficients a_j for various wavelengths λ

λ, nm	Coefficient no. a_j	Coefficient value for A-92	Coefficient value for A-76
–	a_0	87	109
880	a_1	–223	283
890	a_2	164	460
905	a_3	–95	–315
916	a_4	170	–221
928	a_5	197	167
940	a_6	259	–223
952	a_7	–181	–121
964	a_8	–146	75
978	a_9	118	243
990	a_{10}	218	315
1000	a_{11}	–193	–219
1014	a_{12}	–267	139
1026	a_{13}	47	–72
1038	a_{14}	–101	104
1050	a_{15}	160	284

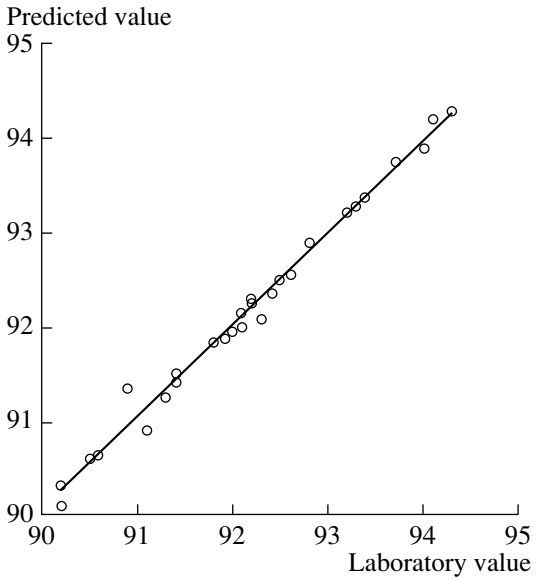


Fig. 3. Predicted versus laboratory values of the octane number for A-92 gasoline.

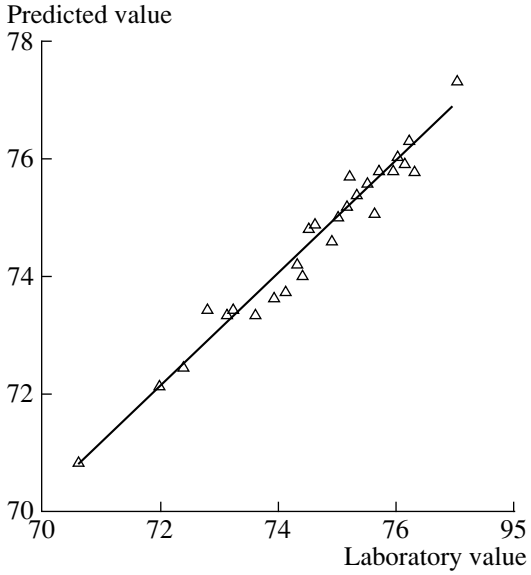


Fig. 4. Predicted versus laboratory values of the octane number for A-76 gasoline.

optical part of the setup. With an eye to the simplification of the measuring setup, we studied the dependence of the octane number accuracy on the number of wavelengths used in experiment. For this purpose, the octane number was predicted with a reduced number of wavelengths in regression analysis. The results of this study are shown in Fig. 5. The standard error in octane number evaluation W versus the number of wavelengths M is plotted for A-76 and A-92 gasolines for the uniform distribution of measuring channels within the spectral range. Note that such a reduction in the number of wavelengths is equivalent to an increase in the spectral gap $\Delta\lambda_f$. Figure 5 shows that an almost ultimate accuracy is attained at eight wavelengths. Moreover, for the A-92 gasoline, the accuracy was better than 0.2 (in units of octane number) even at five wavelengths used in the measurements.

The octane numbers obtained by regression analysis of the absorption factors at five wavelengths in the 890–940 nm range were compared with those obtained in the entire spectral range of 880–1050 nm (the wavelengths are listed in Table 1). The associated results are presented in Table 2.

As follows from Table 2, for both gasolines, the octane number evaluated in the range of 890–950 nm is less accurate as compared to the case when five wavelengths are uniformly spaced in the entire spectral range.

Table 2. Standard error W of the octane number estimated in the two spectral ranges

Range	A-76	A-92
880–1050 nm	$W = 0.28$	$W = 0.09$
890–940 nm	$W = 0.49$	$W = 0.18$

Thus, the number of wavelengths to be used in analysis should be determined for each specific problem; i.e., it depends on the petroleum product and the given prediction accuracy. Note that octane number control of high-octane A-92 gasoline with an accuracy of 0.2–0.3 (an important applied problem) can be performed with five wavelengths in the short-wavelength part of the range. Such conditions are optimum in terms of analysis and instrument simplicity.

Thus, we can design a laser analyzer of petroleum products that consists of a set of injection lasers with specific wavelengths. Laser spectral octanometers are of particular interest due to their potentially high performance (accuracy, etc.) and small dimensions [7, 8]. Remote control of petroleum products during production is another possible application of laser systems, since fiberoptic elements are readily compatible with injection lasers.

The range of 890–940 nm, corresponding to the short-wavelength peak of gasoline absorption, seems to be promising for the optimization of the optical system. It is in this range where practically feasible injection lasers, which are considered as candidates for radiation sources in new versions of the optical analyzer for gasolines, can operate. In particular, injection lasers with multiple quantum wells may cover the considered spectral range by a set of InGaAs and AlGaAs radiators [9, 10]. Superluminescent (light-emitting) diodes with the emission spectrum corresponding to the above $\Delta\lambda_p$ values may also act as efficient radiation sources for spectral analyzers. However, the manufacturing of semiconductor lasing crystals with appropriate spectral parameters is as yet a challenge, thus making the imple-

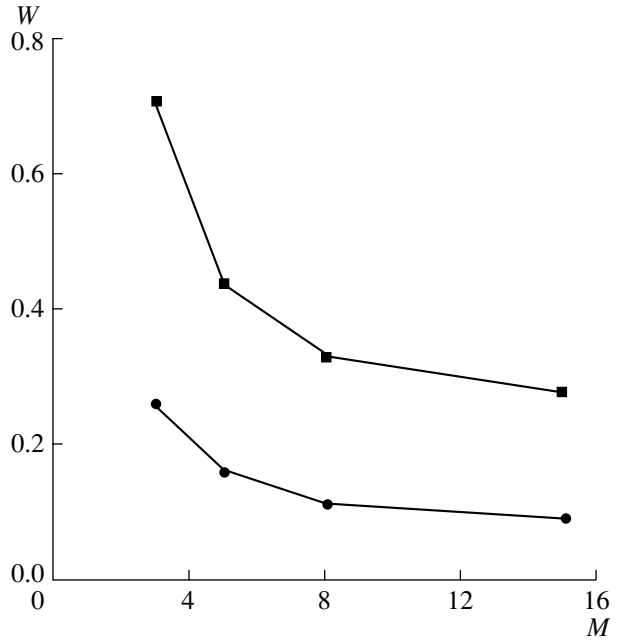


Fig. 5. Standard error W of the octane number versus the number of wavelengths M for (■) A-76 and (●) A-92 gasolines.

mentation of the laser systems difficult. It is hoped that the rapid progress in semiconductor materials technology will provide us with the materials and component basis that are necessary for laser analyzer prototyping.

Along with the proper selection of the spectral range, the correct estimation of the octane number strongly depends on the accuracy with which the gasoline absorption factor is measured. Inexact measurement of the absorption factor causes errors in the octane number obtained by repetitive measurements of the same sample, thus giving a large spread of the octane number. The absolute error in octane number estimation ΔQ can be approximated by

$$\Delta Q \approx \left[\sum_{j=1}^{k} (a_j \Delta x_j)^2 \right]^{1/2}, \qquad (3)$$

where Δx_j is the absolute error involved in the absorption factor of a sample.

Let us assume that the errors in the absorption factor measured at different wavelengths are approximately equal. The experimentally found coefficients a_j lie typically between 100 and 300 (Table 1). The number of wavelengths k is between 5 and 15. Then, it follows from (3) that the octane number can be measured with an accuracy of 0.1 or better (which is foreseen to be an industrial requirement in the future) if the error in the absorption factor should not exceed 10^{-4}.

Such an estimation is confirmed experimentally. As was noted, the spectrometer with a thermal radiator measures the absorption factor accurate to about 3×10^{-5} (for a 5-cm-long optical cell). In such a spectroscopic scheme, the standard deviation ΔQ of the octane number is about 0.03.

CONCLUSIONS

The developed spectrum-selective system, combined with multicomponent regression analysis, allows us to obtain the octane numbers of gasolines with a high accuracy. The key factors having an effect on the resulting error in octane number evaluation were discovered.

Note that the potentialities of the technique extend beyond gasoline analysis. It can be applied to nearly all kinds of petroleum products absorbing in the near-infrared region. However, the use of the spectral systems for estimating other physicochemical parameters of condensed media depends on the problem posed, the nature of the products to be studied, and their spectral characteristics. However, based on the experimental data, we may say that our optical system enables the control of the knock rating of fuels and can be thought of as the starting point for the development of rapid analysis methods.

REFERENCES

1. J. J. Kelly, C. H. Berlow, T. M. Jinguji, and J. B. Callis, Anal. Chem. **61**, 313 (1989).
2. J. J. Kelly and J. B. Callis, Anal. Chem. **62**, 1444 (1990).
3. V. N. Korolev and A. V. Marugin, in *Proceedings of the Scientific Conference on Radiophysics Devoted to the 85th Anniversary of Lobachevsky State University, Nizhni Novgorod, 1998*, p. 20.
4. A. A. Andronov, Yu. V. Kol'tsov, V. N. Korolev, and A. V. Marugin, in *Proceedings of the 3rd All-Russia Conference "Methods and Instrumentation for Measuring Physical Magnitudes," Nizhni Novgorod, 1998*, Part 8, p. 28.
5. V. Yu. Kol'tsov, V. N. Korolev, and S. A. Kusakin, RF Patent No. 2,094,776 (27 October 1997).
6. G. Seber, *Linear Regression Analysis* (Wiley, New York, 1977; Mir, Moscow, 1980).
7. F. Wittrgrafe, M. Hoogerland, and J. Woerdman, Meas. Sci. Technol. **2**, 304 (1991).
8. V. N. Korolev, A. V. Marugin, A. V. Kharchev, and V. B. Tsaregradskiĭ, Zh. Tekh. Fiz. **59** (8), 38 (1989) [Sov. Phys. Tech. Phys. **34**, 863 (1989)].
9. Y. Arakawa and A. Yariv, IEEE J. Quantum Electron. **21**, 1666 (1985).
10. C. Viappiani and G. Rivera, Meas. Sci. Technol. **1**, 1257 (1990).

Translated by A. Sidorova

Technical Physics, Vol. 45, No. 9, 2000, pp. 1182–1185. Translated from Zhurnal Tekhnicheskoĭ Fiziki, Vol. 70, No. 9, 2000, pp. 96–99.

Leakage in $LiNbO_3$ Channel H-Waveguides

L. I. Sotskaya

Institute of Applied Optics, Belarussian Academy of Sciences, Mogilev, 212793 Belarus

E-mail: ipo@physics.belpak.mogilev.by

Received October 12, 1999

Abstract—Attenuation of leaky modes in $LiNbO_3$ channel H-waveguides is studied theoretically. It is shown that leakage significantly affects the parameters of waveguides made of *X*- and *Y*-cut crystals. An approximate expression for attenuation coefficients is obtained from the solution to the scalar eigenvalue problem. © *2000 MAIK "Nauka/Interperiodica".*

Proton exchange is a promising technology for fabricating $LiNbO_3$ optical waveguides [1]. However, in some cases, their modes experience significant attenuation [2]. While, in planar waveguides with the basic cuts, the attenuation can be minimized by selecting appropriate exchange conditions [2–4], in channel waveguides, losses are unavoidable. The reason is that channel devices guide only the modes whose propagation constants β satisfy the condition $\varepsilon_e^{(s)} < \mathrm{Re}(k_0^{-2}\beta^2) < \varepsilon_o^{(s)}$, where $k_0 = 2\pi\lambda_0^{-1}$ is the wave number in free space and $\varepsilon_o^{(s)}$ and $\varepsilon_e^{(s)}$ are the eigenvalues of the crystal permittivity tensor outside the waveguide. Such modes are leaky because of the hybrid nature of their fields [5]. In contrast to absorption or scattering, the leakage effect is the result of two-dimensional restrictions imposed on the mode fields in channel waveguides. Leakage is also associated with crystal anisotropy. Hence, at the given distribution of the waveguide permittivity, the leakage cannot be minimized by technological means. Therefore, evaluating the leaky-wave attenuation coefficients is equivalent to estimating the lower boundary of $|\mathrm{Im}\beta|$. Although these estimates are of fundamental importance, they have not been made so far. This gap can be bridged with the recent rigorous method for calculating the characteristics of leaky modes in anisotropic channel waveguides [6]. In this paper, we use this method to study the leaky-wave attenuation coefficients in proton-exchanged and annealed proton-exchanged channel waveguides made of *X*-, *Y*-, and *Z*-cut $LiNbO_3$ crystals. Leakage is shown to significantly affect the characteristics of the *X*- and *Y*-cut waveguides. An approximate analytical estimate of their attenuation coefficients is obtained from the solution to the scalar eigenvalue problem.

Figure 1 illustrates the cross section of the studied waveguides. The region $y > 0$ is occupied by a homogeneous isotropic dielectric with a permittivity ε_d, and the region $y < 0$ is a $LiNbO_3$ crystal having a permittivity tensor ε. We will restrict the discussion to the basic cuts under the assumption that the tensor ε is diagonal and has the eigenvalues $\varepsilon_{o,e}(x, y) = \varepsilon_{o,e}^{(s)} + \Delta\varepsilon\varphi_{o,e}(\xi, \eta)$, where $\Delta\varepsilon$ is a constant and the functions $\varphi_{o,e}(\xi, \eta)$ specify the waveguide shape. Their arguments are $\xi = xM^{-1}$ and $\eta = yM^{-1}$, where M is a scaling factor. If the fields depend on time and z coordinate as $\exp[i(\omega t - \beta z)]$, the transverse magnetic field components of the channel waveguide modes are given by [7]

$$\begin{gathered}\varepsilon_{yy}\frac{\partial}{\partial y}\left[\frac{1}{\varepsilon_{zz}}\left(\frac{\partial}{\partial x}H_y - \frac{\partial}{\partial y}H_x\right)\right] \\ -\frac{\partial^2}{\partial x^2}H_x - \frac{\partial^2}{\partial x\partial y}H_y - (k_0^2\varepsilon_{yy} - \beta^2)H_x = 0, \\ \varepsilon_{xx}\frac{\partial}{\partial x}\left[\frac{1}{\varepsilon_{zz}}\left(\frac{\partial}{\partial x}H_y - \frac{\partial}{\partial y}H_x\right)\right] \\ +\frac{\partial^2}{\partial y^2}H_y + \frac{\partial^2}{\partial x\partial y}H_y + (k_0^2\varepsilon_{xx} - \beta^2)H_y = 0\end{gathered} \tag{1}$$

throughout the space, where ε_{jj} are the components of the tensor ε, which are equal to ε_d at $y > 0$.

In the region $y < 0$, $\varepsilon_{zz} = \varepsilon_o$ and the ε_{xx} and ε_{yy} components depend on the crystal cut: $\varepsilon_{xx} = \varepsilon_e$ and $\varepsilon_{yy} = \varepsilon_o$ for the *X* and *Y* cuts, while $\varepsilon_{xx} = \varepsilon_o$ and $\varepsilon_{yy} = \varepsilon_e$ for the *Z* cut.

Leaky modes are more conveniently found if system (1) is replaced by an equivalent system of integro-differential equations for field components inside the waveguide [6]. Curves *1* and *2* in Figs. 1 and 2 show results of its numerical solution. In our calculations, we use the model $\varphi_o = \alpha\varphi_e$,

$$\begin{gathered}\varphi_e = \frac{1}{4}\left\{\mathrm{erf}\left[\frac{B}{\gamma}\left(\xi + \frac{F}{2}\right)\right] - \mathrm{erf}\left[\frac{B}{\gamma}\left(\xi - \frac{F}{2}\right)\right]\right\} \\ \times\{\mathrm{erf}[B(\eta + 1)] - \mathrm{erf}[B(\gamma - 1)]\},\end{gathered} \tag{2}$$

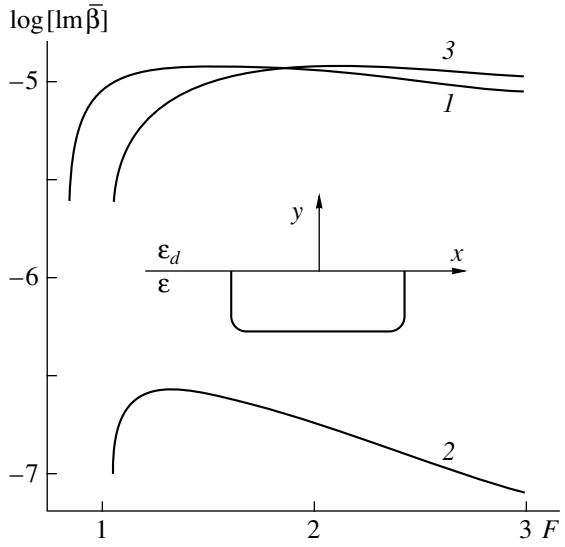


Fig. 1. Cross section of the waveguides (insert) and propagation coefficients in the unannealed waveguides vs. their relative width: (*1*) *X*, (*2*) *Z*, and (*3*) *Y* cuts of $LiNbO_3$.

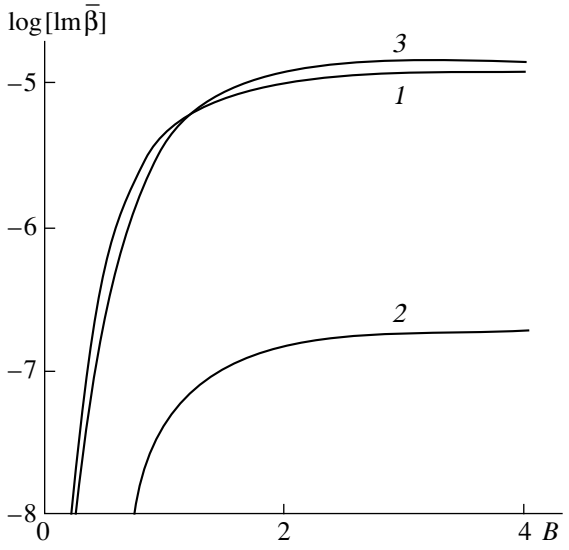


Fig. 2. $\mathrm{Im}\bar{\beta}$ (*B*) for the annealed $LiNbO_3$ waveguides with the (*1*) *X*, (*2*) *Z*, and (*3*) *Y* cuts at $F = 2$.

which approximately describes the annealed proton-exchanged waveguides [8, 9]. Here, $B = MD^{-1}$, and D and γ are the coefficients depending on annealing temperature and time and also on the anisotropy of proton diffusion [9]. According to (2), when $D \longrightarrow 0$ and $B \longrightarrow \infty$ (no annealing), $\varphi_e \equiv 1$ in the region bounded by $|x| < 0.5FM$ and $|y| < M$; outside this region, $\varphi_e \equiv 0$. Thus, the parameters M and F define the dimensions of the unannealed proton-exchange waveguides, which are rectangular in the framework of model (2). In our calculations, we set $\gamma = 0.84$ for the *X* and *Y* cuts of the crystal and 1.2 for the *Z* cut, $\varepsilon_d = 1$, $\varepsilon_o^{(s)} = 5.2$, $\varepsilon_e^{(s)} = 4.84$, $V = k_0 M\sqrt{\Delta\varepsilon} = 2.5$, $\Delta\varepsilon = 0.45$, and $\alpha = -0.25$ [2, 9]. Figure 1 plots the normalized attenuation coefficients $\mathrm{Im}\bar{\beta}$ ($\bar{\beta} = k_0^{-1}\beta$) of leaky modes against the relative width F of the unannealed waveguides. These curves refer to the single-mode waveguides (the single-mode range is $0.86 < F < 3.11$ for the *X* and *Y* cuts and $1.05 < F < 3.79$ for the *Z* cut). As shown in Fig. 1, the attenuation decreases as the waveguide becomes wider (this behavior could be expected, because the modes of planar waveguides described by the diagonal tensor ε are not leaky [5]), or when F approaches the critical values. However, these features can hardly be used to minimize the attenuation, since either the waveguides become multimode (which is undesirable) or their parameters approach the critical values. A more efficient means to lower $|\mathrm{Im}\beta|$ is annealing. From Fig. 2, as D grows and, hence, the maximum increment in ε_e, equal to $\Delta\varepsilon\mathrm{erf}(0.5FB\gamma^{-1})\mathrm{erf}(B)$, decreases, the attenuation rapidly declines. This can qualitatively be explained by weaker coupling between the extraordinary and ordinary waves, which causes leakage [10].

According to Figs. 1 and 2, modes generated in the *X*- and *Y*-cut $LiNbO_3$ waveguides experience a much higher attenuation than those in the *Z*-cut crystals (in particular, at $\lambda_0 = 0.63$ μm and $F = 2$, curves *1* and *2* in Fig. 1 correspond to the attenuations 9.6 and 0.15 dB/cm, respectively). This feature results from the inequalities

$$\Delta\varepsilon(\varepsilon_o^{(s)})^{-1} \ll 1, \quad (\varepsilon_o^{(s)} - \varepsilon_e^{(s)})(\varepsilon_o^{(s)})^{-1} \ll 1,$$
$$\Delta\varepsilon(\varepsilon_e^{(s)} - \varepsilon_d)^{-1} \ll 1,$$

which are valid for these waveguides. This allows us to analyze system (1) asymptotically by a technique similar to that applied in [5, 10] to optical fibers. Eventually, we obtain the approximation

$$\mathrm{Im}\beta = k_0(\Delta\varepsilon)^3 \mathrm{Im}b(2V^4\sqrt{\varepsilon_e^{(s)}})^{-1}, \tag{3}$$

where Imb depends on the crystal cut.

For the *Z* cut,

$$\mathrm{Im}b = \left(V^2\varepsilon_o^{(s)}\int_{-\infty}^{\infty}\int_{-\infty}^{0} u^2 d\xi d\eta\right)^{-1}$$
$$\times \int_{-\infty}^{\infty}\left\{\left[\frac{\partial^2 u}{\partial\xi\partial\eta}\frac{\partial}{\partial\eta}\mathrm{Im}(H_y)\right]_{\eta=0} + V^2\int_{-\infty}^{\infty}\mathrm{Im}(H_y)\right.$$
$$\left.\times\left[\frac{\partial u}{\partial\eta}\frac{\partial}{\partial\xi}(\varphi_o - \varphi_e) - \frac{\partial}{\partial\xi}\left(u\frac{\partial}{\partial\eta}\right)\right.\right. \tag{4}$$

$$+ (\varphi_o - \varphi_e - A)\frac{\partial^2 u}{\partial\xi\partial\eta}\Big]d\eta\Big\}d\xi,$$

where $A = (\varepsilon_e^{(s)} - \varepsilon_o^{(s)})(\Delta\varepsilon)^{-1}$.

For the *X* and *Y* cuts,

$$\mathrm{Im}b = -\left(\varepsilon_o^{(s)}\int_{-\infty}^{\infty}\int_{-\infty}^{0}u^2 d\xi d\eta\right)^{-1}\int_{-\infty}^{\infty}\int_{-\infty}^{0}\mathrm{Im}(H_x)\frac{\partial}{\partial\eta}$$
$$\times\Big[\frac{\partial}{\partial\xi}(u\varphi_e) + (A-\varphi_o)\frac{\partial u}{\partial\xi}\Big]d\xi d\eta. \quad (5)$$

In formulas (4) and (5), the values of $H_{x,y}$ mean the leaky-wave magnetic field components determined in the principal order of perturbation theory. They obey the integral equations

$$H_{x,y}(\xi,\eta) = F_{x,y}(\xi,\eta)$$
$$+\int_{-\infty}^{\infty}\int_{-\infty}^{0}\Gamma(\xi,\xi',\eta,\eta')H_{x,y}(\xi',\eta')d\xi' d\eta', \quad (6)$$

$$F_x = (\varepsilon_o^{(s)})^{-1}\int_{-\infty}^{\infty}\Big\{\Big[\frac{\partial^2 u(\xi',\eta')}{\partial\xi'\partial\eta'}\frac{\partial G}{\partial\eta'}\Big]_{\eta'=0}$$
$$-V^2\int_{-\infty}^{\infty}G\frac{\partial}{\partial\eta'}[\varphi_0(\xi',\eta')]\frac{\partial u(\xi',\eta')}{\partial\xi'}d\eta'\Big\}d\xi', \quad (7)$$

$$F_y = -V^2(\varepsilon_o^{(s)})^{-1}$$
$$\times\int_{-\infty}^{\infty}\int_{-\infty}^{0}G\frac{\partial}{\partial\xi'}[\varphi_0(\xi',\eta')]\frac{\partial u(\xi',\eta')}{\partial\xi'}d\eta' d\xi',$$

$$\Gamma = -V^2 G\varphi_o(\xi',\eta'), \quad (8)$$

$$G = 0.25i[H_0^{(2)}(z_1) - H_0^{(2)}(z_2)],$$

$$z_n = V\sqrt{|b_o + A|[(\xi'-\xi)^2 + (\eta' + (-1)^n\eta)^2]},$$

where $H_0^{(2)}(z)$ is a Hankel function of the second kind.

The function $u(\xi,\eta)$, which enters into expressions (4), (5), (7), and (8), is the solution of the scalar eigenvalue problem for the guided waves:

$$\frac{\partial^2 u}{\partial\xi^2} + \frac{\partial^2 u}{\partial\eta^2} + V^2(\varphi_e - b_0)u = 0,$$
$$u|_{\eta=0} \equiv 0, \quad \lim_{\xi^2+\eta^2\to\infty} u = 0. \quad (9)$$

Note that, in $LiNbO_3$ H-waveguides, the refractive index of the extraordinary wave increases to the greatest extent [1]. This means that $|\varphi_o| \ll \varphi_e$. If an increase in the refractive index of the ordinary wave is completely neglected (assuming that $\varphi_o \equiv 0$), then expressions (3), (4), (6), and (8) yield $H_y \equiv 0$ and $\mathrm{Im}\beta = 0$. In this approximation, modes of the *Z*-cut channel waveguides do not leak. This conclusion can also be derived rigorously by virtue of the fact that, for the *Z* cut at $\partial\varphi_o/\partial\xi = 0$, the extraordinary and ordinary waves are uncoupled [11]. At the same time, the condition $\varphi_o \equiv 0$ does not eliminate leakage in the *X*- and *Y*-cut waveguides. Here, these waves are coupled through mutual transformation on reflection from the interface $y = 0$ [12]. This coupling is taken into account by the first term in (7). As a result, when $|\varphi_o| \longrightarrow 0$, equation (6) for $H_x(\xi,\eta)$ turns into the direct computational formula

$$H_x(\xi,\eta) = (\varepsilon_o^{(s)})^{-1}\int_{-\infty}^{\infty}\Big[\frac{\partial^2 u(\xi',\eta')}{\partial\xi'\partial\eta'}\frac{\partial G}{\partial\eta'}\Big]_{\eta'=0}d\xi'. \quad (10)$$

Thus, the attenuation of the leaky modes in the *Z*-cut channel $LiNbO_3$ H-waveguides is relatively low, because $|\varphi_0|$ is small. At the same time, the smallness of $|\varphi_0|$ provides a simple approximation for the attenuations of the modes in the *X*- and *Y*-cut waveguides. One should solve problem (9) and then take advantage of formulas (3), (5), and (10). Curves *3* in Figs. 1 and 2 were calculated in this approximation. The values of b_0 and functions $u(\xi,\eta)$ were found by the method of variational separation of variables [13]. An increased error near the critical conditions is typical of this method [13], which explains the noticeable divergence between curves *1* and *3* in Fig. 1 near the critical *F*'s. Nevertheless, expressions (3), (5), and (10) can be thought of as good estimates.

Note in conclusion that the results presented above are only tentative, because the components of the permittivity tensor were described in terms of the approximate model. A more adequate model should allow for the presence of several crystal phases in the waveguide [2, 3, 14]. However, at present, an appropriate model of this kind is absent.

ACKNOWLEDGMENTS

This work was supported by the Foundation for Basic Research of Republic of Belarus.

REFERENCES

1. L. D. Hutcheson, *Integrated Optical Circuits and Components* (Marcel Dekker, New York, 1987).
2. S. Chen, P. Baldi, M. P. De Micheli, *et al.*, J. Lightwave Technol. **12**, 862 (1994).
3. Yu. N. Korkishko and V. A. Fedorov, Zh. Tekh. Fiz. **66** (5), 86 (1996) [Tech. Phys. **41**, 453 (1996)].
4. Y. S. Son, H. J. Lee, S. Y. Yi, and S. Y. Shin, Proc. SPIE **1374**, 23 (1990).

5. A. W. Snyder and J. D. Love, *Optical Waveguide Theory* (Chapman and Hall, London, 1983; Radio i Svyaz', Moscow, 1987).
6. A. B. Sotsky and L. I. Sotskaya, Opt. Quantum Electron. **31**, 733 (1999).
7. A. B. Sotskiĭ, L. I. Sotskaya, and Yu. D. Stolyarov, Izv. Vyssh. Uchebn. Zaved., Radiofiz. **30**, 1470 (1987).
8. K. M. Kissa, P. G. Suchoski, and D. K. Lewis, J. Lightwave Technol. **13**, 1521 (1995).
9. U. Hempelmann, H. Herrmann, G. Mrozynsky, *et al.*, J. Lightwave Technol. **13**, 1750 (1995).
10. M. Lu and M. M. Fejer, J. Opt. Soc. Am. A **10**, 246 (1993).
11. V. A. Karpenko, Radiotekh. Élektron. (Moscow) **29**, 843 (1984).
12. F. I. Fedorov and V. V. Filippov, *Reflection and Refraction of Light by Transparent Crystals* (Nauka i Tekhnika, Minsk, 1976).
13. A. B. Sotskiĭ, L. I. Sotskaya, and Yu. D. Stolyarov, Radiotekh. Élektron. (Moscow) **34**, 1158 (1989).
14. Yu. N. Korkishko and V. A. Fedorov, Zh. Tekh. Fiz. **69** (3), 47 (1999) [Tech. Phys. **44**, 307 (1999)].

Translated by A. Khzmalyan

Technical Physics, Vol. 45, No. 9, 2000, pp. 1186–1191. Translated from Zhurnal Tekhnicheskoĭ Fiziki, Vol. 70, No. 9, 2000, pp. 100–105.

ACOUSTICS, ACOUSTOELECTRONICS

Acousto-optic Filter of Nonpolarized Electromagnetic Radiation

V. B. Voloshinov, V. Ya. Molchanov, and T. M. Babkina

Moscow State University, Vorob'evy gory, Moscow, 119899 Russia

e-mail: volosh@osc162.phys.msu.su

Received May 18, 1999; in final form, November 10, 1999

Abstract—An acousto-optic tunable filter that is insensitive to the polarization of incident light is studied both theoretically and experimentally. The fundamental possibility of designing an acousto-optic filter of nonpolarized light is demonstrated. The filter operates in the visible range of electromagnetic waves with λ = 480–750 nm and provides a high spatial and temporal resolution. The number of resolvable spots of the filtered image reaches a value of 400×400 with the spectral bandwidth of the device being $\Delta\lambda = 1.7$ nm. The results of the processing of optical images formed by arbitrarily polarized light beams are discussed. An acousto-optic filter based on a TeO_2 crystal can be used in optics and spectroscopy in the processing of light beams with an arbitrary polarization, as well as in analyzing optical images formed by light beams whose polarization varies with time. © *2000 MAIK "Nauka/Interperiodica".*

INTRODUCTION

It is well known that acousto-optic filters provide spectral filtering and optoelectronic processing of collimated and uncollimated light beams and optical images [1–4]. The interest in acousto-optic filters is explained by the diversity of possible applications of these devices, their simple structure, narrow spectral bandwidth, and their ability to be easily electronically tuned in a wide range of optical wavelengths. Unfortunately, not all characteristics of acousto-optic filters satisfy the practical needs. For example, tunable filters are sensitive to the polarization of incident light, because, for the operation of these devices, it is necessary to use linearly polarized radiation [4–7]. However, the use of polarizers often leads to a decrease in the incident light beam power by 50% or more [8, 9]. The problem of filtering nonpolarized light can be solved in several ways [8–10]. One of the possible methods is based on the use of two detectors and an additional electron processing of two optical images at the filter output. This method requires a fairly complicated filtering system and is rather time-consuming.

In this paper, we propose a scheme of filtering for arbitrarily polarized optical signals. Our method is free of the disadvantages characteristic of other methods. It is based on the use of additional passive optical elements such as prisms and mirrors. These optical elements divide an arbitrarily polarized incident light beam into ordinary and extraordinary components and direct them to the acoustic cell at angles that satisfy the Bragg conditions for the ordinary (*o*) and extraordinary (*e*) polarized light beams, respectively. As a result, at the filter output, we obtain a single diffracted light beam, which is quite convenient from the practical point of view.

PRINCIPLE OF OPERATION OF PARATELLURITE-BASED FILTERS

The mechanism of electromagnetic radiation filtering in tunable acousto-optic filters is based on the selectivity of the anisotropic Bragg scattering of light by an acoustic wave [1–4]. The spectral filtering of optical signals is possible with the use of a filter made on the basis of a paratellurite crystal and characterized by a noncollinear geometry of the acousto-optic interaction [4–7]. The selection of the light wavelengths occurs as a result of the Bragg scattering of the optical beam by the acoustic wave propagating in paratellurite near the [110] direction in the $(1\bar{1}0)$ plane. With the use of this acoustic mode, one obtains a high acousto-optical quality M_2 of the crystal [1–3] and a diffraction efficiency close to unity at low levels of the control electric power *P*.

The Bragg diffraction can be considered as the scattering of photons by phonons, which satisfies the energy and momentum conservation laws. For the scheme of filtering in use, the wave vector relationships have the form [1–3]

$$\begin{aligned} \mathbf{k}_i^o - \mathbf{K} - \Delta\mathbf{K} - \Delta\mathbf{k} &= \mathbf{k}_d^e, \\ \mathbf{k}_i^e + \mathbf{K} + \Delta\mathbf{K} + \Delta\mathbf{k} &= \mathbf{k}_d^o, \end{aligned} \quad (1)$$

1063-7842/00/4509-1186$20.00 © 2000 MAIK "Nauka/Interperiodica"

where $\mathbf{k}_i$ and $\mathbf{k}_d$ are the wave vectors of the incident and diffracted light for the ordinary and extraordinary polarized beams, respectively; $\mathbf{K}$ denotes the acoustic wave vectors; and $\Delta\mathbf{K}$ and $\Delta\mathbf{k}$ are the variations of the acoustic vector and the mismatch vectors.

According to Eqs. (1), the vector $\mathbf{k}_d^o$ is formed as the sum of the wave vectors of the incident light wave and the acoustic wave, and the vector $\mathbf{k}_d^e$ is formed as the difference between these wave vectors. The wave vector diagram illustrating Eqs. (1) is shown in Fig. 1. In this diagram, one can see the [001] optical axis of the crystal and the Bragg angles Θ_B^o and Θ_B^e. The acoustic waves propagate in the crystal at an angle α to the [110] axis in the plane of the acousto-optic interaction, as shown in Fig. 1.

For the magnitudes of the wave vectors of the interacting beams, the following relationships are valid: $k_d = 2\pi n_0/\lambda$, $K = 2\pi f/V$, and $\Delta K = 2\pi\Delta f/V$, where V is the phase velocity of sound and f and Δf are the frequency of the acoustic wave and its variation in the case of the violation of the matching conditions [5]. In the experiment, we used a filter with the ultrasound propagation direction making an angle $\alpha = 10°$ with the [110]-axis. For the selected cut of the TeO_2 crystal and the light wavelength $\lambda = 0.63$ μm, the refraction indices are $n_o = 2.26$ and $n_e = 2.41$ for the ordinary and extraordinary waves, respectively; the phase velocity of sound reaches the value $V = 7.08 \times 10^4$ m/s.

The wave number of incident light has the form $k_i^e = 2\pi n_i/\lambda$, where n_i is the refraction index of the extraordinary light wave ($n_0 \le n_i \le n_e$). For the geometry of the acousto-optic interaction under study, the parameter n_i is expressed as [7]

$$n_i = \frac{n_o n_e}{\sqrt{n_o^2 \sin^2(\Theta_B^e + \alpha) + n_e^2 \cos^2(\Theta_B^e + \alpha)}}. \qquad (2)$$

For light with the extraordinary polarization, the dependence of the acoustic frequency on the Bragg angle has the form [8]

$$f = \frac{V}{\lambda}(n_i \sin\Theta_B^e - \sqrt{n_o^2 - n_i^2 \cos^2\Theta_B^e}). \qquad (3)$$

From the wave vector diagram, one can calculate the diffraction angle Θ_d^o:

$$\Theta_d^o = \arccos\left(\frac{n_i}{n_o}\cos\Theta_B^e\right). \qquad (4)$$

Using Eqs. (3) and (4) for light with the wavelength $\lambda = 0.63$ nm, we can plot the frequency dependences of the Bragg angle $\Theta_B^e(f)$ and the diffraction angle $\Theta_d^o(f)$

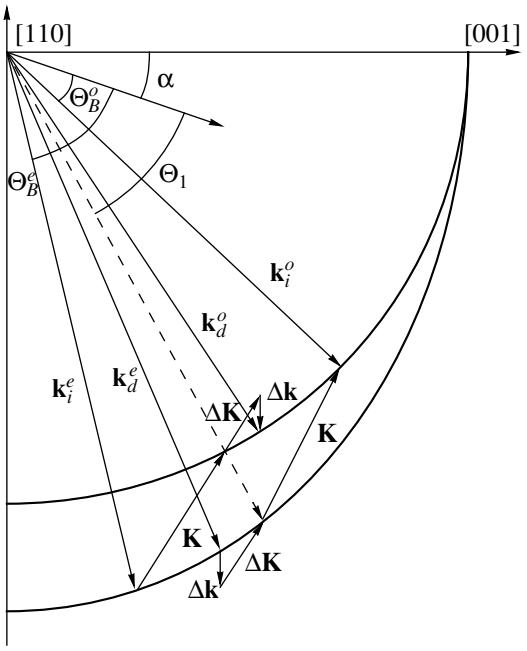


Fig. 1. Wave vector diagram of the acousto-optic interaction for extraordinary polarized light.

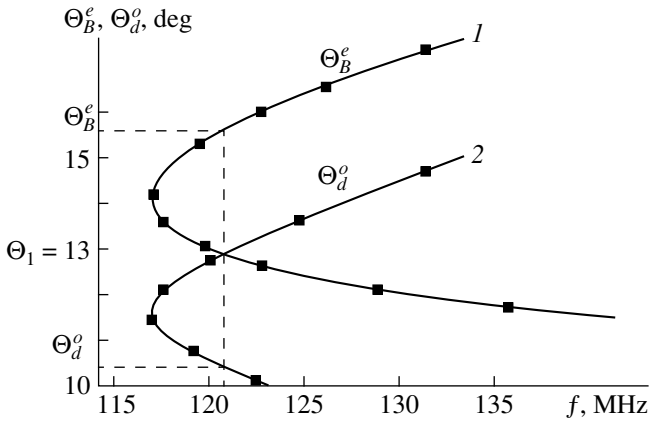


Fig. 2. Frequency dependence of the Bragg angle and the diffraction angle for extraordinary polarized incident light.

for the case of an extraordinary polarized incident light beam. These dependences are shown in Fig. 2, where curve *1* refers to the incident light beam, and curve *2* refers to the diffracted beam. The plots correspond to the fulfillment of the Bragg condition, i.e., to the zero mismatch. In the same way, one can calculate the frequency dependences for the ordinary polarized incident light. In this case, the dependence of the Bragg angle of incidence will have the form similar to that shown in Fig. 2, but curve *1* will correspond to diffracted light, and curve *2* will be for the incident light.

DIFFRACTION OF ARBITRARILY POLARIZED LIGHT

From Fig. 2 and the results of the analysis, we can infer that the frequency dependences of the angles of incidence of the ordinary and extraordinary polarized light beams intersect when the frequency of ultrasound is $f_1 = 120.3$ MHz. The intersection occurs at the angle of incidence of light $\Theta_1 = 13°$. If light propagates at this angle to the acoustic wave front, the Bragg matching condition is automatically fulfilled for both ordinary and extraordinary polarized light [8, 9]. In this case, the acousto-optic cell becomes insensitive to the polarization of light, because the light beams with each of the two polarizations simultaneously satisfy the Bragg condition and interact with ultrasound with maximum efficiency [8–10]. The described diffraction regime is used in the known acousto-optic filter intended for analyzing arbitrarily polarized light beams [8]. The specific feature of this filter is the presence of two diffracted light beams propagating at the output of the acousto-optic cell symmetrically with respect to the transmitted light beam. From Fig. 2, one can see that, when the sound frequency is $f = f_1$, the diffraction angle of the ordinary polarized diffracted light beam is $\Theta_d^o = 10.4°$, while the extraordinary polarized diffracted light beam propagates at the angle $\Theta_d^e = 15.6°$ to the acoustic wave front.

At the frequency of the acoustic wave $f = f_1$, one more diffraction regime that is of practical interest can be realized [9]. This regime was used in our study described below. From Fig. 2, it follows that, at $f = f_1$, the Bragg condition is fulfilled for the extraordinary polarized incident light when the angle of incidence is $\Theta_B^e = 15.6°$. Evidently, in this case the diffraction angle is equal to $\Theta_d^o = \Theta_1 = 13°$. On the other hand, for the ordinary polarized incident light, the Bragg condition is fulfilled at the same ultrasonic frequency when the angle of incidence is equal to $\Theta_B^o = 10.4°$. In this case, the diffraction angle is also equal to $\Theta_d^e = 13°$. Thus, the diffraction angles for the ordinary and extraordinary polarized light beams will be equal. Hence, in Fig. 1, both diffracted light beams will have collinear wave vectors directed along the dashed line. Therefore, we can conclude that the diffracted light beams propagate in the crystal in one direction [9].

Thus, by directing two orthogonally polarized light beams to the cell at the angles $\Theta_B^o = 10.4°$ and $\Theta_B^e = 15.6°$ and using ultrasound of frequency $f_1 = 120.3$ MHz, one can observe a simultaneous diffraction of the ordinary and extraordinary polarized incident light beams. As a result of the diffraction, three light beams will emerge at the output of the acousto-optic cell. Two of them are transmitted light beams of zero diffraction order, and the third beam consists of two diffracted beams of +1 and –1 diffraction orders. The diffracted beam has an arbitrary polarization, because it is a sum of the diffracted components of two initial waves, and the direction of its polarization always coincides with that of the initial light beam.

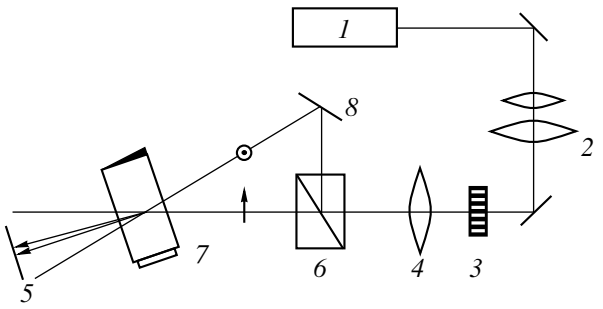


Fig. 3. Schematic diagram of the experimental setup.

INTERACTION IN THE CASE OF THE VIOLATION OF THE MATCHING CONDITION

One of the most important characteristics of acousto-optic filtering is the spectral bandwidth $\Delta\lambda$ of the acousto-optic filter [1]. It is known that the spectral resolution $R = \lambda/\Delta\lambda$ of an acousto-optic device depends on the divergence of the optical beam and the boundedness of the region in which the diffraction of light occurs [1–3]. In the case of a noncollinear acousto-optic interaction, this region is determined by the width of the acoustic column l, which in turn depends on the length of the piezoelectric transducer l_0. The spectral bandwidth of the acousto-optic filter is calculated by the formula $\Delta\lambda(\Delta f/f)$, where Δf is the diffraction frequency bandwidth near the acoustic frequency of the Bragg matching condition [1]. The frequency bandwidth Δf can be most easily calculated at a fixed wavelength of light λ with the use of the wave vector diagrams shown in Fig. 1 that allow for the mismatch vectors $\Delta\mathbf{k}$ [8].

In the acousto-optic interaction, the mismatch vectors $\Delta\mathbf{k}$, being orthogonal to the boundaries of the acoustic column, are related to the acoustic column width l by the condition $\Delta k \le 0.8\pi/l$ [1]. The width of the acoustic column depends on the angle ψ between the vectors of the phase and group velocities of sound, $l = l_0\cos\psi$. When the sound wave vector is oriented along the direction $\alpha = 10°$, the acoustic energy "walk-off" angle is equal to $\psi = 57°$ [7]. Evidently, the limitation for the magnitude of the mismatch vector takes the form

$$\Delta k \le 0.8\pi/l_0\cos\psi. \quad (5)$$

From the wave vector diagrams of Fig. 1, we obtain the relation between the magnitude of the vector $\Delta\mathbf{k}$ and the vector $\Delta\mathbf{K}$. Thus, the wave vector diagram and Eq. (5) provide the expression for the acoustic frequency bandwidth

$$\Delta f \le \frac{0.8K}{l_0\cos\psi}\frac{n_i V\cos(\psi-\Theta_B^e)-\lambda f\sin\psi}{n_i V\sin\Theta_B^e-\lambda f}. \qquad (6)$$

Assuming that the birefringence of the material is small and taking into account the condition $\Delta K \gg K$, from Eq. (6) we derive the expression for the acoustic frequency bandwidth Δf and for the spectral bandwidth of the filter

$$\Delta\lambda \le \frac{0.8\lambda^2}{l_0\cos\psi\Delta n\sin^2(\Theta_B^2+\alpha)}, \qquad (7)$$

where $\Delta n = n_e - n_o$.

Estimates by Eq. (7) show that, for the selected geometry of the interaction and for the transducer length $l_0 = 1.2$ cm, the spectral bandwidth of the filter is $\Delta\lambda = 1.7$ nm.

It should be noted that the regime of acousto-optic interaction under study is close to the case of a wide-aperture diffraction [4]. Therefore, the selected geometry of interaction provides not only narrow bandwidths, but also sufficiently wide angular apertures for the filtering of light beams. In principle, this makes it possible to process the light beams forming the image without a loss in the optical radiation power. However, the main advantage of the regime of interaction used in our study is that it is insensitive to the polarization of the incident light.

EXPERIMENTAL STUDY OF DIFFRACTION

The acousto-optic cell used for filtering of optical images was made on the basis of a paratellurite single crystal. Evidently, for such purposes, any birefringent material can be used if the desired diffraction regime can be realized in it. Our choice of a paratellurite crystal was determined by the high acousto-optical quality of this material [1–4].

The acoustic waves were generated in paratellurite by a piezoelectric transducer made of an *x*-cut lithium niobate crystal; the length of the transducer was $l_0 = 1.5$ cm, and its width was $d = 0.5$ cm. The experimental setup is schematically represented in Fig. 3. To study the acousto-optic cell, we used a He-Ne laser with the wavelength $\lambda = 0.63$ μm as a source of light. In the processing of incoherent images, the source of light was an incandescent lamp *1* with a continuous spectrum of optical frequencies. The light beam from the laser or the lamp was broadened by an optical system *2* and directed to a transparency *3*. A lens *4* formed the image of the transparency on the screen *5* or on the window of a matrix image detector. In the path of the light beam, a polarizing beam splitter *6* was placed. Upon transmission through the splitter, the ordinary polarized light beams were directed to the acousto-optic cell *7*, and the extraordinary polarized components of incident light were directed first to a mirror *8* and then to the crystal. The mirror and the cell of the filter were tuned in such a way that the Bragg matching condition was automatically fulfilled for each of the two orthogonally polarized incident light beams at the same ultrasonic frequency.

The transparency was illuminated with a light beam, and an electric signal of frequency $f_1 = 120.3$ MHz was supplied from a high-frequency oscillator to the cell. As a result, two diffracted light beams propagating in one direction were observed at the output of the acousto-optic cell, as shown in Fig. 3. The polarization of this light coincided with the initial one, which demonstrated the ability of the filter to provide the processing of arbitrarily polarized light beams.

The preliminary experimental studies carried out with the source of coherent light with $\lambda = 0.63$ μm showed that, when the control high-frequency power was $P = 1.2$ W, the diffraction efficiency was $T = 90\%$. In the experiment, we determined the dependence of the Bragg angles of incidence on the acoustic frequency for the ordinary and extraordinary polarized incident light. The experimental points shown in Fig. 2 agree well with the computational data. Thus, the experiment confirmed that a wide-aperture geometry of acousto-optic interaction was realized in the cell, and the selected crystal cut allowed one to work with arbitrarily polarized light beams. Using a laser, we measured the diffraction frequency bandwidth Δf by the level of a 3-dB decrease in the diffraction efficiency. The frequency bandwidth was found to be equal to $\Delta f = 320$ Hz, which at the acoustic frequency $f = 120.3$ MHz corresponded to the spectral bandwidth of the filter $\Delta\lambda = 1.7$ nm.

ACOUSTO-OPTIC PROCESSING OF IMAGES IN NONPOLARIZED LIGHT

In the processing of optical beams in the case of incoherent light, the transparency *3* was a photographic film with a recorded pattern. The lens that formed the optical image had the focal length $F = 7.5$ cm. If the angular divergence of light was less than 6°, the transmitted and diffracted light beams did not overlap at the filter output. In this case, the matrix detector at the filter output recorded only filtered light beams.

At the ultrasonic frequency $f_1 = 120$ MHz, the screen displayed a filtered optical image formed by optical beams with the wavelengths corresponding to the red spectral region. The variations in the acoustic frequency caused changes in the color of the processed image. For example, green optical beams were observed at the ultrasonic frequency $f = 150$ MHz,

Fig. 4. Example of a filtered optical image.

while, at f = 200 MHz, violet beams dominated in the filtered image. When the acoustic frequency was smoothly varied from 100 to 200 MHz, a continuous change of color from red to violet was observed on the screen.

It is known that the maximal number of resolvable spots N of the optical image depends on the angular $\Delta\Theta$ and linear a apertures of the filtering device, as well as on the mean wavelength of the spectral bandwidth [11–14]:

$$N \le a\Delta\Theta/\lambda. \tag{8}$$

It should be noted that formula (8) is valid when the transparency is illuminated by coherent light. The calculation by this formula shows that, for the filter under study with the angular and linear apertures $\Delta\Theta = 6°$ and a = 0.8 cm and for the light wavelength λ = 633 nm, the number of resolvable spots is $N \le 1250$. On the other hand, in the case of incoherent light, the maximal number of resolvable spots will decrease with increasing spectral bandwidth of the filter $\Delta\lambda$ [10, 13]. If the bandwidth of the filtering device is known, one can calculate the angular range of the light beams forming a single spot of the image:

$$\Delta\varphi \approx \frac{(n_e - n_0)\Delta\lambda}{\lambda} \times \sqrt{\sin^4(\Theta_B^e + \alpha) + \sin^2 2(\Theta_B^e + \alpha)}. \tag{9}$$

It was found that, for the type of acousto-optic interaction under study, this angular range is $\Delta\varphi = 3 \times 10^{-4}$. Therefore, the real number of resolvable spots of the filtered image obtained at the output of the acousto-optic filter does not exceed the value $N \le \Delta\varphi = 330$; i.e., it is several times less than in the case of coherent light. In spite of this fact, the filtering device studied in this paper allows one to obtain optical images of a sufficiently good quality.

Figure 4 shows an image obtained with the proposed acousto-optic filter. The quality of the filtered image is reasonably good. It is evident that the resulting picture is largely determined by the spectral bandwidth of the acousto-optic filter. Hence, for the acousto-optic filtering of images, it is expedient to use filters with a narrow spectral bandwidth $\Delta\lambda$. In the case of the high spectral resolution of the filter, the described method of real-time image processing is most attractive for practical use.

SUMMARY

Thus, we have studied a special regime of acousto-optic interaction in a paratellurite crystal. This regime provides the filtering of an arbitrarily polarized light beam. Owing to the use of additional optical elements and a special scheme of the light beam propagation in the crystal, the fulfillment of the Bragg conditions is simultaneously provided for ordinary and extraordinary polarized light beams. This method allows the filtering of an arbitrarily polarized light beam forming the optical image. The spectral resolution obtained in the visible range $R > 300$ proves to be sufficiently high for obtaining good-quality images. The advantage of the proposed method over the known methods is that a single diffracted light beam is observed at the output of the filtering device. Therefore, the filtered image can be recorded with the use of only one detecting device. This fact considerably simplifies both the filtering system as a whole and the subsequent work with the filtered optical images.

REFERENCES

1. V. I. Balakshiĭ, V. N. Parygin, and L. E. Chirkov, *Physical Principles of Acousto-optics* (Radio i Svyaz', Moscow, 1985).
2. A. Korpel, *Acousto-optics* (Marcel Dekker, New York, 1988; Mir, Moscow, 1993).
3. J. Xu and R. Stroud, *Acousto-optics Devices* (Wiley, New York, 1992).
4. I. Chang, Proc. SPIE **90**, 12 (1976).
5. V. B. Voloshinov and O. V. Mironov, Opt. Spektrosk. **68**, 452 (1990) [Opt. Spectrosc. **68**, 264 (1990)].
6. D. Suhre, V. Gottlieb, L. Taylor, *et al.*, Opt. Eng. **31**, 2118 (1992).
7. V. Voloshinov, Ultrasonics **31** (105), 333 (1993).

8. V. Voloshinov, V. Molchanov, and J. Mosquera, Opt. Laser Technol. **28**, 119 (1996).
9. V. Voloshinov, in *Proceedings of I ARL Acousto-optic Tunable Filter Workshop, ARL-SR-54, 1997*, pp. 55–63.
10. L. Denes, B. Kaminsky, M. Gottlieb, *et al.*, in *Proceedings of the I ARL Acousto-optic Tunable Filter Workshop, ARL-SR-54, 1997*, pp. 179–188.
11. L. V. Balakin, V. I. Balakshiĭ, V. B. Voloshinov, *et al.*, Zh. Tekh. Fiz. **61** (10), 100 (1991) [Sov. Phys. Tech. Phys. **36**, 1135 (1991)].
12. I. B. Belikov, V. B. Voloshinov, L. N. Magdich, *et al.*, Pis'ma Zh. Tekh. Fiz. **10** (20), 1225 (1984) [Sov. Tech. Phys. Lett. **10**, 517 (1984)].
13. A. Goutzoulis and D. Pape, *Design and Fabrication of Acousto-optic Devices* (Marcel Dekker, New York, 1994).
14. D. Suhre and E. Villa, Appl. Opt. **37**, 2340 (1998).

Translated by E. Golyamina

Technical Physics, Vol. 45, No. 9, 2000, pp. 1192–1195. Translated from Zhurnal Tekhnicheskoĭ Fiziki, Vol. 70, No. 9, 2000, pp. 106–109.

RADIOPHYSICS

Thermostabilization of Magnetostatic Wave Frequency and Group Velocity in a Cubic Ferrite Film

V. V. Shagaev

Research Institute of Electronic Engineering Materials, Kaluga, 248650 Russia

E-mail: root@niimet.kaluga.su

Received November 16, 1998; in final form, July 26, 1999

Abstract—The conditions for thermostability of the group velocity and frequency of magnetostatic waves in a ferromagnetic layer were studied theoretically. The temperature dependence of the bias field parameters was determined from the conditions for simultaneous thermostability of the frequency and group velocity. Particular emphasis was placed on ferrite films with cubic anisotropy. *© 2000 MAIK "Nauka/Interperiodica".*

INTRODUCTION

Thermostability of magnetostatic wave (MSW) devices based on ferrite films (FF) is a serious problem that should be solved during development of these devices [1]. The temperature drift of MSW device performance is mainly due to the temperature dependence of the magnetic parameters of FF [2, 3]. A method for thermostabilizing the operating frequencies of MSW devices based on the effect of FF demagnetization was developed in [4–6]. In addition, compensating changes in the bias field were used to decrease the temperature drift of the frequency [7, 8].

In this work, we consider a method of simultaneous thermostabilization of the group velocity (GV) and frequency of MSW. The method is based on compensating changes in the bias field intensity and inclination with respect to the film plane. An FF model with cubic uniaxial magnetic anisotropy is used for the theoretical analysis. The analysis is performed for wave vector magnitudes falling within a so-called magnetostatic range [9], which is of particular practical importance.

GEOMETRY OF THE MODEL AND BASIC RELATIONSHIPS FOR AN ANISOTROPIC FILM

We consider a cubic crystal FF with one of the symmetry planes and the axis of uniaxial magnetic anisotropy perpendicular to the film. Let the film be magnetized to saturation by an external magnetic field **H** directed parallel to this plane of symmetry. Thus, all vectors used in the calculations lie in the same plane, which facilitates solution of the problem. It should be noted, however, that a continuous series of crystallographic directions lying in the symmetry planes of the cubic crystal can be considered within the context of this model.

The geometry of the model is shown in Fig. 1. The equilibrium magnetization vector **M**, the field **H**, the crystallographic direction [001], and the normal **n** to the film are in the same crystallographic plane (either (010) or (110)). A right-handed orthogonal coordinate system *xyz* is used. The *z*-axis is directed along the vector **M**; the *y*-axis is parallel to the film. Thus, the angle γ between the *y*-axis and the crystallographic direction [010] is either 0 or $\pi/4$. The angles of inclination of the vectors **M**, **H**, and the crystallographic direction [001] to the plane of the film are denoted as Θ, Θ_H, and δ, respectively.

Let the energy of the FF magnetic anisotropy be described by the expression $[K_c(\beta_1^2\beta_2^2 + \beta_1^2\beta_3^2 + \beta_2^2\beta_3^2) + K_u\sin^2\psi]$, where K_c and K_u are the first constants of cubic and uniaxial anisotropy, respectively; β_1, β_2, and β_3 are the cosines of the angles between the vector **M** and the cubic crystal axes [100], [010], and [001], respectively; and ψ is the angle between **M** and the axis of uniaxial normal anisotropy ($\psi = \pi/2 - \Theta$). The anisotropic properties of FF are taken into account using the tensor of effective demagnetizing factors of anisotropy [9]. Standard calculations give the following expressions for the nonzero tensor components:

$$MN_{xx} = -\frac{3}{8}H_c(4-r)[1-\cos 4(\Theta-\delta)] - H_u\cos^2\Theta, \quad (1a)$$

$$MN_{yy} = -\frac{3}{2}H_c r[1-\cos 2(\Theta-\delta)], \quad (1b)$$

$$MN_{zz} = -\frac{1}{8}H_c\{(4-r)[3+\cos 4(\Theta-\delta)] + 4r\cos 2(\Theta-\delta)\} - H_u\sin^2\Theta, \quad (1c)$$

$$MN_{xz} = \frac{1}{8}H_c[2r\sin 2(\Theta-\delta) + (4-r)\sin 4(\Theta-\delta)] - \frac{1}{2}H_u\sin 2\Theta, \quad (1d)$$

1063-7842/00/4509-1192$20.00

where $H_c = K_c/M$ and $H_u = 2K_u/M$ are the intensities of the cubic and uniaxial anisotropy fields, respectively, and $r = (1/2)(1 - \cos 4\gamma)$.

The temperature drift of the equilibrium orientation of the vector **M** is described by the derivative $d\Theta/dT$, where T is the temperature. The angle Θ is determined from the equation for the external field projection onto the x-axis:

$$H\sin(\Theta_H - \Theta) = MN_{xz} + \frac{1}{2}4\pi M\sin 2\Theta. \qquad (2)$$

Differentiation with respect to T yields

$$\frac{d\Theta}{dT} = \left(\frac{\sin 2(\Theta-\delta)}{4}[r + (4-r)\cos 2(\Theta-\delta)]\right.$$
$$\times\left(H_c\alpha_H - \frac{dH_c}{dT}\right) + \frac{\sin 2\Theta}{2}\left(4\pi M_{\text{eff}}\alpha_H - \frac{d(4\pi M_{\text{eff}})}{dT}\right)$$
$$+ H_z\frac{d\Theta_H}{dT}\Bigg)\Bigg(H_z + \frac{1}{2}H_c[r\cos 2(\Theta-\delta)$$
$$\left. + (4-r)\cos 4(\Theta-\delta)] + 4\pi M_{\text{eff}}\cos 2\Theta\right)^{-1}, \qquad (3)$$

where

$$H_z = H\cos(\Theta_H - \Theta), \qquad (4)$$

$\alpha_H(1/H)(dH/dT)$ is the temperature coefficient of the bias field, $4\pi M_{\text{eff}} = (4\pi M - H_u)$ is the effective magnetization, the derivative dH_c/dT describes the temperature dependence of the cubic anisotropy field, and the derivative $d(4\pi M_{\text{eff}})/dT$ allows the temperature dependences of both the saturation magnetization and the uniaxial anisotropy field to be taken into account.

In the dispersion equations considered below, the magnitude and direction of the external magnetic field are specified by the parameters H_z and Θ. Equations (2) and (4) relate these values to H and Θ_H.

To perform calculations, we need to know the derivative dH_z/dT. Differentiating Eq. (4) and taking into account Eqs. (2) and (1d), we obtain

$$\frac{dH_z}{dT} = \left(\frac{1}{8}H_c[2r\sin 2(\Theta-\delta) + (4-r)\sin 4(\Theta-\delta)]\right.$$
$$\left. + 4\pi M_{\text{eff}}\frac{\sin 2\Theta}{2}\right)\left(\frac{d\Theta}{dT} - \frac{d\Theta_H}{dT}\right) + H_z\alpha_H. \qquad (5)$$

TEMPERATURE COEFFICIENTS OF FREQUENCY AND GROUP VELOCITY

We consider the MSW with the wave vector **k** directed along the y-axis. In the exchangeless approximation, the dispersion equations are derived from the magnetostatic equations, taking into account the electrodynamic boundary conditions at the film surface [10]. Two types of dispersion equations are obtained from the calculations.

The value of the angle Θ sets no limits on the propagation of volume MSW (VMSW) in the film. The dispersion equation for VMSW is

$$kd = \left|\frac{f^2 - f_l^2}{f^2 - f_t^2}\right|^{1/2}\left(L\arctan\left[\frac{2\left|(f^2 - f_l^2)(f^2 - f_t^2)\right|^{1/2}}{f_l^2 + f_t^2 + f_m^2 - 2f^2}\right]\right.$$
$$\left. Lp + \frac{\pi}{2}(1-L) + n\pi\right) \equiv U_{\text{v}}(f, f_l, f_t, f_m), \qquad (6)$$

where

$$f_l^2 = g^2[H_z + M(N_{xx} - N_{zz} - 4\pi\sin^2\Theta) + 4\pi M\cos^2\Theta]$$
$$\times[H_z + M(N_{yy} - N_{zz} - 4\pi\sin^2\Theta)], \qquad (7)$$

$$f_t^2 = g^2[H_z + M(N_{xx} - N_{zz} - 4\pi\sin^2\Theta)]$$
$$\times[H_z + M(N_{yy} - N_{zz} - 4\pi\sin^2\Theta) + 4\pi M], \qquad (8)$$

$$f_m = 4\pi Mg\cos\Theta; \qquad (9)$$

d is the FF thickness; f is the frequency of the nth wave mode ($n = 0, 1, \ldots$); $g = 2.8$ MHz/Oe is the gyromagnetic ratio; if $f_t > f_l$ (forward VMSW), $L = 1$; if $f_t < f_l$ (backward VMSW), $L = -1$; if $2f^2 \le f_l^2 + f_t^2 + f_m^2$, $p = 0$; otherwise, $p = \pi$. The arc tangent function is taken on the interval $(-\pi/2, \pi/2)$, where it is continuous. The frequencies of VMSW fall in the range from f_l to f_t.

If $f_m^2 > |f_t^2 - f_l^2|$, surface MSW (SMSW) can propagate in the film. The dispersion equation for SMSW is

$$kd = \frac{1}{2}\left|\frac{f^2 - f_l^2}{f^2 - f_t^2}\right|^{1/2}$$
$$\times\ln\left(1 + \frac{4[(f^2 - f_l^2)(f^2 - f_t^2)]^{1/2}}{f_m^2 - [(f^2 - f_l^2)^{1/2} + (f^2 - f_t^2)^{1/2}]^2}\right) \qquad (10)$$
$$\equiv U_3(f, f_l, f_t, f_m).$$

The frequencies of SMSW fall within the range $\max\{f_l, f_t\} < f < (1/2f_m)[(f_l^2 + f_t^2 + f_m^2)^2 - 4f_l^2 f_t^2]^{1/2}$. If $f_t > f_l$, the magnitude of the wave vector corresponding to the lower limit of this frequency interval (i.e., $f = f_t$) is $kd = 2(f_t^2 - f_l^2)/(f_m^2 - f_t^2 + f_l^2)$. For a frequency f corresponding to the given magnitude of the wave vector, the temperature coefficient is determined as $\alpha_f = (1/f)(df/dT)$.

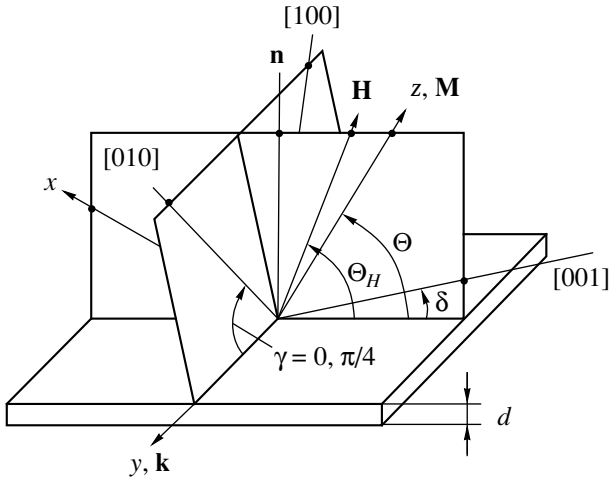


Fig. 1. Model of an anisotropic ferromagnetic film.

The expression for α_f can be derived from the dispersion equation:

$$\alpha_f = -\frac{1}{f^2(\partial U_j/\partial f^2)} \times\left(\frac{\partial U_j}{\partial f_l^2}\frac{d f_l^2}{dT} + \frac{\partial U_j}{\partial f_t^2}\frac{d f_t^2}{dT} + \frac{\partial U_j}{\partial f_m^2}\frac{d f_m^2}{dT}\right), \quad (11)$$

where $j = \nu, s$ for VMSW and SMSW, respectively. For the sake of brevity, the arguments of the functions U_j are omitted. The partial derivatives of U_j are calculated from Eqs. (6) and (10), and the derivatives with respect to temperature, from Eqs. (7)–(9). The derivatives $d(MN_{qq})/dT$ ($q = x, y, z$) are calculated from Eqs. (1a)–(1c) (taking into account dH_c/dT, dH_u/dT, $d\Theta/dT$), whereas $d\Theta/dT$ and dH_z/dT are calculated from Eqs. (3) and (5), respectively. Analytic expressions for α_f are obtained by substituting the obtained expressions into Eq. (11). The parameters α_H and $d\Theta_H/dT$ enter into these analytic expressions as linear functions.

The group velocity is determined from the dispersion relation $f(k)$ by the formula $V_g = 2\pi(\partial f/\partial k)$ [11]. Taking into account that the dispersion equations are given in the form $kd(f)$, we can write the expression for GV:

$$V_g = 2\pi d\left(\frac{\partial(kd)}{\partial f}\right)^{-1}. \quad (12)$$

The temperature coefficient for GV is given by the formula $\alpha_g = (1/V_g)(dV_g/dT)$, which can be easily brought to the form

$$\alpha_g = \alpha_f + f\frac{\partial \alpha_f}{\partial f}. \quad (13)$$

Thus, analytic expressions for α_g can be obtained from the expressions for α_f. It should be noted that the dependences of α_f and α_g on kd are given in parametric form (through the parameter f and the dispersion equation relating kd to f). Computational investigation of the conditions for thermostability was performed using the formulas obtained above.

ANALYSIS OF THE CONDITIONS FOR THERMOSTABILITY

The conditions for simultaneous thermostability of the frequency and GV can be written as a set of two simultaneous equations in two variables α_H and $d\Theta_H/dT$:

$$\alpha_f = 0, \quad \alpha_g = 0. \quad (14)$$

The variables α_H and $d\Theta_H/dT$ enter into the expression for α_f and α_g as linear combinations. Therefore, the given set of simultaneous equations can be solved analytically as a set of two linear equations. However, the expressions thus obtained are rather cumbersome, so that numerical solution is preferable.

As an example, we consider the conditions for thermostability when the frequency and GV of an MSW are specified, whereas the wave vector magnitude can be assigned arbitrarily. Thus, the parameter kd can be regarded as variable, and the field parameters can be sought for as functions of kd. To simplify the analysis of the problem and still obtain some information on the effect of cubic anisotropy, we considered only two crystallographic orientations of the magnetization vector **M**: along the fourth-order symmetry axis and along the third-order symmetry axis (in a cubic crystal, these are the hard and easy magnetic axes, respectively). Conditions (14) are met in these two directions. Thus, according to the model geometry (Fig. 1), the possible values of parameters r and $(\Theta - \delta)$ in Eqs. (1a)–(1d) are (1) for $\mathbf{M} \parallel \langle 100\rangle$, either $r = 0$ and $(\Theta - \delta) = 0, \pi/2, \pi, (3/2)\pi$ or $r = 1$ and $(\Theta - \delta) = 0, \pi$; (2) for $\mathbf{M} \parallel \langle 111\rangle$ $r = 1$ and $(\Theta - \delta) = \arccos(\pm 1/3^{1/2})$, $\arccos(\pm 1/3^{1/2}) + \pi$.

For numerical calculations, the values of the magnetic parameters of yttrium iron garnet ($Y_3Fe_5O_{12}$) films were selected (these films are widely used in MSW devices) [2]: $4\pi M_{\text{eff}} = 1750$ G, $H_c = -42$ Oe, $d(4\pi M_{\text{eff}})/dT = -4.15$ G/K, $dH_c/dT = 0.46$ Oe/K (here and in Fig. 2, K is degrees Kelvin).

The fundamental mode of forward VMSW was considered, and the frequency and GV were taken to be $f = 10^4$ MHz and $V_g/(2\pi d) = 6 \times 10^8$ s^{-1}. First, the dependences $\Theta(kd)$ and $H_z(kd)$ were calculated. For this purpose, the set of two simultaneous equations $U_\nu = kd$ and $(\partial U_\nu/\partial f)^{-1} = V_g/(2\pi d)$ in two variables Θ and H_z and with the variable parameter kd was solved. The function U_ν was determined by Eq. (6) for the fundamental mode. The expressions (7)–(9) for f_l, f_t, and f_m were substituted into the equations. The dependences $\Theta(kd)$ and

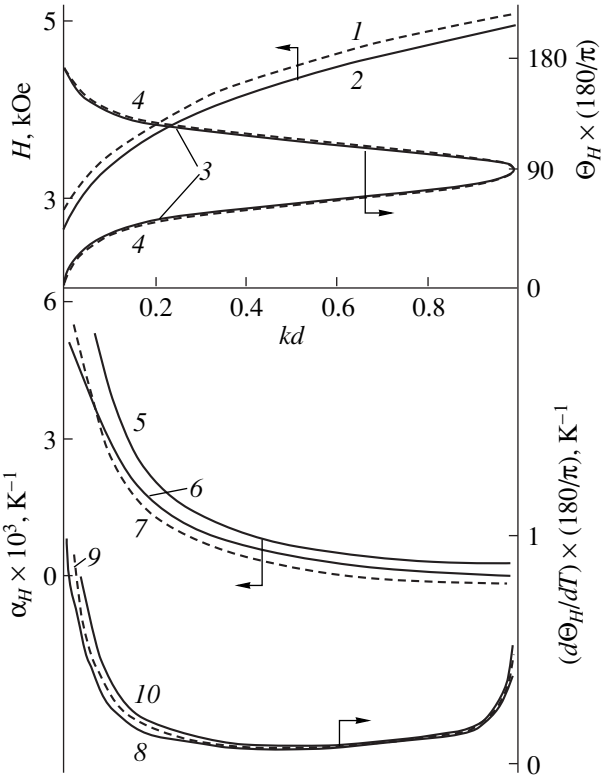


Fig. 2. Dependence of the external magnetic field parameters on *kd* for given values of the frequency and group velocity of the fundamental mode of the forward volume MSW (solid lines, **M** || ⟨111⟩; dashed lines **M** || ⟨100⟩).

$H_z(kd)$ thus obtained were further substituted into Eqs. (14) for calculating the parameters Θ_H and $d\Theta_H/dT$ and into Eqs. (2) and (4) for calculating the parameters H and Θ_H.

The results of the calculations are shown in Fig. 2. On the interval $0 \le \Theta_H \le \pi$, the function $\Theta_H(kd)$ has two branches symmetric about the straight line $\Theta_H = \pi/2$ (curve *4* for **M** || ⟨100⟩ and curve *3* for **M** || ⟨111⟩). It should be noted that a single curve $H(kd)$ corresponds to each pair of symmetric branches $\Theta_H(kd)$ (curve *1* for **M** || ⟨100⟩ and curve *2* for **M** || ⟨111⟩). In addition, the curves $\alpha_H(kd)$ for **M** || ⟨1000⟩ (curve *7*) corresponding to symmetric curves $\Theta_H(kd)$ are identical, whereas the curves $\alpha_H(kd)$ differ at both $d\Theta_H(kd) \le \pi/2$ (curves *5* and *10*, respectively) and $\pi/2 < \Theta_H(kd) < \pi$ (curves *6* and *8*, respectively; curve *8* corresponds to the function $(-1)d\Theta_H(kd)/dT$).

Thus, calculations based on actual values of the magnetic parameters of FF showed that simultaneous thermostabilization of the frequency and GV of an MSW by the proposed method is theoretically possible. The equations obtained in this work allow various modifications of the method for thermostabilization of MSW characteristics to be analyzed and the most effective and practicable of these to be selected. It should be noted that these equations were derived under the assumption of relatively large anisotropy fields. This makes the model described in this work applicable to cubic crystal films regardless of the values of their magnetic parameters.

REFERENCES

1. W. S. Iskhak, Proc. IEEE **76**, 171 (1988).
2. Yu. M. Yakovlev and S. Sh. Gendelev, *Single-Crystal Ferrites in Radio Electronics* (Sov. Radio, Moscow, 1975).
3. V. V. Shagaev, Fiz. Tverd. Tela (St. Petersburg) **40**, 2089 (1998) [Phys. Solid State **40**, 1892 (1998)].
4. A. S. Beregov, E. V. Kudinov, and I. N. Ereshchenko, Élektron. Tekh., Ser. 1: Élektronika SVCh **1** (395), 19 (1987).
5. Yu. K. Fetisov, Zh. Tekh. Fiz. **57** (12), 2393 (1987) [Sov. Phys. Tech. Phys. **32**, 1451 (1987)].
6. A. N. Slavin and Yu. K. Fetisov, Zh. Tekh. Fiz. **58**, 2210 (1988) [Sov. Phys. Tech. Phys. **33**, 1343 (1988)].
7. J. D. Adam, IEEE MTT-S Int. Microwave Symp. Dig., 160 (1979).
8. J. D. Adam and M. R. Daniel, IEEE Trans. Magn. **17**, 2951 (1981).
9. A. G. Gurevich, *Magnetic Resonance in Ferrites and Antiferromagnetics* (Nauka, Moscow, 1973).
10. A. S. Beregov, Izv. Vyssh. Uchebn. Zaved., Radioélektron. **27** (10), 9 (1984).
11. A. S. Beregov, Radiotekh. Élektron. (Moscow) **28** (1), 127 (1983).

Translated by K. Chamorovskiĭ

Technical Physics, Vol. 45, No. 9, 2000, pp. 1196–1200. Translated from Zhurnal Tekhnicheskoĭ Fiziki, Vol. 70, No. 9, 2000, pp. 110–113.

A Self-Consistent Method for Describing Curvilinear Charged Beams

N. D. Naumov

Received October 12, 1999

Abstract—A self-similar solution to the self-consistent gasdynamics equations is used to derive envelope equations for a thin curvilinear charged beam propagating in an external electromagnetic field. © *2000 MAIK "Nauka/Interperiodica".*

INTRODUCTION

The necessity of developing self-consistent models of charged-particle beam motion in external electromagnetic fields arises in connection with practical problems of the formation and transportation of beams and electron rings. One method that allows for the effect of a self-induced field on the transverse beam dynamics is based on envelope equations. These equations are known for straight and ring-shaped charged-particle beams [1–3]. In this paper, we derive the general envelope equations for curvilinear charged-particle beams subjected to an external electromagnetic field. A particular example of this beam configuration is an electron beam injected at a nonzero angle with respect to the geomagnetic field. This problem is of practical interest, because electron beams are used to study the ionosphere [4].

Here, we will construct a model of a steady-state low-current beam, i.e., a beam whose current is much less than the Alfvén limiting current. In this case, the particle motion is relativistic only in the longitudinal direction. The macroscopic beam description uses the stationary Euler equation for a charged fluid in an external electromagnetic field [5]

$$(\mathbf{V}\nabla)\gamma\mathbf{V} + \frac{1}{mn}\nabla p = \frac{e}{m}\left(\mathbf{E} + \mathbf{E}_{\exp} + \frac{1}{c}[\mathbf{V}(\mathbf{B} + \mathbf{B}_{\exp})]\right). \quad (1)$$

Here, e and m are the particle charge and mass, respectively; $\gamma = (1 - u^2/c^2)^{-1/2}$, where u is the longitudinal beam velocity; n is the particle number density; p is the fluid pressure due to the beam emittance; and $\mathbf{E}$ and $\mathbf{B}$ are the self-induced electric and magnetic fields, respectively. In the steady state, one can neglect the effect of the collective field on the longitudinal particle motion and allow for its effect on the transverse motion alone. This problem can be solved analytically only for a thin beam where the ratios of the maximum cross-sectional dimension to the beam curvature radius and to its torsion radius are small. In this case, an approximate solution to the Euler equation (1), accurate to first-order terms in a small parameter, exists. This solution belongs to the class of self-similar solutions to the gasdynamics equations that describe charged fluid motion with particle velocities proportional to the distance from the center of symmetry [6].

CURVILINEAR STRIP BEAM

First, we illustrate the method by a relatively simple problem of the propagation of a cold strip beam of nonrelativistic charged particles in crossed electric and magnetic fields $\mathbf{E}_0 = E_0\mathbf{e}_y$ and $\mathbf{B}_0 = B_0\mathbf{e}_y$. In this case, the trajectory $\mathbf{Y}(s)$ of the axial particle is a trochoid.

It is convenient to solve the problem in terms of the curvilinear coordinates s, q, and z:

$$\mathbf{x} = \mathbf{Y}(s) + q\mathbf{n} + z\mathbf{e}_z,$$

where s is the length of the trajectory measured from the place of the beam injection and $\mathbf{t}$, $\mathbf{n}$, and $\mathbf{e}_z$ are the Frénet trihedral vectors connected with curve $\mathbf{Y}(s)$.

Substitution of the expression for the velocity $\mathbf{v} = u\mathbf{t}$ into the particle motion equation in the orthogonal fields yields the following formula for the trochoid curvature: $k_1 = k - eB_0/mcu$, where $k = eE_{01}/mu^2$ and E_{01} is the external electric field component along the vector $\mathbf{n}$, i.e.,

$$\mathbf{E}_0 = E_{01}\mathbf{n} + E_{02}\mathbf{t}.$$

The E_{02} component determines the variation in the axial particle velocity during beam propagation: $mu\dot{u} = eE_{02}$; here and below, the dot symbol denotes differentiation with respect to s.

Writing Eq. (1) in curvilinear coordinates yields the following equations for the components of the gasdynamic beam velocity $\mathbf{V} = U\mathbf{t} + W\mathbf{n}$:

$$\frac{U}{\sigma}\left(\frac{\partial U}{\partial s} - k_1 W\right) + W\frac{\partial U}{\partial q} = \frac{e}{m}E_{02} - kuW, \quad (2)$$

1063-7842/00/4509-1196$20.00

$$\frac{U}{\sigma}\frac{\partial W}{\partial s} + W\frac{\partial W}{\partial q} + \frac{k_1}{\sigma}U^2 = \frac{e}{m}(E + E_{01}) - kuU, \quad (3)$$

where $\sigma = 1 - k_1 q$.

For the class of charged fluid motion under consideration, the particle velocity is proportional to the distance from the center of symmetry; i.e., the beam velocity components should be sought for in the form

$$U = u(1 + \Gamma\xi), \quad W = u\dot{a}\xi. \quad (4)$$

A similar representation can be used for the particle number density:

$$n = n_0\frac{a_0 u_0}{au}(1 + \nu\xi)\vartheta(1 - \xi^2), \quad (5)$$

where $\vartheta(x)$ is the Heaviside step function and the constants n_0, a_0, and u_0 are determined by the beam injection conditions. Here, $\xi = q/a$ is the self-similar variable, $2a$ is the cross-sectional beam dimension, and Γ and ν are certain functions of s. Within the accuracy of our approach, the collective field can be assumed to be equal to the electric field of a uniform beam $E = 4\pi enq$. Corrections that would allow for the beam curvature and particle number density are terms of the second order of smallness.

We substitute expressions (4) into Eqs. (2) and (3) and expressions (5) into the continuity equation $\mathrm{div} n\mathbf{V} = 0$. As result, we obtain the following equations for functions a, Γ, and ν accurate to first-order terms:

$$\ddot{a} + \frac{\dot{u}}{u}\dot{a} + k_1^2 a + (k_1 + k)\Gamma = \omega^2\frac{u_0}{u^3}a_0, \quad (6)$$

$$\dot{\Gamma} + 2\frac{\dot{u}}{u}\Gamma = k\dot{a} - k_1 a\frac{\dot{u}}{u}, \quad (7)$$

$$\dot{\nu} + \dot{\Gamma} - 2\frac{\dot{a}}{a}\Gamma = 2k_1\dot{a}, \quad (8)$$

where $\omega = \sqrt{4\pi n_0 e^2/m}$ is the plasma frequency of the beam.

This result illustrates the familiar advantage of the approach based on the self-similarity principle, which reduces the problem from solving a system of partial differential equations to a simpler problem of integrating a system of ordinary differential equations.

Equation (7) can be formally integrated in the general form

$$\Gamma = \frac{1}{u^2}\left[\Gamma_0 u_0^2 + \int_0^s (k\dot{a}u^2 - k_1 auu)ds\right],$$

where the initial characteristics u_0 and Γ_0 of the longitudinal beam velocity are governed by the beam injection conditions. Therefore, in this case, the cross-sectional beam dimension can be found as a solution to a single integro-differential equation.

EXPANSION OF THE EXTERNAL FIELD

Before constructing an approximate solution to the Euler equation (1) in the general case, let us analyze the structure of the external electromagnetic field near the beam axis. The beam axis, defined as the trajectory of the axial particle $\mathbf{Y}(s)$, is determined as a solution to the equation of motion of a single particle in the given external field. We assume that this solution is known.

It is convenient to consider the beam motion in terms of the curvilinear coordinates q_1, q_2, and s:

$$\mathbf{x} = \mathbf{Y}(s) + q_1\mathbf{n} + q_2\mathbf{b}.$$

Here, $\mathbf{t}$, $\mathbf{n}$, and $\mathbf{b}$ are the Frénet trihedral vectors connected with the curve $\mathbf{Y}(s)$. Let us consider the expansion of the external field in the Taylor series about this curve. If a vector field $\mathbf{A}$ is represented as

$$\mathbf{A} = A_1\mathbf{n} + A_2\mathbf{b} + A_3\mathbf{t},$$

then, in the curvilinear coordinates introduced above, the conditions $\mathrm{div}\mathbf{A} = 0$ and $\mathrm{curl}\mathbf{A} = 0$ take the form

$$k_1 A_1 - \dot{A}_3 - LA_3 = \sigma\left(\frac{\partial A_1}{\partial q_1} + \frac{\partial A_2}{\partial q_2}\right), \quad (9)$$

$$\frac{\partial A_1}{\partial q_2} = \frac{\partial A_2}{\partial q_1}, \quad (10)$$

$$LA_1 + \dot{A}_1 - k_2 A_2 + k_1 A_3 = \sigma\frac{\partial A_3}{\partial q_1}, \quad (11)$$

$$k_2 A_1 + LA_2 + \dot{A}_2 = \sigma\frac{\partial A_3}{\partial q_2}, \quad (12)$$

where $\sigma = 1 - k_1 q_1$; k_1 and k_2 are the curvature and torsion of curve $\mathbf{Y}(s)$, respectively; and L stands for the operator

$$L = k_2\left(q_2\frac{\partial}{\partial q_1} - q_1\frac{\partial}{\partial q_2}\right).$$

Conditions (9) and (10) reveal the structure of the expansion of functions A_i about the curve $\mathbf{Y}(s)$:

$$A_1 = C_1 + a_1 q_1 + a_2 q_2 + \alpha_1 q_1^2 + 2\alpha_2 q_1 q_2 + \alpha_3 q_2^2,$$

$$A_2 = C_2 + a_2 q_1 - (a_1 + \dot{C}_3 - k_1 C_1)q_2 + \alpha_2 q_1^2 + 2\alpha_3 q_1 q_2 + \alpha_4 q_2^2,$$

$$A_3 = C_3 + a_3 q_1 + a_4 q_2 + \beta_1 q_1^2 + 2\beta_2 q_1 q_2 + \beta_3 q_2^2,$$

where C_i, a_i, α_i, and β_i are certain functions of s.

In order to solve the problem, it is sufficient to find the functions a_i. When substituted into conditions (9)–(12), these expansions yield the functions a_3 and a_4 in the general case:

$$a_3 = \dot{C}_1 + k_1 C_3 - k_2 C_2, \quad a_4 = \dot{C}_2 - k_2 C_1.$$

Therefore, accurate to first-order terms, the expansion of the external field about the beam axis has the form

$$B_1 = B_{01} + f_1 q_1 + f_2 q_2,$$

$$B_2 = B_{02} + f_2 q_1 - (f_0 + f_1) q_2, \quad f_0 = \dot{B}_{03} - k_1 B_{01},$$

$$B_3 = B_{03} + (\dot{B}_{01} + k_1 B_{03} - k_2 B_{02}) q_1 + (\dot{B}_{02} - k_2 B_{01}) q_2,$$

$$E_1 = E_{01} + g_1 q_1 + g_2 q_2,$$

$$E_2 = E_{02} + g_2 q_1 - (g_0 + g_1) q_2,$$

$$g_0 = \dot{E}_{03} - k_1 E_{01},$$

$$E_3 = E_{03} + g_3 q_1 + g_4 q_2, \quad g_3 = \dot{E}_{01} + k_1 E_{03} - k_2 E_{02},$$

$$g_4 = \dot{E}_{02} - k_2 E_{01}.$$

Functions f_1, f_2, g_1, and g_2 are determined by the particular external field configuration. For example, for a ring-shaped beam in a modified betatron, we have $f_1 = nk_1 B_{02}$ and $f_2 = 0$, and for a double spiral stellatron, we have $f_1 = k_1 C_0 \sin\chi$ and $f_2 = k_1 C_0 \cos\chi$. Here, n is the external field index; C_0 is a constant; and $\chi = lk_1 s$, where l is an integer.

SOLUTION OF THE GASDYNAMICS EQUATIONS

Substituting the expression for the gasdynamic beam velocity

$$\mathbf{V} = V_1 \mathbf{n} + V_2 \mathbf{b} + U\mathbf{t}$$

into Eq. (1) leads to the following equations for the functions V_i and U:

$$M\gamma V_1 + \frac{\gamma}{\sigma} U(k_1 U - k_2 V_2) = \frac{e}{m} E_1 + \frac{e}{mc}(V_2 B_3 - U B_2) + F_n, \tag{13}$$

$$M\gamma V_2 + \frac{\gamma}{\sigma} k_2 U V_1 = \frac{e}{m} E_2 + \frac{e}{mc}(U B_1 - V_1 B_3) + F_b, \tag{14}$$

$$M\gamma U - \frac{\gamma}{\sigma} k_1 U V_1 = \frac{e}{m} E_3 + \frac{e}{mc}(V_1 B_2 - V_2 B_1), \tag{15}$$

where the terms F_n and F_b are due to the self-induced field and emittance of the beam and

$$M = V_i \frac{\partial U}{\partial q_i} + \frac{U}{\sigma}\left(\frac{\partial}{\partial s} + L\right).$$

As for a strip beam, the collective field can be approximated by the electromagnetic field of a uniform beam. In the general case, we should consider a beam with an elliptical cross section. As a result, in the transverse coordinates x_i related to the axes of symmetry of the beam cross section, F_i take the form [7]

$$F_1 = x_1 u^2 \left[\frac{h}{a(a+b)} + \frac{H}{a^4}\right],$$

$$F_2 = x_2 u^2 \left[\frac{h}{b(a+b)} + \frac{H}{b^4}\right].$$

Here, $h = 4|e|I/mu^3\gamma^3$ and $H = E/u$, where I, E, a, and b are the beam current, emittance, and cross-section semiaxes, respectively. As the beam propagates, the orientation of its cross section changes; hence, the coordinate axes x_i rotate with respect to the unit vectors $\mathbf{n}$ and $\mathbf{b}$ by a certain angle ψ:

$$x_1 = q_1 \cos\psi + q_2 \sin\psi,$$

$$x_2 = q_2 \cos\psi - q_1 \sin\psi.$$

Accordingly, V_i should be represented in terms of the gas velocity components Λ_i in the new coordinate system as

$$V_1 = \Lambda_1 \cos\psi - \Lambda_2 \sin\psi - q_2\Omega,$$

$$V_2 = \Lambda_2 \cos\psi + \Lambda_1 \sin\psi + q_1\Omega,$$

where $\Omega = u\dot{\psi}$ is the angular velocity of the beam rotation as a whole with respect to the Frénet trihedral.

As the beam propagates in the external field, not only do its cross-sectional dimensions change, but an internal transverse rotational motion of the beam in coordinates x_i may also occur. Therefore, one should use the following expressions for Λ_i in terms of self-similar variables $\chi = x_1/a$ and $\eta = x_2/b$:

$$\Lambda_1 = u(\dot{a}\xi - \kappa a\eta), \quad \Lambda_2 = u(\dot{b}\eta + \kappa b\xi). \tag{16}$$

Here, κ is a function of s, which characterizes the charged-particle motion along the elliptic stream lines within the beam. The self-similar expressions for the longitudinal beam velocity and particle number density have the form

$$U = u(1 + \Gamma_1\xi + \Gamma_2\eta), \tag{17}$$

$$n = \frac{I}{\pi abu|e|}(1 + \nu_1\xi + \nu_2\eta)\vartheta(1 - \xi^2 - \eta^2), \tag{18}$$

where Γ_i and ν_i are certain functions of s.

BEAM ENVELOPE EQUATIONS

Substituting the expression for the velocity $\mathbf{v} = u\mathbf{t}$ into the relativistic equation of motion of a charged particle in an external electromagnetic field leads to the following formula for the curvature of the trajectory $\mathbf{Y}(s)$: $k_1 = \tau_1 E_{01} + \tau_2 B_{02}$, where E_{01} and B_{02} are the external field components at the beam axis, $\tau_1 = e/\gamma mu^2$, and $\tau_2 = -u\tau_1/c$. During beam propagation, the component E_{03} determines the variation in the axial particle veloc-

ity ($mu(\gamma\dot{u} + u\dot{\gamma}) = eE_{03}$) and the relationship $E_{02} = -uB_{01}/c$ is kept.

If we substitute expressions (16) and (17) into Eqs. (13)–(15) and take into account only the first-order terms, then we obtain the following system of ordinary differential equations for the functions a, b, ψ, κ, and Γ_i, which govern the variations in the dimensions and velocity of the beam as it propagates in the external field:

$$\ddot{a} + D\dot{a} + \left(\lambda_1 + \kappa k_2\frac{a}{b} + k_1^2\cos^2\psi + G_0\sin^2\psi\right)a = \kappa k_3 b - \mu_1\Gamma_1 + \frac{h}{a+b} + \frac{H}{a^3}, \tag{19}$$

$$\ddot{b} + D\dot{b} + \left(\lambda_1 + \kappa k_2\frac{b}{a} + k_1^2\sin^2\psi - G_0\cos^2\psi\right)b = \kappa k_3 a - \mu_2\Gamma_2 + \frac{h}{a+b} + \frac{H}{b^3}, \tag{20}$$

$$2ab(\ddot{\psi} + D\dot{\psi}) + (b\dot{a} + a\dot{b})(2\dot{\psi} - \kappa_3) + \frac{1}{\gamma u}\times\frac{d}{ds}\gamma\kappa u(a^2 + b^2) = \mu_1 a\Gamma_2 - \mu_2 b\Gamma_1 + abF_0, \tag{21}$$

$$\frac{1}{\gamma u}\frac{d}{ds}\gamma\kappa u(a^2 - b^2) - (k_3 + k_2)(b\dot{a} - a\dot{b}) + k_1 ab\sin 2\psi = \lambda_2 ab + \mu_1\Gamma_2 a + \mu_2\Gamma_1 b, \tag{22}$$

$$\dot{\Gamma}_1 + D(2\Gamma_1 + k_1 a\cos\psi) + \left(\kappa - k_2\frac{a}{b}\right)\Gamma_2 = \mu_3\dot{a} + \mu_4(\kappa b + a\dot{\psi}) + \tau_1(g_3\cos\psi + g_4\sin\psi)a, \tag{23}$$

$$\dot{\Gamma}_2 + D(2\Gamma_2 - k_1 b\sin\psi) - \left(\kappa - k_2\frac{b}{a}\right)\Gamma_1 = \mu_4\dot{b} - \mu_3(\kappa a + b\dot{\psi}) + \tau_1(g_4\cos\psi - g_3\sin\psi). \tag{24}$$

Here, we use the following notation:

$$D = \frac{\dot{\gamma}}{\gamma} + \frac{\dot{u}}{u}, \quad \kappa_i = \tau_2 B_{0i}, \quad k_3 = k_2 - \kappa_3 + 2\dot{\psi},$$

$$\lambda_1 = \lambda + G_1\cos 2\psi + \left(G_2 - \frac{1}{2}F_0\right)\sin 2\psi,$$

$$\lambda = \dot{\psi}(\kappa_3 - \dot{\psi}) - \kappa^2,$$

$$\lambda_2 = (G_0 + 2G_1)\sin 2\psi + (F_0 - 2G_2)\cos 2\psi,$$

$$F_0 = \tau_2 f_0, \quad G_0 = \tau_1 g_0,$$

$$G_1 = \tau_1 g_1 + \tau_2 f_2, \quad G_2 = \tau_1 g_2 - \tau_2 f_1,$$

$$\mu_1 = \mu_3 + k_1\cos\psi, \quad \mu_2 = \mu_4 - k_1\sin\psi,$$

$$\mu_3 = k\cos\psi + \kappa_1\sin\psi, \quad \mu_4 = \kappa_1\cos\psi - k\sin\psi,$$

$$k = \tau_1 E_{01}.$$

Equations for the functions ν_i, which characterize the variation in the degree of inhomogeneity of the particle number density as the beam propagates in the external field, can be derived by substituting expression (18) into the continuity equation; for the sake of brevity, these equations are omitted.

When there is no external electric field, Eqs. (23) and (24) have the trivial solution $\Gamma_i = 0$; i.e., accurate to terms of the first order of smallness, a possible solution to Eq. (15) has the form $U = u$, where u is the longitudinal beam velocity. Therefore, the equations for the envelope of a monoenergetic beam in an inhomogeneous magnetic field have a simpler form.

A further possibility of simplifying the problem occurs when the drift approximation can be applied to a beam propagating in a weakly inhomogeneous magnetic field. In this case, the contribution from the terms produced by the gradient of the external field is of the second order of smallness. Indeed, the drift approximation condition $|\nabla\mathbf{B}_{\exp}|/k_1 B_{\exp} \sim \varepsilon$ implies that $r|\nabla\mathbf{B}_{\exp}| \sim \varepsilon k_1 r B_{\exp}$, where r is the characteristic cross-sectional beam dimension. Therefore, the equations for the envelope of a monoenergetic beam in a weakly inhomogeneous field are much simpler than system (19)–(24):

$$\ddot{a} + \left(\lambda + \kappa k_2\frac{a}{b} + k_1^2\cos^2\psi\right)a = \kappa k_3 b + \frac{h}{a+b} + \frac{H}{a^3}, \tag{25}$$

$$\ddot{b} + \left(\lambda + \kappa k_2\frac{b}{a} + k_1^2\sin^2\psi\right)b = \kappa k_3 a + \frac{h}{a+b} + \frac{H}{b^3}, \tag{26}$$

$$2ab\ddot{\psi} + (b\dot{a} + a\dot{b})(2\dot{\psi} - \kappa_3) + \frac{d}{ds}\kappa(a^2 + b^2) = 0, \tag{27}$$

$$\frac{d}{ds}\kappa(a^2 - b^2) - (k_3 + k_2)(b\dot{a} - a\dot{b}) + k_1^2 ab\sin 2\psi = 0. \tag{28}$$

When the magnetic field is uniform, the beam axis is a helical line whose curvature and torsion are independent of s: $k_1 = |k_0|\sin\alpha$ and $k_2 = k_0\cos\alpha = \kappa_3$, where $k_0 = \tau_2 B_0$ and α is the angle between the magnetic field and the direction of the beam injection. Equation (27) provides a relationship between the rotation of the helical beam as a whole with respect to the Frénet trihedral and the internal motion of the charged fluid:

$$\dot{\psi} = \frac{1}{2ab}\left[2a_0 b_0\left(\dot{\psi}_0 - \frac{k_2}{2}\right) + \kappa_0(a_0^2 + b_0^2) - \kappa(a^2 + b^2)\right] + \frac{k_2}{2}.$$

The envelope equations can be used to estimate the characteristics of a beam injected into an external elec-

tromagnetic field. The results obtained above remain applicable in the region where the cross-sectional beam dimensions increase insignificantly due to the effect of the space charge.

REFERENCES

1. I. M. Kapchinskiĭ, *Particle Dynamics in Linear Resonant Accelerators* (Atomizdat, Moscow, 1966).
2. E. P. Lee and R. K. Cooper, Part. Accel. **7**, 83 (1976).
3. V. P. Sarantsev and É. A. Perel'shteĭn, *Collective Acceleration of Ions by Electron Rings* (Atomizdat, Moscow, 1979).
4. *Artificial Particle Beams in Space Plasma Studies*, Ed. by B. Grandal (Plenum, New York, 1982; Mir, Moscow, 1985).
5. R. C. Davidson, *Theory of Nonneutral Plasmas* (Benjamin, New York, 1974; Mir, Moscow, 1978).
6. N. D. Naumov, Zh. Tekh. Fiz. **67** (7), 103 (1997) [Tech. Phys. **42**, 814 (1997)].
7. J. D. Lawson, *The Physics of Charged-Particle Beams* (Clarendon, Oxford, 1977; Mir, Moscow, 1980).

Translated by A. Khzmalyan

Technical Physics, Vol. 45, No. 9, 2000, pp. 1201–1206. Translated from Zhurnal Tekhnicheskoĭ Fiziki, Vol. 70, No. 9, 2000, pp. 114–118.

Sputtering of A^3B^5 Materials (GaP, GaAs, GaSb, InP, and InSb) by 2- to 14-keV N_2^+ Ions

I. P. Soshnikov and N. A. Bert

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

E-mail: ipsosh@beam.ioffe.rssi.ru

Received March 11, 1998; in final form, September 13, 1999.

Abstract—Investigations of the general characteristics and distinctive features of sputtering of A^3B^5 materials (GaP, GaAs, GaSb, InP and InSb) under bombardment with N_2^+ ions have been carried out. From the experimental data, dependences of the sputtering yield of these materials on the incidence angle and ion energy have been obtained and the surface relief patterns produced by target etching have been studied. It has been shown that the dependence on energy of the sputtering yield for GaP, GaAs, and InP can be adequately described by the Haffa–Switkovski formula for binary materials and Yudin's approximation for elemental targets. Sputtering of GaSb and InSb proceeds in the surface layer recrystallization mode, and the sputtering yield agrees with calculations based on Onderlinden's model. From a comparison of the experimental and calculated dependences, the surface bonding energies have been determined. © *2000 MAIK "Nauka/Interperiodica".*

The processes involved in irradiation of A^3B^5 materials (GaP, GaAs, GaSb, InP, and InSb) with nitrogen ions, and ion sputtering in particular, attract considerable attention of researchers [1], which is explained by the prospects of employing them in various technologies. Special attention is paid to modifying the properties and structure of the irradiated targets and the synthesis of new materials [1–4]. At the same time, little is known so far about the dependence of the sputtering yield of A^3B^5 materials on the incidence angle and energy of nitrogen ions and surface relief patterns, which are indispensable for understanding the distinctive features of the processes involved and important for many applications (for example, preparing samples for transmission microscopy and forming topological patterns) [5].

In this work, we present the results of investigations of the dependence of the sputtering yield of GaP, GaAs, GaSb, InP, and InSb on energy (E) in the 2- to 14-keV range and the incidence angle of N_2^+ ions, as well as the surface relief patterns formed as a result of sputtering.

For the experiments, single crystal samples of area $S \sim 0.3$–1.5 cm^2 were cut from standard wafers of gallium phosphide, gallium arsenide, gallium antimonide indium phosphide with a (100) surface orientation, and indium antimonide with a (111) surface orientation. After degreasing and removal of the defect layer, the samples were placed in a monochromatic beam of N_2^+ ions of energy from 1 to 15 keV and density $j \sim 100$–250 μA/cm^2 and irradiated to fluences of $\Phi \sim 10^{18}$–10^{19} ion/cm^2. It should be noted that compared to the installation described earlier [6, 7] the ion sputtering installation in this study was supplemented with a Wien filter at the output of the ion source, which prevented other ions (N^+, O_2^+) from striking the sample surface. The temperature of the control samples, measured with a thermocouple immediately after turning off the ion beam, was usually found to be in the interval 50–70°C.

The sputtering yield Y was determined by measuring the thickness of the sputtered layer [7–9]. For measurements of the sputtered layer thickness, part of the sample was masked prior to sputtering. After exposure, the mask was removed and the step of height h that was produced at the boundary between the masked and irradiated regions was measured using electron microscopy and interference optical microscopy. The relation between the sputtered layer thickness h and the sputtering coefficient Y can be fitted with expression

$$Y = \frac{2\rho N_A}{eM}\frac{h}{\tau j \cos\Theta}, \quad (1)$$

where j is the ion current density measured with a Faraday cup, τ is the time of exposure to the ion beam, Θ is the ion incidence angle measured relative to the normal to the target surface, ρ is the density of the target material, M is the molar mass of the target, N_A is Avogadro's number, and e is the elementary charge.

1063-7842/00/4509-1201$20.00

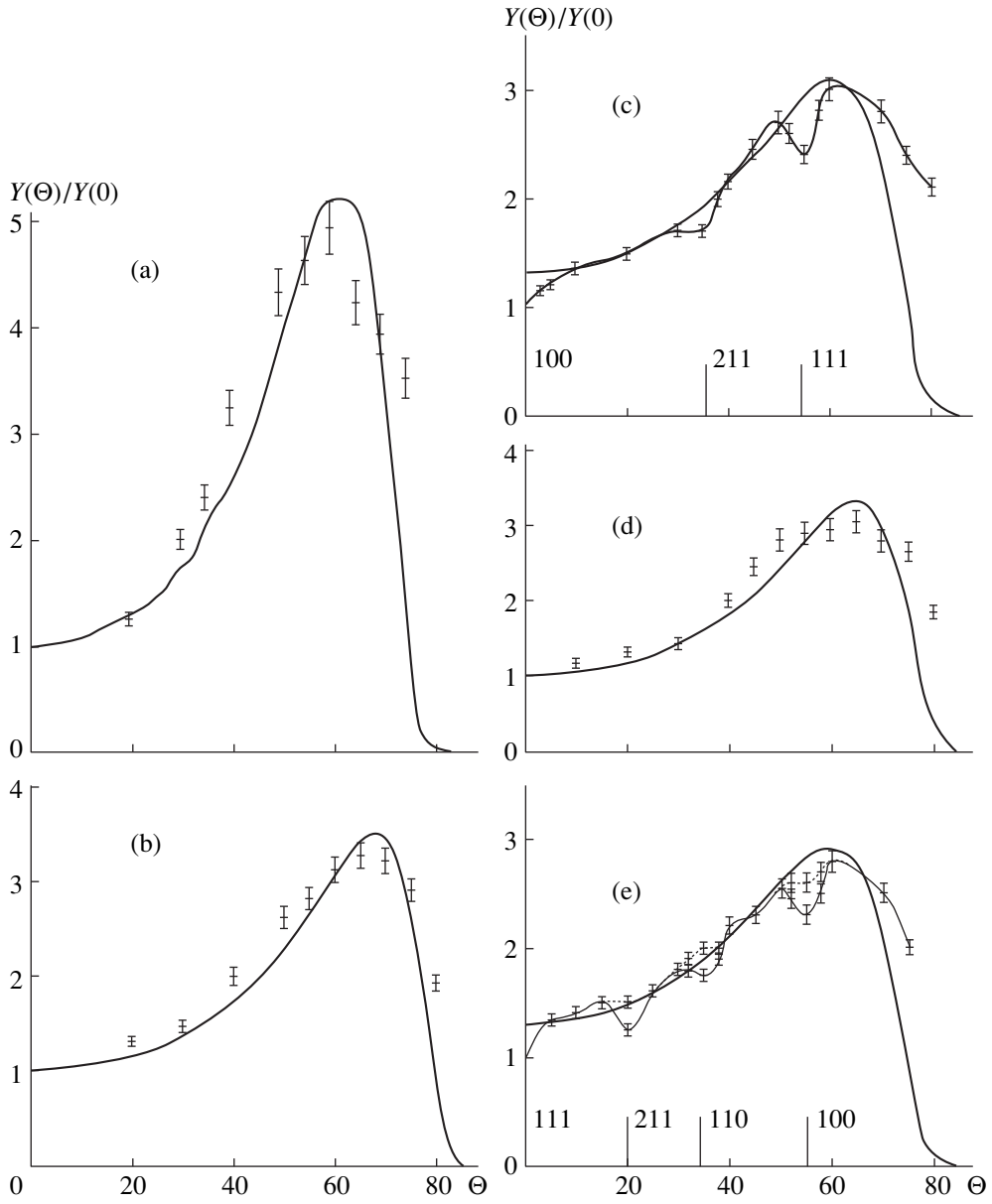


Fig. 1. Angular dependence of sputtering yield under bombardment with 5-keV molecular nitrogen ions: (a) gallium phosphide; (b) gallium arsenide; (c) gallium antimonide; (d) indium phosphide; (e) indium antimonide.

Note that the incidence angle was changed by tilting the sample about the (110) crystallographic axis (Fig. 1).

The surface morphology of the samples was studied using scanning electron microscopy methods [10].

The results of measuring the angular dependences of the sputtering yield of the materials studied are presented in Fig. 2. For combinations target–GaP ion $\leftarrow N_2^+$, GaAs $\leftarrow N_2^+$, and InP $\leftarrow N_2^+$, the dependences have the shape typical of amorphous targets, with one peak at 65–75° relative to the normal to the surface. In the range of angles $\Theta \le 75°$, the dependences are closely approximated by curves calculated by Yamamura's formula [11]

$$Y(\Theta)/Y(0) = \cos^{-f}\Theta \exp[g(1-\cos^{-1}\Theta)], \quad (2)$$

where f and g are parameters defined by the expressions

$$f = (1-\cos\Theta_{max})^{-1}, \quad g = p(1-\cos\Theta_{max})^{-1};$$

Θ_{max} is the maximum angle of sputtered particles; and p is an exponent in the approximate expression $Y \sim \cos^p\Theta$ for the interval close to the normal to the surface.

Values obtained from a comparison of theory and experiment are given in Table 1.

$Y = f(\Theta)$ curves for the combinations GaSb ← N_2^+ and InSb ← N_2^+ have deviations whose minimums lie in the range of ion incidence angles corresponding to low-index crystallographic directions [100], [211], and [111]. This type of angular dependence indicates that the surface being sputtered is monocrystalline; i.e., the recrystallization temperature of the systems GaSb ← N_2^+ and InSb ← N_2^+ is lower than the minimum control temperature $T = 50°C$. From a comparison of sputtering patterns for bombardments with argon ions [5, 7, 9, 12] and molecular nitrogen, it is seen that the recrystallization temperature of gallium antimonide and indium antimonide is lower for bombardment with lighter ions. This result agrees with existing concepts of the dependence of the critical temperature on the mass of the incident particles [13].

Apparently, different types of sputtering produce the specific surface topologies observed in experiments. On the surfaces of GaP, GaAs, and InP, bombardment produces isolated cones characteristic of sputtering of amorphous semiconductor targets. An example of this surface relief is shown in Fig. 2a. The density and size of the cones produced by sputtering with N_2^+ ions are significantly less than in the case of the inert gases argon and xenon [5, 7, 8, 14–18]. At the same time, multifaceted pyramids forming on the surface of gallium antimonide and indium antimonide are typical of sputtering of monocrystalline targets [19]. In addition, it should be noted that sputtering of gallium antimonide at angles $\Theta \sim 40° \pm 3°$ produces a facet-type relief (Fig. 2c) with a characteristic direction perpendicular to the ion beam projection. In [20], formation of a similar specific morphology was observed for ion–target combinations GaP ← N_2^+ as a result of irradiation at angles around $40° \pm 3°$. Results of measuring the sputtering yield as a function of the energy of incident particles for the systems being studied are shown in Fig. 3. For gallium arsenide and phosphides of gallium and indium, the sputtering yield in the range of energies of N_2^+ ions from 2 to 14 keV is a monotonically increasing function. This type of energy dependence of the sputtering yield also agrees with theoretical predictions for GaP, GaAs, and InP and is described by the Haffa–Switkovski formula [21]

$$Y = U_s^{-1}[C_a^s(U_s)_a(Y)_a + C_b^s(U_s)_b(Y)_b], \qquad (3)$$

where Y is the net sputtering yield; U_s is the averaged surface bonding energy; C_a^s, C_b^s are surface atomic concentrations; and $(U_s)_a$, $(U_s)_b$ and $(Y)_a$, $(Y)_b$ are the surface bonding energies and sputtering yields of elemental materials a and b. The product $(U_s)_i(Y)_i$ was calculated using Yudin's approximation [22]

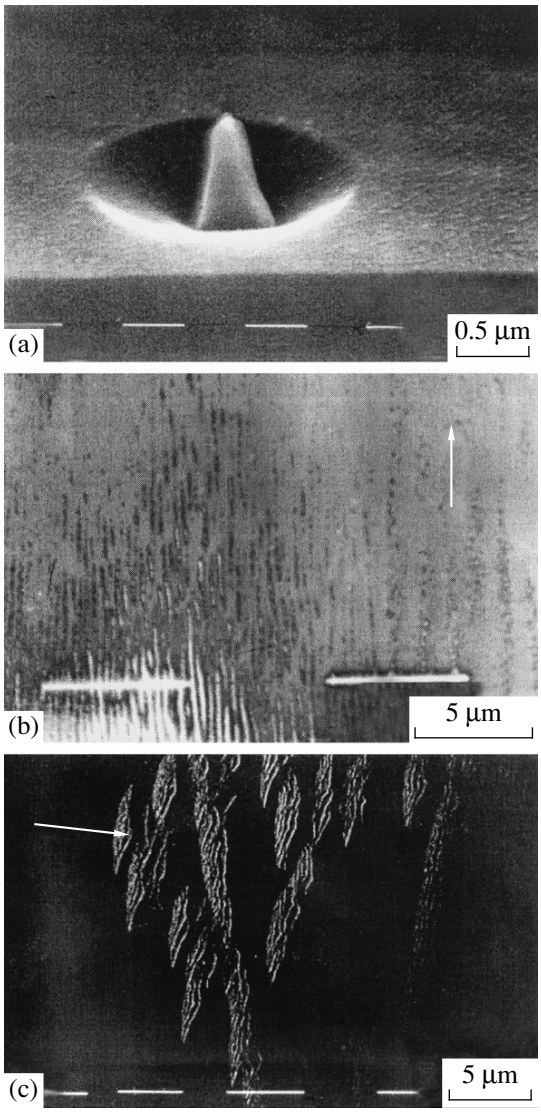


Fig. 2. Surface relief of GaAs (a, b) and InSb (c) targets after bombardment with N_2^+ ions. $E = 5$ keV; $\Phi \sim 0.5 \times 10^{19}$ ion/cm^2 at angles $\Theta = 0°$, 75°, and 40°, respectively. Arrows (b, c) indicate projections of the ion incidence direction.

$$Y = \frac{2Y_{\max}(E/E_{\max})^{1/2}}{1 + E/E_{\max}}, \qquad (4)$$

where $E_{\max}$ is the energy of bombarding particles for which a maximum value of the sputtering yield $Y_{\max}$ is observed; the quantities $E_{\max}$ and $Y_{\max}$ are determined by the formulas

$$E_{\max} = 0.3E_{\text{tf}} = \frac{0.3Z_1Z_2(1 + M_i/M_t)}{6.92 \times 10^6 a}, \qquad (5)$$

$$Y_{\max} = \Lambda(Z_1, Z_2)\frac{n\pi a^2 E_{\text{tf}}}{U_s}, \quad (6)$$

where E_{tf} is the Thomas–Fermi energy; a is a screening radius of the interaction potential; and Λ is the material constant defined by the formula

$$\Lambda(Z_1, Z_2) = \Lambda(Z_2) - 4.65 \times 10^{-12}(Z_1 - 18), \quad (7)$$

$$\Lambda(Z_2) = 1.3 \times 10^{-10} Z_2^{1/2} \times \left[1 + 0.25\sin\left(2\pi\frac{Z_2 + Z + Z_0/4}{Z_0}\right)\right], \quad (8)$$

where $Z_0 = 8$, $Z = 0$ at $Z_2 \le 18$ and $Z_0 = 18$, $Z = 2$ at $Z_0 > 18$.

The surface concentrations C_a^s and C_b^s were estimated using the known Petterson–Shirn–Zigmund equation [23, 24]

$$\frac{C_a^s}{C_a^s} = \frac{C_a^v(U_s)_b}{C_a^v(U_s)_a}\left[\frac{M_b}{M_a}\right]^{1-2m}, \quad (9)$$

where C_i^v and M_i ($i = a, b$) are bulk concentrations and masses of the target atoms and m is a parameter of the Coulomb screening potential.

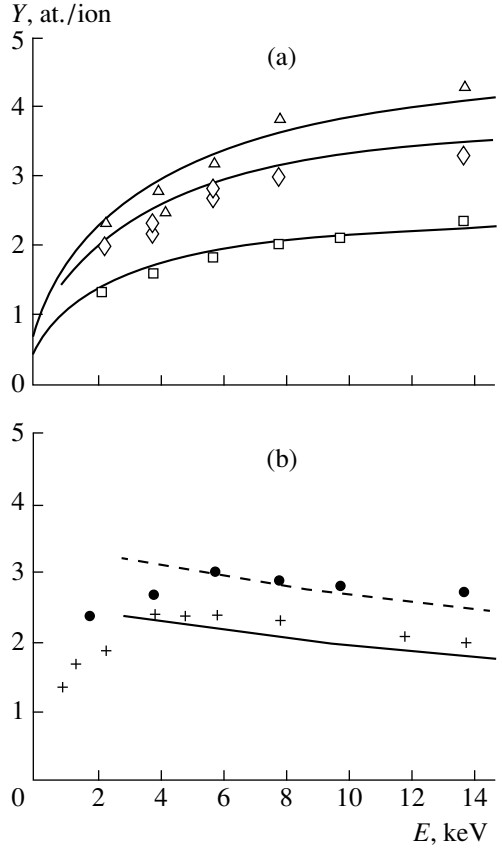


Fig. 3. Dependence on energy of the sputtering yield under bombardment with N_2^+ ions. Squares, GaP; diamonds, GaAs; triangles, InP; plus signs, GaSb; solid circles, InSb.

For estimates, values of the surface bonding energies of the target atoms $(U_s)_a$ and $(U_s)_b$ were taken equal to the corresponding energies for compound materials $(U_s)_a$ and $(U_s)_b$ [25, 26]. The average energy of the surface barrier was determined as a fitting parameter U_s. The values of U_s obtained from a comparison of experiment and theory are given in Table 2. Comparison of these values shows that the averaged surface bonding energies for GaAs and InP are approximately equal to the atomization energies and exceed the values obtained earlier for sputtering by argon ions [8]. For gallium phosphide, the experimental value of the surface barrier height is larger than the atomization energy of this material and is close to the atomization energy of GaN crystals (U_s = 4.4 eV). The observed difference in energy values for GaP, GaAs, and InP is apparently related to the formation on the target surface of a layer containing A^3–N compounds, in which the bonds are stronger than in the initial sample material.

Because sputtering of gallium antimonide and indium antimonide has features typical of crystalline structures, sputtering of these materials was analyzed using Onderlinden's theory [27, 28], which takes into account the effect of surface crystallinity on sputtering processes:

$$Y_{uvw}(E) = \chi_{uvw}(E)\eta_{uvw}\hat{Y}(E), \quad (10)$$

where η_{uvw} is a fitting parameter; its values for low-index crystallographic directions [100], [110], and [111] are equal to $\eta_{100} = \eta_{111} = 1.3$ and $\eta_{110} = 1.6$ [27]; χ_{uvw} is the noncanalized part of the beam, which, neglecting thermal oscillations, can be determined by the formula [27, 28]

$$\chi_{uvw} = \pi n t_{uvw} p_m^2 = \pi n t_{uvw}^3 \Psi_{uvw}^2 = \pi n t_{uvw}^{3/2}\sqrt{3a^2 Z_1 Z_2 e^2/E} = \sqrt{E_{uvw}^c/E}, \quad (11)$$

where t_{uvw} is the interatomic distance along the [uvw] direction; p_m is the minimum parameter of impact for canalization; Ψ_{uvw} is the critical canalization angle along the [uvw] direction; E_{uvw}^c is the characteristic canalization energy along the [uvw] direction below which canalization has no influence on sputtering; a is the potential screening radius; and e is the electron charge.

In the calculations, the charge numbers were taken as $Z_1 = 14$ and $Z_2 = (Z_a + Z_b)/2$ and the surface concentrations of atoms were equal to bulk concentrations. In the energy range above 3 keV, the considered combination of theories gives results that are in fair agreement with experiment (Fig. 3b). The best agreement between calculations and experiment is obtained for $(U_s)_{\text{GaSb}}$ = 2.4 eV and $(U_s)_{\text{InSb}}$ = 2.1 eV, which is lower than the values for sputtering by argon ions [5, 8]. This discrep-

Table 1. Parameters of the angular dependence of sputtering yield for A^3B^5 materials under bombardment with N_2^+ ions of energy E = 5 keV

Material	p	Θ_{max}, deg	f	g
GaP	3.9	63	7.2	3.32
GaAs	2.2	67	3.72	1.42
InP	1.9	65	3.29	1.39

Table 2. Energies of atomization U_d, amorphization U_{am}, sputtering threshold U_{th}, and surface bonding U_0

Material	U_a, eV	U_0, eV N_2^+	U_0, eV Ar^+	U_{am}, eV	U_{th}, eV (theory)
GaP	3.56	4.2	2.2	0.6	~25
GaAs	3.28	3.4	2.8	0.5	~25
GaSb	3.02	2.4	2.65	0.39	~20
InP	3.29	3.4	2.8	0.45	~30
InSb	2.75	2.1	2.5	0.25	~25

ancy is possibly caused by dependence of the surface bonding energy on the crystallographic orientation of the target surface and/or formation of volatile compounds on the sputtered surface.

In this work, experimental investigations of the dependence of the sputtering yield of A^4B^5 materials on energy and incidence angle of N_2^+ ions have been carried out, and the surface relief patterns formed in the sputtering process have been determined. Analysis of the experimental data has shown that for GaP, GaAs, and InP the Haffa–Switkovski model provides a good description of the dependence on energy of the sputtering yield, with the use of Yudin's approximation for sputtering of elemental materials.

It has been found that sputtering of indium antimonide and gallium antimonide at temperatures >50°C proceeds with recrystallization of the surface. It has been shown that for N_2^+ ions in the energy range above 3 keV Onderlinden's model can be used to describe the sputtering yield of gallium antimonide and indium antimonide. Values of the surface bonding energies for targets sputtered by N_2^+ ions in the amorphous mode obtained from a comparison of theory with experiment exceed the analogous values in the case of bombardment with ions of the inert gases Ar^+ and Xe^+, which is apparently due to the formation of A^3–N bonds with implanted nitrogen atoms. At the same time, values of the surface bonding energies obtained for InSb and GaSb sputtered by bombardment with N_2^+ ions in the mode of radiation defect anneal are, on the contrary, lower than the analogous values determined from analysis of sputtering of these materials by ions of the inert gases Ar^- and Xe^+ in the surface amorphization mode.

REFERENCES

1. N. Herbots, O. Vancauwenberghe, O. C. Hellman, *et al.*, in *Low Energy Surface Interactions*, Ed. by J. W. Rabails (Wiley, Cambridge, 1993), pp. 1–82.
2. N. Herbots, O. Vancauwengerghe, O. C. Hellman, and Y. C. Joo, Nucl. Instrum. Methods Phys. Res. B **59/60**, 326 (1991).
3. L. A. DeLouise, J. Vac. Sci. Technol. A **11**, 609 (1993).
4. W. Katzshner, U. Niggebrügge, R. Löffer, *et al.*, Appl. Phys. Lett. **48**, 230 (1986).
5. J. B. Malherbe, Appl. Surf. Sci. **70/71**, 322 (1993).
6. V. T. Barchenko *et al.*, Izv. Leningr. Élektrotekh. Inst. im V. I. Ul'yanova **303**, 42 (1982).
7. N. A. Bert and I. P. Soshnikov, Fiz. Tverd. Tela (St. Petersburg) **35**, 2501 (1993).
8. H. Andersen and H. Bai, in *Sputtering by Particle Bombardment*, Ed. by R. Behrisch (Springer-Verlag, New York, 1981; Mir, Moscow, 1984), Vol. 1.
9. N. A. Bert, K. Yu. Pogrebitskiĭ, I. P. Soshnikov, and Yu. N. Yur'ev, Zh. Tekh. Fiz. **62** (4), 162 (1992) [Sov. Phys. Tech. Phys. **37**, 449 (1992)].
10. J. I. Goldstein, P. Newburt, and P. Echlin, *Scanning Electron Microscopy and X-ray Microanalysis* (Plenum, New York, 1981; Mir, Moscow, 1984).
11. Y. Yamamura, Radiat. Eff. **80**, 57 (1984).
12. I. L. Singer *et al.*, Surf. Sci. **108**, 7 (1981).
13. G. Betts and G. Vener, in *Sputtering by Particle Bombardment*, Ed. by R. Behrisch (Springer-Verlag, New York, 1983; Mir, Moscow, 1986), Vol. 2.
14. S. R. Bhattacharya, D. Ghose, and D. Basu, Nucl. Instrum. Methods Phys. Res. B **47**, 253 (1990).
15. O. Wada, J. Phys. D **17**, 2429 (1984).
16. I. P. Soshnikov, N. A. Bert, Yu. A. Kudryavtsev, and A. V. Lunev, Poverkhnost, No. 3, 83 (1997).
17. I. P. Soshnikov, N. A. Bert, Yu. A. Kudrjavtsev, and A. V. Lunev, Nucl. Instrum. Methods Phys. Res. B **127/128**, 115 (1997).
18. J. Comas and C. B. Cooper, J. Appl. Phys. **38**, 2956 (1967).
19. J. L. Whitton, in *Erosion and Growth of Solids Stimulated by Atom and Ion Beams*, Ed. by G. Kiriakidis, G. Carter, and J. L. Whitton (Martinus Nijhoff, 1986).

20. V. K. Smirnov and V. I. Bachurin, in *Proceedings of the 11th Conference "Interaction of Ions with a Surface," 1993*, Vol. 2, p. 83.
21. P. K. Haff and Z. E. Switkowski, Appl. Phys. Lett. **29**, 549 (1976).
22. V. V. Yudin, Élektron. Tekh., Ser. Poluprovodn. Prib., No. 6, 3 (1984).
23. W. L. Patterson and G. A. Shirn, J. Vac. Sci. Technol. **4**, 343 (1967).
24. P. Zigmund, in *Sputtering by Particle Bombardment*, Ed. by R. Behrisch (Springer-Verlag, New York, 1981; Mir, Moscow, 1984), Vol. 1.
25. *Thermodynamical Constants of Matter*, Ed. by V. P. Glushko (Akad. Nauk SSSR, Moscow, 1971).
26. *Physicochemical Properties of Semiconductors: A Reference Book*, Ed. by E. V. Novoselova (Nauka, Moscow, 1978).
27. H. Rozendal, in *Sputtering by Particle Bombardment*, Ed. by R. Behrisch (Springer-Verlag, New York, 1981; Mir, Moscow, 1984), Vol. 1.
28. D. Onderlinden, Can. J. Phys. **46**, 739 (1968).

Translated by B. Kalinin

Technical Physics, Vol. 45, No. 9, 2000, pp. 1207–1211. Translated from Zhurnal Tekhnicheskoĭ Fiziki, Vol. 70, No. 9, 2000, pp. 119–123.

EXPERIMENTAL INSTRUMENTS AND TECHNIQUES

A Spread in Electrophysical Parameters of Ferroelectric Piezoelectric Solid Solutions and Its Minimization

L. A. Reznichenko, A. Ya. Dantsiger, S. I. Dudkina, L. A. Shilkina, O. N. Razumovskaya, V. A. Servuli, and I. V. Pozdnyakova

Research Institute of Physics, Rostov State University, pr. Stachki 194, Rostov-on-Don, 344090 Russia

E-mail: Larisa@riphys.rnd.su

Received January 28, 1999; in final form, September 23, 1999

Abstract—A spread in electrophysical parameters of solid solutions based on PZT and niobium oxides is considered for a wide range of component concentration. It is shown that composition fluctuations and their associated solution imperfection cause a significant deviation of the parameters from their mean values. The fluctuations are most pronounced when the crystal structure is unstable (e.g., at morphotropic transitions and near the solubility limits of the components). The optimization of methods for solid solution preparation greatly reduces the effect of process variables on parameter reproducibility. *© 2000 MAIK "Nauka/Interperiodica".*

Parameter reproducibility of active elements in piezoelectric devices is of crucial importance in mass production. An insignificant spread in their parameters within a lot or from lot to lot eliminates the need for tuning individual elements and provides the reliable operation of equipment. These elements are usually made of composite ferroelectrics, such as solid solutions (SS) based on PZT and niobium oxides. The functionality of the material depends on its chemical and physical compositions. Therefore, it is of interest to trace the interrelation between variations of composition and parameter spread.

In this work, we give the statistical treatment of a parameter spread in PZT- and sodium niobate–based systems when the concentration of one component is varied. The statistical characteristics are rms deviation of the mean

$$S_n = \sqrt{\frac{1}{n(n-1)}\sum_{i=1}^{n}(x_1-\bar{x})^2}, \quad (1)$$

absolute deviation

$$\overline{\Delta x} = S_n t n_\alpha, \quad (2)$$

and relative deviation

$$\overline{\delta x} = \frac{\overline{\Delta x}}{x}(100\%), \quad (3)$$

where x is the mean value of a parameter ($\bar{x} = \sum_{i=1}^{n} x_i/n$), x_i is its current value, n is the number of samples in a lot, and tn_α is Student's ratio. The characteristics were estimated at a confidence probability of 0.95 [1].

We studied $NaNbO_3$–$PbTiO_3$ (Fig. 1) and $PbZrO_3$–$PbTiO_3$–$PbW_{1/2}Cd_{1/2}O_3$ (Fig. 2) samples obtained in laboratory conditions[1]. Figures 1 and 2 plot $\bar{x}$, $\overline{\Delta x}$, and $\overline{\delta x}$ vs. $PbTiO_3$ concentration (curves *a*, *b*, and *c*, respectively). These quantities were calculated for three electrophysical parameters: relative permittivity $\varepsilon_{33}^T/\varepsilon_0$ (A), piezoelectric coupling coefficient K_p(*B*), and mechanical Q-factor Q_M(*C*).

It is seen that, in some ranges of $PbTiO_3$ concentration, $\overline{\delta x}$ curves run parallel to the abscissa axis; i.e., the relative deviations are nearly the same (within a little spread). For the most fully studied $NaNbO_3$–$PbTiO_3$ system (Fig. 1), such portions are observed for the tetragonal phase T_2 adjacent to the morphotropic region MR_2 (for all of the parameters), for the very MR_2 (for $\varepsilon_{33}^T/\varepsilon_0$ and K_p), and for rhombic phase R_1 adjacent to MR_1 (for Q_M, this also takes place in a considerable part of MR_1)[2]. In Fig. 1, the ranges where $\overline{\delta x}$ runs parallel to the abscissa are hatched[3]. Here, the concentration curves for the absolute deviations $\overline{\Delta x}$ qualitatively coincide with those for the electrophysical parameters.

[1] Solid solutions were obtained by solid-phase synthesis (800°C, 1.8×10^4 s, 850°C, 1.8×10^4 s) with subsequent individual rapid hot pressing of samples (∅ 12 × 4 mm) at temperatures between 1150 and 1200°C (depending on the composition), a pressure of 19.6 MPa, and isothermal hold for 2.4×10^3 s. Hot pressing methods and equipment are described in [2]. SS parameters were determined with lots of ten samples of each of the compositions.

[2] Phases in $NaNbO_3$–$PbTiO_3$ system are designated in Fig. 1 according to [3].

[3] Double hatching means some spread in $\overline{\delta x}$ for Q_M.

1063-7842/00/4509-1207$20.00 © 2000 MAIK "Nauka/Interperiodica"

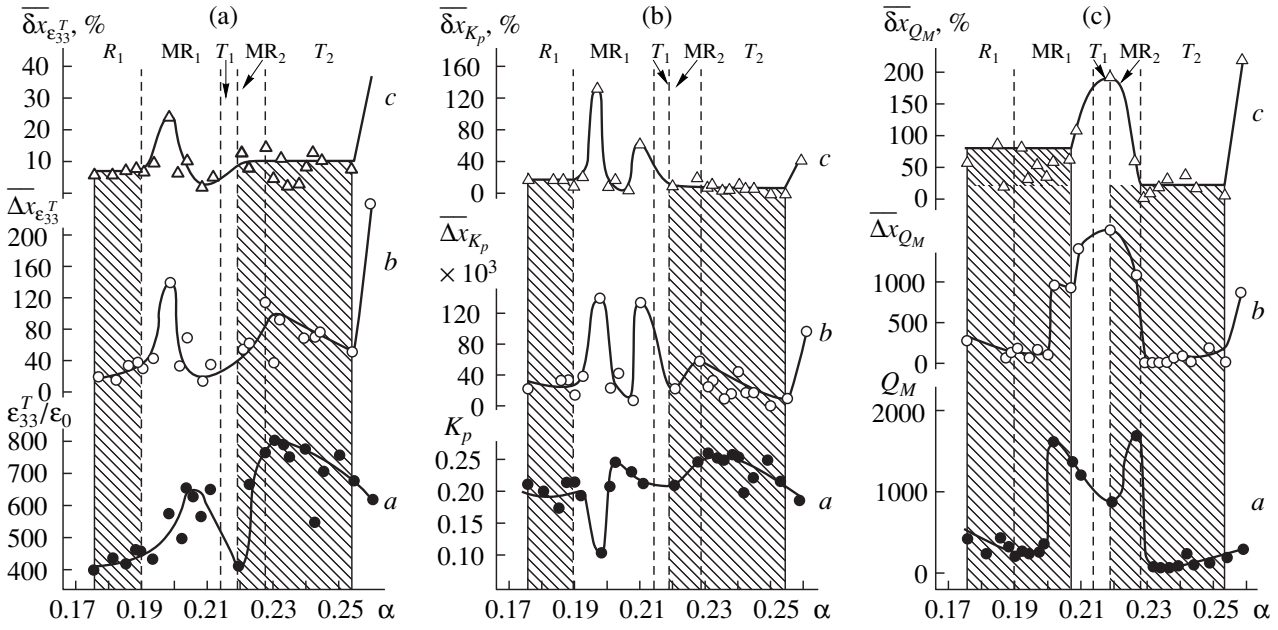


Fig. 1. (*a*) $\bar{x}$, (*b*) $\overline{\Delta x}$, and (*c*) $\overline{\delta x}$ vs. $PbTiO_3$ content in $(1-\alpha)NaNbO_3$–$\alpha PbTiO_3$ systems for (a) $\varepsilon_{33}^T/\varepsilon_0$, (b) K_p, and (c) Q_M.

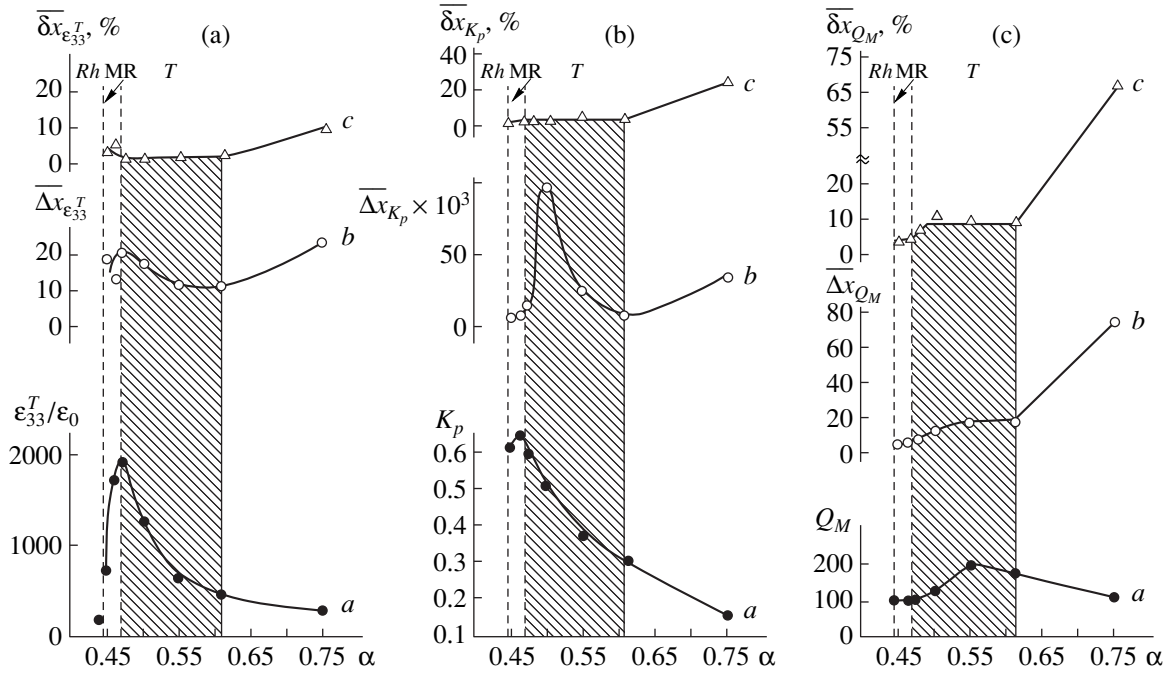


Fig. 2. The same as in Fig. 1 for $\alpha PbTiO_3$–$\beta PbZrO_3$–$\gamma PbW_{1/2}Cd_{1/2}O_3$ systems.

This also follows from analytic expression (3): at $\overline{\delta x} = \overline{\Delta x}/\bar{x} \approx \text{const}$, $\overline{\Delta x} \approx \text{const}\,\bar{x}$; that is, $\overline{\Delta x} \sim \bar{x}$.

In the rest of the $PbTiO_3$ concentration ranges in Fig. 1, $\overline{\delta x}$ varies considerably: up to 24% for $\varepsilon_{33}^T/\varepsilon_0$ and 130% for K_p (inside MR_1), as well as 200% for Q_M (inside MR_2). Note that the $\overline{\delta x}$ variations far exceed those of the parameters within the morphotropic regions. Therefore, the absolute deviations of the parameters $\overline{\Delta x}$ vary in the same manner as $\overline{\delta x}$ in these ranges. This also qualitatively follows from (3): at $\overline{\delta x} = \overline{\Delta x}/\bar{x} \neq \text{const}$, $\overline{\Delta x} = \bar{x}\overline{\delta x}$; i.e., for a slowly varying $\bar{x}\overline{\Delta x} \sim \overline{\delta x}$. This is also supported by the coincidence of the positions of $\overline{\delta x}$ and $\overline{\Delta x}$ maxima and minima. For

the electrophysical parameters, however, the maxima of $\varepsilon_{33}^T/\varepsilon_0$ and K_p coincide with the minima of $\overline{\delta x}$ and $\overline{\Delta x}$; and the minima of K_p and Q_M, with the maxima of $\overline{\delta x}$ and $\overline{\Delta x}$. Such an unusual (in our opinion) run of $\overline{\delta x}$ and $\overline{\Delta x}$ is likely to arise from specific features of SS formation in the morphotropic regions.

Also noteworthy is a narrow range of $PbTiO_3$ concentration in the phase T_2 away from MR where $\overline{\delta x}$, being hitherto almost constant, sharply grows. In this range, $\overline{\Delta x}$ behaves in the same manner, whereas $\varepsilon_{33}^T/\varepsilon_0$ and K_p continue to decrease monotonically.

Similar concentration dependences of parameter spread are also typical of other SS system. For example, in the compositionally simpler $PbZrO_3$–$PbTiO_3$–$PbW_{1/2}Cd_{1/2}O_3$ system (Fig. 2), ranges with $\overline{\delta x} \approx$ const (phase *T*, adjacent to the morphotropic region) are observed for the same parameters. As in Fig. 1, the associated ranges are hatched. Their width is 15 mol % $PbTiO_3$ (cf. 2.5 mol % in the phase T_2 for the former system). The concentration dependences of $\overline{\Delta x}$ (at $\overline{\delta x} \approx$ const) and electrophysical parameters are close to each other (for each of the parameters). Beyond the hatched range of $PbTiO_3$ content, $\overline{\delta x}$ grows considerably (the sharpest growth for Q_M), as in the case of $NaNbO_3$–$PbTiO_3$. Here, $\overline{\Delta x}$ also increases, while the parameters continue to decrease monotonically. Note that, in the $NaNbO_3$–$PbTiO_3$ system with the heterovalent substitution of ions in both cation sublattices, the ranges of constant $\overline{\delta x}$ in the single-phase regions, where the solid solutions are structurally unstable, are narrower than in the PZT system. A possible explanation is the disordering of cations that occupy equivalent crystallographic sites, which eventually causes SS inhomogeneity and imperfection. The morphotropic region in the PZT-based system has been studied less extensively than the niobate one.

When treating experimental data, one should take into account that a parameter spread involves the error in parameter determination, which depends on measuring techniques and equipment, environmental conditions, and operator's skill. For PZT-based SS, the associated values are usually several percent [4] (they reach 10–20% only for Q_M) and remain almost unchanged when the parameters are varied in wide limits (Table 1 [4]). In sodium niobate–based SS, the errors are somewhat larger than those given in [4] (Table 2), yet remaining much lower than $\overline{\delta x}$. Thus, the errors have an insignificant effect on the parameter spread, which depends on the SS position in the phase diagram.

The spread is apparently due to composition fluctuations, which show up most vividly in the regions of phase coexistence [5, 6]. In morphotropic regions, where the parameter values quickly respond to composition variations, any stoichiometry violation leads to a marked change in the phase state and structure. This, in turn, causes noticeable variations of the electrophysical properties, because oxide materials are highly sensitive to composition uniformity and perturbation [7, 8]. As is known, solid-state reactions usually do not proceed to completion; hence, SS are inhomogeneous in composition. After subsequent sintering according to conventional ceramic technology, the reproducibility of

Table 1. Relative deviations of the electrophysical parameters for PZT-based solid solutions

Parameter	Value	Relative deviation, %
$\varepsilon_{33}^T/\varepsilon_0$	250–5000	1
K_p	0.3–0.4	2
	0.4–0.5	1.5
	0.5–0.7	1
Q_M	50–600	10
	600–5000	20

Table 2. Relative deviations of the electrophysical parameters for mass-produced niobate materials

Parameter	Value	Relative deviation, %
$\varepsilon_{33}^T/\varepsilon_0$	120–460	1
K_p	0.30–0.30	5–6
	0.30–0.40	3
Q_m	100–200	15
	200–1000	20

Table 3. Phase ratio in solid solutions of composition $(1 - \alpha)$ $NaNbO_3$–$\alpha PbTiO_3$ at $\alpha = 0.225$

Preparation technique	Sample no.	*P*/*q*	Preparation technique	Sample no.	*P*/*q*
Hot pressing	1	1.2	Conventional ceramic technology	1	1.7
	2	1.3		2	1.4
	3	1.1		3	1.3
	4	1.2		4	1.4
	5	1.3		5	1.3
	6	1.2		6	1.6
	7	1.1		7_{up}	1.4
				7_{down}	1.7

Table 4. Electrophysical parameters of samples made of different PKR-34 hot-pressed material blocks

Lot no.	Block no.	Parameters*								
		$\varepsilon_{33}^T/\varepsilon_0$	K_p	d_{31}, pC/N	g_{31}, mV m/N	$Y_{11}^E \times 10^{-11}$ N/m^2	ρ, g/cm^3	T_C °C	d_{33}^{**}, pC/N	g_{33}^{**}, mV m/N
I	1	450	0.33	36	9.2	1.14	4.5	390	85	21.3
	2	460	0.34	37	9.0	1.15	4.5	388	85	20.3
	3	490	0.32	36	8.2	1.30	4.5	390	85	19.6
II	4	465	0.33	37	9.1	1.17	4.5	390	90	21.3
	5	450	0.32	35	8.8	1.20	4.5	385	85	21.3
	6	400	0.30	37	10.4	1.32	4.5	380	85	24.0
III	7	390	0.31	37	10.7	1.17	4.5	390	90	26.0
	8	385	0.30	36	10.6	1.15	4.5	388	85	24.9

* d_{31} and d_{33}, piezoelectric constants; g_{31} and g_{33}, piezoelectric sensitivities; Y_{11}^E, Young's modulus; ρ, density; and T_C, Curie point.
** g_{33} and d_{33} are obtained by the quasi-static method.

Table 5. Statistical characteristics of hot-pressed samples made of the same PKR-34 material block

	$\varepsilon_{33}^T/\varepsilon_0$	K_p	d_{31}, pC/N	g_{31}, mV m/N		$\varepsilon_{33}^T/\varepsilon_0$	K_p	d_{31}, pC/N	g_{31}, mV m/N
$\bar{x}$	425	0.388	41	10.8	$\overline{\Delta x}$	4.2	0.009	1.2	0.4
S_n	2.0	0.004	0.6	0.2	$\overline{\delta x}$, %	1.0	2.3	2.9	3.7

the properties worsens still further. This is illustrated by data represented in Table 3. It lists the values of the parameter P/q [9] for SS of composition $(1 - \alpha)NaNbO_3$–$\alpha PbTiO_3$ at $\alpha = 0.225$. This parameter was suggested as a measure of the tetragonal distortion of the SS perovskite lattice and the amount of tetragonal phase when the cell parameter ratio c/a is impossible to determine because of an uncertain split of associated diffraction lines. As follows from Table 3, the spread of P/q for various samples and for various areas within the morphotropic region is rather large and depends on the production technology. Samples prepared by hot pressing have the lesser spread of P/q. For samples prepared by this method, a distinct tendency to decreasing the spread of the parameters is observed in going from the laboratory [10] to large-scale production technology [11]. In the latter case, the spread is minimal within a lot and when samples are made of the same block of the material (Tables 4, 5).

Such an effect is explained by the stabilized composition and more perfect crystal structure of the SS.

CONCLUSIONS

(1) The behavior of the absolute, $\overline{\Delta x}$, and relative, $\overline{\delta x}$, parameter deviations depends on the SS position in the phase diagram. In single-phase regions, $\overline{\Delta x}$ variations are similar to those of $\bar{x}$ and $\overline{\delta x}$ is constant. In multiphase regions, the increased sensitivity of the SS to chemical inhomogeneity causes variations of $\overline{\Delta x}$ and $\overline{\delta x}$ much greater than in the former case. Here, $\overline{\Delta x}$ varies in the same manner as $\overline{\delta x}$ and, as a rule, in the direction opposite to the direction of $\bar{x}$ variation, while $\overline{\delta x}$ depends on component concentration in an intricate way.

(2) Heterovalent substitutions of ions in both cation sublattices of niobate-based SS make their composition still more nonuniform and their structure still more imperfect. This extends the parameter spread in comparison with the PZT-based systems.

(3) The spread can be minimized by optimizing the SS production process.

ACKNOWLEDGMENTS

This work was partially supported by the Russian Foundation for Basic Research (grant no. 99-02-17575).

REFERENCES

1. O. N. Kassandrova and V. V. Lebedev, *Processing of Observations* (Nauka, Moscow, 1970).

2. E. G. Fesenko, *Perovskite Family and Ferroelectricity* (Atomizdat, Moscow, 1972).

3. L. A. Reznichenko, A, Ya. Dantsiger, O. N. Razumovskaya, *et al.*, in *Fundamental Problems of Piezoelectronics: Proceedings of the International Conference "Piezotechnology-95"* (Kniga, Rostov-on-Don, 1995), Vol. 2, p. 13.

4. A. Ya. Dantsiger, Doctoral Dissertation (Rostov-on-Don, 1986).

5. V. A. Isupov, Fiz. Tverd. Tela (Leningrad) **22** (2), 172 (1980).

6. M. F. Kupriyanov, G. M. Konstantinov, and A. E. Panich, Ferroelectric Morphotropic Transitions (Rostov. Gos. Univ., Rostov-on-Don, 1992).

7. Yu. D. Tret'yakov and H. Lepis, *Chemistry and Technology of Hardphase Material* (Mosk. Gos. Univ., Moscow, 1985).

8. E. G. Fesenko, A. Ya. Dantsiger, and O. N. Razumovskaya, *Novel Piezoelectric Ceramic Materials* (Rostov. Gos. Univ., Rostov-on-Don, 1983).

9. É. A. Zavadskiĭ and V. M. Ishchuk, *Metastabilities in Ferroelectrics* (Naukova Dumka, Kiev, 1987).

10. A. Ya. Dantsiger, O. N. Razumovskaya, L. A. Reznichenko, and S. I. Dudkina, *High-Efficiency Piezoceramic Materials: A Reference Book* (Kniga, Rostov-on-Don, 1994), Vol. 2.

11. O. N. Razumovskaya, A. N. Klevtsov, V. A. Servuli, *et al.*, Izv. Akad. Nauk, Neorg. Mater. **31**, 843 (1995).

Translated by V. Isaakyan

Technical Physics, Vol. 45, No. 9, 2000, pp. 1212–1213. Translated from Zhurnal Tekhnicheskoĭ Fiziki, Vol. 70, No. 9, 2000, pp. 124–125.

BRIEF COMMUNICATIONS

The Effect of Dopants on the Texture Formation in MBE-Grown Polysilicon Films

T. A. Gracheva, N. D. Malygin, D. V. Shengurov, and V. G. Shengurov

Research Physicotechnical Institute, Lobachevskiĭ State University, Nizhni Novgorod, 603600 Russia

E-mail: shengurov@phys.unn.runnet.ru

Received February 16, 2000

Abstract—The texture of polysilicon films as a function of doping impurities and growth temperature was studied. © *2000 MAIK "Nauka/Interperiodica".*

Polysilicon films (PSFs) are a promising material for micro- and optoelectronic applications. Of special value are the films with preferential grain orientation [1, 2]. PSFs, which are obtained by low-temperature methods, are used in thin-film transistors. Such transistors have a higher carrier mobility in the channel than those based on the films with randomly oriented grains. MBE appears to be one promising low-temperature technique for PSF growth [3–5]. The properties of MBE-grown PSFs are known to depend on various factors, specifically, dopant type [4, 5]. Therefore, it would be of interest to trace the dopant effect on texture type.

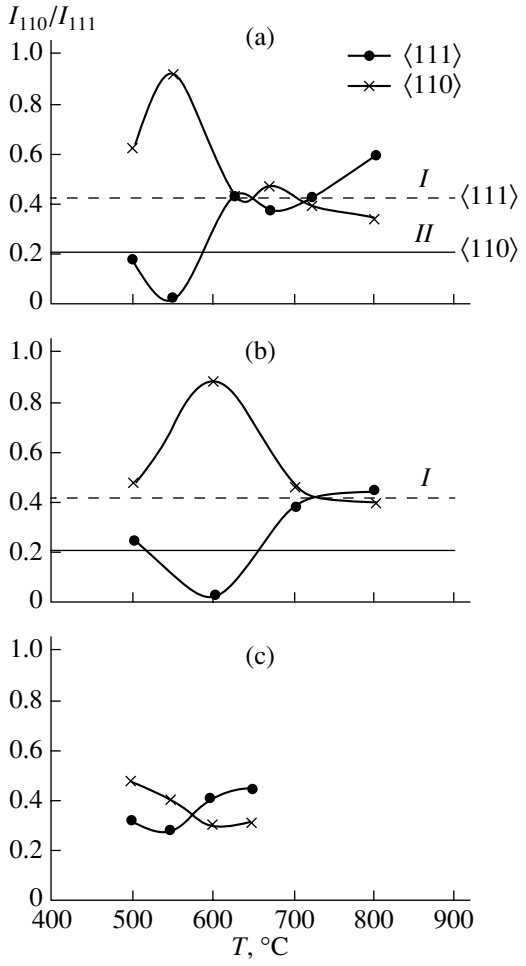
Silicon films ~1.5 μm thick were MBE-grown using volatilizable silicon sources doped by phosphorus or gallium to a concentration of $\sim 1 \times 10^{19}$ cm^{-3}. PSFs lightly doped by boron ($N_B \approx 10^{16}$ cm^{-3}) were also grown for comparison. Silicon wafers covered by ~0.5-μm-thick thermally grown SiO_2 were used as substrates. The PSF growth rate was ~2 μm/h, and the substrate temperature was varied from 500 to 800°C.

The films were studied by X-ray diffractometry. A part of the volume oriented in an (*hkl*) direction was estimated with the orientation factor f_{hkl}:

$$f_{hkl} = \frac{I_{hkl}/F_{hkl}}{\sum\limits_{hkl} I_{hkl}/F_{hkl}},$$

where I_{hkl} and F_{hkl} are, respectively, the scattering intensity and factor measured for each (*hkl*) diffraction plane. The scattering factor was determined for texture-free samples.

Figure 1a shows the degree of texturing (fractions of texture components) vs. substrate temperature T_S for the films grown with the use of the lightly doped source. For T_S between 500 and 550°C, the prevailing direction of texture orientation is ⟨110⟩. Note that, at T_S = 550°C, about 90% of the grain volume is oriented in the direction ⟨110⟩. As T_S increases further, the fraction of the ⟨110⟩ orientation drops and the rate of ⟨111⟩ texture formation rises. For the films heavily doped by phosphorus, the temperature dependence of film orientation is similar (Fig. 1b). Here, the maximum of the ⟨110⟩ orientation and the minimum of the ⟨111⟩ orien-



Degree of PSF texturing vs. substrate temperature for the films doped by (a) boron ($N_B = 10^{16}$ cm^{-3}), (b) phosphorus ($N_P \sim 10^{19}$ cm^{-3}), and (c) gallium ($N_{Ga} \sim 10^{19}$ cm^{-3}). (*I* and *II*) references.

1063-7842/00/4509-1212$20.00 © 2000 MAIK "Nauka/Interperiodica"

tation are shifted by about 50°C toward higher temperatures as compared with Fig. 1a. For the films heavily doped by gallium (Fig. 1c), the ⟨110⟩ maximum shifts toward lower temperatures and the volume fraction of grains with this orientation is smaller than in the previous cases.

The table lists the grain sizes in the variously doped films for different growth temperatures. The grains are seen to coarsen with increasing temperature. For the films lightly doped by boron, coarsening is noticeable at high temperatures. For the phosphorus-doped PSFs, the grains grow faster. In the boron-doped PSFs, intense grain growth occurs at lower temperatures.

Our results suggest that the formation of ⟨110⟩ texture becomes significant in the gallium-doped films. In the films doped by phosphorus, the formation of this texture slows down at low temperatures. This observation may be considered within a model where dopant atoms are viewed as surfactants [6]. It appears that dopants influence the grain orientation and size through their tendency to surface segregation: phosphorus is less prone to segregation on the growth surface than gallium; hence, the former activates grain boundary migration to a lesser extent.

Thus, PSFs grown by MBE with volatilizable silicon sources exhibit ⟨110⟩ and ⟨111⟩ textures. At low temperatures (500–600°C), the former orientation prevails, while the latter is dominant in the high-temperature range. As for the effect of the dopants, the lowest temperature of ⟨110⟩ texture formation either rises (phosphorus) or decreases (gallium) in comparison with PSFs lightly doped by boron.

PSF grain size vs. dopant

Dopant	*T*, °C				
	550	600	650	700	800
	grain size, nm				
B	–	90	100	110	130
P	–	115	–	135	150
Ga	75	–	130	–	–

REFERENCES

1. K. T.-Y. Kung and R. Reif, J. Appl. Phys. **62**, 1503 (1987).
2. K. T.-Y. Kung and R. Reif, J. Appl. Phys. **63**, 2131 (1988).
3. D. A. Pavlov, V. G. Shengurov, D. V. Shengurov, and A. F. Khokhlov, Fiz. Tekh. Poluprovodn. (St. Petersburg) **29**, 286 (1995) [Semiconductors **29**, 142 (1995)].
4. D. V. Shengurov, Pis'ma Zh. Tekh. Fiz. **23** (11), 83 (1997) [Tech. Phys. Lett. **23**, 450 (1997)].
5. D. V. Shengurov, D. A. Pavlov, V. N. Shabanov, *et al.*, Fiz. Tekh. Poluprovodn. (St. Petersburg) **32**, 627 (1998) [Semiconductors **32**, 562 (1998)].
6. A. Sakai, T. Tatsumi, and K. Isnida, Phys. Rev. B **47**, 6803 (1993).

Translated by V. Isaakyan

Technical Physics, Vol. 45, No. 9, 2000, pp. 1214–1218. Translated from Zhurnal Tekhnicheskoĭ Fiziki, Vol. 70, No. 9, 2000, pp. 126–130.

BRIEF COMMUNICATIONS

Characteristics of the Heating Dynamics of a Graphite Conductor Taking into Account the Skin Effect

P. V. Novikov* and G. N. Churilov**

** Krasnoyarsk State Technical University, Krasnoyarsk, 660074 Russia*

*** Kirensky Institute of Physics, Siberian Division, Russian Academy of Sciences, Akademgorodok, Krasnoyarsk, 660036 Russia*

E-mail: churilov@iph.krasn.ru

Received June 12, 1999; in final form, January 17, 2000

Abstract—Calculations for a one-dimensional model of RF heating of a cylindrical graphite conductor have been carried out. The heating dynamics are analyzed in the general form. Conductor temperature profiles and the times for heating up to the graphite sublimation temperature as a function of current and frequency have been obtained. A model of conductor heating with partial return of the energy irradiated by the conductor surface has been considered. Frequency and current ranges have been determined to carry out this graphite sublimation method in a chamber with reflecting walls. The problem is associated with carbon vapor production and subsequent synthesis of fullerenes and other carbon structures. © *2000 MAIK "Nauka/Interperiodica".*

INTRODUCTION

Metals in the form of finely dispersed powders find an ever expanding use in technology. Heating conductors in the form of wires by passing high pulsed currents [1] is one of the effective ways of obtaining these materials. In 1990, arc sputtering of carbon was employed by a group of German physicists, which led them to the discovery of a method for fullerene synthesis [2]. The method consisted in the sputtering of carbon by a direct or alternating current arc in a helium atmosphere at a pressure of 100 torr. Afterwards, many other fullerene synthesis methods were developed. But in spite of the great diversity of these methods, their productivity did not exceed several grams per hour [3–9]. The common feature of all these methods is that carbon is first transformed into a plasma at a temperature of 6000–7000 K and then, during subsequent cooling (usually in a helium gas atmosphere), fullerene molecules are formed. The sublimation temperature of graphite in vacuum is about 4000 K. Unfortunately, the carbon plasma temperature at which fullerene molecules can be formed is unknown. We assume that this temperature is below 6000 K and that fullerenes can be synthesized from carbon vapor produced by graphite sublimation. The heating necessary for graphite sublimation can be effected by passing a current through a graphite rod.

It is evident that, when a direct current is passed through a graphite rod, the temperature near its axis exceeds that at the surface due to radiation losses. Sublimation will start first in cracks and voids in the graphite before expanding to other regions, which will lead to mechanical destruction of the graphite rod. Therefore, RF heating should be used, because due to the skin effect the joule heat will then be released mostly in the conductor surface layer [10].

In this paper, we consider the dynamics of heating a cylindrical conductor by RF currents to the temperature at which sublimation starts.

THE MODEL

Heating a conductor in vacuum is described by a nonstationary equation for heat conductivity in cylindrical coordinates [11]

$$c\rho\frac{\partial T}{\partial t} = \frac{1}{r}\frac{\partial}{\partial r}\left(\lambda r\frac{\partial T}{\partial r}\right) + q_V(r), \tag{1}$$

where $c = c_V$ is the specific heat, ρ is the density, λ is the heat conductivity, and $q_V(r)$ is the power of the volume heat sources.

Let us denote the boundary conditions at the conductor axis as

$$-\lambda\frac{\partial T}{\partial r}\bigg|_{r=0} = 0 \tag{2}$$

and at the conductor surface as

$$-\lambda\frac{\partial T}{\partial r}\bigg|_{r=r_0} = -q_r(T)|_{r=r_0}, \tag{3}$$

where $q_r(T) = \varepsilon\sigma_c T^4$ is the radiation flux density from the conductor surface; σ_c is the Stefan–Boltzmann constant; ε is the integrated emittance; and the initial condition is $T(r, t = 0) = T_0$ where $T_0 = 293$ K.

1063-7842/00/4509-1214$20.00

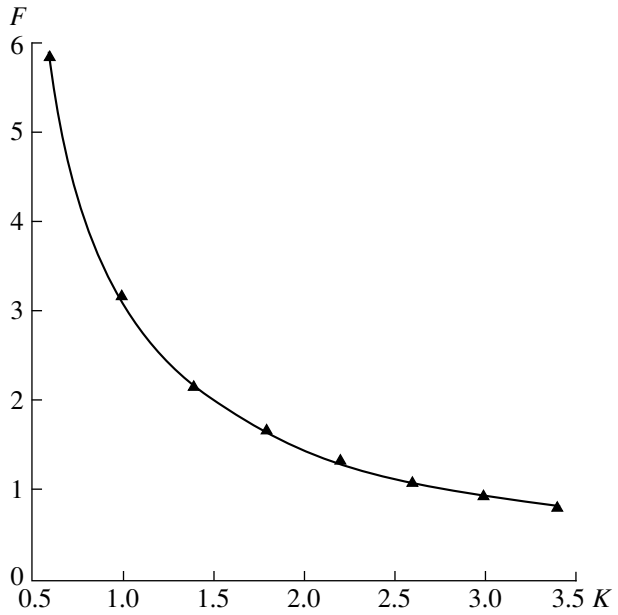


Fig. 1. Time for reaching the sublimation temperature F as a function of the dimensionless parameter K ($r_0 = 5 \times 10^{-2}$ m, $\Delta = 0.14$, $\alpha = 0$, and F is time in arb. units).

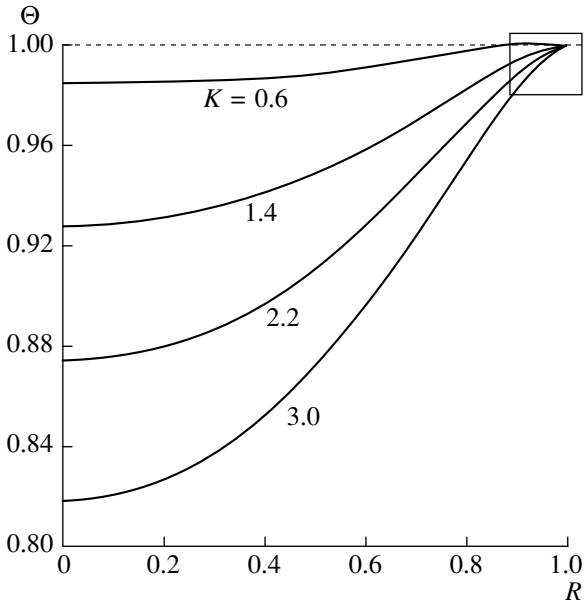


Fig. 2. Temperature profiles at the start of sublimation for different values of the parameter K ($r_0 = 5 \times 10^{-2}$ m, $\Delta = 0.14$, $\alpha = 0$). The portion in the rectangle is shown scaled up in Fig. 3.

In equation (1), the joule power of volume sources is determined as follows [10]:

$$q_V(r) = \frac{\overline{j^2}(r)}{\sigma} = \frac{Q}{\pi r \delta}\exp(-2(r_0 - r)/\delta), \qquad (4)$$

where $\delta = \sqrt{2/(\mu\sigma\omega)}$ is the skin layer thickness, $Q = I^2/(2\pi\sigma\delta r_0)$ is the thermal power per unit length of the conductor, r_0 is the conductor radius, σ is the electrical conductivity, $\mu \approx \mu_0$ is the magnetic permeability, I is the effective current, and ω is the current frequency.

Let us introduce normalized variables of the transfer process

$$R = \frac{r}{r_0}, \quad F = \frac{at}{r_0^2}, \quad \Theta = \frac{T}{T_e}, \quad \Delta = \frac{\delta}{r_0},$$
$$K = \frac{Q}{\lambda T_e}, \quad \xi(\Theta) = \frac{r_0}{\lambda T_e} q_r(T), \qquad (5)$$

where R and F are the spatial and time coordinates ($a = \lambda/c\rho$ is the thermal diffusivity); Θ is the relative temperature; Δ is the relative skin layer thickness; and K and ξ are parameters characterizing the specific power Q and the density q_r of the radiation flux from the surface, respectively.

In this way, the number of initial parameters can be reduced, and a solution in a general form can be obtained. With the new variables, the system (1)–(3) can be written as

$$\frac{\partial\Theta}{\partial F} = \frac{1}{R}\frac{\partial}{\partial R}\left(R\frac{\partial\Theta}{\partial R}\right) + \frac{K}{\pi R\Delta}\exp(-2(1 - R)/\Delta), \qquad (6)$$

$$\left.\frac{\partial\Theta}{\partial R}\right|_{R=0} = 0, \quad \left.\frac{\partial\Theta}{\partial R}\right|_{R=1} = -\xi(\Theta)|_{R=1}. \qquad (7)$$

The solution contains two parameters, K and Δ. The current, frequency, and time can be expressed through the generalized parameters as follows:

$$I = r_0\sqrt{2\pi\sigma\Delta K\lambda T_e},$$
$$\omega = \frac{2}{(r_0\Delta)^2\mu\sigma}, \qquad (8)$$
$$t = \frac{Fr_0^2}{a}.$$

CALCULATION RESULTS

Computer calculations were performed by the finite-difference method. The iteration algorithm provided second-order accuracy of the spatial and time steps [11]. Profiles of the conductor temperature at the moment when sublimation started on its surface and the time elapsed until this moment were calculated for different values of K and Δ. As the parameter K and, correspondingly, the power Q is increased, the time of heating to the sublimation temperature T_e decreases (Fig. 1). The radial temperature gradient rises (Fig. 2). Under these conditions, overheating of the conductor's central region does not occur when the sublimation temperature is reached on the surface.

Due to radiation losses, the temperature maximum occurs not at the conductor surface but close to it

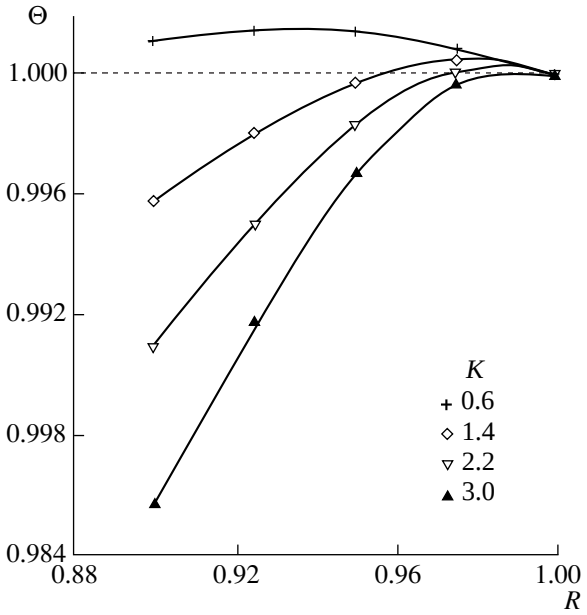


Fig. 3. Temperature profiles near the surface at the start of sublimation for different values of the parameter K ($r_0 = 5 \times 10^{-2}$ m, $\Delta = 0.14$, $\alpha = 0$).

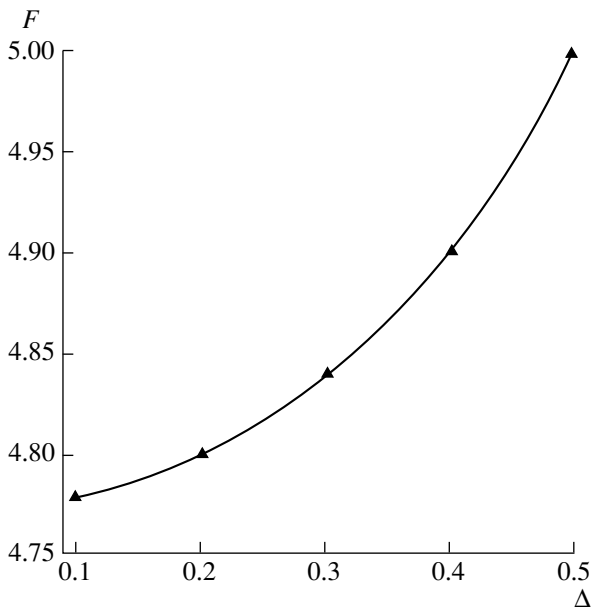


Fig. 4. Time taken to reach the sublimation temperature F as a function of the dimensionless parameter Δ ($K = 0.7$, $r_0 = 5 \times 10^{-2}$ m, $\alpha = 0$, and F is time in arb. units).

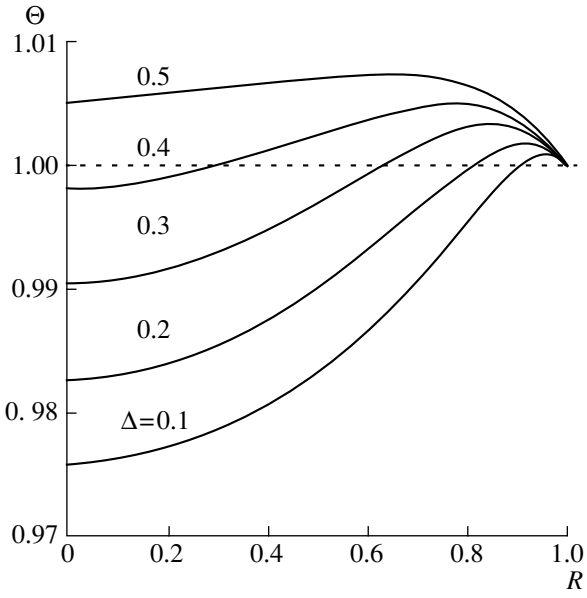


Fig. 5. Temperature profiles at the start of sublimation for different values of the parameter Δ ($K = 0.7$, $r_0 = 5 \times 10^{-2}$ m, $\alpha = 0$).

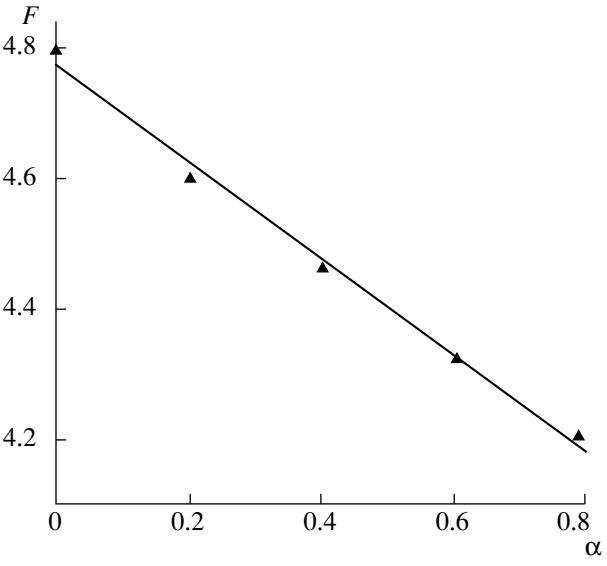


Fig. 6. Time taken for reaching the sublimation temperature F as a function of the coefficient α ($K = 0.7$, $\Delta = 0.14$, $r_0 = 5 \times 10^{-2}$ m, and F is time in arb. units).

(Fig. 3). As the parameter K increases, the maximum shifts towards the surface because of increasing heat release.

To evaluate the experimental potentialities of graphite sublimation as a means of synthesizing new carbon structures, we used the following values of the thermophysical parameters [12] for carbon, which were assumed to be constant during heating: specific heat c = 2.1 J/(g K), density $\rho = 2.1$ g/cm^3, heat conductivity λ = 2.66 W/(cm K), integrated emittance $\varepsilon = 0.56$, electrical conductivity $\sigma = 1.13 \times 10^4$ S/cm^2, and sublimation temperature $T_e = 4473$ K. For a graphite conductor of radius 5×10^{-2} m, the parameter range $K = 0.4$–3.4, according to (8), corresponds to a current range of 3.5–10.0 kA, and the time necessary to reach sublimation is 3.0–0.5 s (Fig. 1). If the current is reduced, by the time the sublimation temperature is reached at the surface, the conductor will be heated almost uniformly through-

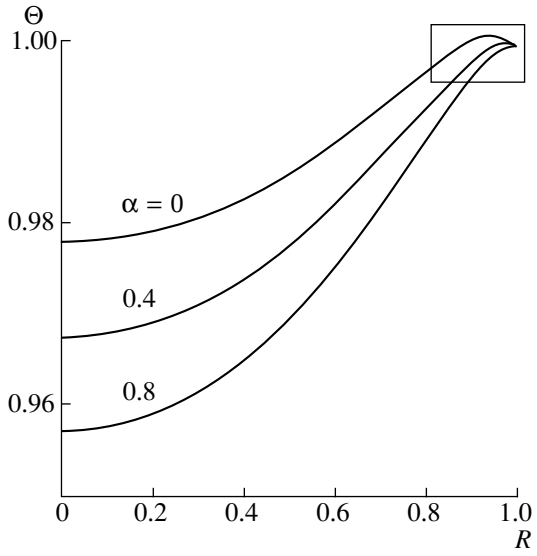


Fig. 7. Temperature profiles at the start of sublimation for different values of the parameter α ($K = 0.7$, $\Delta = 0.14$, $r_0 = 5 \times 10^{-2}$ m). The portion in the rectangle is shown scaled up in Fig. 8.

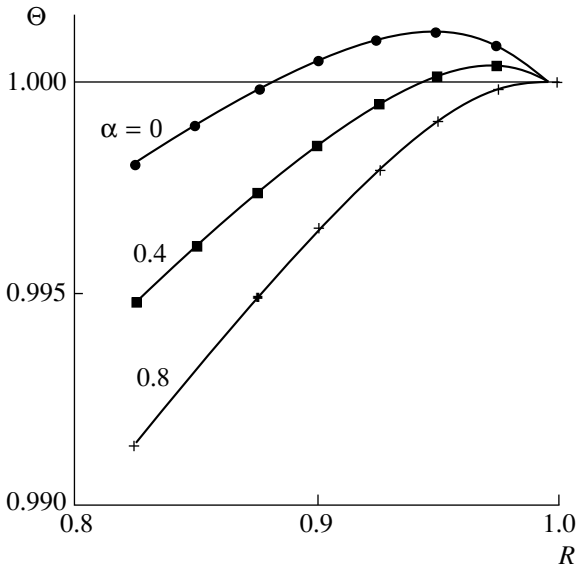


Fig. 8. Temperature profiles near the surface at the start of sublimation for different values of the parameter α ($K = 0.7$, $\Delta = 0.14$, $r_0 = 5 \times 10^{-2}$ m).

out its volume. If the current is increased, the heating time becomes too short and the process approaches the pulse heating described in [1, 13].

With increasing parameter Δ (which corresponds to decreasing frequency ω), the rod heating time increases, since both the volume heat release (see equation (6)) and the radial temperature gradient (Fig. 5) decrease due to the fact that with decreasing frequency the skin layer spreads over the conductor volume.

For a graphite conductor of radius 5×10^{-2} m, the range of 0.1–0.5 of the parameter Δ corresponds to the frequency range of 900–35 kHz. The calculations show that even when skin layer thickness is 0.4–0.5 of the conductor radius the inner conductor regions are considerably overheated (Fig. 5). Thus, a further decrease in the current frequency will result in an increase in the skin layer thickness that will make the use of RF heating unwarranted.

The calculations and estimations presented have shown that to produce the high currents corresponding to sublimation is, in practice, a very complicated task. Therefore, we changed the model of heating and calculated temperature profiles for a system with the following boundary condition at the surface:

$$-\lambda \left.\frac{\partial T}{\partial r}\right|_{r = r_0} = -(1-\alpha) q_r(T)|_{r = r_0}. \qquad (9)$$

Experimentally, this boundary condition can be created, for example, using a cylindrical heating chamber in which the walls have reflectance α.

The higher the reflectance of the walls, the shorter the time of heating to the sublimation temperature T_e (Fig. 6) and the higher the radial temperature gradient (Fig. 7). Moreover, at higher values of the reflectance α, the temperature maximum approaches the surface (Fig. 8); and at α = 0.6–0.8, it is found at the conductor surface.

The partial heat recovery also reduces power consumption. Thus, for a conductor of radius 5×10^{-2} m at α = 0.8, the reasonable range of the parameter K is 0.15–2.2, corresponding to a current of 2.0–8.0 kA and a time necessary to reach sublimation of 8.0–0.5 s.

CONCLUSIONS

(1) The dynamics of conductor RF heating has been analyzed in the general form as a function of the generalized parameters.

(2) Conductor temperature profiles and the times of heating to the sublimation temperature for currents of 2–10 kA in the frequency range of 35–900 kHz have been calculated using parameters corresponding to artificial graphite.

(3) Specific ranges of the control parameters have been determined to carry out graphite sublimation with reflecting walls at α = 0.8.

REFERENCES

1. V. I. Petrosyan and É. E. Dagman, Prib. Tekh. Éksp., No. 2, 213 (1990).
2. W. Kratschmer, D. R. Huffman, *et al.*, Chem. Phys. Lett. **170**, 167 (1990).
3. J. P. Hare, H. W. Kroto, and R. Taylor, Chem. Phys. Lett. **177**, 394 (1991).
4. W. A. Scrivens and J. M. Tour, J. Org. Chem. **57**, 6932 (1992).

5. V. P. Bubnov, I. S. Krainskiĭ, E. É. Laukhina, *et al.*, Izv. Akad. Nauk, Ser. Khim., No. 5, 805 (1994).
6. V. P. Afanas'ev, A. A. Bogdanov, G. A. Dyuzhev, *et al.*, Zh. Tekh. Fiz. **67** (2), 125 (1997) [Tech. Phys. **42**, 234 (1997)].
7. L. P. F. Chibante, A. Thess, J. M. Alford, *et al.*, J. Phys. Chem. **97**, 8696 (1993).
8. J. B. Howard, J. T. McKinnon, M. E. Johnson, *et al.*, J. Phys. Chem. **96**, 6657 (1992).
9. G. N. Churilov, L. A. Solovyov, Y. N. Churilova, *et al.*, Carbon **37**, 427 (1999).
10. I. E. Tamm, *The Principles of Electricity Theory* (Nauka, Moscow, 1966).
11. A. A. Samarskiĭ, *The Theory of Difference Schemes* (Nauka, Moscow, 1989).
12. É. N. Marmer, *Graphitic Carbon Materials: A Handbook* (Metallurgiya, Moscow, 1985).
13. V. S. Vorob'ev, Zh. Tekh. Fiz. **66** (1), 35 (1996) [Tech. Phys. **41**, 17 (1996)].

Translated by M. Astrov

Technical Physics, Vol. 45, No. 9, 2000, pp. 1219–1222. Translated from Zhurnal Tekhnicheskoĭ Fiziki, Vol. 70, No. 9, 2000, pp. 131–134.

BRIEF COMMUNICATIONS

Chemical Analysis of Products Obtained by Nanosecond Laser Ablation

M. N. Libenson, G. D. Shandybina, and A. L. Shakhmin

St. Petersburg State Institute of Fine Mechanics and Optics (Technical University), ul. Sablinskaya 14, St. Petersburg, 197101 Russia

E-mail: libenson@beam.ifmo.ru

Received November 18, 1999

Abstract—The chemical analysis of a laser jet was performed with a combination of laser back mass transfer and X-ray photoelectron spectroscopy (XPS). It was revealed that, as compared to the source material, films deposited from the jet under laser mass transfer in air show increased oxidizability and changed chemical composition. *© 2000 MAIK "Nauka/Interperiodica".*

INTRODUCTION

Ablation in a broad sense means mass removal from a solid surface (for instance, a decrease in the glacier mass due to thawing, thinning of a spacecraft body as it passes through the Earth atmosphere, and so on). This term has become common in the literature concerned with different fields of laser technology, microelectronics, and medicine, as well as with fundamental physical research.

Laser ablation is a phenomenon of removing a microscopic amount of material from the surface under the action of laser emission. It may proceed through different mechanisms. They are primarily evaporation, thermoelastic damage, shock wave propagation, photodynamic mechanism [1], photochemical mechanism [2], and others.

The nature of ablation processes depends to some extent on the target material and environment composition. However, more heavily they depend on the laser radiation characteristics: intensity, pulse duration, wavelength, number of pulses, etc. In certain ranges of laser parameters (for instance, at short pulses and high intensities), the fundamental physical mechanisms of laser ablation are yet to be fully understood.

Pulse laser ablation is a powerful method of surface treatment. High rates of heating and cooling along with the possibility to scan the focused radiation both over the surface and in the bulk, make it possible to carry out specific treatments in air. Under certain irradiation conditions, thermochemical processes in the vapor phase may have a profound effect on the result of ablation. In this paper, we study vapor-phase redox processes during short-pulse laser ablation in air. Metals and nontransparent semiconductor films are ablated due to evaporation and the formation of a near-surface plasma. The chemical composition of a laser jet resulting when sharp-focused pulsed laser radiation scans the target surface is also investigated.

EXPERIMENTAL METHODS AND RESULTS

In recent years, the laser ablation of a target material and its subsequent transfer to and condensation on a substrate placed at a small distance away from or directly against the target are the subject of much investigation. This technology is applied for fabricating film elements. Two types of laser transfer are recognized. The first type is direct transfer, when a radiation, having passed through a transparent substrate, evaporates a material deposited on it and the evaporated material condenses on the second substrate (acceptor) positioned in the direction of the laser beam. The second type is back laser transfer, where, after passing through a transparent acceptor, the laser radiation is focused on the target surface and the evaporated material flows in the direction opposite to that of the beam and condenses on the acceptor. In the former case, only thin films on a transparent substrate can be used, while, in the latter, both thin films and massive wafers are applied.

In this work, we employed back laser oblation to study vapor-phase redox processes. These processes attend the short-pulse laser ablation of such materials as massive titanium and aluminum plates, as well as chromium films and films of iron oxide deposited on glass substrates.

The experimental setup is presented in Fig. 1. After passing through a transparent acceptor mounted 50 μm above the target surface, a pulse single-mode YAG : Nd laser radiation was focused on the target surface to a spot 50 μm in diameter and uniformly scanned an area of 7×7 mm. The pulse width was 250 ns.

For each of the materials, the ablation power thresholds were estimated. In order of increasing thresholds, the materials are arranged as follows: chromium, titanium, iron oxide, and aluminum. In experiments, the luminous flux densities were 2–2.5 times greater than

1063-7842/00/4509-1219$20.00

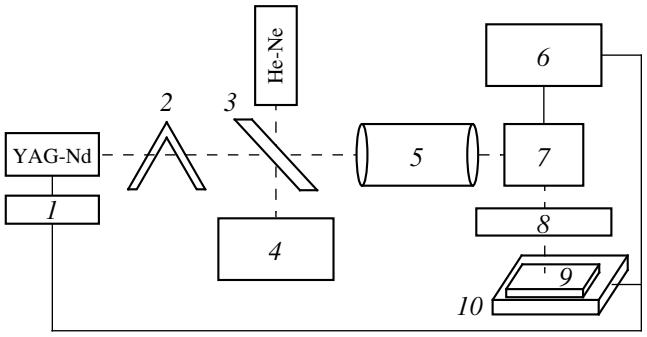


Fig. 1. Experimental setup: (*1*) pulse driver, (*2*) Fresnel attenuator, (*3*) transparent plate, (*4*) power meter, (*5*) optical system, (*6*) computer, (*7*) scanner, (*8*) substrate (acceptor), (*9*) target, and (*10*) operating table.

the ablation threshold of the material. The irradiation causes the formation of an intense jet. The acceptor was positioned in the region of its maximum intensity.

The chemical composition of the films deposited on the acceptor was examined with XPS. This places certain constraints on the size of the films (the irradiated area should be no less than 5 mm in diameter in view of the low spatial resolution of the XPS technique). The required area was delineated by scanning the laser spot over the sample surface. The scan rate was chosen in such a way as to prevent spot overlapping in "rows" and "columns."

The electronic spectra were obtained with an ESCA-5400 Perkin–Elmer spectrometer. The operating pressure was kept at 10^{-9} torr.

As is known, chemical analysis of dielectrics may distort the electronic spectra because of charging the material. This charge should be neutralized. To estimate the possible distortion of the spectrum, the spectra from the glass acceptor were obtained with and without charge neutralization. The neutralization was accomplished through bombarding the samples by low-energy (~20 eV) electrons. In the former case, the broadening of the spectral lines at the base was observed. In the latter case, the spectrum only shifted along the energy scale. Thus, the measurements without charge neutralization allowed us to find the undistorted spectral lines and resolve their components.

We obtained the electronic spectra for the source massive titanium and source aluminum foil, as well as for the films condensed from these sources, and resolved their components.

Similar spectra were obtained for the source films of iron oxide and chromium, as well as for the associated films produced by laser ablation.

Depth profiling by means of XPS and ion etching was made. It was found that Ar^+ ion etching to a depth of 2000 Å removes carboniferous contaminants and metal oxides from the metal surface.

For the massive samples, the depth of ablation per pulse was as great as 2–3 μm; that is, not only the con-

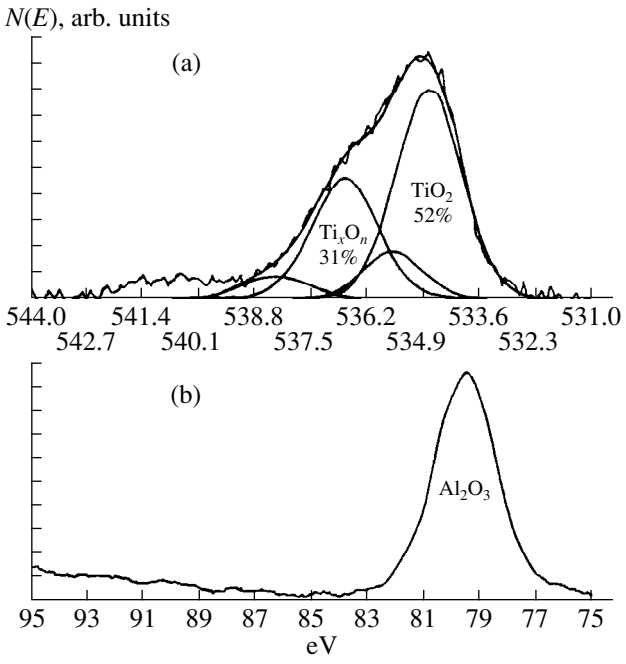


Fig. 2. Electronic spectra of the deposited films: (a) titanium target resolved Ti components) and (b) aluminum foil (resolved Al components).

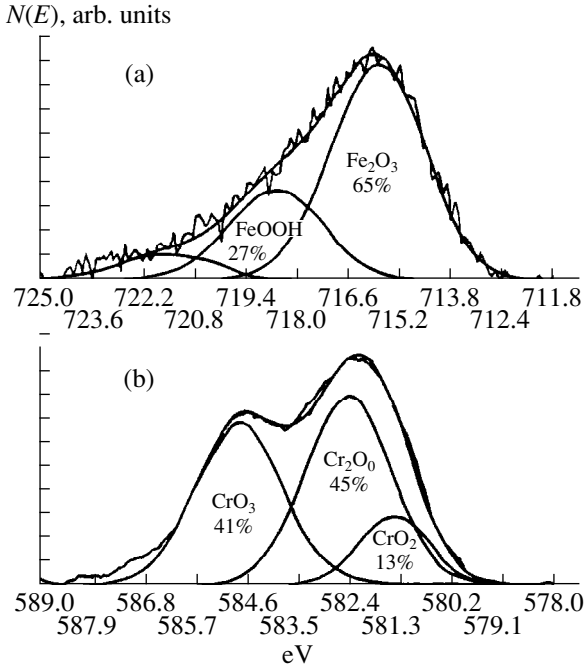


Fig. 3. Electronic spectra of the deposited films. Target is (a) iron oxide film on glass substrate (resolved Fe components) and (b) chromium film on glass substrate (resolved Cr components).

taminants but also the pure metal are removed from the surface in this case.

The photoelectron spectra of the titanium and aluminum films deposited in the jet area showed a great concentration of oxygen (O ≈ 50–56%), the metals (Al, Ti ≈ 12–25%), carbon (C ≈ 12–14%), and the substrate elements (the deposited films replicated the spatial structure of the scan area and were not continuous).

The resolution of the spectral components allowed the determination of the degree of oxidation of metals that combined with oxygen in the deposited films.

The Ti and Al components of the electronic spectra taken from the deposited titanium and aluminum films are presented in Fig. 2. Titanium is present in different oxidized forms (TiO_2 ≈ 50% and Ti_xO_n ≈ 30%), and Al is mainly oxidized to Al_2O_3 (the effect of charging is ignored) [3, 4].

The XPS spectra taken from the source iron oxide film show that it consists of different oxides and the pure metal (31% Fe_2O_3, 27% FeO, and 29% Fe).

The composition of the film produced by the short-pulse laser ablation of the source iron oxide film in air was radically different. The resolved spectrum of the deposited film is shown in Fig. 3a. In this film, pure iron was absent, the fraction of iron oxidized to the greatest extent (Fe^{+3}) doubled, and compounds of iron with a hydroxyl group appeared (27% FeOOH and 65% Fe_2O_3).

The XPS spectrum from the evaporated chromium film (Fig. 3b) revealed not only stable chromium oxide (Cr_2O_3), typical of thermal oxidation, but also unstable oxides of maximum valence: Cr^{+4} and Cr^{+6} (46% Cr_2O_3, 41% CrO_3, and 13% CrO_2).

DISCUSSION

Short-pulse laser ablation in air, which is attended by vapor-phase reactions, was found to promote the oxidation of the products deposited from the erosion jet. It is remarkable that each of the examined materials behaves in a different manner (Fig. 4).

The chromium film contains several oxides up to the one with the deepest degree of oxidation: very unstable chromic anhydride CrO_3, which rapidly decomposes under the conditions of thermal oxidation.

The composite semiconductor iron oxide film radically changes its composition after laser ablation in air. The metallic component completely oxidizes, the oxide components present in the source film oxidize to the maximum degree, and the compounds containing hydroxyl groups appear.

It appear that the plasma favors the oxidation process. As a result, the source chromium film takes on the semiconducting properties, and the source semiconductor film becomes dielectric.

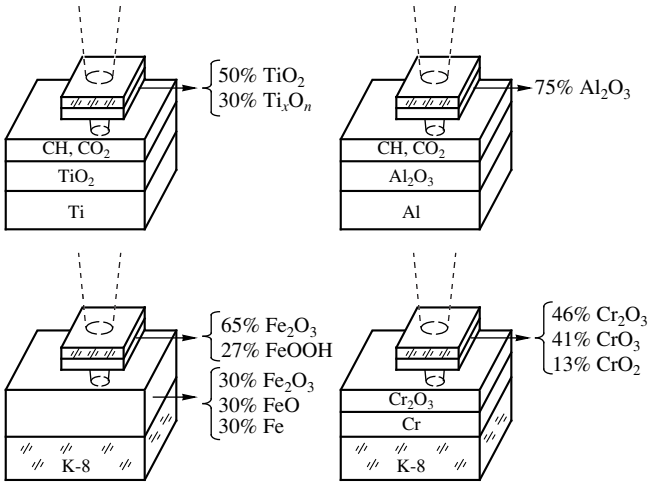


Fig. 4. Chemical compositions of the target and deposited films.

For the massive materials, the tendency is the same but less obvious. Namely, the evaporated films are free of the pure metal (Ti or Al) component. The aluminum fully oxidizes to the higher oxide Al_2O_3, while the titanium may have lower oxides as well.

Thus, laser ablation in air is accompanied by extensive vapor-phase redox processes. Their essential role in the development of instabilities along the scanning path [5] is also confirmed. In fact, a decrease in the volume occupied by the oxide molecules as compared to that of the corresponding atoms decreases the near-surface pressure and causes the backward flow of the vapor-phase-reaction products onto the irradiated surface. In turn, this produces absorptivity feedback; hence, the instabilities along the beam path.

Within the concept put forward in [5], the situation with the oxide films remained unclear: they have initial oxide components but demonstrate distinct structure self-organization. Our data show that all of the components of the iron oxide film oxidize to higher oxides. Thus, the model of reactive laser ablation with partial condensation of products obtained by jet–air interaction applies to this material as well.

CONCLUSIONS

Note that the method of laser back transfer, used to study the chemical composition of the erosion jet, is efficient in controlling the degree of oxidation of films. In [6], the highest degree of oxidation/reduction under instability was achieved on laser-irradiated silicon–molybdenum films. It is noteworthy that the chemical composition of the films formed by reactive laser ablation differs from that of the source film. Moreover, it cannot be obtained under conditions of steady-state thermal oxidation.

It should be noted that the results were obtained for single-pulse target irradiation. Multipulse irradiation may drastically change the situation, which calls for specific investigation.

ACKNOWLEDGMENTS

We are grateful to P.E. Sobolev and V.A. Chuĭko for their assistance.

This work was supported by the Russian Foundation for Basic Research (grant no. 97-02-18289a).

REFERENCES

1. S. I. Kudryashov, A. A. Karabutov, S. V. Kuznetsov, *et al.*, Izv. Ross. Akad. Nauk, Ser. Fiz. **60** (3), 2 (1996).
2. N. Emil and P. E. Sobol, *Phase Transformations and Ablation in Laser-Treated Solids* (Wiley, New York, 1995).
3. V. I. Nefedov, *X-ray Electron Spectroscopy of Chemical Compounds: A Handbook* (Khimiya, Moscow, 1984).
4. J. F. Moulder, W. F. Stickle, P. E. Sobol, and K. D. Bomben, *Handbook of X-ray Photoelectron Spectroscopy* (Physical Electronics, Eden Prairie, 1995).
5. M. N. Libenson, A. V. Lysenko, V. A. Chuĭko, and G. D. Shandybina, Izv. Ross. Akad. Nauk, Ser. Fiz. **61**, 1491 (1997).
6. M. N. Libenson and G. D. Shandybina, Pis'ma Zh. Tekh. Fiz. **21** (5), 9 (1995) [Tech. Phys. Lett. **21**, 168 (1995)].

Translated by B. Malyukov

Technical Physics, Vol. 45, No. 9, 2000, pp. 1223–1225. Translated from Zhurnal Tekhnicheskoĭ Fiziki, Vol. 70, No. 9, 2000, pp. 135–138.

BRIEF COMMUNICATIONS

Local Measurements of Magnetic Characteristics of High-Temperature Superconducting Films

Yu. N. Pirogov and E. A. Protasov

Moscow Engineering Physics Institute (Technical University), Kashirskoe sh. 31, Moscow, 115409 Russia

Received November 15, 1999

Abstract—A new experimental technique for measuring magnetic characteristics of high-temperature superconducting (HTS) films is described. The measurement system includes a laser providing for local measurements. The measured magnetic characteristics of the HTS films are of scientific and practical interest. © *2000 MAIK "Nauka/Interperiodica".*

The magneto-optical investigation technique is currently widely used to study processes of penetration and capture of the magnetic flux in high-temperature superconducting (HTS) films. The operating principle and design of the polarization-optical microscope are described in several papers [1–4].

The capabilities of this method depend primarily on the choice of magneto-optical film (MOF). The first studies utilized domain films in which the easy magnetization axis was perpendicular to the film plane. Thus, only the normal component of the magnetic induction can be observed in such a configuration. In further studies, ferrite-garnet films with "easy plane" anisotropy were used. The easy magnetization axis in these films lies in the film plane or at a certain small angle to it. By using such films, it is possible to observe both the normal and tangential components of the magnetic induction.

The magneto-optical method made it possible to directly observe magnetic-flux penetration into a sample, to investigate the effect of the defect structure in a superconductor on the distribution and value of the captured magnetic flux, and to study the distribution of surface currents. However, in order to comprehensively understand the influence of defects on the processes of interaction of high-temperature superconductors (HTS) with a magnetic field, local measurements of their magnetic characteristics should be made. A diaphragm which cut a part of the pattern was used for this purpose in [5, 6], and the intensity of the light interacting with an MOF and passing through the diaphragm was measured with a photodetector (PD). However, a feature of this system is that, when the region in which the measurements are conducted is reduced, the light intensity incident on the PD decreases because of a decrease in the diaphragm diameter, leading to a deteriorating signal-to-noise ratio.

In this paper, a somewhat different method for local measurements of HTS magnetic characteristics is proposed. The setup described below has a number of advantages. An additional light source (a laser) is used for measurements. Good noise characteristics are achieved with a He–Ne laser with a power of several milliwatts.

Moreover, this setup makes it possible to direct the laser beam to the necessary region of an HTS sample by directly observing the beam adjustment on a monitor screen, and the monitoring can be carried out during the entire experiment.

Finally, the laser beam can be focused, so that measurements can be made in a region with dimensions of 15–20 μm.

SETUP DESIGN

Figure 1 shows a schematic diagram of the setup. It included a nitrogen optical cryostat, a magneto-optical visualizer of magnetic fields, a PD unit, a temperature sensor, a magnetic-field source, and a measurement system. An HTS film was pasted onto a heat sink of the vacuum nitrogen cryostat with Ramsay grease. A reflecting silver layer was deposited by evaporation onto a magneto-optical ferrite-garnet film grown on a substrate of gallium–gadolinium garnet. The magneto-optical transducer manufactured in this way was pressed mechanically to the HTS film. The temperature sensor and heater coil were mounted in the heat sink, making it possible to control and measure the temperature in a wide range (77–300 K).

The optical part of the setup included an OSV-1 white-light source *1*, beam-splitting cubes *2* and *15*, a polarizer *3*, a semitransparent mirror *4*, and an analyzer *5*. After passing through the entrance window *6* of the nitrogen cryostat *7* and passing twice through the MOF *8*, the light beam was reflected by the silver coating and entered the lens *9*, which formed a full or partial MOF image on the array of a video camera *10*. The signal from the video camera was fed to the input of a video

1063-7842/00/4509-1223$20.00

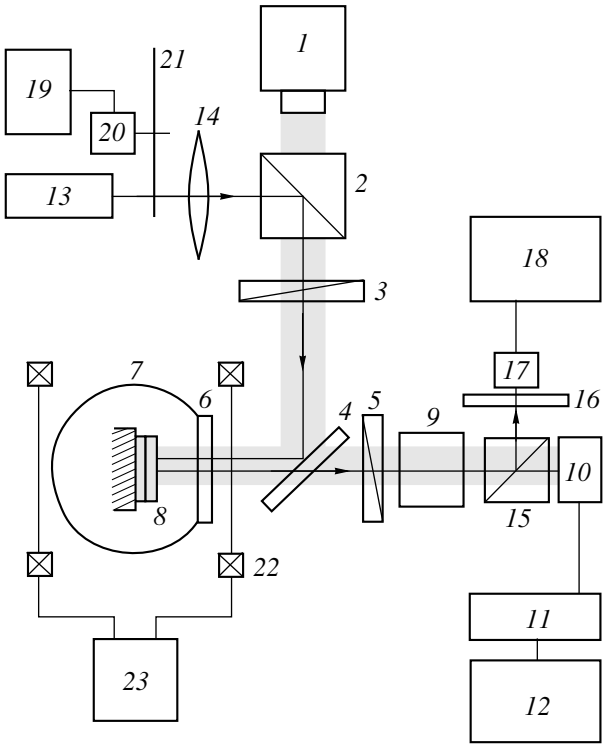


Fig. 1. Schematic diagram of the setup: (*1*) OSV-1 white-light source; (*2*) beam-splitting cube; (*3*) polarizer; (*4*) semitransparent mirror; (*5*) analyzer; (*6*) entrance window; (*7*) nitrogen cryostat; (*8*) magneto-optical film; (*9*) lens; (*10*) video camera; (*11*) video tape recorder; (*12*) TV set; (*13*) He–Ne laser; (*14*) lens; (*15*) beam-splitting cube; (*16*) red-light filter; (*17*) photodetector; (*18*) RV7-32 selective microvoltmeter; (*19*) GZ-33 generator; (*20*) electric motor; (*21*) disk with four holes; (*22*) Helmholtz coils; and (*23*) TEK-5020 power supply or GZ-109 sound generator.

tape recorder *11* and then to a TV set *12* or a PC video port.

The measuring part consisted of the following components. The beam of the He–Ne laser *13* was focused by a lens *14*. Its further path coincided with that of the white-light beam, and, on being reflected by the beam-splitter *15*, the beam passed through a red filter *16*. The laser light intensity was detected by the PD *17*, whose signal was measured by a PB7-32 selective microvoltmeter *18*. The laser beam was modulated using a system consisting of a GZ-33 generator *19*, an electric motor *20*, and a disk *21* with four holes. The modulation frequency was 40 Hz. Thus, it was possible to direct the laser beam to the local HTS region of interest and to control the adjustment procedure by the image on the screen. The PD signal was proportional to the rms density of the magnetic flux in the region determined by the area of the laser-beam cross section. This area was approximately 1 mm in size without additional focusing for possible sample dimensions of 10×10 mm.

The magnetic field H, which was normal to the sample surface, was formed by Helmholtz coils *22* ensuring a field inhomogeneity over the sample of no higher than 0.1%. The coils were energized by a TEK-5020 power supply unit *23* and provided magnetic fields of up to 300 Oe. An exciting ac field with a frequency $f = 20$–1000 Hz was produced by connecting the coils to a GZ-109 sound generator.

EXPERIMENT

Similar to other works, we observed the dynamics of magnetic-flux penetration into the sample. We found that the magnetic flux penetrates into it primarily at the sample edges and at defect locations. When the magnetic field is removed, the flux is captured most strongly in defect-free regions, whereas defective regions capture almost no magnetic flux.

An $Y_1Ba_2Cu_3O_7$ film (0.3 μm thick, deposited by evaporation onto a 0.5-μm-thick strontium titanate substrate using the pulsed laser technique) was studied experimentally. The sample's dimensions were 10×5 mm. Figure 2 shows the magneto-optical pattern of the sample corresponding the state of a superconductor with a captured magnetic flux. Light regions correspond to the areas where the magneto-optical film is magnetized. The three most typical regions were investigated. The strongest capture was observed in regions *1* and *2*, and region *3* was defective.

As an example, we present the results of studying magnetic-flux shielding by an HTS sample. The sample was cooled to $T = 78$ K at $H = 0$. Experiments were performed at an ac magnetic field $H = 2.4$ Oe at $f = 40$ Hz, and the laser beam was not modulated. A signal from the PD was measured by the selective microvoltmeter at the second harmonic.

Figure 3 presents the data obtained for regions *1–3* and the integral characteristic, which were obtained by illuminating the entire sample with the white-light beam with subsequent focusing of the MOF-reflected radiation at the PD. The integral characteristic, which is usually measured, shows that the width of the transition from the state in which the sample fully shields the magnetic field to the state in which the magnetic flux fully penetrates into the sample is ≈4 K (between the levels of 0.1 and 0.9). However, the transition is much narrower for each local region. Thus, the wide transition in the integral characteristic is determined by the fact that the magnetic flux penetrates into different HTS regions at different temperatures and regions with wide transitions also being present.

Let us analyze the dependences for regions *1–3*. Figure 3 shows that the shielding properties of regions *1* and *2* are almost identical. The flux begins penetrating at $T = 88$ K, and the transition width is ≈1 K. In the defective region, the magnetic flux penetrates into the sample at 87 K (i.e., one degree lower compared to regions *1* and *2*), and the transition is wider (≈2 K). This indicates that the shielding properties of the defective region are inferior to those of regions *1* and *2*. Note that,

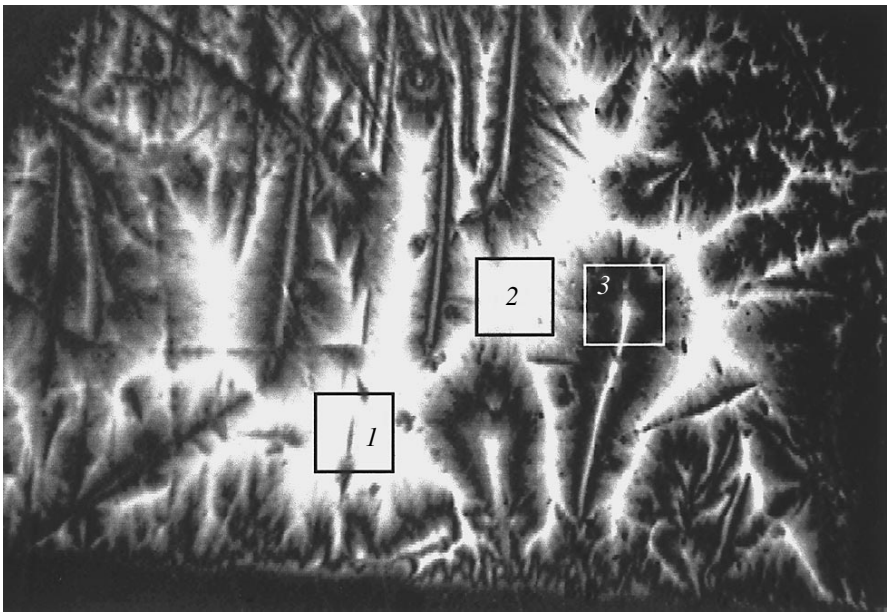


Fig. 2. Magneto-optical pattern of the sample under study corresponding to the state of a superconductor with a captured magnetic flux. The sample's dimensions are 10×5 mm.

in measurements in region *3*, the signal at first rises and, approximately at 88.5 K, begins falling. This occurs due to the fact that, at the moment when region *3* of the HTS stops shielding the magnetic field, there are still some regions in the sample that continue shielding.

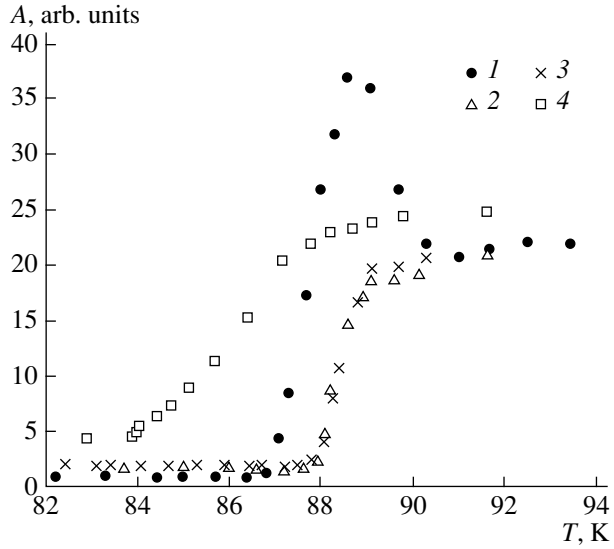


Fig. 3. Amplitude A of the photodetector signal vs. temperature. The signal amplitude is proportional to the r.m.s. value of the magnetic flux density in the region irradiated by the laser beam: (*1*) defect region, (*2*) first region, (*3*) second region, and (*4*) integral characteristic.

Therefore, the magnetic flux concentrates in the defective region. At $T = 89$ K, "stronger" regions stop shielding the field and the magnetic flux spreads over the entire sample. Since regions *1* and *2* have the highest shielding characteristics and are the last to let the magnetic flux penetrate into them, no drops in the curves are observed. A similar behavior of superconductors was observed in [4].

REFERENCES

1. O. S. Esikov and E. A. Protasov, Pis'ma Zh. Tekh. Fiz. **15** (20), 11 (1989) [Sov. Tech. Phys. Lett. **15**, 791 (1989)].
2. A. A. Polyanskiĭ, V. K. Vlasko-Vlasov, M. V. Indenbom, *et al.*, Pis'ma Zh. Tekh. Fiz. **15** (22), 1 (1989) [Sov. Tech. Phys. Lett. **15**, 872 (1989)].
3. A. I. Belyaeva, S. V. Voĭtsenya, V. P. Yur'ev, *et al.*, Sverkhprovodimost: Fiz., Khim., Tekh. **3**, 1189 (1990).
4. V. K. Vlasko-Vlasov, M. V. Indenbom, V. I. Nikitenko, *et al.*, Sverkhprovodimost: Fiz., Khim., Tekh. **3**, 1199 (1990).
5. M. V. Indenbom, N. F. Kolesnikov, M. P. Kulakov, *et al.*, Physica C (Amsterdam) **166**, 486 (1990).
6. V. K. Vlasko-Vlasov, M. V. Indenbom, V. I. Nikitenko, *et al.*, Sverkhprovodimost: Fiz., Khim., Tekh. **5**, 2017 (1992).

Translated by A. Seferov

Technical Physics, Vol. 45, No. 9, 2000, pp. 1226–1229. Translated from Zhurnal Tekhnicheskoĭ Fiziki, Vol. 70, No. 9, 2000, pp. 139–142.

BRIEF COMMUNICATIONS

Extended Temperature Range of Absorptive Optical Bistability in Semiconductors

T. M. Lysak and V. A. Trofimov

Moscow State University, Vorob'evy gory, Moscow, 119899 Russia

Received December 1, 1999

Abstract—Optical bistability based on the temperature dependence of the absorption coefficient in a semiconductor is considered. It is theoretically demonstrated that the upper bound of the thermostat temperatures at which the bistability is possible can be raised by a factor of three as against its conventional value reported in the literature. This can be achieved in the case of the Auger recombination of free electrons in semiconductors. © *2000 MAIK "Nauka/Interperiodica".*

Optical bistability received much attention in the literature, being regarded as a potential basis for creating optical switches (processors) [1–4]. In recent years, we have seen the revival of interest in the phenomenon owing to efforts in the field of "optical" winchesters and the recognition of the fact that the speed of electronic processors is likely to reach its physical limit in the coming years. Accordingly, it seems worthwhile to seek new ways of implementing optical bistability and to analyze the potential of well-known strategies. One of the strategies rests on absorptive optical bistability (AOB), i.e., the temperature dependence of the absorption coefficient. Its main disadvantage is higher power consumption per state transition compared with other types of optical bistability [1, 4]. Another demerit is that AOB requires heat removal, depending on thermostat temperature and other factors. On the other hand, if optical absorption and subsequent heat release are negligible, the switching time may be as short as 10^{-9} s [1]. Furthermore, it is possible to organize parallel operation of many thousands of optical processors. Such a system could compete with existing electronic computers if the power consumption per state transition were sufficiently low. Indeed, a shorter switching time can be achieved in a semiconductor with optical bistability based on the dependence of the absorption coefficient on free electron concentration, compared with bistability based on the temperature shift of the fundamental-absorption edge. However, in the former approach, the problem of heat removal may arise if indirect transitions are involved. In our opinion, the difficulty would be even more serious if exciton transitions were used. Finally, the mathematical model employed in this study allows one to extend the results to the effect of short laser pulses that change the free-electron concentration and vibrational temperature of a semiconductor. Thus, the results reported below may find fairly wide application.

This study demonstrates that the range of the thermostat temperatures at which AOB is possible can be expanded threefold, compared with the well-known values [4]. This considerably alleviates heat removal and reduces power consumption per state transition, since the temperature of the bistable element is allowed to increase appreciably. This is achieved due to nonlinear recombination of free electrons, which is common in practice [5, 6]. The relaxation time of free electrons is assumed to be a linear or quadratic function of the concentration. The results for nonlinear relaxation rate are compared with those for constant relaxation rate, the latter corresponding to a weakly excited semiconductor [5, 6].

Consider the interaction of laser radiation with a thin layer of a semiconductor, assuming that the process goes uniformly across the width. We examine only the near-axis region of the beam, transverse diffusion and heat conduction being approximated by a heat-sink term. This approximation is frequently used [1–4] for qualitatively analyzing such systems. Also, we take into account the optical generation of free electrons from, e.g., an impurity level. It is assumed that the depletion of the level is negligible and that the transition is saturated due to the dynamic Burstein–Moss effect [5, 6]. The free carriers are generated under the action of laser radiation whose wavelength is close to the fundamental absorption edge. Conduction band electrons recombine through indirect transitions, which heats up the semiconductor. It is precisely such semiconductors that are examined in experiments on temperature-dependent AOB.

1063-7842/00/4509-1226$20.00

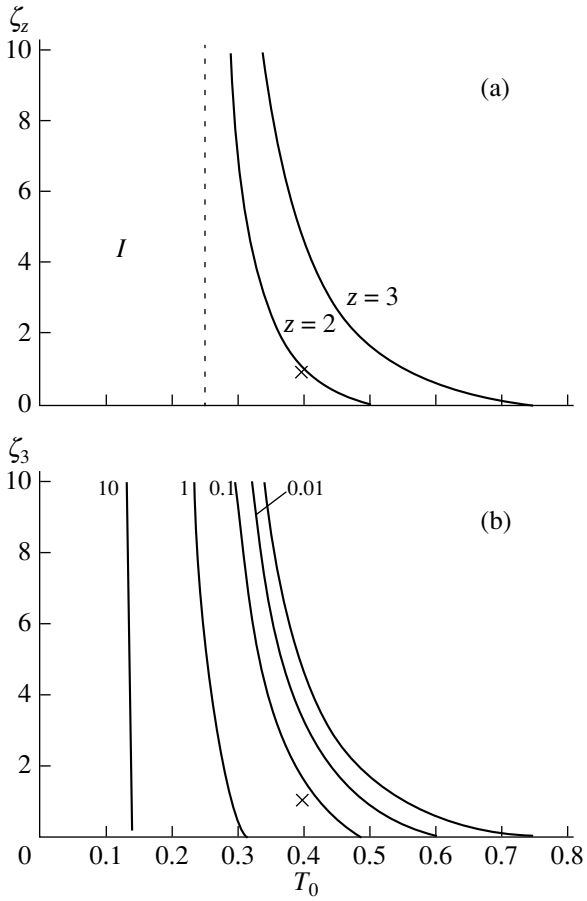


Fig. 1. AOB region I in the (T_0, ζ_z) plane, where $\zeta_z = z(\tau_D/q)(q/\tau_{pm})^{1/z}$. Panel (a) refers to the absence of transition saturation. The solid curves depict $\zeta_z^*(T_0)$ for $z = 2$ and $z = 3$. The broken curve is the boundary of the AOB region in the case of "linear" recombination. Panel (b) refers to the presence of the saturation with $z = 3$. The curves depict $\xi_z^*(T_0)$ for different values of ξ_z, indicated by the numerals.

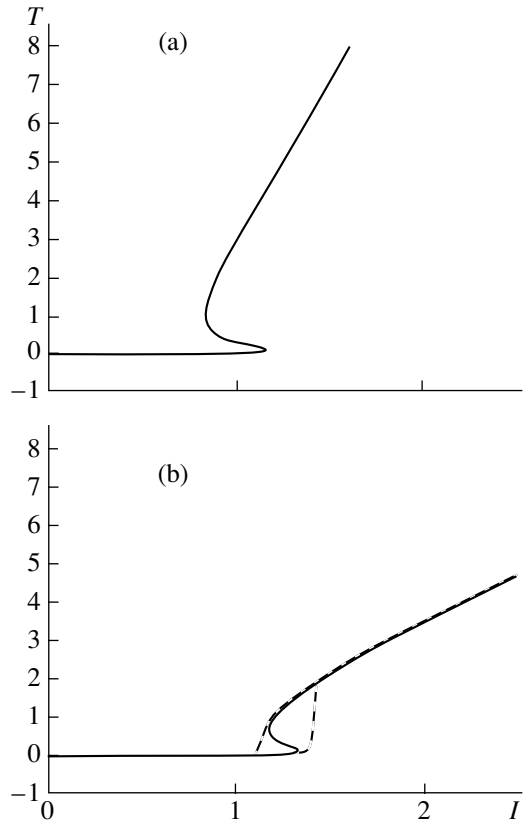


Fig. 2. Bistable T–I characteristics for the (a) absence and (b) presence of the dynamic saturation of the transition. The parameter values, corresponding to the cross in Fig. 1, are $\tau_{pm} = 0.3$, $T_0 = 0.4$, $\beta = 1$, $\tau_D = 1$, and $q = 10$. The solid curves refer to stationary states, whereas the broken curves depict the response to the trapezoidal pulse.

The above assumptions lead to a lumped-parameter model described by the dimensionless equations

$$\frac{dn}{dt} = \delta(n, T)I - \frac{n^z}{\tau_{pm}} - \frac{n}{\tau_D}, \qquad \frac{dT}{dt} = q\frac{n^z}{\tau_{pm}} - T \tag{1}$$

subject to the initial conditions

$$T|_{t=0} = 0, \quad n|_{t=0} = 0. \tag{2}$$

Here, τ_{pm} is the free-carrier relaxation time, and T is the difference between the actual temperature of the semiconductor and its unperturbed value T_0. The temperature difference is normalized to the energy gap expressed in units of temperature. The parameter z controls the dependence of the recombination time on the free-carrier concentration: if $z = 1$, the dependence is flat, and if $z = 3$, it corresponds to Auger recombination.

The heat-sink term in the second equation of system (1) refers to heat removal from the laser-beam axis. The variable t is time normalized to the characteristic time of the heat removal. The variable n is the free-carrier concentration normalized to its maximum attainable value under the stated conditions (at $t = 0$). The quantity q indicates what fraction of absorbed energy is converted into heat during electron recombination. The parameter τ_D refers to the diffusive withdrawal of free carriers from the near-axis irradiated region. The function $I(t)$ describes the waveform of the laser pulse.

Depending on irradiation conditions, the absorption coefficient may have one of the following forms [5, 6]:

$$\delta(n, T) \equiv \begin{cases} e^{-\frac{1}{T - T_0}} & \text{(3a)} \\ (1 - n)e^{-\frac{1}{T + T_0}}. & \text{(3b)} \end{cases}$$

The second form allows for the dynamic saturation of the transition.

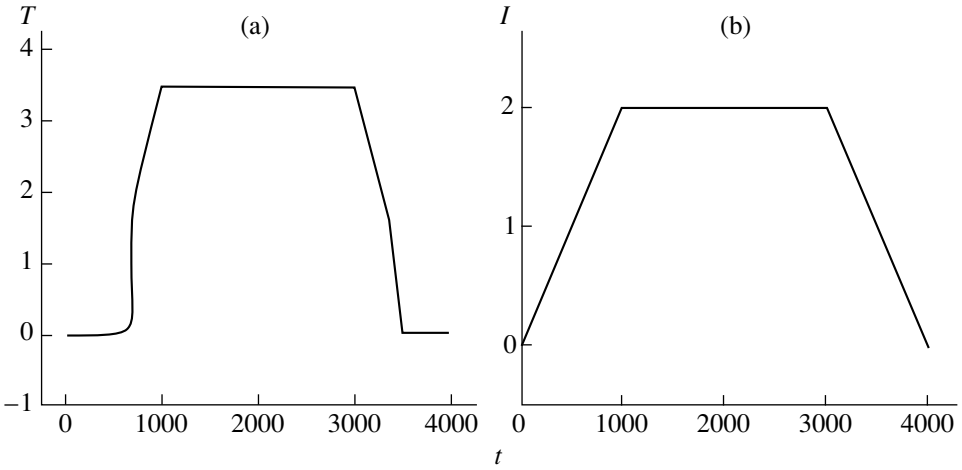


Fig. 3. Effect of a trapezoidal laser pulse: (a) the evolution of the temperature and (b) the intensity waveform of the pulse. The parameter values are the same as in Fig. 2.

With absorption coefficient (3a), AOB arises if

$$\zeta_z = z(\tau_D/q)(q/\tau_{pm})^{1/z} < \zeta_z^*(T_0), \tag{4}$$

where $\zeta_z^*(T_0)$ is the limit value of $\zeta_z(T_0)$ for which the inequality

$$\zeta_z T^{\frac{z-1}{z}}\left(1-\frac{T}{(T+T_0)^2}\right)+\left(1-z\frac{T}{(T+T_0)^2}\right)<0 \tag{5}$$

has a solution.

Inequality (5) indicates that a nonzero value of $\zeta_z^*(T_0)$ exists if the thermostat temperature T_0 meets the constraint

$$T_0 < z/4. \tag{6}$$

Consequently, with a constant relaxation time (z = 1), AOB exists if $T_0 < 0.25$ [4]. Note that the T–I characteristic is bistable for any $z \geq 1$, τ_{pm}, τ_D, and q if the ambient temperature T_0 is below 0.25, which is a well-known fact. More importantly, AOB may occur also for $0.25 \leq T_0 < z/4$ ($0.25 \leq T_0 < 0.75$ at $z = 3$) provided that the interaction parameters satisfy conditions (4), (5), and

$$1 > \frac{\zeta^z}{(\zeta^*(T_0))^z} \equiv \frac{1}{q^{z-1}}\left(\frac{z}{\zeta^*(T_0)}\right)^z\frac{\tau_D^z}{\tau_{pm}}. \tag{7}$$

The dramatic expansion of the AOB region is illustrated by Fig. 1a. It shows the dependences $\zeta_z^*(T_0)$ at z = 2 and z = 3. Note that, in the presence of dynamic transition saturation [with absorption coefficient (3b)], the energy fraction needed q for AOB to occur is larger. For example, at a fixed value of the parameter $\zeta_z = (\tau_{pm}/q)^{1/z}$, AOB exists if condition (4) is met, but $\zeta_z^*(T_0)$ now denotes the limit value of $\zeta_z(T_0)$ for which the inequality

$$\frac{\xi_z T^{1/z}}{1-\xi_z T^{1/z}}\left(\frac{\zeta_z}{z}T^{\frac{z-1}{z}}+1\right)+\zeta_z T^{\frac{z-1}{z}}\left(1-\frac{T}{(T+T_0)^2}\right) + \left(1-z\frac{T}{(T+T_0)^2}\right)<0 \tag{8}$$

has a solution. In (8), the parameter ξ_z refers to the density of the optical energy that is released during free-electron relaxation.

To clarify the picture, Fig. 1b shows the boundaries of the AOB region at z = 3 and different values of ξ_z.

Figure 2 substantiates the possibility of switching by presenting T–I characteristics in the (a) absence and (b) presence of dynamic saturation of the transition for the parameter values corresponding to the cross in Fig. 1. Both of the curves are bistable, but the presence of the saturation is characterized by a smaller difference between the switching temperatures and higher switching intensities.

Switching is also evidenced by Fig. 3a. It depicts the evolution of the temperature T when a trapezoidal pulse (Fig. 3b) is incident on the semiconductor. For convenience, Fig. 2b includes the corresponding T–I dependence (broken curves). The graphs testify to the existence of AOB and switching waves provided that conditions (5) and (6) are satisfied.

To sum up, the nonlinear dependence of recombination on concentration makes it possible to dramatically expand the range of the ambient temperatures at which AOB occurs. This may lead to designing more efficient devices. Also, one can reduce power consumption per state transition, since less power is required to maintain a given temperature of the thermostat.

ACKNOWLEDGMENTS

This study was supported by the Russian Foundation for Basic Research (grant no. 99-01-01233). Support was also received within the program "Universities of Russia: Basic Research Projects."

REFERENCES

1. H. Gibbs, *Optical Bistability: Controlling Light with Light* (Academic, New York, 1985; Mir, Moscow, 1988).
2. *Transverse Pattens in Nonlinear Optics*, Ed. by N. N. Rozanov, Proc. SPIE **1840** (1992).
3. *Optical Computing: Digital and Symbolic*, Ed. by R. Arratoon (Marcel Dekker, New York, 1989; Mir, Moscow, 1993).
4. N. N. Rozanov, *Optical Bistability and Hysteresis in Distributed Nonlinear Systems* (Nauka, Moscow, 1997).
5. R. A. Smith, *Semiconductors* (Cambridge Univ. Press, Cambridge, 1978; Mir, Moscow, 1982).
6. V. L. Bonch-Bruevich and S. G. Kalashnikov, *The Physics of Semiconductors* (Nauka, Moscow, 1990).

Translated by A. Sharshakov