Technical Physics, Vol. 47, No. 6, 2002, pp. 655–659. Translated from Zhurnal Tekhnicheskoĭ Fiziki, Vol. 72, No. 5, 2002, pp. 1–4.

THEORETICAL AND MATHEMATICAL PHYSICS

Transition and Diffraction Radiations from a Charge Particle Interacting with an Anisotropically Conducting Ball

I. I. Kalikinskiĭ

Astrakhan State Pedagogical University, Astrakhan, 414056 Russia

Received August 27, 2001

Abstract—The problems of transition radiation from a charged particle interacting with an anisotropically conducting ball and of diffraction radiation from a charged particle flying close to this ball are solved. The energy, spectra, and polarization of the emitted radiation are determined. *© 2002 MAIK "Nauka/Interperiodica".*

INTRODUCTION

At present, there are many papers devoted to transition radiation, which was discovered as early as 1946 by Ginzburg and Frank [1] (an extensive bibliography on this problem is given in [2]). The problem of transition radiation from a charged particle interacting with an anisotropically conducting plane (straight-line anisotropy) was solved by Barsukov and Naryshkina [3]. The problem of transition radiation from a charged particle interacting with a ring grid and a radial grid (cylindrical anisotropy) was addressed in my earlier papers [4, 5]. It is also of interest to consider the problem of transition radiation in the interaction between a charged particle and a ball with a spherically anisotropic conductivity. This problem, being obviously important from a theoretical standpoint, has, at the same time, significant astrophysical applications (the author is grateful to Ya.N. Istomin for drawing attention to the latter point).

Thus, the problem of transition and diffraction radiations from a charged particle interacting with a radially conducting ball refers, in particular, to the study of active galactic nuclei, which are, according to present opinion, black holes surrounded by accretion disks, whose magnetic fields compress the magnetic fields of the hole to nearly the field of a monopole (the magnetic induction being on the order of 10^4 G). In such a field, charged particles move freely along the magnetic lines (the conductivity in the radial direction is nearly infinite) and cannot propagate across the lines (the conductivity along spherical surfaces is nearly zero). Hence, transition and diffraction radiations associated with active galactic nuclei can be described using a radially conducting ball model, which can also be applied to magnetized neutron stars, whose magnetic fields, although dipolar, have an induction on the order of 10^{12} G.

The objective of this paper is to investigate the problem of transition and diffraction radiations from a charged particle interacting with a ball with a spherically anisotropic conductivity.

FORMULATION OF THE PROBLEM AND SOLUTION OF THE INHOMOGENEOUS VECTOR WAVE EQUATION

We consider a ball with a spherically anisotropic conductivity, in which case we must distinguish between the following two possible cases. In the first case, a ball consists of nested, electrically insulated, perfectly conducting spherical layers. It is easily understood that, with respect to an external field, such a ball is equivalent to a perfectly conducting ball. The corresponding boundary condition is

$$E_\Theta\big|_{R = R_0} = 0, \tag{1}$$

where R_0 is the ball radius.

In the second case, the conductivity of the ball is infinite in the radial direction, but it is zero in the transverse directions. The corresponding boundary condition has the form

$$E_R\big|_{R = R_0} = 0. \tag{2}$$

Let a point particle with charge q move parallel to the z-axis with velocity $\mathbf{v}(0, 0, -v)$. The moving charged particle generates a current with density $\mathbf{j}(0, 0, -j)$, where

$$j = qv\delta(x - x_0)\delta(y - y_0)\delta(z + vt) \tag{3}$$

and it can be assumed that $y_0 = 0$. The particle trajectory intersects the $X0Y$ plane at the point $M_0(x_0, y_0)$, in which case we have $r_0 = \sqrt{x_0^0 + y_0^0} = R_0\sin\Theta_0$ (with $0 \le \Theta \le \pi/2$) for transition radiation and $r_0 > R_0$ for diffraction radiation. We denote by $\mathbf{r}$ the distance from the point $M(x, y)$ to the axis of a cylindrical coordinate system and obtain

$$\mathbf{M}_0\mathbf{M} = \mathbf{r}' = \mathbf{r} - \mathbf{r}_0$$

and

$$r' = \sqrt{r^2 + r_0^2 - 2rr_0\cos\varphi}, \tag{4}$$

where φ is the polar angle.

1063-7842/02/4706-0655$22.00

We expand the current density in a Fourier integral, e.g.,

$$j = \int_{-\infty}^{\infty} j_\omega e^{-i\omega t} d\omega, \tag{5}$$

with the Fourier coefficients

$$j_\omega = \frac{q}{2\pi} e^{-i\frac{\omega}{v}z} \frac{\delta(r')}{2\pi r'}. \tag{6}$$

Using the formula [6]

$$\frac{\delta(r')}{r'} = \int_0^{\infty} J_0(\lambda r')\lambda d\lambda \tag{7}$$

and the addition theorem for Bessel functions [7]

$$J_0(\lambda r') = \sum_{m=-\infty}^{\infty} J_m(\lambda r) J_m(\lambda r_0) e^{im\varphi}, \tag{8}$$

we obtain

$$j_\omega = \frac{q}{4\pi^2} e^{-i\frac{\omega}{v}z} \sum_{m=-\infty}^{\infty} e^{im\varphi} \int_0^{\infty} J_m(\lambda r) J_m(\lambda r_0)\lambda d\lambda. \tag{9}$$

In order to solve the electrodynamic problem as formulated, we start with the following Maxwell equations for the Fourier coefficients:

$$\nabla \times \mathbf{E}_\omega = \frac{i\omega}{c}\mathbf{H}_\omega, \quad \nabla \times \mathbf{H}_\omega = \frac{4\pi}{c}\mathbf{j}_\omega - \frac{i\omega}{c}\mathbf{E}_\omega. \tag{10}$$

Choosing H_r and H_φ as potentials ($H_z = 0$), we solve Eqs. (10) in cylindrical coordinates. The potentials satisfy the equation

$$\Delta \mathbf{H}_\omega + \frac{\omega^2}{c^2}\mathbf{H}_\omega = -\frac{4\pi}{c}\nabla \times \mathbf{j}_\omega \tag{11}$$

or, equivalently, the equations

$$(\Delta \mathbf{H}_\omega)_\varphi + \frac{\omega^2}{c^2} H_{\omega\varphi} = \frac{4\pi}{c}\frac{\partial j_\omega}{\partial r}, \tag{12}$$

$$(\Delta \mathbf{H}_\omega)_r + \frac{\omega^2}{c^2} H_{\omega r} = -\frac{4\pi}{c}\frac{1}{r}\frac{\partial j_\omega}{\partial \varphi}, \tag{13}$$

$$(\nabla \times \mathbf{j}_\omega)_z = 0, \quad H_{\omega z} = 0.$$

Using the recurrence formulas for Bessel functions [7], we arrive at

$$\frac{4\pi}{c}\frac{\partial j_\omega}{\partial r} = \frac{q}{2\pi c} e^{-i\frac{\omega}{v}z} \sum_{m=-\infty}^{\infty} e^{im\varphi} \times \int_0^{\infty} J_m(\lambda r_0)[J_{m-1}(\lambda r) - J_{m+1}(\lambda r)]\lambda^2 d\lambda, \tag{14}$$

$$\frac{-4\pi}{c}\frac{1}{r}\frac{\partial j_\omega}{\partial \varphi} = -\frac{iq}{2\pi c} e^{-i\frac{\omega}{v}z} \sum_{m=-\infty}^{\infty} e^{im\varphi} \times \int_0^{\infty} J_m(\lambda r_0)[J_{m-1}(\lambda r) + J_{m+1}(\lambda r)]\lambda^2 d\lambda. \tag{15}$$

We represent the sought-for solution to the inhomogeneous equations (12) and (13) in the form

$$H^{(0)}_{\omega\varphi} = \frac{q}{2\pi c} e^{-i\frac{\omega}{v}z} \sum_{m=-\infty}^{\infty} e^{im\varphi} \times \int_0^{\infty} [B_m(\lambda) J_{m+1}(\lambda r) + C_m(\lambda) J_{m-1}(\lambda r)]\lambda d\lambda, \tag{16}$$

$$H^{(0)}_{\omega r} = \frac{iq}{2\pi c} e^{-i\frac{\omega}{v}z} \sum_{m=-\infty}^{\infty} e^{im\varphi} \times \int_0^{\infty} [B_m(\lambda) J_{m+1}(\lambda r) - C_m(\lambda) J_{m-1}(\lambda r)]\lambda d\lambda. \tag{17}$$

Turning to the familiar relationships [8]

$$(\Delta \mathbf{H}_\omega)_\varphi = \Delta H_{\omega\varphi} - \frac{H_{\omega\varphi}}{r^2} + \frac{2}{r^2}\frac{\partial H_{\omega r}}{\partial \varphi}, \tag{18}$$

$$(\Delta \mathbf{H}_\omega)_r = \Delta H_{\omega r} - \frac{H_{\omega r}}{r^2} - \frac{2}{r^2}\frac{\partial H_{\omega\varphi}}{\partial \varphi}, \tag{19}$$

we solve Eqs. (12) and (13) in the same way as was done in [4, 5]. As a result, we obtain

$$B_m(\lambda) = -C_m(\lambda) = \frac{\lambda J_m(\lambda r_0)}{\lambda^2 + \frac{\omega^2}{v^2}(1-\beta^2)}, \tag{20}$$

where $\beta = v/c$.

Thus the set of inhomogeneous equations (12) and (13) is solved. The solution to the electrodynamic problem is sought for in the form

$$\mathbf{H}_\omega = \mathbf{H}^{(0)}_\omega + \mathbf{H}^{(1)}_\omega, \tag{21}$$

where $\mathbf{H}^{(0)}_\omega$ is the magnetic field of a charged particle (a solution to the inhomogeneous vector wave equation) and $\mathbf{H}^{(1)}_\omega$ is the magnetic field of transition or diffraction radiation (a solution to the homogeneous vector wave equation).

The latter solution was obtained in [4, 5]:

$$H_{\omega\varphi}^{(1)} = \frac{q}{2\pi c}\sum_{m=-\infty}^{\infty} e^{im\varphi} \times \int_0^\infty [\tilde{\mathscr{D}}_m(\lambda) J_{m+1}(\lambda r) + \tilde{\mathscr{E}}_m(\lambda) J_{m-1}(\lambda r)] e^{i\kappa z}\lambda d\lambda, \tag{22}$$

$$H_{\omega r}^{(1)} = \frac{iq}{2\pi c}\sum_{m=-\infty}^{\infty} e^{im\varphi} \times \int_0^\infty [\tilde{\mathscr{D}}_m(\lambda) J_{m+1}(\lambda r) - \tilde{\mathscr{E}}_m(\lambda) J_{m-1}(\lambda r)] e^{i\kappa z}\lambda d\lambda, \tag{23}$$

where $\kappa = \sqrt{(\omega^2/c^2) - \lambda^2}$. For $z > 0$, we have $\mathrm{Im}\,\kappa > 0$, and for $z < 0$, we must replace the tilde ~ by the double tilde ≈ and $e^{i\kappa z}$ by $e^{-i\kappa z}$.

The coefficients $\tilde{\mathscr{D}}_m(\lambda)$ and $\tilde{\mathscr{E}}_m(\lambda)$ can be found from the boundary conditions (1) and (2) and the condition

$$\nabla \cdot \mathbf{H}_\omega^{(1)} = 0. \tag{24}$$

SOLUTION TO THE PROBLEM OF TRANSITION AND DIFFRACTION RADIATIONS FROM A CHARGED PARTICLE INTERACTING WITH A PERFECTLY CONDUCTING BALL

From condition (24), we obtain

$$\tilde{\mathscr{D}}_m(\lambda) = -\tilde{\mathscr{E}}_m(\lambda)[=\mathscr{D}_m(\lambda)], \tag{25}$$

$$\tilde{\tilde{\mathscr{D}}}_m(\lambda) = -\tilde{\tilde{\mathscr{E}}}_m(\lambda). \tag{26}$$

The boundary condition (1) has the form

$$(E_{\omega r}\cos\Theta - E_{\omega z}\sin\Theta)_{R=R_0} = 0, \tag{27}$$

where

$$E_{\omega r} = \frac{ic}{\omega}\left(-\frac{\partial H_{\omega\varphi}}{\partial z}\right),$$
$$E_{\omega z} = \frac{ic}{\omega}\left[\frac{1}{r}\frac{\partial}{\partial r}(rH_{\omega\varphi}) - \frac{1}{r}\frac{\partial H_{\omega r}}{\partial \varphi}\right], \tag{28}$$

$$H_{\omega\varphi} = H_{\omega\varphi}^{(0)} + H_{\omega\varphi}^{(1)}, \quad H_{\omega r} = H_{\omega r}^{(0)} + H_{\omega r}^{(1)}. \tag{29}$$

We substitute expressions (29) for the Fourier coefficients $H_{\omega\varphi}$ and $H_{\omega r}$ into relationships (27) and (28) and pass over from cylindrical coordinates to spherical ones. As a result, we obtain

$$\begin{aligned} E_{\omega\Theta m} = {} & -\frac{ic}{\omega}B_m(\lambda)e^{-i\frac{\omega}{v}R\cos\Theta}\left\{\frac{2i\omega}{v}J_m'(\lambda R\sin\Theta)\cos\Theta\right. \\ & - 2\lambda J_m''(\lambda R\sin\Theta)\sin\Theta - \frac{2}{R\sin\Theta}J_m'(\lambda R\sin\Theta)\sin\Theta \\ & \left. + \frac{2m^2}{\lambda(R\sin\Theta)^2}J_m(\lambda R\sin\Theta)\sin\Theta\right\} \\ & + \frac{ic}{\omega}D_m(\lambda)e^{i\kappa R\cos\Theta}\left\{2i\kappa J_m'(\lambda R\sin\Theta)\cos\Theta\right. \\ & + 2\lambda J_m''(\lambda R\sin\Theta)\sin\Theta + \frac{2}{R\sin\Theta}J_m'(\lambda R\sin\Theta)\sin\Theta \\ & \left. - \frac{2m^2}{\lambda(R\sin\Theta)^2}J_m(\lambda R\sin\Theta)\sin\Theta\right\}. \end{aligned} \tag{30}$$

We expand the quantities $E_{\omega\Theta m}$ in a series in associated Legendre functions:

$$E_{\omega\Theta m} = -\frac{ic}{\omega}\sum_{n=0}^{\infty} e_{mn}(R)\mathscr{P}_n^{(m)}(\cos\Theta). \tag{31}$$

Then, the boundary condition becomes

$$e_{mn}(R_0) = 0. \tag{32}$$

From this condition, we determine the unknown coefficients $D_{mn}(\lambda)$:

$$D_{mn}(\lambda) = B_m(\lambda)\frac{\mathscr{G}_{mn}(\lambda)}{\mathscr{F}_{mn}(\lambda)}, \tag{33}$$

where

$$\begin{aligned} \mathscr{F}_{mn}(\lambda) = {} & \int_0^\pi \mathscr{P}_n^{(m)}(\cos\Theta)\sin\Theta d\Theta e^{i\kappa R_0\cos\Theta} \\ & \times\left\{2i\kappa J_m'(\lambda R_0\sin\Theta)\cos\Theta + 2\lambda J_m''(\lambda R_0\sin\Theta)\sin\Theta\right. \\ & + \frac{2}{R_0\sin\Theta}J_m'(\lambda R_0\sin\Theta)\sin\Theta \\ & \left. - \frac{2m^2}{\lambda(R_0\sin\Theta)^2}J_m(\lambda R_0\sin\Theta)\sin\Theta\right\}, \end{aligned} \tag{34}$$

$$\mathscr{G}_{mn}(\lambda) = \int_0^\pi \mathscr{P}_n^{(m)}(\cos\Theta)\sin\Theta d\Theta e^{-i\frac{\omega}{v}R_0\cos\Theta}$$

$$\times\left\{\frac{2i\omega}{v}J'_m(\lambda R_0\sin\Theta)\cos\Theta - 2\lambda J''_m(\lambda R_0\sin\Theta)\sin\Theta\right. \quad (35)$$
$$-\frac{2}{R_0\sin\Theta}J'_m(\lambda R_0\sin\Theta)\sin\Theta$$
$$\left.+\frac{2m^2}{\lambda(R_0\sin\Theta)^2}J_m(\lambda R_0\sin\Theta)\sin\Theta\right\}.$$

We thus arrive at the desired coefficients:

$$\mathcal{D}_{mn}(-\lambda) = B_m(-\lambda)\frac{\mathcal{G}_{nm}(-\lambda)}{\mathcal{F}_{mn}(-\lambda)} = (-1)^{m+1}\mathcal{D}_{mn}(\lambda). \quad (36)$$

SOLUTION TO THE PROBLEM OF TRANSITION AND DIFFRACTION RADIATIONS FROM A CHARGED PARTICLE INTERACTING WITH A RADIALLY CONDUCTING BALL

For a radially conducting ball, the corresponding boundary condition has the form

$$(E_{\omega r}\sin\Theta + E_{\omega z}\cos\Theta)_{R=R_0} = 0, \quad (37)$$

where the Fourier coefficients $E_{\omega r}$ and $E_{\omega z}$ can be determined from formulas (28).

The coefficients $D_{mn}(\lambda)$ can be obtained in the same way as was done in the previous section:

$$\mathcal{D}_{mn}(\lambda) = B_m(\lambda)\frac{\overline{\mathcal{G}}_{mn}(\lambda)}{\overline{\mathcal{F}}_{mn}(\lambda)}, \quad (38)$$

where

$$\overline{\mathcal{G}}_{mn}(\lambda) = \int_0^\pi \mathcal{P}_n^{(m)}(\cos\Theta)\sin\Theta d\Theta e^{-i\frac{\omega}{v}R_0\cos\Theta}$$

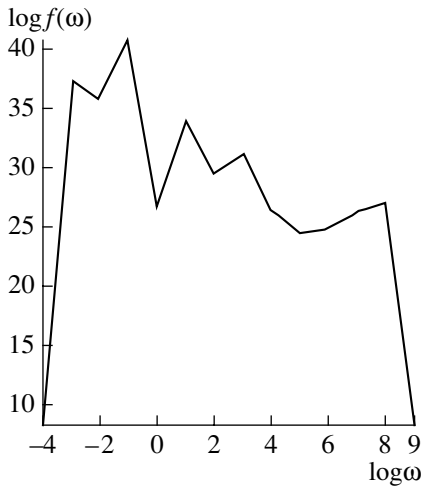


Figure.

$$\times\left\{\frac{2i\omega}{v}J'_m(\lambda R_0\sin\Theta)\sin\Theta + 2\lambda J''_m(\lambda R_0\sin\Theta)\cos\Theta\right. \quad (39)$$
$$+\frac{2}{R_0\sin\Theta}J'_m(\lambda R_0\sin\Theta)\cos\Theta$$
$$\left.-\frac{2m^2}{\lambda(R_0\sin\Theta)^2}J_m(\lambda R_0\sin\Theta)\cos\Theta\right\},$$

$$\overline{\mathcal{F}}_{mn}(\lambda) = \int_0^\pi \mathcal{P}_n^{(m)}(\cos\Theta)\sin\Theta d\Theta e^{i\kappa R_0\cos\Theta}$$

$$\times\left\{2i\kappa J'_m(\lambda R_0\sin\Theta)\sin\Theta - 2\lambda J''_m(\lambda R_0\sin\Theta)\cos\Theta\right. \quad (40)$$
$$-\frac{2}{R_0\sin\Theta}J'_m(\lambda R_0\sin\Theta)\cos\Theta$$
$$\left.+\frac{2m^2}{\lambda(R_0\sin\Theta)^2}J'_m(\lambda R_0\sin\Theta)\cos\Theta\right\}.$$

We thus arrive at the relationship

$$\mathcal{D}_{mn}(-\lambda) = (-1)^{m+1}\mathcal{D}_{mn}(\lambda). \quad (41)$$

CALCULATION OF THE RADIATION INTENSITY

In expressions (22) and (23), we switch from Bessel functions to Hankel functions. Using the circumvention rules for Hankel functions and their asymptotics [7] and, then, applying the saddle point method, we obtain the following expression for the spectral radiation intensity in the wave zone, where $H^{(1)}_{\omega\varphi mn} = E^{(1)}_{\omega\Theta mn}$:

$$\frac{dW_\omega}{d\Omega} = c\,\mathrm{Re}E^{(1)}_{\omega\Theta mn}H^{(1)*}_{\omega\varphi mn}R^2$$
$$= \frac{q^2\omega^2}{\pi^2c^3}\cos^2\Theta\left|\mathcal{D}_{mn}\left(\frac{\omega}{c}\sin\Theta\right)\right|^2. \quad (42)$$

The field in the wave zone is linearly polarized; moreover, the electric vector lies in a plane determined by the *z*-axis and the position vector of the observation point.

ANALYSIS OF THE RESULTS OBTAINED

Formulas (20), (33), (38), and (42) imply that the radiation intensity falls off to zero as $r_0 \longrightarrow \infty$ and also as $R_0 \longrightarrow 0$. It is expedient to present the results of a calculation of the radiation intensity for a black hole with a mass of 10^8 solar masses ($R_0 = 3\times10^{11}$ m) and a charged particle with different energies: 10^{10}, 10^{11}, …,

Table

$\log\omega$	$f(\omega)$
1	3.2×10^{19}
2	1.4×10^{15}
3	1.5×10^{15}
4	1.6×10^{7}
5	2.0×10^{6}
6	3.5×10^{3}
7	1.4×10^{2}
8	0.95

10^{20} eV. It should be noted that the frequency at which the radiation intensity is the highest ($\omega \sim 10$ s^{-1}) is independent of the particle energy.

The values of the function

$$f(\omega) = \frac{4\pi^2 c^5}{q^2 v^4}\frac{dw_\omega}{d\Omega}$$

calculated for $n = 0$, $m = 0$, $\Theta = \pi/4$, and $\tau_0 = R_0$ are summarized in the table. The plot of the function $f(\omega)$ calculated for $\omega = 10, 10^2, \ldots, 10^8$ s^{-1} is shown in the figure.

ACKNOWLEDGMENTS

I would like to thank Ya.N. Istomin for bringing to my attention the astrophysical applications of the results obtained.

REFERENCES

1. V. L. Ginzburg and I. M. Frank, Zh. Éksp. Teor. Fiz. **16**, 15 (1946).
2. L. A. Vardanyan and I. G. Melkumova, in *Bibliography of Works on Transient Radiation of Charged Particles (1945–1982)*, Ed. by G. M. Garibyan (Erevanskiĭ Fiz. Inst., Yerevan, 1983).
3. K. A. Barsukov and L. G. Naryshkina, Radiofizika **8**, 936 (1965).
4. I. I. Kalikinskiĭ, Zh. Tekh. Fiz. **61** (9), 20 (1991) [Sov. Phys. Tech. Phys. **34**, 978 (1991)].
5. I. I. Kalikinskiĭ, Zh. Tekh. Fiz. **65** (10), 131 (1995) [Tech. Phys. **40**, 1047 (1995)].
6. D. Ivanenko and A. Sokolov, *Classical Theory of Field* (Gostekhizdat, Moscow, 1949).
7. I. S. Gradshteyn and I. M. Ryzhik, *Table of Integrals, Series, and Products* (Fizmatgiz, Moscow, 1963; Academic, New York, 1980).
8. L. D. Landau and E. M. Lifshitz, *Electrodynamics of Continuous Media* (Nauka, Moscow, 1982; Pergamon, New York, 1984).

Translated by G. Shepekina

Technical Physics, Vol. 47, No. 6, 2002, pp. 660–665. Translated from Zhurnal Tekhnicheskoĭ Fiziki, Vol. 72, No. 5, 2002, pp. 5–10.

THEORETICAL AND MATHEMATICAL PHYSICS

Focusing of an X-ray Spherical Wave by a Uniaxially Bent Crystal: Johann Scheme and Logarithmic Spiral

T. Tchen

Lomonosov State Academy of Fine Chemical Technology, pr. Vernadskogo 86, Moscow, 117571 Russia

e-mail: docent 65@mtu-net.ru, ttchen@e-mail.ru

Received September 14, 2001

Abstract—Focusing of a spherical wave in the well-known Johann scheme is considered theoretically. An analytical formula for the size of the crystal bent surface reflecting radiation in the Johann scheme is derived. The intensity distribution near the focus is found. The spherical aberration of the diffracted beam is analyzed. Back scattering is shown to minimize the aberration. Spectral characteristics of an Johann spectrometer are discussed. The focusing of a spherical wave by a crystal bent into a logarithmic spiral is considered. It is shown that the Johann scheme is a specific case of a logarithmic spiral. *© 2002 MAIK "Nauka/Interperiodica".*

INTRODUCTION

X-ray radiation focusing with the Johann scheme [1] is a most promising way of producing beams with a large collection solid angle. In this scheme, reflecting planes make the same angle with plane harmonics of an incident spherical wave at any point of the crystal surface. The spatial and spectral parameters of X-ray beams obtained with the Johann scheme have been studied theoretically and experimentally in a number of works [2–8]. Specifically, the theory of a Johann focusing spectrometer has been elaborated in [2, 3, 7]. However, the spherical aberration of the diffracted beam has not been considered as yet.

In this work, we took into account spherical aberration and came to results differing from those obtained in [2, 3, 7].

1. DIFFRACTION REFLECTION RANGE IN THE JOHANN SCHEME

Consider X-ray spherical wave focusing by a uniaxially bent crystal in the Bragg geometry [9]. We assume that the diffraction reflection of the wave is coherent elastic; that is, $\mathbf{k}_0 + \mathbf{h} = \mathbf{k}_h$, $\mathbf{k}_h^2 = \mathbf{k}_0^2 = k^2$. Here, $\mathbf{k}_0$ and $\mathbf{k}_h$ are the wave vectors of the incident and diffracted waves. The vector $\mathbf{h}$ of the reciprocal lattice of a bent crystal is given by

$$\mathbf{h}(h_x, h_y, h_z) = \mathbf{h}_0 - \nabla(\mathbf{h}_0\mathbf{u}), \qquad (1)$$

where $\mathbf{h}_0$ is the vector of the undistorted crystal: $h_0 = 2k\sin\Theta_B$, $k = 2\pi/\lambda$, λ is the wavelength of the incident radiation, and Θ_B is the Bragg angle. Parenthesized vectors mean the scalar product.

The atomic displacement vector $\mathbf{u}(u_x, u_y, u_z)$ for an elastically bent crystal has the form

$$u_x = -xz/R_x, \quad u_z = x^2/2R_x, \qquad (2)$$

where R_x is the bending radius of the crystal.

In (2), terms quadratic in z are neglected. We will consider only those beams emitted by the source S in the plane of diffraction scattering of the crystal (Fig. 1). Unlike [9], the second- and third-order terms in x will be left in the expressions for the direction cosines γ_{0x} and γ_{0z} of the wave vector $\mathbf{k}_0$ (x is the point of incidence of an arbitrary beam on the X axis). The related expres-

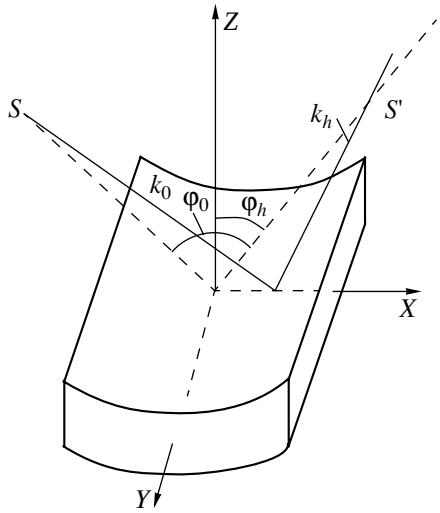


Fig. 1. One-dimensional focusing of a spherical wave by a uniaxially bent crystal. *S*, point source; *S*', its image.

1063-7842/02/4706-0660$22.00

sions are then written as follows:

$$\begin{aligned}\gamma_{0x} &= \sin\varphi_0 + x\cos^2\varphi_0/L_0 - x^2\sin\varphi_0\cos^2\varphi_0/L_0^2 \\ &- x^2\sin\varphi_0\cos\varphi_0(\cos\varphi_0 - L_0/R_x)/2L_0^2 \\ &+ x^3\sin^2\varphi_0\cos^2\varphi_0/L_0^3 - x^3\cos^3\varphi_0(\cos\varphi_0 \\ &- L_0/R_x)/2L_0^3 + x^3\sin^2\varphi_0\cos\varphi_0(\cos\varphi_0 - L_0/R_x)/L_0^3,\end{aligned} \quad (3)$$

$$\begin{aligned}\gamma_{0z} &= -\cos\varphi_0 + x\sin\varphi_0\cos\varphi_0/L_0 + x^2/2R_xL_0 \\ &- x^2\sin^2\varphi_0\cos\varphi_0/L_0^2 + x^2\cos^2\varphi_0(\cos\varphi_0 - L_0/R_x)/2L_0^2 \\ &- x^3\sin\varphi_0/2R_xL_0^2 + x^3\sin^3\varphi_0\cos\varphi_0/L_0^3 \\ &- x^3\sin\varphi_0\cos^2\varphi_0(\cos\varphi_0 - L_0/R_x)/2L_0^3 \\ &- x^3\sin\varphi_0\cos^2\varphi_0(\cos\varphi_0 - L_0/R_x)/L_0^3,\end{aligned} \quad (4)$$

$$\begin{aligned}h_x &= -2k\sin\Theta_B(x\cos\Psi/R_x + \sin\Phi), \\ h_z &= 2k\sin\Theta_B(\cos\Psi - x\sin\Psi/R_x).\end{aligned} \quad (5)$$

Here, Ψ is the angle between the reflecting planes and the crystal surface.

From Fig. 2, the angle between the beam incident from the point P and the normal to the surface at this point is given by

$$\Omega = \pi/2 - (\Theta_B + \delta\Theta).$$

Taking into account that

$$\begin{aligned}\cos\Omega &= -(\mathbf{k}_0\mathbf{h})/kh, \\ \gamma_{0x} &= (x + L_0\sin\varphi_0)/|SP|, \\ \gamma_{0z} &= (z - L_0\cos\varphi_0)/|SP|\end{aligned} \quad (6)$$

and putting $\delta\Theta = \Delta\Theta = |\chi_{hr}|/\sin 2\Theta_B$, we come to an expression for the range x_{eff} of symmetrical diffraction on the crystal surface (in this range, the diffracted beam does not fall outside the region of total reflection):

$$x_{\text{eff}} = \Delta\Theta\left|\sin\Theta_B/L_0 - 1/R_x\right|^{-1}, \quad L_0 \neq R_x\sin\Theta_B. \quad (7)$$

For the Johann scheme ($L_0 = R_x\sin\Theta_B$), we find with (3)–(6)

$$x_{\text{eff, Johann}} = R_x(2\Delta\Theta\tan\Theta_B)^{1/2}. \quad (8)$$

Here, $\Delta\Theta$ is the half-width of the Bragg reflection curve and χ_{hr} is the Fourier component of the X-ray polarizability. From (7) and (8), it is seen that the location on the Rowland circle in the Johann scheme allows one to increase considerably (by two or three orders of magnitude) the collection angle in comparison with the case $L_0 \neq R_x\sin\Theta_B$.

Now we will seek for an expression for the distance $|SP|$ (Fig. 2) as a power series in x. Consider two cases. In case (a), we include the curvature of the crystal surface, assuming that $z \approx x^2/R_x$. Then, it is easy to see that

$$\begin{aligned}|SP| &\cong L_0[1 + x\sin\varphi_0/L_0 + x^2\cos^2\varphi_0/2L_0^2 \\ &- x^2\cos\varphi_0/2L_0R_x - \sin\varphi_0(1/L_0 - \cos\varphi_0/R_x)x^3/2L_0^2 \\ &+ \sin^3\varphi_0x^3/2L_0^3].\end{aligned} \quad (9)$$

For $L_0 = R_x\cos\varphi_0$, the quadratic and cubic terms vanish; hence, fourth-order terms in x must be taken into account. Similarly,

$$\begin{aligned}|S'P| &\cong L_h[1 + x\sin\varphi_h/L_h + x^2\cos^2\varphi_h/2L_h^2 \\ &- x^2\cos^2\varphi_h/2L_hR_x - \sin\varphi_h(1/L_h - \cos\varphi_h/R_x) \\ &\times x^3/2L_h^2 + \sin^3\varphi_hx^3/2L_h^3].\end{aligned} \quad (10)$$

In case (b), the crystal is bent so that the reflecting plane remains flat ($z \approx 0$). Then,

$$\begin{aligned}|SP| &= L_0[1 + x\sin\varphi_0/L_0 + x^2\cos^2\varphi_0/2L_0^2 \\ &- x^3\sin\varphi_0\cos^2\varphi_0/2L_0^3].\end{aligned} \quad (11)$$

For the distance $|S'P|$, we will come to an expression similar to (11) making the changes $L_0 \longrightarrow L_h$ and $\varphi_0 \longrightarrow \varphi_h$. The expression for the intensity distribution at the focus will depend on which bending model [(a) or (b)] we use. For (a), it is the Pearsey integral; for (b), the Airy function.

2. SPHERICAL ABERRATION OF THE REFLECTED BEAM

From the preceding section, it follows that the diffracted beam in the Johann scheme is subjected to spherical aberration. Let $|S'P| = L_h$ and $|OS'| = L_{h(0)}$. Consider the symmetrical geometry of diffraction. Then, for case (a) we have

$$L_h(x) \cong L_{h(0)}(1 + x\cos\Theta_B/(R_x\sin\Theta_B));$$

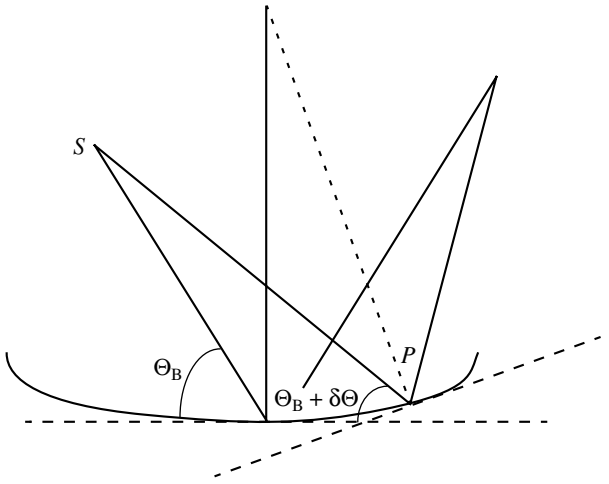


Fig. 2. On the calculation of the diffraction reflection zone of the crystal.

that is, the spherical aberration is significant for heavily bent crystals, when $x \le R_x$. For back scattering in the Johann scheme, the aberration is insignificant up to third-order terms in x inclusive.

In case (b) (the reflecting surface curvature is neglected), spherical aberration is present from the first- to third-order terms inclusive. Its effect can be minimized by using the back scattering geometry. Then, the second-order aberration alone is left.

3. WAVE INTENSITY DISTRIBUTION AT THE FOCUS

With allowance for aberration, we will find the intensity of the diffracted beam at the focus. According to the theory of dynamic Bragg diffraction in thick crystals [2, 3, 7, 10], the amplitude of the diffracted wave at the entrance plane of the crystal ($z \approx 0$) is given by (up to insignificant factors)

$$E_h(x, 0) = (2k)^{1/2}|\chi_{hr}|\sigma_h C\gamma_0 \exp(ikL_0)\exp(i\pi/2) \times \exp(-ik\cos\varphi_h x^2/2R_x)[8\pi L_0 \sin(\varphi_0 + \varphi_h)]^{-1} \times \int_{-\infty}^{+\infty} d\check{y} R(\check{y}) \exp(-i\check{y}\sigma_h x), \quad (12)$$

where

$$\check{y} = \{\Delta\Theta_B - \chi_0(\cos\varphi_0 + \cos\varphi_h)/2\sin(\varphi_0 + \varphi_h)\} \times \sin 2\Theta_B/|\chi_{hr}|$$

is the normalized angular variable (with allowance for refraction), $\Delta\Theta_B$ is the deviation from the exact Bragg angle, $\sigma_h = k\chi_h/2\cos\Theta_B$, and C is the polarization factor.

In (12), we took into account that $L_0 = R_x\cos\varphi_0$ in the Johann scheme. For the amplitude coefficient of plane wave reflection by a slightly bent crystal, $R(\check{y})$, we will use its asymptotic expression coincident with the coefficient of reflection by a perfect unbent crystal:

$$R(\check{y}) = \check{y} - (\check{y}^2 - 1)^{1/2}.$$

Using the Huygens–Fresnel X-ray diffraction and X-ray optical principles and taking into account the results of Section 1, we find the intensity of the diffracted wave at a point ξ_p near the focus:

$$I_h(\xi_p) \sim \left| k^2 C\chi_{hr}\sigma_h (8\pi^{3/2} R_x^2 \sin^2 2\Theta_B)^{-1} \int_{-\infty}^{+\infty} d\check{y} R(\check{y}) \times \int_{-\infty}^{+\infty} dx \exp\{ik[\xi_p\gamma_h x/L_h - \check{y}\sigma_h x/k + \cos^2\varphi_h x^2/2L_h - \cos\varphi_h x^2/2R_x - x^3 \sin\varphi_h \cos^2\varphi_h/2L_h^2]\}\right|^2. \quad (13)$$

The integral over the coordinate x in (13) can be reduced to the Airy function:

$$I_h(\xi_p) \sim \left| k^2 C\chi_{hr}\sigma_h (4\pi R_x^2 \sin^2 2\Theta_B)^{-1} \times \int_{-\infty}^{+\infty} d\check{y} R(\check{y}) (3A_2)^{-1/3} \Phi(A_1/\{3A_2\}^{1/3})\right|^2, \quad (14)$$

where

$$\Phi(t) = (\pi)^{-1/2} \int_0^{+\infty} \cos(u^3/3 + ut) du = (2^{-1}\pi^{-1/2}) \int_{-\infty}^{+\infty} \exp\{i(ut + u^3/3)\} du$$

is the Airy function. For symmetric diffraction, $A_1 = -k\xi_p/R_x + \check{y}\sigma_h$ and $A_2 = -k\cos\Theta_B/2R_x^2$. The integral over $\check{y}$ in (13) is taken by the method of stationary phase. The stationary point is $\check{y}_{st} \approx 0$.

The intensity $I_h(\xi_p)$ vanishes at $t_{min} = -2.33811$. With the spherical aberration taken into account, the geometrical focus position ($\xi_p = 0$) and the position of the focused wave intensity maximum are not coincident. The largest maximum of the Airy function, 0.9494, is at $t_{max} \cong -1.02$. For the (220) reflection of the CuK_α radiation from a silicon crystal ($R_x \approx 1$ m), the coordinate of the intensity maximum is $\xi_p^{max} \approx -0.1$ μm.

Now let us derive a formula for the diffraction width of the maximum of function (14). Note that $\Phi(-2.34) = -0.0024$ and $\Phi(4.00) = 0.0016866$. Then, the half-width of the intensity maximum is given by

$$\Delta\xi_p \approx 6.34 R_x^{1/3}\lambda^{2/3}\cos\Theta_B^{1/3}/(2\pi)^{2/3}(2/3)^{1/3} \approx (3\pi)^{1/3} R_x^{1/3}\lambda^{2/3}(\cos\Theta_B)^{1/3}. \quad (15)$$

For the parameters of the Johann scheme that were mentioned above, the focus width was estimated with (15) at $\Delta\xi \approx 0.6$ μm. The half-intensity focus width is ≥0.1 μm. The diffraction broadening calculated by the relationship $\Delta\xi = \lambda R_x/x_{eff}$ [2, 3, 7] (without allowance for spherical aberration) is $\Delta\xi \approx 1.5 \times 10^{-2}$ μm.

4. JOHANN FOCUSING SPECTROMETER

The spectral characteristics of a Johann spectrometer calculated in view of (15) are several orders of magnitude lower than those obtained in [2, 7]. The spectral

resolution can be found from the Bragg condition

$$\Delta\lambda/\lambda = \cot\Theta_B \Delta\Theta_\xi,$$

where $\Delta\Theta_\xi \approx \Delta\xi_p/L_h$ and $L_h = R_x \sin\Theta_B$.

For an extended source of length d, only its length $d_{coh} \sim 2\Delta\Theta L_0 = 2\Delta\Theta R_x \sin\Theta_B$ takes part in coherent Bragg reflection. To be definite, we consider here symmetrical diffraction; $\Delta\Theta$ is the half-width of the Bragg reflection curve. Then, setting $\Delta\xi_p$ equal to sum (15) of the focus diffraction width and d_{coh}, we find the spectral resolution

$$\Delta\lambda/\lambda \sim \cot\Theta_B\{(3\pi\cos\Theta_B)^{1/3}\lambda^{2/3}/(R_x^{2/3}\sin\Theta_B) + 2\Delta\Theta\}. \quad (16)$$

For the MoK_α radiation ($\lambda = 1.54$ Å) reflected from a silicon crystal with $R_x = 1$ m, estimator (16) yields $\Delta\lambda/\lambda \approx 4.3 \times 10^{-5}$. Note that for these numerical values, the second term in (16) makes a major contribution.

For back scattering ($\Theta_B \approx \pi/2$), the spectral resolution,

$$(\Delta\lambda/\lambda)_{\pi/2} \le (3\pi)^{1/3}\lambda^{2/3}(\Delta\Theta)^{4/3}/R_x^{2/3} + 2(\Delta\Theta)^2 \quad (17)$$

is several times higher than at $\Theta_B \neq \pi/2$. However, if the crystal is bent only slightly ($R_x \approx 1$ m),

$$(\Delta\lambda/\lambda)_{\pi/2} \le 2(\Delta\Theta)^2 \approx 1.8 \times 10^{-5}.$$

Note that formula (17) yields the upper limit of the spectral resolution for back scattering. The half-width of the total reflection range for back diffraction is $\Delta\Theta \sim |\chi_{hr}|^{1/2}$.

From Fig. 3 it follows that, for a given source size, the resolution of a Johann spectrometer is the best when the bending radius of a crystal meets the condition $R_x > d/(2\Delta\Theta\sin\Theta_B)$. For $R_x \gg d/(2\Delta\Theta\sin\Theta_B)$, the effect of the source size is negligibly small. The spectral resolution in this case is on the order of 10^{-7}.

The Johann scheme is more sensitive to the source size than the Cauchois scheme [11] for the Laue geometry. However, the characteristics of a Johann spectrometer are no worse than those of a Cauchois spectrometer [12].

5. FOCUSING OF A SPHERICAL WAVE BY A CRYSTAL BENT INTO A LOGARITHNIC SPIRAL

A surface shaped into a logarithmic spiral is one of the best in terms of focusing efficiency. For a logarithmic spiral, the angle between any beam emitted from a point S and the surface (at the point of incidence) is constant (Fig. 4). If the source–surface distance S is varied over some limits, the focusing effect of the spiral is retained (see below).

Consider symmetric diffraction. The size of the crystal surface fragment taking part in the reflection of

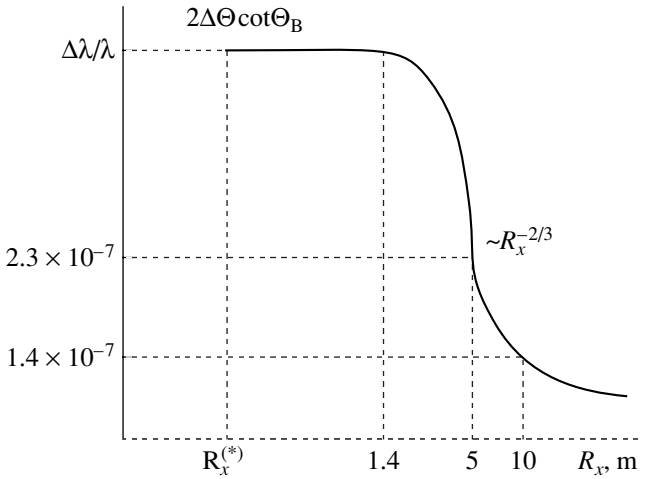


Fig. 3. Spectral resolution of a Johann focusing spectrometer vs. radius of curvature of the crystal. $d = 10$ μm = const. The plateau corresponds to radii $R_x \gg (3\pi\cos\Theta_B)^{1/2}\lambda/[(2\Delta\Theta\sin\Theta_B)^{3/2}]$. $R_x^{(*)}$ is the critical radius satisfying the condition of slightly bent reflecting surfaces.

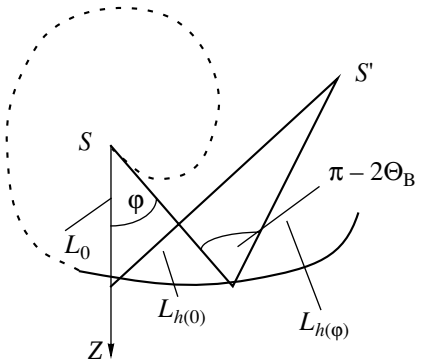


Fig. 4. Focusing of a spherical wave by a crystal shaped into a logarithmic spiral.

an incidence beam is

$$x_{eff} = (1 + \beta^2)^{1/2} L_0 \exp(\beta\varphi)\{\exp(\beta\Delta\varphi) - 1\}/\beta. \quad (18)$$

For symmetric reflection, $\beta = \cot\Theta_B$ and $\Delta\varphi$ is the incident beam divergence. From (18), it follows that for the symmetric back reflection ($\Theta_B = \pi/2$, $\cos\Theta_B \le \Delta\Theta$) of a spherically divergent beam, the reflecting surface size is limited only by the maximal divergence of the incident beam and the crystal–source spacing:

$$x_{eff,\,\Theta = \pi/2} \sim L_0\Delta\varphi. \quad (19)$$

For the source far away from the crystal ($L_0 \approx 5$ m, $\Delta\varphi \approx 3 \times 10^{-3}$), $x_{eff,\,\Theta = \pi/2} \approx 1.5$ cm. Thus, we have arrived at a formula for a focusing X-ray lens made of a crystal bent into a logarithmic spiral. From Fig. 4, the

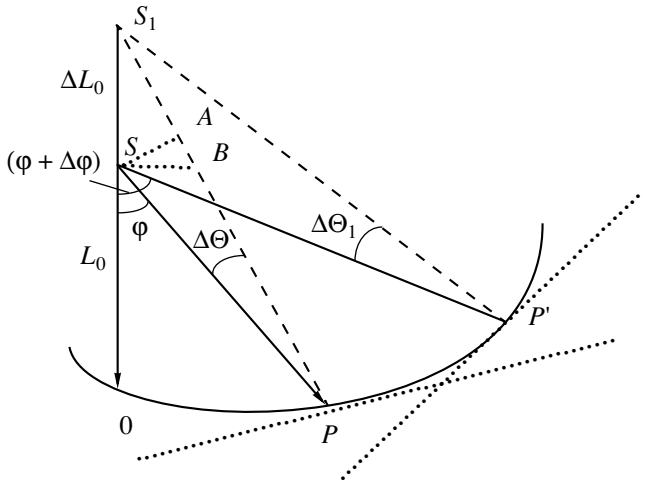


Fig. 5. On the effect of source (S) displacement on the focusing of a spherical wave by a logarithmic spiral.

spacing between the point source S and its image S' is

$$|SS'| = L_S = (L_0^2 + L_{h(0)}^2 \pm 2L_0 L_{h(0)}\cos 2\Theta_B)^{1/2}. \quad (20)$$

The geometric aberrations are minimized and focusing proceeds by the "point-to-point" principle if the equality

$$L_S = \{L_0^2 \exp(2\beta\varphi) + L_{h,\varphi}^2 \pm 2L_0 L_{h,\varphi}\exp(\beta\varphi)\cos 2\Theta_B\}^{1/2} \quad (21)$$

is valid for any arbitrary angle φ.

From (20) and (21), two focusing approaches are possible. In the first one, $\beta\varphi \ll 1$, so that the effect of aberrations is minor and $L_{h,\varphi} \geq L_{h,0}$. The second approach seems to be of greater interest, and we will concentrate on it. Conditions for this approach are the following:

$$L_0 = L_{h,\varphi}, \quad L_{h(0)} = L_{h,\varphi}\exp(\beta\varphi). \quad (22)$$

Clearly, equalities (20) and (21) are also valid and point focusing takes place. Bearing in mind that the equation of a logarithmic spiral has the form

$$R = R(\varphi) = (1 + \beta^2)^{1/2} L_0 \exp(\beta\varphi), \quad (23)$$

we find the sought lens formula from (22):

$$1/L_0 + 1/L_{h(0)} = [\exp(\beta\varphi) + 1]/(R\sin\Theta_B). \quad (24)$$

For small φ ($\varphi \ll \beta^{-1}$), we find from (23) and (24)

$$1/L_0 + 1/L_{h(0)} \cong 2/(R\sin\Theta_B), \quad (25)$$

where $L_{h(0)} \approx L_{h,\varphi} = L_0 = R\sin\Theta_B$. The result obtained means that the Johann scheme is a specific case of a logarithmic spiral. Note that, with conditions (22) met, the point S' is the asymptotic pole of another logarithmic spiral.

Consider now the effect of aberrations on focusing. With (22) in mind, we differentiate distance L_S (21) with respect to φ:

$$dL_s/d\varphi = \beta L_0^2 \exp(\beta\varphi)[\exp(\beta\varphi) \pm \cos 2\Theta_B]/L_s. \quad (26)$$

From (26), the aberration along the length SS' can be found. For $L_0 \approx 1$ m, $0 < \varphi \leq 10^{-5}$, and $\Theta_B \approx 23.65°$, we obtain $\Delta L_S \approx 20$ μm. From (26), it is seen that the aberration is a linear function of φ if $\beta\varphi \ll 1$.

The aberration is absent at $\beta = 0$, which corresponds to $\Theta_B = \pi/2$ in the symmetric case, and at $\varphi = \ln(\cos 2\Theta_B)/\beta$. Now we will show that a displacement of the source S does not disturb focusing under certain conditions. From Fig. 5,

$$\xi = \Delta L_0 \sin(\varphi - \Delta\Theta) = L_0 \exp(\beta\varphi)\sin(\Delta\Theta).$$

Let $\Delta\Theta$ be the angular half-width of the reflection curve. Then, the source displacement is given by

$$\Delta L_0 \cong L_0 \exp(\beta\varphi)\sin(\Delta\Theta)/|\sin(\varphi - \Delta\Theta)|. \quad (27)$$

It is easy to see from (27) that the source displacement $\Delta L_0 \geq L_0$ (for $\beta\varphi \ll 1$) if $\varphi \ll \Delta\Theta$. If $\varphi \sim \Delta\Theta$, the source–crystal spacing may be varied over wide limits without considerably affecting the focusing ($\Delta L_0 > L_0$). Finally, for $\Delta\Theta \ll \varphi \ll 1$, we have $\Delta L_0 \ll L_0$.

For the focusing to remain unchanged when the source is displaced to the position S_1, it is necessary that the angle $S_1P'S$ for the beam SP' (Fig. 5) be equal to $\Delta\Theta_1 = \Delta\Theta$. Taking into account that

$$\Delta L_0 \approx L_0 \exp\{\beta(\varphi + \Delta\varphi)\} \times \sin(\Delta\Theta_1)/\sin\{(\varphi + \Delta\varphi) - \Delta\Theta_1\},$$

we find

$$\tan(\Delta\Theta_1) = \sin(\varphi + \Delta\varphi)\sin(\Delta\Theta)/[\exp(\beta\Delta\varphi)\{\sin(\varphi - \Delta\Theta)\} + \cos(\varphi + \Delta\varphi)\sin(\Delta\Theta)]. \quad (28)$$

The displacement of the source S to the point S_1 does not disturb focusing, $\Delta\Theta_1 \approx \Delta\Theta$. Specifically, for $\Delta\varphi \ll \varphi \ll 1$ and $\varphi \geq \Delta\Theta$, we find from (28) that $\Delta\Theta_1 \cong \Delta\Theta$. In this case, $\Delta L_0 \leq L_0$; in other words, a displacement of a point source in a rather wide range has a negligible effect on the focusing. Let us estimate the maximal size of the source in the direction normal to L_0, $\xi_\perp$, at which focusing still takes place. From Fig. 5,

$$\xi_\perp \approx \Delta L_0 \tan[(\varphi + \Delta\varphi) - \Delta\Theta_1] = L_0 \exp[\beta(\varphi + \Delta\varphi)] \times \sin(\Delta\Theta_1)/\cos[(\varphi + \Delta\varphi) - \Delta\Theta]. \quad (29)$$

If $(\varphi + \Delta\varphi) \sim \Delta\Theta_1$ in (29), the transverse size of the source $\xi_\perp > L_0\Delta\Theta_1$. For $\Delta\Theta_1 = \Delta\Theta \sim 10^{-5}$ and $L_0 \approx 1$ m, $\xi_\perp > L_0\Delta\Theta \approx 10$ μm.

Basically, for beams with large $\varphi \leq \pi/2cL_0$, the maximum possible source size increases greatly. In Section 4, we found that the spectral characteristics of the beam reflected drastically degrade in the Johann scheme with

an extended source ($d > d_{coh} \cong 2L_0\Delta\Theta$). However, it has just been shown that the use of sources with a size $> d_{coh}$ does not disturb the focusing if the crystal surface has the form of a logarithmic spiral. This is an undeniable advantage of a logarithmic spiral over a cylindrically bent crystal surface in the Johann scheme, since the former shape extends the selection of the sources.

To conclude, we note that crystals, in general, are easy to make into the logarithmic spiral shape (see, e.g., [13]).

REFERENCES

1. H. H. Johann, Z. Phys. **69**, 185 (1931).
2. K. T. Gabrielyan, F. N. Chukhovskiĭ, and Z. G. Pinsker, Zh. Tekh. Fiz. **50** (1), 3 (1980) [Sov. Phys. Tech. Phys. **25**, 1 (1980)].
3. F. N. Chukhovskiĭ, Metallofizika **3** (5), 3 (1981).
4. D. B. Wittry and S. Sun, J. Appl. Phys. **67**, 1633 (1990).
5. W. Z. Chang and D. B. Wittry, J. Appl. Phys. **74**, 2999 (1993).
6. F. N. Chukhovskii, W. Z. Chang, and E. Förster, J. Appl. Phys. **77**, 1843 (1995).
7. K. T. Gabrielyan, F. N. Chukhovskiĭ, and D. I. Piskunov, Zh. Éksp. Teor. Fiz. **3**, 834 (1989).
8. J. Witz, Acta Crystallogr. A **25**, 30 (1969).
9. T. Chen, V. A. Bushuev, and R. N. Kuz'min, Zh. Tekh. Fiz. **60** (10), 60 (1990) [Sov. Phys. Tech. Phys. **35**, 1148 (1990)].
10. F. N. Chukhovskii, K. T. Gabrielyan, and P. V. Petrashen', Acta Crystallogr. A **34**, 610 (1978).
11. Y. Cauchois, J. Phys. Radium **3**, 320 (1932).
12. K. T. Gabrielyan, F. N. Chukhovskiĭ, and Z. G. Pinsker, Zh. Tekh. Fiz. **50**, 1641 (1980) [Sov. Phys. Tech. Phys. **25**, 956 (1980)].
13. P. M. De Wolff, *Selected Topics on X-ray Crystallography*, Ed. by J. Bouman (North-Holland, Amsterdam, 1951).

Translated by V. Isaakyan

Technical Physics, Vol. 47, No. 6, 2002, pp. 666–672. Translated from Zhurnal Tekhnicheskoĭ Fiziki, Vol. 72, No. 5, 2002, pp. 11–18.

THEORETICAL AND MATHEMATICAL PHYSICS

2D Laue Focusing of X-ray Radiation by a Double-Crystal System

T. Tchen

Lomonosov Sate Academy of Fine Chemical Technology, pr. Vernadskogo 86, Moscow, 117571 Russia

e-mail: docent65@mtu-net.ru, ttchen@e-mail.ru

Received October 1, 2001

Abstract—The theory of the dynamic 2D focusing of an X-ray wave upon its Laue diffraction by a system of two bent crystals is developed. The reflecting planes of the crystals are perpendicular to their surfaces. 2D focusing is shown to be highly sensitive to the bending radius of the crystals and to the difference in their thicknesses. The effect of astigmatism on 2D Laue focusing is described, and conditions for stigmatic focusing are found. One-dimensional focusing inside the crystals is discussed. The spectral characteristics of a double-crystal 2D-focusing spectrometer are considered. © *2002 MAIK "Nauka/Interperiodica".*

INTRODUCTION

Dynamic Laue focusing of X rays has been studied both theoretically and experimentally for the case of a narrow beam diffracted by perfect (flat) crystals [1–4], as well as a spherical wave diffracted by both perfect [5–12] and elastically bent crystals [13–20]. It has been shown that the spherical wave can be focused both inside and outside the crystals. Even in the pioneering works [1, 3, 4], the effect of focusing was applied for the spectral decomposition of X-ray radiation. The theory of a bent-crystal Laue spectrometer was developed in [19]. The transfer of X rays with magnifying and demagnifying bent-crystal Laue lenses was demonstrated in [21–24]. The focusing of X rays by two bent crystals was studied in [25].

However, in all the works cited, the object of investigation was 1D focusing. In this work, we suggest a theory of dynamic focusing of a spherical X-ray wave Laue-diffracted by two bent crystals. With a proper arrangement of the crystals, the 2D focusing of the radiation with the formation of the focus takes place.

2D FOCUSING OF A SPHERICAL WAVE IN A VACUUM FOR THE CASE OF LAUE DIFFRACTION BY TWO BENT CRYSTALS

Let a spherical X-ray wave be incident on a system of two elastically bent crystals (Fig. 1). We will use the parabolic decomposition for the spherical wave. Let the crystals be so bent that their reflecting planes coincide with the normal cross sections, which remain undistorted. Then, each of the plane harmonics of the spherical wave "sees" the crystal unbent at the point of incidence. Therefore, the amplitude coefficient of plane-wave reflection (ACPWR) can be given by the expression for the case of an undistorted (flat) crystal. The plane of diffraction scattering by the first crystal will be referred to as the meridional plane. Then, the diffraction plane of the second crystal makes an angle $\approx\Theta_B^{(1)}$ with the sagittal plane of the first crystal and the sagittal plane of the second crystal makes an angle $\approx\Theta_B^{(2)}$ with the meridional plane of the first crystal. Here, $\Theta_B^{(1,2)}$ are the Bragg angles for the first and the second crystal, respectively. In the sagittal plane, the Fourier transform $R(k_y, t)$ of the Green function $G_h^{(y)}$ $(y_T - y, t)$ is roughly equal to unity. Then, the Green function

$$G_h^{(y)}(y_T - y, t) = (2\pi)^{-1}\int_{-\infty}^{+\infty} dk_y \exp[-ik_y(y_T - y)]R(k_y, t)$$

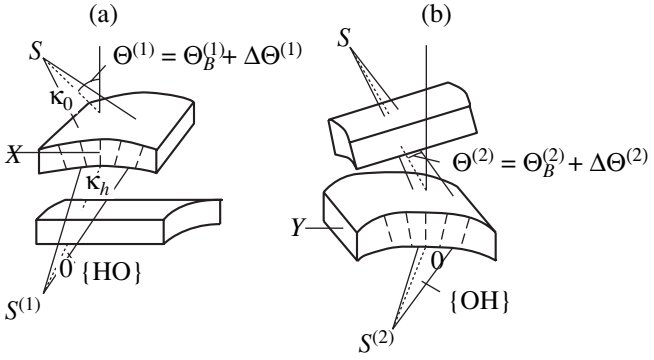


Fig. 1. 2D focusing of a spherical X-ray wave Laue-diffracted in a system of two bent crystals. (a) Focusing in the {HO} plane, $|OS^{(1)}| = L_{hh}^{(1)}$; (b) focusing in the {OH} plane, $|OS^{(2)}| = L_{hh}^{(1)}$. $S^{(1,2)}$, images of the source *S*.

behaves like the delta function. With this in mind and taking into account the Huygens–Fresnel X-ray diffraction and X-ray optical principles, we find the intensity of the doubly diffracted spherical wave at a distance L_{hh} to the second crystal in a vacuum:

$$I_{hh}(x, y) = I_h^{(1)}(x) I_h^{(2)}(y), \tag{1}$$

where $I_h^{(1)}(x)$ and $I_h^{(2)}(y)$ are the intensity distributions in the {HO} and {OH} planes, respectively (Fig. 1), and (x, y) are the coordinates of the point of observation. We have

$$I_h^{(1,2)} = (\gamma_{h(1,2)}^2/2\pi\kappa(L_0^{(1,2)} + L_h^{(1,2)})L_0^{(1,2)}L_h^{(1,2)}\alpha_0^{(1,2)}\alpha_h^{(1,2)}) \times \int_{-\infty}^{+\infty} d\check{y}_{(1,2)} T(\check{y}_{(1,2)}, t_{(1,2)}) \exp\{i\varphi_{(1,2)}(\check{y}_{(1,2)}, t_{(1,2)})\}, \tag{2}$$

$$\varphi_{(1,2)}(\check{y}_{(1,2)}, t_{(1,2)}) = -\sigma_{h(1,2)}^2 \check{y}_{(1,2)}^2 (1/\alpha_0^{(1,2)} + 1/\alpha_h^{(1,2)})/2\kappa + \sigma_{h(1,2)}\check{y}_{(1,2)}\gamma_{(1,2)}(\Delta\Theta^{(1,2)}/\alpha_0^{(1,2)} - \gamma_{h(1,2)} \times \{\psi - \psi_0\}/L_h^{(1,2)}\alpha_h^{(1,2)}), \tag{3}$$

where

$$\sigma_{h(1,2)} = \kappa\chi_{h(1,2)}/2\sin\Theta_B^{(1,2)}, \quad \psi - \psi_0 = x - x_0 \text{ for } \varphi_{(1)} \quad \text{and} \quad \psi - \psi_0 = y - y_0 \text{ for } \varphi_{(2)}. \tag{4}$$

In the integral in (2),

$$R(\check{y}_{(1,2)}, t_{(1,2)}) = (\chi_h/\chi_{\bar{h}})^{1/2} \times \exp\{i\varepsilon_{(1,2)}t_{(1,2)}/2\}/2(\check{y}_{(1,2)}^2 + 1)^{1/2} \tag{5}$$

is the ACPWR by flat crystals [26],

$$\varepsilon_{(1,2)} = \kappa\chi_{0(1,2)}/\gamma_{0(1,2)} - C\kappa(\chi_{h(1,2)}\chi_{\bar{h}(1,2)})^{1/2} \times (\check{y}_{(1,2)} \pm (\check{y}_{(1,2)}^2 + 1)^{1/2})/\gamma_{h(1,2)}, \tag{6}$$

$\check{y}_{(1,2)} = \Delta\Theta^{(1,2)}\sin 2\Theta_B^{(1,2)}/(\chi_h\chi_{\bar{h}})^{1/2}$ is the normalized angular variable (without allowance for refraction), $\Delta\Theta^{(1,2)}$ are the deviations from the Bragg angle for the first and the second crystal,

$$\psi_{0(1,2)} = L_h^{(1,2)}(1 - \gamma_{h(1,2)}^2)^{1/2} + (1/\gamma_{h(1,2)}^2 - 1)^{1/2} t_{(1,2)}, \tag{7}$$

$$\alpha_0^{(1,2)} = \gamma_{0(1,2)}^2/L_0^{(1,2)} - \gamma_{0(1,2)}/R_{(1,2)}, \quad \alpha_h^{(1,2)} = \gamma_{h(1,2)}^2/L_h^{(1,2)} + \gamma_{h(1,2)}/R_{(1,2)}, \tag{8}$$

γ_0 and γ_h are the direction cosines for the incident and diffracted waves, $R_{(1,2)}$ are the bending radii of the crystals in the planes of diffraction scattering,

$$L_h^{(1)} = L_{hh}^{(1)} + L_{12}/\gamma_{h(1)} + t_2/\gamma_{h(2)}, \quad L_0^{(2)} = L_0^{(1)} + (t_1 + L_{12})/\gamma_{h(1)}, \tag{9}$$

$L_h^{(2)} = L_{hh}^{(2)}$ is the distance from the second crystal to the source image in the {OH} plane, $L_{hh}^{(1)}$ is the distance from the second crystal to the image in the {HO} plane, L_{12} is the spacing between the crystals, $\kappa = 2\pi/\lambda$, λ is the incident wavelength, $t_{(1,2)}$ are the thicknesses of the crystals, $\chi_{h,\bar{h}}$ and χ_0 are the Fourier components of X-ray polarizability, and C is the polarization factor. The signs $\pm$ in (6) correspond to two modes (α and β) of the wave field in the crystals (the minus sign corresponds to the weakly absorbable mode), and the indices (1) and (2) correspond to the first and the second crystal.

Intensity (2) is calculated with the method of stationary phase (here, it is taken into account that the function $(\check{y}^2 + 1)^{-1/2}$ varies much more slowly than the exponential). With regard for the ACPWR phase, we find the first-, second-, third-, and fourth-order derivatives of the eikonal function:

$$\begin{aligned}
\partial\Phi_{(1,2)}/\partial\check{y}_{(1,2)} &= -\sigma_{h(1,2)}^2\check{y}_{(1,2)}(1 - \alpha_0^{(1,2)} + 1/\alpha_h^{(1,2)})/\kappa \\
&+ \sigma_{h(1,2)}\gamma_{h(1,2)}(\Delta\Theta^{(1,2)}/\alpha_0^{(1,2)} - \gamma_{h(1,2)}(\Psi - \Psi_0)/L_h^{(1,2)}\alpha_h^{(1,2)}) \\
&- t_{(1,2)}\tan\Theta_B^{(1,2)}\sigma_{h(1,2)}\{1 \pm \check{y}_{(1,2)}/(\check{y}_{(1,2)}^2 + 1)^{1/2}\}, \\
\partial^2\Phi_{(1,2)}/\partial\check{y}_{(1,2)}^2 &= -\sigma_{h(1,2)}^2(1/\alpha_0^{(1,2)} + 1/\alpha_h^{(1,2)})/\kappa \\
&\pm \tan\Theta_B^{(1,2)}t_{(1,2)}\sigma_{h(1,2)}/(\check{y}_{(1,2)}^2 + 1)^{3/2}, \\
\partial^3\Phi_{(1,2)}/\partial\check{y}_{(1,2)}^3 &= (\pm 3\tan\Theta_B^{(1,2)}t_{(1,2)}\sigma_{h(1,2)} \\
&\times \check{y}_{(1,2)})/(\check{y}_{(1,2)}^2 + 1)^{5/2}, \\
\partial^4\Phi_{(1,2)}/\partial\check{y}_{(1,2)}^4 &= \{\pm 3\tan\Theta_B^{(1,2)}t_{(1,2)}\sigma_{h(1,2)} \\
&\times (1 - 4\check{y}_{(1,2)}^2)\}/(\check{y}_{(1,2)}^2 + 1)^{7/2}; \\
\Phi_{(1,2)} &= \varphi_{(1,2)} + \varepsilon_{\check{y}(1,2)}t_{(1,2)}/2,
\end{aligned} \tag{10}$$

where $\varepsilon_{\check{y}(1,2)}$ is the $\check{y}_{1,2}$-dependent part of $\varepsilon_{(1,2)}$ [see (6)].

The stationary points $\check{y}_{st}$ are determined from the condition $\partial\Phi_{(1,2)}/\partial\check{y}_{(1,2)} = 0$. For the beams to intersect, that is, for focusing in the {HO} and {OH} planes to take place, it is also necessary that the second derivative $\partial\Phi_{(1,2)}/\partial\check{y}_{(1,2)}$ vanish ($\partial^2\Phi_{(1,2)}/\partial\check{y}_{(1,2)} = 0$). The point $\check{y}_{st(1,2)} = 0$ is stationary for both planes, so that the wave is focused at the geometrical focus with the coordinates found from the following equations:

$$\begin{aligned}
1/\alpha_0^{(1,2)} + 1/\alpha_h^{(1,2)} &= \pm\kappa\tan\Theta_B^{(1,2)}t_{(1,2)}/\sigma_{h(1,2)}, \\
\Psi - \Psi_0 &= L_h^{(1,2)}\alpha_h^{(1,2)}(\Delta\Theta^{(1,2)}/\alpha_0^{(1,2)} \\
&- t_{(1,2)}\tan\Theta_B^{(1,2)}/\gamma_{h(1,2)})/\gamma_{h(1,2)}.
\end{aligned} \tag{11}$$

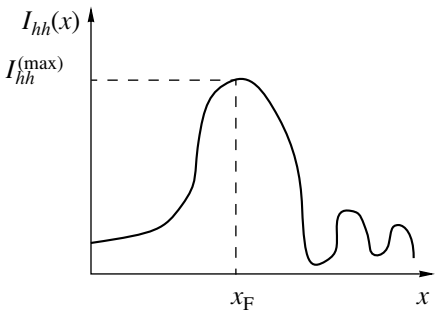


Fig. 2. Intensity distribution in the {HO} plane for the doubly diffracted wave.

$I_{hh}^{(\max)} \cong 0.675\gamma_{(1)}^2/\{2\pi\kappa(L_0^{(1)}L_h^{(1)})L_0^{(1)}L_h^{(1)}\alpha_0^{(1)}\alpha_h^{(1)}\}$.

The 2D (point) focus is the point of intersection of two focal lines passing through the points $S^{(1)}$ and $S^{(2)}$. The focal line passing through the point $S^{(1)}$ makes an angle $\Theta^{(2)}$ with the Y axis.

When the second crystal is rotated about the normal to its surface through an angle φ, the second focal line, passing through the point $S^{(2)}$, also rotates through the same angle. In the limit $\varphi \longrightarrow 90°$, the second focal line transforms into a new focal line arising because of the focusing of the convergent beam by the second crystal. At $\varphi = 90°$, the first focal line disappears. 2D focusing disappears, and 1D focusing into a line $L_h^{(2)}$ distant from the second crystal takes place. In this case, the geometrical conditions for focusing have the form

$$1/\alpha_0 + 1/\alpha_h = \pm\kappa[\tan\Theta_B^{(1)}t_1/\sigma_{h(1)} + \tan\Theta_B^{(2)}t_2/\sigma_{h(2)}], \quad (12)$$

where $\alpha_h = \gamma_{h(1)}^2/L_h^{(2)} + \gamma_{h(1)}/R_1 + \gamma_{h(2)}/R_2$.

For the convergent beam to be focused by the second crystal, it is necessary that $\xi \sim x_{(2)}/R_2 > \phi \sim x_{(2)}/L_h^{(2)}$, where ξ is the angle between the cross sections normal to the crystal surface and ϕ is the angle of beam convergence.

When the second crystal is rotated through an angle $\approx 2\Theta_B^{(2)}$ counterclockwise about the X axis, the focus shifts to another point along a circular arc of radius $\approx L_{hh}^{(2)}$. The angle between the directions to the initial and final positions of the focus is $4\Theta_B^{(2)}$.

From (11) and (12), it follows that two focuses corresponding to the α and β modes of the wave field may form behind the second crystal if the crystals are not too thick.

It is of interest to consider 2D Laue focusing in the case when one focal line arises because of α mode focusing and the other, because of β mode focusing. Conditions for α–β focusing have the form

$$\begin{aligned} 1/\alpha_{0,(+)}^{(1)} + 1/\alpha_{h,(+)}^{(1)} &= \kappa\tan\Theta_B^{(1)}t_1/\sigma_{h(1)}, \\ 1/\alpha_{0,(-)}^{(2)} + 1/\alpha_{h,(-)}^{(2)} &= -\kappa\tan\Theta_B^{(2)}t_2/\sigma_{h(2)}, \end{aligned} \quad (13)$$

where the signs $\pm$ correspond to the different modes.

Note that the second crystal is exposed to the beam (wave) with the angle of divergence $\sim|\chi_{hr}|^{1/2}$ (in the plane of diffraction scattering of this crystal) but only a small part of the beam diffracts. However, the effect of focusing more than compensates for the transmission loss. At $L_0 \approx 1$ m, $L_h \leq 1$ m, $\chi_{hr} \sim 10^{-5}$, and the focus width $\Delta\Psi \approx 10$ μm, 2D focusing provides roughly a 800-fold gain in the collection angle compared with the case of 1D focusing.

Let us find the intensity distribution in the neighborhood of the focus (we will consider the focusing of the weakly absorbable mode for definiteness). Intensity (1) in view of (2) and (10) is expressed as

$$\begin{aligned} I_{hh}(x, y) = (16)^{-1}\gamma_{h(1)}^2\gamma_{h(2)}^2[4\pi^2\kappa^2(L_0^{(1)} + L_h^{(1)})(L_0^{(2)} + L_h^{(2)}) \\ \times L_0^{(1)}L_0^{(2)}L_h^{(1)}L_h^{(2)}\alpha_0^{(1)}\alpha_0^{(2)}\alpha_h^{(1)}\alpha_h^{(2)}]^{-1} \\ \times I_p^{(1)}(A_1^{(1)}, A_2^{(1)})I_p^{(2)}(A_1^{(2)}, A_2^{(2)}), \end{aligned} \quad (14)$$

where $I_p(A_1, A_2) = \int_{-\infty}^{+\infty}\exp\{i[A_2\tau + A_1\tau^2 + \tau^4]\}d\tau$ is the Pearsey integral [27],

$$\begin{aligned} \tau &= \check{y}_{(1,2)}C_{(1,2)}/(4!)^{1/4}, \\ C_{(1,2)} &= (3\tan\Theta_B^{(1,2)}t_{(1,2)}\sigma_{h(1,2)})^{1/4}, \\ A_1^{(1,2)} &= (6)^{1/2}(\partial^2\Phi_{(1,2)}/\partial\check{y}_{(1,2)}^2)\big|_{\check{y}=0}/C_{(1,2)}^2, \\ A_2^{(1,2)} &= C_{(1,2)}^{-1}\delta(\partial\Phi_{(1,2)}/\partial\check{y}_{(1,2)})\big|_{\check{y}=0} \\ &= -\sigma_{h(1,2)}\gamma_{h(1,2)}^2C_{(1,2)}^{-1}\delta\psi/L_h^{(1,2)}\alpha_h^{(1,2)}, \\ \delta\Psi &= \Psi - \Psi_F, \end{aligned} \quad (15)$$

and Ψ_F is the position of the geometrical focus.

The intensity distribution has a peak with weak oscillations on the right of it (Fig. 2).

Let the diffraction width of the focus be defined as a distance over which the intensity drops twice. Then, taking into account the properties of the Pearsey integral [27] and putting $A_1 = 0$, we come to the following estimates of the diffraction sizes of the focus in the {HO} and {OH} planes:

$$\begin{aligned} A_2 &\approx 1.8, \\ \Delta\Psi^{(1,2)} &\cong 1.8L_h^{(1,2)}\alpha_h^{(1,2)}C_{(1,2)}/\sigma_{h(1,2)}\gamma_{h(1,2)}^2. \end{aligned} \quad (16)$$

The focus size is seen to be $\Delta\Psi \sim t_{(1,2)}^{1/4}$. For the (220) reflection of CuK_α radiation from a silicon crystal of thickness $t \cong 600$ μm with $R = -2$ m, $\sin\Theta_B \cong 0.4$, and

$L_h \cong 0.65$ m, formula (16) yields $\Delta\Psi \approx 11$ μm. From (16), an interesting consequence follows: the diffraction broadening of the focus is smaller for a system of flat crystals. This feature of a flat-crystal scheme, however, shows up only for $L_h \geq R\gamma_h$.

Exactly at the geometrical focus ($A_1 = A_2 = 0$), the Pearsey integral absolute value equals 1.813. Its maximum (roughly 2.7) is observed at $A_1 \cong -2$ and $A_2 = 0$, not at the geometrical focus. The maximal intensity of the doubly diffracted wave is reached if

$$1/\alpha_0 + 1/\alpha_h = \kappa\tan\Theta_B t/\sigma_h + \kappa(2\tan\Theta_B t^{1/2})/\sigma_h^{3/2}. \quad (17)$$

Relationship (17) implies that the contribution from the second term on the right-hand side is significant if $t \leq 2\Lambda/\pi < \Lambda$, where $\Lambda = \lambda\cos\Theta_B/|\chi_{hr}|$ is the extinction length. For dynamic Laue focusing, the shift of the intensity maximum from the geometrical focus can be neglected.

FOCUSING OF A SPHERICAL WAVE INSIDE CRYSTALS

Consider the case $L_0^{(1)} \neq R_1\gamma_{0(1)}$. For the weakly absorbable mode, the eikonal function Φ in the first crystal is given by

$$\Phi = -\sigma_{h(1)}^2\breve{y}_{(1)}^2/2\kappa\alpha_0^{(1)} - \sigma_{h(1)}\breve{y}_{(1)}x + \sigma_{h(1)}z(\breve{y}_{(1)}^2 + 1)^{1/2}/\gamma_{h(1)} + \sigma_{h(1)}\gamma_{h(1)}\Delta\Theta^{(1)}\breve{y}_{(1)}/\alpha_0^{(1)}. \quad (18)$$

Let $\gamma_0 = \gamma_h$. Then, from (18) it follows that the weakly absorbable wave is focused inside the first crystal at a depth

$$Z_F^{(1)} = \sigma_{h(1)}\cot\Theta_B^{(1)}/\kappa\alpha_0^{(1)}. \quad (19)$$

The thickness of the first crystal must exceed $Z_F^{(1)}$. Varying the distance $L_0^{(1)}$, one can vary the position of the focus inside the crystal. Note in passing that this circumstance might be used in studying the defect structure of bent crystals.

Putting $Z_F^{(1)} = t_1$ in (19), we come to the equation for a cylindrical Laue lens [22] as a specific case:

$$\cos\Theta_B/L_0 - 1/R_1 = \chi_h/2t_1\sin^2\Theta_B.$$

The condition for kinematic focusing at the exit surface of the first crystal, which follows from the lens equation, has the form $L_0 = R_1\cos\Theta_B$. The wave weakly absorbable in both crystals is focused inside the second crystal at a depth $Z_F^{(2)} < t_2$:

$$Z_F^{(2)} = \sigma_{h(2)}\cot\Theta_B^{(2)}/\kappa\alpha_0^{(2)}.$$

It is of interest that either crystal and also the entire double crystal system may serve as a demagnifying (in the longitudinal direction) X-ray diffraction lens. In fact, a slight displacement of the source by a distance ΔL_0 along the incident beam direction causes the focus to shift by a distance

$$\Delta Z_F = |\chi_{hr}|\cos\Theta_B/2\sin^2\Theta_B(\cos\Theta_B - L_0/R)^2\Delta L_0 \quad (20)$$

for the symmetric geometry of diffraction. As follows from (20), the longitudinal coefficient of image transfer may be very low (~10^{-4}–10^{-5}). For example, for sources in the immediate vicinity of the crystal surface (or in the case of a flat crystal), the longitudinal coefficient of image transfer is proportional to ~$|\chi_{hr}|$. Note that a weakly absorbable Bloch wave focused in the first crystal is confined once again beyond the first crystal at a distance $L_h^{(1)} = 2\sin^2\Theta_B^{(1)}\cos\Theta_B^{(1)}t_1/|\chi_{hr(1)}| - \cos^2\Theta_B^{(1)}/\alpha_0^{(1)}$.

Let the planes of diffraction scattering of the crystals coincide. Then, in the second crystal, the mode weakly absorbable in both is focused at a depth

$$\begin{aligned} Z_F^{(2)} &= \pm|\chi_{hr(1,2)}|(L_{12} - L_h^{(1)})/(\sin\Theta_B^{(2)}\sin2\Theta_B^{(2)}) \\ &= \pm|\chi_{hr(2)}|L_{12}/(\sin\Theta_B^{(2)}\sin2\Theta_B^{(2)}) \\ &+ |\chi_{hr(2)}|\sin\Theta_B^{(1)}\sin2\Theta_B^{(1)}t_1/(\sin\Theta_B^{(2)}\sin2\Theta_B^{(2)}|\chi_{hr(1)}|) \\ &\pm|\chi_{hr(2)}|\cos^2\Theta_B^{(1)}/(\sin\Theta_B^{(2)}\sin2\Theta_B^{(2)}\alpha_0^{(1)}). \end{aligned} \quad (21)$$

Here, the upper plus refers to the case $L_{12} > L_h^{(1)}$; the lower minus, to $L_{12} < L_h^{(1)}$. It is seen that in a system of flat crystals with $\Theta_B^{(1)} = \Theta_B^{(2)}$ and $\chi_{hr(1)} = \chi_{hr(2)}$, the doubly diffracted wave is focused in the second crystal at a depth

$$Z_F^{(2)} = \pm t_1 + |\chi_{hr}|(L_0 + L_{12})/(\sin\Theta_B\sin2\Theta_B). \quad (22)$$

For $(L_0 + L_{12}) \ll t_1\sin\Theta_B\sin2\Theta_B/|\chi_{hr}|$ and $L_{12} < L_h^{(1)}$, the wave will be focused in the second crystal at a depth equal to the thickness of the first crystal. This result is consistent with those obtained in [1–4].

When focusing a spherical wave inside, an elastically bent crystal acts as an X-ray diffraction lens. The formula for such a lens can be written as

$$1/L_0 + 1/Z_F = 1/F, \quad (23)$$

where the focal length F is

$$\begin{aligned} F &= |\chi_{hr}|L_0R/\{(|\chi_{hr}| + \sin\Theta_B\sin2\Theta_B)R - 2\sin^2\Theta_BL_0\} \\ &\approx |\chi_{hr}|/\{2\sin^2\Theta_B|\cos\Theta_B/L_0 - 1/R|\}. \end{aligned}$$

In the case of a plane wave ($L_0 \gg R\cos\Theta$), we have $F \approx R|\chi_{hr}|/2\sin^2\Theta_B$. Hence, for slightly bent crystals with $R \approx 1$ m, $F \sim |\chi_{hr}|$. For bending radii $R < \lambda\sin\Theta_B\sin2\Theta_B/|\chi_{hr}|^2$, the focusing takes place at depths shorter than the extinction length. Obviously, if a spherical wave passes through a set of bent crystals roughly meeting the Bragg condition and having thicknesses $t_i \geq \Lambda_i$, the diffraction focus may form in each of the

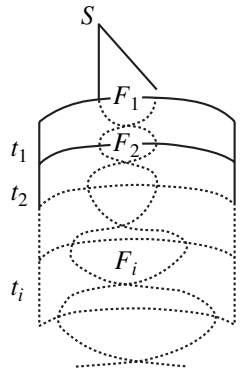


Fig. 3. Multiple focusing of the weakly absorbable mode of the spherical X-ray wave upon Laue diffraction by crystals of different atomic composition. *S*, point source; $F_{1, 2, ..., i}$, foci.

crystals (Fig. 3). For this to occur, it is necessary that the wave weakly absorbable in the first crystal be weakly absorbed in the others.

EFFECT OF ASTIGMATISM ON THE FOCUSING OF THE SPHERICAL WAVE BY THE DOUBLE-CRYSTAL SYSTEM

From Eqs. (11), the 2D focused beam is subjected to astigmatism. Let diffraction be symmetric ($\gamma_0 = \gamma_h = \gamma$),

Effect of astigmatism on the transmission of doubly focussed beam

R_2, m	ΔL_{hh}, cm	ΔL, μm	$10^{-4} N$	N
–2.100	8.099	1285	6.50	285
–2.050	4.676	742	2.17	35.5
–2.040	3.989	633	1.58	15.78
–2.030	3.300	523	1.08	3.2
–2.020	2.611	414	0.68	0.35
–2.010	2.317	367	0.53	6.68
–2.000	1.230	195	0.15	22.70
–1.990	0.538	85		
–1.985	0.192	30		
–1.984	0.123	20		
–1.983	0.054	9		
–1.982	–0.016	3		
–1.981	–0.085	13		
–1.980	–0.154	24		

Note: $t_1 = t_2 = 400$ μm, $L_0^{(1)} = 1$ m, $L_{12} = 1$ cm, $R_1 = -2$ m, $\sigma_h = 4.5899 \times 10^5$ m^{-1}, $L_x^{(2)} = L_y^{(2)} = 2$ cm, $L_{hh}^{(1)} = 63.05$ cm, and $L_h^{(1)} = 64.24$ cm. (220) reflection of CuK_α radiation from silicon crystals. $\Delta L_x = \Delta L_y = \Delta L$ are the sizes of aberration spots.

the crystals have the same structure ($\sigma_{h(1)} = \sigma_{h(2)}$), and the Bragg angles coincide. Then, the distances $L_{hh}^{(1, 2)}$ from the second crystal to the source images $S^{(1, 2)}$ are given by

$$L_{hh}^{(1)} = L_h^{(1)} - (L_{12} + t_2)/\gamma, \quad L_{hh}^{(2)} = L_h^{(2)},$$

$$L_h^{(1, 2)} = \gamma G_{(1, 2)} R_{(1, 2)}/(R_{(1, 2)} - G_{(1, 2)}), \qquad (24)$$

$$G_{(1, 2)} = \kappa \sin\Theta_B t_{(1, 2)}/\sigma_h - (\gamma/L_0^{(1, 2)} - 1/R_{(1, 2)})^{-1}.$$

For arbitrary $L_0^{(1)}$, L_{12}, $t_{(1, 2)}$, and $R_{(1, 2)}$, the doubly diffracted beam converges not to a point but to the focal spot, whose size depends on the astigmatism (see table). Let us suppose that the beam illuminates the entire surface area $\approx L_x^{(2)} L_y^{(2)}$ of the second crystal. Then, the size of the focal spot located in the middle between the focal lines is $\approx (\Delta L_{hh} L_x^{(2)}/2L_{hh}^{(1)})(\Delta L_{hh} L_y^{(2)}/2L_{hh}^{(1)})$. Here, we assume that $\Delta L_{hh} = L_{hh}^{(2)} - L_{hh}^{(1)} > 0$.

The diffraction broadening of the focus in the {HO} and {OH} planes calculated by formula (16) are $\Delta\Psi \approx 10$ μm. Astigmatism reduces the collection angle by $N \approx \Delta L_x \Delta L_y/(\Delta\Psi)_{(1)}(\Delta\Psi)_{(2)}$ times. The values of the coefficient N for various R_2 are listed in the table. It is seen that the astigmatism and, hence, the collection angle depend on the bending radius R_2. For example, the collection angle grows by two orders of magnitude when R_2 changes from –2.000 to –1.984 m.

Now we will show that the degree of astigmatism depends on the difference in the crystal thicknesses. Let the source be placed near the surface of the first crystal ($L_0^{(1)} \longrightarrow 0$) and $R_{(1, 2)} \longrightarrow \infty$ (flat crystals). Then, from (24), the astigmatism is expressed as

$$\Delta L_{hh} = (\kappa \sin\Theta_B \gamma \Delta t/\sigma_h) + (L_{12} + t_2)/\gamma. \qquad (25)$$

It is seen that ΔL_{hh} is a linear function of the difference $\Delta t = t_2 - t_1$ in the thicknesses for $L_{12} = \text{const}$ and $t_2 = \text{const}$.

With $\Delta t \cong 1$ μm, $\sin\Theta_B \cong 0.4$, and $\chi_h \sim 10^{-5}$, we have from (25) $\Delta L_{hh} \approx 3$ cm. In this case, the collection angle drops by three or four orders of magnitude compared with the stigmatic situation. The result obtained is of practical importance, since the effect of focusing can be applied to measure the crystal thickness with a micron (or even submicron) accuracy.

DOUBLE-CRYSTAL FOCUSING LAUE SPECTROMETER

Now let us discuss the spectral characteristics of the double-crystal system when it is used as a double-crystal Laue spectrometer with diffraction focusing. In the presence of astigmatism, we have, in essence, two spectrometers with different spectral resolutions. Let the

condition when astigmatism is absent be fulfilled: $\Delta L_{hh} = L_{hh}^{(2)} - L_{hh}^{(1)} = 0$. Taking into account that $d\Theta = \delta\lambda \tan\Theta_B/\lambda$ (according to the Bragg equation) and that the linear dispersion is given by

$$D_\Psi = \tan\Theta_B(L_h^{(1,2)} - L_0^{(1,2)}\alpha_h^{(1,2)}/\alpha_0^{(1,2)})/\lambda\gamma, \quad (26)$$

we find for the spectral resolution in the {OH} and {HO} planes

$$(d\lambda/\lambda)_{(1,2)} \cong \cot\Theta_B\gamma\Delta\Psi^{(1,2)} \times \left|L_h^{(1,2)} - L_0^{(1,2)}\alpha_h^{(1,2)}/\alpha_0^{(1,2)}\right|^{-1}. \quad (27)$$

For $\alpha_0 = \alpha_h$, $L_0 = 1$ m, $\gamma = 0.91$, $R = -2$ m, $L_h = 0.48$ m, and $\delta\Psi = 10$ μm, the spectral resolution is estimated from (27) at $d\lambda/\lambda \approx 4 \times 10^{-5}$. As follows from (27), the spectral resolution varies as $t^{1/4}$ and the lower limit is reached at $t \sim \Lambda$. Also, we have from (27) that in the symmetric geometry of diffraction ($\alpha_0 = \alpha_h$), the resolution for the flat crystals is the worst if $L_0 = L_h$.

If a plane wave is incident ($L_0 \gg R\gamma$) and $\alpha_0 = \alpha_h$, the spectral resolution is on the order of 10^{-6} ($L_0 \approx 10$ m).

When the spherical wave source is in the immediate vicinity of the surface of the first crystal ($L_0^{(1)} \longrightarrow 0$), the resolution is the best if $L_h^{(1,2)} \approx -R_{(1,2)}\gamma_{h(1,2)}$. The minus sign means that the first crystal presents its convex surface to the source ($R < 0$).

In the absence of astigmatism, the decomposition into the spectrum takes place along the line making the angles Ψ_1 and Ψ_2 with focal lines *1* and *2* (Fig. 4). Here, $\sin\Psi_1/\sin\Psi_2 = \Delta x/\Delta y$ and $\Psi = \Psi_1 + \Psi_2$. The spectral resolution is given by $\max\{(d\lambda/\lambda)_1, (d\lambda/\lambda)_2\}$.

FOCUSING OF A PLANE WAVE AND SYNCHROTRON RADIATION

A practical case of special interest is the incidence of a plane X-ray wave ($L_0^{(1)} \gg R_1\gamma_{0(1)}$) and synchrotron radiation. In the latter case, the radiation has an extremely small divergence in one plane and a considerable divergence in the plane normal to the former.

Focusing conditions (11) remain valid for the plane wave if the change $\alpha_0^{(1,2)} \longrightarrow -\gamma_0^{(1,2)}/R_{(1,2)}$ is made. Then, the linear dispersions and the spectral resolutions in the {HO} and {OH} planes equal each other:

$$D_\Psi^{(p1)} \cong \tan\Theta_B^{(1,2)}L_0^{(1,2)}R_{(1,2)}(1/L_h^{(1,2)} + 1/\gamma R_{(1,2)})/\lambda,$$
$$(d\lambda/\lambda)_{(1,2)}^{p(1)} \cong \frac{(\Delta\Psi)_{(1,2)}\cot\Theta_B^{(1,2)}}{\{L_0^{(1,2)}(1/\gamma + R_{(1,2)}/L_h^{(1,2)})\}}. \quad (28)$$

Clearly, when numerically estimating the parameters from (28), one must put $L_0 \gg R\gamma$. Formulas (28) characterize the spectral properties of the plane wave being focused by the double-crystal spectrometer.

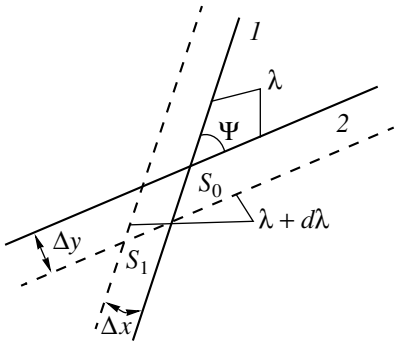


Fig. 4. On the spectral decomposition of the spherical wave upon Laue diffraction (astigmatism is absent). (*1*) Focal line normal to the {HO} plane and (*2*) focal line normal to the {OH} plane. Δx and Δy, diffraction broadening of the focus in the {HO} and {OH} planes, respectively; S_0, point source image (focus) for the wavelength λ meeting the Bragg condition; and S_1, image for $\lambda + d\lambda$.

For the plane wave, the astigmatism is given by

$$\Delta L_{hh} = -\gamma R_2(1 + \sigma_h R_2/(\kappa\sin\Theta_B t_2)) + \gamma R_1(1 + \sigma_h R_1/(\kappa\sin\Theta_B t_1)) + (L_{12} + t_2)/\gamma. \quad (29)$$

Let us show that ΔL_{hh} strongly depends on the difference in the bending radii $\Delta R = R_2 - R_1 \ll R_1$. In fact, with $t_1 = t_2 = 400$ μm, $\gamma \cong 0.91$, $\sigma_h = 4.5899 \times 10^5$ m^{-1}, $R_1 \gg 1$ m, $\lambda = 1.54$ Å, $L_{12} \approx 1$ cm, and $\Delta R \approx 1$ cm, we find that $\Delta L_{hh} \approx 1.47$ cm. The collection angle in this case drops by one or one and a half orders of magnitude in comparison with the case when the beam is focused into a "point." For $L_{12} \gg \Delta R$, the value of ΔL_{hh} depends on ΔR only slightly and the astigmatism depends only on the distance L_{12}. As follows from (29), the astigmatism can be diminished (at $R_2 > R_1$) if the convex surface of the first crystal is presented to the incident wave ($R_1 < 0$). Thus, with small distances between the crystals ($L_{12} < \Delta R$), the point focusing of the plane wave is highly sensitive to the difference in the bending radii of the crystals.

Figure 4 plots the astigmatism against the bending radius of the first crystal. For the crystals of the same thickness, the astigmatism is the least if the bending radii coincide. Formula (29) has been derived for crystalline plates made of the same material. From (29), the astigmatism is absent in the specific case $R_1 = R_2 = R$ and $\Delta t = t_2 - t_1 < 0$, when the intercrystal spacing is given by

$$L_{12}^{(p1)} = -(|\chi_{hr}|\gamma^2 R^2\Delta t/2\sin^2\Theta_B t_1 t_2 + t_2). \quad (30)$$

From (30), with $R = 2$ m, $\sin\Theta_B = 0.4$, $t_1 = 400$ μm, $t_2 = 300$ μm, and $\chi_h \sim 10^{-5}$, the intercrystal spacing at which the image of a remote source is stigmatic equals $L_{12}^{(p1)} \approx 8.7$ cm.

For the crystals of the same thickness and bending radius, the astigmatism depends linearly on the intercrystal spacing $\Delta L_{hh} = (L_{12} + t_2)/\gamma$. In this case, the collection angle will be high if the crystals are very close to each other.

Let the bending radii be the same ($R_1 = R_2 = R$) but the thicknesses differ. The astigmatism depends on the difference in the thicknesses $\Delta t = t_2 - t_1 \ll t_1$ as follows:

$$\Delta L_{hh} \cong \gamma\sigma_h R^2 \Delta t/(\kappa \sin\Theta_B t_1^2) + (L_{12} + t_2)/\gamma. \quad (31)$$

We will estimate ΔL_{hh} for the same numerical parameters as before. Then, the first term in (31) is $1.6 \times 10^2 \Delta t$. For $\Delta t = 1$ μm and $L_{12} \approx 1$ cm, its contribution is much less than from the second term. With $L_{12} \approx 1$ cm, the ΔL vs. Δt dependence becomes significant if the bending radii are sufficiently large ($R \geq 10$ m).

Finally, let the radiation incident on the first crystal diverge weakly in the plane of scattering and have a finite divergence in the perpendicular plane (synchrotron radiation). The geometrical conditions for the 2D Laue focusing of synchrotron radiation are given by

$$\begin{aligned} -R_1/\gamma + 1/\alpha_h^{(1)} &= \kappa \tan\Theta_B^{(1)} t_1/\sigma_{h(1)}, \\ 1/\alpha_0^{(2)} + 1/\alpha_h^{(2)} &= \kappa \tan\Theta_B^{(2)} t_2/\sigma_{h(2)}, \\ x - x_0 &= -L_h^{(1)}\alpha_h^{(1)}(\Delta\Theta^{(1)} R_1 + t_1 \tan\Theta_B^{(1)})/\gamma^2, \\ y - y_0 &= L_h^{(2)}\alpha_h^{(2)}(\Delta\Theta^{(2)}/\alpha_0^{(2)} - t_2 \tan\Theta_B^{(2)}/\gamma)/\gamma. \end{aligned} \quad (32)$$

It follows that the nth diffraction harmonic λ/n, along with the fundamental harmonic λ, is focused to the focal point.

REFERENCES

1. V. L. Indenbom, I. Sh. Slobodetskiĭ, and K. G. Truni, Zh. Éksp. Teor. Fiz. **66**, 1110 (1974) [Sov. Phys. JETP **39**, 542 (1974)].
2. É. V. Suvorov and V. I. Polovinkina, Pis'ma Zh. Éksp. Teor. Fiz. **20** (5), 326 (1974) [JETP Lett. **20**, 145 (1974)].
3. V. L. Indenbom and É. V. Suvorov, Pis'ma Zh. Éksp. Teor. Fiz. **23**, 485 (1976) [JETP Lett. **23**, 441 (1976)].
4. V. L. Indenbom, É. V. Suvorov, and I. Sh. Slobodetskiĭ, Zh. Éksp. Teor. Fiz. **71**, 359 (1976) [Sov. Phys. JETP **44**, 187 (1976)].
5. A. M. Afanas'ev and V. G. Kon, Fiz. Tverd. Tela (Leningrad) **19**, 1775 (1977) [Sov. Phys. Solid State **19**, 1035 (1977)].
6. V. G. Kon, Metallofizika **10** (2), 78 (1988).
7. V. G. Kon, Fiz. Tverd. Tela (Leningrad) **19**, 3567 (1977) [Sov. Phys. Solid State **19**, 2085 (1977)].
8. V. V. Aristov, A. A. Snigiriv, A. M. Afanas'ev, *et al.*, Acta Crystallogr. A **42**, 426 (1986).
9. V. G. Kohn and A. H. Toneyan, Acta Crystallogr. A **42**, 441 (1986).
10. V. G. Kon, Kristallografiya **24**, 712 (1979) [Sov. Phys. Crystallogr. **24**, 408 (1979)].
11. V. V. Aristov, V. G. Kon, and A. A. Snigirev, Kristallografiya **31**, 1059 (1986) [Sov. Phys. Crystallogr. **31**, 626 (1986)].
12. L. V. Levonyan, Pis'ma Zh. Tekh. Fiz. **7**, 269 (1981) [Sov. Tech. Phys. Lett. **7**, 116 (1981)].
13. F. N. Chukhovskiĭ and P. V. Petrashen', Dokl. Akad. Nauk SSSR **222**, 599 (1975) [Sov. Phys. Dokl. **20**, 314 (1975)].
14. P. V. Petrashen' and F. N. Chukhovskiĭ, Zh. Éksp. Teor. Fiz. **69**, 477 (1975) [Sov. Phys. JETP **42**, 243 (1975)].
15. P. V. Petrashen' and F. N. Chukhovskiĭ, Pis'ma Zh. Éksp. Teor. Fiz. **23**, 385 (1976) [JETP Lett. **23**, 347 (1976)].
16. F. N. Chukhovskiĭ and P. V. Petrashen', Dokl. Akad. Nauk SSSR **228**, 1087 (1976) [Sov. Phys. Dokl. **21**, 303 (1976)].
17. P. V. Petrashen' and F. N. Chukhovskiĭ, Kristallografiya **21**, 283 (1976) [Sov. Phys. Crystallogr. **21**, 152 (1976)].
18. F. N. Chukhovskiĭ, Metallofizika **2** (6), 3 (1980).
19. K. T. Gabrielyan, F. N. Chukhovskiĭ, and Z. G. Pinsker, Zh. Tekh. Fiz. **50**, 1641 (1980) [Sov. Phys. Tech. Phys. **25**, 956 (1980)].
20. K. T. Gabrielyan and F. N. Chukhovskiĭ, Zh. Tekh. Fiz. **52**, 2127 (1982) [Sov. Phys. Tech. Phys. **27**, 1309 (1982)].
21. V. I. Kushnir and É. V. Suvorov, Pis'ma Zh. Éksp. Teor. Fiz. **32**, 551 (1980) [JETP Lett. **32**, 534 (1980)].
22. V. I. Kushnir and É. V. Suvorov, Fiz. Tverd. Tela (Leningrad) **24**, 1606 (1982) [Sov. Phys. Solid State **24**, 918 (1982)].
23. V. I. Kushnir and E. V. Suvorov, Phys. Status Solidi A **69**, 483 (1982).
24. V. I. Kushnir, V. M. Kaganer, and E. V. Suvorov, Acta Crystallogr. A **41**, 17 (1985).
25. K. T. Gabrielyan and G. O. Demirchyan, Kristallografiya **34**, 22 (1989) [Sov. Phys. Crystallogr. **34**, 11 (1989)].
26. Z. G. Pinsker, *X-ray Crystal Optics* (Nauka, Moscow, 1982).
27. T. Pearsey, Philos. Mag. **37**, 311 (1946).

Translated by V. Isaakyan

Technical Physics, Vol. 47, No. 6, 2002, pp. 673–681. Translated from Zhurnal Tekhnicheskoĭ Fiziki, Vol. 72, No. 5, 2002, pp. 19–27.

GASES AND LIQUIDS

Capillary Oscillations of an Emitting Charged Viscous Drop of Finite Conductivity

A. I. Grigor'ev, S. O. Shiryaeva, and V. A. Koromyslov

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

e-mail: shir@uniyar.ac.ru

Received September 17, 2001

Abstract—Capillary oscillations of a charged drop of a viscous incompressible liquid with finite conductivity emitting electromagnetic waves are considered. A dispersion relation for the capillary oscillations has been derived and analyzed using a linear approximation to oscillation amplitudes. © *2002 MAIK "Nauka/Interperiodica".*

INTRODUCTION

Investigating the capillary oscillations and stability of a charged drop of a real liquid is of interest in connection with numerous scientific and technological applications (see, for example, [1–4] and the references therein). Most theoretical investigations carried out in a linear approximation to the oscillation amplitude [1–4], and all investigations with the use of approximations of a higher order of smallness (see [5]), dealt with drops of an perfectly conducting inviscid liquid. The phenomenon of electromagnetic emission by a vibrating drop was discussed in only one study [6], which also dealt with an perfectly conducting inviscid liquid. It appears interesting, therefore, to study capillary oscillations of a viscous drop with a finite transfer rate of its electric charge taking into account radiation energy losses.

1. Consider a spherical drop of a viscous incompressible liquid with a density ρ, coefficient of kinematic viscosity ν, conductivity σ, permittivity ε_1, and surface tension γ, whose surface is in capillary oscillations of an infinitesimally small amplitude caused by thermal motion of the drop molecules. A drop having a radius R and carrying a charge Q is placed in a vacuum. An equation of the drop free surface in a spherical coordinate system with the origin in its center of mass can be written in the form $r = R + \xi(\Theta, t)$, where ξ is the distortion of the equilibrium spherical drop surface caused by the capillary wave motion, $|\xi| \ll R$; Θ is the spherical angular coordinate; and t is the time.

To simplify the subsequent presentation and calculations, we introduce dimensionless variables $R = 1$, $\rho = 1$, and $\gamma = 1$. (All dimensionless quantities we denote by the same symbols as the corresponding dimensional ones.) Other quantities are expressed in units of their characteristic values and have lower asterisks, viz.,

$$r_* = R; \quad t_* = R^{3/2}\rho^{1/2}\gamma^{-1/2}; \quad u_* = R^{-1/2}\rho^{-1/2}\gamma^{1/2};$$

$$P_* = R^{-1}\gamma; \quad s_* = R^{-3/2}\rho^{-1/2}\gamma^{1/2}; \quad Q_* = R^{3/2}\gamma^{1/2};$$

and

$$\nu_* = R^{1/2}\rho^{-1/2}\gamma^{1/2}.$$

Here, u is the liquid velocity, P is the liquid pressure, and s is the oscillation frequency.

A set of electrohydrodynamic equations describing the system considered with a field E induced by the charge Q has the form

$$\frac{d\mathbf{u}}{dt} \equiv \frac{\partial \mathbf{u}}{\partial t} + (\mathbf{u}\cdot\nabla)\mathbf{u} = -\nabla P_1 + \nu\Delta\mathbf{u}; \quad \nabla\cdot\mathbf{u} = 0; \tag{1.1}$$

$$\nabla\cdot\mathbf{D}_j = 0; \quad \mathbf{D}_j = \varepsilon_j\mathbf{E}_j; \tag{1.2}$$

$$\Delta\mathbf{E}_j - \frac{1}{c^2}\frac{\partial \mathbf{E}_j}{\partial t^2} = 0; \quad j = 1, 2. \tag{1.3}$$

Subscript 1 denotes quantities relating to the drop, and subscript 2, to the ambient; ε_j is the permittivity (in the following we assume that $\varepsilon_2 = 1$ and $\varepsilon_1 \equiv \varepsilon$); $\mathbf{u}(\mathbf{r}, t)$ is the liquid velocity field and $P_1(\mathbf{r}, t)$ is the pressure in the liquid drop in the presence of an external electric field; and c is the propagation speed of electromagnetic waves in vacuum.

At the drop free surface described by an equation in the form

$$F(\mathbf{r}, t) \equiv r - 1 - \xi(\Theta, t) = 0, \tag{1.4}$$

the following boundary conditions must be satisfied:

—the kinematic condition

$$\frac{dF}{dt} = \frac{\partial F}{\partial t} + \mathbf{u}\cdot\nabla F = 0; \tag{1.5}$$

—the dynamic condition for the tangential components of the stress tensor

$$(\Pi_{2\tau} - \Pi_{1\tau}) - \nu[\boldsymbol{\tau}\cdot(\mathbf{n}\cdot\nabla)\mathbf{u} + \mathbf{n}\cdot(\boldsymbol{\tau}\cdot\nabla)\mathbf{u}] = 0;$$
$$\Pi_\tau = \frac{\varepsilon}{4\pi}E_n E_\tau; \tag{1.6}$$

where E_n and E_τ are the normal and tangential components of the electric field strength and $\mathbf{n}$ and $\boldsymbol{\tau}$ are unit vectors normal and tangential to the drop surface (1.4), respectively;

—the dynamic boundary condition for the normal component of the stress tensor

$$-(P_1 - P_2) + 2\nu\mathbf{n}\cdot(\mathbf{n}\cdot\nabla)\mathbf{u} - P_E + P_\gamma = 0; \tag{1.7}$$

where P_2 is the pressure from the ambient and P_E and P_γ are the pressures exerted by electric and surface tension forces, respectively;

—the conditions of discontinuity of the normal component of the electric induction and continuity of the tangential components of the electric field

$$r = 1 + \xi: D_{2n} - D_{1n} = 4\pi\kappa; \quad E_{2\tau} = E_{1\tau}; \tag{1.8}$$

and the balance equation for the surface charge density

$$\frac{\partial\kappa}{\partial t} - \sigma(\mathbf{n}\cdot\mathbf{E}_1) + \mathrm{div}_\Sigma\kappa\mathbf{u}_t + \kappa b\mathbf{E}_\tau = 0;$$
$$r \longrightarrow \infty: D_2 \longrightarrow 0; \tag{1.9}$$
$$r \longrightarrow 0: D_1 \longrightarrow 0;$$

where $\kappa(\Theta, t)$ is the surface density of the electric charge, b is the mobility of charged particles, $\mathbf{u}_\tau$ and $\mathbf{E}_\tau$ are vectors in a plane tangential to the drop surface, and $\mathrm{div}_\Sigma\mathbf{a}$ is the surface divergence of the vector $\mathbf{a}$.

In addition, we impose additional conditions following from obvious requirements of invariability of the drop volume and position of its center of mass and conservation of the total drop charge:

$$\int_V r^2 dr\sin\Theta d\Theta d\phi = \frac{4}{3}\pi R^3; \tag{1.10}$$
$$V = [0 \le r \le R + \xi(\Theta, t),\ 0 \le \Theta \le \pi];$$

$$\oint_S \xi(\Theta, t)\mathbf{n}_r\sin\Theta d\Theta d\phi = 0; \tag{1.11}$$
$$S = [r = R + \xi(\Theta, t),\ 0 \le \Theta \le \pi];$$

$$\oint_S \kappa(\Theta, t)dS = Q; \tag{1.12}$$
$$S = [r = R + \xi(\Theta, t),\ 0 \le \Theta \le \pi].$$

The set of Eqs. (1.1); (1.3), together with conditions (1.5)–(1.12), is a mathematical formulation of the problem being solved.

2. We shall seek a solution to the problem using the scalarization technique described in detail in [7], where it is shown, in particular, that the liquid velocity field $\mathbf{u}(\mathbf{r}, t)$ in a drop, being of the first order of smallness in $|\xi(\Theta, t)|$, can be presented in the form of a sum of three orthogonal vector fields:

$$\mathbf{u}(\mathbf{r}, t) = \hat{\mathbf{N}}_1\Psi_1(\mathbf{r}, t) + \hat{\mathbf{N}}_2\Psi_2(\mathbf{r}, t) + \hat{\mathbf{N}}_3\Psi_3(\mathbf{r}, t); \tag{2.1}$$

where $\Psi_j(\mathbf{r}, t)$ are scalar functions determined by the field $\mathbf{u}(\mathbf{r}, t)$ and $\hat{\mathbf{N}}_i$ are vector operators meeting the orthogonality conditions

$$\hat{\mathbf{N}}_j^+\cdot\hat{\mathbf{N}}_i = 0 \quad \text{at} \quad i \ne j \tag{2.2}$$

and the conditions of commutativity with the Laplace operator Δ

$$\Delta\hat{\mathbf{N}}_i = \hat{\mathbf{N}}_i\Delta; \quad i = 1, 2, 3. \tag{2.3}$$

It is convenient to choose the operators $\hat{\mathbf{N}}_i$ in the following form:

$$\hat{\mathbf{N}}_1 = \nabla; \quad \hat{\mathbf{N}}_2 = \hat{\mathbf{N}}_1\times\mathbf{r} = \nabla\times\mathbf{r};$$
$$\hat{\mathbf{N}}_3 = \hat{\mathbf{N}}_1\times\hat{\mathbf{N}}_2 = \nabla\times(\nabla\times\mathbf{r}); \tag{2.4}$$

$$\hat{\mathbf{N}}_1^+ = -\nabla; \quad \hat{\mathbf{N}}_2^+ = \mathbf{r}\times\nabla; \quad \hat{\mathbf{N}}_3^+ = (\mathbf{r}\times\nabla)\times\nabla. \tag{2.5}$$

The operators $\hat{\mathbf{N}}_j^+$ and $\hat{\mathbf{N}}_j$ are Hermitian adjoints. Operator $\hat{\mathbf{N}}_1$ separates the potential component of the liquid motion; operator $\hat{\mathbf{N}}_2$, the eddy toroidal component about axis OZ; and the operator $\mathbf{N}_3$, the eddy poloidal component.

According to [7], the toroidal component of the velocity field does not influence the motion of the drop free surface and can be omitted in the present analysis. Substituting expansion (2.1) into Eqs. (1.1) and using conditions (2.2) and (2.3) to find scalar functions $\Psi_1(\mathbf{r}, t)$ and $\Psi_3(\mathbf{r}, t)$ and introducing additions of the first order of smallness in $|\xi|$ to the pressure field in the drop $p_1(\mathbf{r}, t)$, we obtain a set of equations

$$\Delta\Psi_j(\mathbf{r}, t) - \frac{1}{\nu}(1 - \delta_{j1})\frac{\partial\Psi_j(\mathbf{r}, t)}{\partial t} = 0; \quad j = 1, 3; \tag{2.6}$$

$$p_1(\mathbf{r}, t) = -\frac{\partial}{\partial t}\Psi_1(\mathbf{r}, t). \tag{2.7}$$

Assume that the time dependence of scalar functions $\Psi_j(\mathbf{r}, t)$ and distortions of the equilibrium drop surface $\xi(\Theta, t)$ are exponential:

$$\Psi_j(\mathbf{r}, t) \sim \exp(st); \quad \xi(\Theta, t) \sim \exp(st).$$

We write the solution to set (2.6) in the form of an expansion in Legendre polynomials:

$$\Psi_1(\mathbf{r}, t) = \sum_{n=2}^{\infty} C_n^{(1)} r^n P_n(\mu)\exp(st), \quad \mu \equiv \cos\Theta \tag{2.8}$$

and

$$\Psi_3(\mathbf{r}, t) = \sum_{n=2}^{\infty} C_n^{(3)} \frac{i_n\left(\sqrt{\frac{s}{\nu}} r\right)}{i_n\left(\sqrt{\frac{s}{\nu}}\right)} P_n(\mu)\exp(st), \quad (2.9)$$

where $i_n(x)$ are the modified first-kind spherical Bessel functions [9].

The function $\xi(\Theta, t)$ describing the drop surface distortion can also be presented in the form of a series in Legendre polynomials:

$$\xi(\Theta, t) = \sum_{n=2}^{\infty} Z_n P_n(\mu)\exp(st). \quad (2.10)$$

The range of n from $n = 2$ to $n \longrightarrow \infty$ in the expansions above is defined by conditions (1.10) and (1.11).

Boundary conditions (1.5)–(1.7) of the first order of smallness in $|\xi|$ after implementation of the scalarization procedure take the form

$$r = 1: \frac{\partial \xi(\Theta, \varphi, t)}{\partial t} = \frac{\partial \Psi_1}{\partial r} - \frac{1}{r}\Delta_\Omega \Psi_3; \quad (2.11)$$

$$\boldsymbol{\tau} \equiv \mathbf{e}_\Theta: \left\{(\Pi_{2\Theta} - \Pi_{1\Theta}) - \nu\left[\frac{\partial}{\partial \Theta}\left\{2\frac{\partial}{\partial r}\left(\frac{\Psi_1}{r}\right) + \frac{\partial^2 \Psi_3}{\partial r^2} - \frac{1}{r^2}(2 + \Delta_\Omega)\Psi_3\right\}\right.\right.$$
$$\left.\left. + \frac{1}{\sin\Theta}\frac{\partial}{\partial \varphi}\left\{\frac{\partial \Psi_2}{\partial r} - \Psi_2\right\}\right]\right\}\Bigg|_{r=1} = 0, \quad (2.12)$$

where Δ_Ω is the angular component of the Laplace operator,

$$r = 1: \left\{-p_1 + 2\nu\left[\frac{\partial^2 \Psi_1}{\partial r^2} - \Delta_\Omega\left(\frac{\partial}{\partial r}\left(\frac{\Psi_3}{r}\right)\right)\right] - p_E - (2 + \Delta_\Omega)\xi\right\} = 0. \quad (2.13)$$

3. To find an expression describing the electric field induced in the ambient medium by a charged conducting spherical drop with the surface perturbed by capillary wave motion, it is necessary to solve the set of Eqs. (1.2) and (1.3) subject to boundary conditions (1.8) and (1.9) with additional condition (1.12).

It is reasonable to seek a solution to this problem in the form of a superposition in spherical coordinates with the origin at the center of mass of the drop:

$$\mathbf{E}_j = \mathbf{E}_2^{(0)}(r) + \mathbf{E}_j^{(1)}(r, \Theta, t), \quad j = 1, 2, \quad (3.1)$$

where $\mathbf{E}_2^{(0)}(r)$ is the electric field strength in the vicinity of the unperturbed sphere and $\mathbf{E}_j^{(1)}(r, \Theta, t)$ is an additional contribution to the field strength resulting from capillary oscillations of the drop surface and having the same order of smallness as the surface perturbation, that is, $\mathbf{E}_j^{(1)} \sim \xi$.

By virtue of the linearity of the problem being solved and expansion (3.1), vectors $\mathbf{E}_2^{(0)}$ and $\mathbf{E}_j^{(1)}$ are solutions to sets of equations analogous to (1.2) and (1.3).

In the zero-order approximation, the electric field strength inside the drop is $\mathbf{E}_1^{(0)}(r) = 0$ and the field strength $\mathbf{E}_2^{(0)}(r)$ outside the drop can be found by solving the problem

$$\Delta \mathbf{E}_2^{(0)} \equiv \text{grad}\,\text{div}\mathbf{E}_2^{(0)} - \text{curl}\,\text{curl}\mathbf{E}_2^{(0)} = 0; \quad (3.2)$$

$$\text{div}\mathbf{E}_2^{(0)} = 0; \quad (3.3)$$

$$r = 1: \mathbf{E}_2^{(0)}\mathbf{n} = 4\pi\kappa_0; \quad \mathbf{E}_2^{(0)}\boldsymbol{\tau} = 0; \quad (3.4)$$

$$\oint_S \kappa(\Theta, t)dS = Q; \quad S = [r = R + \xi(\Theta, t);\ 0 \le \Theta \le \pi];$$
$$r \longrightarrow \infty: \mathbf{E}_2^{(0)}(\mathbf{r}, t) \longrightarrow 0, \quad (3.5)$$

where κ_0 is the charge density at the unperturbed drop surface.

We shall seek $\mathbf{E}_2^{(0)}$ in the form

$$\mathbf{E}_2^{(0)} = -\nabla\Phi.$$

In this case, Eq. (3.2) subject to (3.3) is satisfied identically and Eq. (3.3) transforms into the Laplace equation for the potential Φ:

$$\Delta\Phi = 0. \quad (3.6)$$

Due to the central symmetry of the problem in a zero-order approximation, Eq. (3.6) takes the form of an ordinary second-order differential equation with a solution

$$\Phi = -\frac{A}{r}.$$

Then, for the field strength $\mathbf{E}_2^{(0)}$ we obtain

$$\mathbf{E}_2^{(0)} = \frac{A}{r^2}\mathbf{e}_r. \quad (3.7)$$

The second boundary condition (3.4) is satisfied for arbitrary values of the constant A due to $\mathbf{e}_r \perp \boldsymbol{\tau}$. The magnitude of this constant $A = Q$ is determined from

the first of conditions (3.4) and (3.5). Finally, the expression for the field strength $\mathbf{E}_2^{(0)}$ takes the form

$$\mathbf{E}_2^{(0)}(r) = -\frac{Q}{r^2}\mathbf{e}_r. \tag{3.8}$$

Consider the determination of $\mathbf{E}_j^{(1)}$ as a problem of the first order of smallness. To simplify the subsequent calculation, we omit the superscript in $\mathbf{E}_j^{(1)}$ indicating the order of smallness. An additional contribution of the first order of smallness to the electric field strength $\mathbf{E}_j$ is determined by Eqs. (1.2) and (1.3). Let us scalarize these equations by representing vector $\mathbf{E}_j$ in the form of an expansion:

$$\mathbf{E}_j = \hat{\mathbf{N}}_1\Phi_j^{(1)} + \hat{\mathbf{N}}_2\Phi_j^{(2)} + \hat{\mathbf{N}}_3\Phi_j^{(3)}, \tag{3.9}$$

where operators $\hat{\mathbf{N}}_i$ are defined by expressions (2.4) and (2.5) and satisfy conditions (2.2) and (2.3) and $\Phi_j^{(m)}$ are arbitrary scalar functions.

With due regard for (2.2), Eq. (3.3) transforms into the Laplace equation for function $\Phi_j^{(1)}$:

$$\begin{aligned}\nabla\mathbf{E}_j &= -\hat{\mathbf{N}}_1^+\cdot\mathbf{E}_j = -\hat{\mathbf{N}}_1^+(\hat{\mathbf{N}}_1\Phi_j^{(1)} + \hat{\mathbf{N}}_2\Phi_j^{(2)} + \hat{\mathbf{N}}_3\Phi_j^{(3)})\\ &= -\hat{\mathbf{N}}_1^+\cdot\hat{\mathbf{N}}_1\Phi_j^{(1)} = \nabla\cdot\nabla\Phi_j^{(1)} = \Delta\Phi_j^{(1)} = 0.\end{aligned} \tag{3.10}$$

Due to commutativity of the operators $\hat{\mathbf{N}}_i$ and the Laplace operator (2.3), condition (1.3) transforms into three scalar equations. Substituting expansion (3.9) into (1.3) gives

$$\sum_{m=1}^{3}\hat{\mathbf{N}}_m\left[\Delta\Phi_j^{(m)} - \frac{1}{c^2}\frac{\partial^2\Phi_j^{(m)}}{\partial t^2}\right] = 0.$$

Multiplying the last expression from the left subsequently by $\hat{\mathbf{N}}_m^+$ and taking into account conditions (2.2) and (2.3), we obtain a set of scalar equations

$$\Delta\Phi_j^{(m)} - \frac{1}{c^2}\frac{\partial^2\Phi_j^{(m)}}{\partial t^2} = 0; \quad m = 1, 2, 3. \tag{3.11}$$

Inasmuch as we are solving a problem of electromagnetic emission by a drop, it is natural to assume that $\Phi_j^{(m)} \sim \exp(st)$. Then, Eqs. (3.11) reduce to Helmholtz equations.

From Eq. (3.11) at $m = 1$ and from Eq. (3.10), we obtain $(s^2/c^2)\Phi_j^{(1)} = 0$; the complex frequency s of the oscillations being different from zero, the scalar function $\Phi_j^{(1)} = 0$. Thus, the strength $\mathbf{E}_j$ of the electric field produced by the surface oscillation of a charged drop can be written as follows:

$$\mathbf{E}_j = \hat{\mathbf{N}}_2\Phi_j^{(2)} + \hat{\mathbf{N}}_3\Phi_j^{(3)}. \tag{3.12}$$

Functions $\Phi_j^{(m)}$ are solutions to the Helmholtz equation and have the form

$$\Delta\Phi_j^{(m)} - k^2\Phi_j^{(m)} = 0; \quad k^2 \equiv s^2/c^2; \quad m = 2; 3. \tag{3.13}$$

It is easy to show that, due to the axial symmetry of the problem, the electric field strength $\mathbf{E}_j$ both inside the drop ($j = 1$) and outside ($j = 2$) has only a poloidal component $\sim\hat{\mathbf{N}}_3\Phi_j^{(3)}$ in (3.12), the toroidal component, $\sim\hat{\mathbf{N}}_2\Phi_j^{(2)}$, being $\Phi_j^{(2)} \equiv 0$. Ultimately, for the field $\mathbf{E}_j$ in a spherical coordinate system, we have

$$\mathbf{E}_j \equiv \hat{\mathbf{N}}_3\Phi_j^{(3)} \equiv -\frac{1}{r}\Delta_\Omega\Phi_j^{(3)}\mathbf{e}_r + \frac{1}{r}\frac{\partial}{\partial r}r\frac{\partial\Phi_j^{(3)}}{\partial\Theta}\mathbf{e}_\Theta. \tag{3.14}$$

We shall seek a solution to Eq. (3.13) at $j = 2$ (for the ambient medium) in the form of a superposition of traveling waves:

$$\Phi_2^{(3)} = \sum_{n=2}^{\infty} D_n\frac{h_n(kr)}{h_n(k)}P_n(\mu); \quad \mu \equiv \cos\Theta, \tag{3.15}$$

where $h_n(kr)$ is the modified third-kind spherical Bessel function [8].

We shall seek a solution to Eq. (3.12) at $j = 1$ (inside the drop) in the form of a superposition of standing waves:

$$\Phi_1^{(3)} = \sum_{n=2}^{\infty} A_n\frac{i_n(kr)}{i_n(k)}P_n(\mu), \tag{3.16}$$

where $i_n(z)$ is the modified first-kind spherical Bessel function.

Substituting (3.15) and (3.16) into (3.14), we find for the field $\mathbf{E}_j$

$$\begin{aligned}&\mathbf{E}_j(\mathbf{r}, t)\\ &= \sum_n\left[E_{rn}^{(j)}n(n+1)P_n(\mu)\mathbf{e}_r + E_{\Theta n}^{(j)}\frac{dP_n(\mu)}{d\Theta}\mathbf{e}_\Theta\right]\exp(st),\end{aligned} \tag{3.17}$$

with the components for the ambient medium ($j = 2$)

$$\begin{aligned}E_{rn}^{(2)} &\equiv D_n\frac{h_n(kr)}{rh_n(k)};\\ E_{\Theta n}^{(2)} &\equiv D_n\left\{\frac{h_n(kr)}{rh_n(k)} + \frac{d}{dr}\left[\frac{h_n(kr)}{h_n(k)}\right]\right\}\end{aligned} \tag{3.18}$$

and for the drop liquid

$$E_{rn}^{(1)} = A_n \frac{i_n(kr)}{r i_n(k)};$$
$$E_{\Theta n}^{(1)} \equiv A_n \left\{ \frac{i_n(kr)}{r i_n(k)} + \frac{d}{dr}\left[\frac{i_n(kr)}{i_n(k)}\right]\right\}. \quad (3.19)$$

For the condition of equality of tangential components of the field strength **E** at the drop surface to be satisfied, we take into account that the total electric field outside the drop is a superposition of term (3.8) of the zeroth order of smallness and additional contribution (3.17) at $j = 2$ of the first order of smallness. The expression for a unit vector $\boldsymbol{\tau}_\Theta$ of the tangent to the perturbed drop surface is written in the form

$$\boldsymbol{\tau}_\Theta = \frac{1}{r}\frac{\partial \xi}{\partial \Theta}\mathbf{e}_r + \mathbf{e}_\Theta.$$

Ultimately, from the condition of equality of the tangential components, in the first order of smallness in $|\xi|$ we have

$$r = 1: QZ_n \frac{\partial P_n(\mu)}{\partial \Theta} + E_{\Theta n}^{(2)} \frac{dP_n(\mu)}{d\Theta} = E_{\Theta n}^{(1)} \frac{dP_n(\mu)}{d\Theta}. \quad (3.20)$$

Substituting (3.18) and (3.19) into (3.20), it is easy to find relationships between the unknown coefficients A_n, D_n, and Z_n:

$$A_n = D_n f_n(k) + Z_n Q I_n^{-1}(k);$$
$$f_n(k) \equiv H_n(k) I_n^{-1}(k); \quad H_n(k) \equiv 1 + \frac{d}{dr}\left[\frac{h_n(kr)}{h_n(k)}\right]_{r=1}; \quad (3.21)$$
$$I_n(k) \equiv 1 + \frac{d}{dr}\left[\frac{i_n(kr)}{i_n(k)}\right]_{r=1}.$$

For all modes of the capillary oscillations of the drop which are not damped by viscosity (at $\nu < 1$), their dimensionless frequencies are much less than the normalized electrodynamic constant c; therefore, when calculating the intensity of electromagnetic radiation from a vibrating drop, it is natural to use asymptotic expressions (at $k \equiv (s/c) \longrightarrow 0$) for I_n, H_n, and f_n:

$$I_n(k) = (n+1) + O(k^2);$$
$$H_n(k) = -n - (2n+1)k + O(k^2); \quad (3.22)$$
$$f_n(k) = -\frac{n}{n+1}\left(1 + \frac{2n+1}{n}k\right) + O(k^2); \quad k^2 \equiv s^2/c^2.$$

Let us write complete expressions for the projections of the electric field strength $\mathbf{E}_j$ onto the axes:

$$E_{2r} = \frac{Q}{r^2} + \sum_n D_n \frac{1}{r}\frac{h_n(kr)}{h_n(k)} n(n+1) P_n(\mu)\exp(st);$$
$$E_{2\Theta} = \sum_n D_n \frac{1}{r}\left[\frac{h_n(kr)}{h_n(k)} + r\frac{d}{dr}\left(\frac{h_n(kr)}{h_n(k)}\right)\right]\frac{dP_n(\mu)}{d\Theta}\exp(st);$$
$$E_{1r} = \sum_n \{D_n f_n + Z_n Q I_n^{-1}\} \times \frac{1}{r}\frac{i_n(kr)}{i_n(k)} n(n+1) P_n(\mu)\exp(st);$$
$$E_{1\Theta} = \sum_n \{D_n f_n + Z_n Q I_n^{-1}\}\frac{1}{r}\left\{\frac{i_n(kr)}{i_n(k)} + r\frac{d}{dr}\left[\frac{i_n(kr)}{i_n(k)}\right]\right\}\frac{dP_n(\mu)}{d\Theta}\exp(st). \quad (3.23)$$

From the discontinuity condition (1.8) of the normal component of the electric induction vector at the free surface of a charged drop, one can find a correction $\kappa_1(\Theta, t)$ to the equilibrium surface charge density $\kappa_0 = \Theta/4\pi$ arising from distortion ξ of the drop sphericity:

$$r = 1: \kappa_1(\Theta, t) = \frac{1}{4\pi}\{-QZ_n(2 + \varepsilon I_n^{-1} n(n+1)) + D_n n(n+1)(1 - \varepsilon f_n)\} P_n(\mu)\exp(st). \quad (3.24)$$

Using the condition of charge balance at the drop surface given by Eq. (1.9), we find a relation between the coefficients D_n, Z_n, $C_n^{(1)}$, and $C_n^{(3)}$:

$$D_n = Z_n B_n G_n(k) + C_n^{(1)} G_n(k) + C_n^{(3)} G_n(k) I_n\left(\sqrt{\frac{s}{\nu}}\right);$$
$$G_n(k, \varepsilon, \sigma, b, s) \equiv \frac{Q}{s - f_n(k)\lambda(\varepsilon, \sigma, b, s)};$$
$$B_n(\varepsilon, \sigma, b, s) \equiv \frac{2s + n\lambda(\varepsilon, \sigma, b, s)}{(n+1)n}; \quad (3.25)$$
$$\lambda(\varepsilon, \sigma, b, s) \equiv \varepsilon s + 4\pi\sigma + bQ(n+1).$$

Here, $I_n(x)$ is the same function as that defined in (3.21) but of different argument: $x \equiv (\sqrt{s/\nu})$. Let us derive an expression for the electromagnetic field pressure on the drop surface. Using a linear approximation with respect to small parameter $|\xi|$, we have

$$p_E = \frac{Q}{8\pi}\sum_n [L_n^{(1)} Z_n + L_n^{(2)} C_n^{(1)} + L_n^{(3)} C_n^{(3)}] P_n(\mu)\exp(st);$$
$$L_n^{(1)} \equiv \frac{Q}{4\pi}[2Q - B_n G_n(k) n(n+1)];$$
$$L_n^{(2)} \equiv -\frac{Q}{4\pi} G_n(k) n(n+1); \quad (3.26)$$

$$L_n^{(3)} \equiv -\frac{Q}{4\pi} G_n(k) I_n\left(\sqrt{\frac{s}{\nu}}\right) n(n+1).$$

To write the dynamic boundary condition for the tangential components of the stress tensor (1.6) using a linear approximation in $|\xi|$, it is sufficient to find the component $\Pi_{2\tau}$ because, according to (1.6) and (3.23), $\Pi_{1\tau}$ is of the second order of smallness in $|\xi|$:

$$r = 1\colon \Pi_{2\tau} \equiv \frac{Q}{4\pi}\left[QZ_n \frac{dP_n(\mu)}{d\Theta} + E_{2\Theta}\right]\exp(st)$$

$$\equiv \frac{Q}{4\pi}\sum_n \left\{ Z_n[Q + G_n B_n H_n(k)] + C_n^{(1)} G_n H_n(k) \right. \quad (3.27)$$

$$\left. + C_n^{(3)} G_n H_n(k) I_n\left(\sqrt{\frac{s}{\nu}}\right)\right\}\frac{dP_n}{d\Theta}\exp(st).$$

4. Substituting expansions (2.8)–(2.10) into (2.11)–(2.13) and taking into account (3.26) and (3.27), we obtain a set of equations for obtaining unknown coefficients in expressions (2.8)–(2.10):

$$Z_n N_n^{(1)} + C_n^{(1)} N_n^{(2)} + C_n^{(3)} N_n^{(3)} = 0;$$
$$Z_n M_n^{(1)} + C_n^{(1)} M_n^{(2)} + C_n^{(3)} M_n^{(3)} = 0;$$
$$Z_n T_n^{(1)} + C_n^{(1)} T_n^{(2)} + C_n^{(3)} T_n^{(3)} = 0;$$
$$N_n^{(1)} \equiv 3; \quad N_n^{(2)} \equiv -n; \quad N_n^{(3)} \equiv -n(n+1);$$
$$M_n^{(1)} \equiv \frac{Q}{4\pi}[Q + G_n B_n H_n(k)];$$
$$M_n^{(3)} \equiv \frac{Q}{4\pi} G_n H_n(k) I_n\left(\sqrt{\frac{s}{\nu}}\right)$$
$$- \nu J_n\left(\sqrt{\frac{s}{\nu}}\right) - \nu(n-1)(n+2);$$
$$M_n^{(2)} \equiv \frac{Q}{4\pi} G_n H_n(k) - 2\nu(n-1);$$
$$T_n^{(1)} \equiv (n-1)(n+2) + \frac{Q}{4\pi}[2Q - B_n G_n n(n+1)]; \quad (4.1)$$
$$T_n^{(2)} \equiv s + 2\nu n(n-1) - \frac{Q}{4\pi} G_n(k) n(n+1);$$
$$T_n^{(3)} \equiv 2\nu\left[I_n\left(\sqrt{\frac{s}{\nu}}\right) - 2\right] n(n+1)$$
$$- \frac{Q}{4\pi} G_n(k) I_n\left(\sqrt{\frac{s}{\nu}}\right) n(n+1);$$

$$I_n\left(\sqrt{\frac{s}{\nu}}\right) \equiv \sqrt{\frac{s}{\nu}} \frac{i_{n+1}\left(\sqrt{\frac{s}{\nu}}\right)}{i_n\left(\sqrt{\frac{s}{\nu}}\right)} + (n+1)$$
$$\equiv g_n\left(\sqrt{\frac{s}{\nu}}\right) + (n+1);$$

$$J_n\left(\sqrt{\frac{s}{\nu}}\right) \equiv \left[\frac{d^2}{dr^2}\frac{i_n\left(\sqrt{\frac{s}{\nu}}\right)}{i_n\left(\sqrt{\frac{s}{\nu}}\right)}\right]_{r=1}$$
$$\equiv \frac{s}{\nu} - 2g_n\left(\sqrt{\frac{s}{\nu}}\right) + n(n+1).$$

Coefficients B_n, G_n, and H_n are defined by relationships (3.22) and (3.25). The set of equations (4.1) is homogeneous and has a solution only when a determinant composed of the coefficient by the unknowns Z_n, $C_n^{(1)}$, and $C_n^{(3)}$ is zero. Equating the determinant to zero, we obtain the dispersion relation for the capillary oscillations:

$$(s - 2g_n\nu)\left\{ s^2 + n(n-1)(n+2) + \frac{Q^2}{2\pi} n \right.$$
$$\left. - \frac{Q}{4\pi} B_n G_n n^2 (n+1) \right\} + 2s^2\nu(n-1)(2n+1)$$
$$- s^2 \frac{Q}{4\pi} G_n\left[H_n g_n - \frac{s}{c}(n+1)(2n+1)\right]$$
$$+ s\nu\frac{Q}{4\pi} G_n\{2g_n n^2(n+1) - n(n-1)(n+1)^2 \quad (4.2)$$
$$+ 4H_n g_n n + H_n n(n^2-1)(n-2)\}$$
$$- 4s\nu^2 g_n n(n-1)(n+2) + \frac{Q}{4\pi} G_n g_n n$$
$$\times\left\{ H_n\left[\frac{Q^2}{2\pi} - (n-1)(n+2)\right] - \frac{Q^2}{4\pi} n(n+1)\right\} = 0.$$

At $c \longrightarrow \infty$, $\nu \neq 0$, σ = const, b = const, and ε = const, this dispersion relation is reduced to that for a charged drop with finite conductivity derived in [9].

It can be shown that for an perfectly conducting liquid in the limit of low viscosity (at $s \gg \nu$), dispersion

relation (4.2) is reduced to

$$s^2 + 2s\nu(n-1)(2n+1) + \frac{s}{c}\frac{Q^2}{4\pi}(n+1)(2n+1) + n(n-1)\left[(n+2) - \frac{Q^2}{4\pi}\right] = 0. \quad (4.3)$$

It is easy to see that at $\nu = 0$ and $c \longrightarrow \infty$, Eq. (4.3) is reduced to the dispersion relation for a charged drop of perfectly conducting inviscid liquid derived by Rayleigh [10]. At $\nu \neq 0$ and $c \longrightarrow \infty$, Eq. (4.3) is reduced to that for a charged drop of perfectly conducting liquid with low viscosity [11]. If the constant c is finite, the third term in (4.3) determines the damping constant χ_n of capillary oscillations of a charged drop of perfectly conducting liquid related to the emission of electromagnetic radiation by this drop:

$$\chi_n = \frac{Q^2}{8\pi c}(n+1)(2n+1). \quad (4.4)$$

Numerical calculations with the use of complete dispersion relation (4.2) allowed us to obtain dependences of the real and imaginary components of the frequency of drop oscillations on the characteristic physical parameters of the phenomenon illustrated in Figs. 1–5. Figure 1 shows the dependences of the real and imaginary components of the dimensionless frequency on the dimensionless conductivity σ at $n = 2$, $\nu = 0.01$, $b = 0.5$, $Q = 1$, and $\varepsilon = 80$. Curve *1* corresponds to capillary oscillations of the drop; curves *2–6*, to aperiodically damped poloidal liquid motion in the drop. Curve *7* characterizes damped aperiodic liquid motion associated with relaxational redistribution of the charge over the drop surface. It is seen that the increase in conductivity results in a linear increase in the damping constant in branch *7*, corresponding to the relaxation motion while not affecting the damping constants of the capillary oscillations and poloidal liquid motions. Neither does the frequency of capillary oscillations depend on the conductivity. It is interesting to note that high-frequency capillary waves can superpose on fast-decaying low-frequency waves resulting from interaction between poloidal liquid motion and liquid motion associated with the movement of charge carriers. In Fig. 2, these motions are represented by branches *8*. Strictly speaking, the liquid motions described by branches *8* are not wavelike because their damping constants in all cases are appreciably in excess of the frequencies; therefore, these motions should be interpreted as aperiodic.

Figure 2 displays dependences of the real and imaginary components of the dimensionless frequency on the dimensionless viscosity ν at $n = 2$, $\sigma = 25$, $b = 0.5$, $Q = 1$, and $\varepsilon = 80$. The numbers of curves *1–8* correspond to the notation in Fig. 1. Curves *9* and *10* correspond to damping constants of capillary–poloidal liquid motion. It is seen that with rising viscosity, the damping constants of branches *1–6* increase.

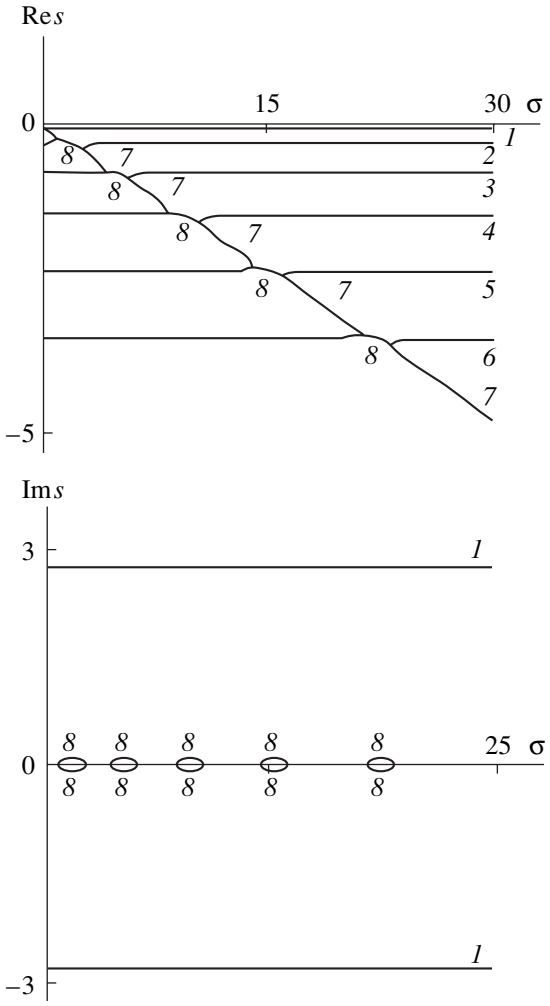


Fig. 1. Real and imaginary components of the dimensionless oscillation frequency as functions of the dimensionless liquid conductivity σ.

In this case, as in Fig. 1, curve *7*, corresponding to relaxation liquid motion, interacts with branches *2–6*, corresponding to poloidal motion, and with branch *10*, corresponding to capillary-poloidal motion. In Fig. 2b, the region of interaction between relaxation branch *7* and poloidal branches *2–5* is shown on a larger scale. Also on a larger scale, Fig. 2c shows the region of interaction between branch *7* and capillary–poloidal branch *10*. It should be noted that the damping constant of relaxation motion *7* associated with the charge redistribution does not vary with viscosity.

In the general case, for a liquid with finite conductivity, the damping constant χ_n of capillary oscillations should depend on the conductivity σ and surface mobility b of charge carriers. Numerical calculations show that the damping constant χ_n of capillary oscillations of a drop associated with emission of electromagnetic radiation by the drop (this damping constant can be extracted from the total damping constant assuming $n = 0$) grows linearly with σ and b in the ranges of con-

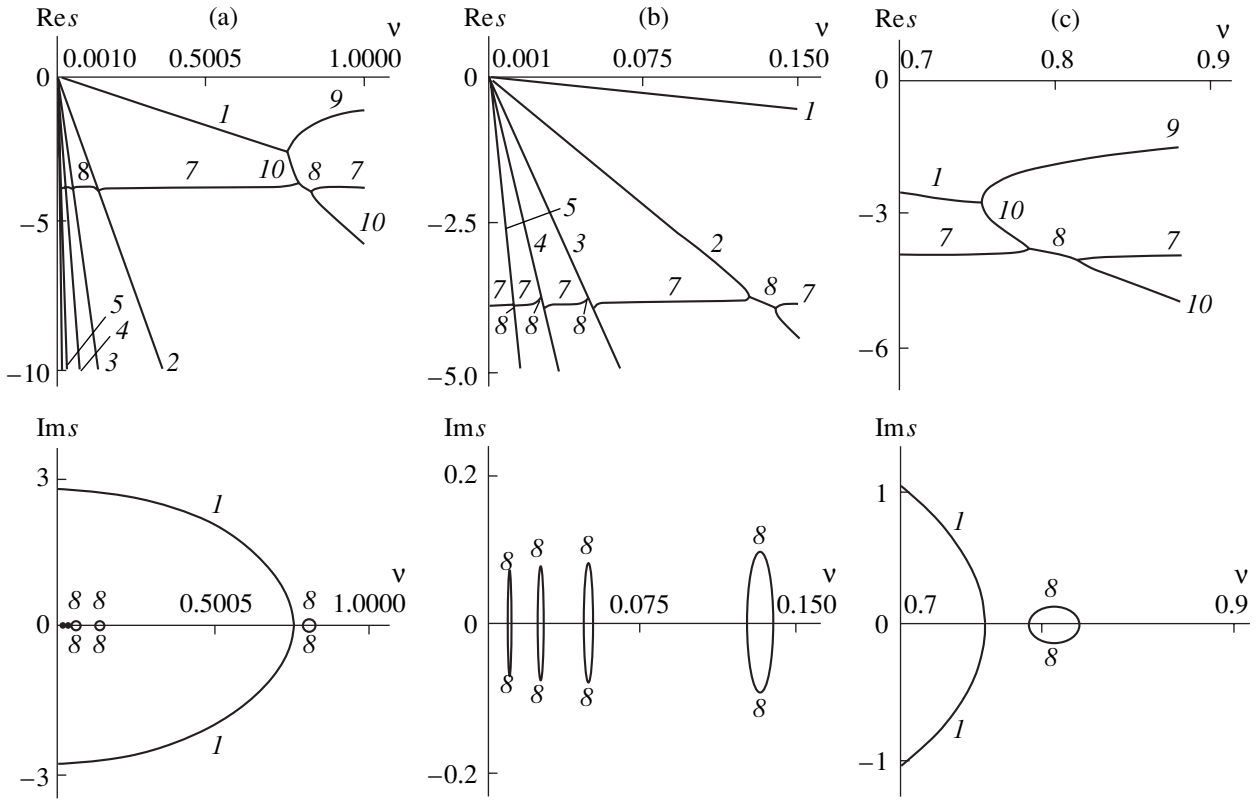


Fig. 2. Real and imaginary components of the dimensionless oscillation frequency as functions of the dimensionless liquid viscosity ν.

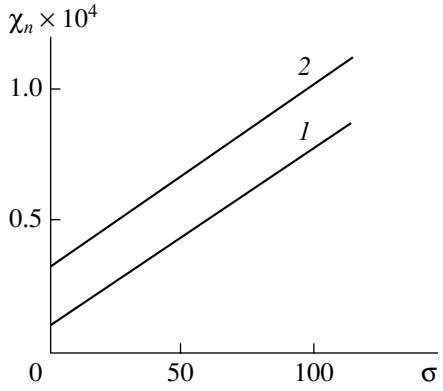


Fig. 3. Dependence of the dimensionless damping constant χ_n of capillary oscillations of a drop related to the electromagnetic drop radiation on the dimensionless liquid conductivity σ. $n = 100$, $Q = 1$, $b = 0.5$, and $\varepsilon = 80$; ν: (*1*) 0, (*2*) 10^{-9}.

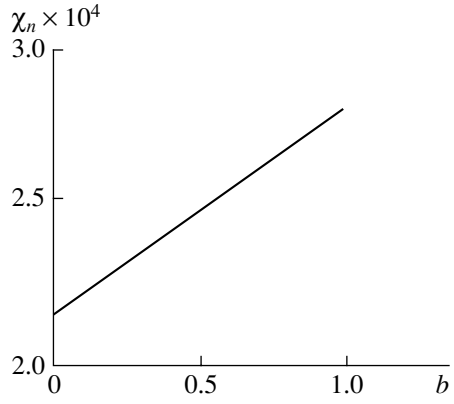


Fig. 4. Dependence of the dimensionless damping constant χ_n of capillary oscillations of the drop related to the electromagnetic emission from the drop on the dimensionless surface charge mobility *b*. $n = 100$, $Q = 1$, $\sigma = 25$, $\varepsilon = 80$, and $\nu = 0$.

ductivity and the charge carrier mobility, which is of interest from the viewpoint of possible applications in geophysics. This can be seen in Figs. 3 and 4, where the corresponding dependences are given for a mode with $n = 100$. Such a high mode is chosen in order for the intensity of electromagnetic radiation, which in this case is characterized by the damping constant $\sim n^2$, to have magnitudes corresponding to applications in geophysics (for example, to electromagnetic radiation from storm clouds [6] or from Saint Elmo's fire arising at the surface of airplanes and causing intensive radiointerference [12]). The intensity of electromagnetic radiation depends on the damping constant related to the electromagnetic radiation and varies as $\sim\chi_n \exp(2\chi_n t)$. In Fig. 3, curve *2* represents the dependence $\chi_n = \chi_n(\sigma)$ calculated numerically for a liquid with low viscosity

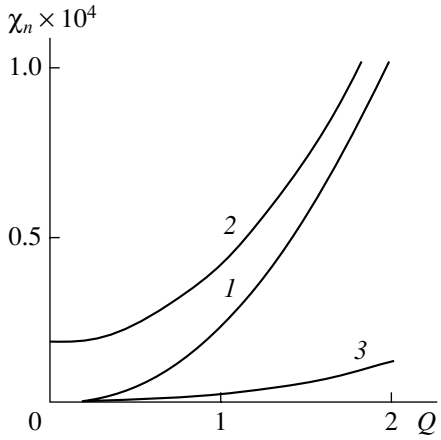


Fig. 5. Dependence of the dimensionless damping constant χ_n of capillary oscillations of the drop related to the electromagnetic emission from the drop on the dimensionless drop charge Q. $n = 100$, $\varepsilon = 80$; (*1*) $\nu = 0$, $b = 0.5$, and $\sigma = 25$; (*2*) $\nu = 10^{-9}$, $b = 0.5$, and $\sigma = 25$; (*3*) $\nu = 0$, $b = 0$, and $\sigma = 0$.

($\nu = 10^{-9}$). It is seen that the contribution to the total damping constant due to viscosity does not depend on σ and the viscosity only causes a vertical shift of the dependence $\chi_n = \chi_n(\sigma)$ for an inviscid liquid.

A numerical calculation using Eq. (4.2) shows that, as the drop charge Q (or the surface charge density κ) increases at finite conductivity σ and mobility b, the damping constant χ_n (Fig. 5) both for an ideal liquid (curve *1*) and for a liquid with low viscosity (curve *2*) grows as a square of Q. Curve *3* in Fig. 5 is the dependence on Q of the damping constant of oscillations related to the electromagnetic emission by a charged drop of inviscid nonconducting liquid ($Q \neq 0$, $\nu = 0$, $\sigma = 0$, and $b = 0$). It appears interesting that the damping constant is nonzero for a purely dielectric liquid, where the charge is "trapped" on the drop surface; that is, a vibrating charged dielectric drop is capable of emitting electromagnetic radiation, although of appreciably lower intensity.

The viscosity ν has no effect on the functional dependence $\chi_n = \chi_n(\sigma, b, Q)$ of the damping constant of capillary oscillations related to emission of electromagnetic radiation by the drop; however, with increasing viscosity, frequencies of surface oscillations decrease until they completely cease and the drop stops emitting electromagnetic radiation at the corresponding frequency. Numerical calculations show that oscillations of a mode with $n = 100$ cease at $\nu \approx 0.05$; of the principal mode with $n = 2$, at $\nu \approx 0.75$.

CONCLUSION

Because of the finite charge redistribution rate and energy loss by capillary oscillations through emission, the damping constants of all modes of capillary oscillations turn out to be greater compared with a liquid drop of infinite conductivity and the structure of the arising liquid motion is more complicated.

The contribution to the damping rate due to the emission of electromagnetic radiation from a charged drop grows with increasing conductivity of the liquid and mobility of the charge carriers according to a linear law. As conductivity of the liquid increases, poloidal eddy motion in the drop interacts with near-surface liquid motion caused by the charge redistribution. The viscosity affects only the damping constants of capillary and poloidal liquid motions in the drop and makes possible their interaction with the near-surface motion caused by the charge redistribution accompanying drop vibrations.

ACKNOWLEDGMENTS

The work was supported by a grant from the President of the Russian Federation (grant no. 00-15-9925).

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. A. I. Grigor'ev, Zh. Tekh. Fiz. **70** (7), 22 (2000) [Tech. Phys. **45**, 543 (2000)].
4. D. F. Belonozhko and A. I. Grigor'ev, Élektrokhim. Obrab. Met., No. 4, 17 (2000).
5. S. O. Shiryaeva, D. F. Belonozhko, V. B. Svetovoĭ, and A. I. Grigor'ev, Preprint No. 31, IMI RAN (Institute of Microelectronics, Russian Academy of Sciences, Yaroslavl, 2001).
6. V. I. Kalechits, I. E. Nakhutin, and P. P. Poluéktov, Dokl. Akad. Nauk SSSR **262**, 1344 (1982).
7. S. O. Shiryaeva, A. É. Lazaryants, *et al.*, Preprint No. 27, IMI RAN (Institute of Microelectronics, Russian Academy of Sciences, Yaroslavl, 1994).
8. *Handbook of Mathematical Functions*, Ed. by M. Abramowitz and I. A. Stegun (National Bureau of Standards, Washington, 1964; Nauka, Moscow, 1979).
9. A. I. Grigor'ev and S. O. Shiryaeva, Izv. Akad. Nauk, Mekh. Zhidk. Gaza, No. 5, 107 (1997).
10. C. D. Hendricks and J. M. Schneider, Am. J. Phys. **31**, 450 (1963).
11. A. I. Grigor'ev and A. É. Lazaryants, Izv. Akad. Nauk SSSR, Mekh. Zhidk. Gaza, No. 5, 11 (1991).
12. A. I. Grigor'ev, I. D. Grigor'eva, and S. O. Shiryaeva, J. Sci. Explos. **5**, 163 (1991).

Translated by N. Mende

Technical Physics, Vol. 47, No. 6, 2002, pp. 682–689. Translated from Zhurnal Tekhnicheskoĭ Fiziki, Vol. 72, No. 5, 2002, pp. 28–34.

GASES AND LIQUIDS

Disintegration of a Drop in an External Electrostatic Field

V. A. Koromyslov, A. I. Grigor'ev, and M. V. Rybakova

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

e-mail: grig@uniyar.ac.ru

Received September 17, 2001

Abstract—Regular features of the disintegration of both a drop of a perfectly conducting liquid and a drop of a dielectric liquid into two or three parts in an external uniform electric field are studied using the principle of minimizing the potential energy of the final state of a closed system with spontaneous processes. © *2002 MAIK "Nauka/Interperiodica".*

1. INTRODUCTION

The problem of the stability of a liquid drop in a uniform electrostatic field is of considerable interest for various physical, geophysical, and engineering applications (see, e.g., reviews [1, 2] and the literature cited therein). As a rule, in solving particular physical or engineering problems associated with the possible disintegration of a drop in an electrostatic field, two main questions ought to be answered. (i) Will the drop be unstable, and, as a result, will it disintegrate? (ii) What is the disintegration mechanism: will the drop disintegrate into several parts of comparable size or will it lose excess charge and generate a large number of highly charged small drops [1, 2]?

One of the problems arising in determining the disintegration mechanism is associated with the fact that, under the action of a uniform electrostatic field, a drop, though stable against disintegration, stretches out to take on a nearly spheroidal shape [3, 4]. This effect complicates analytic treatment because the relevant mathematical expressions are quite lengthy. Solving the problem of the instability of a spheroidal drop by classical hydrodynamic methods (e.g., by mode-by-mode analysis) is usually laborious, and the solution can only be obtained for drops having a small eccentricity [3, 5].

An approach to analyzing the problem under discussion on the basis of the energy principle is simpler. Although this approach cannot be used to obtain exact quantitative characteristics of the disintegration process, it yields reliable qualitative predictions as applied to the problems of the disintegration of a highly charged drop [6–8] and the disintegration of a highly deformed nonspheroidal drop in an external electrostatic field. The problem of the disintegration of an uncharged drop with an arbitrary electrical conductivity in an external uniform electrostatic field **E** in the case of natural spheroidal deformations (such that the equilibrium shape of a drop in the field **E** is close to a prolate spheroid) still remains unsolved.

In the context of what was said above, we investigate the possible disintegration of a drop of an inviscid incompressible liquid into two or three parts. We consider a drop in an external uniform electrostatic field **E** between the plates of a plane capacitor and assume that the drop has the shape of a prolate spheroid stretched out along the field **E**. The remaining basic assumptions are as follows. (i) The drop, together with the electrodes creating the external field and the source of electromotive force (emf), forms a closed system. (ii) In accordance with the principle of minimum potential energy of the final state of a closed system, the change in the total potential energy of the system during the disintegration of the drop is extreme. (iii) The kinetic energy of the parent drop before the disintegration and that of the daughter drops after the disintegration are both zero. (iv) The total volume of a liquid in the parent and daughter drops remains unchanged during the disintegration. (v) The liquid in the drop is in equilibrium with its saturated vapor, so that the mass of the liquid lost by evaporation is negligible.

Hence, we consider a spherical drop of radius R, with the surface tension σ, in a uniform electrostatic field **E**, along which the drop stretches out to take on an equilibrium spheroidal shape with the eccentricity e, satisfying the relationships

$$e = \sqrt{1 - a^2/b^2}, \quad a = R(1 - e^2)^{-1/3},$$

$$b = R(1 - e^2)^{1/6}.$$

Here, a and b are the major and minor semiaxes of the spheroid. In what follows, we consider two opposite limiting cases in which the liquid in the drop is (i) a perfect dielectric or (ii) a perfect conductor.

2. DIELECTRIC LIQUID DROP

2.1. Disintegration of a Drop into Two Parts

The potential energy of the surface-tension forces of a spheroidal drop has the form

$$U_0 = 2\pi\sigma R^2(1 - e^2)^{1/3}\left(1 + \frac{\arcsin e}{e(1 - e^2)^{1/2}}\right).$$

The potential energy of a dielectric spheroid in a uniform electrostatic field **E** is described by the expression [9]

$$U_E = = -\frac{1}{2}\mathbf{P}\cdot\mathbf{E} = -\frac{1}{2}\frac{ab^2}{3}\frac{\varepsilon-1}{1+(\varepsilon-1)k}E^2,$$

where **P** is the dipole moment of a dielectric spheroid polarized by the field **E**, ε is the dielectric constant of a liquid, and the depolarization factor k is equal to [9]

$$k = \frac{1-e^2}{e^3}(\mathrm{Arth}e - e).$$

The total potential energy of a spheroidal dielectric drop in an external uniform electrostatic field **E** is

$$U_0 = 2\pi\sigma R^2(1-e^2)^{1/3} \times\left(1+\frac{\arcsin e}{e(1-e^2)^{1/2}}\right)-\frac{1}{6}\frac{\varepsilon-1}{1+(\varepsilon-1)k}R^3E^2.$$

We suppose that, in a sufficiently strong field **E**, the drop is unstable and disintegrates into two daughter drops of comparable size having a volume equal to the sum of the volumes of spheres of radii R_1 and R_2. Since the parent drop is a perfect dielectric, the daughter drops remain uncharged. Assuming that the eccentricities e_1 and e_2 of the daughter drops depend only on the strength of the external field, we can write the following expression for the potential energy of a system of two daughter drops in the field **E**:

$$U = 2\pi\sigma R_1^2(1-e_1^2)^{1/3}\left(1+\frac{\arcsin e_1}{e_1(1-e_1^2)^{1/2}}\right) + 2\pi\sigma R_2^2(1-e_2^2)^{1/3}\left(1+\frac{\arcsin e_2}{e_2(1-e_2^2)^{1/2}}\right) \qquad (1)$$
$$-\frac{1}{6}\frac{(\varepsilon-1)E^2R_1^3}{1+(\varepsilon-1)k_1}-\frac{1}{6}\frac{(\varepsilon-1)E^2R_2^3}{1+(\varepsilon-1)k_2}-\frac{2p_1p_2}{r^3}.$$

Here, the first and second terms describe the energy of the surface tension of the daughter drops, the third and fourth terms describe the energy of their polarization in the field **E**, the fifth term describes the energy of the dipole interaction between the polarized daughter drops, r is the distance between their centers, and their dipole moments p_1 and p_2 are defined as

$$p_j = -\frac{1}{3}\frac{\varepsilon-1}{1+(\varepsilon-1)k_j}R_j^3E; \quad j = 1, 2. \qquad (2)$$

In expression (1), the parameter r implicitly determines the degree of deformation of the parent drop. During the disintegration, the parent drop is distorted into a dumbbell shape and then splits at the neck into two daughter drops of comparable size.

We introduce the notation $V_1/V = n$, where V_1 is the volume of the first daughter drop and V_1 is the volume of the parent drop. Then, from the volume conservation condition, we find

$$V_2 = (1-n)V \Rightarrow R_1 = n^{1/3}R, \quad R_2 = (1-n)^{1/3}R.$$

We substitute this relationship into expression (1) and take into account expression (2) to obtain

$$U = 2\pi\sigma R^2 n^{2/3}(1-e_1^2)^{1/3}\left(1+\frac{\arcsin e_1}{e_1(1-e_1^2)^{1/2}}\right) + 2\pi\sigma R^2(1-n)^{2/3}(1-e_2^2)^{1/3}\left(1+\frac{\arcsin e_2}{e_2(1-e_2^2)^{1/2}}\right) \qquad (3)$$
$$-\frac{1}{6}\frac{(\varepsilon-1)nE^2}{1+(\varepsilon-1)k_1}-\frac{1}{6}\frac{(\varepsilon-1)(1-n)E^2}{1+(\varepsilon-1)k_2} -\frac{2(n-n^2)}{9r^3}\frac{(\varepsilon-1)^2}{(1+(\varepsilon-1)k_1)(1+(\varepsilon-1)k_2)}R^6E^2.$$

In order to simplify the calculations, we nondimensionalize expression (3) by dividing by $2\pi\sigma R^2$ and introduce the notation

$$U = \frac{U}{2\pi\sigma R^2} \quad x \equiv \frac{r}{R}, \quad w \equiv \frac{E^2R}{2\pi\sigma}. \qquad (4)$$

With this notation, expression (3) becomes

$$U = n^{2/3}(1-e_1^2)^{1/3}\left(1+\frac{\arcsin e_1}{e_1(1-e_1^2)^{1/2}}\right) + (1-n)^{2/3}(1-e_2^2)^{1/3}\left(1+\frac{\arcsin e_2}{e_2(1-e_2^2)^{1/2}}\right) \qquad (3a)$$
$$+\frac{1}{6}\frac{(\varepsilon-1)nw}{1+(\varepsilon-1)k_1}-\frac{1}{6}\frac{(\varepsilon-1)(1-n)w}{1+(\varepsilon-1)k_2} -\frac{2(n-n^2)(\varepsilon-1)^2w}{9x^3(1+(\varepsilon-1)k_1)(1+(\varepsilon-1)k_2)}.$$

It should be taken into account that the eccentricities e_1 and e_2 depend on the strength of the external electric field [4]:

$$e_1 = \frac{3(\varepsilon-1)}{2^{3/2}(\varepsilon+2)}w^{1/2}n^{1/6},$$

$$e_2 = \frac{3(\varepsilon-1)}{2^{3/2}(\varepsilon+2)}w^{1/2}(1-n)^{1/6}.$$

Let us determine the value of the parameter n at which expression (3a) has a minimum. To do this, we

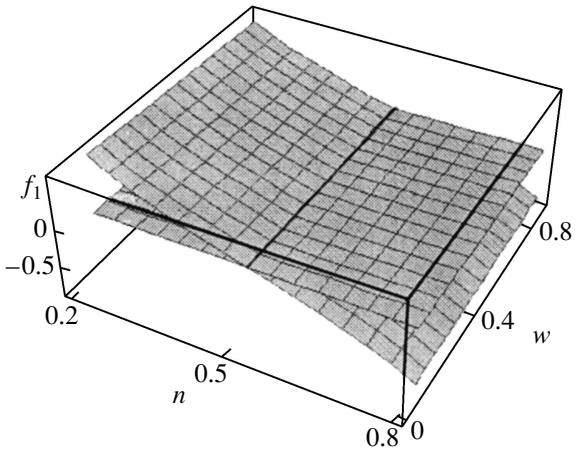


Fig. 1. Surface $f_1 = f_1(n, w)$, which determines the dependence of the first derivative of the potential energy of two dielectric drops on the dimensionless volume n and on the parameter w characterizing the strength of the external electric field. The calculations were carried out for $\varepsilon = 10$.

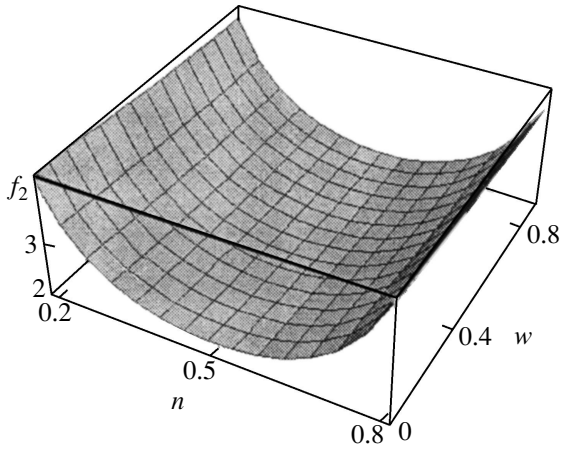


Fig. 2. Dependence of the second derivative $f_2 = f_2(n, w)$ of the potential energy of two dielectric drops on the dimensionless volume n and on the parameter w. The calculations were carried out for $\varepsilon = 10$.

require that the following two conditions be satisfied:

$$\frac{\partial U}{\partial n} = 0 \quad \text{and} \quad \frac{\partial^2 U}{\partial n^2} > 0. \tag{5}$$

The first of conditions (5) yields

$$\begin{aligned} f_1(n, w) \equiv \frac{\partial U}{\partial n} &= \frac{15 - 14e_1^2}{18n^{1/3}(1 - e_1^2)^{2/3}} \\ &+ \frac{(9 - 8e_1^2)\arcsin e_1}{18n^{1/3}e_1(1 - e_1^2)^{7/6}} - \frac{15 - 14e_2^2}{18(1 - n)^{1/3}(1 - e_2^2)^{2/3}} \\ &- \frac{(9 - 8e_2^2)\arcsin e_2}{18(1 - n)^{1/3}e_2(1 - e_2^2)^{7/6}} - \frac{1}{6}\frac{(\varepsilon - 1)w}{1 + (\varepsilon - 1)k_1} \\ &+ \frac{(\varepsilon - 1)^2 n Dk_1 w}{6(1 + (\varepsilon - 1)k_1)^2} + \frac{1}{6}\frac{(\varepsilon - 1)w}{1 + (\varepsilon - 1)k_2} \\ &+ \frac{(\varepsilon - 1)^2(1 - n)Dk_2 w}{6(1 + (\varepsilon - 1)k_2)^2} \\ &- \frac{2(1 - 2n)(\varepsilon - 1)^2 w}{9x^3(1 + (\varepsilon - 1)k_1)(1 + (\varepsilon - 1)k_2)} \\ &+ \frac{2(n - n^2)(\varepsilon - 1)^3 w}{9x^3(1 + (\varepsilon - 1)k_1)(1 + (\varepsilon - 1)k_2)} \\ &\times \left(\frac{Dk_1}{1 + (\varepsilon - 1)k_1} + \frac{Dk_2}{1 + (\varepsilon - 1)k_2}\right) = 0, \end{aligned} \tag{6}$$

where

$$Dk_1 \equiv \frac{\partial k_1}{\partial n} = \frac{3e_1 + (e_1^2 - 3)\operatorname{arctanh} e_1}{6ne_1^3},$$

$$Dk_2 \equiv \frac{\partial k_2}{\partial n} = -\frac{3e_2 + (e_2^2 - 3)\operatorname{arctanh} e_2}{6(1 - n)e_2^3}.$$

Because of the complexity of Eq. (6), it can only be solved numerically. The sought-for function $w(n)$ is a curve of the intersection of the surface $f_1 = f_1(n, w)$ and the $z \equiv 0$ plane. In constructing the surface $f_1 = f_1(n, w)$, there is an uncertainty in the choice of x. The related numerical calculations were carried out under the assumption that the daughter drops touch one another, in which case the x value is equal to the sum of their major semiaxes,

$$x = n^{1/3}(1 - e_1^2)^{-1/3} + (1 - n)^{1/3}(1 - e_2^2)^{-1/3}.$$

Figure 1 shows the intersection of the surface $f_1 = f_1(n, w)$ calculated for $\varepsilon = 10$ and the $z = 0$ plane. We can see that the necessary disintegration condition can be satisfied only for $n = 0.5$, which corresponds to the disintegration of a parent drop into two daughter drops of equal volume.

Figure 2 shows the surface $f_2(n, w) = \partial^2 U/\partial n^2$ in the n and w coordinates. The second of conditions (5) (a sufficient condition) is seen to hold for any n and w values. The explicit analytic expression for the function $f_2 = f_2(n, w)$ is fairly involved; therefore, we do not write it out here. Hence, the two daughter drops into which a dielectric parent drop disintegrates should have the same volume; this corresponds to a "symmetric" disintegration mechanism.

2.2. Disintegration of a Drop into Three Parts

In this section, we assume that two daughter drops of equal volume detach from the opposite sides of an unstable dielectric parent drop in the field **E**. The potential energy of the final state of such a system has the form

$$U = 2\pi\sigma R_1^2(1-e_1^2)^{1/3}\left(1+\frac{\arcsin e_1}{e_1(1-e_1^2)^{1/2}}\right)$$
$$-\frac{1}{6}\frac{\varepsilon-1}{(\varepsilon-1)k_1}R_1^3E^2-\frac{2p_2p_2}{(2r)^3}$$
$$+2\left\{2\pi\sigma R_2^2(1-e_2^2)^{1/3}\left(1+\frac{\arcsin e_2}{e_2(1-e_2^2)^{1/2}}\right)\right. \quad (7)$$
$$\left.-\frac{1}{6}\frac{\varepsilon-1}{1+(\varepsilon-1)k_2}R_2^3E^2-\frac{2p_2p_2}{r^3}\right\}.$$

Here, the first term describes the energy of the surface tension of the central daughter drop, the second term accounts for the energy of its polarization in the field **E**, and the third term describes the energy of the dipole interaction between the two side daughter drops. In parentheses, the first term is the energy of the surface tension of the two side daughter drops, the second term is the energy of their polarization in the field **E**, and the third term is the energy of the dipole interaction of the two side daughter drops with the remainder of the parent drop. The subscript 1 refers to the central daughter drop, and the subscript 2 stands for the side daughter drops. The dipole momenta of the drops are again described by expression (2), in which r is now the distance between the centers of a side daughter drop and the remainder of the parent drop.

We assume that the volume V_2 of a daughter drop is n times smaller than the volume V of the parent drop. From the volume conservation condition, we obtain the relationships

$$R_1 = R(1-2n)^{1/3}, \quad R_2 = Rn^{1/3}.$$

Substituting these relationships into expression (7) and taking into account expression (2), we find

$$U = 2\pi\sigma R^2(1-2n)^{2/3}(1-e_1^2)^{1/3}\left(1+\frac{\arcsin e_1}{e_1(1-e_1^2)^{1/2}}\right)$$
$$-\frac{1}{6}\frac{(\varepsilon-1)(1-2n)}{1+(\varepsilon-1)k_1}R^3E^2-\frac{1}{36r^3}\frac{(\varepsilon-1)^2}{(1+(\varepsilon-1)k_2)^2}n^2R^6E^2$$
$$+2\left\{2\pi\sigma R^2n^{2/3}(1-e_2^2)^{1/3}\left(1+\frac{\arcsin e_2}{e_2(1-e_2^2)^{1/2}}\right)\right.$$

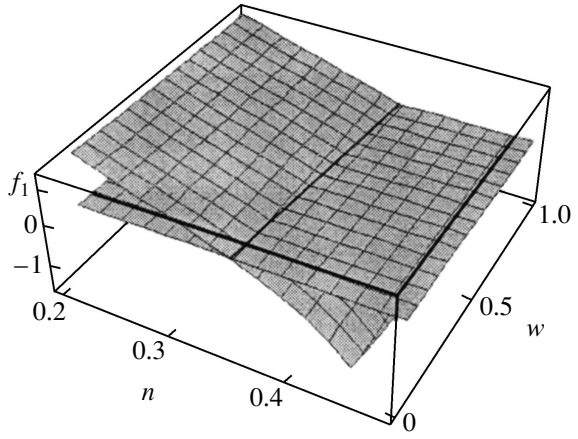


Fig. 3. Same as in Fig. 1, but for disintegration into three dielectric drops.

$$-\frac{1}{6}\frac{(\varepsilon-1)n}{1+(\varepsilon-1)k_2}R^3E^2-\frac{2(n-2n^2)}{9r^3}$$
$$\left.\times\frac{(\varepsilon-1)^2}{(1+(\varepsilon-1)k_1)(1+(\varepsilon-1)k_2)}R^6E^2\right\}.$$

As before, we must take into account the fact that the eccentricities e_1 and e_2 depend on the strength of the external electric field [4]:

$$e_1 = \frac{3(\varepsilon-1)}{2^{3/2}(\varepsilon+2)}w^{1/2}(1-2n)^{1/6},$$
$$e_2 = \frac{3(\varepsilon-1)}{2^{3/2}(\varepsilon+2)}w^{1/2}n^{1/6}.$$

In order to simplify the calculations, we nondimensionalize the potential energy of the final state of the system in the above manner. Then, we use relationships (4) to arrive at

$$U = (1-2n)^{2/3}(1-e_1^2)^{1/3}\left(1+\frac{\arcsin e_1}{e_1(1-e_1^2)^{1/2}}\right)$$
$$+2n^{2/3}(1-e_2^2)^{1/3}\left(1+\frac{\arcsin e_2}{e_2(1-e_2^2)^{1/2}}\right)$$
$$-\frac{1}{6}\frac{(\varepsilon-1)(1-2n)w}{1+(\varepsilon-1)k_1} \quad (8)$$
$$-\frac{1}{3}\frac{(\varepsilon-1)nw}{1+(\varepsilon-1)k_2}-\frac{1}{36r^3}\frac{(\varepsilon-1)^2n^2w}{(1+(\varepsilon-1)k_2)^2}$$
$$-\frac{4(n-2n^2)}{9r^3}\frac{(\varepsilon-1)^2w}{(1+(\varepsilon-1)k_1)(1+(\varepsilon-1)k_2)}.$$

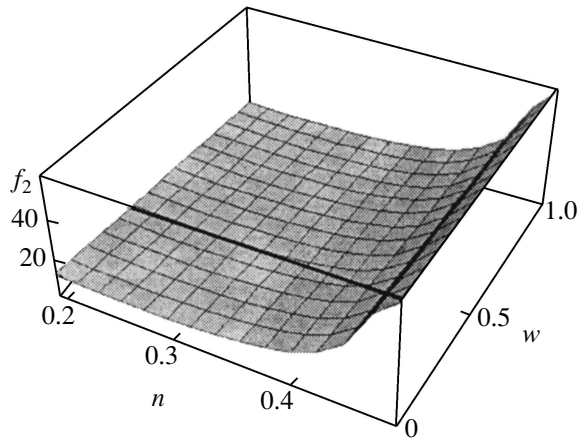


Fig. 4. Same as in Fig. 2, but for disintegration into three dielectric drops.

In order to determine the position where the function $U = U(n)$ is minimum, we require that conditions (5) be satisfied. The first of conditions (5) yields

$$f_1(n, w) \equiv \frac{\partial U}{\partial n} = -\frac{15 - 14e_1^2}{9(1-2n)^{1/3}(1-e_1^2)^{2/3}}$$
$$-\frac{(9-8e_1^2)\arcsin e_1}{9(1-2n)^{1/3}e_1(1-e_1^2)^{7/6}} + \frac{15-14e_2^2}{9n^{1/3}(1-e_2^2)^{2/3}}$$
$$+\frac{(9-8e_2^2)\arcsin e_2}{9n^{1/3}e_2(1-e_2^2)^{7/6}} + \frac{1}{3}\frac{(\varepsilon-1)w}{1+(\varepsilon-1)k_1}$$
$$+\frac{(\varepsilon-1)^2(1-2n)Dk_1 w}{6(1+(\varepsilon-1)k_1)^2} - \frac{1}{3}\frac{(\varepsilon-1)w}{1+(\varepsilon-1)k_2}$$
$$+\frac{(\varepsilon-1)^2 nDk_2 w}{6(1+(\varepsilon-1)k_2)^2} - \frac{(\varepsilon-1)^2 nw}{18x^3(1+(\varepsilon-1)k)^2} \quad (9)$$
$$+\frac{(\varepsilon-1)^3 n^2 wDk_2}{18x^3(1+(\varepsilon-1)k)^2} - \frac{4(1-4n)(\varepsilon-1)^2 w}{9x^3(1+(\varepsilon-1)k_1)(1+(\varepsilon-1)k_2)}$$
$$+\frac{4(n-2n^2)(\varepsilon-1)^3 w}{9x^3(1+(\varepsilon-1)k_1)(1+(\varepsilon-1)k_2)}$$
$$\times\left(\frac{Dk_1}{1+(\varepsilon-1)k_1} + \frac{Dk_2}{1+(\varepsilon-1)k_2}\right) = 0,$$

where

$$Dk_1 \equiv \frac{\partial k_1}{\partial n} = -\frac{3e_1 + (e_1^2-3)\operatorname{arctanh} e_1}{3(1-2n)e_1^3},$$

$$Dk_2 \equiv \frac{\partial k_2}{\partial n} = \frac{3e_2 + (e_2^2-3)\operatorname{arctanh} e_2}{6ne_2^3}.$$

As in the case of disintegration of a dielectric drop into two parts, Eq. (9) was solved numerically. In constructing the function $f_1(n, w)$, we assume that the side daughter drops touch the central daughter drop; i.e., that

$$x = (1-2n)^{1/3}(1-e_1^2)^{-1/3} + n^{1/3}(1-e_2^2)^{-1/3}.$$

The surface $f_1 = f_1(n, w)$ intersected by the $z = 0$ plane is shown in Fig. 3. The calculations were carried out for $\varepsilon = 10$. The intersection of the surface $f_1 = f_1(n, w)$ by the $z = 0$ plane is the sought-for solution to the equation $f_1(n, w) = 0$. As may be seen, in order for disintegration to occur, it is necessary that $n \approx 0.33$, which corresponds to the three daughter drops having an approximately the same volumes. The volume of the side daughter drops is seen to increase insignificantly with the parameter w. Figure 4 shows the surface $f_2(n, w) = \partial^2 U/\partial n^2$ calculated for $\varepsilon = 10$. One can see that the second of conditions (5) is satisfied for any n and w values. The analytic expression for the function $f_2 = f_2(n, w)$ is not presented here because of its complexity. Hence, a dielectric drop disintegrates into three parts by a symmetric mechanism, as is the case with disintegration into two parts.

3. PERFECTLY CONDUCTING DROP

It is well known [1, 10] that, in a strong uniform electrostatic field, a perfectly conducting parent drop disintegrates, emitting a large number of smaller (by two orders of magnitude) daughter drops. For this reason, we restrict ourselves to analyzing the relevant disintegration mechanism. We assume that, during the disintegration, two substantially smaller daughter drops with the same volume and with the charges Q equal in magnitude but opposite in sign detach from the opposite sides of an ellipsoidal parent drop with the eccentricity e_1. Because of their small radii and the high Laplace's pressure in them, the side daughter drops can be assumed to be spherical, in which case the depolarization factor for them reduces to $k_2 = 1/3$ [9].

The energy of a conductor in a uniform electrostatic field is described by the expression [9]

$$U = -\frac{1}{2}\mathbf{P}\cdot\mathbf{E},$$

where the electric dipole moment $\mathbf{P}$ is

$$\mathbf{P} = \mathbf{E}\frac{V}{4\pi k} = \frac{\mathbf{E}R^3}{3k}.$$

Consequently, the potential energy of the final state

of a system of three drops has the form

$$U = -\frac{E^2R_1^3}{6k_1} + 2\pi\sigma R_1^2(1-e_1^2)^{1/3}\left(1+\frac{\arcsin e_1}{e_1(1-e_1^2)^{1/2}}\right) + 2\left\{-\frac{E^2R_2^3}{2} + 4\pi\sigma R_2^2 + U_1 + U_2 + U_3\right\} + U_4. \quad (10)$$

Here, R_1 is the radius of the central daughter drop, R_2 is the radius of the side daughter drops, and the potential energy U_1 of the dipole interaction between the central drop and a side drop is

$$U_1 = \frac{2\mathbf{p}_1\mathbf{p}_2}{r^3}, \quad \mathbf{p}_1 = \frac{\mathbf{E}R_i^3}{3k_i} \Rightarrow U_1 = -\frac{2R_1^3R_2^3}{3r^3k_1}E^2, \quad (11)$$

where p_1 is the dipole moment of the central drop, p_2 is the dipole moment of a daughter drop, and r is the distance between their centers. In expression (10), the electrostatic energy U_2 of a charged daughter drop is equal to

$$U_2 = \frac{Q^2}{2C} = \frac{Q^2}{2R_2}. \quad (12)$$

In order to determine the energy U_3 of the electrostatic interaction between a charged daughter drop and the polarized remainder of the parent drop, we assume that the polarized drop is a dipole. The total polarization charge of half of a spheroid can be obtained by integrating the surface charge density γ:

$$Q^* = \frac{Eb^2}{4k_1}.$$

The distance between the center of the drop and the centroid of the charge is

$$r_{c_x} = \frac{2a}{3}.$$

The axial symmetry of the problem implies that $r_{c_y} = 0$ and $r_{c_z} = 0$.

Hence, the energy of the electrostatic interaction between the polarized parent drop and a charged daughter drop is described by the expression

$$U_3 = -\frac{QQ^*}{(r+r_{c_x})} + \frac{QQ^*}{(r-r_{c_x})} = -\frac{QER_1^3}{3k_1}\left(r^2 - \frac{4R_1^3}{9(1-e^2)^{2/3}}\right)^{-1}. \quad (13)$$

In expression (10), the term U_4 is the potential energy of a dipole formed by the two side charged daughter drops in the field **E**:

$$U_4 = -\mathbf{pE} = -2QEr. \quad (14)$$

Substituting formulas (11)–(14) into expression (10), we obtain

$$U = -\frac{E^2R_1^3}{6k_1} + 2\pi\sigma R_1^2(1-e_1^2)^{1/3}\left(1+\frac{\arcsin e_1}{e_1(1-e_1^2)^{1/2}}\right) + 2\left\{-\frac{E^2R_2^3}{2} + 4\pi\sigma R_2^2 - \frac{2R_1^3R_2^3}{3r^3k_1}E^2 + \frac{Q^2}{2R_2} - \frac{QER_1^3}{3k_1}\left(r_1 - \frac{4R_1^2}{9(1-e_1^2)^{2/3}}\right)^{-1}\right\} - 2QEr. \quad (15)$$

We assume that the volume V_2 of a side drop is n times smaller than the volume V of the parent drop. From the volume conservation condition, we can obtain the relationships

$$R_1 = (1-2n)^{1/3}R, \quad R_2 = n^{1/3}R,$$

in which case we have

$$e_1 = \frac{3w^{1/2}}{2^{3/2}}(1-2n)^{1/6}$$

and the depolarization factor k_1 has the same form as for a dielectric drop.

Under the assumption that the charge of the daughter drops does not depend explicitly on their volume, the potential energy of the system under consideration can be rewritten as

$$U = -\frac{E^2R^3(1-2n)}{6k_1} + 2\pi\sigma R^2(1-2n)^{2/3}(1-e_1^2)^{1/3} \times\left(1+\frac{\arcsin e_1}{e_1(1-e_1^2)^{1/2}}\right) - E^2R^3n + 8\pi\sigma R^2n^{2/3} - \frac{4(n-2n^2)}{3r^3k_1}R^6E^2 + \frac{Q^2}{n^{1/3}R} + \frac{2(1-2n)QER^3}{3k_1}\left(r^2 - \frac{4(1-2n)^{2/3}R^2}{9(1-e_1^2)^{2/3}}\right)^{-1} - 2QEr.$$

After nondimensionalizing according to (4), we take into account the identity $Q \equiv Q/ER^2$ to obtain

$$U = (1-2n)^{2/3}(1-e_1^2)^{1/3}\left(1+\frac{\arcsin e_1}{e_1(1-e_1^2)^{1/2}}\right) + 4n^{2/3} - \frac{w(1-2n)}{6k_1} - wn - \frac{4(n-2n^2)w}{3x^3k_1} + \frac{wQ^2}{n^{1/3}}$$

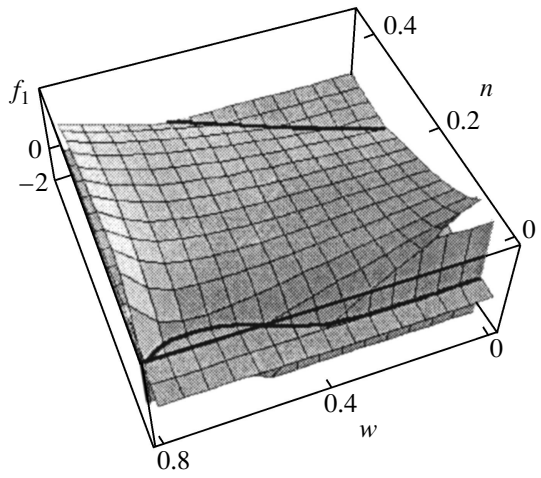


Fig. 5. Same as in Fig. 1, but for disintegration into three perfectly conducting drops.

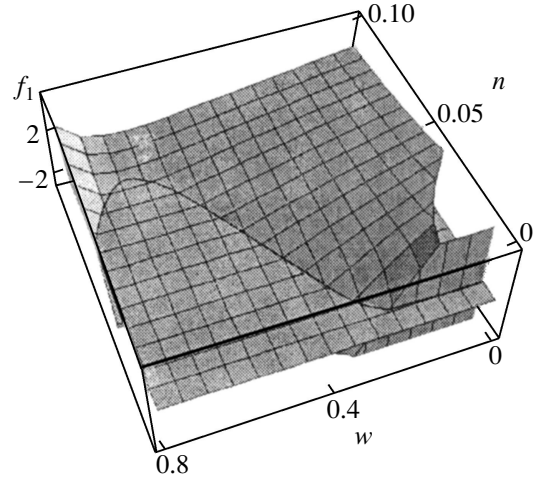


Fig. 6. A magnified fragment of Fig. 5: the region of small n values is reproduced on an enlarged scale.

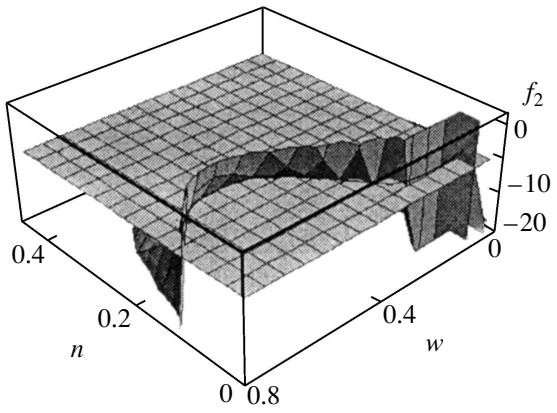


Fig. 7. Same as in Fig. 4, but for disintegration into perfectly conducting drops.

$$+\frac{2(1-2n)Qw}{3k_1}\left(x^2-\frac{4(1-2n)^{2/3}}{9(1-e_1^2)^{2/3}}\right)^{-1}-2Qwx.$$

The position where the function $U = U(n, Q)$ is minimum is determined by the relationships

$$\left(\frac{\partial U}{\partial n}\right)_Q = 0;\quad \left(\frac{\partial U}{\partial Q}\right)_n = 0;\quad \begin{vmatrix} A & B \\ B & C \end{vmatrix} > 0, \qquad (16)$$

in which $A = \partial^2U/\partial n^2$, $C = \partial^2U/\partial Q^2$, $B = \partial^2U/\partial Q\partial n$, and

$$\left(\frac{\partial U}{\partial Q}\right)_n = \frac{2Qw}{n^{1/3}}-\frac{2w(1-2n)}{3k_1}$$

$$\times\left(x^2-\frac{4(1-2n)^{2/3}}{9(1-e_1^2)^{2/3}}\right)^{-1}-2xw = 0.$$

Using these relationships, we arrive at the following expression for the charge of a daughter drop:

$$Q = n^{1/3}\left[x-\frac{1-2n}{3k_1}\left(x^2-\frac{4(1-2n)^{2/3}}{9(1-e_1^2)^{2/3}}\right)^{-1}\right]. \qquad (17)$$

From the second of relationships (16), we find

$$f_1(n,w) = \left(\frac{\partial U}{\partial n}\right)_Q = -\frac{15-14e_1^2}{9(1-2n)^{1/3}(1-e_1^2)^{2/3}}$$

$$-\frac{(9-8e_1^2)\arcsin e_1}{9(1-2n)^{1/3}e_1(1-e_1^2)^{7/6}}+\frac{8}{3n^{1/3}}+\frac{w}{3k_1}$$

$$+\frac{(1-2n)wDk_1}{6k_1^2}-w-\frac{4w(1-4n)}{3x^3k_1}+\frac{4w(n-2n^2)Dk_1}{3x^3k_1^2} \qquad (18)$$

$$-\frac{Q^2w}{3n^{4/3}}+\frac{2wQ}{3}\left[\frac{2k_1+(1-2n)Dk_1}{k_1^2}\left(x^2-\frac{4(1-2n)^{2/3}}{9(1-e_1^2)^{2/3}}\right)^{-1}\right.$$

$$\left.+\frac{16(3-e_1^2)(1-2n)^{2/3}}{81k_1(1-e_1^2)^{5/3}}\left(x^2-\frac{4(1-2n)^{2/3}}{9(1-e_1^2)^{2/3}}\right)^{-2}\right] = 0,$$

$$Dk_1 \equiv \frac{\partial k_1}{\partial n} = -\frac{3e_1+(e_1^2-3)\operatorname{arctanh} e_1}{3(1-2n)e_1^3}.$$

Equations (17) and (18) were solved numerically, the distance between the centers of the parent drop and a side daughter drop being defined as

$$x = (1-2n)^{1/3}(1-e_1^2)^{-1/3}+n^{1/3}.$$

The surface $f_1 = f_1(n, w)$ intersected by the $z = 0$ plane is shown in Figs. 5 and 6. The intersection of the surface $f_1 = f_1(n, w)$ by the $z = 0$ plane is the sought-for solution to the equation $f_1(n, w) = 0$. A sufficient condi-

tion for existence of the minimum is very complicated, so we do not write it out here. Figure 7 shows the intersection of the numerically calculated surface

$$f_2(n, w) = \begin{vmatrix} A & B \\ B & C \end{vmatrix},$$

which characterizes the sufficient condition, and the $z = 0$ plane. As may be seen in Fig. 7, the condition $f_2(n, w) > 0$ is satisfied only for small n values.

Hence, although Eqs. (17) and (18) at fixed w have two solutions (one in the region of small n values and the other in the region of large n values), the only possible mechanism for the disintegration of a perfectly conducting drop is by the emission of small daughter drops, in which case the radius of the produced drops increases with electric field strength.

4. CONCLUSION

A dielectric drop can only disintegrate into parts of comparable size. A drop of a perfectly conducting liquid in a strong electrostatic field **E** disintegrates, emitting many very small daughter drops whose radii are two orders of magnitude smaller than the radius of the parent drop.

REFERENCES

1. A. I. Grigor'ev and S. O. Shiryaeva, Izv. Akad. Nauk SSSR, Mekh. Zhidk. Gaza, No. 3, 3 (1994).
2. D. F. Belonozhko and A. I. Grigor'ev, Élektrokhim. Obrab. Met., No. 4, 17 (2000).
3. G. Taylor, Proc. R. Soc. London, Ser. A **280**, 383 (1964).
4. A. I. Grigor'ev, S. O. Shiryaeva, and E. I. Belavina, Zh. Tekh. Fiz. **59** (6), 27 (1989) [Sov. Phys. Tech. Phys. **34**, 602 (1989)].
5. S. O. Shiryaeva, Zh. Tekh. Fiz. **69** (8), 28 (1999) [Tech. Phys. **44**, 894 (1999)].
6. V. A. Koromyslov, A. I. Grigor'ev, and S. O. Shiryaeva, Zh. Tekh. Fiz. **68** (8), 31 (1998) [Tech. Phys. **43**, 904 (1998)].
7. S. I. Shchukin and A. I. Grigor'ev, Zh. Tekh. Fiz. **70** (4), 1 (2000) [Tech. Phys. **45**, 381 (2000)].
8. S. I. Shchukin and A. I. Grigor'ev, Zh. Tekh. Fiz. **70** (6), 27 (2000) [Tech. Phys. **45**, 694 (2000)].
9. L. D. Landau and E. M. Lifshitz, *Electrodynamics of Continuous Media* (Nauka, Moscow, 1982; Pergamon, New York, 1984).
10. A. I. Grigor'ev and S. O. Shiryaeva, J. Phys. D **23**, 1361 (1991).

Translated by O. Khadin

Technical Physics, Vol. 47, No. 6, 2002, pp. 690–694. Translated from Zhurnal Tekhnicheskoĭ Fiziki, Vol. 72, No. 5, 2002, pp. 35–40.

GAS DISCHARGES, PLASMA

Deposition of a-C:H Films in a DC Glow Discharge with a Magnetron Plasma Localized near the Anode

E. A. Konshina

Vavilov State Optical Institute, All-Russia Research Center, Birgevaya Linia 12, St. Petersburg, 199034 Russia

e-mail: konshina@soi.spb.su

Received March 23, 2001; in final form, August 21, 2001

Abstract—A multielectrode CVD system having a region of crossed magnetic and electric fields is used for the deposition of a-C:H films in a dc glow-discharge plasma. The *I–V* characteristics of the discharge are taken in the pressure range from 0.004 to 0.1 Pa. The effect of magnetic field on the *I–V* characteristics is investigated. The film deposition rate as a function of discharge power, hydrocarbon vapor pressure, conductivity and potential of the substrate, and inert gas is studied. The results obtained are discussed in terms of the adsorption mechanism of film condensation in a plasma. The effect of deposition rate on the film properties is analyzed. © *2002 MAIK "Nauka/Interperiodica"*.

INTRODUCTION

CVD in a glow discharge plasma is common in the production of amorphous hydrogenated carbon (a-C:H) films. In this method, hydrocarbon decomposition products condense in a positive-ion plasma. The glow discharge plasma is produced with rf-biased [1–7] or dc [8–10] diode-type systems. Inertia of ions (in comparison with electrons) in an rf discharge decreases their effect on the growing film. The efficiency of gas ionization can be improved by placing a grid negatively biased with respect to the rf plasma near the substrate [11] and also by using a magnetic field perpendicular to the rf plasma electric field [12].

In this work, we study the performance of a multielectrode device designed for the deposition of a-C:H films. A particular feature of this device is that a magnetron plasma localized near the anode coexists with a dc glow discharge plasma in a quasi-closed volume. The effect of interelectrode voltage, discharge power, gas pressure, and other factors on the rate of a-C:H film deposition is analyzed.

EXPERIMENTAL

(i) *Device for film deposition in a glow discharge plasma*. The schematic representation of the multielectrode device for the deposition of a-C:H films is shown in Fig. 1. Localized plasma *1*, produced by crossed magnetic and electric fields, sustains a glow discharge in volume *2*. We used annular permanent magnet *3* and two planar annular electrodes: anode *4* and cathode *5*. The magnetic field intensity near the cathode surface was about 600 G. Additional electrode *6* and substrate-carrying electrode *7* were mounted on glass cylinders with a diameter of about 130 mm. The cylinder walls confine the plasma inside the quasi-closed volume of the vacuum chamber. The ring-shaped gap between the magnet and the cathode is used as a gas inlet. Such a design of the device provides uniform gas distribution and efficient gas consumption, as well as reduces the film contamination by foreign impurities. Cathode *4* and anode *5* are placed in the neighborhood of the permanent magnet to provide maximum intensity and uniformity of the magnetic field in the crossed field region. The spacing between cathode *4* and substrate holder *7* is about 50 mm in the absence of the additional electrode. This electrode is placed 30 mm away from the cathode. This device allows the deposition of films with a thickness nonuniformity of no more than 20% on substrates of diameter up to 100 mm.

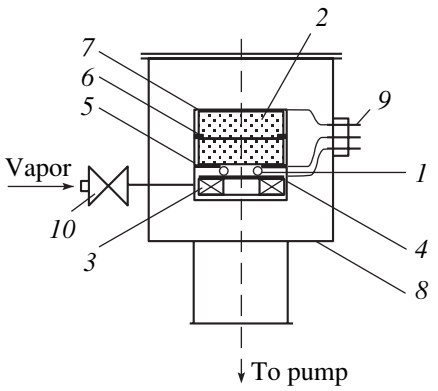


Fig. 1. Multielectrode device: *1*, localized plasma; *2*, glow discharge plasma; *3*, permanent magnet; *4*, cathode; *5*, anode; *6*, additional electrode; *7*, substrate holder; *8*, vacuum chamber; *9*, current leads; and *10*, valve.

1063-7842/02/4706-0690$22.00 © 2002 MAIK "Nauka/Interperiodica"

The device is placed into vacuum chamber *8* having current leads *9* and adjustable valve *10* for gas supply (Fig. 1). In the experiments, the anode was positively biased with a standard dc power unit. The cathode was under the ground potential. The substrate holder was either negatively biased with a dc power unit or under the ground potential. The additional electrode was used as a substrate holder when a-C:H films were deposited at a glancing angle [13]. The vacuum chamber was evacuated to $(1–5) \times 10^{-3}$ Pa by a rotary backing pump and a turbomolecular pump. The gas pressure in the vacuum chamber was measured by a vacuum ionization gage.

The vapor of the material to be deposited was delivered to the vacuum chamber through the adjustable valve and the receiver pre-evacuated by the backing pump. When the voltage is applied to the electrodes, a high-density toroidal plasma arises as a result of the effective electron capture by a magnetic trap, followed by gas ionization in the discharge gap between the anode and the substrate holder. Ion collisions with the solid surface confining the plasma generate extra electrons which also take part in gas ionization. a-C:H films were deposited on polished copper and glass substrates covered by thin conducting and semiconducting layers.

(ii) *Performance of the device*. Figure 2 shows the ion current I through substrate holder *7* (Fig. 1) vs. acetylene pressure P in the vacuum chamber. The glow-discharge acetylene plasma was obtained with the device described above. When the plasma is sustained by the plasma localized in the device, the ion current varies only slightly with the gas pressure (varying from 0.01 to 0.04 Pa) at a constant voltage U (curves *a* and *b* in Fig. 2). An increase in the voltage from 600 to 800 V leads to a twofold increase in the current through the substrate holder. In the absence of the permanent magnet, the glow discharge was initiated at a higher pressure, 0.06 Pa (curve *c* in Fig. 2), and at a higher voltage, 1200 V. In this case, the ion current depends more strongly on the gas pressure and increases with it. Figure 2 suggests that the localized plasma generated in the region of the crossed electric and magnetic fields has a substantial effect on the properties of the discharge.

The experimental voltage dependences of the ion current to the substrate are presented in Fig. 3a. The curves increase following the law $I \sim U^a$. The exponent a defining the gas ionization efficiency varies nonmonotonically in the pressure range from 0.004 to 0.1 Pa (Fig. 3b). The gas ionization is more effective at pressures near 0.03 Pa. The ion current decreases both when the pressure decreases (due to a reduction in the gas molecule concentration in the quasi-closed volume) and when the pressure grows up to 0.05 Pa. With the additional electrode (at the ground potential) introduced, the ion current to the substrate decreases (line *4*) in comparison with its value at the same pressure (line *3* in Fig. 3a). The same slope of straight lines *3* and *4* in Fig. 3a indicates that the gas ionization efficiency remains constant.

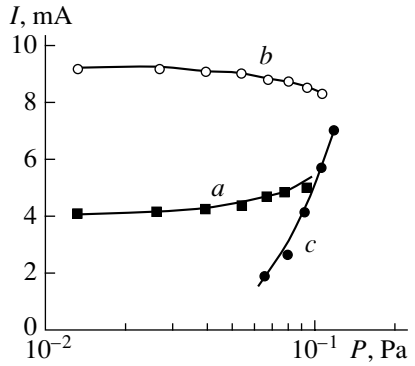


Fig. 2. Ion current through the substrate holder vs. acetylene pressure in the vacuum chamber. Glow discharge is sustained (*a*, *b*) by the localized plasma and (*c*) without it. U = 600 (*a*), 800 (*b*), and 1200 V (*c*).

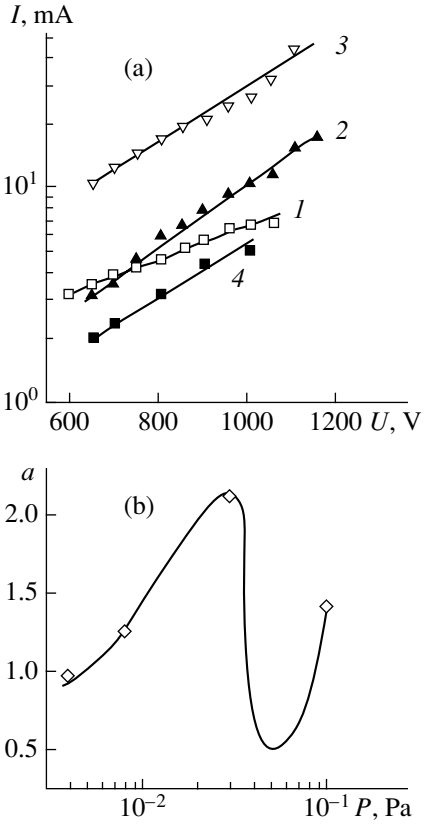


Fig. 3. (a) Ion current vs. voltage for the three-electrode arrangement. Acetylene pressure P = 0.004 (*1*), 0.03 (*2*), 0.05 (*3*), and 0.05 Pa (with an additional electrode) (*4*); (b) exponent a in the law $I \sim U^a$ vs. pressure.

The performance analysis of our multielectrode device showed that it operates in a wider pressure range (from 0.004 to 0.1 Pa) in comparison with standard diode-type systems. This greatly extends the range of operating conditions for the deposition of a-C:H films

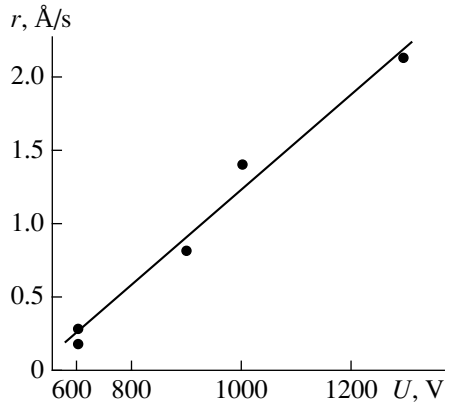


Fig. 4. Rate of a-C:H film deposition on the glass substrates vs. interelectrode voltage.

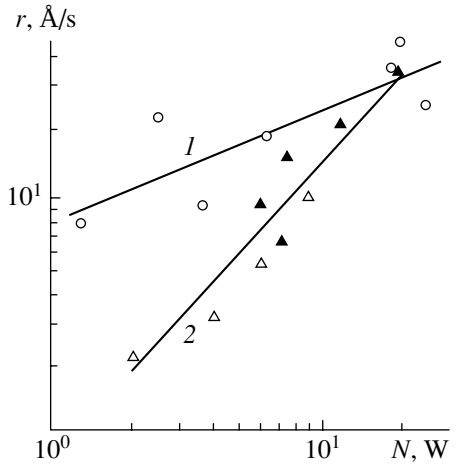


Fig. 5. Rate of a-C:H film deposition on the copper substrates vs. discharge power at acetylene pressures $P \geq 0.1$ (○) and $P \leq 0.05$ Pa (▲). △, deposition from the acetylene–argon plasma.

in a glow discharge plasma. The ion current to the substrate can be varied from 2 to 40 mA at interelectrode voltages from 600 to 1200 V. Then, the discharge power N dissipated by positive ions at the substrate holder will vary between 1.2 and 44 W.

FACTORS GOVERNING THE RATE OF a-C:H FILM DEPOSITION

(i) *The effect of glow discharge parameters*. Chemical reactions and physical processes at the surface of a-C:H films during their deposition in a low-temperature hydrocarbon plasma were considered in terms of the adsorbed layer model [14]. This model assumes that CH_3 plasma radicals are physically adsorbed on the surface and then pass into the chemisorbed state as a result of cross linking due to energetic ions. The surface coverage depends on the number of surface states and surface temperature. When a-C:H films are deposited with the described device, their surface is continuously bombarded by positive ions whose energy depends largely on the interelectrode voltage. The ion energy may be high enough to consolidate the condensate either by cross linking or by decomposing weakly bound particles with subsequent surface diffusion and desorption of the decomposition products. The balance of the processes conducive to film deposition and etching defines the deposition rate. The film deposition rate can also be varied by properly selecting the glow discharge parameters (gas pressure, ion current to the substrate, and interelectrode voltage), which control the number of ions and their energy.

The rate (r) of film deposition on the surface of the glass substrates in the acetylene plasma as a function of the interelectrode voltage is shown in Fig. 4. The deposition rate was defined as the ratio of the a-C:H film thickness to the deposition time. The thickness was measured by an MII-4M microinterferometer with an accuracy of 10%. The thickness of a-C:H films was found to be 0.1–0.3 μm. As the voltage increases from 700 to 1300 V, the deposition rate varies from 0.25 to 2 Å/s.

The experimental dependences of the deposition rate on the discharge power for the a-C:H films obtained at a vapor pressure of ≈0.1 Pa (straight line *1*) and ≈0.05 Pa (straight line *2*) are shown in Fig. 5. The films were deposited on the copper substrates from the acetylene and acetylene–argon plasma at room temperature. When N is varied from 1 to 20 W, the rate varies from 1 to 35 Å/s. The different slopes of straight lines *1* and *2* in Fig. 5 demonstrate the effect of pressure on the rate of film condensation. At high pressures (line *1*), the fraction of energetic ions involved in the condensation decreases because of energy losses due to collisions with gas molecules. This must decrease the rate of ion-stimulated cross linking and lead to the formation of polymer-like a-C:H films containing a considerable amount of bound hydrogen. For low pressure (line *2* in Fig. 5), the power dissipated by the substrates approaches the power applied to the discharge because the energy lost by the ions in the plasma is smaller. The content of hydrogen in the films obtained under these conditions decreases, and the interelectrode voltage defines the ratio of carbon atoms in the sp^3 and sp^2 states in the film structure [15].

(ii) *The effect of an inert gas*. The addition of an inert gas to acetylene decreases the deposition rate of the a-C:H films for the same pressure in the vacuum chamber. The film deposition rates in the acetylene–argon plasma are shown in Fig. 5 (symbols △). The reduction of the rates is caused by a decrease in the density of hydrocarbon radicals in the plasma and the enhancement of the desorption of weakly bound particles because of surface bombardment by argon ions, which do not carry the material being deposited and do not combine with the condensate.

The dependence of the deposition rate on the acetylene volume concentration in a mixture with krypton is

shown in Fig. 6 for a constant discharge power of 1.8 W and a pressure of 0.05 Pa. The rate of film deposition on the copper substrates can be reduced from 4 to 0.5 Å/s by decreasing the relative content of acetylene in the mixture to 20%. The ratio between the carboniferous particles and inert gas ions is responsible for two competitive processes: deposition and etching of the film. The carboniferous particles (neutral molecules, radicals, and ions) present in the plasma favor film growth. On the contrary, film bombardment by inert gas ions etches the surface during the condensation and lowers the deposition rate.

(iii) *The effect of substrate conductivity.* Figure 7 illustrates the effect of surface conductivity on the deposition rate at an acetylene pressure in the vacuum chamber of about 0.05 Pa. For the same discharge power, the rate of film deposition on the surface of a transparent conducting indium–tin oxide (ITO) layer is one order of magnitude higher than that on the surface of a-Si:C:H semiconductor layer with a resistivity of about 10^{12} Ω cm.

Space charge is produced on the substrate surface when it is bombarded by positive ions in a dc glow discharge plasma. Charge leakage from the surface depends on the substrate conductivity, as well as on the thickness and resistivity of the growing film. For a metallic substrate, the formula $d = R_D/\rho_F$ yields a film thickness d at which the surface charge is negligible. Here, R_D is the resistance of the discharge gap and ρ_F is the resistivity of the film [16]. For $R_D \approx 3 \times 10^5$ Ω, the film of thickness ≈0.3 μm and a resistivity $\rho_F = 10^8$ Ω cm does not prevent charge leakage. As the resistivity grows, the critical thickness and the deposition rate decrease. Therefore, the surface conductivities of the substrate and condensate in a glow discharge plasma should be considered as important factors affecting the kinetics of a-C:H film deposition.

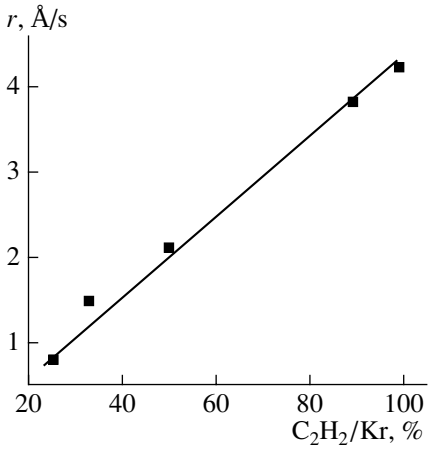


Fig. 6. Rate of a-C:H film deposition on the copper substrates vs. acetylene volume content in the acetylene–argon mixture.

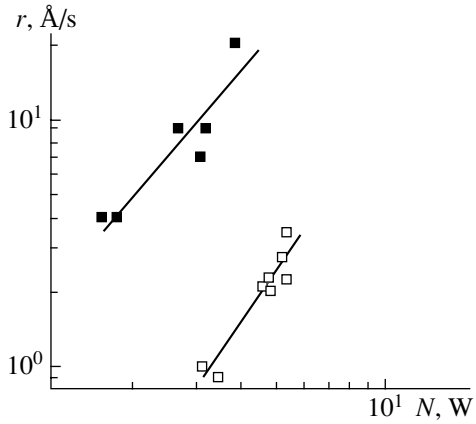


Fig. 7. Rate of a-C:H film deposition on the In_2SnO_3 transparent conducting electrode (■) and a-Si:C:H semiconducting layer (□) at the acetylene pressure $P \approx 0.05$ Pa vs. discharge power.

PROPERTIES OF a-C:H FILMS AND THEIR APPLICATIONS

A correlation between the deposition rate and the properties of a-C:H films was studied by the author in [17–20]. Here, we will briefly review the basic results. When the rate of a-C:H film deposition on glass substrates ranges from 0.4 to 2.5 Å/s, the refractive index varies from 2.4 to 2.0 and the absorption coefficient decreases from 0.3 to 0.1 at a wavelength of 633 nm [17, 18]. A fourfold decrease in the deposition rate results in a decrease in the optical gap width from 2.1 to 1.1 eV [19] if argon is added to acetylene in amounts of 50–70 vol. %. The resistivity of the films can be varied by six orders of magnitude, starting from ~10^7 Ω cm, if the deposition rate ranges from 1 to 10 Å/s [20]. An increase in the absorption in the visible range for a-C:H films obtained from acetylene is accompanied by optical gap narrowing and a rise in the film conductivity. The correlation between the absorption coefficient at a wavelength of 633 nm, optical gap width, and resistivity as functions of the deposition rate is discussed in [21].

a-Si:H deposits obtained by the above method have various optical, electrical, surface, and other properties and can be used in different optical devices. The mechanically strong and chemically stable amorphous coatings, which are transparent in the IR region and have the refractive index 2–2.4 and an absorption coefficient of ≈0.02 at a wavelength of 10.6 μm, raise the optical breakdown threshold of laser mirrors. The effect of various factors on the breakdown threshold of copper mirrors with such a coating is analyzed at length in [22]. For this purpose, the optimal deposition rate is about 2 Å/s.

The thin transparent films with a low refractive index (about 1.6) obtained from toluene and octane were used as layers orienting liquid crystals (LCs) [13].

The interaction of LC molecules with the relatively smooth surface of a-C:H layers is associated with dispersion intermolecular interaction at the solid body–liquid crystal interface. To deposit such films, the device with the additional electrode was used. In this arrangement, the substrates were tilted to the electrode surface. The ions struck the surface at a glancing angle, which modified greatly the conditions for their interaction with the growing film. An appreciable difference in the orienting properties of the layers deposited at negative and floating potentials was found. When the ions were involved in the condensation (the surface is under a negative bias), the films aligned LC molecules parallel to the substrate surface [23]. At the same time, the films obtained at a floating potential oriented the same LC molecules normally to the surface in the same device geometry [24]. IR spectroscopy studies showed that a-C:H orienting layers provide stable orientation of LC molecules [25].

Black a-C:H films, absorbing in the visible range, were used for optical isolation between writing and reading light beams in optically addressed LC reflection modulators. The light-absorption efficiency of a-C:H films on a-Si:H and a-Si:C:H optically sensitive semiconducting layers was demonstrated in [26]. A 1-μm-thick a-C:H film with an absorption coefficient of $\approx 5 \times 10^4$ cm^{-1} at a wavelength of 633 nm reduces the incident light intensity by a factor of 100. The application of a-C:H absorbing films in LC modulators would be a step forward in solving the problem of optical isolation [27].

CONCLUSION

This paper generalizes the results of investigations into a-C:H film fabrication technology. The films are obtained in a dc glow-discharge plasma with an original multielectrode device with the magnetron plasma localized near the anode. The device allows one to deposit the films in a wider range of working pressures (from 0.1 to 0.004 Pa) in comparison with standard diode systems. The dependences of the deposition rate on the interelectrode voltage, discharge power, and relative content of an inert gas in the mixture with a hydrocarbon are obtained. Knowing them, one can control the process of condensation. By varying the deposition rate from 0.25 to ≈30 Å/s, one can modify the structure and properties of the films from polymer- to diamond-like. The results obtained are expected to be of interest in the areas of applied physics where thin a-C:H films are used.

REFERENCES

1. I. P. Anderson, S. Berg, H. Norstrom, *et al.*, Thin Solid Films **63**, 155 (1979).
2. D. S. Whitmel and R. Williamson, Thin Solid Films **35**, 255 (1976).
3. L. Holland and S. M. Ojha, Thin Solid Films **58**, 107 (1979).
4. B. Discler, A. Bubenzer, and P. Koidl, Solid State Commun. **48**, 105 (1983).
5. D. I. Jones and A. D. Stewart, Philos. Mag. B **45**, 432 (1982).
6. K. Enke, H. Dimigen, and H. Hubach, Appl. Phys. Lett. **36**, 291 (1980).
7. I. Watanabe, S. Hasegawa, and Y. Kurata, Jpn. J. Appl. Phys. **21**, 856 (1982).
8. D. R. McKenzie, R. C. McPhedran, and D. J. H. Cockayne, Thin Solid Films **106**, 247 (1983).
9. B. Meyerson and F. W. Smith, J. Non-Cryst. Solids **35/36**, 435 (1980).
10. A. V. Balakov and E. A. Konshina, Zh. Tekh. Fiz. **52**, 810 (1982) [Sov. Phys. Tech. Phys. **27**, 521 (1982)].
11. T. Mori and Y. Namba, J. Vac. Sci. Technol. A **1**, 23 (1983).
12. C. Weismantel, K. Bewilogua, K. Bzeuer, *et al.*, Thin Solid Films **96**, 31 (1982).
13. E. A. Konshina, V. M. Tolmachev, A. I. Vangonen, and D. A. Fatkulina, Opt. Zh. **64** (5), 88 (1997) [J. Opt. Technol. **64**, 476 (1997)].
14. W. Jacob, Thin Solid Films **326**, 1 (1998).
15. E. A. Konshina, Opt.-Mekh. Prom-st **54** (2), 15 (1987) [Sov. J. Opt. Technol. **54**, 80 (1987)].
16. M. Ludwig, Appl. Opt. **25**, 3977 (1986).
17. E. A. Konshina and V. A. Tolmachev, Zh. Tekh. Fiz. **65** (1), 175 (1995) [Tech. Phys. **40**, 97 (1995)].
18. V. A. Tolmachev and E. A. Konshina, Diamond Relat. Mater. **5**, 1397 (1996).
19. E. A. Konshina, Fiz. Tekh. Poluprovodn. (St. Petersburg) **33**, 469 (1999) [Semiconductors **33**, 451 (1999)].
20. E. A. Konshina, Zh. Tekh. Fiz. **70** (3), 87 (2000) [Tech. Phys. **45**, 374 (2000)].
21. E. A. Konshina and N. A. Feoktistov, J. Phys. D **34**, 1131 (2001).
22. E. A. Konshina, Zh. Tekh. Fiz. **68** (9), 59 (1998) [Tech. Phys. **43**, 1062 (1998)].
23. E. A. Konshina, Kristallografiya **40**, 1074 (1995) [Crystallogr. Rep. **40**, 999 (1995)].
24. E. A. Konshina, V. M. Tolmachev, and A. I. Vangonen, Kristallografiya **43**, 107 (1998) [Crystallogr. Rep. **43**, 98 (1998)].
25. E. A. Konshina and A. I. Vangonen, Opt. Zh. **65** (7), 39 (1998) [J. Opt. Technol. **65**, 543 (1998)].
26. E. A. Konshina and A. P. Onokhov, Zh. Tekh. Fiz. **69** (3), 80 (1999) [Tech. Phys. **44**, 340 (1999)].
27. A. P. Onokhov, E. A. Konshina, N. A. Feoktistov, *et al.*, Ferroelectrics **246**, 259 (2000).

Translated by M. Astrov

Technical Physics, Vol. 47, No. 6, 2002, pp. 695–697. Translated from Zhurnal Tekhnicheskoĭ Fiziki, Vol. 72, No. 5, 2002, pp. 41–43.

SOLID-STATE ELECTRONICS

Preparation, Properties, and Applications of Laser-Prepared Intercalated Structures in Nanoelectronics

I. M. Budzulyak, I. I. Grigorchak, B. K. Ostafiychuk, and L. S. Yablon

Stefanik University, Ivano-Frankovsk, 76025 Ukraine

e-mail: bo inter@pu.if.ua

Received May 14, 2001

Abstract—Experimental investigation of the structure and electrical properties of layered gallium selenide and indium selenide crystals laser-intercalated by copper and gallium is discussed. © *2002 MAIK "Nauka/Interperiodica".*

Considering intercalation (i.e., the embedding of ions, atoms, or molecules into "guest" sites) as a branch of crystal engineering, one can suppose that it can form a conceptual basis for the nanotechnology of low-dimension materials. Nanotechnologies are expected to replace today's microtechnologies, whose potential has been exhausted. There is no doubt that incipient physical nanoelectronics will closely be related to a breakthrough in developing new processes and active devices. The intercalation methods known can be subdivided by convention into three groups: thermal exposure, element-selective, and electrochemical [1]. It is meant that an intercalant occupies guest vacancies, penetrating from its liquid phase, from its vapor phase (when being incorporated into a hetero-organic complex) by a direct or an exchange reaction, or in the form of an ion (radical ion) during electrode polarization, respectively. At present, a wide spectrum of intercalates can be obtained. However, they are inappropriate for rapidly changing process conditions, especially in the case of multistage processes, and for effectively scanning sample interiors by guest components, because the associated processes are lasting and selective. In a number of cases, the above methods imply thermal load on the "host" material because of the low vapor pressure of some metals, concentration threshold of intercalation [2], and kinetic difficulties associated with mass transfer in solids. However, heating deteriorates the stability of the crystal lattice.

The aim of this study is to eliminate these disadvantages by applying a new intercalation method. We suggest to enhance the process of intercalation by using laser irradiation of a thin absorbing film of a guest component deposited on one of the host material faces, for example, perpendicularly to the substrate, up to formation of the layered crystal structure. The irradiation (shots with an energy density of 2–20 J/cm^2) can be performed from both the film side and the opposite side if the host material is transparent to laser radiation. Absorbing the laser radiation, the guest vaporizes in 2–8 ms and, due to high concentration and temperature gradients, penetrates into guest sites which are situated between atomic planes coupled by weak van der Waals forces. Thus, by varying the film thickness and the radiation energy density, one can minimize the thermal load on the host material and the amount of the guest can conveniently be controlled by the number of laser shots. Since films of some intercalants are easy to deposit to a desired thickness and with a given distribution by well-known methods, there appears the possibility of scanning one or more intercalants over the sample to form complex nanostructures. Unlike the conventional intercalation methods, the kinetic barriers to effective intercalation are usually overcome under laser irradiation. Therefore, the laser approach expands the range of elements that can be introduced into a given host material.

In order to experimentally substantiate our method, we used indium and gallium selenides as host materials having layered crystal structures. From Bridgman-grown ingots, we cut samples in the form of $8 \times 4 \times 0.5$ mm rectangular parallelepipeds, two faces of which were perpendicular to the crystallographic C axis and the remaining four were parallel to it. On one of the four faces, a 1- to 3-μm-thick metal (Cu, Ga) film was deposited by thermal evaporation in a vacuum. Next, the samples were irradiated by the pulses of a free-running Nd laser. The shot duration was 2 ms, the radiation wavelength $\lambda = 1.06$ μm, and the energy density of the radiation was varied from 2 to 20 J/cm^2. The laser beam was directed onto the sample both from the film side and from the opposite side at an angle of $90° \pm 10°$. The irradiation from the sample side was performed only for gallium monoselenide, which is transparent to this radiation. Mass spectrometric analysis shows that the amount of both the Cu and Ga intercalates increases with increasing number of intercalating shots irrespective of the side of irradiation. The results of laser intercalation are summarized in the table. It is interesting that, under the same process conditions, the amount of

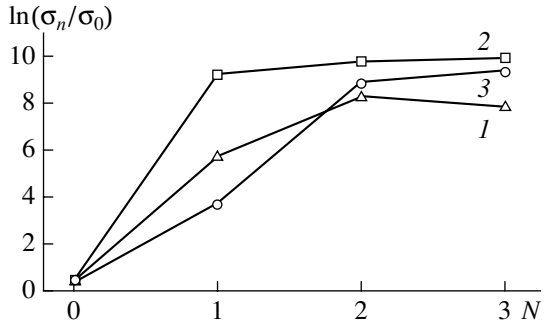


Fig. 1. Conductivity perpendicularly to the layers for (*1*) InSe and GaSe irradiated (*2*) on the side of the metallized surface and (*3*) on the opposite side vs. number of shots.

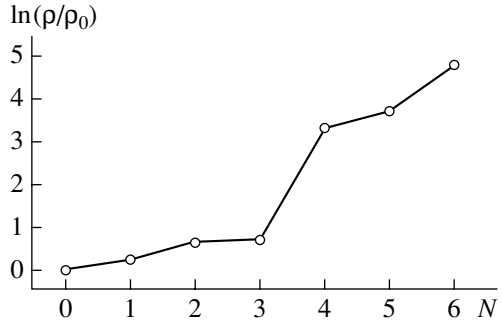


Fig. 2. Resistivity perpendicularly to the GaSe layers vs. number of shots (Ga intercalation).

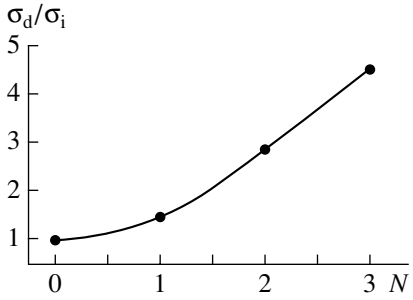


Fig. 3. Ratio of the dark conductivity perpendicularly to the layers to the conductivity observed under illumination (48 h ahead) as a function of the number of intercalating shots.

the Ga intercalate is noticeably less than that of the Cu one. This may be related to the higher diffusion coefficient and the higher penetrability of Cu as compared to those of Ga in the matrices used.

X-ray diffraction analysis of the laser-intercalated structures was performed with a double-crystal diffractometer (CuK_α radiation) by the Bond method. After the third intercalation by Cu, the layer spacing in InSe increased by 0.0026 Å, while that in GaSe irradiated from the side opposite to the metallized surface decreased by 0.0069 Å. The irradiation of reference (uncovered) GaSe samples under the same conditions increases the layer spacing by 0.0060 ± 0.0002 Å. Similar "contracting" effects have been also observed when CoO_2 was intercalated by Li [3] and $TiTe_2$ by Ag [4].

In the case of the laser intercalation by Cu, the conductivity of both compounds along the crystallographic *C* axis ($\sigma_{\|C}$) varies as shown in Fig. 1. After the third intercalcalating shot, the value of $\sigma_{\|C}$ for InSe increases by more than three orders of magnitude and for GaSe by nearly 2×10^4 times. These conductivity values exceed (by ≤20%) the corresponding values for the InSe and GaSe reference samples irradiated under the same conditions. The difference in the kinetics of Cu intercalation when the samples are irradiated on the side of the metallized surface and on the opposite side is also noticeable. The value of the resistivity perpendicular to the layers ($\rho_\perp$) is greatly affected by the concentration of ionized intrinsic or extrinsic (for example, introduced by doping with Dy) acceptors. For example, if the conductivity of the initial crystals is decreased 70 times and 3×10^2 times, the ratio of $\rho_{\perp 0}/\rho_{\perp n}$ (here, 0 and *n* are subscripts numbering the shots) drops to 220–500 and 10–25, respectively. In the latter case, the functions $\rho_\perp(x)$ become nonmonotonic. Unlike copper, the introduction of gallium, which is physically similar to the host metal (a useful expedient for the eloboration of a convenient model approximating the phenomenon of intercalation), causes the conductivity of GaSe to drop with increasing guest gallium content (Fig. 2). It should be noted that the changes in ρ/ρ_0 between intercalating shots 2 and 3, as well as between shots 4 and 5, correspond to the case when the laser beam falls on the gallium film, while those between shots 1 and 2, 3 and 4, and 5 and 6 are observed when the beam strikes the opposite side. The introduction of Ga causes the ac resistivity in the direction perpendicular to the layers to increase throughout the frequency range investigated (from 0.1 Hz to 20 MHz), the increase being smaller in the high-frequency part.

The study of relaxation processes incovered exotic effects, namely, variation of the dark resistivity in the direction perpendicular to the GaSe and InSe layers with the amount of copper intercalate (Fig. 3). These effects are related to intercalation-induced negative residual photoconductivity and its temperature quenching [5–7]. In addition, we found narrow guest concentration intervals wherein optically enhanced growth of the intergrain (interface) barriers is observed. This may serve as the initial step toward the production of charged-coupled nanostructures and photovaricaps.

As follows from Fig. 4, the initially (before the Ga intercalation) symmetric *I–V* characteristics of the samples become of the diode type during the Ga intercalation. This effect is much less noticeable in the case of the laser intercalation by Cu. In addition, in ⟨Cu⟩GaSe, we found N-shaped sections, which appear even after the first intercalating shot in the range of field strengths between 20 and 55 V/cm. These sections are absent in

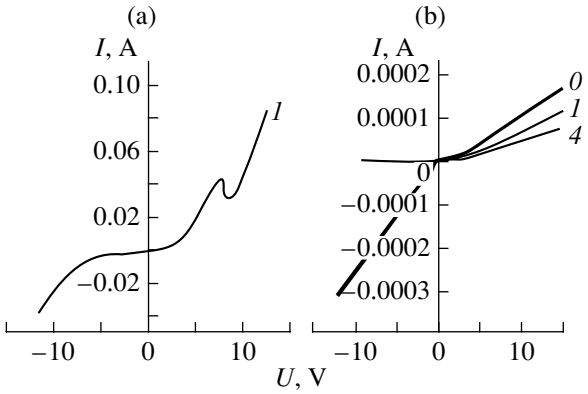


Fig. 4. *I–V* characteristics of the GaSe structures intercalated by (a) Cu and (b) Ga (figures at the curves are the numbers of the shots).

⟨Cu⟩InSe (for these samples, the symmetry of the characteristics remains unchanged) and ⟨Ga⟩GaSe. Note that the number of the N-shaped sections, as well as the values of the excess currents, are increasing functions of the guest concentration and the field strength responsible for the appearance of these sections shifts toward higher values when the number of the intercalating shots grows. It is therefore likely that the shape of the *I–V* characteristics is associated with the intercalation-induced formation of interlayer barriers with matrix compensation in the case of Ga⟨GaSe⟩. For Cu⟨GaSe⟩, the concentration of shallow impurities increases, as a result of which the thickness of these barriers becomes comparable to the de Broglie wavelength; i.e., they become tunnel-transparent. On the other hand, the N-shaped *I–V* characteristics may indicate the validity of the model proposed in [8]. If so, the intercalation of gallium monoselenide by copper may change the layer interaction integral $\varepsilon_\perp$ and the position of the Fermi level to the point where the applied voltage per layer becomes greater than $\varepsilon_\perp$.

Energy conditions of laser intercalation

Energy density of radiation, J/cm^2	Result of intercalation
2 ± 0.2	Unconvincing signs of intercalation
5 ± 0.5	Noticeable intercalation with nonuniform distribution over guest sites
9.5 ± 0.5	Uniform distribution of intercalate
15 ± 0.5	Thermal load becomes noticeable
20 ± 0.5	Thermal load increases the amount of guest metal changes insignificantly

We by no means think that our interpretation of the experimental results is exhaustive. This work will hopefully stimulate further advances in this field of nanoelectronics.

ACKNOWLEDGMENTS

The authors thank Prof. Z.D. Kovalyuk for the samples used in the experiments, as well as S.V. Gavrilyuk and V.V. Netyaga for the assistance in performing the X-ray and electrical measurements.

REFERENCES

1. J. M. Thomas, in *Physicochemical Aspects of Intercalation Phenomena* (Dordrecht, 1984), NATO ASI, pp. 521–543.
2. V. M. Kochkin and Y. N. Dmitriev, Chem. Rev. **19**, 1 (1994).
3. A. Honders, J. M. Kinderen, A. H. Heeren, *et al.*, Solid State Ionics **4**, 205 (1984).
4. A. N. Titov, Neorg. Mater. **33**, 534 (1997).
5. A. Sh. Abdinov, A. G. Kyazym-zade, and A. A. Akhmedov, Fiz. Tekh. Poluprovodn. (Leningrad) **12**, 1759 (1978) [Sov. Phys. Semicond. **12**, 1041 (1978)].
6. A. Sh. Abdinov and A. G. Kyazym-zade, Fiz. Tekh. Poluprovodn. (Leningrad) **9**, 1970 (1975) [Sov. Phys. Semicond. **9**, 1290 (1975)].
7. A. Sh. Abdinov and A. G. Kyazym-zade, Fiz. Tekh. Poluprovodn. (Leningrad) **10**, 81 (1976) [Sov. Phys. Semicond. **10**, 47 (1976)].
8. S. M. Artemenko and A. F. Volkov, Fiz. Tverd. Tela (Leningrad) **23**, 2153 (1981) [Sov. Phys. Solid State **23**, 1257 (1981)].

Translated by Yu. Vishnyakov

Technical Physics, Vol. 47, No. 6, 2002, pp. 698–702. Translated from Zhurnal Tekhnicheskoĭ Fiziki, Vol. 72, No. 5, 2002, pp. 44–49.

SOLID-STATE ELECTRONICS

Effect of Gamma Irradiation on Photoconversion Characteristics of Metal/Gallium Arsenide Barrier Structures with Textured Interface

N. L. Dmitruk, O. Yu. Borkovskaya, R. V. Konakova, I. B. Mamontova, S. V. Mamykin, and D. I. Voitsikhovskiy

Institute of Semiconductor Physics, National Academy of Sciences of Ukraine, Kiev, 03028 Ukraine

e-mail: nicola@dep39.semicond.kiev.ua

Received October 16, 2001

Abstract—The effect of ^{60}Co gamma irradiation at doses of 10^3–2×10^5 Gy on the photoconversion and dark *I–V* characteristics of Au/GaAs surface-barrier solar cells (SCs) is studied. The morphology of the interface microrelief is varied to reach the highest photoconversion efficiency. Of the two types of microrelief morphology (dendritic and quasi-grating) obtained by the chemical anisotropic etching of *n*-(100)GaAs, the latter is more promising, particularly for SCs designed for space application, since the associated SCs offer higher efficiency and radiation resistance. © *2002 MAIK "Nauka/Interperiodica".*

INTRODUCTION

GaAs-based Schottky-barrier solar cells are promising for space application because they may offer a high efficiency-to-weight ratio and have a higher resistance to hard radiation (fast protons and neutrons) than Si cells [1–4]. Moreover, in a certain dose range, irradiation by fast electrons and ^{60}Co gamma quanta causes radiation-induced ordering or gettering of active recombination centers (the so-called low-dose effect). This effect shows up as an increase in the diffusion length of charge carriers and a decrease in the recombination rate at the interface [5, 6]. The interface here serves the function of a sink for intrinsic and radiation-induced defects. Therefore, both the amount of the low-dose effect and the dose range of its existence depend on the degree of structure perfection at the interface and on the presence of mechanical stress fields in the interface region [7]. One of the most promising ways of increasing the efficiency of the cells is the use of a textured active interface. In deciding on a particular microrelief morphology and a technique for its formation, one usually takes into consideration the effect of the microrelief on the optical and recombination parameters of the structure, i.e., on the photoconversion efficiency. However, for space application, the effect of the microrelief morphology on the radiation resistance of the structures should also be taken into account. In this work, we try to establish such a correlation.

We studied the effect of ^{60}Co γ radiation on the photoconversion efficiency of metal Au/GaAs barrier structures with various morphologies of the interface microrelief obtained by chemical anisotropic etching of the GaAs surface.

EXPERIMENTAL

Au/GaAs barrier structures were fabricated with the planar technology by thermally evaporating a semitransparent gold film onto *n*-GaAs wafers (N_d = 1–3 × 10^{16} cm^{-3}). The substrate temperature was kept at about 120°C. Ohmic contacts were formed on the back side of the wafer by alloying indium. Two types of the GaAs(100) surface microrelief promising for use in photodetectors [8] and solar cells [9] were investigated. The dendritic microrelief was obtained by etching in concentrated HNO_3 acid. The quasi-grating one, a set of *V*-shaped grooves oriented along the [110] direction and having a period varying over the surface within certain limits, was produced by etching in a 2HF + $2H_2SO_4$ + $1H_2O$ mixture. The development of the microrelief was controlled by varying the etching conditions (duration and temperature). The surface of flat samples was treated in a polishing etchant. Prior to the gold deposition, the oxide layers forming during the anisotropic etching were etched off in an HCl solution; i.e., the surface finish of the wafers was the same. Gold was deposited simultaneously on all GaAs wafers and on a satellite quartz wafer used to measure the thickness and optical parameters of the gold film by multiple-angle incident ellipsometry, as well as by reflection and transmission spectroscopy. The surface morphology was studied by a Nanoscope "Dimension 3000" scanning atomic force microscope (AFM) with the silicon nitride tip operating in the tapping mode. The parameters of the semiconductor and the interface were determined from dark *I–V* characteristics and high-frequency (1 MHz) *C–V* characteristics.

We also recorded the spectral dependences of the short-circuit current and *I–V* characteristics under illu-

1063-7842/02/4706-0698$22.00

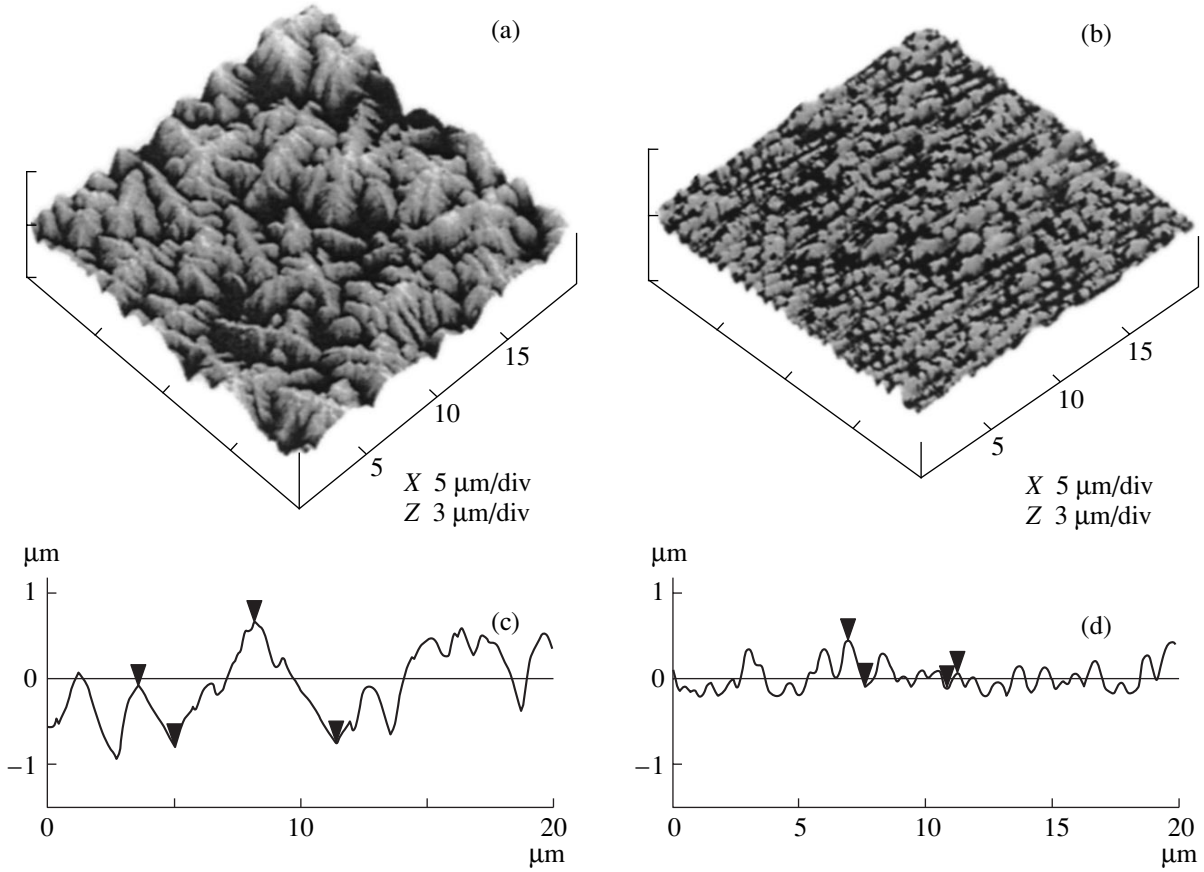


Fig. 1. Microrelief of GaAs surface: (a) dendritic and (b) quasi-grating. (c, d) Corresponding profiles (for (d), the section passes perpendicular to the grooves).

mination using a solar simulator based on an incandescent lamp with a tungsten filament. The simulator was calibrated with a reference silicon solar cell. The ^{60}Co gamma irradiation was performed with an MPX-γ-25M setup in the 10^3–2×10^5 Gy dose range.

RESULTS AND DISCUSSION

Figure 1 shows the AFM images of the two GaAs surface morphologies and the corresponding microrelief profiles. From this figure, one can determine both the statistical parameters of the microrelief and its geometrical features (in particular, the inclination of the facets to the sample plane, an increase in the surface area, etc.), which should be taken into account in analyzing the spectrum of the radiation transmitted into the semiconductor. From this spectrum, the interface recombination parameters are determined [9].

As the quasi-grating microrelief develops, both the depth (h_{max}) and the groove spacing (a) increase. For example, h_{max} = 0.3 and 1.2 μm and $a \approx 0.6$ and 1.4–2.0 μm for two different times of anisotropic etching, respectively. The inclination of the facets to the sample plane also increases from 32°–42° to 45°–48°. At the same time, the dendritic microrelief evolves from pyramidal–prismatic to purely dendritic [10] and is more complex. In particular, microplanes set at an angle of 40°–45° and at a higher angle of about 65° are observed. Because of this, the field in the areas with a higher curvature is enhanced more strongly and the thermal–field mechanism makes a greater contribution to current transport in the barrier structures [11]. This follows from the forward branches of the *I–V* characteristics presented in Fig. 2 for the structures with a different interface microrelief. Hence, along with the thermionic and recombination mechanisms of carrier transport [12], the thermal–field component should also be taken into account. The effect of this component shows up as an apparent decrease in the barrier height in certain surface regions. The inclusion of the barrier height distribution over the diode surface [13], which is broader in the case of the dendritic morphology, enables us to explain both the current rise in the initial region of the *I–V* characteristic and the smaller ideality factor n (in the range of high biases) for the structures with the developed microrelief. However, the latter can also be related to a thinner transition layer at the interface and/or to a higher density of interface states, which are filled through the exchange with the metal.

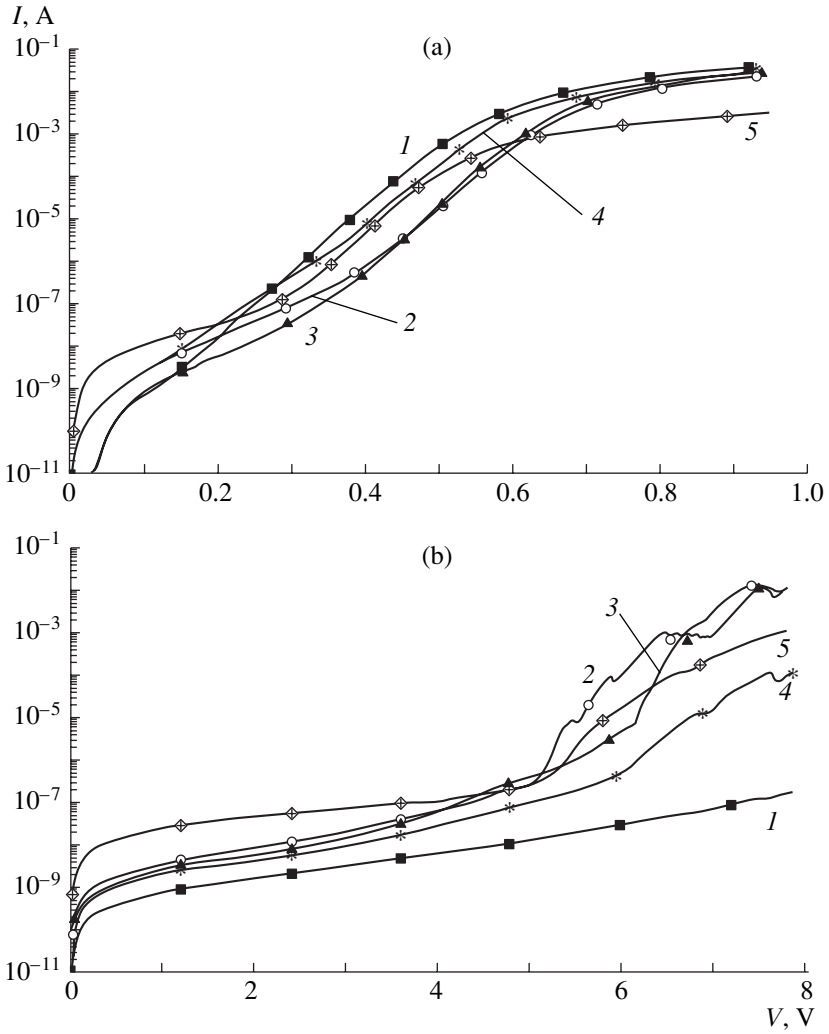


Fig. 2. (a) Forward and (b) reverse dark *I–V* characteristics of Au/GaAs diodes with (*1*) flat and (*2–5*) textured interface. (*2*, *3*) Quasi-grating microrelief and (*4*, *5*) dendritic microrelief (*3* and *5* correspond to the developed microrelief).

Figure 3a shows the *I–V* characteristics under illumination, and Fig. 3b demonstrates the spectra of the short-circuit photocurrent in the surface-barrier structures. As is seen, the short-circuit current I_{sc} grows in comparison with that for the flat structure as the microrelief develops (approximately three times and 3.2 times for the quasi-grating and developed dendritic microreliefs, respectively). However, the open-circuit voltage V_{oc} increases to a lesser extent: by a factor of ≈1.3 for the quasi-grating microrelief and by factors of ≈1.25 and ≈1.16 for the incipient and developed dendritic microrelief, respectively. The structures with the dendritic microrelief exhibit a lower fill factor *FF* of the *I–V* characteristics under illumination than the flat structures and those with the quasi-grating microrelief. Accordingly, the quasi-grating structures have the highest efficiency.

It turned out that the effect of gamma irradiation is more pronounced in the structures with the developed microrelief. The low-dose effect (which decreases the recombination rate at the interface and the ideality factor *n*, as well as increases the minority carrier diffusion length L_p and contact barrier height φ_b) is observed in the structures with the dendritic relief and developed quasi-grating microrelief at doses lower than 10^4 Gy (Fig. 4). Among the parameters of the *I–V* characteristics under illumination, the short-circuit current I_{sc} is the most sensitive to the irradiation, as in the case of Si- and GaAs-based flat structures irradiated by fast electrons and protons [1–4]. Figure 5 depicts the spectral dependences of the relative photocurrent when the diodes with the developed interface microrelief were irradiated by improving (≈10^3 Gy) and degrading (≈2×10^5 Gy) doses. It is evident that the basic effect of the gamma irradiation, which is virtually wavelength-independent, is on the interface recombination rate [14]. The *I–V* characteristics under illumination start to degrade from the irradiation dose 10^5 Gy. The struc-

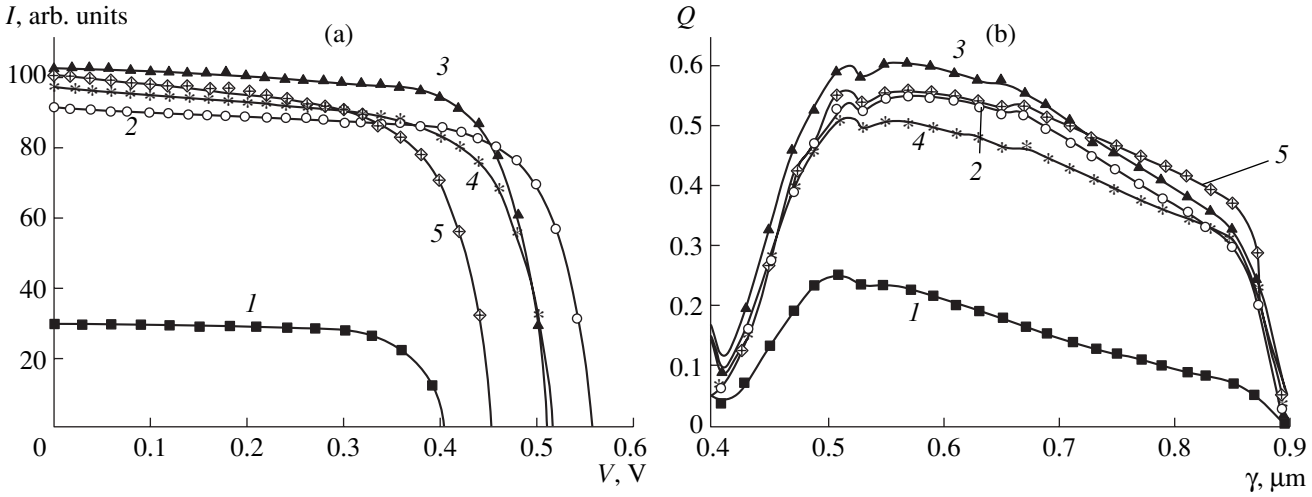
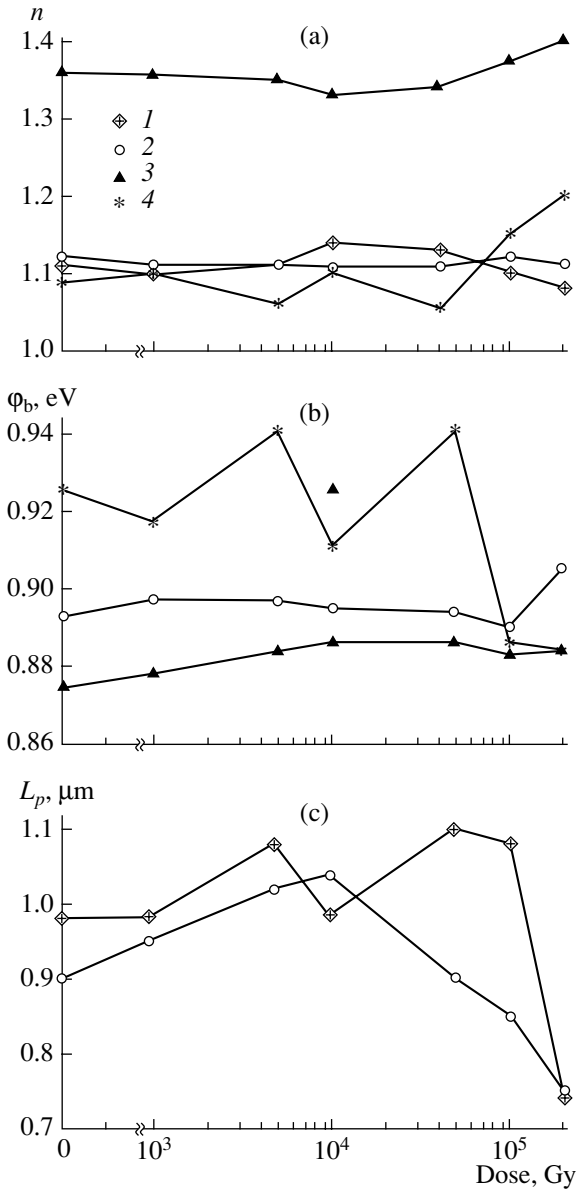


Fig. 3. (a) *I–V* characteristics under illumination and (b) spectral dependences of the short-circuit photocurrent (external quantum efficiency) of Au/GaAs photoconverters with (*1*) flat and (*2–5*) textured interface. Figures mean the same as in Fig. 2.



tures with the developed dendritic microrelief are the least resistant. Here, the efficiency drops by about 5% from the initial value at a dose of 2×10^5 Gy. The quasi-grating microrelief imparts to the structures the highest resistance: the efficiency drops by no more than 2% from the initial value.

Thus, the dendritic microrelief of the metal/GaAs interface with the facets tilted at angles ≥60° provides higher values of the short-circuit photocurrent. In spite of this, it is less suitable for SCs designed for space application than the quasi-grating microrelief (at least

Fig. 4. Dose dependences of (a) ideality factor *n*, (b) barrier height φ_b, and (c) minority carrier diffusion length for diodes with the textured interface. (*1*, *2*) Quasi-grating microrelief and (*3*, *4*) dendritic microrelief (*2* and *4* correspond to the developed microrelief).

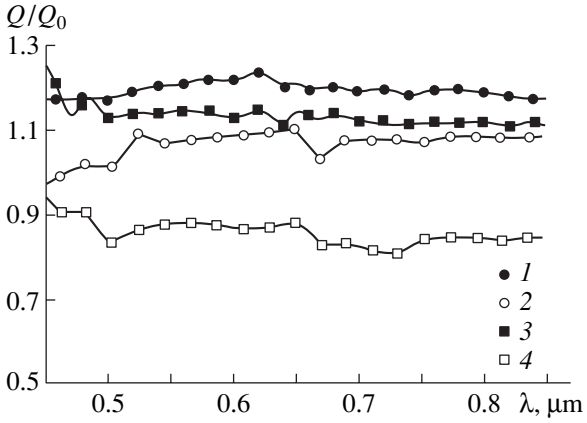


Fig. 5. Spectral dependences of the ratio between the photocurrents in the irradiated and nonirradiated structures with (*1*, *2*) the developed quasi-grating and (*3*, *4*) dendritic interface microreliefs. Dose equals (*1*, *3*) 10^3 and (*2*, *4*) 2×10^5 Gy.

without special passivating treatments [14, 15]) in terms of efficiency and radiation resistance.

REFERENCES

1. J. Wysocki, J. Appl. Phys. **34**, 2915 (1963).
2. H. Gebert and J. Edmund, J. Electrochem. Soc. **125**, 678 (1978).
3. G. M. Grigor'eva, V. A. Grilikhes, K. N. Zvyagina, *et al.*, Geliotekhnika, No. 1, 8 (1989).
4. M. Ya. Bakirov, Geliotekhnika, No. 6, 32 (1991).
5. O. Yu. Borkovskaya, N. L. Dmitruk, R. V. Konakova, *et al.*, Radiat. Eff. **42**, 249 (1979).
6. O. Yu. Borkovskaya, N. L. Dmitruk, R. V. Konakova, *et al.*, Fiz. Tekh. Poluprovodn. (Leningrad) **20**, 1640 (1986) [Sov. Phys. Semicond. **20**, 1028 (1986)].
7. O. Yu. Borkovskaya, N. L. Dmitruk, V. G. Litovchenko, and O. N. Mishchuk, Fiz. Tekh. Poluprovodn. (Leningrad) **23**, 207 (1989) [Sov. Phys. Semicond. **23**, 129 (1989)].
8. N. L. Dmitruk, O. Yu. Borkovskaya, O. I. Mayeva, and I. B. Mamontova, Sens. Actuators A **75**, 151 (1999).
9. N. L. Dmitruk, O. Yu. Borkovskaya, I. B. Mamontova, and S. V. Mamykin, Sol. Energy Mater. Sol. Cells **60**, 379 (2000).
10. T. Ya. Gorbach, E. V. Pidlisnyĭ, and S. V. Svechnikov, Optoélektron. Poluprovodn. Tekh. **13**, 34 (1988).
11. O. Yu. Borkovskaya, T. Ya. Gorbach, N. L. Dmitruk, and O. N. Mishchuk, Élektron. Tekh., Ser. 2, No. 5, 50 (1989).
12. O. Yu. Borkovskaya, N. L. Dmitruk, D. I. Voĭtsikhovskiy, *et al.*, Poverkhnost, No. 8, 87 (1999).
13. N. L. Dmitruk, O. Yu. Borkovskaya, I. N. Dmitruk, *et al.*, in *Proceedings of ICFSI-8, Sapporo, 2001*.
14. N. L. Dmitruk, O. Yu. Borkovskaya, and I. B. Mamontova, Geliotekhnika, No. 5, 3 (1998).
15. N. L. Dmitruk, O. Yu. Borkovskaya, and I. B. Mamontova, Zh. Tekh. Fiz. **69** (6), 132 (1999) [Tech. Phys. **44**, 726 (1999)].

Translated by M. Lebedev

Technical Physics, Vol. 47, No. 6, 2002, pp. 703–707. Translated from Zhurnal Tekhnicheskoĭ Fiziki, Vol. 72, No. 5, 2002, pp. 50–55.

OPTICS, QUANTUM ELECTRONICS

New Mechanisms of Optical Writing/Reading in Magnetic Media

A. V. Nikolaev, E. P. Nikolaeva, V. N. Onishchuk, and A. S. Logginov

Moscow State University, Vorob'evy gory, Moscow, 119899 Russia

e-mail: asl@osc.phys.ru

Received June 18, 2001; in final form, December 4, 2001

Abstract—Optical detection of magnetic features of 0.1 μm or less in size, as well as their generation and motion due to laser shots, is described. The physical effects discovered make it possible to implement the basic functions of a memory device (writing, shift in a storage register, and reading) by purely optical means. These effects can form a basis for designing novel ultra-high-density solid-state memories with optical access and control. © *2002 MAIK "Nauka/Interperiodica".*

INTRODUCTION

In digital memories with optical writing/reading, the bit size is defined largely by the effective area of interaction between the optical radiation and a record medium, i.e., by the diameter of a focused laser beam [1–3]. This is equally true for both commercial memory devices (compact disks, videodisks, magnetooptical disks, etc.) and new-generation memories under development (3D optical memory based on photorefractive and photochromic materials). However, because of the diffraction-related constraints imposed on the degree of focusing, a bit size of much less than the recording laser wavelength is extremely difficult to obtain even with special optical tools and various buffer structures that limit the size of the bit being recorded [1, 3]. In addition, such approaches require sophisticated facilities and are very expensive. Moreover, a decrease in the bit size inevitably complicates data readout schemes. In conventional methods of optical reading (effect of induced luminescence, phase-change effect, and change in the plane of polarization and/or in the intensity of transmitted or reflected light according to whether "0" or "1" is recorded), a decrease in the bit size causes a decrease in the valid signal intensity against the optical background and noise. Various special methods are used to solve this problem, such as land–groove recording, partial response signaling and maximum likelihood detection (PRML), magnetically induced superresolution (MSR), etc. [1–7]. It is believed that, when applied to optical memory with a magnetic active (recording) medium (magnetooptical disks), these writing/reading methods will shrink the bit (magnetic domain) size down to 0.1–0.2 μm [3–6].

The restrictions and difficulties associated with the design of new-generation memory devices necessitate the search for novel materials and physical mechanisms that would admit a change in the material parameters over a submicron area due to optical action and subsequent detection of this change by optical means.

In this work, we report experiments on direct optical detection of magnetic features 0.1 μm or less in size, as well as study the laser generation and control of these features. The physical effects employed can form a basis for designing ultra-high-density solid-state magnetic memories with optical reading/writing.

EXPERIMENTAL TECHNIQUES

Optical detection of magnetic features with a size below the optical resolution of conventional magnetooptical microscopy based on the Faraday or Kerr effects can be provided with dark-field microscopy. This method is intended for visualizing the phase structure of objects that cannot be distinguished from the environment optically. Here, the illuminating beam does not fall into the objective lens and the image is formed by the light scattered by phase inhomogeneities of the sample. Due to magnetooptical effects, the magnetization distribution in the sample is merely the phase structure that can be visualized in dark field [8]. Because of the diffraction principle of imaging, dark-field microscopy (widely used in various investigations but poorly developed as applied to magnetic phenomena) can basically detect structures of size beyond the optical resolution of conventional microscopy (ultramicroscopy). The use of the dark-field geometry in optical microscopy allows the detection of mesoscopic magnetic features 0.1 μm or less in size. Of special importance is the fact that dark-field microscopy does not require scanning and is well compatible with standard methods for studying the fast dynamics of magnetic structures (e.g., streak photography) [9], whereas scanning tunnel microscopy has fundamental limitations when applied to dynamics investigations.

1063-7842/02/4706-0703$22.00

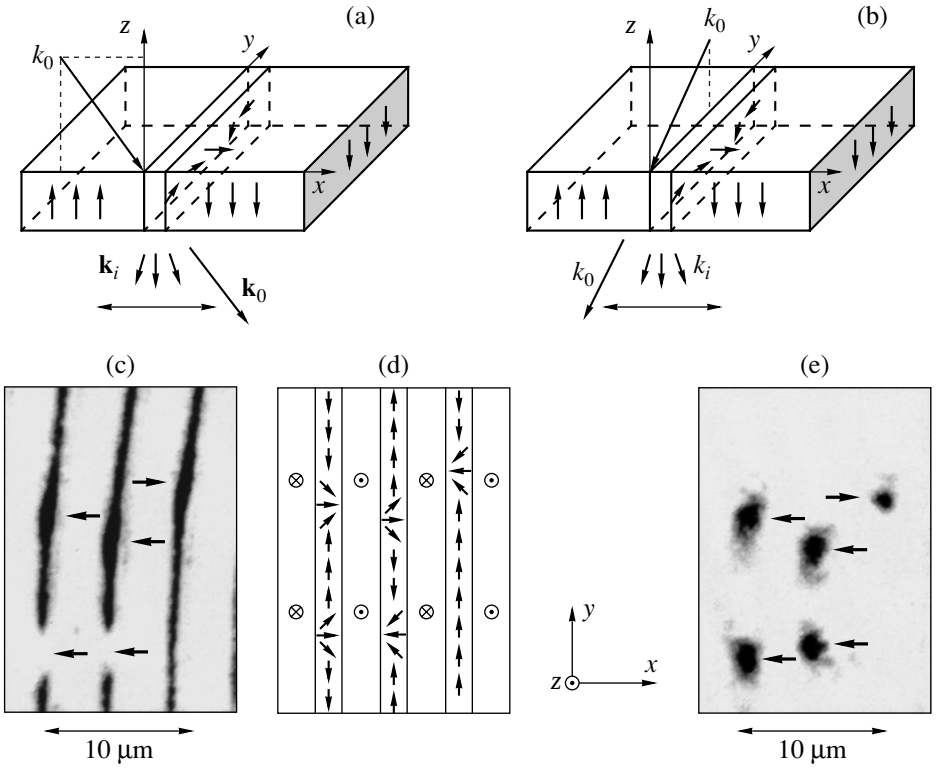


Fig. 1. (a, b) Optical scheme for one-sided dark-field illumination and magnetization distribution in domains, domain walls, and Bloch lines. The plane of light incidence is (a) perpendicular and (b) parallel to the DW plane. $\mathbf{k}_0$, wavevector of incident radiation; $\mathbf{k}_i$, wavevector of scattered radiation. (c, e) Typical images of DWs and Bloch lines when the light is incident (c) normal and (e) parallel to the DW plane (opposite contrast). (d) Schematic magnetization distribution in the sample.

The detection of submicron magnetic features with dark-field microscopy was first demonstrated by Thiaville and Miltat, who succeeded in direct optical observation of vertical Bloch lines (VBLs) in garnet ferrite films with perpendicular anisotropy [10]. VBLs are stable elements of a domain wall (DW) structure and are 0.1 μm or less in size [11]. In the polarized anisotropic dark-field observation (PADO) method [10], the illuminating beam strikes the sample at some angle and bypasses the objective lens, while the image is formed by the light diffracted by magnetic inhomogeneities (Fig. 1a). In the PADO geometry, the plane of light incidence is perpendicular to the DW plane. In this case, the DW appears as a bright fringe on the black background and local variations of the brightness are treated as VBL locations (Fig. 1c; opposite contrast, VBLs are marked by arrows). With the angle of light incidence fixed, the VBL contrast (bright or dark) depends on the VBL magnetic topology and magnetization direction in adjacent domains (the topology shown in Fig. 1d is that of the DWs with VBLs in Fig. 1c) [12].

Combining the PADO method with streak photography (which provides the observation of instantaneous dynamic configurations of the domain structure [13]) has allowed us to detect and comprehensively study the VBL controllable generation and movement by locally applying a laser radiation [14, 15]. Experiments were carried out with uniaxial magnetic garnet ferrite films of composition $(BiTm)_3(FeGa)_5O_{12}$ which had the following parameters: crystallographic orientation (111), saturation magnetization $4\pi M_s$ = 173 G, collapse field H_{col} = 126 Oe, period of the equilibrium strip domain structure 8.5 μm, thickness 7.5 μm, and quality factor Q = 3.8.

GENERATION AND MOVEMENT OF SUBMICRON MAGNETIC FEATURES DUE TO LASER IRRADIATION

From experiments [14, 15], as the energy of a laser shot (duration 10 ns, wavelength λ = 540 nm, maximum peak energy $W_{max} \sim 10^{-6}$ J, and spot diameter on the film 3–4 μm) grows, the shot causes the (1) movement of existing VBLs toward a DW, (2) generation of a pair of VBLs on DWs that were initially free of the lines, and (3) irreversible changes in the configuration of both individual DWs and the domain structure as a whole.

Figure 2 illustrates the VBL movement under the action of focused laser shots. The initial domain structure with a specially introduced VBL is shown in Fig. 2a (VBL is marked by an arrow, symbol ⊗ shows the site of the laser spot). With a shot of energy $W \approx$

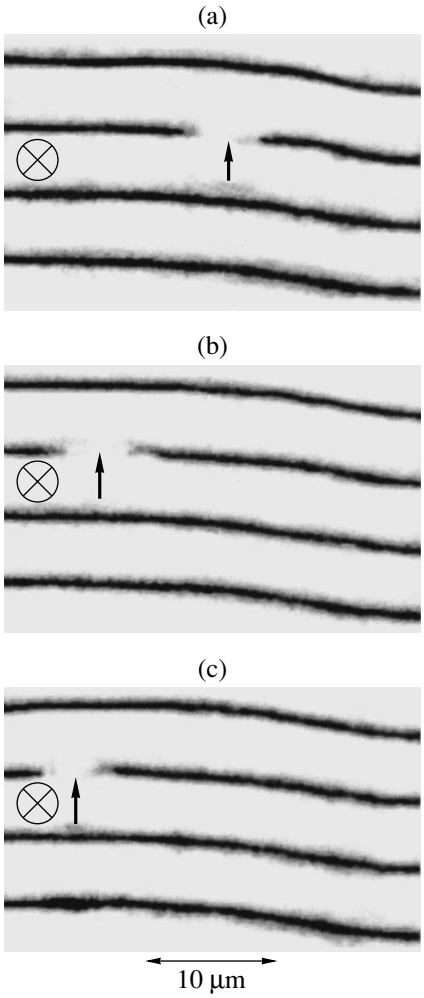


Fig. 2. VBL movement under laser irradiation: (a) initial position and (b, c) shifts after the sequential application of two shots. Dark-field image (opposite contrast).

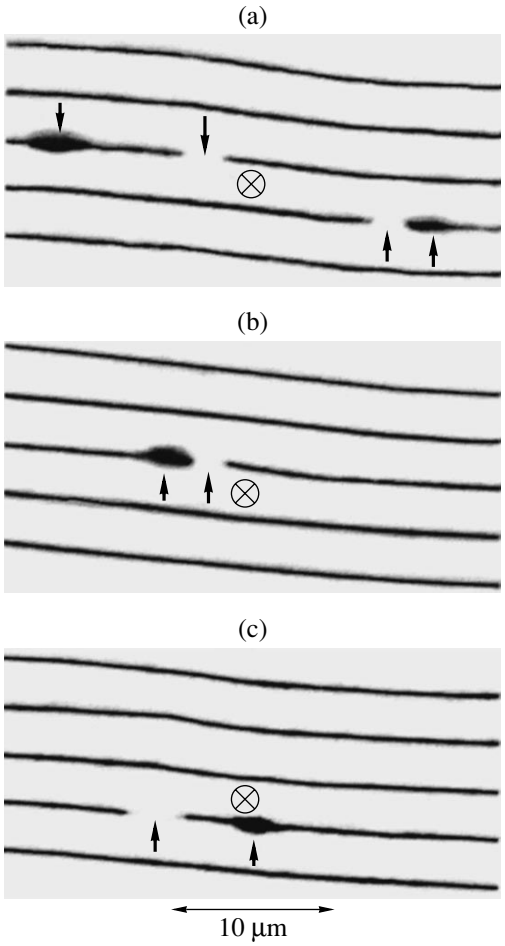


Fig. 3. Typical dark-field images (opposite contrast) of Bloch lines (marked by arrows) originating in DWs due to the laser shot.

0.1 W_{max} applied, the VBL shifts (Fig. 2b). From Figs. 2b and 2c, it is seen that the travel of the VBL decreases as it approaches the laser spot. This effect, as well as the conditions for uniform and reproducible VBL movement, calls for further investigation.

At $W \approx 0.2\ W_{max}$, the shot generates a pair of VBLs on DWs that were initially free of the lines. Figures 3a–3c show typical images of localized VBLs produced by the shot. The contrast is seen to change at the walls where VBL pairs have been generated.

Experimental conditions under which the VBL generation is controllable and reproducible are discussed at length in [15].

The series of streak photos in Fig. 4 demonstrates the movement of the DW and the generation of a pair of VBLs due to the shot. The upper photos, showing the "instantaneous" configurations of the domain structure, were obtained in the direct-illumination mode (using the Faraday effect) with crossed polarizers. A dark area is seen to appear near the laser spot, and the DW nearest to the spot bends. The rate of bending is the highest within several tens of nanoseconds after the shot application. Then, the DW relaxes to its initial shape within 1500–2000 ns and the dark area disappears (in this area, the Faraday effect is weak, since the laser heating reduces the saturation magnetization). The lower photos are the dark-field images of the same process in the same area. Here, the DW bend appears as a local change in the DW brightness. VBLs (marked by arrows) near the laser spot become distinct 300 ns after the shot has been applied and reach their final positions within the same 1500–2000 ns noted above.

Based on the experimental data obtained in [14, 15], we put forward a physical model of VBL generation and movement due to a laser shot. Central in this model is the thermal effect of the shot, which locally reduces the saturation magnetization of the film in the laser spot region. A local change in the magnetization upon heating and subsequent cooling modifies radically the distribution of stray fields in this region and causes nearby DWs to shift. When the velocity of a DM exceeds the

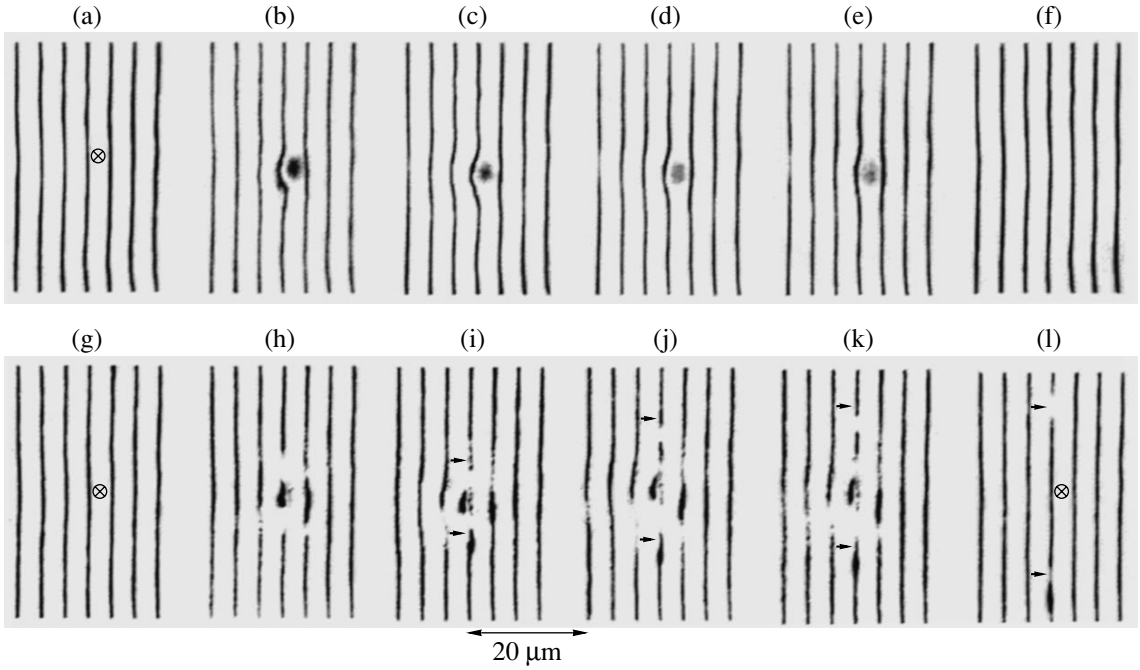


Fig. 4. Instantaneous dynamic configurations of the domain structure at various time instants after the application of the "writing" laser shot. (e, g) Before the pulse application; (b, h) 100, (c, i) 300, (d, j) 700, and (e, k) 1000 ns after the pulse; and (f, l) after the dynamic processes have been completed. (a)–(f) Bright-field images with crossed polarizers; (g)–(l) dark-field images (opposite contrast).

critical value, a horizontal VBL originates in it. This horizontal line emerges on the film surface, producing a pair of VBLs. If the velocity of the DW is below the critical value, the Bloch lines may advance.

The above effect of Bloch line reproducible generation allows us to formulate a concept of optical writing of a submicron magnetic bit of information. For the optical generation of a magnetic feature of size ≈0.1 μm, there is no need to shrink the focused laser beam as much as possible or to produce an artificial potential relief to specify the size of a laser-written domain (which is typical of attempts to raise the thermomagnetic writing density on magnetooptical disks). The experimental results described above suggest that a "magnetic" bit can optically be written as a feature of the DW structure rather than as a domain. In this case, there is no need for a correlation between the scale of a feature recorded and the size of the laser spot. Here, the "size" of a bit, a mesoscopic object inside a DW, is specified by its physical nature and not by the characteristic scale of the "writer."

OPTICAL READOUT OF SUBMICRON MAGNETIC STRUCTURES

In experiments on the generation and motion of VBLs, the lines can conveniently be detected with the PADO method, which allows one to visualize and study the motion not only of the VBLs but also of the domain wall. However, if a VBL is viewed as an information bit, only the signal from the VBL is valid for readout. In this case, such a feature of the PADO method as the representation of a VBL in the form of a dark or bright region (according to the magnetization distribution in the line) against the background of the DW complicates the readout scheme and requires three-level (bright VBL, dark VBL, domain wall) discretization of the optical signal detected. Moreover, the study of the mechanism behind VBL and DW dark-field imaging [12] has shown that the change in the contrast in the VBL vicinity (see Fig. 1) depends on the inclination of the DW with VBLs rather than on the magnetization distribution. The inclination can be related to different physical mechanisms [12]. The indirect character of imaging in the PADO mechanism (in the geometry of Fig. 1a) greatly complicates the interpretation of diffraction images observed experimentally in terms of micromagnetism, does not allow exact determination of the VBL topology, and makes it possible to visualize VBLs only for a limited class of samples [12].

The aforesaid explains the extensive search for direct magnetooptical methods that would be sensitive to the magnetization distribution in Bloch lines and provide their simple yet reliable detection in various samples.

Our investigations have shown that VBLs can optically be detected with other dark-field geometries [16, 17]. In Fig. 1b, as in the conventional PADO method, only scattered light falls into the objective lens because of the inclined incidence of the beam. In this

case, however, the plane of light incidence is parallel to the DWs. Neither domains nor DWs without VBLs take part in imaging, and the DW images become almost free of background noise.

Figures 1c and 1e are the dark-field images of a fixed region that were obtained when the plane of light incidence was normal and parallel to the DW plane (conventional and modified PADO geometry, respectively; VBLs are indicated by arrows). In the latter case, magnetooptical diffraction by the DWs is absent and the DWs exhibit no contrast (are not visualized). However, with the angle of incidence chosen properly, the Bloch lines (bright spots) are seen (Fig. 1e). Figure 1e differs from an image obtained with the conventional PADO method in that (1) Bloch lines are visualized as bright spots irrespective of the magnetic topology and (2) their sizes along and across the DW are nearly the same.

Further experimentation is needed to gain better insight into the physics of VBL visualization. Yet we can note that the Bloch line imaging is consistent with predictions of the magnetooptical diffraction model developed in [10]. This model considers diffraction by magnetization distribution in a VBL and treats a DW as a one-dimensional phase edge (for polarized light) and the very VBLs as point diffraction centers. Our results are in qualitative agreement with this model; namely, (1) diffraction by DWs is absent, (2) the contrast for Bloch lines of different topology is the same, and (3) the ratios of the VBL sizes along and across a DW are similar.

From the above, we can hypothesize that, in the PADO geometry considered, direct magnetooptical diffraction by magnetization distribution in a Bloch line, rather than the distortion of the DW vertical profile, is responsible for VBL visualization.

CONCLUSION

The experiments described above illustrate the physical mechanisms providing optical detection of magnetic features 0.1 μm or less in size, as well as those responsible for their generation and advance. The physical effects discovered make it possible to implement the basic functions of a memory device (writing, shift in a storage register, and reading) by purely optical means. These effects can form a basis for designing novel ultra-high-density solid-state VBL-based memories with optical access and control.

Owing to such properties as nonvolatility and high record density (up to 1 and even 128 Gbit/cm^3 [18, 19]), VBL memories have extensively been studied since the mid-1980s. However, they have not been implemented because associated magnetically controlled read/write devices, as well as devices advancing VBLs, are not feasible. The effects of VBL optical generation and detection that were discovered in this work may help to implement ultra-high-density magnetic memories with optical access that would outperform the magnetooptical memories currently available in many parameters.

ACKNOWLEDGMENTS

This work was supported by the Program "Russian Universities: Basic Research" and the Russian Foundation for Basic Research (grant no. 96-15-96620).

REFERENCES

1. M. Mansuripur and G. Sincerbox, Proc. IEEE **85**, 1780 (1997).
2. S. Kawata, Proc. IEEE **87**, 2009 (1999).
3. *The Future of Data Storage Technologies*, Ed. by S. C. Esener and M. H. Kryder (National Technology Research Institute, Baltimore, 1999); http://itri.loyola.edu/reports.htm.
4. M. Kaneko, K. Aratani, and M. Ohta, Jpn. J. Appl. Phys. **31**, 568 (1992).
5. M. Birukawa, N. Miyatake, and T. Suzuki, IEEE Trans. Magn. **34**, 438 (1998).
6. H. Fuji, T. Okumura, S. Maeda, *et al.*, IEEE Trans. Magn. **35**, 2047 (1999).
7. G. Mathew, B. Farhang-Boroujeny, and C. Y. Ng, IEEE Trans. Magn. **36**, 2098 (2000).
8. B. Kuhlow and M. Lambeck, Physica B & C (Amsterdam) **80**, 374 (1975).
9. A. S. Logginov, A. V. Nikolaev, and V. V. Dobrovitski, IEEE Trans. Magn. **29**, 2590 (1993).
10. A. Thiaville, F. Boileau, J. Miltat, *et al.*, J. Appl. Phys. **63**, 3153 (1988).
11. A. P. Malozemoff and J. C. Slonczewski, *Magnetic Domain Walls in Bubble Materials* (Academic, New York, 1979; Mir, Moscow, 1982).
12. A. Thiaville, J. Ben Youssef, Y. Nakatani, *et al.*, J. Appl. Phys. **69**, 6090 (1991).
13. L. P. Ivanov, A. S. Logginov, and G. A. Nepokoĭchitskiĭ, Zh. Éksp. Teor. Fiz. **84**, 1006 (1983) [Sov. Phys. JETP **57**, 583 (1983)].
14. A. S. Logginov, A. V. Nikolaev, V. N. Onishchuk, *et al.*, Pis'ma Zh. Éksp. Teor. Fiz. **66**, 398 (1997) [JETP Lett. **66**, 426 (1997)].
15. A. S. Logginov, A. V. Nikolaev, E. P. Nikolaeva, and V. N. Onishchuk, Zh. Éksp. Teor. Fiz. **117**, 571 (2000) [JETP **90**, 499 (2000)].
16. A. S. Logginov, A. V. Nikolaev, and V. N. Onishchuk, in *Proceedings of the All-Russia Scientific Conference "Physics of Condensed State," Sterlitamak, 1997*, Vol. 2, p. 50.
17. V. I. Belotelov, A. S. Logginov, and A. V. Nikolaev, Radiotekh. Élektron. (Moscow) **46**, 870 (2001).
18. S. Konishi, IEEE Trans. Magn. **19**, 1838 (1983).
19. G. Ashton, *Solid State Memory Study Mid-Term Report* (National Media Laboratory, St. Paul, 1993).

Translated by V. Isaakyan

Technical Physics, Vol. 47, No. 6, 2002, pp. 708–712. Translated from Zhurnal Tekhnicheskoĭ Fiziki, Vol. 72, No. 5, 2002, pp. 56–60.

OPTICS, QUANTUM ELECTRONICS

Multiple-Exposure High-Sensitivity Holographic Interferometry for Studying Moving Objects

A. M. Lyalikov

Kupala State University, Grodno, 230023 Belarus

Received July 23, 2001; in final form, November 29, 2001

Abstract—A method for increasing the sensitivity of measurements through aberration compensation upon reconstruction of interferograms from two multiple-exposure holograms is proposed. At the early stage of object investigation, the holographic structures recorded at certain time instants are rerecorded by two coherent beams on new image carriers. In this case, the interference moiré method is employed to monitor the equality of the vectors of the holographic structures rerecorded. At the final stage, the new nonlinear holograms are processed in an optical analyzer of conjugate holograms with the use of incoherent light. The hologram thus reconstructed offers a high sensitivity of measurements and is free of aberrations. The method is tested by visualizing the temperature-field variations in a glass substrate with a conducting coating. *© 2002 MAIK "Nauka/Interperiodica".*

INTRODUCTION

Holographic interferometry offers significant advantages over the conventional method, such as compensation of aberrations due to the optical elements of an interferometer, arbitrary adjustment of the interference fringes, and sensitivity control at the stage of hologram optical processing [1–3]. To increase the sensitivity of measurements in holographic interferometry, one can take advantage of the nonlinear properties of holograms and reconstruct higher diffraction orders [4, 5] or rerecord holograms [6, 7]. Rerecording in combination with the nonlinear recording of secondary holograms makes it possible to substantially improve the sensitivity of measurements by effectively compensating aberrations in recording and optical processing systems [8–10].

The advantages of the holographic methods become even more evident in studies of fast processes. Various states of the object can be recorded on several image carriers [11] or on a single one [12, 13]. The latter technique, where the reference beam can move, enables one to reduce experimentation costs, since several states of the object are recorded on the same carrier. Apparently, the number of records per carrier is limited. Therefore, the recording of a large number of holographic structures corresponding to various states of the object requires a number of carriers. In this case, the optical processing of several multiple-exposure holograms involves difficulties when object different states recorded on different carriers are compared. Consider, for example, a pair of multiple-exposure holograms in a two-channel interferometric correlator [14], where the filtering diaphragm discriminates waves diffracted by the holographic structures of the object states being compared. Under these conditions, the resulting interference pattern, visualizing the difference in the states, will additionally be distorted by inhomogeneities in the substrates of the hologram carriers. Moreover, the recording of several holographic structures on a single carrier complicates markedly the diffraction spectrum of the multiple-exposure hologram, which is observed in the second focal plane of the objective lens. In the case of the nonlinear recording of such a hologram, the diffraction spectrum is so complicated by crosstalk-induced extra maxima [15] that it becomes impossible to discriminate between waves diffracted by given holographic structures and spurious ones. Therefore, the nonlinear properties of multiple-exposure holograms cannot improve the measurements sensitivity.

In this work, we developed a method for optically processing multiple-exposure holograms that offers increased sensitivity, compensates aberrations, and enables one to use even incoherent light at the final stage of processing.

DESCRIPTION OF THE METHOD

Let us assume that N states of the object are holographically recorded on each of several carriers at various time instants. The resulting intensity distribution on the carrier is the sum of the intensities from each exposure recording a particular holographic structure:

$$I = I_1 + I_2 + \ldots + I_N. \quad (1)$$

If a multiple-exposure hologram is recorded under linear conditions, its amplitude transmission is directly proportional to the resulting intensity [3]. Let us record two multiple-exposure holograms with the amplitude

1063-7842/02/4706-0708$22.00

transmissions τ and τ' given by

$$\tau = \tau_0 + \sum_{l=1}^{N} \cos[2\pi(\xi_l x + \eta_l y) + \varepsilon_l + \varphi], \quad (2)$$

$$\tau' = \tau'_0 + \sum_{l=1}^{N} \cos[2\pi(\xi_l x + \eta_l y) + \varepsilon'_l + \varphi], \quad (3)$$

where τ_0 and τ'_0 are constants; $\xi_l = \cos\gamma_{lx}/T_l$ and $\eta_l = \cos\gamma_{ly}/T_l$ are the spatial frequencies of the lth holographic structure; γ_{lx} and γ_{ly} are the angles between the fringes and the axes *0X* and *0Y*, respectively; T_l is the period of the fringes in the lth holographic structure; ε_l and ε'_l are the object-related phase shifts at the instants of the lth exposures of the first and second holograms, respectively; and φ is the phase distortion caused by aberrations in the recording system (holographic interferometer). The coordinate frame *X0Y* lies in the plane of the hologram. The spatial frequencies ξ_l and η_l are directly proportional to the projection of the vector of the lth holographic structure. This vector lies in the plane of the hologram and is orthogonal to the fringes. Note that the recording of multiple-exposure holograms necessitates a change in the fringe orientation or period for each exposure to avoid overlap between the diffraction orders of the reconstructed waves during the optical processing of multiple-exposure holograms. This can be provided, for example, by varying the angle between the object and reference beams prior to each exposure [15].

It has been demonstrated [16–19] that an increase in the sensitivity of measurements for obtaining high-quality low-noise interferograms implies the use of incoherent illumination. The associated method is implemented in two stages. The first stage involves the rerecording of each of the holographic structures written in complex conjugate orders from a multiple-exposure hologram on two new carriers. The rerecording also implies verification of the strict equality of the rerecorded holographic structure vectors. At the second stage, this pair of holograms, recorded under nonlinear conditions, is processed in an optical analyzer of conjugate holograms with the use of an incoherent light source [3].

Multiple-exposure holograms are rerecorded by two coherent beams in a rerecording device (Fig. 1a) which can be based on any two-element interferometer [17]. To approximately equate the vectors of the holographic structures rerecorded, we make two holes in filter *3* of spatial frequencies, which is located in the second focal plane of objective lens *2* (Fig. 1a). The holes are centered at the points with coordinates (ξ_1, η_1) and (ξ_2, η_2) relative to the principal optical axis (Fig. 1b). Such a spatial filter transmits waves with spatial frequencies approximately equal to $\xi_1 = \cos\alpha_1/\lambda$, $\eta_1 = \cos\beta_1/\lambda$ and $\xi_2 = \cos\alpha_2/\lambda$, $\eta_2 = \cos\beta_2/\lambda$. Here, $\cos\alpha_1$, $\cos\beta_1$, $\cos\alpha_2$, and $\cos\beta_2$ are the direction cosines of the waves and α_1 and α_2 (β_1 and β_2) are the angles between the direction of propagation and the *0X* (*0Y*) axis. For the rerecording of the first holographic structure corresponding, for example, to the bth exposure, we employ two coherent waves with the spatial frequencies ξ_{01}, η_{01} and ξ_{02}, η_{02} meeting the conditions

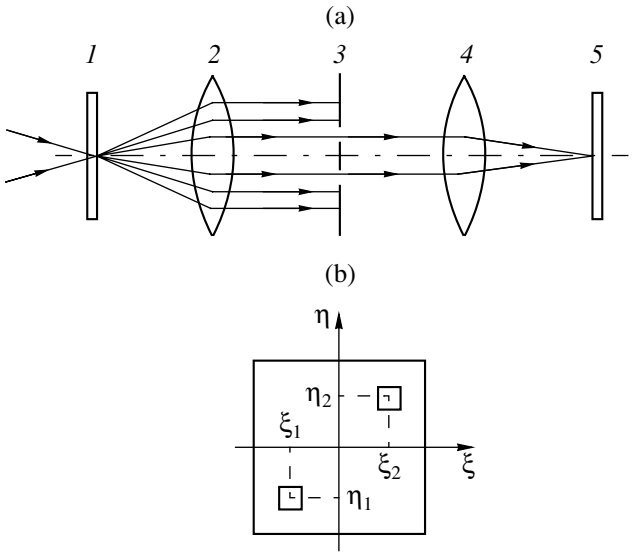


Fig. 1. (a) Optical scheme of the device for rerecording multiple-exposure holograms and (b) spatial-frequency filter.

$$\begin{aligned} \xi_{01} &= \xi_1 - \xi_b, \quad \eta_{01} = \eta_1 - \eta_b, \\ \xi_{02} &= \xi_2 + \xi_b, \quad \eta_{02} = \eta_2 + \eta_b, \end{aligned} \quad (4)$$

where $\xi_b = \cos\gamma_{bx}/T_b$, $\eta_b = \cos\gamma_{by}/T_b$, and T_b is the period of the fringes in the bth holographic structure that makes angles γ_{bx} and γ_{by} with the *0X* and *0Y* axes, respectively, in the plane of the multiple-exposure hologram. In this case, the spatial-frequency filter selects the waves

$$\begin{aligned} A_1 &= a_b \exp\{i[2\pi(\xi_1 + \delta\xi_1)x \\ &+ 2\pi(\eta_1 + \delta\eta_1)y + \varepsilon_b + \varphi + \psi_0]\}, \\ A_2 &= a_b \exp\{i[2\pi(\xi_2 + \delta\xi_2)x \\ &+ 2\pi(\eta_2 + \delta\eta_2)y - \varepsilon_b - \varphi + \psi_0]\}, \end{aligned} \quad (5)$$

where $\delta\xi_1$, $\delta\eta_1$, $\delta\xi_2$, and $\delta\eta_2$ are adjustment errors due to an insufficient accuracy in satisfying condition (4) and the function ψ_0 takes into consideration optical inhomogeneities of the substrate of the carrier of multiple-exposure hologram (2). The adjustment errors are related to the finite sizes of the holes in the spatial-frequency filter. Waves (5) form a rerecorded holographic structure in plane *5*, which is optically coupled with multiple-exposure hologram *1* by objective *4*. The

amplitude transmission of the rerecorded *b*th holographic structure is represented as

$$\tau_b = \{1 + \cos[2\pi(\xi_1 - \xi_2 + \delta\xi)x + 2\pi(\eta_1 - \eta_2 + \delta\eta)y + 2(\varepsilon_b + \varphi)]\}^{-\gamma/2}, \quad (6)$$

where $\delta\xi = \delta\xi_1 - \delta\xi_2$, $\delta\eta = \delta\eta_1 - \delta\eta_2$, and γ is the contrast ratio.

The projections of the rerecorded holographic structure vector onto the *0X* and *0Y* axes are derived from expression (6):

$$K_x = 2\pi(\xi_1 - \xi_2 + \delta\xi), \quad K_y = 2\pi(\eta_1 - \eta_2 + \delta\eta). \quad (7)$$

One can record the rerecorded hologram again under nonlinear conditions ($\gamma \neq -2$) to provide its reconstruction in higher diffraction orders. Note that rerecorded hologram (6) is free of the inhomogeneities of the substrate under original multiple-exposure hologram (2).

For rerecording the second holographic structure corresponding, for example, to the *c*th exposure, we use two coherent waves with the spatial frequencies ξ'_{01}, η'_{01} and ξ'_{02}, η'_{02} meeting the conditions

$$\xi'_{01} = \xi_1 - \xi_c, \quad \eta'_{01} = \eta_1 - \eta_c, \quad \xi'_{02} = \xi_2 + \xi_c, \quad \eta'_{02} = \eta_2 + \eta_c, \quad (8)$$

where $\xi_c = \cos\gamma_{cx}/T_c$, $\eta_c = \cos\gamma_{cy}/T_c$, and T_c is the period of the fringes in the *c*th holographic structure making angles γ_{cx} and γ_{cy} with the *0X* and *0Y* axes, respectively, in the plane of the multiple-exposure hologram.

Note that the *c*th holographic structure can be located in either the first or the second multiple-exposure hologram. When the holographic structures belong to different carriers, aberrations are the most difficult to compensate. In this case, the *c*th holographic structure is rerecorded in the following way. The second multiple-exposure hologram placed in position *1* (Fig. 1a) is illuminated by two coherent waves with condition (8) met. In this case, the spatial-frequency filter selects the waves

$$A'_1 = a_c \exp\{i[2\pi(\xi_1 + \delta\xi'_1)x + 2\pi(\eta_1 + \delta\eta'_1)y + \varepsilon'_b + \varphi + \psi'_0]\},$$
$$A'_2 = a_c \exp\{i[2\pi(\xi_2 + \delta\xi'_2)x + 2\pi(\eta_2 + \delta\eta'_2)y - \varepsilon'_c - \varphi + \psi'_0]\}, \quad (9)$$

where $\delta\xi'_1$, $\delta\eta'_1$, $\delta\xi'_2$, and $\delta\eta'_2$ are the adjustment errors arising for the same reasons as in the former case and the function ψ'_0 takes into account optical inhomogeneities in the substrate under multiple-exposure hologram (3).

The rerecording of the second holographic structure involves the replacement of multiple-exposure holograms and the adjustment of the optical system for the new condition for the spatial frequencies of the incident waves to be satisfied. Therefore, the misadjustments upon rerecording of the first and the second holographic structures are different. However, the vector of the holographic structure rerecorded depends on the difference in the errors rather than on their absolute values. Then, to equate the vectors of the holographic structures rerecorded, one can match the interference field rerecording the second holographic structure with the already rerecorded first holographic structure [19]. This can be accomplished by placing the new carrier in the same position *5* (Fig. 1a) after rerecording the first holographic structure on it. The equality of the vectors of the second interference holographic structure being rerecorded and the first already rerecorded holographic structure is monitored by the width of the moiré fringes. The moiré pattern represents the low-frequency modulation of these periodic structures, which can be observed on rerecorded holographic structure *5* in both transmitted and scattered light. By additionally adjusting one of the beams illuminating second multiple-exposure hologram *1*, one can vary the moiré pattern and obtain a fringe of an infinite width. Then, the carrier with the first rerecorded holographic structure is removed and the second holographic structure is recorded on the new carrier placed in position *5*. An infinitely wide fringe in the moiré pattern corresponds to equal differences in the adjustment errors:

$$\delta\xi_1 - \delta\xi_2 = \delta\xi'_1 - \delta\xi'_2, \quad \delta\eta_1 - \delta\eta_2 = \delta\eta'_1 - \delta\eta'_2. \quad (10)$$

Thus, the amplitude transmission of the *c*th rerecorded holographic structure has the form

$$\tau_c = \{1 + \cos[2\pi(\xi_1 - \xi_2 + \delta\xi)x + 2\pi(\eta_1 - \eta_2 + \delta\eta)y + 2(\varepsilon'_c + \varphi)]\}^{-\gamma/2}. \quad (11)$$

As long as equalities (10) are satisfied, the projections of the vectors of the rerecorded holographic arrays in expressions (6) and (11) equal each other, which provides equality of the vectors themselves.

One can process the pair of holograms (6) and (11) in an optical analyzer of conjugate holograms [3] with the use of incoherent light. For example, the selection of the $\pm n$th diffraction orders in the record plane makes it possible to observe an interference pattern adjusted to an infinitely wide fringe. Such a pattern visualizes the object dynamics in the time interval between the *b*th and *c*th exposures upon the recording of the original multiple-exposure holograms:

$$I_{cb} = 1 + \cos[4n(\varepsilon_b - \varepsilon_c)]. \quad (12)$$

The sensitivity of such an interferogram is increased $4n$ times. The aberrations in the recording and optical processing systems, as well as the inhomogeneities of the substrates, are completely compensated.

EXPERIMENTAL RESULTS

The above method for improving the sensitivity and eliminating aberrations was tested in experiments on visualizing the temperature field in a glass substrate with a transparent conducting coating near one of the electrodes. The voltage applied to the electrodes was varied with time by a certain law. The substrate with the conducting coating was placed in the object arm of the holographic interferometer. The state of the substrate was holographically recorded at certain time instants. In one experiment, we recorded 15 states of the substrate (five holographic structures on each of three carriers) by varying the orientation of the fringes in the structures. When the interference patterns visualizing the temperature-field variations in the substrate for a time between two arbitrary exposures were reconstructed by the conventional methods without increasing the sensitivity, the curvature of the fringes was so small that it was impossible to estimate the temperature-field variations even qualitatively. When the interference patterns were reconstructed in the interferometric correlator with two different multiple-exposure holograms, the interference fringes were heavily distorted by substrate inhomogeneities. Figure 2 shows such an interferogram adjusted to an infinitely wide fringe and reconstructed with the holographic structures recorded at the corresponding time instants. The holographic structures were taken from different multiple-exposure holograms. It is seen from the interferograms (the electrode is on the right) that the aberrations significantly exceed the desired signal. To compensate for the inhomogeneities in the multiple-exposure hologram substrates and to increase the sensitivity, we rerecorded the holographic structures obtained at the time instants of interest on new carriers by the method proposed. The new holograms exhibited the nonlinearity allowing for the use of higher diffraction orders. Figure 3 shows interferograms reconstructed through the rerecorded holograms in the optical analyzer of conjugate holograms with an incoherent light source. We achieved an eightfold increase in the sensitivity when reconstructing the interferograms in the diffraction orders ± 2. By exactly superimposing the rerecorded holograms in the optical analyzer, we reconstructed the interferogram adjusted to the infinitely wide fringe (Fig. 3a). Note that such an adjustment of the interference pattern with complete compensation of aberrations necessitates the equality of the vectors of the holographic structures. The adjustment to finite-width fringes was accomplished by slightly rotating one of the holograms relative to another about the optical axis of the device. However, interferograms reconstructed in such a way (Fig. 3b) can suffer from uncompensated aberrations introduced by systems recording multiple-exposure holograms and by those rerecording holographic structures. These uncompensated aberrations depend on the angle of rotation and on the form of the function φ.

Fig. 2. Interferogram reconstructed from two multiple-exposure holograms in the interferometric correlator.

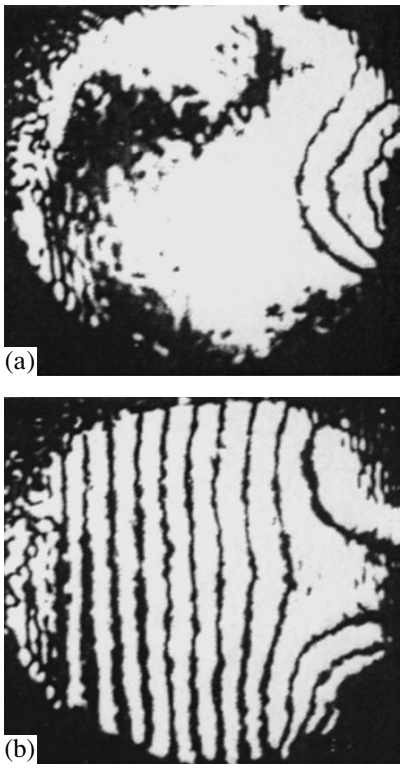


Fig. 3. Interferograms adjusted to (a) an infinitely wide fringe and (b) finite-width fringes and reconstructed from two rerecorded holograms in the optical analyzer of conjugate holograms.

CONCLUSION

In the interferograms obtained, the inhomogeneities in the substrates of the multiple-exposure holograms are totally eliminated. The increase in the sensitivity is sufficient for quantitative estimate of the variations of the temperature fields in the glass substrate. If the sen-

sitivity turns out to be insufficient, the holographic structures rerecorded on the new carriers [see expressions (6) and (11)] can additionally be rerecorded in coherent or incoherent light by the method presented in [17] to achieve the desired sensitivity.

ACKNOWLEDGMENTS

The work was carried out within the Interinstitute Program "Beam Interactions" and supported by the Ministry of Education of Belarus.

REFERENCES

1. C. M. Vest, *Holographic Interferometry* (Wiley, New York, 1979; Mir, Moscow, 1982).
2. Yu. I. Ostrovskiĭ, M. M. Butusov, and G. V. Ostrovskaya, *Holographic Interferometry* (Nauka, Moscow, 1977).
3. A. K. Beketova, A. F. Belozerov, A. N. Berezkin, *et al.*, *Holographic Interferometry of Phase Objects* (Nauka, Leningrad, 1979).
4. O. Bryngdahl and A. W. Lohmann, J. Opt. Soc. Am. **58**, 141 (1968).
5. K. Matsumotto and M. J. Takashima, J. Opt. Soc. Am. **60** (1), 30 (1970).
6. J. Schwider, in *Proceedings of the III All-Union School on Holography, Leningrad, 1972*, p. 247.
7. I. S. Zeĭlikovich and S. A. Pul'kin, Opt. Spektrosk. **53**, 588 (1982) [Opt. Spectrosc. **53**, 349 (1982)].
8. V. A. Afanas'eva, L. T. Mustafin, and V. A. Seleznev, Opt. Spektrosk. **37**, 788 (1974).
9. Shufen Fu and Jianwen Chen, Opt. Commun. **67**, 417 (1988).
10. D. Apostol, D. Barbulesku, I. I. Komisarova, *et al.*, Zh. Tekh. Fiz. **58**, 2156 (1988) [Sov. Phys. Tech. Phys. **33**, 1309 (1988)].
11. S. V. Vasiljev, A. Y. Ivanov, and V. I. Nedolugov, Proc. SPIE **2340**, 454 (1994).
12. R. J. Collier, C. B. Burckhardt, and L. H. Lin, *Optical Holography* (Academic, New York, 1971; Mir, Moscow, 1973).
13. Yu. N. Zakharov and S. N. Mensov, *Physical Foundations and Applied Problems of Holography* (Leningr. Inst. Yad. Fiz., Leningrad, 1984), pp. 138–141.
14. *Optical Data Processing: Applications*, Ed. by D. Casasent (Springer-Verlag, New York, 1978; Mir, Moscow, 1980).
15. *Handbook of Optical Holography*, Ed. by H. J. Caulfield (Academic, New York, 1978; Mir, Moscow, 1982), Vol. 2.
16. I. S. Zeĭlikovich, Opt. Spektrosk. **49**, 396 (1980) [Opt. Spectrosc. **49**, 215 (1980)].
17. I. S. Zeĭlikovich and A. M. Lyalikov, Usp. Fiz. Nauk **161** (1), 143 (1991) [Sov. Phys. Usp. **34**, 74 (1991)].
18. A. M. Lyalikov, Pis'ma Zh. Tekh. Fiz. **24** (12), 72 (1998) [Tech. Phys. Lett. **24**, 487 (1998)].
19. I. S. Zeĭlikovich, A. M. Lyalikov, and V. V. Sigov, USSR Inventor's Certificate No. 1368624, Byull. Izobret., No. 3 (1988).

Translated by A. Chikishev

Technical Physics, Vol. 47, No. 6, 2002, pp. 713–719. Translated from Zhurnal Tekhnicheskoĭ Fiziki, Vol. 72, No. 5, 2002, pp. 61–66.

OPTICS, QUANTUM ELECTRONICS

Soliton Formation from a Gaussian Pulse in an Optical Fiber

I. V. Dzedolik and A. I. Dzedolik

Taurida Vernadsky National University, Simferopol, 95007 Ukraine
e-mail: dzedolik@crimea.edu

Received September 24, 2001

Abstract—Propagation of a frequency-modulated Gaussian pulse along an optical fiber with a gradient refractive index is studied. It is shown that, after propagating a certain distance, the pulse undergoes compression. If, at the point of maximum compression, the pulse intensity is exactly equal to the threshold intensity, then either a bright or dark vortex soliton can form. *© 2002 MAIK "Nauka/Interperiodica".*

INTRODUCTION

At present, most of the methods for information transfer along optical fibers are based on pulse coding. Both the need of pulse shortening with the purpose of the enhancing information transfer rate and the necessity of lengthening the transmission line require increasing the power of the pulse transmitted along an optical fiber. In this case, the pulse propagation along the fiber can be substantially affected by nonlinear effects.

The formation of an envelope soliton in an optical fiber from a pulse with a bounded spectrum and a peak power of ~1 W was predicted theoretically in [1] and demonstrated experimentally in [2]. Conditions for the formation of bright and dark solitons in a fiber were studied in [3]. In [4], it was shown that, in an optical fiber, a stable laser pulse can evolve into a soliton. The stability of soliton pulses generated by exciting two modes with different group velocities in an optical fiber was studied analytically in [5]. In [6], the dynamics of a vortex soliton generated due to the self-focusing of cw laser radiation in a nonlinear medium was investigated numerically.

It is well known [7, 8] that a frequency-modulated pulse with a Gaussian envelope in a fiber undergoes compression. In the course of compression, the pulse shortens and its intensity increases. As a result, even if the initial pulse intensity is lower than the critical one, it can reach the threshold value at which a soliton forms. In this paper, we consider the dynamics of the amplitude and phase of a frequency-modulated Gaussian pulse in an optical fiber and show that the pulse compression can result in the formation of a bright or dark vortex soliton.

FIELD EQUATIONS

Let us consider the dynamics of an optical pulse propagating in a transversely nonuniform medium with axial symmetry and dispersion, namely, in an optical fiber with a nonuniform refractive index. We assume that, in the cross section $z = 0$, the waveguide conditions for a pulse with a given electric field profile are satisfied; i.e., the transverse profile of the pulse electric field is described by the guided modes, the radiative modes have been already emitted, and the guiding conditions for the fundamental harmonic in a fiber are fulfilled. In the transverse direction, the Sommerfeld radiation conditions for all the components of the electromagnetic field **E** and **B** are satisfied: $\lim_{r \to \infty} \{ r(k\psi_j + \partial\psi_j/\partial r)\} = 0$.

For silica glass, from which optical fibers are produced, the nonlinear response time is $\tau_{NL} \sim 10^{-15}$ s [7]. Generally, the pulse duration T_0 can be comparable with the nonlinear response time τ_{NL}. In a fiber along which the pulse propagates, the nonlinear response is usually small as compared to the linear one. Hence, the optical nonlinearities can be approximated by expanding the polarization vector in series in the electric field strength,

$$\mathbf{P} = \mathbf{P}_L + \mathbf{P}_{NL}, \tag{1}$$

where $P_{NL} \ll P_L$.

For silica glass, the largest contribution to the nonlinear response comes from the term proportional to the third power of the electric field strength,

$$\mathbf{P}(\mathbf{r}, t) = \int_0^\infty dt_1 \chi_1(\mathbf{r}, t_1)\mathbf{E}(\mathbf{r}, t - t_1)$$
$$+ \int_0^\infty\int_0^\infty\int_0^\infty dt_1 dt_2 dt_3 \chi_3(\mathbf{r}, t_1, t_2, t_3)\mathbf{E}(\mathbf{r}, t - t_1) \tag{2}$$
$$\times \mathbf{E}(\mathbf{r}, t - t_1 - t_2)\mathbf{E}(\mathbf{r}, t - t_1 - t_2 - t_3).$$

The electromagnetic field of a pulse propagating along an optical fiber satisfies Maxwell's equations for

a dielectric,

$$\nabla \times \mathbf{B} = \frac{1}{c}\frac{\partial \mathbf{D}}{\partial t}, \quad \nabla \cdot \mathbf{D} = 0,$$
$$\nabla \times \mathbf{E} = -\frac{1}{c}\frac{\partial \mathbf{B}}{\partial t}, \quad \nabla \cdot \mathbf{B} = 0, \tag{3}$$

and the constitutive equation $\mathbf{D} = \mathbf{E} + 4\pi\mathbf{P} \equiv \hat{\varepsilon}\,\mathbf{E}$.

From set (1)–(3), we obtain the integrodifferential equation for **E**:

$$\nabla^2\mathbf{E} - \nabla(\nabla \cdot \mathbf{E}) - \frac{1}{c^2}\frac{\partial^2 \mathbf{E}}{\partial t^2} = \frac{4\pi}{c^2}\frac{\partial^2 \mathbf{P}}{\partial t^2}. \tag{4}$$

By virtue of the equation $\nabla\mathbf{D} = 0$, we have $\hat{\varepsilon}\,\nabla\mathbf{E} + (\nabla\hat{\varepsilon}) \cdot \mathbf{E} = 0$. We ignore the polarization effects and assume that $|\hat{\varepsilon}\,\nabla \cdot \mathbf{E}| \gg |(\nabla\hat{\varepsilon}) \cdot \mathbf{E}|$; i.e., $\nabla \cdot \mathbf{E} \approx 0$. Then, from Eq. (4), we obtain

$$\left(\nabla^2 - \frac{1}{c^2}\frac{\partial^2}{\partial t^2}\right)\mathbf{E} = \frac{4\pi}{c^2}\frac{\partial^2}{\partial t^2}(\hat{\chi}_1 + \hat{\chi}_3|\mathbf{E}|^2)\mathbf{E}, \tag{5}$$

where

$$\hat{\chi}_1 = \int_0^\infty dt_1\chi_1(\mathbf{r}, t_1), \quad \hat{\chi}_3 = \int_0^\infty\int_0^\infty\int_0^\infty dt_1dt_2dt_3\chi_3(\mathbf{r}, t_1, t_2, t_3)$$

are the integral operators.

We separate out the second derivative with respect to the longitudinal coordinate in the Laplace operator, $\nabla_\perp^2 + \partial^2/\partial z^2$, and multiply Eq. (5) by the complex conjugated value **E*** and the complex conjugated Eq. (5) by **E** and, then, subtract these equations from one another. After integrating the equation obtained over the cross section, we arrive at

$$\iint_S dS\left\{(\mathbf{E}^* \cdot \nabla_\perp^2\mathbf{E} - \mathbf{E} \cdot \nabla_\perp^2\mathbf{E}^*) + \left[\mathbf{E}^* \cdot \left(\frac{\partial^2}{\partial z^2} - \frac{1}{c^2}\frac{\partial^2}{\partial t^2}\hat{\varepsilon}\right)\mathbf{E} - \mathbf{E} \cdot \left(\frac{\partial^2}{\partial z^2} - \frac{1}{c^2}\frac{\partial^2}{\partial t^2}\hat{\varepsilon}\right)\mathbf{E}^*\right]\right\} = 0, \tag{6}$$

where $\hat{\varepsilon} = 1 + 4\pi\hat{\chi}_1 + 4\pi\hat{\chi}_3|\mathbf{E}|^2$.

Taking into account the Sommerfeld radiation condition

$$\lim_{r\to\infty}\left[r\left(\frac{\omega}{c}\mathbf{E} + \frac{\partial}{\partial r}\mathbf{E}\right)\right] = 0$$

and the second Green's formula, we find that

$$\iint_S dS(\mathbf{E}^* \cdot \nabla_\perp^2\mathbf{E} - \mathbf{E} \cdot \nabla_\perp^2\mathbf{E}^*) = \oint_L dl\left(\mathbf{E}^* \cdot \frac{\partial}{\partial r}\mathbf{E} - \mathbf{E}\frac{\partial}{\partial r}\mathbf{E}^*\right) = 0$$

at $L \longrightarrow \infty$. We represent the electric field vector in the form

$$\mathbf{E}(\mathbf{r}, t) = \mathbf{e}(\mathbf{r}_\perp)Z(z, t)$$

and substitute it into Eq. (6). Then, taking into account that $Z^* \neq 0$, we arrive at the equation for $Z(z, t)$:

$$\frac{\partial^2}{\partial z^2} - \frac{1}{c^2}\frac{\partial^2}{\partial t^2}Z = \frac{4\pi}{c^2}\frac{\partial^2}{\partial t^2}(\hat{\alpha}_1 Z + \hat{\alpha}_3 Z|Z|^2), \tag{7}$$

where

$$\hat{\alpha}_1 = \left(\int_0^{2\pi} d\varphi\int_0^\infty drr\hat{\chi}_1(\mathbf{r}_\perp)|\mathbf{e}(\mathbf{r}_\perp)|^2\right) \times \left(\int_0^{2\pi} d\varphi\int_0^\infty drr|\mathbf{e}(\mathbf{r}_\perp)|^2\right)^{-1},$$

$$\hat{\alpha}_3 = \left(\int_0^{2\pi} d\varphi\int_0^\infty drr\hat{\chi}_3(\mathbf{r}_\perp)|\mathbf{e}(\mathbf{r}_\perp)|^2|\mathbf{e}(\mathbf{r}_\perp)|^2\right) \times \left(\int_0^{2\pi} d\varphi\int_0^\infty drr|\mathbf{e}(\mathbf{r}_\perp)|^2\right)^{-1}.$$

We apply the method of slowly varying amplitudes [7, 8] to analysis of the pulse dynamics in an optical fiber. In this method, the complex amplitudes $A(z, t)$ of the pulse field components (envelopes) are assumed to slowly vary in time and along the longitudinal coordinate as compared to fast oscillations at the fundamental harmonic

$$Z(z, t) = A(z, t)\exp[i(\omega_0 t - \beta z)]. \tag{8}$$

We expand the slowly varying amplitude $A(z, t - t_1)$ in a Taylor series in the t_1 variable (the relaxation time) about the current time t:

$$A(z, t - t_1) = A(z, t) + \sum_{m=1}^{\infty}\frac{(-t_1)^m}{m!}\frac{\partial^m A(z, t)}{\partial t^m}. \tag{9}$$

Taking into account Eqs. (8) and (9), we obtain

$$\hat{\alpha}_1 Z = \left[\alpha_1(z, \omega_0)A + \sum_{m=1}^{\infty} \frac{(-i)^m}{m!} \frac{\partial^m \alpha_1(z, \omega_0)}{\partial \omega^m}\bigg|_{\omega=\omega_0} \frac{\partial^m A}{\partial t^m}\right] \times \exp[i(\omega_0 t - \beta(\omega_0)z)], \quad (10)$$

where

$$\alpha_1(z, \omega_0) = \int_0^{\infty} \alpha_1(z, t_1)\exp(-i\omega_0 t_1)dt_1.$$

For the real frequencies, the function $\alpha_1(z, \omega_0)$ has no singularities [9]. Analogously, we expand the slowly varying amplitude in series in t_1, t_2, and t_3:

$$\hat{a}_3 Z|Z|^2 = \left[\alpha_3(x, \omega_0)A|A|^2 - i\sum_{k=1}^{3} \frac{\partial \alpha_3(z, \omega_0)}{\partial \omega_k}\bigg|_{\omega_k=\omega_0} \frac{\partial A|A|^2}{\partial t} + \frac{1}{2!}\sum_{k=1}^{3}\sum_{m=1}^{3} \frac{\partial^2 \alpha_3(z, \omega_0)}{\partial \omega_k \partial \omega_m}\bigg|_{\omega_{k,m}=\omega_0} \frac{\partial^2 A|A|^2}{\partial t^2} + \ldots\right] \times \exp[i(\omega_0 t - \beta(\omega_0)z)]. \quad (11)$$

Let us consider a pulse with a duration much longer than the nonlinear response time of silica glass ($T_0 \gg \tau_{NL}$). Then, the nonlinear response of the medium can be regarded as inertia-free and the response function can be represented as a product of delta functions,

$$\chi_3(t_1, t_2, t_3) = \chi_3(\omega_0)\delta(t_1)\delta(t_2)\delta(t_3);$$

i.e., we can retain only the first term in expansion (11). We ignore the dependence of the fiber permittivity on the longitudinal coordinate. In this approximation, the equation for the electric field amplitude in a silica fiber takes the form

$$\left[\frac{\partial^2}{\partial z^2} - \frac{1}{c^2}\frac{\partial^2}{\partial t^2} - i2\beta_0\frac{\partial}{\partial z} - i2\omega_0\frac{1}{c^2}\frac{\partial}{\partial t} + \left(\frac{\omega_0^2}{c^2} - \beta_0^2\right)\right]A = \frac{4\pi}{c^2}\left[\alpha_1\frac{\partial^2}{\partial t^2} + i2\alpha_1\omega_0\frac{\partial}{\partial t} - a_1\omega_0^2 + \sum_{m=1}^{\infty}\frac{(-i)^m}{m!}\frac{\partial^m \alpha_1}{\partial \omega^m} \times\left(\frac{\partial^{m+2}}{\partial t^{m+2}} + i2\omega_0\frac{\partial^{m+1}}{\partial t^{m+1}} - \omega_0^2\frac{\partial^m}{\partial t^m}\right)\right]_{\omega_0} A + \alpha_3\frac{4\pi}{c^2} \times\left(|A|^2\frac{\partial^2 A}{\partial t^2} + A\frac{\partial^2|A|^2}{\partial t^2} + 2\frac{\partial A}{\partial t}\frac{\partial|A|^2}{\partial t} + i2\omega_0|A|^2\frac{\partial A}{\partial t} + i2\omega_0 A\frac{\partial|A|^2}{\partial t} - \omega_0^2 A|A|^2\right). \quad (12)$$

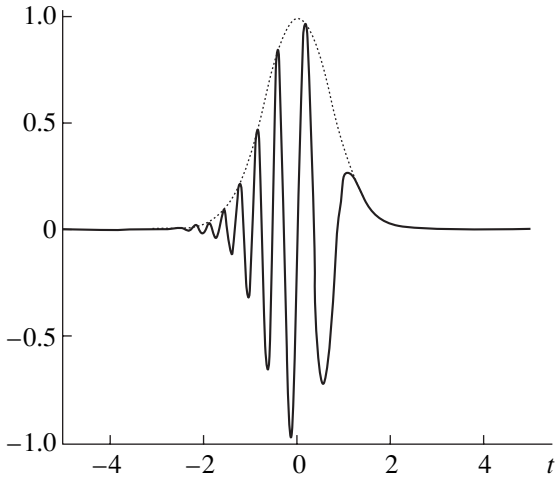


Fig. 1. Pulse with initial frequency modulation ($\nu > 0$) at $z = 0$.

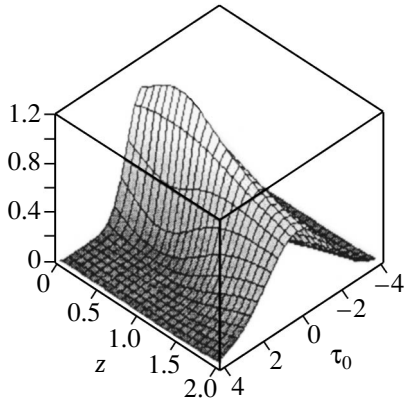


Fig. 2. Pulse envelope at $\gamma = 1$.

Equation (12) describes both linear dispersion (the terms $\sim\partial^m\alpha_1\partial^m A/\partial\omega^m\partial t^m$) and nonlinear effects, such as the breaking of the amplitude envelope (the term $\sim A|A|^2$) and its perturbations (the nonlinear terms with time derivatives of the amplitude).

DYNAMICS OF THE PULSE ENVELOPE

Let us find the solution to Eq. (12) for the envelope of a pulse with given parameters in the cross section $z = 0$ in the linear approximation, assuming $\chi_3 \longrightarrow 0$. The frequency spectrum of the pulse with an amplitude $A_0(t)$ at $z = 0$ is

$$F(\omega) = \int_{-\infty}^{\infty} A_0(t')\exp(-i\omega t')dt'. \quad (13)$$

Taking into account that harmonic waves propagate in a linear medium independently of each other, we will seek a solution to the equation for the longitudinal fac-

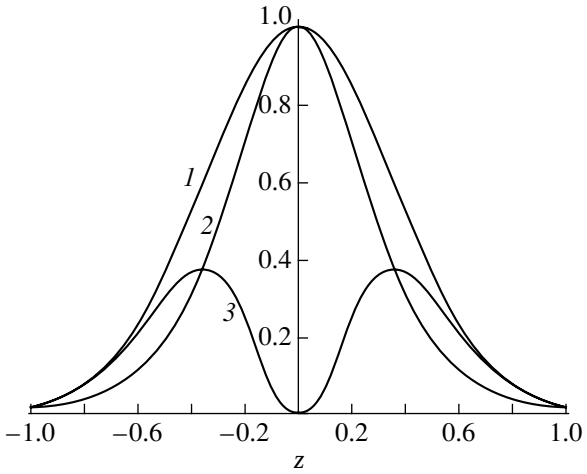


Fig. 3. Longitudinal profile of the pulse energy density: (*1*) Gaussian pulse, (*2*) bright soliton, and (*3*) dark soliton.

tor $Z(z, t)$ as a superposition of the harmonics $F(\omega)\exp\{i[\omega t - \beta(\omega)z]\}$ [7, 8]:

$$Z(z, t) = \frac{1}{2\pi}\int_{-\infty}^{\infty} F(\omega)\exp\{i[\omega t - \beta(\omega)z]\}d\omega. \quad (14)$$

We substitute expression (13) for the spectrum into Eq. (14) to arrive at the expression for $Z(z, t)$ with allowance for the boundary conditions:

$$Z(z, t) = \frac{1}{2\pi}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} A_0(t') \times \exp\{i[\omega(t - t') - \beta(\omega)z]\}dt'd\omega. \quad (15)$$

The solution to Eq. (15) can be obtained in a closed form by calculating the double integral with allowance for the exact dispersion law $\beta = \beta(\omega)$.

Let us consider the dynamics of a relatively long pulse with a smooth front. We assume that the width of the pulse frequency spectrum $\Delta\omega$ is fairly small. Then, the approximate solutions for $Z(z, t)$ can be found by expanding the propagation constant in a Taylor series about the carrier frequency ω_0:

$$\beta(\omega) = \beta(\omega_0) + \left.\frac{\partial\beta}{\partial\omega}\right|_{\omega=\omega_0}(\omega - \omega_0) + \frac{1}{2}\left.\frac{\partial^2\beta}{\partial\omega^2}\right|_{\omega=\omega_0}(\omega - \omega_0)^2 + \ldots. \quad (16)$$

Substituting series (16) into Eq. (15), we obtain expressions for Z in the first, second, etc. approximations of dispersion theory in accordance to the terms retained. After taking the inner integral over ω and, then, the integral over t', we arrive at the solutions in the corresponding approximations of dispersion theory. In the second approximation of dispersion theory, we obtain the following expression for Z:

$$Z(z, t) = \frac{1}{2\pi}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} A_0(t')\exp(-i\omega t)dt' \times \exp\left\{i\left[\left(t - \frac{\partial\beta}{\partial\omega}z\right)(\omega - \omega_0) - \frac{1}{2}\frac{\partial^2\beta}{\partial\omega^2}z(\omega - \omega_0)^2\right]\right\} \times d(\omega - \omega_0)\exp[i(\omega_0 t - \beta(\omega_0)z)] \equiv A(z, t)\exp[i(\omega_0 t - \beta_0 z)]. \quad (17)$$

If a Gaussian pulse is initially frequency modulated, its envelope at $z = 0$ can be represented in the form [7, 8]

$$A_0(t) = \exp(-t^2/T_0^2 - i\omega_0 v t^2). \quad (18)$$

When the frequency parameter is positive ($v > 0$), the frequency decreases toward the pulse trailing edge (Fig. 1); in the opposite case ($v < 0$), it increases toward the trailing edge.

Substituting expression (18) for $A_0(t)$ into Eq. (17), we obtain the envelope of a frequency-modulation pulse in the fiber core in the cross section z:

$$A(z, t) = \frac{1}{[(1 - \gamma\bar{z})^2 + \bar{z}^2]^{1/4}} \times \exp\left\{-\frac{(t - \bar{z}L_D/v_g)^2}{T_0^2[(1 - \gamma\bar{z})^2 + \bar{z}^2]}\right\}\exp\{i\vartheta\}, \quad (19)$$

where

$$\vartheta = \frac{(t - \bar{z}L_D/v_g)^2[\bar{z}(1 + \gamma^2) - \gamma]}{T_0^2[(1 - \gamma\bar{z})^2 + \bar{z}^2]} - \frac{1}{2}\arctan\left(\frac{\bar{z}}{1 - \gamma\bar{z}}\right)$$

is the correction to the pulse phase, $\gamma = v\omega_0 T_0^2$ is the initial frequency modulation depth, $\bar{z} = z/L_D$ is the normalized fiber length, $L_D = T_0^2/4D$ is the dispersion length over which the pulse duration increases by a factor of $\sqrt{2}$,

$$D = \frac{1}{2}\frac{\partial^2\beta}{\partial\omega^2}$$

is the dispersion factor, and $v_g = (\partial\beta/\partial\omega)^{-1}$.

Figure 2 shows the dynamics of the amplitude $|A|$ of a frequency-modulated pulse propagating in the fiber core for $\gamma = 1$ [see Eq. (19)]. The pulse amplitude increases over the normalized distance $\bar{z} = 0.5$, on which the pulse undergoes compression in the normalized time $\tau_0 = (t - \bar{z}L_D/v_g)/T_0$. After passing the point of maximum compression, the pulse spreads out. The rate of the pulse envelope deformation increases with

increasing the coefficient γ. For example, at $\gamma = 10$, a pulse with initial frequency modulation is not compressed at all.

For $\gamma = 1$ and $D > 0$ (normal dispersion), the duration of a frequency-modulated pulse propagating in the fiber core, $2T = 2T_0[(1 - \gamma z/L_D)^2 + (z/L_D)^2]^{1/2}$, has a minimum at the distance $z_{\min} = \gamma T_0^2/4D(1 + \gamma^2)$. Therefore, in a linear medium, the frequency-modulated pulse is "focused" in time. Due to compression, the pulse energy density increases, so that nonlinear effects can come into play if the energy density reaches a level at which the refractive index begins to depend on the electric field strength.

BRIGHT AND DARK SOLITONS

Let us consider in more detail the effects of linear dispersion and the nonlinear breaking of the amplitude envelope, ignoring other processes. We go over to the frame of reference related to the pulse ($\zeta = z$ and $\tau = t - z/v$, where $v = \text{const}$) and switch to the new variables in Eq. (12). The velocity of the moving frame of reference can be found by equating to zero the coefficient by the first derivative with respect to τ in the equation obtained:

$$v = \frac{\beta_0}{\omega_0(1 + 4\pi\alpha_1)/c^2 + 2\pi(\omega_0^2/c^2)\partial\alpha_1/\partial\omega}. \quad (20)$$

It follows from (20) that the velocity of the new frame of reference is determined by the frequency of the fundamental harmonic ω_0 and the dispersion of the dielectric susceptibility of the fiber core $\partial\alpha_1/\partial\omega$. We write the equation for the pulse amplitude in the form

$$b_1\frac{\partial^2 A}{\partial z^2} + b_2\frac{\partial^2 A}{\partial \tau^2} - i\frac{\partial A}{\partial z} + b_3 A + b_4 A|A|^2 = 0, \quad (21)$$

where

$$b_1 = \frac{1}{2\beta_0},$$

$$b_2 = \frac{1}{2\beta_0 c^2}\left[\frac{c^2}{v^2} - 1 - 4\pi\alpha_1 + 8\pi\omega_0\frac{\partial\alpha_1}{\partial\omega} + 2\pi\omega_0^2\frac{\partial^2\alpha_1}{\partial\omega_0}\right],$$

$$b_3 = \frac{\omega_0^2(1 + 4\pi\alpha_1)/c^2 - \beta_0^2}{2\beta_0}, \quad b_4 = \frac{2\pi\alpha_3\omega_0^2}{\beta_0 c^2}.$$

When the rate of dispersion spreading is exactly equal to the rate of the nonlinear breaking of the amplitude envelope, we can find the steady-state pulse shape [7] assuming that

Fig. 4. Wave surface of a bright soliton.

Fig. 5. Wave surface of a dark soliton.

$$A(z, \tau) = \tilde{A}(\tau)\exp(-i\Gamma z). \quad (22)$$

Substituting expression (22) into Eq. (21), we arrive at the equation

$$g\frac{d^2\tilde{A}}{d\tau^2} = h\tilde{A} - \tilde{A}^3, \quad (23)$$

where $g = b_2/b_4$ and $h = (\Gamma^2 b_1 + \Gamma - b_3)/b_4$.

In the given approximation, for $(d\tilde{A}/d\tau)_0 = 0$, $g > 0$, and $h > 0$, we obtain a solution to Eq. (23) in the form of a bright soliton,

$$\tilde{A}(\tau) = \frac{\sqrt{2}\sqrt{h}}{\cosh\left(\sqrt{\frac{h}{g}}\tau\right)}. \quad (24)$$

The pulse intensity ($I \sim v\tilde{A}^2$) is proportional to

$$I \sim \frac{2vh}{[(1 - \gamma\bar{z}_s)^2 + \bar{z}_S^2]^{1/2}}$$

$$\times \exp\left\{-\frac{2(t - \bar{z}_S L_D/v_g)^2}{T_0^2[(1 - \gamma\bar{z}_S)^2 + \bar{z}_S^2]}\right\}\frac{1}{\cosh^2(\sqrt{h/g}\tau)},$$

where $\bar{z}_s$ is the fiber cross section in which the soliton forms.

If $g < 0$ and $(d\tilde{A}/d\tau)_0 = h/\sqrt{2|g|}$, then the solution to Eq. (23) has the form of a kink wave,

$$\tilde{A}(\tau) = \sqrt{h}\tanh\left(\sqrt{\frac{h}{2|g|}}\tau\right), \quad (25)$$

at $\tilde{A}^2 < h$ or an "antikink" at $\tilde{A}^2 > h$. In the latter case, the field intensity is proportional to

$$I \sim \frac{vh}{[(1-\gamma\bar{z}_s)^2 + \bar{z}_S^2]^{1/2}}$$
$$\times \exp\left\{-\frac{2(t-\bar{z}_S L_D/v_g)^2}{T_0^2[(1-\gamma\bar{z}_S)^2+\bar{z}_S^2]}\right\}\left[1 - \frac{1}{\cosh^2(\sqrt{h/2|g|}\tau)}\right].$$

In the center of the kink or antikink (at $\tau = 0$), the amplitude is zero; i.e., a dark soliton is formed [1, 3, 7–9]. The parameter g depends on both the first and second frequency derivatives of the dielectric susceptibility, i.e., on the sign of the dispersion and on the nonlinear properties of the fiber (χ_3).

A steady-state pulse (soliton) in a fiber has the form

$$A(z, t) = \tilde{A}(t - z/v)\exp\{i[\omega_0 t - (\beta_0 + \Gamma)z]\}, \quad (26)$$

where $\tilde{A}(t - z/v)$ is described by expression (24) or (25).

We denote the pulse duration by $\tau_S = \sqrt{|g|/h}$. Then, we can express the soliton amplitude A_0 and the propagation constant Γ in terms of the duration. Solving the equation for Γ,

$$b_1\Gamma^2 + \Gamma - b_3 - b_2/\tau_S^2 = 0,$$

we obtain

$$\Gamma_\pm = -\beta_0 \pm \left[\frac{\omega_0^2}{c^2}(1 + 4\pi\alpha_1) + \frac{2\beta_0 b_2}{\tau_S^2}\right]^{1/2} \equiv -\beta_0 + \tilde{\Gamma}.$$

For the pulse amplitude, we have $A_0 = \sqrt{h} = \sqrt{b_2/b_4\tau_S^2}$. We will describe a soliton pulse propagating along a fiber by the expression

$$A(z, t) = \tilde{A}(t - z/v)\exp[i(\omega_0 t - \tilde{\Gamma}z)]. \quad (27)$$

The soliton velocity, phase, and amplitude depend on the type of dispersion in the fiber:

$$\left(\sim\frac{\partial\alpha_1}{\partial\omega}, \frac{\partial^2\alpha_1}{\partial\omega^2}\right).$$

The soliton amplitude and phase also depend on the soliton duration (τ_S).

Figure 3 shows the longitudinal energy density profiles for a Gaussian pulse, $W \sim \exp(-2\tau^2/T_0^2)$; a bright soliton, $W \sim \cosh^{-2}(-\tau/\tau_S)$; and a dark soliton, $W \sim \exp(-2\tau^2/T_0^2)[1 - \cosh^{-2}(-\tau/\tau_S)]$. The wave surfaces of the pulses are shown in Figs. 4 and 5.

We assume that the soliton forms from a pulse excited on one of the modes of an optical fiber with a parabolic refractive index $n = \sqrt{\varepsilon_1(1 - 2\Delta R^2)}$ [10]:

$$\psi = w^{l/2}\exp(-w/2)L_N^l(w)A(z, t)$$
$$\times \exp[i(\omega_0 t - \beta_0 z + \vartheta(z, t) + \kappa l\varphi)], \quad (28)$$

where $w = VR^2$, $R = r/\rho$, $V^2 = 2\rho^2(\omega_0/c)^2\varepsilon_1\Delta$, ρ is the core radius, and $L_N^l(w)$ ($N = 0, 1, 2, \ldots$) is the generalized Laguerre polynomial.

When the mode azimuthal index l is not zero, the real and imaginary parts of the pulse electric field are both zero on the fiber axis and the phase is indefinite at $r = 0$. In this case, a nonsteady optical vortex over the pulse length $\sim\tau_S$ is formed. A nodal line formed by the singular points of the electric field moves along the fiber axis. Thus, both the bright and dark solitons formed from a single pulse are nonsteady optical vortices, whose axis is the nodal line. The vortex axis can be regarded as a string (see, e.g., review [11]). Moreover, the dark soliton has one more singular point in its center; thus, for a dark soliton, the field singularity on the axis is doubly degenerated.

DISCUSSION

Let us estimate the optical fiber length over which a Gaussian pulse with initial frequency modulation ($\gamma = 1$) and an intensity lower than the threshold one ($I_G \sim v_g A_0^2 < v\tilde{A}^2$) evolves into a soliton. At $v_g = v$, $g > 0$, and $h = 1/2$, the compression of a frequency-modulated Gaussian pulse in the linear regime results in the formation of a bright soliton over the length $z = L_D/2$. For a pulse with $T_0 = 10^{-12}$ s at $\gamma = 1$ and $\partial^2\beta/\partial\omega^2 = 10^{-26}$ s^2/m, the fiber length over which the bright soliton forms is $z = 25$ m. If the intensity of the compressed pulse exceeds the threshold value for soliton formation, then the Gaussian pulse spreads out after passing the point of maximum compression [7]. At $g < 0$ and $(d\tilde{A}/d\tau)_0 = h/\sqrt{2|g|}$, all other parameters being the same, a dark soliton can be formed over a fiber length of $z = 25$ m provided that $h = 1$.

REFERENCES

1. A. Hasegawa and F. Tappert, Appl. Phys. Lett. **23**, 142 (1973).
2. L. F. Mollenauer, R. H. Stolen, and J. P. Gordon, Phys. Rev. Lett. **45**, 1045 (1980).

3. M. Jain and N. Tzoar, Opt. Lett. **3** (5), 202 (1978).
4. K. J. Blow and D. Wood, Opt. Commun. **58** (5), 349 (1986).
5. Yu. S. Kivshar', Kvantovaya Élektron. (Moscow) **17**, 1603 (1990).
6. C. T. Law, X. Zhang, and G. A. Swartzlander, Opt. Lett. **25** (11), 55 (2000).
7. S. A. Akhmanov, V. A. Vysloukh, and A. S. Chirkin, *The Optics of Femtosecond Pulses* (Nauka, Moscow, 1988).
8. M. B. Vinogradova, O. I. Rudenko, and A. P. Sukhorukov, *Theory of Waves* (Nauka, Moscow, 1990).
9. L. D. Landau and E. M. Lifshitz, *Electrodynamics of Continuous Media* (Nauka, Moscow, 1982; Pergamon, New York, 1984).
10. A. W. Snyder and J. D. Love, *Optical Waveguide Theory* (Chapman and Hall, London, 1983; Radio i Svyaz', Moscow, 1987).
11. A. Yu. Morozov, Usp. Fiz. Nauk **162** (8), 83 (1992) [Sov. Phys. Usp. **35**, 671 (1992)].

Translated by N. Ustinovskiĭ

Technical Physics, Vol. 47, No. 6, 2002, pp. 720–725. Translated from Zhurnal Tekhnicheskoĭ Fiziki, Vol. 72, No. 5, 2002, pp. 67–72.

OPTICS, QUANTUM ELECTRONICS

Gradient Light-Absorbing SiO_x/Me Coatings for Display Panels

I. Z. Indutnyy*, P. E. Shepeliavyi*, E. V. Michailovskaya*, C. W. Park, J. B. Lee**, and Y. R. Do****

**Institute of Semiconductor Physics, National Academy of Sciences of Ukraine, Kiev, 03028 Ukraine*
e-mail: indutnyy@isp.kiev.ua
***Corporate R&D Center, Samsung SDI Ltd., 575, Shing-Dong, Paldal-Gu, Suwon City, Kyungki-Do, Korea, 442-390*

Received October 22, 2001

Abstract—Thin (<0.5 μm) gradient SiO_x/Ti coatings with a low coefficient of specular reflection (≤1%) in the visible part of the spectrum are obtained by thermal evaporation *in vacuo*. The optical parameters and the distribution of the components across the coating are studied. The optical properties of the gradient layers are simulated, and the results of simulation are compared with experimental data. These coatings can be used in fabricating black (light-absorbing) matrices for cathode-ray tubes and flat-panel displays. © *2002 MAIK "Nauka/Interperiodica".*

INTRODUCTION

Light-absorbing (black) matrices are widely used in modern cathode-ray tubes and various flat-panel displays. In matrix screens, the space between pixels is filled with a black coating which absorbs environmental light and the radiation scattered by adjacent pixels. The absence of the optical background improves the image contrast.

The coatings used in black matrices must have low coefficients of diffuse (R_d) and specular R_{sp} reflection in the visible part of the spectrum ($R \le 1\%$), be opaque (optical density $D > 3$–4), and offer a sufficiently high conductivity. In addition, the black coating in modern flat-panel displays must be of a submicron thickness. Low reflection, high optical density, and submicron thickness are impossible to combine in homogeneous coatings. For example, for the reflection from the coating–screen glass interface to be low, the refractive indices n of the coating and the substrate must differ insignificantly and the absorption coefficient k must be low. However, the optical density of a submicron coating will be low if k is high. This problem can be solved with the use of multilayer or gradient films whose composition (hence, optical properties) varies in the direction normal to the substrate. Gradient coatings for the black matrices can be applied by various techniques. Among them are controllable vacuum evaporation of two components from two evaporators [1] and rf sputtering of a metal (Me) in a variable-composition atmosphere [2] to form MeO–Me coatings. One more way suggested by us for applying SiO–Cr insulator–metal coatings [3] consists in vacuum evaporation of the component mixture from a single evaporator. Because of a slight difference in the temperatures of evaporation for SiO and Cr, the process conditions can be selected so that the composition of the resulting gradient layer varies gradually from SiO to Cr. Such coatings provide desired optical and electrical properties at thicknesses from 0.3 to 0.6 μm. Unfortunately, the process of chromium vacuum deposition and subsequent lithography of the coating are not environmentally safe.

In this work, we study the properties of chromium-free gradient insulator–metal coatings for matrix panels.

SIMULATION OF OPTICAL PROPERTIES OF GRADIENT INSULATOR–METAL COATINGS

To make use of the single-evaporator method, it is necessary to take a pair of components that have close temperatures of evaporation *in vacuo*. As an insulator, we took silicon monoxide, because good SiO_x layers ($x = 1.2$–1.3 in our experiments [4]) can be obtained by thermal evaporation *in vacuo*. Moreover, the refractive index of thermally evaporated SiO_x films differs from that of panel glasses only slightly and the coefficient of absorption in the visible part of the spectrum is low. Along with Cr, which was used as a metallic component in our early experiments, one can try Fe, Ti, Co, V, and Ni. The temperatures of evaporation of Mo, Ta, amd W exceed that of SiO insignificantly. However, in the molten state, these elements vigorously react with the conventional materials of resistive evaporators; therefore, SiO_x/Ti gradient coatings were studied in this work.

The coefficient of specular reflection is the basic optical property of a black coating. Let us calculate the value of R for a SiO_x/Ti heterogeneous layer in which

the composition and the optical constants n and k vary from those typical of SiO_x to those of Ti. Figure 1 shows the cross section of the gradient layer (medium *2*) on a glass substrate (medium *1*) in air (medium *3*). Let a radiation of intensity I_0 be incident from the side of medium *1* normally to the surface. We partition inhomogeneous layer *2* by planes parallel to the substrate into N quasi-homogeneous layers of the same thickness $\Delta z = z_N/N$. It is assumed that each jth layer has its own complex refractive index (the medium is nonmagnetic)

$$\hat{n}_j = n_j + ik_j.$$

We adopt that the first layer ($0 < z < z_1$) consists of SiO_x and one or several of the last layers ($z_m < z < z_N$) consist of Ti. The distribution of $\hat{n}_j$ in the intermediate region ($z_1 < z < z_m$) is arbitrary.

The reflection and transmission in such a structure are most convenient to simulate with the Abeles matrix theory, which is presented for insulators in [5] and inhomogeneous absorbing structures in [6]. Abeles showed that the amplitude coefficients of a planar wave at the boundaries of a homogeneous layer can linearly be related through the characteristic matrix. Under our assumptions, the characteristic matrix of the j layer has the form

$$M_j = \begin{vmatrix} \cos\left(\frac{2\pi}{\lambda}\hat{n}_j\Delta z\right) & -\frac{i}{\hat{n}_j}\sin\left(\frac{2\pi}{\lambda}\hat{n}_j\Delta z\right) \\ -i\hat{n}_j\sin\left(\frac{2\pi}{\lambda}\hat{n}_j\Delta z\right) & \cos\left(\frac{2\pi}{\lambda}\hat{n}_j\Delta z\right) \end{vmatrix}. \quad (1)$$

The matrix of the entire structure is found by multiplying the matrices for the homogeneous layers:

$$M_N = \prod_{j=1}^{N} M_j = \begin{vmatrix} m_{11} & m_{12} \\ m_{21} & m_{22} \end{vmatrix}. \quad (2)$$

The amplitude reflection and transmission factors (r and t, respectively) can be expressed through the elements of characteristic matrix (2):

$$r = \frac{(m_{11} + m_{12}n_a)n_g - (m_{21} + m_{22}n_a)}{(m_{11} + m_{12}n_a)n_g + (m_{21} + m_{22}n_a)},$$
$$t = \frac{2n_g}{(m_{11} + m_{12}n_a)n_g + (m_{21} + m_{22}n_a)}, \quad (3)$$

where n_g and n_a are the refractive indices of the glass and air, respectively.

Finally, the energy factors of reflection and transmission are given by the amplitude factors multiplied by their complex conjugates:

$$R = rr^*; \quad T = \frac{n_a}{n_g}tt^*. \quad (4)$$

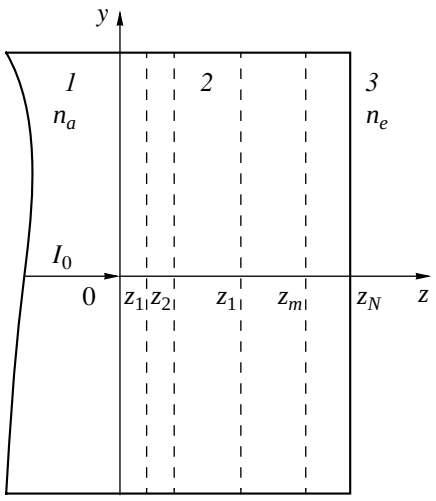


Fig. 1. Cross section of gradient layer *2* on the glass substrate *1*. n_g, refractive index of the glass; n_a, refractive index of air.

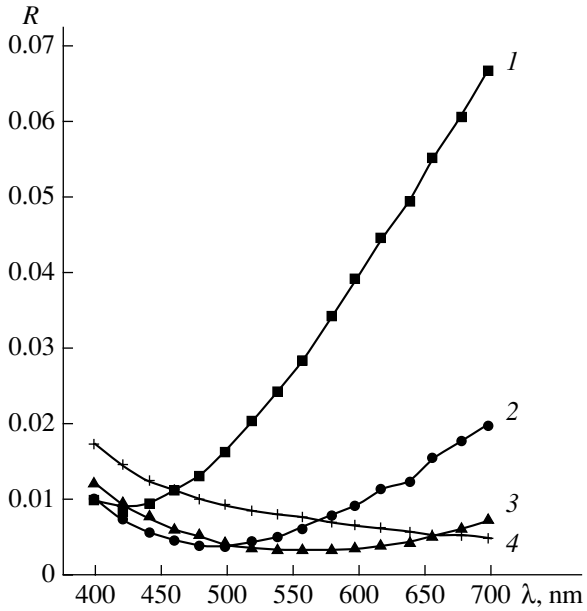


Fig. 2. Spectral dependence of the specular reflection coefficient R at the interface between the glass and the SiO_x/Ti gradient layer. $z_m - z_1 = 150$ (*1*), 250 (*2*), 350 (*3*), and 1000 nm (*4*). $N = 200$. n and k vary linearly with layer thickness.

Figure 2 shows the simulated wavelength dependence of the reflection from the glass/inhomogeneous SiO_x/Ti multilayer structure in the visible range. Curves *1–4* correspond to different thicknesses of the gradient region ($z_m - z_1$). The thickness of the metal layer ($z_N - z_m$) was selected such that the transmission was less than 10^{-5}. It should be noted that the metal layer is unnecessary if the intermediate (gradient) layer is more than 250 nm thick, because the gradient layer alone provides the desired optical density in this case. In the

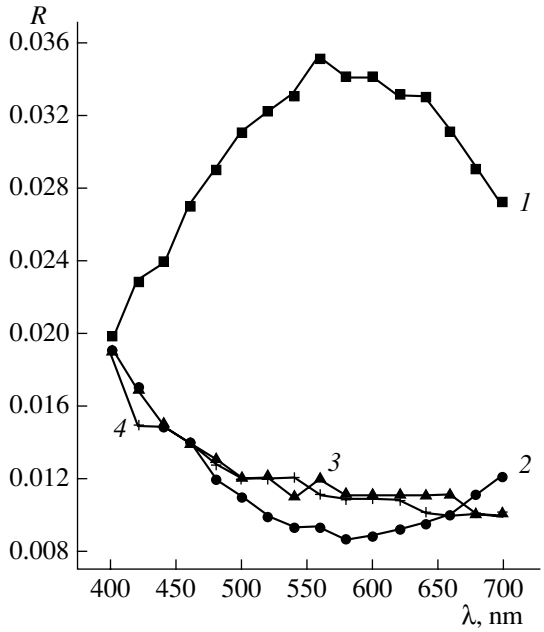


Fig. 3. Spectral dependence of the specular reflection coefficient R at the interface between the glass and the SiO_x/Ti gradient layer. Designations are the same as in Fig. 2. k varies exponentially with layer thickness.

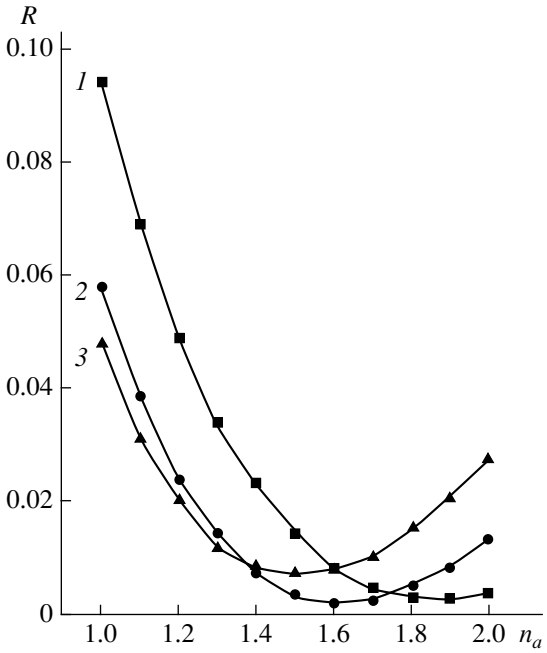


Fig. 4. Specular reflection coefficient R at the interface between the glass and the SiO_x/Ti gradient layer vs. refractive index n_g of the glass substrate. λ = 400 (*1*), 560 (*2*), and 700 nm (*3*).

simulation, we assumed that the optical constants n and k of the gradient layer vary linearly with thickness from the values typical of SiO_x (near the substrate) to those of Ti at the point z_m. The optical constants for Ti were taken from [17], and those for thermally evaporated SiO_x were measured by the procedure described in [8]. In experiments, we used silicon monoxide from Cerac Inc.; the refractive index used in the calculations was that of BK7 glass (Schott Glass). As follows from Fig. 2, the coefficient of specular reflection in the visible range depends markedly on the gradient layer thickness. It is relatively high for small thicknesses, then drops with growing thickness, and reaches a minimum in the interval $450 \le \lambda \le 650$ nm at the thickness $z_m - z_1 \approx 350$ nm. With a further increase in $z_m - z_1$, the reflection varies insignificantly and slightly increases only at $z_m - z_1 \approx 1000$ nm throughout the spectral range.

In the real gradient coatings, the optical constants may vary nonlinearly in the intermediate layer. In [3], we studied the distribution of the optical constants in an inhomogeneous black SiO/Cr layer. It was found that the refractive index in this layer varies linearly from the value for SiO to that for Cr, while the absorption coefficient is well approximated by a power law with an exponent of about 4.5 or by an exponential. It should be noted that the refractive index of such structures varies within a narrow interval, especially in the short-wave range (at λ = 400 nm, n for $SiO_{1.2}$ equals 1.9 and n_{Ti} = 2.1, while k varies from 1.4 to 2.95); hence, the optical properties of the layer are controlled largely by the gradient of the absorption coefficient.

Figure 3 shows the spectral dependence of the reflection from the SiO_x/Ti structure when the refractive index varies with thickness linearly and the absorption coefficient varies exponentially. As for the linear dependence of k, R decreases with increasing $z_m - z_1$ and is minimal in the visible range at a somewhat lesser thickness of the gradient layer, $z_m - z_1$ = 250 nm. As $z_m - z_1$ grows further, R remains virtually unchanged. The minimal value of R in this case ($R_{min} \approx 0.008$) is 2.5 times as large as R_{min} for the case when both constants vary linearly with thickness.

The reflection from the transparent substrate–gradient film interface must also depend on the difference in the optical constants at the interface (in our case, at the interface between SiO_x and the substrate material). Figure 4 shows R vs. refractive index of the substrate for several wavelengths from the visible range. The optical constants vary linearly with the gradient layer thickness, and the optimal value of $z_m - z_1$ is 350 nm. As follows from Fig. 4, $R \le 0.01$ throughout the visible range when the refractive index lies in the interval $1.50 \le n \le 1.75$, that is, when it is somewhat lower than that of $SiO_{1.2}$. A similar result is obtained for the exponential distribution of k across the gradient layer, but the optimal value of the refractive index of the substrate here is higher: $1.6 \le n \le 2.0$.

Thus, the results of simulation indicate that the spectral dependence and the value of the specular reflection at the interface between the transparent substrate and the inhomogeneous (gradient) SiO_x/Me film

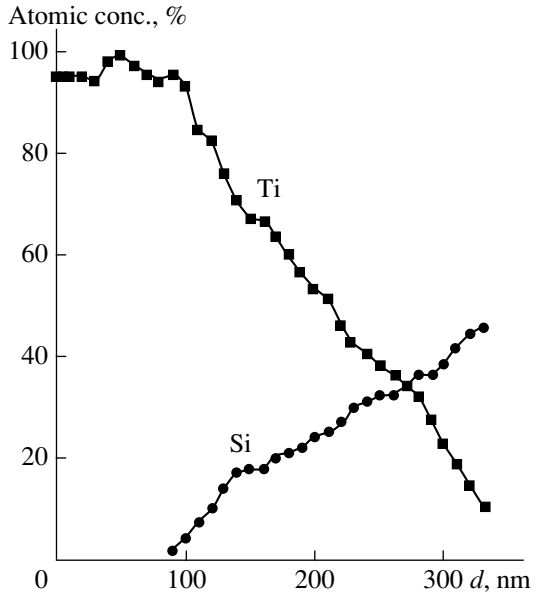


Fig. 5. Distributions of Ti and Si over the SiO_x/Ti cross section. The origin is at the coating–air interface.

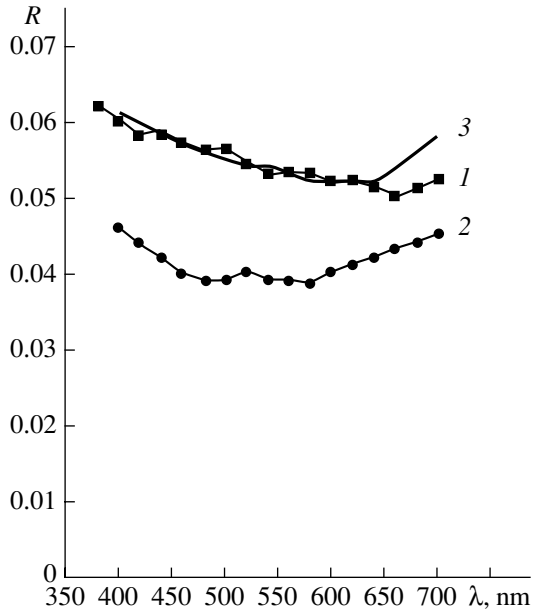


Fig. 6. (*1*) Spectral dependence of the specular reflection from the glass substrate with the SiO_x/Ti gradient layer measured from the side of the substrate. (*2*) Reflection from the air–substrate interface. (*3*) Calculated specular reflection from the SiO_x/Ti coating measured from the side of the substrate (the Maxwell–Garnett effective medium model with the gradient of the metallic phase).

are specified basically by two factors: the thickness of the variable-composition region and the difference in the refractive indices of the transparent substrate and the SiO_x film. The value of R also depends, though to a lesser extent, on the variation of the optical constants with gradient region thickness. However, given the properties of the substrate and $z_m - z_1$, the minimal value of the reflection coefficient can be varied in a wide range by varying the optical constant distribution law.

EXPERIMENTAL

Samples used were prepared by thermally evaporating *in vacuo* (residual pressure 2×10^{-3} Pa) a SiO + Ti powder mixture from a single evaporator on a glass substrate or cathode-ray tube panel. The temperature at which Ti is vigorously evaporated *in vacuo* is slightly above that of SiO. Therefore, when the temperature of the evaporator is smoothly raised, first SiO sublimates and is deposited, then SiO and Ti are codeposited, and finally Ti alone is deposited. Varying the evaporation conditions (rate), as well as the weight and composition of the charge in the evaporator, one can attain a desired distribution of the components and thickness of the black layer. During the deposition, the thickness of the layer was monitored by a KIT-1 quartz meter. After the deposition, it was measured with an MII-4 microinterferometer. The thickness of the layers studied was between 0.3 and 0.6 μm. The depth profiling of the component atoms was accomplished with a VG Microlab 310-F Auger spectrometer. The reflection of the normally incident light was measured with a KSVU-23 spectrophotometer; the diffuse reflection at an angle of incidence of 45°, with a test bench (Samsung SDI Ltd).

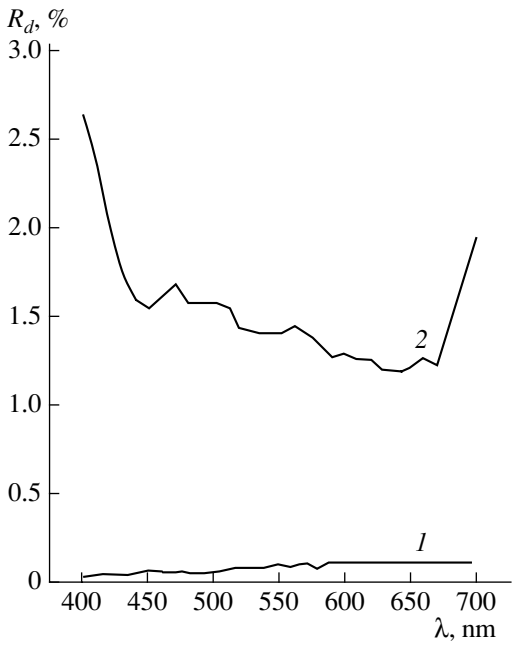


Fig. 7. Spectral dependence for the coefficient R_d of diffuse reflection from light-absorbing (*1*) SiO_x/Ti and (*2*) graphite coatings.

Figure 5 shows the distributions of Ti and Si across the 330-nm-thick SiO_x/Ti coating obtained. The thin

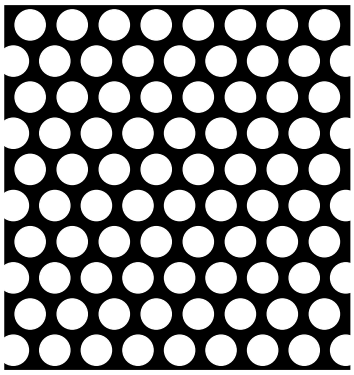


Fig. 8. Fragment of the SiO_x/Ti-based black matrix on the screen of the cathode-ray tube. The hole diameter is 110 μm.

layer near the coating–substrate interface consists almost totally of SiO_x: the concentration of Ti atoms is less than 10%. At a distance from the substrate, the Ti concentration grows and the outer layer of the coating (about 90 nm thick) consists almost totally of Ti. The variable-composition region is 240 nm thick, and, as follows from Fig. 5, the metal concentration in it varies with thickness nearly linearly. The optical density of the thickness is more than 4, and its sheet resistance measured by the four-point probe method is close to 30 Ω/□.

Curve *1* in Fig. 6 shows the spectral dependence of the specular reflection coefficient R_{sp} of the coating (conditions are the same as in Fig. 5). The reflection was measured on the side of the substrate. In such a measuring scheme, the reflection from the front surface of the substrate R_s (curve *2*) is imposed on that from the coating–substrate interface R. The value of $R = R_{sp} - R_s$ is small (≈0.01) throughout the visible range; that is, the coating is achromatic and effectively absorbs the incident radiation.

Figure 7 demonstrates the spectral dependence of the diffusion reflection R_d for the coating applied on the cathode-ray tube panel (curve *1*). For comparison, the diffusion reflection for a standard graphite coating is also plotted (curve *2*). It is seen that R_d for the black insulator–metal coating is one order of magnitude less throughout the visible range and does not change after thermal heating at 450°C for 1 h, which is used in the fabrication of cathode-ray tubes.

DISCUSSION

From Auger spectra recorded after the variable-composition coating had been argon-etched, we found the distribution of the metal atoms across the depth. In these calculations, the distribution of the optical constants n and k over the cross section should be used. To derive an accurate relationship between the metal concentration and the optical constants, it is necessary to know the chemical composition and structure of the layer. Upon the evaporation, Ti may combine with silicon to form silicide nanoinclusions, which change the composition of the matrix or enter into the matrix as metallic inclusions. The fraction of Ti agglomerated increases with its concentration, as demonstrated by measurements of the SiO/Me composite resistivity as a function of the metal concentration [9].

Let us assume that a gradient SiO_x/Ti layer represents, in the first approximation, a two-phase amorphous SiO_x/metallic Ti structure in any cross section. Such an approach was adopted in [10] to describe the optical properties of SiO_2/Ag(Au) composites over a wide range of concentrations (10–90%). As in [10], the effective optical constants of the SiO_x/Ti composite will be analyzed in terms of the Maxwell–Garnett effective-medium theory. Since neither the sizes nor the shapes of the inclusions in the two-phase structure are known (moreover, they may be concentration-dependent), we will use the Maxwell–Garnett formula for spherical inclusions and assume that the optical constants of both phases do not depend on the inclusion size. Thus, knowing the optical constants for SiO_x and Ti and taking into account that the component atoms are distributed in the two-phase structure linearly (Fig. 5), one can calculate the distribution of the optical constants of the SiO_x/Ti structure over the cross section of the layer deposited. Substituting the functions $\hat{n}_j(z)$ into formulas (1)–(4) yields the distribution of the specular reflection coefficient (curve *3* in Fig. 6). Unlike Figs. 2 and 3, here the reflection from the front boundary of the substrate is taken into account. In the visible range, the calculated and measured specular reflection coefficients of the glass substrate are seen to agree well; hence, the assumptions made are valid.

Usually, black matrices based on graphite coatings are fabricated only with lift-off photolithography, because chemical etchants for graphite are absent. The insulator–metal black matrices can be fabricated by the direct photolithography method. This method was employed in this work to prepare the SiO_x/Ti black matrices. The basic process steps in their production are the following.

(1) Application of a SiO_x/Ti black coating on the inner surface of a 17-inch cathode-ray tube panel by thermal evaporation *in vacuo*. This was made on a standard vacuum station for panel aluminizing. The uniformity of the coating thickness was ±7%.

(2) The application of an AZ-HKT-501 positive photoresist layer by spinning. The rate of rotation of the tube panel was 246 rpm.

(3) Photoresist drying, exposure, and development.

(4) Curing of the protective mask and etching of the black layer.

(5) Removal of the protective mask.

A fragment of the SiO_x/Ti black gradient matrix is shown in Fig. 8. It has been found that these matrices

offer a number of advantages over graphite ones: a low coefficient of diffuse reflection (one order of magnitude lower than that of a graphite matrix), excellent physicochemical properties (e.g., high mechanical strength and high adhesion) and their stability against thermal and chemical treatments, extremely low hygroscopicity and the absence of gas release, and high reproducibility of the coating properties.

The use of direct photolithography makes it possible to apply our process for fabricating black matrices intended for flat-panel displays (for example, for LCD filters).

REFERENCES

1. W. H. Southwell, US Patent No. 4934788 (1990).
2. I. Hiroshi, US Patent No. 5976639 (1999).
3. P. E. Shepelyavyi, E. V. Mikhailovskaya, I. Z. Indutnyy, *et al.*, Opt. Tekh., No. 2 (6), 16 (1995).
4. V. Ya. Bratus', V. A. Yukhimchuk, L. I. Berezhinskiĭ, *et al.*, Fiz. Tekh. Poluprovodn. (St. Petersburg) **35**, 854 (2001) [Semiconductors **35**, 821 (2001)].
5. M. Born and É. Wolf, *Principles of Optics* (Pergamon, Oxford, 1969; Nauka, Moscow 1970).
6. R. Jacobson, in *Physics of Thin Films: Advances in Research and Development*, Ed. by G. Haas, M. H. Francombe, and R. W. Hoffman (Academic, New York, 1975; Mir, Moscow, 1978), Vol. 8.
7. P. B. Johnson and R. W. Christy, Phys. Rev. B **9**, 5056 (1974).
8. I. Z. Indutny and A. I. Stetsun, Proc. SPIE **2113**, 55 (1993).
9. C. W. Park, J. B. Lee, Y. R. Do, *et al.*, Semicond. Phys., Quantum Electron. Optoelectron. **3**, 496 (2000).
10. R. W. Cohen, G. D. Cody, M. D. Coutts, *et al.*, Phys. Rev. B **8**, 3689 (1973).

Translated by V. Isaakyan

Technical Physics, Vol. 47, No. 6, 2002, pp. 726–730. Translated from Zhurnal Tekhnicheskoĭ Fiziki, Vol. 72, No. 5, 2002, pp. 73–78.

**OPTICS,
QUANTUM ELECTRONICS**

Multiskan-Based Double Synthetic Aperture for Locating the Illuminance Boundary of a Weak Optical Signal

B. G. Podlaskin, E. G. Guk, and E. V. Nosenko

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

Received November 14, 2001

Abstract—A model of double synthetic aperture based on Multiskan photodetectors is suggested. The model aperture makes it possible to locate and track the illuminance boundary of a weak optical signal in real time. With such an aperture, the boundary position can be found by integrally measuring the median of the photocurrent distribution for the modified signals. © *2002 MAIK "Nauka/Interperiodica".*

The problem of locating the faint optical boundary between two media is central to the practice of optical data processing. Difficulties arise, for example, in locating petroleum spots on the sea surface [1, 2], as well as the sea shelf and ocean currents [3]. This problem is also of great importance in the area of environmental protection, where the concentration of various gases is measured with gas-analyzing indicator tubes [4]. Here, the location of the optical boundary is difficult because of high spatial inhomogeneities.

In practice, the optical boundary (front) is located by differential methods: from the position of the peak of the signal first derivative, from the position of the point where the signal intersects a (fairly arbitrarily) preset level, or from the position of the zero of the second-order derivative. All these methods suffer from the same disadvantage: the need for differentiation with respect to time. When both the first and the second derivatives of the signal are used, the position of the front depends on its shape and, especially, on the noise level. If the position is found from the point of intersection, the location accuracy depends on the stability of this level, variation of the constant component of the optical distribution, contrast ratio, and the width and shape of the transition region. That is why the location accuracy for the optical boundary of weak signals in the presence of noise (i.e., under real conditions) is low, within 5–50% [1, 2].

Refractometry methods, where four-quadrant photodiodes are used as position detectors, provide the highest accuracy [5]. However, in this case, too, the degree of discrimination of two media with close refractive indices is insufficient.

In a number of works, it has been shown [6, 7] that optical distribution parameters can be found with a high accuracy by using integral methods of image processing. In this case, the coordinates of image informative signs are integrally estimated from the positions of the optical signal medians and by calculating sets of quantiles or integral moments of the signals. According to published data [8], the distribution median adequately estimates the position of spatially localized optical signals. An advantage of the integral methods is the full use of the input energy and noise integration over narrower space and time domains, because there is no need for scanning the entire range of optical signal distribution. For example, the integral estimation of the position of a light source locates its median with an accuracy of 0.02–0.002% of the field of vision [7].

The transition from scanning methods to integral functionals implies the use of photodetectors that can implement the integral method of signal conversion throughout the interval where the signal distribution is defined. Today, an integrated coordinate Multiskan photodetector [9] is the only device satisfying these requirements.

A Multiskan photodetector is a silicon structure consisting of an array of back-to-back-connected photodiodes. The diodes are under a linearly distributed potential when a dc voltage is applied to a resistive divider. The operation of a Multiskan as a variable-aperture device that forms an integral signal takes advantage of its integral aperture characteristic. The position U_0 of its inflection point depends on the potential of the resistive divider and the output potential in such a way that the photocurrent from the Multiskan output is given by [9]

$$I = \alpha\int_0^L f(x)\tanh\frac{x-\xi_0}{A_x}dx$$

or, in terms of the applied voltage,

$$I = \beta\int_{-E}^{E} f(u)\tanh\frac{u-U_0}{A_u}du. \qquad (1)$$

1063-7842/02/4706-0726$22.00

Here, L is the length of the Multiskan photosensitive layer; ξ_0, spatial position of the inflection point in the aperture characteristic; E, voltage applied to the resistive layer; A_x, aperture spatial width; A_u, aperture width in terms of voltage; α and β, proportionality coefficients (hereafter, $\alpha = \beta = 1$); and $f(x)$ and $f(u)$, optical signal distribution in terms of coordinates x and u, respectively.

For an optical signal, the determination of the illuminance front with the integral method poses fundamental difficulties, unlike the case of locating a localized optical spot. When the front is faint, one usually deals with optical distributions having a significant density on both sides of a front to be located. In this case, the position of the distribution median is virtually independent of the optical front coordinate.

The aim of this work is to obviate this difficulty. The problem of finding the position of the weak signal front is reduced to the measurement of the median of the modified signal. The modification is made by a set of photodetectors having a double synthetic aperture.

Two-aperture synthesis is a signal processing method whereby transformations equivalent to the addition or subtraction of optical signals (with assigning given weighting factors to them and introducing a necessary spatial shift of the optical distributions into these operations) are performed with one or several photodetectors. Such modifications sometimes make it possible to find the coordinates of the informative signs of the initial distributions from the integral characteristics of the signals modified.

Previously, the method of double synthetic aperture was aimed at improving the location accuracy of Multiskans in the presence of intense background illumination [10]. However, the problem of locating the optical front has the specific features noted above, which require a radically different approach to be employed. In this case, it is necessary to transform the initial distribution so that the position of the median of the new distribution is directly related to the position of the front. Clearly, the desired integral procedure must be capable of revealing this differential sign without scanning the interval where the illuminance function is defined and without subsequent signal filtering and smoothing.

The integral methods of signal formation allow the location of informative signs in the tracking mode; that is, the entire time of observation t' is completely spent on the valid signal formation. It is in this sense that the term "real-time operation" is applied to these methods.

For the system to operate in the tracking (real-time) mode, it is necessary that an informative sign (optical front coordinate in our case) be found not by scanning the entire realization. Instead, it must be formed as a continuous current value U^0_{control}, which is automatically set up (within the interval $\pm E$) as a potential corresponding to the position of the front. For this purpose, one needs to relate the difference current I and the coordinates $(U_0 + \varepsilon)$ and $(U_0 - \varepsilon)$ of the inflection points in the Multiskan aperture characteristics. We will find the control voltage U_{control} by integrating the Multiskan instantaneous output current over time:

$$U_{\text{control}} = \frac{1}{C}\int_0^{t'} I_\Sigma dt,$$

where C is the capacitance of the integrator.

It is obvious that U_{control} reaches the steady-state value at $I_\Sigma = 0$. By definition, this corresponds to the position of the optical distribution median. Thus, the value of U^0_{control} can be found from the condition $I_\Sigma = 0$.

This condition is the universal condition for finding the optical distribution median when the photodetectors operate in real time. Since I_Σ is the sum of the currents I_1 and I_2 of two Multiskans, the value and the variation of I_Σ depend on the law by which $f(u)$ transforms into I_1 and I_2. Namely, they depend on the relative shift of the potentials across the Multiskan resistive dividers, positions of the inflection points in the aperture characteristics, transconductances of the I–V characteristics, current gains, and other parameters.

To track the illuminance front position with Multiskans, we adopt the following model of formation of the currents I_1 and I_2. Let two Multiskans (M_1 and M_2) be placed in parallel along the mutual voltage axis u and be shifted by Δx so that their limits of integration are shifted relative to each other by some value $\pm\varepsilon$. The relative positions of the inflection points in the aperture characteristics are defined by the potentials $+\varepsilon$ and $-\varepsilon$ applied to the mutual buses of M_1 and M_2 ($U_{\text{control}} = 0$).

Then, the currents $I_1(U_{\text{control}})$ and $I_2(U_{\text{control}})$ are given by

$$\begin{aligned} I_1 &= \int_{-E+\varepsilon}^{E+\varepsilon} f(u)\tanh\frac{u - U_{\text{control}} - \varepsilon}{A_u}du, \\ I_2 &= \int_{-E-\varepsilon}^{E-\varepsilon} f(u)\tanh\frac{u - U_{\text{control}} + \varepsilon}{A_u}du \end{aligned} \tag{2}$$

and the equation for current balance has the form

$$\int_{-E+\varepsilon}^{E+\varepsilon} f(u)\tanh\frac{u - U_{\text{control}} - \varepsilon}{A_u}du - \int_{-E-\varepsilon}^{E-\varepsilon} f(u)\tanh\frac{u - U_{\text{control}} + \varepsilon}{A_u}du = 0. \tag{3}$$

Figure 1 depicts a connection diagram realizing a double synthetic aperture where a spatial shift of Multiskans causes a shift in the supply voltages.

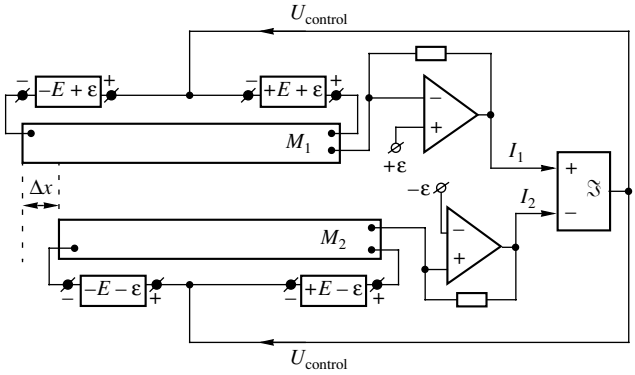


Fig. 1. Connection diagram for Multiskan photodetectors operating in the synthetic aperture mode for locating the illuminance front.

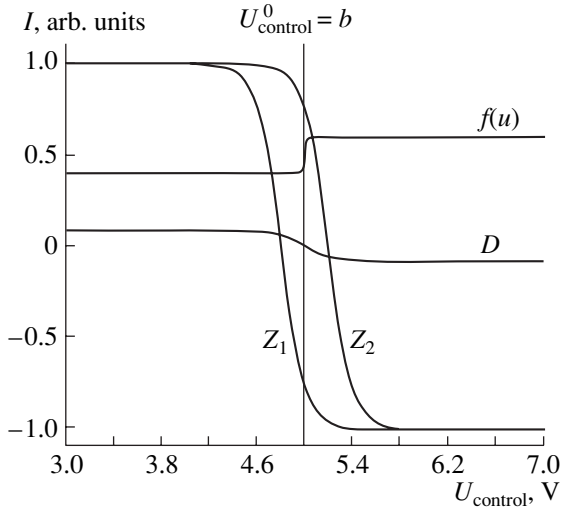


Fig. 2. Results of analysis of the current balance equation. Z_1 and Z_2, Multiskan aperture characteristics; $f(u)$, function simulating the input distribution; and D, distribution of the difference current $I_1 - I_2$ over the illuminance front region.

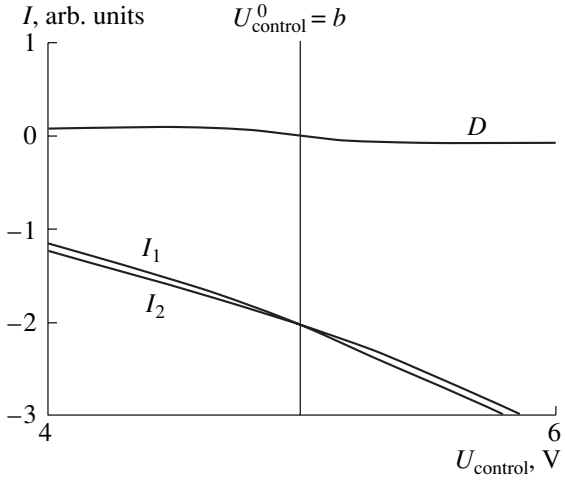


Fig. 3. Behavior of the currents I_1 and I_2, as well as of their difference D, in the vicinity of the illuminance front.

When the Multiskans are identically illuminated by the optical signal $f(u)$, the value of $U_{control}$ starts to vary because of the difference photocurrent until the positions of the aperture characteristics relative to the input signal become such that the total photocurrent vanishes.

Since such two-aperture synthesis is intended for locating the optical signal front, we define $f(u)$ as

$$f(u) = p + q\tanh\frac{u-b}{B}. \tag{4}$$

Here, the parameter b specifies the position of the front in the interval $\pm E$ and the parameter B specifies the steepness of the front. The parameters q and p characterize the illuminance drop and the constant component of the optical signal, respectively. Then, Eq. (3) takes the form

$$\int_{-E+\varepsilon}^{E+\varepsilon}\left[p + q\tanh\frac{u-b}{B}\right]\tanh\frac{u-U_{control}-\varepsilon}{A_u}du - \int_{-E-\varepsilon}^{E-\varepsilon}\left[p + q\tanh\frac{u-b}{B}\right]\tanh\frac{u-U_{control}+\varepsilon}{A_u}du = 0. \tag{5}$$

We will solve this equation for $U_{control}$, setting the front position parameter $b = 5$ for definiteness. The other parameters are as follows: voltages applied to the resistive dividers $E = \pm 10$ V, shift $\varepsilon = 0.2$ V, width of the transition region of the Multiskan aperture characteristic $A_u = 0.2$ V, optical front width $B = 0.02$ V, $q = 0.1$, and p (the parameter including the constant component and optical distribution contrast) = 0.5.

Figure 2 reflects the results of analysis of Eq. (5) for current balance. Shown are the function $f(u)$, which simulates the initial optical distribution for $b = 5$, and the aperture characteristics Z_1 and Z_2 of the Multiskans M_1 and M_2, respectively, which are shifted by a value $u = 2\varepsilon$ relative to each other. The curve $D(U_{control}) = I_1(U_{control}) - I_2(U_{control})$ demonstrates that the value $D(U^0_{control}) = 0$, which is the solution to Eq. (5), is at the point $U^0_{control} = b$, which corresponds to the specified position of the front within the given calculation accuracy.

The behavior of each of the curves I_1 and I_2 near their intersection is shown in Fig. 3. The currents are seen to have a single point of intersection; that is, the solution found is unique. Let us consider the behavior of the difference curve D and of the integral currents I_1 and I_2 (Figs. 3, 4) in order to physically substantiate the coincidence of the solution to Eq. (5) with the position b of the optical front. In the regions where the curve $D(U_{control})$ saturates (on the right and on the left of the inflection point), the difference $|I_1 - I_2|$ equals the sum of the currents in the end portions shifted relative to each other in the Multiskan space, i.e., in the portions

from $(-E - \varepsilon)$ to $(-E + \varepsilon)$ and from $(+E - \varepsilon)$ to $(+E + \varepsilon)$. This saturation current will be denoted by ΔI. The constancy of ΔI both on the right and on the left of the front position indicates that the photocurrents from the Multiskans cancel each other in the interval from $(-E + \varepsilon)$ to $(+E - \varepsilon)$, where the inflection points of the aperture characteristics Z_1 and Z_2 are beyond the front region. The presence of ΔI causes U_{control} to vary and thus Z_1 and Z_2 to move along the u axis. When one of the characteristics reaches the front position coordinate, an extra photocurrent directed oppositely to ΔI arises due to the difference in the illuminance on the right and on the left of the front. As the characteristics move further and become symmetric about the front coordinate, ΔI vanishes, the process of charge accumulation on the integrator $\Im$ is completed, and a value of U^0_{control} corresponding to the front position forms.

Consider in greater detail the derivative $dD(U_{\text{control}})/dU_{\text{control}}$, which clarifies how the variation of the integral difference $I_1 - I_2$ provides the passage of the curve D through the zero. It is this derivative that allows us to find the photocurrent distribution whose median specifies the front position. Also, this derivative illustrates how the optical front position (the differential sign of the initial distribution that is the most stable against variation of the front width and shape, as well as against space–time noise [8]), can be found by using the integral procedure of median finding.

Figure 4 shows a family of curves $dD(U_{\text{control}})/dU_{\text{control}}$ for various front widths B. As B grows, the peak height decreases and the curve broadens. However, the broadening takes place much more slowly than the front width B increases. In the interval $B = 0.2$–2.0 V (the tenfold extension of the front), the effective value of the difference aperture increases less than twofold. This is associated with the properties of the Multiskan aperture characteristics, which are close in shape to hyperbolic tangents, whose difference rapidly approaches zero.

The shape of the curve $dD(U_{\text{control}})/dU_{\text{control}}$ is severely affected by the shift ε and the steepness A_u of the aperture characteristics. Both parameters define the size of the spatial "window" produced by the Multiskan aperture characteristics shifted relative to each other. Within this window, the difference photocurrent is integrated and the integral procedure of locating the optical front is performed.

Since the aperture characteristics are rather smooth (the transition region width at a level of 0.8 of the saturation current is 0.4 V), a large location error may arise if b is close to the interval extremes $+E$ and $-E$.

Figure 5 shows the relative error $\Delta = (U^0_{\text{control}} - b)/b$ throughout the interval ±10 V for the signal front width $B = 0.2$ V. In the absence of noise, the location error is within the given calculation accuracy over the entire operating interval. At a distance of 0.5% from the device edge, the accuracy improves to 1%. As the front

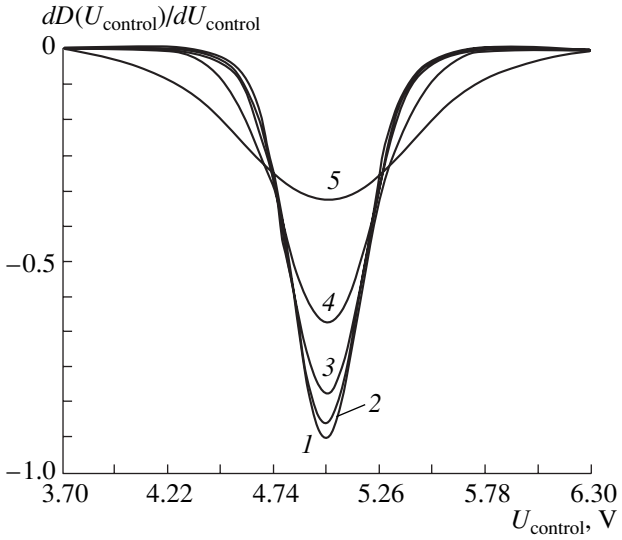


Fig. 4. Run of the derivative $dD(U_{\text{control}})/dU_{\text{control}}$ for the front width B (*1*) 0.02, (*2*) 0.05, (*3*) 0.1, (*4*) 0.2, and (*5*) 0.5 V.

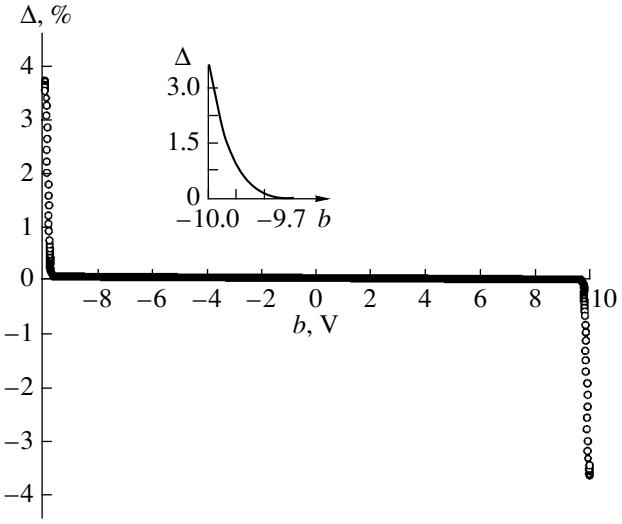


Fig. 5. Relative error $\Delta = (U^0_{\text{control}} - b)/b$%, in the interval ±10 V for the front width $B = 0.2$ V. The insert shows the relative error at the boundary of the operating range.

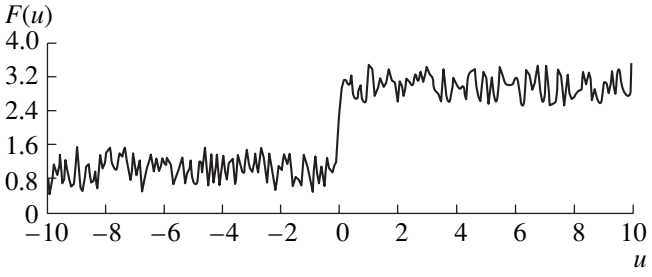


Fig. 6. Realization of the noisy signal $F(u)$.

Table

N	σ
1	1.997×10^{-3}
2	1.239×10^{-3}
10	1.934×10^{-4}

width grows, the error at the limits of the interval increases and the region where the error falls into the specified range shrinks.

The region of accurate calculation can be extended by decreasing the parameter ε. However, such a decrease and thereby a decrease in the geometrical shift Δx of the devices reduce the difference current $D(U_{\text{control}})$. In this case, the time of front location grows and the location accuracy worsens.

A necessary estimate of the model suggested is the stability of the solution against spatial inhomogeneities and noise. For this purpose, we used a computer model that imposes a random noise-simulating process on the signal $f(u)$. Our goal was to obtain a basic estimate of the noise effect on the accuracy of the method and compare the resulting accuracy with that of the conventional time scanning methods without considering the noise spectrum.

The model of the noisy signal $F(u)$ is depicted in Fig. 6. The signal-to-noise ratio was estimated from the relationship between the illuminance drop and the width of the noise track N. In model calculations, the latter was set equal to 1, 2, and 10. For all the values, the rms error σ was calculated based on 100 samples (see the table).

For comparison with the conventional methods, we calculated this error for $N = 1$ when finding the front coordinate from the peak position of the first derivative of $F(u)$. The value of σ obtained was equal to 8×10^{-2}, or 40 times as large as that in our integral method.

CONCLUSION

The use of a double synthetic aperture for locating the illuminance front of a weak optical signal by the integral method allows the following conclusions.

A combination of an electrical bias and a relative spatial shift of Multiskan photodetectors makes it possible to elaborate a technique for locating a signal differential sign in real time (tracking mode) without scanning the interval where the initial function is defined.

This technique provides a unique solution which corresponds to the illuminance front position. This solution does not depend on the front width and optical distribution contrast. It offers a high accuracy over a substantial area of the photosensitive layer of the photodetectors.

The technique reduces the effect of noise on the location accuracy by integrating noise over the interval 2ε. For a signal-to-noise ratio of 1, the rms error is 2×10^{-3}.

Thus, this work discovers the high operational functionality of a Multiskan-based synthetic aperture. Namely, it cannot only locate local signals but also track the current position of the illuminance front of optical signals.

REFERENCES

1. I. V. Aleshin, É. G. Goncharov, A. Z. Zurabyan, *et al.*, Opt. Zh. **65** (12), 132 (1998) [J. Opt. Technol. **65**, 1061 (1998)].
2. A. Z. Zurabyan, Opt. Zh. **65** (11), 67 (1998) [J. Opt. Technol. **65**, 924 (1998)].
3. G. Rockwell and G. C. Kineke, J. Geophys. Res. **100**, 2321 (1995).
4. B. G. Podlaskin, V. F. Tul'vert, and E. A. Chekulaev, Prib. Tekh. Éksp., No. 4, 169 (1994).
5. I. M. Belousova, A. V. Bukanov, S. N. Leonov, and A. G. Skepko, Opt.-Mekh. Prom-st, No. 12, 15 (1990).
6. B. G. Podlaskin, L. Z. Dich, and N. A. Tokranova, Pis'ma Zh. Tekh. Fiz. **20** (2), 30 (1994) [Tech. Phys. Lett. **20**, 60 (1994)].
7. B. G. Podlaskin and N. A. Tokranova, in *Proceedings of the 4th International Conference on Ultraprecision in Manufacturing Engineering on Progress in Precision Engineering and Nanotechnology, 1997*, Vol. 1, p. 140.
8. A. N. Shiryaev, *Probability* (Nauka, Moscow, 1989).
9. K. F. Berkovskaya, N. V. Kirillova, B. G. Podlaskin, *et al.*, Zh. Tekh. Fiz. **53**, 2015 (1983) [Sov. Phys. Tech. Phys. **28**, 1237 (1983)].
10. B. G. Podlaskin, A. V. Vasil'ev, E. G. Guk, and N. A. Tokranova, Zh. Tekh. Fiz. **70** (10), 110 (2000) [Tech. Phys. **45**, 1339 (2000)].

Translated by V. Isaakyan

Technical Physics, Vol. 47, No. 6, 2002, pp. 731–736. Translated from Zhurnal Tekhnicheskoĭ Fiziki, Vol. 72, No. 5, 2002, pp. 79–84.

RADIOPHYSICS

Anisotropic Disk Dielectric Resonator with Conducting End Faces

Yu. V. Prokopenko and Yu. F. Filipov

Usikov Institute of Radiophysics and Electronics, National Academy of Sciences of Ukraine, ul. Akad. Proskura 12, Kharkov, 61085 Ukraine

e-mail: prokopen@ire.kharkov.ua

Received July 16, 2001; in final form, November 29, 2001

Abstract—Electrodynamic analysis of a disk dielectric resonator with finite-conductivity end faces that is excited at the whispering-gallery mode is carried out. The inverse problem of determining the components of the permittivity tensor of the anisotropic uniaxial crystal, loss tangent of the dielectric material, and surface resistance of the conductor is formulated and solved. The frequency spectra and energy characteristics of a Teflon disk resonator with copper-covered end faces are studied in the 8-mm wave range. © *2002 MAIK "Nauka/Interperiodica".*

INTRODUCTION

The advance into the millimeter-wave range suggests the development of appropriate microwave devices. The new-generation compact devices are being developed around disk dielectric resonators (DDRs), which are well compatible with passive and active elements of hybrid circuits and offer a high Q factor (significantly higher than that of metallic cavity resonators). DDRs operating at lower oscillation modes are used in stable centimeter-wave oscillators and filters. Such resonators are made of high-permittivity ($\varepsilon \geq 100$) insulators. It has been shown [1, 2] that low-permittivity ($\varepsilon = 2$–10) cylindrical dielectric rods can support slowly decaying azimuthal oscillations. These oscillations are generated inside the dielectric when waves with high azimuthal indices n are incident on the side surface of a DDR (with the radius of curvature r_0) at the angle of total internal reflection. The presence of the caustic and an outside region where the field decays exponentially provide a high Q value of the resonator.

Anomalously low attenuation was first observed in acoustic waves traveling along the inner surface of an annular gallery. That is why azimuthal oscillations that occur when the condition $k'r_0 < n < \sqrt{\varepsilon}\,k'r_0$ is met were called whispering-gallery oscillations [3]. Here, $k' = \omega'/c$, c is the velocity of light, and ω' is the real part of the complex frequency $\omega = \omega' - i\omega''$ (where $\omega'' \geq 0$).

Whispering-gallery oscillations are very sensitive to the surface condition of films and laminated structures, which are today the subject of extensive research in solid-state physics. This is because the interface quality is crucial for the operation of many physical devices.

Anisotropic materials with a small (no greater than 5×10^{-5}) loss tangent $\tan\delta$, such as uniaxial single-crystalline quartz, sapphire, and rutile, have made it possible to increase significantly the unloaded Q value of DDRs. In particular, a Q factor of sapphire resonators as high as ~10^5 at room temperature, ~10^7 at 77 K, and ~10^9 at 4 K [4] has been obtained. The unique set of parameters, high polarization stability, and ease of excitation of DDRs made of these single crystals have stimulated applied research in the millimeter-wave range. A number of devices that effectively stabilize the oscillator frequency [5], as well as new-generation filters [6, 7], have been developed. Such resonators are also used in studying high-frequency electron paramagnetic resonance [8] and in electron spin resonance spectroscopy [9].

In contactless measurements of insulator parameters [10, 11] and those of single-crystal high-temperature superconductors [12], a sample studied is placed into the electromagnetic field of a resonator. A significant drawback here is that a rigorous electrodynamic approach to the analysis of experimental results is absent. The use of samples with perfectly conducting end faces makes it possible to study the spectral characteristics of resonator free oscillations [13].

DISK DIELECTRIC RESONATOR

In a DDR made of an anisotropic uniaxial single crystal with its anisotropy axis parallel to the DDR axis, the permittivity tensor has the form

$$[\varepsilon_{ij}] = \begin{cases} \delta_{ij} & r > r_0 \\ \begin{pmatrix} \varepsilon_\perp & 0 & 0 \\ 0 & \varepsilon_\perp & 0 \\ 0 & 0 & \varepsilon_z \end{pmatrix}, & r \leq r_0, \end{cases}$$

1063-7842/02/4706-0731$22.00

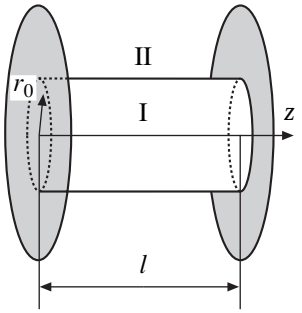


Fig. 1. DDR with conducting end faces.

where ε_z and $\varepsilon_\perp$ are, respectively, the components of $[\varepsilon_{ij}]$ in the directions parallel and perpendicular to the crystal optical axis and δ_{ij} is the Kronecker delta.

The fields of monochromatic oscillations in these resonators are given by [13]

$$E_z = D_n G_E(r)\cos(k_z z)\exp i(n\varphi - \omega t);$$

$$H_z = C_n G_H(r)\sin(k_z z)\exp i(n\varphi - \omega t);$$

$$q^2 E_\varphi = \frac{1}{r}\frac{\partial E_z}{\partial\varphi\partial z} - ik\frac{\partial H_z}{\partial r}; \quad q^2 E_r = \frac{\partial^2 E_z}{\partial r\partial z} + i\frac{k}{r}\frac{\partial H_z}{\partial\varphi};$$

$$q^2 H_\varphi = ik\varepsilon_\perp\frac{\partial}{\partial r}E_z + \frac{1}{r}\frac{\partial^2 H_z}{\partial\varphi\partial z};$$

$$q^2 H_r = \frac{\partial^2 H_z}{\partial r\partial z} - i\frac{k\varepsilon_\perp}{r}\frac{\partial E_z}{\partial\varphi}.$$

Here,

$$G_j = \begin{cases} J_n(q_j r), & r \le r_0 \\ \dfrac{J_n(q_j r_0)}{H_n^{(1)}(q_0 r_0)} H_n^{(1)}(q_0 r), & r \ge r_0 \end{cases};$$

the subscript j stands for H or E; $k = \omega/c$; the azimuthal subscript n equals 0, 1, 2, …; k_z is the axial component of the wave vector; $q_H = \sqrt{\varepsilon_\perp k^2 - k_z^2}$; $q_E = \sqrt{\varepsilon_z/\varepsilon_\perp}\,q_H$ and $q_0 = \sqrt{k^2 - k_z^2}$ are the radial components of the wave vector inside and outside the insulator, respectively; C_n and D_n are constant factors; and $q^2 = q_H^2$ when $r \le r_0$ and $q^2 = q_0^2$ when $r \ge r_0$.

The frequency $\omega_p = \omega_p' - i\omega_p''$ ($\omega_p'' \ge 0$) for the pth mode of the resonant oscillation is found from a solution to the characteristic equation

$$(\alpha_H - \alpha_0)(\varepsilon_z\alpha_E - \alpha_0) = a^2, \qquad (1)$$

where

$$\alpha_j = \frac{1}{q_j r_0}\frac{J_n'(q_j r_0)}{J_n(q_j r_0)}, \quad \alpha_0 = \frac{1}{q_0 r_0}\frac{H_n^{(1)'}(q_0 r_0)}{H_n^{(1)}(q_0 r_0)},$$

$$a = \frac{kk_z n}{q_0^2 q_H^2 r_0^2}(1 - \varepsilon_\perp),$$

prime denotes differentiation with respect to the argument, and $J_n(x)$ and $H_n^{(1)}(x)$ are the nth-order Bessel and Hankel functions of the zero kind. The subscript p equals nsm, where n, s, and m are the azimuthal, radial, and axial subscripts, respectively.

For resonators with perfectly conducting end faces (Fig. 1), the axial component of the wave vector is $k_z = m\pi/l$, where l is the height of the resonator and $m = 0, 1, 2, \ldots$.

A DDR supports independent *EH* and *HE* oscillations. When the condition

$$\left|(\alpha_H - \alpha_0)^{1/2}(\varepsilon_z\alpha_E - \alpha_0)^{-1/2}\right| \gg 1$$

is met, the free oscillations of the resonator are of the *HE* type, the constant D_n is determined from the excitation condition, and

$$C_n = -D_n\frac{J_n(q_E r_0)}{J_n(q_H r_0)}\frac{a}{\alpha_H - \alpha_0}.$$

Otherwise, the resonator mode is *EH*, the constant C_n is determined from the excitation condition, and

$$D_n = -C_n\frac{J_n(q_H r_0)}{J_n(q_E r_0)}\frac{a}{\varepsilon_z\alpha_E - \alpha_0}.$$

If the axis of the uniaxial crystal makes an angle Θ with the geometrical axis of the resonator, the permittivity tensor takes the following form in the cylindrical coordinates (r, φ, z):

$$[\varepsilon_{ij}] = \begin{pmatrix} \varepsilon_+ + \varepsilon_-\cos 2\varphi & -\varepsilon_-\sin 2\varphi & \varepsilon_{xz}\cos\varphi \\ -\varepsilon_-\sin 2\varphi & \varepsilon_+ - \varepsilon_-\cos 2\varphi & -\varepsilon_{xz}\sin\varphi \\ \varepsilon_{xz}\cos\varphi & -\varepsilon_{xz}\sin\varphi & \varepsilon_{zz} \end{pmatrix},$$

where $2\varepsilon_\pm = \varepsilon_z\sin^2\Theta + \varepsilon_\perp(\cos^2\Theta \pm 1)$, $2\varepsilon_{xz} = (\varepsilon_z - \varepsilon_\perp)\sin 2\Theta$, and $\varepsilon_{zz} = \varepsilon_z\cos^2\Theta + \varepsilon_\perp\sin^2\Theta$.

In the cylindrical coordinates, the permittivity tensor components depend on the azimuth angle φ. This dependence splits the *EH* and *HE* oscillations into the ordinary and extraordinary modes, as follows from experiments [14]. The spectral characteristics of these modes provide information on the permittivity components along the principle lattice directions, which is necessary for designing high-performance millimeter-wave and optical integrated circuits.

For the case when the end faces are not perfectly conducting planes, a rigorous theory does not exist. We

calculated the resonance frequencies with the magnetic-wall method, variational methods, and mode-matching method by appropriately approximating the electromagnetic fields outside the resonator with allowance for their singularities at the insulator edges [15]. In [16, 17], experimental identification of the resonance frequencies with the azimuthal and radial indices was suggested. In this case, the radial and azimuthal components of the wave vector are described by the characteristic equation presented above and the effective height of the resonator is defined as $l_{\text{eff}} = m\pi/k_z$.

DDR WITH IMPEDANCE END FACES

The quadratic relationship obtained from the Maxwell equations for resonant oscillations in resonators with perfectly and imperfectly conducting end faces is integrated over the entire space [18]. Using the radiation condition at $r \longrightarrow \infty$ (outgoing wave), the continuity of the tangential electric and magnetic field components at the curved surface of the resonator, and the impedance boundary condition at the end faces of the resonator, we arrive at an integral equation that defines a shift in the resonance frequency due to imperfect conductivity of the DDR end faces (Fig. 1) [18]:

$$(\omega - \omega_p^*)\int_V [(\mathbf{H}\mathbf{H}_p^*) + (\mathbf{E}[\varepsilon_{ij}]\mathbf{E}_p^*)]dV$$

$$= i\frac{c^2}{2\pi}\zeta\int_S ([\mathbf{e}_x \times [\mathbf{e}_z \times \mathbf{H}]]\mathbf{H}_p^*)dS,$$

where $\mathbf{e}_z$ is the unit vector along the DDR axis; ω, $\mathbf{E}$, and $\mathbf{H}$ are the frequency and the electric and magnetic field vectors for the resonator with imperfectly conducting end faces; and ω_p, $\mathbf{E}_p$, and $\mathbf{H}_p$ are the frequency and the electric and magnetic field vectors for the pth mode of the resonator with perfectly conducting end faces (asterisk means complex conjugation). Vectors in parentheses denote the scalar product. On the left-hand side of this formula, integration is performed over the entire volume V; on the right-hand side, over the surface S of the end faces having the surface impedance $\zeta = R_S + iX_S$.

Representing the vectors $\mathbf{E}$ and $\mathbf{H}$ as a linear combination of the respective self-fields $\mathbf{E}_p$ and $\mathbf{H}_p$ of the DDR with perfectly conducting end faces, we obtain the following system of equations:

$$(\omega - \omega_p^*)W_p\beta_p = -i\zeta\sum_p I_{p'p}^2\beta_{p'}.$$

Here, $\beta_{p'}$ is the expansion coefficient and

$$W_p = \frac{1}{8\pi}\int_V [(\mathbf{H}_p\mathbf{H}_p^*) + (\mathbf{E}_p[\varepsilon_{ij}]\mathbf{E}_p^*)]dV$$

is the energy of the pth electromagnetic mode in the resonator with perfectly conducting end faces. The parameter

$$I_{p'p}^2\zeta = \frac{c^2}{16\pi^2}\zeta\int_S ([\mathbf{e}_z \times [\mathbf{e}_z \times \mathbf{H}_{p'}]]\mathbf{H}_p^*)dS$$

takes into account losses in the imperfectly conducting end faces and describes the interaction between DDR modes. The consistency condition for the system of homogeneous linear equations for $\beta_{p'}$ yields a characteristic equation for the natural frequencies of the DDR with finite-conductivity end faces:

$$\det\left\|(\omega - \omega_p^*)\delta_{p'p} + i\zeta\frac{I_{p'p}^2}{W_p}\right\| = 0. \qquad (2)$$

This relationship defines the shift between the resonance frequencies of a DDR with imperfectly conducting end faces and that with perfectly conducting faces for a particular oscillation mode. The shift in the real part of the resonance frequency is specified by the imaginary part of the surface impedance (reactance) X_S. The reactance specifies the nondissipative energy stored in the surface layers of the end faces. The surface resistance R_S is responsible for the period-averaged Joule loss $I_{pp}^2 R_S$ in the impedance end faces of the resonator.

When the surface impedance ζ is small ($\zeta \ll 1$) and the mode interaction is neglected ($I_{p'p}^2 = 0$ if $p' \neq p$), relationship (2) becomes

$$(\omega - \omega_p^*) + i\zeta\frac{I_{pp}^2}{W_p} = 0, \qquad (3)$$

where I_{pp} is the surface current on the finite-conductivity end faces induced by the electromagnetic wave.

Relationship (3) can be used to calculate the surface impedance ζ of the end faces from the DDR resonance frequency ω measured.

DDR WITH LOCAL INHOMOGENEITIES

Radial–azimuthal inhomogeneities in the DDR isotropic material that are included in the permittivity $\tilde{\varepsilon}(r, \varphi)$ also change the spectral characteristics of the resonance oscillations. The shift in the resonance frequencies due to the inhomogeneities is described by the integral equation

$$(\omega - \omega_p)\int_V [(\mathbf{H}\mathbf{H}_p^*) + \varepsilon(r)(\mathbf{E}_p^*\mathbf{E})]dV$$

$$= \omega\int_V [\varepsilon(r) - \tilde{\varepsilon}(r, \varphi)](\mathbf{E}_p^*\mathbf{E})dV.$$

Inhomogeneities in the form of radial slots split the frequencies of degenerate resonance oscillations that depend on the azimuthal coordinate as $\sin n\varphi$ and $\cos n\varphi$, which is observed experimentally. The split is a function of the number and arrangement of the slots. This effect can be used for characterizing local inhomogeneities in the form of insulating, semiconductor, or high-temperature superconductor films at microwaves [19, 20].

Double disk resonators excited at whispering-gallery modes are used in ultrasensitive detectors [21]. At a certain spacing between two identical dielectric resonators, the frequency degeneration of symmetric and asymmetric oscillations with contiguous axial indices occurs. A weak inhomogeneity (e.g., a tilt of the end face) removes the degeneracy, producing two orthogonal oscillations with close frequencies modulated by orthogonal envelopes. This property can be used for spatially separation of two oscillations with close frequencies in the millimeter and submillimeter ranges [22].

Q FACTOR OF A DDR

The unloaded Q factor Q_0^Y of a DDR with impedance end faces for the Y_p mode is given by

$$\frac{1}{Q_0^Y} = \frac{1}{Q_\mathrm{d}^Y} + \frac{1}{Q_\mathrm{rad}^Y} + \frac{1}{Q_\mathrm{c}^Y}. \tag{4}$$

Here, Y means the *EH* or *HE* oscillation mode and Q_rad^Y, Q_d^Y, and Q_c^Y are the Q factors due to radiation, dielectric loss, and loss in the conductive end faces, respectively.

For whispering-gallery oscillations, which occur when $|q_0 r_0| \le n \le \min\{|q_\mathrm{H} r_0|, |q_E r_0|\}$, the relationship $\omega' \gg \omega''$ is valid. This corresponds to low radiation losses. The radiation Q factor of the resonator at the resonance frequency ω of the Y_p mode is $Q_\mathrm{rad}^Y = \omega'/2\omega''$. The Q factor due to the dielectric loss is given by [17]

$$Q_\mathrm{d}^Y = \frac{1}{\tan\delta}(1 + R_0^Y),$$

where, in the general case, $R_0^Y = W_\mathrm{II}^Y / W_\mathrm{I}^Y$ and W_I^Y and W_II^Y are the electromagnetic energies of the Y_p oscillation inside (I) and outside (II) the dielectric, respectively. In a DDR (irrespective of whether the end faces are conducting or not), this relationship takes the form

$$R_0^{EH} = -\frac{|J_n(q_H r_0)|^2}{|H_n^{(1)}(q_0 r_0)|^2} \times \frac{\left(\frac{k_z^2}{q_0^2} + 1 - \frac{\sin(k_z l)}{2k_z l}\right)\Phi_1 + \left(\frac{k_z^2}{q_0^2} + \frac{1}{2} - \frac{\sin(k_z l)}{4k_z l}\right)\Phi_2}{\left(\frac{k_z^2}{q_H^2} + 1 - \frac{\sin(k_z l)}{2k_z l}\right)F_1^H + \left(\frac{k_z^2}{q_H^2} + \frac{1}{2} - \frac{\sin(k_z l)}{4k_z l}\right)F_2^H}$$

for the *EH* modes and

$$R_0^{HE} = -\frac{1}{\varepsilon_z}\frac{|J_n(q_E r_0)|^2}{|H_n^{(1)}(q_0 r_0)|^2} \times \frac{\left(\frac{k_z^2}{q_0^2} + 1 + \frac{\sin(k_z l)}{2k_z l}\right)\Phi_1 + \left(\frac{k_z^2}{q_0^2} + \frac{1}{2} + \frac{\sin(k_z l)}{4k_z l}\right)\Phi_2}{\left(\frac{k_z^2}{q_H^2} + 1 + \frac{\sin(k_z l)}{2k_z l}\right)F_1^E + \left(\frac{k_z^2}{q_H^2} + \frac{1}{2} + \frac{\sin(k_z l)}{4k_z l}\right)F_2^E}$$

for the *HE* modes, where

$$F_1^j = \left(1 - \frac{n^2}{q_j^2 r_0^2}\right)|J_n(q_j r_0)|^2 + |J_n'(q_j r_0)|^2,$$

$$F_2^j = \frac{1}{q_j r_0}[J_n(q_j r_0)J_n'^*(q_j r_0) + J_n^*(q_j r_0)J_n'(q_j r_0)],$$

$$\Phi_1 = \left(1 - \frac{n^2}{q_0^2 r_0^2}\right)|H_n^{(1)}(q_0 r_0)|^2 + |H_n^{(1)'}(q_0 r_0)|^2,$$

$$\Phi_2 = \frac{1}{q_0 r_0}[H_n^{(1)}(q_0 r_0)H_n^{(1)'*}(q_0 r_0) + H_n^{(1)*}(q_0 r_0)H_n^{(1)'}(q_0 r_0)].$$

According to [23], the Q factor Q_c^Y of the resonator due to the loss in the conducting end faces for the Y_p mode is

$$Q_\mathrm{c}^Y = \frac{2\pi}{c}\frac{k' l R_Y}{R_S},$$

where

$$R_{EH} = 1 + \frac{F_1^H + \frac{1}{2}F_2^H - \frac{|J_n(q_H r_0)|^2}{|H_n^{(1)}(q_0 r_0)|^2}\left(\Phi_1 + \frac{1}{2}\Phi_2\right)}{\frac{k_z^2}{q_H^2}(F_1^H + F_2^H) - \frac{k_z^2}{q_0^2}\frac{|J_n(q_H r_0)|^2}{|H_n^{(1)}(q_0 r_0)|^2}(\Phi_1 + \Phi_2)},$$

$$R_{HE} = 1 + \frac{F^E - \frac{|J_n(q_E r_0)|^2}{|H_n^{(1)}(q_0 r_0)|^2}\left[\left(\frac{k_z^2}{q_H^2}+1\right)\Phi_1 + \frac{1}{2}\left(\frac{2k_z^2}{q_H^2}+1\right)\Phi_2\right]}{\varepsilon_\perp \varepsilon_z \frac{k^2}{q_H^2}(F_1^E + F_2^E) - \frac{k^2}{q_H^2}\frac{|J_n(q_E r_0)|^2}{|H_n^{(1)}(q_0 r_0)|^2}(\Phi_1 + \Phi_2)},$$

$$F^E = \varepsilon_z\left[\left(\frac{k_z^2}{q_H^2}+1\right)F_1^E + \frac{1}{2}\left(\frac{2k_z^2}{q_H^2}+1\right)F_2^E\right].$$

The energy characteristics of DDRs can be used to determine $\tan\delta$ of the insulator and R_S of the conductor at microwave frequencies from the measured parameters of the resonator (r_0, l, $[\varepsilon_{ij}]$) and its natural frequencies (ω) for a particular mode (Y_p).

MICROWAVE CHARACTERISTICS OF MATERIALS

The theory developed above can be used for material characterization at microwaves. The experimentally found spectral and energy characteristics of a DDR with the copper end faces operating at higher modes were reported in [17]. The resonator was made of sheet Teflon and had $r_0 = 3.9$ cm and $l = 0.71$ cm (Fig. 1).

For uniaxial crystals, the tensor $[\varepsilon_{ij}]$ has only two unknown components, ε_z and $\varepsilon_\perp$. To evaluate them, it is sufficient to experimentally find three pairs of the frequencies ω of the resonant modes and their azimuthal, n, and axial, m, indices ($m = 2l/\lambda_r$, where λ_r is the wavelength in the resonator) and to solve the self-consistent system of fundamental equations emerging from characteristic equation (1). A solution to this system uniquely determines the parameters ε_z and $\varepsilon_\perp$ of the resonator material.

The permittivity of the resonator under study found by the above method was $\varepsilon = 2.04 \pm 0.03$ (for an isotropic medium, $\varepsilon_\perp = \varepsilon_z = \varepsilon$). The radiation loss in the resonator was neglected, because $Q^Y_{\text{rad}} \gg Q^Y_{\text{d}}$ for whispering-gallery oscillations.

Experimental values of the unloaded Q factor Q_0^Y and formula (4) allow one to find the loss tangent $\tan\delta$ of the insulator and the surface resistance R_S of the end faces as a function of frequency. In general, the unloaded Q factor Q_0^Y depends on the environmental conditions. Therefore, with relationship (4), one can experimentally determine $\tan\delta$ and R_S as functions of, for example, temperature and pressure.

The loss tangent $\tan\delta$ of the DDR material was measured to be $(3 \pm 0.5) \times 10^{-4}$ at the room conditions.

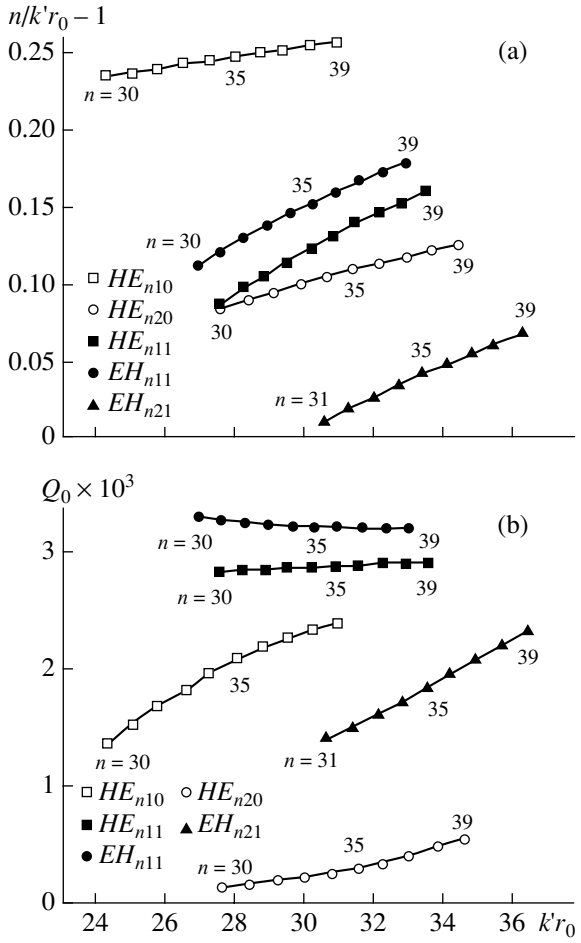


Fig. 2. (a) Slowing factor and (b) unloaded Q of a DDR with copper end faces vs. frequency.

The surface resistance R_S of pure electrolytic copper (with a surface roughness of less than 0.05 μm, which is much less than the skin depth ≈0.35 μm) in the form of a 35-μm-thick sheet used to clad the glass–cloth laminate was found to be $(5.5 \pm 0.2) \times 10^{-2}$ Ω.

SPECTRAL AND POWER CHARACTERISTICS OF THE DDR

The microwave parameters of the insulator and of the conductor on the end surfaces were used to derive the spectral and power characteristics of the DDR (Fig. 2). The frequency dependence of the slowing factor $\xi = n/k'r_0 - 1$ was obtained from relationship (3) for the reactance $X_S = R_S$. The unloaded Q factor Q_0 was calculated from formula (4). Figure 2 shows results for DDRs with Y_p oscillation modes having the azimuthal indices n from 30 to 39, radial indices $s = 1$ and 2, and axial indices $m = 0$ and 1.

The resonance frequencies of the resonator with copper end faces differ from those for the resonator with perfectly conducting surfaces by 10–15 MHz. The

power loss in the copper end faces decreased the unloaded Q factor 10 to 100 times.

CONCLUSION

DDRs excited at whispering-gallery oscillation modes can easily be combined with active and passive elements of millimeter-wave hybrid circuits. High-Q DDRs with finite-conductivity end faces are of interest for measuring the low surface resistance of superconductors at millimeter waves [18].

An integral equation describing the effect of finite conductivity of the end faces on resonant oscillations is derived and studied. The frequency dependences of the slowing factor and of the unloaded Q of a Teflon disk resonator with copper end faces are presented.

A self-consistent system of equations that is based on characteristic relationship (1) and includes the azimuthal and axial indices measured at the resonance frequencies of whispering-gallery modes is constructed. Its solution specifies the components of the permittivity tensor $[\varepsilon_{ij}]$ of an anisotropic uniaxial crystal whose optical axis is aligned with the resonator axis. The electrodynamic analysis of the DDR made it possible to evaluate the loss tangent of the insulator and the surface impedance of the conducting end faces from the energy characteristics measured.

Our method may find wide recognition in designing millimeter- and submillimeter-wave devices and in measuring microwave parameters of materials.

REFERENCES

1. S. N. Vlasov, Radiotekh. Élektron. (Moscow) **11** (3), 572 (1967).
2. J. R. Wait, Radio Sci. **2**, 1005 (1967).
3. Lord Rayleigh, Philos. Mag. **20**, 1001 (1910).
4. V. B. Braginsky, Kh. S. Bagdasarov, F. V. Bulygin, and V. S. Ilchenko, Pis'ma Zh. Tekh. Fiz. **11**, 427 (1985) [Sov. Tech. Phys. Lett. **11**, 177 (1985)].
5. V. I. Minakova and P. R. Stankov, Radiotekh. Élektron. (Moscow) **33**, 1696 (1987).
6. X. H. Jiao, P. Guillon, P. Auxtmery, *et al.*, IEEE Trans. Microwave Theory Tech. **35**, 1169 (1987).
7. Yu. M. Bezborodov, T. N. Naritnik, and V. B. Fedorov, *Microwave Filters Built on Dielectric Resonators* (Tekhnika, Kiev, 1989).
8. G. Annino, M. Cassetari, J. Longo, *et al.*, Rev. Sci. Instrum. **70**, 1797 (1999).
9. G. Annino, M. Cassetari, J. Longo, *et al.*, Chem. Phys. Lett. **281**, 306 (1997).
10. J. Krupka, D. Cros, M. Auborg, *et al.*, IEEE Trans. Microwave Theory Tech. **42**, 56 (1994).
11. N. Klien, C. Zuccaro, U. Dahne, *et al.*, J. Appl. Phys. **78**, 6683 (1995).
12. N. T. Cherpak, Yu. F. Filipov, E. V. Izhik, *et al.*, Funct. Mater. **5**, 61 (1998).
13. V. N. Egorov and I. N. Maltseva, Élektron. Tekh., Ser. I: Élektron. SVCh, No. 1, 3 (1984).
14. A. Ya. Kirichenko, Yu. V. Prokopenko, Yu. F. Filipov, and N. T. Cherpak, Radiotekh. Élektron. (Moscow) **33**, 300 (1989).
15. K. N. Tsibizov, S. A. Borisov, and Yu. M. Bezborodov, Zarubezhn. Radioélektron., No. 1, 21 (1981).
16. Yu. F. Filipov and S. N. Kharkovsky, Izv. Vyssh. Uchebn. Zaved., Radiofiz. **33**, 1304 (1990).
17. A. A. Barannik, Yu. V. Prokopenko, Yu. F. Filipov, and N. T. Cherpak, Radiofiz. Élektron. (Khar'kov) **5** (3), 104 (2000).
18. Yu. V. Prokopenko and Yu. F. Filipov, Radiofiz. Élektron. (Khar'kov) **6** (1), 30 (2001).
19. Yu. F. Filipov, S. N. Kharkovsky, and A. Ya. Kirichenko, Microwave Opt. Technol. Lett. **10**, 124 (1995).
20. N. T. Cherpak, Yu. F. Filipov, S. N. Kharkovsky, and A. Ya. Kirichenko, Int. J. Infrared Millim. Waves **14**, 614 (1993).
21. H. Peng, IEEE Trans. Microwave Theory Tech. **44**, 848 (1996).
22. Yu. F. Filipov and S. N. Kharkovsky, Pis'ma Zh. Tekh. Fiz. **16** (20), 24 (1990) [Sov. Tech. Phys. Lett. **16**, 775 (1990)].
23. Yu. V. Prokopenko, Yu. F. Filipov, and N. T. Cherpak, Radiofiz. Élektron. (Khar'kov) **4** (2), 50 (1999).

Translated by A. Khzmalyan

Technical Physics, Vol. 47, No. 6, 2002, pp. 737–741. Translated from Zhurnal Tekhnicheskoĭ Fiziki, Vol. 72, No. 5, 2002, pp. 85–89.

RADIOPHYSICS

Transformation of Surface Magnetostatic Waves Channeled by a Step Bias Field

A. Yu. Annenkov, S. V. Gerus, and S. I. Kovalev

Institute of Radio Engineering and Electronics (Fryazino Branch), Russian Academy of Sciences, pl. Vvedenskogo 1, Fryazino, Moscow oblast, 141190 Russia

e-mail: svg318@ms.ire.rssi.ru

Received November 8, 2001

Abstract—The propagation of magnetostatic waves in a waveguide ferromagnetic channel formed by exposing a ferromagnetic film to a step bias is simulated. Both weak and strong channel couplings are considered. The dispersion characteristics and the distributions of wave functions are calculated. It is shown that the amplitudes of "half-waves" fitting into the channel width alternately increase and decrease with increasing frequency and the half-waves pass into the channel with the largest bias. *© 2002 MAIK "Nauka/Interperiodica".*

INTRODUCTION

To date, the propagation of magnetostatic waves (MSWs) in simple rectangular magnetic waveguides [1] and in more complex two-channel ones [2] has been simulated. The substantial evolution of the wave functions when the MSW frequency changes has been discovered. However, the shape of the dispersion curves for the bias configurations considered remained virtually the same as for Damon–Eshbach surface magnetostatic waves (SMSWs) [3]. In this paper, we consider a bias configuration causing essential changes in the shape of the MSW dispersion curves.

MATHEMATICAL MODEL

A ferromagnetic film is exposed to a tangential magnetic (bias) field with the configuration shown in Fig. 1. The film plane coincides with the *xy* coordinate plane. The bias field is directed along the *x* axis and has a step along this axis. The distribution of the field along the other axes is assumed to be roughly uniform. Such an approximation is valid, because the nonuniformity along the width (*z* axis) can be neglected if the width *d* of the ferromagnetic film is much smaller than the characteristic sizes (h_A and h_B) of the field nonuniformity along the waveguide width (*x* axis). An MSW propagates in the direction of the *y* axis.

The approach used to solve the problem numerically is based on the magnetostatic approximation of the Maxwell equations. With regard for the equation of magnetic moment motion, this approximation yields the expression for magnetic potential Ψ specifying the spatial distribution of the MSW variable magnetic field:

$$\mathrm{div}[\mu(x, z)\mathrm{grad}\Psi] = 0. \qquad (1)$$

Here, $\mu(x, z)$ is the dynamic permeability tensor. It involves both the jumps of the magnetic parameters on the ferromagnetic film surface and nonuniformities induced by the magnetostatic field shown in Fig. 1. The anisotropy of the ferromagnet and the exchange interaction are neglected. Equation (1) is solved numerically by the finite difference method. First, we choose a rectangular contour away from the waveguide, so that the potential Ψ on it can be set equal to zero. Then, conditions of "magnetic" or "electrical" walls are written for this contour and a rectangular mesh is superimposed on the domain of interest. Finally, Eq. (1), represented in integral form on the mesh, is reduced to the problem of a quadratic matrix operator. This method was detailed in [4, 5]. Eventually, we obtain the wave numbers and distributions of the scalar potential Ψ of the MSW modes.

Although the distribution of the wave is calculated for the complex channel *C* as a whole (Fig. 1), the behavior of a wave can be described via the interaction of waves in two rectangular channels *A* and *B* with the parameters h_A, H_A and h_B, H_B, respectively. In this case, the modes of the entire waveguide can be considered as the hybridization of modes in separate rectangular channels, where the propagation of the SMSWs is mul-

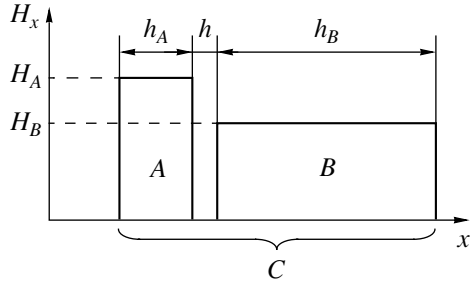


Fig. 1. Distribution of the bias field over the channel width. The channel *C* is formed by the combination of the channels *A* and *B*.

1063-7842/02/4706-0737$22.00 © 2002 MAIK "Nauka/Interperiodica"

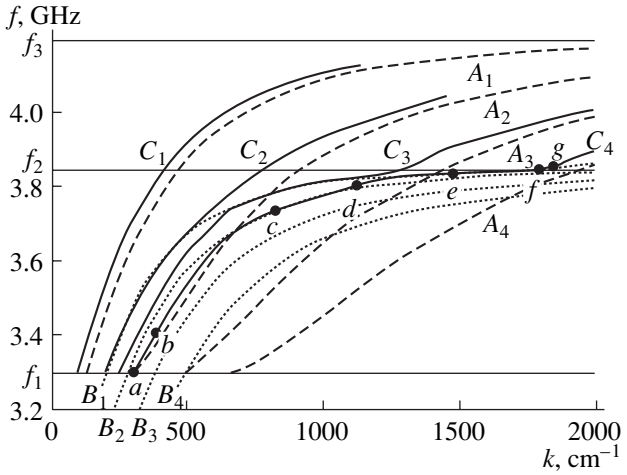


Fig. 2. Dispersion curves for the first four modes of the channels *A*, *B*, and *C*. d = 18 μm, $4\pi M$ = 1850 G, h_A = 0.1 mm, h_B = 0.3 mm, H_A = 570 Oe, and H_B = 445 Oe.

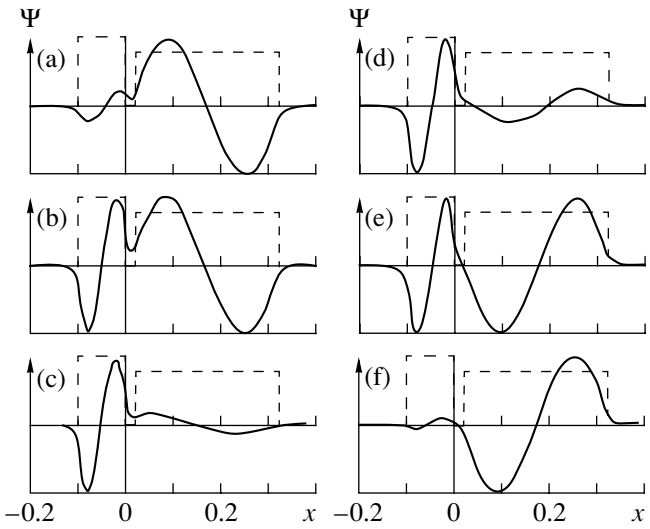


Fig. 3. Distribution of the potential for the (a, b, and c) third and (d, e, and f) fourth modes of the channel *C* for $kh \gg 1$ near the cross point at f = 3686 GHz, and k = 690 cm^{-1}. f = 3640 (a, d), 3687 (b), 3685 (e), and 3730 GHz (c, f).

timode because of the finite aperture. We will show that the shapes of the dispersion curves for the complex waveguide can be controlled by the parameters of the magnetic field nonuniformity.

In Fig. 2, the two sets of the SMSW dispersion curves are shown by dashed lines (channel *A*) and dotted lines (channel *B*) for the first four modes of either rectangular channel. The mode numbers coincide with the numbers of half-waves fitting into the channel width. The two sets are shifted relative to each other both along the vertical axis due to the different bias fields and along the horizontal one because of the different channel widths. The values of the parameters h_A, H_A, h_B, and H_B are chosen so as to provide the crossing of curves from various sets. The modes of the composite channel *C* arise from the interaction and hybridization of the initial modes in the channels *A* and *B*. As a result, waves with a complex distribution of the wave functions over the width of the entire channel form (Figs. 3, 4). In contrast to the simple rectangular channel, the number of half-waves fitting into the width of the composite channel *C* no longer indicates the number of its mode.

The mathematical model employed can be used for both weak and strong interactions.

DISCUSSION

MSWs occupy the frequency range from γH_B to $f_3 = \gamma(H_A + 2\pi M)$, where M is the saturation magnetization and γ is the gyromagnetic ratio. Surface waves, i.e., those not oscillating in the bulk of the film, have frequencies from $f_1 = \gamma\sqrt{H_A(H_A + 4\pi M)}$ to f_3. In this paper, we consider this type of MSW.

Note also that in the range $\gamma\sqrt{H_B(H_B + 4\pi M)} < f < f_1$, there exist surface–bulk MSWs with a complex distribution of the wave function along the waveguide width and thickness. They are surface in the channel *B*, and bulk in the channel *A*. In the frequency range $\gamma H_B < f < \gamma\sqrt{H_B(H_B + 4\pi M)}$, there exist bulk modes localized either over the total width of the channel *C* or over its part *B*.

Let us analyze the dispersion curves and magnetostatic potential distribution for surface waves propagating in our waveguide.

(i) *Weak coupling.* To begin with, consider the case where the interaction between the channels *A* and *B* is weak. This case may be realized by separating the channels from each other [2] to meet the condition $kh \gg 1$. Then, the dispersion curves of the waveguide *C* sequentially coincide with the segments of the curves corresponding to the channels *A* and *B* in accordance with their crossings in Fig. 2 (the crossings of the dashed and dotted curves).

The curve C_1 of the first mode coincides with the curve A_1 if $kh \gg 1$. The wave function is represented by the half-wave occupying the total width of the waveguide *C* below the frequency $f_2 = \gamma(H_B + 2\pi M)$. Above this frequency, the wave propagates almost exclusively in the *A* channel, whereas in the *B* channel, the wave function decays like in a vacuum but with a smaller decrement.

The dispersion curve C_2 of the second mode coincides with the curve B_1 in the range of small wave numbers. The wave potential is represented by two antiphase half-waves in the regions *A* and *B*. In this case, the amplitude Ψ_B of the half-wave in the *B* channel is much greater than that (Ψ_A) in the *A* channel. Their ratio depends on the channel coupling parameter,

which is proportional to exp(–*kh*). At the crossover, the amplitudes of the half-waves are the same. Above the point where the curves B_1 and A_2 intersect, the curve C_2 coincides with the curve A_2 (Fig. 2). Both half-waves gradually pass into the *A* channel, and the *B* channel contains only the decay function, since the wave solution for the channel *B* is absent above the curve B_1. The features of the wave function distribution near the crossover will be described later.

The dispersion curve C_3 of the third mode initially coincides with the curve B_2. Then, it merges with the segments of the curves A_2 and B_1 and ends, coinciding with the curve A_3. The potential Ψ undergoes the following changes while moving along the dispersion curve from bottom up. Its amplitude is the greatest in the channel whose dispersion curve coincides with the dispersion curve of the channel *C* at a particular frequency. For example, at small frequencies, two large amplitude half-waves are in the *B* channel and one small-amplitude half-wave is in the *A* channel (Fig. 3a). Above the first crossover point, one of the half-waves passes into the channel *A*. In this case, the amplitude Ψ_A of the second mode having formed in this channel reaches its peak value and the amplitude Ψ_B of the half-wave remaining in the channel *B* goes down to its minimal value, depending on the channel coupling parameter. After the second crossover has been passed, the number of half-waves in both channels does not change but the half-waves exchange amplitudes. Above the third crossover, all three half-waves gather in the channel *A*, whereas the wave solution in the channel *B* disappears.

The fourth mode C_4 behaves in a similar way, with the only difference that it begins with the mode A_2 of the channel *A* and then follows the curves B_2, A_3, B_1, and A_4. At small wave numbers, either channel contains two half-waves (Figs. 3d, 3c, and 3f). With increasing wave number, all the half-waves gather in the channel *A*. The amplitudes of the half-waves in either channel increase and decrease in turn, depending on which curve coincides with the dispersion curve of the fourth mode at a given frequency. This is illustrated in Fig. 5, where the dashed curves show the ratio Ψ_A/Ψ_B in the case of weak coupling. In essence, this curve characterizes the channel coupling parameter. The coupling is maximal in the frequency range where $\Psi_A/\Psi_B \approx 1$, which corresponds to the crossover ranges in Fig. 2.

Modes with higher numbers behave in a similar way. The dispersion curve of the complex waveguide *C* has inflection points, as well as convex and concave segments located in the ranges where the MSW dispersion curves of the channels *A* and *B* cross. The dispersion curves of the simple rectangular channels do not have such features. Consider the crossover ranges where the hybridization of the modes belonging to the channels *A* and *B* is the most distinct. At the cross points, the dispersion curves are separated by a distance

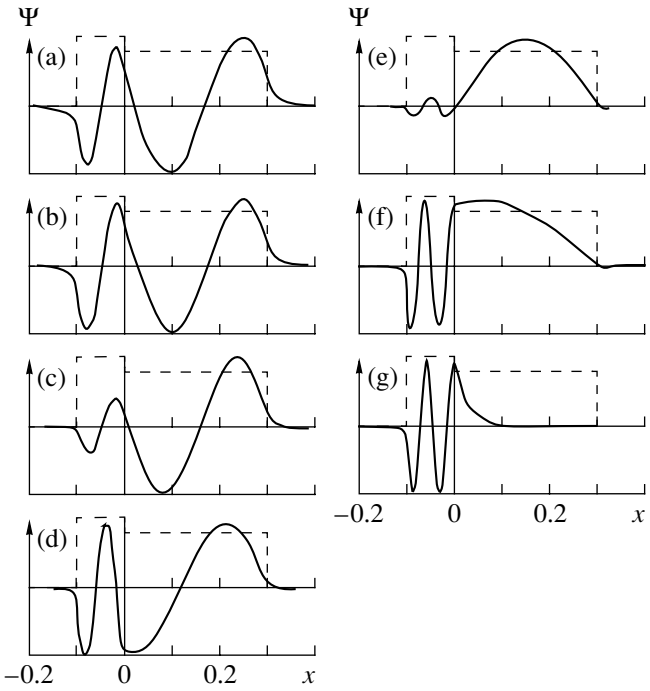


Fig. 4. Distribution of the potential for the fourth mode of the channel *C* at the characteristic points (a)–(g) marked in Fig. 2.

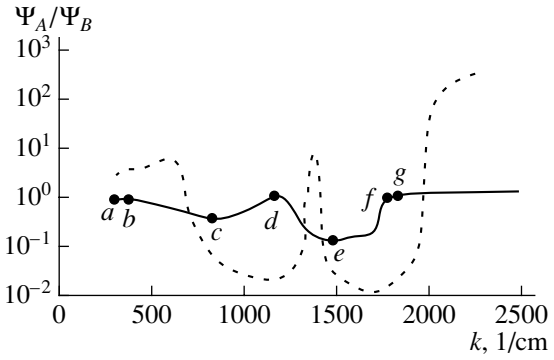


Fig. 5. Ratio of the potential amplitudes in the channels *A* and *B* vs. wave number.

proportional to the coupling parameter [6]. The resulting dispersion curve of the waveguide *C* is concave if it is located to the upper left of the crossover point. Otherwise, the curve is convex. The amplitudes of the half-waves in the channels *A* and *B* are the same (Figs. 3b, 3e). However, the higher frequency (concave) curve corresponds to the distribution of the Ψ function such that the half-waves in the different channels are in phase (Fig. 3b), whereas the lower frequency (convex) dispersion curve corresponds to the distribution for which the waves are in antiphase (Fig. 3e). Accordingly, inphase (in the aforementioned sense) half-waves have a higher frequency than antiphase ones.

In the above cases, as the frequency rises, the numbers of *A*(*B*)-channel modes crossed by the dispersion

curve of the channel *C* increase (decrease). However, this is not a general rule. The situation where the dispersion curves cross twice (as A_3 and B_4 in Fig. 2) or where two or more curves of one channel cross the curve of the other channel are also possible. Then, the numbers of *A*- and *B*-channel modes vary nonmonotonically. The general rules are the following.

When approaching the crossover range, the half-waves of the potential Ψ either pass from the channel *B* to the channel *A* or do not change their location at all. The number of half-waves in the channel *B* may increase only if an additional inphase half-wave appears. This half-wave arises, according to the above-mentioned rule, at the inflection points of the dispersion curve *C*; however, it is not regular in the sense that its left part does not cross the abscissa axis but is adjacent to the half-wave from the neighboring channel *A*. Therefore, it is convenient to take the number of zeros of the Ψ function, rather than the number of half-waves, as the parameter most fully characterizing the modes of the complex channel.

The number of zeros of the wave function is constant and equals $N - 1$, where N is the number of the *C*-channel mode.

With increasing frequency, all the half-waves are found in the channel *A*, regardless of how the curves *A* and *B* cross.

(ii) *Strong coupling.* Now we turn to the case where $h = 0$ and the channels *A* and *B* are located close to each other. In Fig. 2, the dispersion curves of the first four modes of the waveguide *C* (solid lines) are shown. These curves are displaced from their unperturbed prototypes, namely, from the dispersion curves of the waveguides *A* and *B*. As the wave number increases, the dispersion curves either cross their prototypes (as takes place for the mode C_2) or move away from them (mode C_3). The dispersion curves as if interact: they repulse each other, shift, and change shape. Not only their slope but also their curvature changes. The reconfiguring of the curves is accompanied by a substantial reconfiguring of the wave function Ψ.

Consider these metamorphoses at length with the SMSW fourth mode. In the range of small wave numbers, the dispersion curve of the composite channel C_4 is located between the second modes B_2 and A_2 (Fig. 2, points *a* and *b*). With increasing frequency, the curve C_4 shifts towards B_2 and then returns to A_2. In Figs. 4a and 4b, the distribution of the function Ψ for these points is shown. As is seen, less than two half-waves fit into the *A*-channel width, whereas more than two half-waves fit into that of the channel *B*. As a result, the waveguide *C* supports four half-waves of the fourth mode. The point *a* is closer to the dispersion curve A_2 than to B_2, whereas the point *b* is almost exactly between them (Fig. 2). This affects the distribution of the potential Ψ. In the plot of the potential distribution for the point *a* (Fig. 4a), almost two half-waves fit into the narrow channel. Then, these half-waves shift towards the channel *B* (Fig. 4b).

The dispersion curve A_2 gradually approaches B_2 as the wave number grows. At the same time, the potential distribution in the waveguide *C* approaches the form when either channel contains strictly two half-waves (Fig. 4c). This occurs at the point *c* located not at the point where the dispersion curves A_2 and B_2 intersect but much to the right of this point (Fig. 2) because of the strong channel interaction.

A further redistribution of the magnetostatic potential over the width of the complex waveguide with increasing frequency proceeds in a way similar to that for weak coupling between the waveguides *A* and *B*. The half-waves gradually pass from the broad channel to the narrow one. The broad channel becomes evanescent, and the potential distribution in it takes the form of an exponentially vanishing function (Fig. 4g and the point *g* in Fig. 2). However, the strong channel interaction smears the crossover range and changes the shape of the wave functions.

In Fig. 5, the solid line shows the ratio Ψ_A/Ψ_B of the amplitudes of the half-waves. In contrast to the case of weak coupling (dashed line), strong coupling causes the amplitude Ψ_B to exceed Ψ_A almost in the whole range of wave numbers. The wave numbers above the point *f* can be disregarded, since the waveguide *B* is evanescent for them. For definiteness, the amplitude Ψ_B is taken to be equal to the value of the wave function Ψ at the center of the waveguide. In the case of strong coupling, the ratio Ψ_A/Ψ_B can no longer be considered a measure of channel interaction as for weak coupling. The channel interaction is strong both near and far from the cross points. It redistributes the amplitudes in favour of the channel *B* and flattens the curve Ψ_A/Ψ_B at small wave numbers.

The curvature of the concave segments of the dispersion curve becomes smaller. As in the case of weak coupling, these segments (Fig. 2: points *d*, *f*) correspond to the inphase neighboring half-waves with the same amplitudes in the different channels (Fig. 5; points *d*, *f*). Only if the coupling is strong are these half-waves not separated by a dip and make up a single half-wave (Figs. 4d, and 4f).

The antiphase half-waves near the boundary between the channels are shown in Figs. 4c and 4e. For weak coupling, these waves correspond to the point in the low-frequency (convex) dispersion curve in the crossover region and the amplitudes of the half-waves in the channels *A* and *B* are equal to each other. In the case of strong coupling, the positive curvature is small at these points; i.e., the crossover region is smeared greater than for the inphase half-waves. The amplitudes of the half-waves are not equal; moreover, their ratio tends to the minimum (Fig. 5; points *c*, *e*).

CONCLUSION

It is shown that dispersion curves that differ from the exponential form, which is typical of rectangular waveguides, can be constructed by appropriately choosing the configuration of the bias field which forms the MSW waveguide. The combination of two closely spaced channels with different widths and bias fields makes it possible to produce convex and concave segments of certain curvature in given segments of the dispersion curves.

The distribution of the wave function in either channel of the double waveguide is a periodic function characterized by the number of half-waves. With increasing frequency, the half-wave amplitudes in the channels alternately increase and decrease, so that the wave concentrates in either one or the other channel, depending on the frequency. The half-waves pass from one channel to the other. In the high-frequency range, the entire wave occupies the channel with the maximal bias field. The number of zeros of the wave function is constant throughout the range where the quasi-surface MSWs exist.

If the spacing between the channels is large (weak coupling), the amplitudes of the half-waves are the same at the points where the modes of the initial channels intersect. If neighboring half-waves from the different channels are in phase, a concave segment is observed in the dispersion curve. In the opposite case, the dispersion curve is convex. The scale of these changes decreases with decreasing channel spacing (increasing coupling). However, this takes place in a dissimilar manner in either channel. The curvature of the convex and concave segments in the dispersion curves decreases, and their locations change.

REFERENCES

1. M. Miyazaki, K. Yashiro, and S. Ohkawa, IEEE Trans. Microwave Theory Tech. **33**, 421 (1985).
2. A. Yu. Annenkov, S. V. Gerus, and S. I. Kovalev, Zh. Tekh. Fiz. **68** (2), 91 (1998) [Tech. Phys. **43**, 216 (1998)].
3. R. W. Damon and J. R. Eshbach, J. Phys. Chem. Solids **19**, 308 (1961).
4. I. V. Vasil'ev and S. I. Kovalev, Radiotekh. Élektron. (Moscow) **38**, 2174 (1993).
5. A. Yu. Annenkov, I. V. Vasil'ev, S. V. Gerus, and S. I. Kovalev, Zh. Tekh. Fiz. **65** (4), 71 (1995) [Tech. Phys. **40**, 330 (1995)].
6. W. H. Louisell, *Coupled Mode and Paramagnetic Electronics* (Wiley, New York, 1960; Inostrannaya Literatura, Moscow, 1963).

Translated by M. Fofanov

Technical Physics, Vol. 47, No. 6, 2002, pp. 742–751. Translated from Zhurnal Tekhnicheskoĭ Fiziki, Vol. 72, No. 5, 2002, pp. 90–100.

**SURFACES,
ELECTRON AND ION EMISSION**

Impurity Composition and Cleaning of the Cylindrical Surface of CVD Single-Crystal Tungsten

B. M. Zykov, V. N. Zykova, V. N. Lebedev, and Yu. K. Udovichenko

Sukhumi Physicotechnical Institute, Academy of Sciences of Abkhazia, Sukhumi, 384914 Georgia

Received October 3, 2001

Abstract—Secondary-ion mass spectrometry (SIMS), Auger electron spectroscopy (AES), and low-energy electron diffraction (LEED) methods are used to determine the chemical composition on the surface and in the near-surface layer of CVD cylindrical single-crystal tungsten, which is applied as the emitter material in thermionic thermal-to-electrical energy converters in space-borne nuclear power plants. Results obtained are compared with those for reference perfect plane-parallel W(110). *© 2002 MAIK "Nauka/Interperiodica".*

INTRODUCTION

Single-crystal tungsten with the (110) close-packed face on its surface has been shown [1, 2] to be a candidate material to be used as a collector in thermionic thermal-to-electric energy converters (TECs), which are used in space-borne nuclear power plants. Unlike Mo(110) [3–5], Nb(110) [6, 7], or $Mo_{97.5}Nb_{2.5}$(110) [8, 9], the suboxide films of which readily oxidize further under TEC operating conditions, films of lower tungsten oxide W_2O seem to be an optimal overlayer for cesium in Me–O–Cs systems used in TECs [1, 2, 10]. The term "optimal" here means that, when used in the Me–O–Cs systems, this film provides a combination of the minimal work function φ_{min} and maximal thermal stability. The temperature T = 600 K, at which the work function of the W_2O/Cs system still remains the lowest under static conditions (i.e., when the O and Cs fluxes are absent once the film has been formed), is the highest of those ever reported for the Me–O–Cs system. Under dynamic conditions of TEC operation, the collector is optimized, as a rule, by properly selecting its temperature T_c and Cs vapor pressure. At an oxygen partial pressure $P_O \sim 10^{-4}$ Pa (a total residual pressure $P_{res} = 10^{-2}$–5×10^{-3} Pa), the W_2O film on W(110) exists in the interval 1425–1625 K; at lower P_O, even from 925 K [2, 10]. The temperatures cited are the operating range of T_c. However, the most significant advantage of the W(110)–O system is that its properties do not change when an excess amount of W is deposited (i.e., the system offers the property of self-recovery). This is because oxygen can dissolve only in the first atomic layer (i.e., at the surface) of Group VIb transition metals (in contrast to Group IVb and Vb metals) [2, 10]. Therefore, when excessive W is deposited on a surface covered by an already existing film of any binary (W–O) or ternary (W–O–Cs) compound, the film emerges on the surface again and cannot be buried by the tungsten unless the process conditions are changed. This removes the problem of degradation of TEC performance due to the transfer of the emitter material (W) to the collector (Group IV(V)b metals or their alloys).

The above properties are typical of W flat single-crystal surfaces. However, in modern space-borne sources of electrical energy, the electrodes have the cylindrical form or, in general, are polyhedral. Today, polyhedral single-crystal tungsten emitters can be produced by CVD. For example, fairly perfect extended hexagonal W single crystals with {110} or {112} [11] faces can be deposited on single-crystal Mo cylinders with the ⟨111⟩ axis by varying the chloride process conditions (WCl_6 vapor pressure, substrate temperature, and oxygen content in the vapor phase). Subsequently, however, the surface cannot be made single-face even if special procedures (selective etching, high-temperature annealing *in vacuo*, etc.) are applied: there always exist at least two, (110) and (112), faces. Since the angle between these faces in the bcc lattice is π/8, microscopically, this surface must consist of inclined-wall terraces and the dihedral angle between them must be 7π/8. Even from these general considerations (the presence of extended one-dimensional grooves and inclined faces), one can expect that the integral emission and adsorption properties of the rough surfaces to differ from those of the (110) and (112) surfaces. Moreover, the surface of CVD tungsten must contain a large concentration of defects, such as facets and screw dislocations [11–14]. Yet, the much lower concentration of defects on the single-crystal tungsten cylindrical surface compared with that on polycrystalline tungsten coatings covering an Nb + 1% Zn alloy (which is produced by the same CVD technology and is widely used as a TEC collector material) implies that the impurity concentration in the former case is lower. This point is of special concern for carbon (the most harmful impurity in emission electronics), which diffuses from bulk defects toward the surface upon heating and can appear on it in amounts far exceeding the bulk concentration,

1063-7842/02/4706-0742$22.00 © 2002 MAIK "Nauka/Interperiodica"

i.e., segregate. Even special cleaning techniques may turn out to be inefficient in this case.

Therefore, studies of oxygen adsorption on, as well as adhesion of, O–Cs binary films to the cylindrical surface of CVD single-crystal tungsten as a collector material are of great importance. It appears that the shape of the surface cannot help influence its properties. Therefore, we will study the properties of a cylindrical single-crystal tungsten emitter. Before proceeding to this study, it is natural to consider the chemical composition (including impurity composition) of the surface and near-surface layers of the material and elaborate methods for obtaining atomically smooth surface.

EXPERIMENTAL TECHNIQUES

A cylindrical fragment of a single-crystal W emitter of diameter ≈20 mm was electrolytically etched (polished) in a 3% solution of NaOH in distilled water and then subjected to high-temperature annealing *in vacuo*. It was bored out (without cooling) to thin the walls down to 1.2 mm. Then, 10×10-mm^2 specimens were cut out from the fragment with a dry diamond disk. Before being loaded into the instruments, one of the specimens was electrolytically etched again in the same solution and rinsed in distilled water to remove the residual etchant. Then, all the specimens were rinsed in methylene chloride to remove oil spots and in boiling distilled water to remove salt contaminants.

Two samples, etched and unetched, were placed into an MS 7201 M secondary-ion mass spectrometer to determine the chemical composition of surface contaminants. The specimens were not heated. The instrument also contained a Mo specimen that was used to precalibrate the ion gun and a reference, a plane-parallel W(110) single-crystal specimen, to perform comparative studies. The latter was prepared according to the requirements for LEED and was identical to that used in [10]. The impurity content in the reference specimen at a depth of 15–20 μm was as follows (as determined by the SIMS method with a Cameca IMS-3F instrument): H, 20; C, 50; O, 20; Na, 300; Al, 0.7; Si, 3.7; K, 100; Ca, 0.6; Ti, 0.2; Ni, 1.5; Cu, 0.8; Mo, 0.01; and Ta, 3 ppm.

During recording of the mass spectra, the pressure of residual gases did not exceed 5×10^{-6} Pa; that of the working gas, $(1–2) \times 10^{-2}$ Pa. The evacuation facilities did not contain hydrocarbons (the working fluid of the steam-jet pump was 5F4É). Extra-pure-grade He and Ar were used as working gases. The ion beam current, 7–18 μA, was kept constant during the experiments. The angle of incidence of the beam on the specimen was 50°; the ion energy, 4 keV; and the spot size on the specimen, ≈0.1 cm^2. The exposure time (including the spectrum recording time) was ≈20 min. First He^+ spectra and then Ar^+ spectra were recorded. The Ar^+/He^+ sputtering ratio was 60 [15].

The spectra were recorded successively as a series of several operators without additional exposure to the ion beam. The reference had to be only purified; therefore, it was exposed to the Ar^+ beam immediately. Three of its spectra were recorded in argon and then four spectra in helium. The same was done for the unetched W. For the etched W, two series of spectra were taken: three spectra in Ar and then three spectra in He (series 1) and five spectra in Ar and four spectra in He (series 2). The last-recorded He spectrum for W(110) was taken as the reference spectrum. The series were recorded so as to study all the specimens in any of the gases without interrupting the instrument operation. The only exception was the etched W: for it, one more series of spectra in both He and Ar was recorded afterwards. During the idle time of the instrument, it was filled with Ar (starting from a high vacuum) up to a pressure of several hundreds of pascals. Prior to recording, the instrument was evacuated and then a desired working gas was delivered to a given pressure.

Upon processing the spectra, only the most intense peaks of nongaseous impurities, namely, C, Na, Al, K, Ca, Cr, Fe, Re, WO, and ReO, were taken into consideration. First, the peak intensities I of these impurities were normalized to the peak intensity of the main tungsten isotope ^{184}W. Then, for each of the series, a change in the peak intensity of these impurities relative to the intensity of the same peaks in the initial spectrum in each of the series (i.e., the ratio I_i/I_0) was determined. In this way, we found by how many times the peak intensity of any impurity changes from spectrum to spectrum in a given series or from series to series for a particular working gas.

The third etched specimen was placed into an original general-purpose electron spectrometer (GPES), which, during evacuation at 625 K to a desired pressure ($P_{res} \le 10^{-8}$ Pa), was not heated so as to correctly compare subsequently the elemental composition of the surface contamination at various heating temperatures. This composition was determined by AES with the electron gun mounted at an angle of 20° to the specimen surface. The energy E_p of the primary beam was 1200–1400 eV; the modulating voltage, 2.5–3.5 V. The spectra were recorded as the first derivative of the backscattered electron distribution with respect to energy by means of a four-grid decelerating field energy analyzer (DFEA) [16]. The structure of the surface contamination film, as well as that of the clean surface, could be monitored by LEED with the electron gun normal to the surface.

Heating of the specimens up to ≈1000 K was accomplished by thermal radiation. To still higher temperatures, the specimen could be heated by electron bombardment of its back surface. Carbon was removed from the surface by oxidation (the oxygen partial pressure P_O was 5×10^{-5} Pa). The source of spectral-grade atomic oxygen is described in [17]. During the cyclic cleaning of the surface, the oxidation time (to remove

carbon) was 2 min at 1100 K and subsequent heating to remove metal oxides lasted 1 min at 1900 K without lowering P_O [18].

IMPURITY COMPOSITION NEAR THE SURFACE

Consider first the cleaning of the reference. Between the first and third Ar spectra, only the peak intensities of C, WO, and ReO decrease: by a factor of 1.4, 2.4, and 1.6, respectively. That of Re remains unchanged. For the remaining impurities, the peaks grow: for Na, K, and Cr most significantly (by a factor of 7, 5, and 4.4, respectively); for the others, 1.2–1.5 times. This means that Na, K, and Cr are accumulated well below the surface. Another situation arises with the He^+ beam, which etches off a much thinner (approximately 60 times) layer from the surface than the Ar^+ beam [15]. Here, in the fourth spectrum, the peak intensities for Na, Al, Ca, Cr, Re, ReO, and WO do not change, while those for C, K, and Fe decrease, respectively, by 1.8, 1.4, and 2.5 times. Hence, the relative surface concentration of the last three impurities is lower than that of the others (Fig. 1a). This spectrum will hereafter be referred to as the reference spectrum.

The etched specimen of the CVD tungsten was first exposed to the He^+ beam (Fig. 1b). From the three spectra of the first series, it follows that the concentration of all the impurities listed above decreases (the decrease being considerable for some of them) the instant the He irradiation starts. For example, even in the second spectrum, the intensity of the Na peak drops 19 times; Al, 6 times; K, 31 times; Ca, 6.8 times; Cr, 3.3 times; and Fe, 11 times. However, for C, Re, ReO, and WO, the peak intensity changes by no more than 10%. In the next spectrum, the peak intensity for Cr, Re, ReO, and WO remains the same, while that for the other impurities grows but does not exceed the initial values. For example, the peaks of Na and Al grow four times; K, 6.8 times; Ca, 3.6 times; and Fe, 2.6 times. Consequently, these impurities are process-related impurities (Na and K are likely introduced by the etchant).

The following series of three spectra in Ar^+ confirms the conclusions drawn for the He^+ irradiation in the first series. The peak intensity for all the impurities increases (except for Re, ReO, and WO, for which it drops): for C, twofold; Na, 9.6 times; Al, four times; K, 3.6 times; Ca, 4.8 times; Cr, 2.8 times; and Fe, 4.8 times relative to the intensity in the first spectrum (Fig. 1c). This means that many of the impurities concentrate much deeper the layer than can be probed with the He^+ beam for a reasonable time.

Even in the Ar^+ series next to the first one, the concentration of the impurities (with the exception for Cr) starts decreasing. The chromium concentration in the last, fifth, spectrum of this series grows 2.6 times relative to its concentration in the first spectrum. Hence, among the impurities considered, chromium is buried at the greatest depth. As for the other impurities, the concentration (peak intensity) in the last spectrum of the second Ar^+ series decreases relative to that in the first spectrum of the same series by a factor of 2.3 for C, 4.9 for Na, 3.3 for Al, 9 for K, 1.2 for Ca, 2 for Fe, and 1.6 for WO and ReO. The intensity of the Re peak does not change. The fact that the impurity content drops with depth suggests that the impurities are localized near the surface and are related most likely to the deposition or subsequent processing of the material.

A comparison of the last spectrum of the specimen in the second Ar^+ series with the last spectrum of the reference shows that they are almost identical in purity (Fig. 1d). The discrepancy in the peak heights is no more than 5–10% for most of the impurities. The exception is K, Ca, and Cr, for which the discrepancy is ≈25%.

Once one He^+ series and two Ar^+ series had been taken from the etched specimen, one more He^+ series of four spectra was recorded. The same was done for W(110). In the former case, the peak intensities are reduced further from spectrum to spectrum (by 25–50% on average) for all the impurities except for Al (10%) and Fe (three-fold decrease). However, the intensity of the WO and ReO peaks grows by 2–3% and that of the Re peak, by 17%. A comparison of the initial spectrum in the first He^+ series (when the surface contamination was the highest) with the last one in the second He^+ series shows that the concentration decreased by a factor of 3.3. for C, 229 for Na, 16 for Al, 111 for K, 18.5 for Ca, 2.7 for Cr, and 18 for Fe (with regard for the intermediate cleaning by the Ar^+ beam). Conversely, the concentration of Re, ReO, and WO rose by 26, 22, and 11%, respectively (Fig. 1e).

At the same time, comparing the final spectrum of the specimen in the He^+ series (Fig. 1e) with the same spectrum for W(110) (Fig. 1a), we see that the concentrations of C, Al, Ca, Cr, Re, ReO, and WO in them coincide within 4–11% (the specimen becomes purer than the reference). The Na and K concentrations in the specimen are higher by a factor of 2.3 and 1.3, respectively, while the Fe concentration in W(110) is 1.3 times that in the specimen. Certainly, these data characterize the surface, rather than volume, purity of the specimen and W(110).

Consider now the unetched specimen. In view of the results described above, He^+ irradiation preceding Ar^+ irradiation was not performed. Here, the peak intensities vary not as considerably as for the etched specimen. For example, the intensity in the last (third) spectrum of the Ar^+ series decreases relative to the initial spectrum by a factor of 1.4 for C, 1.2 for Fe, 1.1 for Re, 1.4 for WO, and 1.6 for ReO. The intensity increases by a factor of 1.7 for Na, 1.1 for Al, 1.1 for K, 1.4 for Ca, and 2.1 for Cr.

Next was the He^+ series of four spectra. Here, the intensity in the fourth spectrum decreases relative to

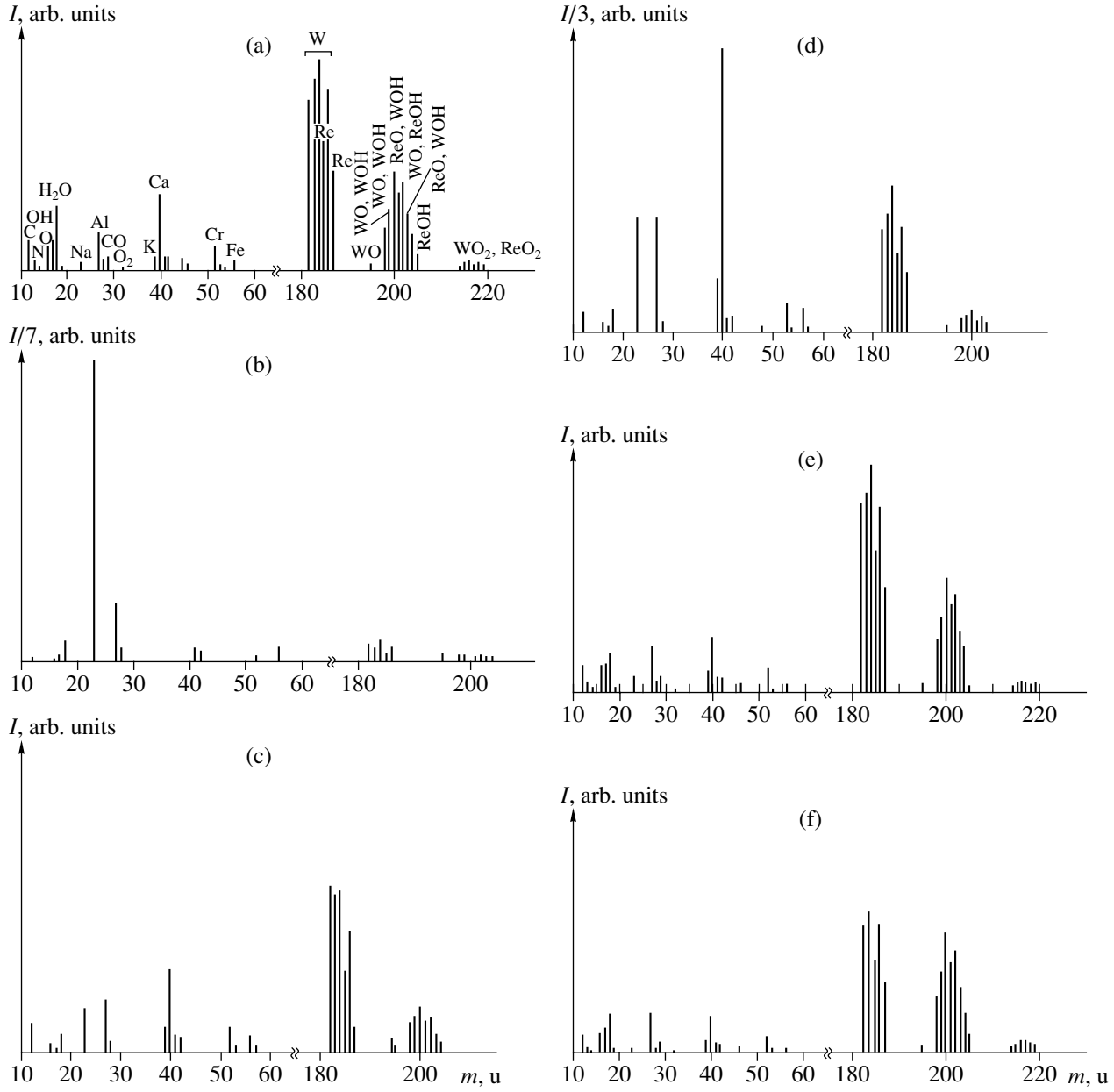


Fig. 1. Secondary-ion mass spectra. (a) Last spectrum in the last He^+ series for the W(110) reference, (b) initial spectrum in the He^+ series for the etched specimen, (c) initial spectrum in the Ar^+ series for the etched specimen after exposure to the He^+ beam, (d) last spectrum in the second Ar^+ series for the etched specimen, (e) last spectrum in the second He^+ series for the etched specimen after exposure to the Ar^+ beam in the second Ar^+ series, and (f) last spectrum of the unetched specimen after exposure to the Ar^+ beam.

that in the initial spectrum by a factor of 1.1 for C and K, 1.3 for Ca, and 1.7 for Fe. However, the intensities of the Al, Cr, WO, and ReO peaks increase by a factor of 1.4, 1.1, 1.7, and 1.5, respectively. A comparison between the last spectrum of this series (Fig. 1f) with the last spectrum for W(110) in the He^+ series (Fig. 1a) shows that the specimen contains 1.3 times more Al, 1.1 times more K, and 1.5 times more WO and ReO. However, the concentration of the other impurities in the specimen is higher by a factor of 1.3 for C, 1.2 for Ca, 1.1 for Cr, 1.3 for Fe, and 1.1 for Re. Thus, the unetched CVD tungsten specimen approaches the single-crystal W(110) reference in purity after the He^+ and Ar^+ irradiations.

The data for the two CVD tungsten specimens show that most of the contaminants are due to the electrolytic polishing in the etchant. Therefore, if the additional treatment pursues no goal other than surface cleaning, it should not be carried out.

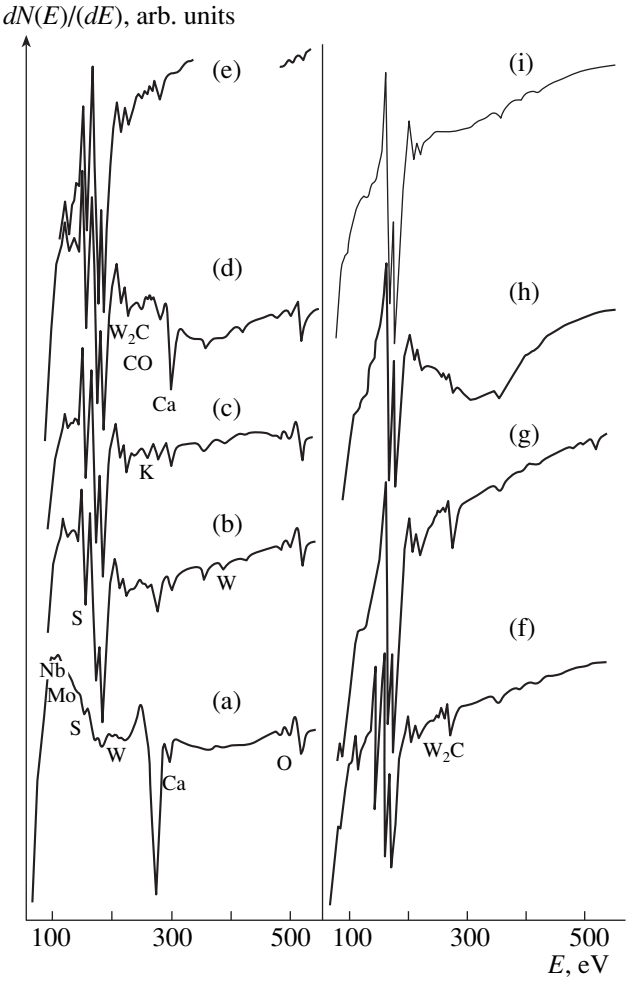


Fig. 2. Auger spectra after heating. T = 575 (a), 1025 (b), 1075 (c), 1225 (d), 1375 (e), 1475 (f), and 1625 (g) K. (h) Second heating at 1125–1175 K, (i) clean surface.

The evaluation of the impurity content in W requires that the relative sensitivity factors (RSFs) be known. Unfortunately, the RSFs only for Al, Cr, and W have been reported in the literature. Therefore, we calculated the absolute impurity content for Al and Cr. From the last spectrum of W(110) in the He^+ series (Fig. 1a), the Al and Cr content in W is 3×10^{-2} and 3.5×10^{-1} at. %. From the same spectrum of the etched specimen (Fig. 1e), we have 3×10^{-2} at.% Al and 3×10^{-1} at. % Cr; for the unetched specimen (Fig. 1f), also 3×10^{-2} and 3×10^{-1} at.%, respectively.

Thus, the purity of CVD single-crystal tungsten (as a TEC emitter material) approaches the purity of a W(110) single crystal, which completely meets requirements for LEED specimens if the ion etch depth is sufficiently large. The additional electrolytic cleaning (polishing) increases the concentration of impurities. However, it is etched specimens of CVD tungsten unexposed to ion beams in the SIMS method that were taken for subsequent AES studies in the GPES.

SURFACE CONTAMINANTS

After GPES evacuation at 625 K down to P_{res} = 10^{-8} Pa and cooling, the specimen was heated again to 575 K. No diffraction pattern is observed in this case. The AES spectrum shows the most intense peak from the thick carbon or hydrocarbon film at 272 eV and the second most intense oxygen peak (515 eV) with its satellites (475 and 490 eV). Next in order of intensity are the calcium peak (293 eV), as well as traces of niobium (105, 133, and 142 eV), sulfur (152 eV), tungsten (169 and 179 eV), and Mo (120 eV) (Fig. 2a).

Upon further heating, the S peak (152 eV) grows starting from 625 K and the W peaks, from 925 K. At 1025 K, the C peak (272 eV) lowers considerably and a faint K peak (253 eV) arises. Its position coincides with one of the peaks from CO or W_2C carbide. The more intense peak from these substances (261 eV) has not yet appeared. The S and W peaks grow substantially (Fig. 2b). At 1075 K, a weak LEED pattern with anomalously moving reflections is observed upon normal incidence of the primary electron beam. However, it still bears little information and cannot be identified. Yet, taking into account that the reflections move in the [01] and $[0\bar{1}]$ directions when E_p is varied, we can assume the presence of prismatic facets arranged along the cylindrical surface of the specimen. At this temperature, the intensity of the Auger peak for Ca (293 eV) exceeds that of the C peak. Simultaneously, the K peak grows, approaching the height of the C peak. The basic O peak (515 eV) is also of high intensity (Fig. 2c).

At 1175 K, the S peak starts and the C peak continues decreasing. At 1225 K, the Ca peak reaches the maximum intensity, which is four times as high as that of the C peak (Fig. 2d). The LEED pattern from the W(112) face represents a fairly distinct (2×2) structure with anomalous reflections which move in the directions mentioned above when E_p is varied (Fig. 3a). At the same time, no additional reflections are seen in the LEED pattern from the W(110) face (this pattern is obtained if the angle of incidence of the primary beam is changed by 20°–30°). The (2×2) pattern is most likely associated with Ca segregation, since when the AES spectrum contains even more intense lines for the other elements, such as S, C, and O, neither this nor any other LEED pattern is observed.

After the appearance of the initial LEED pattern, it becomes evident that, microscopically, the cylindrical surface of the CVD single-crystal tungsten consists of terraces with the (112) face as the base and the (110) faces as the inclined walls. The base is parallel to the cylinder envelope. Since the angle between these faces in the bulk is $\pi/8$, the dihedral angle of the terraces must equal $7\pi/8$.

However, before relating the (2×2) pattern to Ca or CaO, we should consider other processes that may occur in parallel with Ca segregation at 1225 K. At this temperature, carbon is not removed from the transition

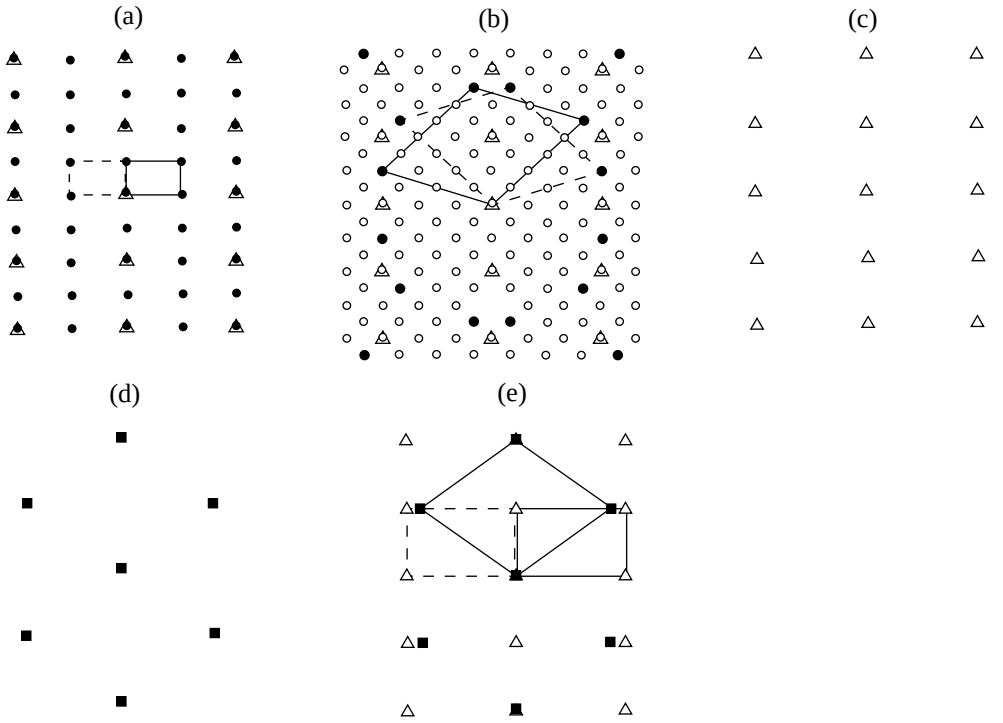


Fig. 3. LEED patterns (squares, reflections from the (110) face; triangles, reflections form the (112) face; filled circles, reflections from the structure; empty circles, moiré reflections). (a) (2×2), CaO + CO on W(112); (b) $c(3 \times 2)$, W_2C(0001) on W(112); (c) clean W(112) face; (d) clean W(110) face; and (e) artificial superposition of the patterns from the clean W(112) and W(110) faces.

metal surface [2]; therefore, the continuing decline in the C peak intensity (even after its sharp drop at 1025 K) and the concurrent growth of the Ca peak can be assigned to the "screening" of C by fast Ca diffusion toward the surface. However, the temperature interval where the Ca peak increases is at the same time optimal for C oxidation [2, 5, 10, 18]. Thus, the surface (more specifically, the (112) face, because carbon does not precipitate on the (110) face) may contain both Ca and CO.

At 1275 K, the Ca and O peaks start to decrease. At 1375 K, the former disappears and the latter has an intensity of roughly one sixth that observed when the Ca intensity was the highest. The (2×2) LEED pattern also disappears. Note that alkaline-earth metals produce strong double chemical bonds with oxygen when oxidizing, so that the oxides on the refractory metal surface do not dissociate. This leads us to the assumption that Ca is desorbed from the surface in tandem with O; in other words, we suppose that CaO, not elemental Ca, is present on the surface. Moreover, the K peak disappears, but extra carbon peaks at 242, 251, and 261 eV arise. They indicate the presence of surface C in the form of W_2C, CO, or both (Fig. 2e) (the presence of CO is indicated by the presence of O peaks, which are absent if pure carbides are the only surface components). The intensity of the C basic peak (272 eV) grows by roughly 2.5 times. This means that upon heating, the calcium at least partially screens the carbon. We cannot argue that the screening is complete, because the temperature 1375 K is close to that of fast carbon diffusion from the volume to the surface and carbide formation [5, 19].

At this temperature (1375 K), the material is intensely sputtered from the surface. This is indicated both indirectly (by a sharp rise in the residual pressure P_{res}) and directly (by the appearance of a deposit on the shutter placed before the specimen is heated). Hence, the actual amount of Ca on the surface is large and can screen C. Since O does not dissolve in the bulk of Group VIb elements (because the associated atoms lack vacant *d* orbitals [10, 20]), CaO oxide may form only when Ca segregated on the surface or at surface defects combines with residual surface oxygen. For $P_{res} = 10^{-8}$ Pa, the partial oxygen pressure $P_O = (1–5) \times 10^{-10}$ Pa; therefore, it can be assumed that the external O flow is absent upon heating of the specimen.

Carbon oxidation and carbide formation take place in a similar way.

The conventional approach to the identification of the (2×2) LEED pattern on W(112) gives the coverage $\Theta = 0.25$ (Fig. 4a). However, with such Θ, the sputtering of CaO cannot be so high. Therefore, as a first approximation, we can assume that the (2×2) pattern is formed not by CaO molecules but by substrate atoms unscreened by the film and producing a surface framework having the (2×2) structure with its interior filled with a film copying the substrate structure. Then, the

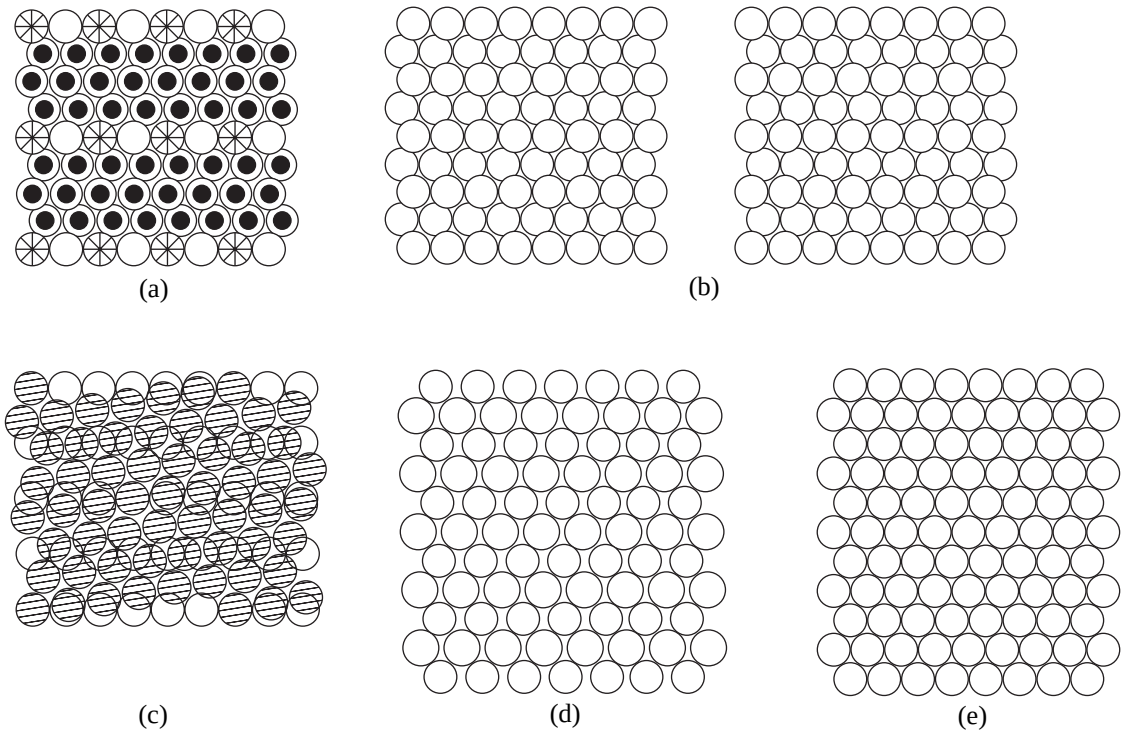


Fig. 4. Arrangement of atoms and molecules (large empty circles, W atoms; figured circles, CO molecules; hatched circles, W_2C molecules; small filled circles, CaO molecules). (a) CaO + CO film on W(112), $\Theta_W = 0.25$, $\Theta_{CO} = 0.25$, $\Theta_{CaO} = 1.5$; (b) clean W(112) face, two types of domains; (c) W_2C film on W(112), $\Theta_{W_2C} = 1.667$, one of the domain types; (d) clean W(110) face; and (e) W(110) face converted to the (111) face of the fcc lattice ((0001) face of the hcp lattice) on the cylindrical surface.

Θ_{CaO} in the uppermost layers of the substrate may be as high as 0.75. In this case, the LEED pattern will not change up to the reflection intensity. The presence of the mobile anomalous reflections indicates that the substrate surface is composed of moderately inclined faces. Therefore, the possibility of CaO molecules being formed in grooves between upper close-packed (CP) atom rows on the W(112) face (Fig. 4b) should also be taken into account. Then, with the same LEED pattern, Θ may reach 1.75. Such a high value may explain the intense sputtering of the film when CaO is desorbed even without considering its presence at defects not appearing in the LEED pattern.

The strange point in the above model of the (2×2) pattern is that W atoms in the CP rows must alternate (or even occupy alternate rows) (Fig. 4a) although the adsorbate (Ca) flux is unlimited. This implies the presence of a tungsten compound in which the spatial configuration of atomic and molecular orbitals prevents the adsorbate from being combined with each of the atoms in the row. CO as the additional adsorbate appears to be the most plausible.

If a surface atom of a transition metal (for example, in CP rows on the (110) and (112) faces of Group VIb elements) has a vacant *d* orbital, a CO molecule in the ground state (i.e., with triple C–O bonds and two occupied C and O orbitals opposite to the linear σ bond, which is the strongest in this triad [21]) can also produce a strong linear donor–acceptor bond via the occupied orbital of the C atom (CO combines with a metal only through the C atom [21]) and the vacant d_{z^2} orbital (i.e., the orbital directed normally to the surface). In other words, a MeCO subcarbonyl molecule may form. This molecule has one occupied O orbital directed normally (both toward and outward) to the substrate. This orbital also can produce a donor–acceptor bond with another ionized adsorbate, e.g., with cesium. Thus, being bonded with the metal in the ground state (in the excited state of a CO molecule, i.e., when one of the two π bonds is broken, Me–C–Me bridge bonds occur largely at defects or if surface atoms of the substrate lack vacant *d* orbitals), a CO molecule will always be directed normally to the surface. Obviously, such a molecule cannot be buried by an external W flux, e.g., from the TEC emitter. In other words, in the case of W mass transfer from the environment, the CO molecule will always "penetrate" into the upper layer of the W atoms. In view of its location on the substrate and the strong tendency to produce bonds to Cs^+, it becomes clear to what extent the work function of the CO-covered TEC collector subjected to the Cs vapor increases. Indeed, the CO contamination has the most adverse effect on the TEC collector [5].

From the difference in the linear sizes of CO and CaO molecules and taking into account the different nature of W–CO and W–CaO bonds, one can suppose that CO molecules, when incorporated into the CO–CaO film, will form ridges. This, in turn, will cause the formation of inclined facets occupied by CaO. In fact, CaO molecules, having strong double Ca–O bonds, can be attached to W atoms only via one of the two remaining occupied orbitals of the O atoms. They form donor–acceptor σ bonds via the vacant d_{z^2} orbital of the W atom. Then, double Ca–O bonds in the CaO molecule, and hence, the molecule itself, will arrange themselves at an angle to the substrate. Naturally, this will affect the packing density of the CP W atom rows (which produce bonds with CO molecules), especially in view of the asymmetry of the two π bonds about the linear σ bond in the ground state of CO. The polarization, rather than chemical, nature of the CaO–W bond is out of the question because of the high temperature (1375 K) of CaO desorption.

The two π C–O bonds, which are asymmetric about the σ C–O bond, must be arranged so that the remaining half-occupied d_{yz} orbital of the W atom in carbonyl WCO (this orbital is directed along a CP row at an angle to the substrate) is placed in the dihedral angle formed by the π bonds. Therefore, these π bonds must be equally directed relative to the σ bond and the CP row of W atoms. Because of this, the W atoms in the CP rows alternate when WCO molecules form.

Thus, if Ca and C (or W_2C) oxidize concurrently, there appears a high probability of a mixed CaO + CO film being formed, where WCO molecules create the framework with the (2×2) structure and $\Theta_{CO} = 0.25$. The value of Θ_{CaO} in this film equals 1.5. Note that the W surface atoms responsible for the coverage 0.25 and the (2×2) structure do not combine with any of the adsorbates (Fig. 4a).

Upon heating to 1475 K, the W(112) face gives the $c(3 \times 2)$ diffraction pattern due to carbide $W_2C(0001)$ (Fig. 3b). The LEED pattern taken from the (110) face shows no extra reflections. On the other hand, the LEED pattern taken from the (110) face of the flat specimen exhibits reflections typical of $W_2C(0001)$ [10, 22, 23]. That is why the assumption that only the W_2C amorphous film is produced on the W(110) inclined faces seems to be unjustified. Unfortunately, the AES method cannot distinguish the states of the W_2C carbide on the (112) and (110) faces of the cylindrical surface, since the relatively wide and, what is more, inclined beam of primary electrons covers both faces. This leads us to the conclusion that the carbon does not precipitate on the inclined W(110) faces forming terraces with the (112) faces as their base. This fact can be explained by the increased energy of an inclined (110) face compared with that of a (112) face (terrace base) and also by the presence of the boundaries (grooves) between the terraces. At the same temperature (1475 K), the O peaks also disappear (Fig. 2f); that is, at this temperature, the surface is totally free of CO.

As the temperature rises further (1525 K), the S peak in the Auger spectrum starts decreasing and completely disappears at 1625 K (Fig. 2g). Still higher temperatures, up to 2000 K, cause no changes in the Auger spectrum and the LEED pattern: the former exhibits the peaks of W and W_2C and the latter, the $c(3 \times 2)$ structure.

It should be noted that the Auger spectra from CO and W_2C differ only slightly; therefore, the associated peaks may superpose. It has been found, however, that the presence of at least the basic O peak at 512–516 eV is a good indication of CO on the surface. To verify this experimental finding in the presence of the C and O peaks in the spectrum, it is necessary to rapidly heat the specimen to 1175–1475 K and record the spectrum immediately after heating (preventing its cooling). In so doing, the C peak at 272 eV is sharply reduced and the O peaks disappear (Fig. 2h). This confirms the above conclusion that CO may be present on the surface when a part of the carbide is oxidized by the residual surface oxygen at moderate heating temperatures.

From the aforesaid, it follows that the CVD tungsten cylindrical surface can be cleaned of all the impurities but carbon even after the first heating *in vacuo*. Carbon combines with the substrate material at temperatures below 1475 K to form carbides.

It has been reported [2, 4, 10, 22, 23] that W_2C carbide dominates on any of the W crystallographic faces except (100) when single-crystal W reacts with C no matter how the latter was deposited (by segregation or from an external source), the most close-packed (0001) face of W_2C being parallel to the substrate. Hence, the same could be expected for the W(112) face.

When analyzing the undoubtedly moiré LEED pattern with $c(3 \times 2)$, we constructed 25 models of the structure with more or less realistic Θ (0.750, 0.883, 0.917, 1.083, 1.167, 1.250, 1.333, 1.583, 1.667, and 1.750) and with scattering centers variously arranged on the W(112) face. Basically, each of the models could be responsible for the LEED pattern observed. However, only two of them [with the reflection coordinates in the reciprocal lattice (i.e., on the LEED pattern) $(\bar{1}/3, 3/2)$, (1/6, 7/4), and (1/2, 1/4) for $\Theta = 0.883$ and $(\bar{1}, 1/2)$, $(\bar{1}/6, 7/4)$, and (5/6, 5/4) for $\Theta = 1.667$] proved to be close to the W_2C low-index faces, $(10\bar{1}0)$ and (0001), respectively. Taking into account the type of chemical bonds in W_2C molecules (three-center W–C–W bridges), the spatial orientation of bonding and antibonding orbitals in C and W atoms, the need for minimizing the energy spent on the rotation of W_2C molecule rows relative to CP rows of W atoms on the (112) face (that is, the rotation of W_2C rows must be as small as possible), and data from the works cited above

(specifically, the data for Θ_{W_2C}), we adopted the model with $\Theta = 1.667$ as the most realistic. For this model, the arrangement of W_2C molecules relative to the substrate is shown in Fig. 4c. The deviation from the perfect $W_2C(0001)$ face, for which the least spacing between W atoms is 0.299 nm [24], along the three principle lattice directions is 1.64, –2.58, and –6.05% (for the last two directions, the interatomic spacings in W_2C molecules are much smaller than on the perfect face). The deviation from Θ_{W_2C} on the perfect face (1.585 for W(112)) is 5.19% (the $W_2C(0001)$ face becomes closer packed). Upon forming the $W_2C(0001)$ film, the rotation of the CP substrate atom rows is 8°50′. Sites of coincidence between W atoms in the $W_2C(0001)$ film and on the W(112) substrate produce the $c(3 \times 2)$ coincidence site (coherent) lattice.

From the model selected, it follows that the W_2C film is formed not only by the uppermost CP W rows on the (112) face but also by the rows in the grooves. The latter are located between the former and are partially (by 5.72%) screened by them (Fig. 4b). Consequently, the formation of the carbide leads to the radical reconstruction of the W(112) face, causing the groove CP substrate atom rows to emerge on the surface. Similar effects have been observed upon the formation of Mo and W suboxides [2, 10, 20].

CARBON REMOVAL AND THE STRUCTURE OF THE CLEANED SURFACE

After 30 cycles of oxidation of the carbon-contaminated surface (by the method described above) and subsequent oxygen evacuation to a residual pressure $P_{res} \leq 10^{-8}$ Pa, followed by several rapid heatings to 1925 K, the Auger spectrum had no peaks other than the W peak (Fig. 2i). Under normal incidence of the beam, the LEED pattern is typical of the clean (112) face of the bcc lattice (Fig. 3c). Taking into consideration that the cylinder envelope was placed in the horizontal plane of the instrument and was at the same time normal to the optical axis of the central electron gun, as well as that the diffraction pattern (reciprocal lattice) was rotated through 90° relative to the atom arrangement in the direct lattice, we can conclude that the CP W atom rows on the (112) face are aligned with the cylinder envelope (Fig. 4b).

When the angle of beam incidence is varied in the plane normal to the cylindrical surface of the specimen, there comes a point where the LEED pattern from the (112) face is no longer observed. Instead, one can see the pattern from the clean (110) plane of the bcc lattice (Fig. 3d). The angle between the (110) and (112) in this lattice is $\pi/8$; that is, when emerging on the surface, these faces make a dihedral angle of $7\pi/8$. Thus, the clean cylindrical surface of single-crystal CVD tungsten consists of terraces with inclined (110) faces as walls.

However, at some intermediate angle of incidence when $\{h0\}$ and $\{0k\}$ reflections ($h = k = \pm1, \pm2, \pm3, \ldots$) in the pattern from the (112) face gradually disappear and the pattern changes to that from the (110) face, one can see that reflections of type $\{2h + 1, 2k + 1\}$ ($h = k = 0, \pm1, \pm2, \ldots$) in both patterns roughly coincide in the first order. However, in the directions [10] and $[\bar{1}0]$, the reciprocal lattice for the perfect (110) face must be 15.47% shorter than the reciprocal lattice for the (112) face (Fig. 3e). Therefore, it can be assumed that one of the faces on the imperfect substrate is distorted. Relying on the general considerations that a complex surface tends to have a relief minimizing its surface, we believe that the probability of distortion is higher for the terrace inclined walls. Their energy is higher than that of the terrace bases, and they have a variable potential relief because of the inclination. In other words, the W atom rows along the bcc lattice constant on the (110) face must be compressed to the close packing state in the first approximation.

The LEED patterns imply only two types of (110) face distortion. In one case, it can be assumed that, when the (110) and (112) faces join (along the grooves), the (110) face is reconfigured in such a way (Fig. 4d) that the directions of its CP rows of two types become coincident (in either domain) with the directions of first-order inclined rows on the (112) face. The inclined rows make an angle of 31°29′ with the CP rows on this face (Fig. 4b). In the other case, the (110) face may even be converted to the (111) face of the fcc lattice or, which is the same, to the (0001) face of the hcp lattice (Fig. 4e). In this case, the LEED (110) and (112) reflections along directions parallel to the *h* axis will coincide. Along directions perpendicular to the *h* axis (i.e., along those parallel to the *k* axis), the first-order reflections will deviate by 5.72%, compressing the LEED pattern from the (110) face relative to that from the undistorted (112) face. Visually, such a deviation is practically indiscernible for the first-order reflections. In this case, the joint of the (110) and (112) faces (terrace groove or ridge) will run along the mutual CP row of W atoms. In view of the considerable decrease in the energy in the groove because of the CP row formation, the latter case of inclined (110) face distortion seems to be more probable.

CONCLUSION

Thus, microscopically, the clean cylindrical surface of single-crystal CVD tungsten consists of terraces. Their bases are the (112) faces coplanar with the cylinder envelope. The walls are the (110) faces making an angle of 157°30′ with the base. The (110) face is distorted, i.e., slightly compressed (by 8.14%), in such a way that it is converted to the (111) face of the fcc lattice [(0001) face of the hcp lattice]. The joint of the distorted (110) and undistorted (112) faces (terrace groove or ridge) runs along mutual CP rows of W atoms.

In terms of impurity content in the near-surface layer (but not on the surface), CVD tungsten single crystals are highly competitive with zone-melted perfect tungsten single crystals, which were prepared according to the requirements for specimens to be studied by modern methods for surface investigation. The basic advantage of the material studied in this work is a relatively low content of carbon impurity—the most harmful impurity for emission electronics. This allows process engineers to obtain an atomically clean surface by applying conventional cleaning procedures. To decrease the surface contamination, especially, contamination by carbon (the clean surface facilitates subsequent purification of the near-surface layer), it would be desirable to formulate more stringent demands on the processing (vacuum) conditions, use hydrocarbon-free evacuation means, and remove carbon from the TEC electrodes by the technique described above directly in a TEC before its start-up.

REFERENCES

1. B. M. Zykov, V. P. Kobyakov, and Yu. I. Nardaya, Vysokochist. Veshchestva, No. 4, 125 (1991).
2. B. M. Zykov and Yu. I. Nardaya, Zh. Tekh. Fiz. **65** (4), 150 (1995) [Tech. Phys. **40**, 372 (1995)].
3. B. M. Zykov and V. K. Tskhakaya, Zh. Tekh. Fiz. **49**, 1700 (1979) [Sov. Phys. Tech. Phys. **24**, 948 (1979)].
4. B. M. Zykov, A. M. Sabel'nikov, and V. K. Tskhakaya, Poverkhnost, No. 6, 48 (1990).
5. B. M. Zykov, A. M. Sabel'nikov, V. K. Tskhakaya, *et al.*, Poverkhnost, No. 4, 65 (1983).
6. B. M. Zykov and V. K. Tskhakaya, Zh. Tekh. Fiz. **50**, 1771 (1980) [Sov. Phys. Tech. Phys. **25**, 1035 (1980)].
7. B. M. Zykov, A. M. Sabel'nikov, and V. K. Tskhakaya, Poverkhnost, No. 9, 22 (1990).
8. B. M. Zykov, A. M. Sabel'nikov, and V. K. Tskhakaya, Poverkhnost, No. 12, 21 (1986).
9. B. M. Zykov, Yu. I. Nardaya, and A. M. Sabel'nikov, Vysokochist. Veshchestva, No. 4, 116 (1991).
10. B. M. Zykov, V. P. Kobyakov, and Yu. I. Nardaya, Vysokochist. Veshchestva, No. 1, 71 (1991).
11. V. P. Kobyakov, Kristallografiya **41**, 552 (1996) [Crystallogr. Rep. **41**, 523 (1996)]; Kristallografiya **41**, 1093 (1996) [Crystallogr. Rep. **41**, 1040 (1996)].
12. S. R. Morrison, *The Chemical Physics of Surfaces* (Plenum, New York, 1977; Mir, Moscow, 1980).
13. M. Roberts and C. McKee, *Chemistry of the Metal–Gas Interface* (Oxford Univ. Press, Oxford, 1978; Mir, Moscow, 1981).
14. J. S. Blakemore, *Solid State Physics* (Cambridge Univ. Press, Cambridge, 1985; Mir, Moscow, 1988).
15. *Methods of Surface Analysis*, Ed. by A. W. Czanderna (Elsevier, New York, 1975; Mir, Moscow, 1979).
16. B. M. Zykov, V. K. Tskhakaya, and V. I. Yarygin, USSR Inventor's Certificate No. 1062803, Otkrytiya Izobret., No. 47, 236 (1983).
17. B. M. Zykov and A. M. Sabel'nikov, Prib. Tekh. Éksp., No. 1, 219 (1991).
18. J. A. Becker, E. J. Hecker, and R. G. Brandes, J. Appl. Phys. **32**, 411 (1961); Y. Viswanath and L. D. Schmidt, J. Chem. Phys. **58**, 4184 (1973).
19. B. M. Zykov, D. S. Ikonnikov, and V. K. Tskhakaya, Fiz. Met. Metalloved. **47**, 336 (1979).
20. B. M. Zykov and A. M. Sabel'nikov, Poverkhnost, No. 10, 61 (1988).
21. I. N. Semenov and K. V. Ovchinnikov, *"Unexpected" Inorganic Compounds* (Khimiya, Leningrad, 1972).
22. R. M. Stern, Appl. Phys. Lett. **5** (11), 218 (1964).
23. D. F. Ollis and M. Boudart, Surf. Sci. **23**, 320 (1970).
24. B. F. Ormont, *Structures of Inorganic Substances* (GITTL, Moscow, 1950).

Translated by V. Isaakyan

Technical Physics, Vol. 47, No. 6, 2002, pp. 752–759. Translated from Zhurnal Tekhnicheskoĭ Fiziki, Vol. 72, No. 5, 2002, pp. 101–108.

**SURFACES,
ELECTRON AND ION EMISSION**

Effect of the Electronic Structure of a Substrate on the Photoinduced Behavior of Adsorbed NO and CO Molecules

T. T. Magkoev* and Y. Murata**

** North-Ossetic State University, Vladikavkaz, 362020 Russia*
e-mail: magkoev@osetia.ru

*** Institute for Solid-State Physics, University of Tokyo, 7-22-1 Roppongi, Minato-ku, Tokyo 106, Japan*

Received October 30, 2001

Abstract—Photoprocesses in systems produced by adsorption of NO and CO molecules on the Pt(111) and Ni(111) surfaces, as well as on the (111) surface of Pt–Ge alloy, is studied by the IR absorption spectroscopy, resonant multiphoton ionization, and UV photoelectron spectroscopy methods. The energy of photons varies between 2.3 and 6.4 eV. The character of the processes depends on the type of the metallic substrate. On the Pt(111) surface, NO molecules dissociate or are desorbed, depending on the degree of coverage. On the Ni(111) surface, the molecules only dissociate. Conversely, NO molecules adsorbed on the (111) surface of the Pt–Ge alloy are only desorbed from the surface. In the CO/Pt(111) and CO/Pt(111)–Ge systems, CO molecules adsorbed on on-top adsorption sites are desorbed under the action of the photons, while those occupying bridging adsorption sites change their properties insignificantly. A model of photoinduced processes is suggested. According to this model, the lifetime of a state excited by charge transfer between the valence band of the metal and the 2π-antibonding molecular orbital plays a decisive part in the occurrence of one or the other of these processes. © *2002 MAIK "Nauka/Interperiodica".*

INTRODUCTION

In recent years, desorption of atoms and molecules from the solid surface that is induced by electron transitions (DIET) in adsorbate–substrate systems has been the subject of much investigation [1–6]. The aim of researchers has been (and remains to be) to obtain further insight into the adsorbate dynamics on the substrate surface. The elucidation of DIET mechanisms would make it possible to tackle a number of fundamental problems, such as the nature of the chemisorptive bond in the ground and excited states, and to find a mechanism whereby the excitation energy of the electronic subsystem is converted to the energy of the translational, rotational, and vibrational motions of atoms and molecules being desorbed [3–6]. Such studies are also of great applied interest because they uncover mechanisms of atomic and molecular transformations on the solid surface, allowing for the control of these transformations in applications [4].

The Antoniewicz model [7] is often used in describing particle desorption due to electron transitions in adsorbate–substrate systems. It assumes that the shape of the potential energy surface (PES) and the lifetime t of an excited adsorbate–metal (A–M) complex (Fig. 1) are basic factors specifying the type of particle desorption. The probability of an adsorbed particle being desorbed rises with increasing lifetime of the excited complex. Eventually, the particle leaves the surface, having gained an energy exceeding the energy of adsorbate–substrate bonds. On the metal surface, however, electron excitations in the valence band relax very quickly and the lifetime is short (about several femtoseconds). Because of this, the disintegration of the excited A–M complex is accompanied by recombination. For polyatomic molecules, dissociation, along with desorption and recombination, may also take place. Like desorption, dissociation implies bond breaking on the surface. Therefore, the study of this process might clarify DIET mechanisms.

The dynamics of adsorbed particles is usually studied with laser radiation that generates photons in the visible and near-UV parts of the spectrum [5, 8–10]. In this case, a valence electron transition responsible for a specific photoinduced response of the adsorption system can be identified. This makes it possible to selectively activate or inhibit specific surface reactions, which cannot be done with thermal excitation, as well as with electron- or ion-induced excitations. Most works studying the energy distribution over the translational, rotational, and vibrational degrees of freedom of molecules being desorbed use laser spectroscopy of atoms and molecules in the gas phase, the method of resonant multiphoton ionization (RMPI) being the most extensively employed [9]. In recent years, this method has become common in studying the dynamics of NO and CO desorption from a number of substrates [5, 8, 10–12]. The choice of these molecules as an object of investigation has been dictated by their practical impor-

1063-7842/02/4706-0752$22.00

tance in heterogeneous catalysis. In addition, they offer a high chemisorptive efficiency with respect to most adsorbents; hence, related adsorption systems can be taken as models for studying the fundamental laws of the interaction of these and other molecules with solid surfaces.

In this work, we studied photoinduced processes in NO/Pt(111), NO/Ni(111), CO/Pt(111), and CO/Ni(111) systems, as well as on the (111) surface of Pt–Ge alloy (hereafter, Pt(111)–Ge alloy), using the methods of IR absorption spectroscopy (IRAS), RMPI, and UV photoelectron spectroscopy (UVPS). From vibrational spectra of the particles on the surface and the energy distribution of molecules being desorbed over degrees of freedom, we develop a model of these processes. They are shown to depend on the substrate metal and on the type of adsorption site. According to theoretical considerations [13], the *d* band of Pt is filled if a small amount of Ge (4% in our case) is added. Therefore, we tried to trace how much a minor change in the electronic properties of the substrate influences the response of the systems to the irradiation. We have found that the process type depends on the intramolecular vibration frequency ν of CO and NO. The particles with higher ν tend to be desorbed because the electrons transit to the unfilled 2π orbital of the molecule. For the particles with lower ν, photodissociation prevails. Also, the type of photoprocess (desorption, dissociation, or recombination) depends markedly on the excited state lifetime, as well as on the intramolecular bond strength and the strength of adsorbate–substrate bonds.

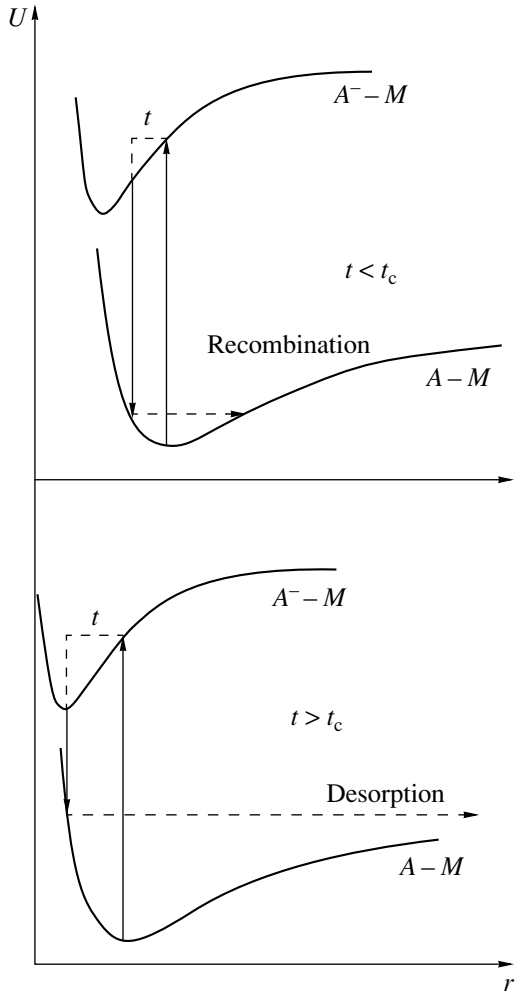


Fig. 1. Antoniewicz desorption model for an adsorbed particle [7]. If the lifetime of an excited state *t* exceeds the critical value t_c ($t > t_c$), the de-excited particle can pass through the potential barrier of attraction to the substrate and be desorbed. Otherwise, recombination takes place. The transitions between the ground and excited states are marked by arrows.

METHODS OF INVESTIGATION

Experiments were carried out in an ultra-high-vacuum chamber designed for IRAS and RMPI investigation and equipped with Auger electron spectroscopy (AES) and low-energy electron diffraction (LEED) facilities. UVPS studies of the Pt(111)–Ge alloy were performed in another ultra-high-vacuum chamber with a synchrotron radiation source. Photoelectron spectra were recorded by means of a double-focussing cylindrical reflector analyzer. An incident UV beam makes an angle of 48° with the normal to the sample surface and 90° with the axis of the analyzer. The residual gas pressure in both chambers was no higher than 2×10^{-10} torr. IR spectra of adsorbed molecules were recorded with a Nicolet-Nexus 870 Fourier-transform spectrometer with a resolution of 4 cm^{-1}. The monochromator and detector of the spectrometer were arranged so that the incident and reflected IR beams polarized in the incidence surface made an angle of 80° with the normal to the sample surface. Such a geometry provides the highest sensitivity to intramolecular vibrations of the adsorbates, since the axes of the CO and NO molecules are aligned with the normal to the sample surface. The processes under study were initiated by the radiation from a Lambda Physics pulsed excimer laser with a pulse repetition rate and duration of 10 Hz and 8 ns, respectively. This radiation generates photons with a wavelength of 193, 248, 352, and 532 nm. The energy density of the laser beam was 1–3 mJ/cm^2 for an angle of incidence of 60°. The distribution of the molecules being desorbed was taken with a Spectra Physics tunable dye laser. In this case, the probing beam ran parallel to the sample surface at a distance of 2 mm from it. The particles were detected by the RMPI method using the standard (1 + 1) scheme to excite the $A^2\Sigma^+ \leftarrow X^2\Pi$ transitions for NO and using the (2 + 1) scheme to excite the $B^1\Sigma^+ \leftarrow X^1\Sigma^+$ transitions for CO.

The Pt(111)–Ge alloy was prepared by evaporating a Ge monolayer on the Pt(111) surface with subsequent annealing at 1100°C. As follows from the AES data, the concentration of Ge in the surface layer of the alloy thus

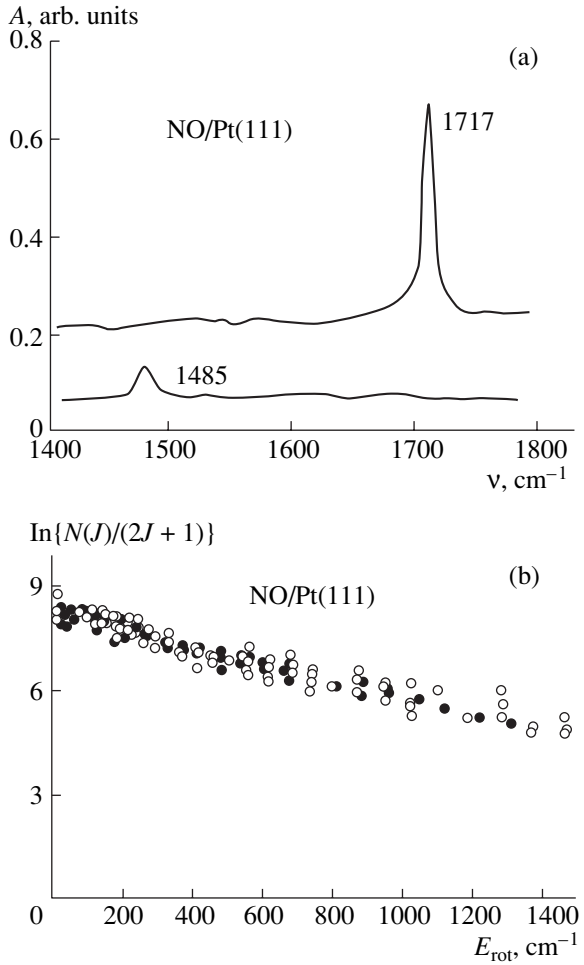


Fig. 2. (a) IR spectra taken from the NO/Pt(111) system for exposures of 0.4 and 3.0 L (the latter provides the full coverage of the substrate kept at 90 K). (b) Distribution of NO molecules photodesorbed from the Pt(111) surface over rotational energies (full coverage). Filled circles, spin-orbital with the quantum number $\Omega = 1/2$; empty circles, spin-orbital with $\Omega = 3/2$. The desorbent photon energy is 6.4 eV.

prepared corresponds to 0.1 of a monolayer. The properties of the alloy were found to be stable against subsequent thermal treatments. Recent STM studies of this alloy have indicated the presence of the (5×5) superstructure, which suggests that the concentration of Ge in the first atomic layer of the Pt(111) surface is 4% [14]. The surfaces of the samples were cleaned by conventional annealing in oxygen and hydrogen, followed by rapid thermal annealing at 800°C for Ni(111) and 1100°C for Pt(111) and Pt(111)–Ge. The samples were then cooled to 90 K, and the NO or CO gases to be adsorbed were delivered to the chamber up to a pressure of 10^{-7} torr. All the IRAS and RMPI measurements were performed at a sample temperature of 90 K (for details, see [14–19]).

EXPERIMENTAL RESULTS

Figure 2a shows the IR absorption spectra for the NO/Pt(111) system at a substrate temperature of 90 K and exposures of 0.4 and 3.0 L (1 L = 10^{-6} torr s). When the substrate coverage by the NO molecules is low, an adsorption band due to intramolecular vibrations of adsorbed NO appears at $\nu = 1485$ cm^{-1}. As the coverage grows, this band shifts to larger wave numbers and its intensity decreases. At the same time, a new absorption band at large ν appears. When the surface is fully covered by NO, the spectrum has a single absorption band at $\nu = 1717$ cm^{-1}. It is assumed that these absorption bands are associated with intramolecular vibrations of NO molecules occupying dissimilar adsorption sites. The lower frequency band corresponds to NO molecules occupying bridge adsorption sites, while the higher frequency one is associated with NO molecules occupying on-top adsorption sites [20–23]. Such an interpretation of the IR spectra, which was accepted long ago, has become the subject of furious debates. The reason is that, according to recent structure investigations, NO molecules, when being adsorbed on the Pt(111) surface, occupy largely octahedral and tetrahedral voids [24–26]. Because of this uncertainty, we will not assign spectrum features to any of the adsorption centers. Instead, NO molecules adsorbed will by convention be referred to as "low-frequency" and "high-frequency" for the sake of convenience.

For the photon energy 6.4 eV, the process occurring in the NO/Pt(111) system depends on the substrate coverage. At low coverages, the photons decrease the intensity of the lower frequency absorption band. Since the desorption signal is absent in this case, one can suppose that the NO molecules adsorbed dissociate under the action of the radiation. When the surface is fully covered by NO, the intensity of the related absorption band (1717 cm^{-1}) also decreases under irradiation but a relatively high RMPI signal is observed in this case; that is, the adsorbate is desorbed. The distribution of the NO molecules being desorbed over rotational degrees of freedom that is derived from the RMPI spectra is shown in Fig. 2b on the semilogarithmical scale. The distribution is seen to be linear, with its slope corresponding to the NO desorption rotational temperature 490 K. Since the substrate temperature is 90 K, this may indicate that NO molecules desorbed from Pt(111) under the action of photons are rotationally excited.

For the CO/Pt(111) system, as well as for NO/Pt(111), two absorption bands are observed. For the fully covered surface, their peaks lie at 1855 and 2105 cm^{-1}. These bands are assigned to the molecules occupying bridging adsorption sites and on-top adsorption sites, respectively. It should be noted that, unlike NO/Pt(111), such an interpretation here has been supported by recent investigations with structure-sensitive methods. The irradiation of CO/Pt(111) with an energy of 6.4 eV decreases the intensity of the high-frequency band, while the low-frequency one does not undergo

noticeable changes. In this case, as follows from the RMPI spectroscopic data, the CO molecules are desorbed. The related RMPI spectrum is depicted in Fig. 3b. The data points are well approximated by the model spectrum (solid curve) for molecules with a rotational temperature of 130 K. Thus, one can suppose that CO molecules desorbed are not rotationally excited, unlike the NO molecules, which have a rotational temperature of 490 K.

The IR spectrum from the NO/Ni(111) system (adsorption at a substrate temperature of 90 K) has a single absorption band corresponding to intramolecular vibrations of adsorbed NO. As the surface concentration of NO (coverage) increases, the position of the peak tends to larger wave numbers, reaching 1585 cm^{-1} for 100% coverage (Fig. 3a). By way of example, Fig. 3a demonstrates the IR spectra taken from the NO/Ni(111) system at exposures of 0.5 and 3.0 L. For the full coverage of the surface, the effect of the photons is negligibly small, whereas for smaller coverages, the intensity of the related absorption band declines and a new peak at 1820 cm^{-1} arises. Similar absorption lines in the high-frequency (above 1800 cm^{-1}) part of the spectrum are also observed when NO molecules are adsorbed on the Ni(111) surface with preadsorbed O and N atoms. Therefore, one can assume that the photons cause the NO molecules to dissociate when the coverage is low, while the properties of the NO/Ni(111) system remain virtually the same for the full coverage. It should be noted that photon energies used in this work do not all trigger the photoinduced processes observed. From the dependences of the process type on the photon energy E, it can be concluded that there exist threshold energies E_{th} for each of the processes. The table correlates E_{th} for desorption and dissociation of NO and CO molecules on the Pt(111) and Ni(111) surfaces with the associated frequencies of intramolecular vibrations.

For the NO/Pt(111)–Ge system, the IR spectrum does not contain the low-frequency band at any coverage, unlike the NO/Pt(111) system. The corresponding spectra for exposures of 0.4 and 3.0 L are demonstrated in Fig. 4a. As the coverage grows, the absorption line becomes more intense and shifts toward larger wave numbers, reaching 1716 cm^{-1} for 100% coverage. The spectrum is similar to that for the NO/Pt(111) system in shape, while differing in intensities and half-widths of the peaks. The photon irradiation lowers the intensities of the peaks, which is due to NO molecule desorption, in accordance with the RMPI data. Like for the NO/Pt(111) system, Fig. 4b shows the distribution of NO molecules desorbed from the Pt(111)–Ge surface over rotational energies. The distribution can be considered as linear, with its slope corresponding to the rotational temperature 410 K for the photodesorbed molecules. It is seen that the distributions of the molecules desorbed from the Pt(111) and Pt(111)–Ge surfaces (Figs. 2b, 4b) agree qualitatively.

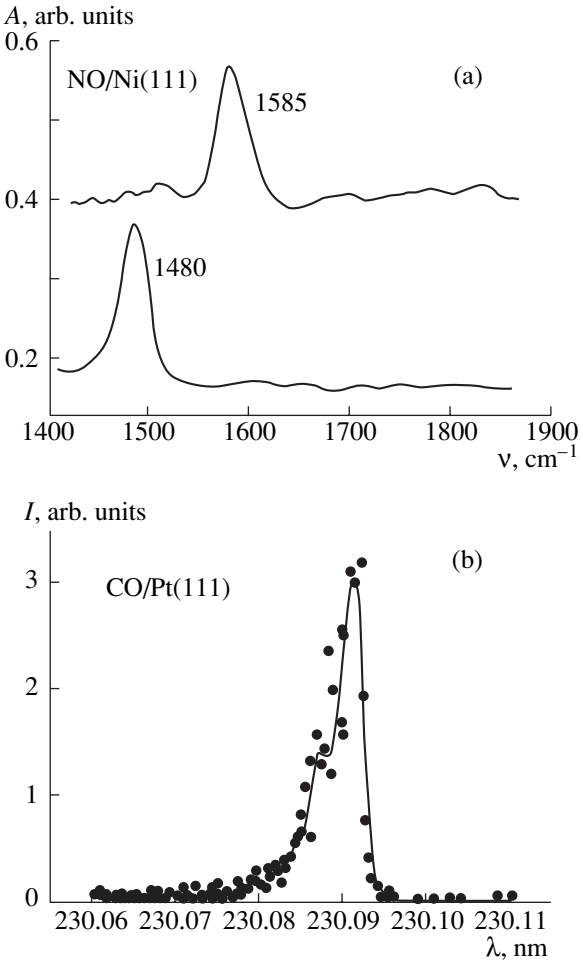


Fig. 3. (a) IR spectra taken from the NO/Ni(111) system for exposures of 0.5 and 3.0 L (the latter provides the full coverage). (b) Experimental RMPI spectrum for CO molecules photodesorbed from the Pt(111) surface by photons of energy 6.4 eV (circles) and the corresponding RMPI spectrum simulated with regard for the rotational temperature of the desorbed molecules (130 K) (solid line).

DISCUSSION

To discuss the results obtained, it is reasonable to invoke the well-known Blyholder electronic model of bonding between NO (CO) molecules and metallic substrates [27], which was supported by subsequent simulations [28, 29] (Fig. 5). For NO and CO, molecular orbitals responsible for chemisorptive bonding with the surface are the 5σ and 2π orbitals, which are of bonding and antibonding nature, respectively, in a free molecule. According to the Blyholder model, these two orbitals produce related bonding and antibonding orbitals when interacting with *d* electrons of the substrate. It is assumed here that the electron density is transferred from the 5σ orbital to the conduction band of the metal and from the conduction band to the 2π-antibonding orbital. According to UVPS data, the occupied bonding state $5\sigma_b$ of a chemisorbed CO molecule and two occu-

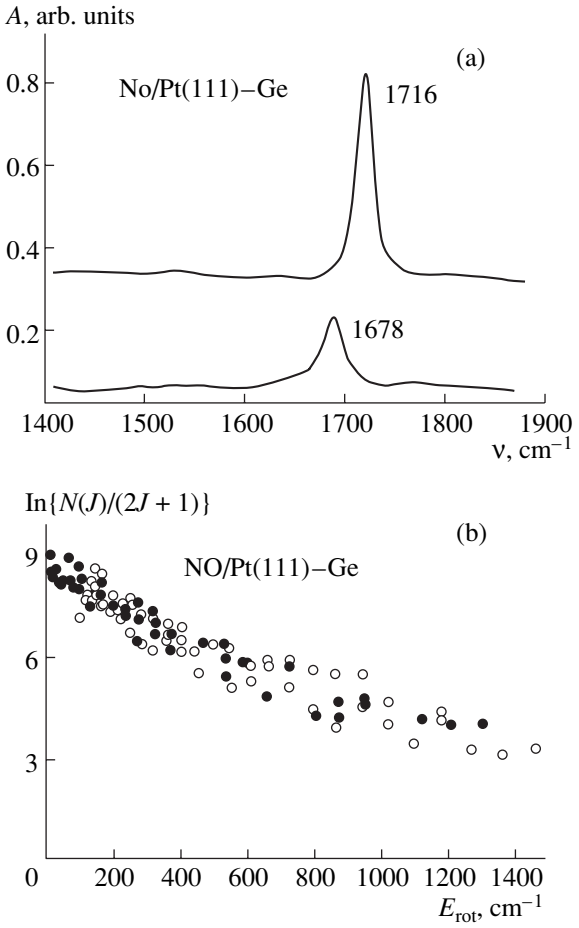


Fig. 4. (a) IR spectra taken from the NO/Pt(111)–Ge system for exposures of 0.4 and 3.0 L. (b) Distribution of NO molecules photodesorbed from the Pt(111)–Ge surface over rotational energies. Ω = 1/2 (filled circles) and 3/2 (empty circles).

pied bonding states, $5\sigma_b$ and $2\pi_b$, of an NO molecule lie at 9.2, 9.6, and 2.7 eV, respectively, below the Fermi level [30, 31]. At the same time, inverse UVPS data indicate that for a CO molecule, there exist two unfilled states lying 1.5 and 4.0 eV above the Fermi level (they are assigned to the $5\sigma_a$ and $2\pi_a$ antibonding orbitals) and one unfilled state for an NO molecule (1.5 eV above the Fermi level), which is associated with the $2\pi_a$ antibonding orbital [31–33]. The energy width of the *d* band of Pt and Ni is 8 and 5 eV, respectively [34, 35]. It should be noted that the energy of transition between the above levels somewhat differs from that measured by the UVPS method because of the final-state effects. These effects result in under- and overestimated binding energies for occupied and unoccupied levels, respectively.

Threshold photon energies E_{th} for the NO and CO photodesorption and photodissociation and the intramolecular vibration frequencies typical of these processes on the Pt(111) and Ni(111) surfaces

System	E_{th}, eV	Photoprocess	ν, cm^{-1}
NO/Pt(111)	2.3	Desorption	1716
CO/Pt(111)	2.3–3.5	"	2105
NO/Pt(111)	5.0	Dissociation	1485
NO/Ni(111)	3.5	"	1480

An electron transition responsible for a specific photoprocess can be found from the tabulated data. The threshold photon energies listed seem to be insufficient for exciting electron transitions between the levels of the molecules adsorbed. To explain the fact that the NO and CO molecules are desorbed at photon energies of 2.3–6.4 and 3.5–6.4 eV, respectively, one has to assume that the photons induce the electron transition from the *d* band of the metal to the unoccupied $2\pi_a$ state, which is behind the photodesorption of both molecules. One could suppose that, for CO photodesorption, the $5\sigma_a$ state, which lies 1.5 eV above the Fermi level E_f (like the $2\pi_a$ orbital for NO), plays a crucial role. However, the fact that the photon energy intervals inducing NO and CO desorptions are not the same (see the table) is against this supposition.

As was indicated above, the behavior of the NO/Pt(111) and NO/Ni(111) systems under the photon irradiation is different for the full surface coverage. In the former case, the molecules are desorbed, while in the latter, the properties of the system change insignificantly. This can be explained by the different effect of the substrate electronic structure on the photoexcitation. Platinum and nickel are isoelectronic transition metals, and their electronic configurations are qualitatively similar. The most pronounced discrepancy is in the energy width of the *d* band: 8 and 5 eV for Pt and Ni, respectively [34, 35]. According to the Antoniewicz model, the difference in the photoprocesses may be due to the different geometry of the PESs in the ground and excited states and to different lifetimes *t* of the excited state. Assuming that the PES geometries for Pt and Ni are similar because of the similarity of their electronic configurations, one can argue that the parameter *t* governs the type of the photoprocesses.

Let us see how the different widths of the *d* band lead to different values of the excited state lifetime. First, we note that the *d* band of Ni is narrower than that of Pt; hence, when they are equally filled, the center of the *d* band in Ni lies closer to the Fermi level. As a result, the interaction between the substrate and the 2π level of the molecule for Ni is stronger than for Pt, because the energy spacing between the 2π level and the center of the *d* band is smaller in the former case [28]. On the other hand, inverse UVPS data indicate that the interaction of the substrate with the 5σ state for Pt is stronger than for Ni [33]. The lifetime of an excited state depends on the force of its interaction with the substrate or on the degree of its localization: if the interaction becomes stronger or the state is delocalized,

the lifetime decreases. As was noted above, the desorption of NO and CO is due to the photon-induced transition of electrons from the metal to the unoccupied 2π orbital of the molecule. Since the interaction between the *d* band and the 2π state for Pt is weaker than for Ni, one can conclude that the lifetime of the related excited state in Pt is larger. That is why the NO molecules are desorbed from the Pt(111) surface and remain on the Ni(111) surface under the action of the photons.

Further evidence in favor of such an interpretation is IR spectroscopy data. According to the Blyholder model, when the molecule is adsorbed on the metal surface, the electron density is transferred from the occupied antibonding 5σ orbital of the molecule to the *d* band of the metal and from the *d* band of the metal to the antibonding 2π orbital of the molecule. The interaction of the latter orbital with the substrate is stronger; hence, the electron density on this orbital is higher and intramolecular bonds weaken. As a result, the intramolecular vibration frequency decreases. At the same time, the interaction of the 5σ orbital with the substrate must not have a considerable effect on the intramolecular bond energy because of its antibonding nature. Note that among orbitals producing chemisorptive bonding, only the 2π orbital weakens intramolecular bonds; hence, a higher frequency of interatomic vibrations in the adsorbed molecule means a weaker interaction of the 2π state with the substrate. This, in turn, increases the lifetime of the related excited state and thus explains the different responses of the NO/Pt(111) and NO/Ni(111) systems to the photon irradiation.

The above considerations also help to explain the dependence of the process on the adsorption site occupied by the molecules. It was indicated that only high-frequency particles respond to the photon action on the NO/Pt(111) and CO/Pt(111) systems. In the case of CO, these are molecules occupying on-top adsorption sites. Such a specific selectivity can be explained if it is remembered that the excited state PESs for the on-top and bridge configurations are almost identical [36]. According to calculations, the PES shape is not responsible for the selectivity. The selectivity effect can be considered in terms of excited state lifetime *t* under the assumption that *t* for high-frequency molecules is larger than for low-frequency ones. The simulation of CO chemisorption on Pt atom clusters counts in favor of this approach. Its results indicate that the antibonding 5σ orbital makes a major contribution to chemisorptive bonding, while the 2π orbital has a minor effect, in the on-top adsorption geometry [37]. Moreover, unlike CO molecules occupying on-top sites, the 2π state of a bridging CO molecule splits into two sublevels, which reflects the stronger interaction of the 2π orbital with the substrate [37]. Consequently, it can be assumed that the lifetime of the excited state associated with the 2π level in the case of on-top adsorption (high-frequency molecules) is larger than for bridge adsorption (low-frequency molecules).

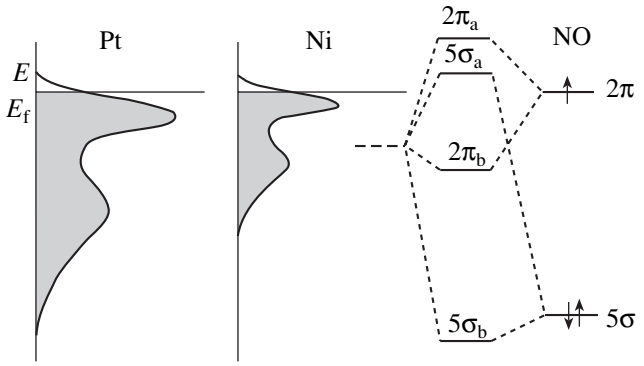


Fig. 5. Energy diagram for the interaction of an NO molecule with the Pt and Ni surfaces. The 5σ and 2π molecular orbitals produce the related bonding, $5\sigma_b$ and $2\pi_b$, and antibonding, $5\sigma_a$ and $2\pi_a$, orbitals when interacting with the valence band of the metal.

Further evidence for the decisive role of the 2π state in the mechanism of the photoinduced processes is apparent from comparing the results obtained for the systems with Pt(111) and Pt(111)–Ge used as substrates. As was noted above, the chemisorptive properties of Pt(111)–Ge differ greatly from those of Pt(111) [16, 18]. Namely, the energy of NO and CO adsorption on Pt(111)–Ge is appreciably lower than on Pt(111). Also, low-frequency molecules observed on Pt(111) are absent on Pt(111)–Ge. Nevertheless, in spite of these differences in the ground state, the distributions of the desorbed molecules over rotational energies for both substrates are qualitatively similar (Figs. 2b, 4b). Figure 6 compares the UVPS spectra for Pt(111) and Pt(111)–Ge, from which it follows that the basic change in the Pt electronic structure upon alloying with Ge is the increase in the occupancy of the *d* band. Because of this, the density of states in the *d* band that are vacant for adsorption decreases and, hence, the charge transfer from the 5σ level of the molecule to the *d* band of the metal markedly decreases in comparison with the pure Pt(111). This allows us to explain the fact that the binding energy of the NO and CO molecules on the Pt(111)–Ge surface is lower than on Pt(111). At the same time, the relative increase in the occupancy of the *d* band of the Pt–Ge alloy is insignificant and has a minor effect on the reverse charge transfer from the *d* band of the metal to the antibonding 2π orbital of the molecule. The fact that, upon adsorption, the electron densities on this orbital are close to each other for both substrates is confirmed by the closeness of the associated intramolecular vibration frequencies (Figs. 2a, 4a). This and also the similarity of the distributions of the molecules desorbed from Pt(111) and Pt(111)–Ge under the photon irradiation over rotational energies suggest that the 2π level of the molecules governs the photodesorption mechanism.

Another photoprocess observed in the systems considered is photodissociation. Note that only the low-fre-

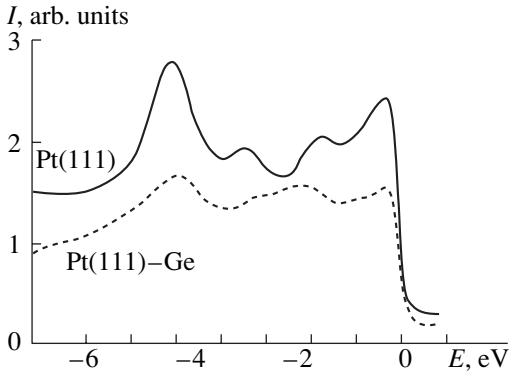


Fig. 6. UVPS spectra for the Pt(111) and Pt(111)–Ge surfaces. The photon energy is 36 eV. The value $E = 0$ corresponds to the Fermi level.

quency particles dissociate on both Pt(111) and Ni(111). As follows from the table, the threshold energy of NO photodissociation is substantially higher than the threshold energy of photodesorption. Therefore, we can suppose that the electronic mechanisms behind the two processes are different. However, the inverse UVPS data for the NO/Pt(111) and NO/Ni(111) systems do not reveal unoccupied states other than $2\pi_a$, so that the electronic mechanism of photodissociation cannot be identified with certainty. Nonetheless, based on the available data, we can make some qualitative suppositions. The absence of the photodesorption signal when the photons act on the low-frequency particles indicates that the photodissociation competes with the recombination of the excited molecules into the potential well of the adsorbate–substrate PES. Within the Antoniewicz model, this is a consequence of the short lifetime of the excited state. In its turn, the lifetime is short because the interaction between the 2π antibonding orbital and the *d* band of the metal for the low-frequency molecules is stronger than for the high-frequency ones. The lower intramolecular vibration frequency means a larger amount of charge on the 2π antibonding orbital, hence, the stronger interaction of this level with the substrate and the smaller lifetime of the excited state. On the other hand, as the charge on the 2π antibonding orbital grows, intramolecular bonds loosen. This may cause the molecules to photodissociate because the related activation barrier is reduced. Results [5], according to which the photodissociation cross section increases with loosening intramolecular bonds for molecules in the gas phase, validate this assumption. Since the photodissociation takes place at low coverages, one can suppose that the accommodation of dissociation products on the surface is also of importance here. The significance of the accommodation follows, in particular, from the following observation. When the NO/Ni(111) system with a totally covered surface is heated, the molecules are first desorbed in amounts providing the exposure of adsorption sites needed for the subsequent accommodation of dissociation products of the remaining molecules [21, 22].

Thus, we conclude that the photoprocesses in the NO(CO)/Pt(111), NO(CO)/Ni(111), and NO(CO)Pt(111)–Ge adsorption systems are governed by the type and surface concentration of the molecules and by the type of substrate. It has been shown that the 2π antibonding orbital is crucial in the mechanisms behind these processes. As the orbital interacts with the substrate, the photoinduced processes change in the order desorption–recombination–dissociation.

REFERENCES

1. H. Guo, P. Saalgranck, and T. Seideman, Prog. Surf. Sci. **62** (9–10), 239 (1999).
2. V. N. Ageev, Prog. Surf. Sci. **47** (1–2), 55 (1994).
3. D. Teillet-Billy, J. P. Gauyacq, and M. Persson, Phys. Rev. B **62**, 13306 (2000).
4. X.-Y. Zhu, Surf. Sci. **390**, 224 (1997).
5. F. M. Zimmermann and W. Ho, Surf. Sci. Rep. **22** (4–6), 127 (1995).
6. P.-T. Howe and H.-L. Dai, Surf. Sci. **451** (1), 12 (2000).
7. P. R. Antoniewicz, Phys. Rev. B **21**, 3811 (1980).
8. S. A. Buntin, L. J. Richter, D. S. King, *et al.*, J. Chem. Phys. **91**, 6429 (1989).
9. X.-L. Zhou, X.-L. Zhu, and J. M. White, Surf. Sci. Rep. **13** (1-2), 73 (1991).
10. H.-L. Dai and W. Ho, *Laser Spectroscopy and Photochemistry on Metal Surfaces* (World Scientific, Singapore, 1995).
11. T. Kluner, H.-J. Freund, and V. Staemmler, Phys. Rev. Lett. **80**, 5208 (1998).
12. G. Eicchorn, M. Richter, K. Al-Shamery, *et al.*, J. Chem. Phys. **111**, 386 (1999).
13. K. Terakura, J. Phys. F **7**, 1773 (1977).
14. K. Fukutani, Y. Murata, T. T. Magkoev, *et al.*, Surf. Sci. **464**, 48 (2000).
15. A. Peremans, K. Mase, and Y. Murata, Phys. Rev. B **47**, 4135 (1993).
16. T. T. Magkoev, K. Fukutani, M.-B. Song, *et al.*, Surf. Sci. **363**, 281 (1996).
17. K. Fukutani, M.-B. Song, and Y. Murata, J. Chem. Phys. **103**, 2221 (1995).
18. K. Fukutani, T. T. Magkoev, Y. Murata, *et al.*, J. Electron Spectrosc. Relat. Phenom. **88–91**, 597 (1998).
19. T. T. Magkoev and Y. Murata, Phys. Status Solidi B **223**, R3 (2001).
20. C. T. Campbell, G. Ertl, and J. Segner, Surf. Sci. **115**, 309 (1982).
21. H.-P. Steinruck, C. Schneider, P. A. Heinmann, *et al.*, Surf. Sci. **208**, 136 (1989).
22. J. C. Chen, W. Erley, and H. Ibach, Surf. Sci. **224**, 215 (1989).
23. P. Fouquet, A. P. Graham, and G. Witte, J. Chem. Phys. **112**, 7600 (2000).
24. L. D. Mapledoram, M. P. Bessent, and D. A. King, Chem. Phys. Lett. **228**, 527 (1994).

25. N. Materer, A. Barbieri, D. Gardin, *et al.*, Phys. Rev. B **48**, 2859 (1993).
26. R. Lindsay, A. Theobald, T. Giessel, *et al.*, Surf. Sci. **405**, L566 (1998).
27. G. Blyholder, J. Phys. Chem. **68**, 2772 (1964).
28. A. Ruban, B. Hammer, P. Stoltze, *et al.*, J. Mol. Catal. **115**, 421 (1997).
29. H. Aizawa and S. Tsuneyuki, Surf. Sci. **399**, L364 (1998).
30. H. P. Bonzel and G. Pirug, Surf. Sci. **62**, 45 (1977).
31. S. Ferrer, K. H. Frank, and B. Reihl, Surf. Sci. **162**, 264 (1985).
32. V. Dose, Surf. Sci. Rep. **5**, 337 (1985).
33. G. Rangelov, N. Memmel, E. Bertel, *et al.*, Surf. Sci. **251/252**, 965 (1991).
34. V. N. Smith, Phys. Rev. B **9**, 1365 (1974).
35. D. A. Papaconstantopoulos, *Handbook of the Band Structure of Elemental Solids* (Plenum, New York, 1986).
36. H. Aizawa and S. Tsuneyuki, Surf. Sci. **363**, 223 (1996).
37. S. Ohnishi and N. Watari, Phys. Rev. B **49**, 14619 (1994).

Translated by V. Isaakyan

Technical Physics, Vol. 47, No. 6, 2002, pp. 760–766. Translated from Zhurnal Tekhnicheskoĭ Fiziki, Vol. 72, No. 5, 2002, pp. 109–114.

EXPERIMENTAL INSTRUMENTS AND TECHNIQUES

Numerical Simulation of Silane Decomposition in an RF Plasma

V. I. Strunin, A. A. Lyakhov, G. Zh. Khudaĭbergenov, and V. V. Shkurkin

Omsk State University, pr. Mira 55, Omsk, 644077 Russia

e-mail: strunin@omsu.omskreg.ru

Received August 3, 2001

Abstract—A model of silane decomposition in a radio-frequency argon plasma is constructed. The concentrations of SiH_4 decomposition products, as well as products of synthesis (higher silanes), are calculated. The role of metastable argon atoms in the formation of SiH_3 radicals and the higher silanes is analyzed. © *2002 MAIK "Nauka/Interperiodica".*

INTRODUCTION

Silane decomposition in a radio-frequency (HF) plasma is a conventional technique used in the production of a-Si:H films, which are usually deposited in PECVD reactors. Recently, a breakthrough in jet plasma-chemical methods has been made. These differ in the ways of activating gaseous reagents. For example, in the electron-beam activation technique [1], a beam of electrons with an energy of several kiloelectronvolts interacts with a free jet of a working gas that passes through a supersonic nozzle. The electrons cause the ionization and dissociation of the SiH_4 molecules with the subsequent deposition of the decomposition products on the substrate. Another technique [2] implies the preliminary decomposition of a silicon-containing gas in a capacitive HF discharge (like in a PECVD reactor) and the subsequent transport of the decomposition products from the discharge region into the vacuum chamber through a nozzle. In this case, the silane decomposition in the discharge region is accompanied by various interradical reactions leading to the formation of disilanes and polysilanes, which change the plasma composition and the amount of film-forming agents. Information on the concentrations of the working gas components and on their behavior under various discharge conditions assists in optimizing the plasma-chemical processes and thus improving the quality of the films.

NUMERICAL MODEL

There is experimental evidence that SiH_3 radicals make the most essential contribution to the formation of device-grade silicon films. Therefore, the simulation of the production and loss of this radical in a plasma are of special interest. It should be noted that the validity of a model to a great extent depends on the plasma components being included. Our model considers the components listed in Table 1 and takes into account the variation of the silane concentration, as well as the transition of the radicals and silane to higher silanes and dust Si_yH_n ($y > 5$).

The chemical kinetics of the processes in a gas discharge can be described by the equations

$$\frac{\partial N_i}{\partial t} + u\nabla N_i = \nabla(D_i\nabla N_i) + \sum_j N_j(n_e K_{ji} + \sum K_{jli}N_l) - N_i\sum_j\left(n_e K_{ij} + \sum_l K_{jli}N_l\right), \quad (1)$$

where N_i is the concentration of the ith component,

Table 1. Parameters of Lenard-Jones potential

Component	$\sigma_i(A)$	$\varepsilon_i(K)$
SiH_4	4.084	207.6
SiH_3	3.943	170.3
SiH_2	3.803	133.1
SiH	3.662	95.8
H	2.5	30.0
H_2	2.915	59.7
Si_2H_3	4.494	318.2
Si_2H_4	4.601	312.6
Si_2H_5	4.717	306.9
Si_2H_6	4.828	301.3
$Si_2H_6^*$	4.828	301.3
$Si_2H_6^{**}$	4.828	301.3
Si_3H_n	5.08	354
Si_4H_n	5.80	354
Si_5H_n	6.50	354
Ar, Ar*	3.44	120

1063-7842/02/4706-0760$22.00

D_i are the diffusion coefficients, n_e is the electron concentration, K_{ji} are the electron-induced reaction rate constants, K_{jli} are the interradical reaction rate constants, and u is the gas flow rate in a plasmatron.

The first term on the right-hand side stands for the diffusion contribution to the concentration, and the second and the third terms describe the production and loss of sort-i particles, respectively. The term $n\nabla N_i$ takes into account the gas flow (convection) in the discharge chamber and makes the solution of Eq. (1) very cumbersome. It ensures the stationarity of the solution. However, in the time interval $[0, \tau]$, where $\tau = M_0/Q$ (M_0 and Q are, respectively, the gas mass and the gas flow rate in the reactor), convection can be disregarded and the problem becomes essentially nonstationary [3]. This approximation is known as the plug-flow regime and holds only when stirring of the reacting gas is negligibly small; otherwise, it is impossible to introduce the parameter τ, which defines the residence time of the gas in the reactor. In experiments, τ is usually 0.1 s, which considerably exceeds the chemical reaction time.

The diffusion coefficients were calculated by the Wilke formula, which is commonly used for multicomponent mixtures:

$$D_i = (n - n_i)\left[\sum_{k=1,\,k\neq i}^{N} n_k/D_{ki}\right]^{-1}, \qquad (2)$$

where n is the total concentration of the mixture and n_i is the concentration of the ith component.

Since the concentrations of radicals in the discharge do not exceed 0.001% of the plasma-forming gas concentration, we may restrict our consideration to the diffusion of the particles in the main components, which are SiH_4 and Ar. The coefficients D_{ki} of binary diffusion were determined with the formula from the molecular–kinetic theory of gases [4] using the Lenard-Jones potential

$$D_{ki} = 1.412308\frac{\sqrt{T^3/\mu_{ki}}}{p\sigma_{ki}^2\Omega_{ki}^*(T_{ki}^*)}\frac{\text{cm}^2}{\text{s}}. \qquad (3)$$

Here, p is the gas pressure expressed in torr, $\mu_{ki} = m_k m_i/(m_k + m_i)$ is the reduced mass of particles of the ith and kth sorts, $\sigma_{ki} = (\sigma_k + \sigma_i)/2$ is the effective collision diameter, $T_{ki}^* = kT/\varepsilon_{ki}$ is the characteristic temperature, $\varepsilon_{ki} = \sqrt{\varepsilon_k\varepsilon_i}$ is the intermolecular potential parameter (the potential well depth), and $Q_{ki}^*(T_{ki}^*)$ is the collision integral for mass transfer normalized to the Ω integral in the hard-sphere model. In our calculations, we used the approximation [5]

$$\Omega_{ki}^*(T_{ki}^*) = A(T_{ki}^*)^{-B} + C\exp(-DT_{ki}^*) + E\exp(-FT_{ki}^*) + G\exp(-HT_{ki}^*), \qquad (4)$$

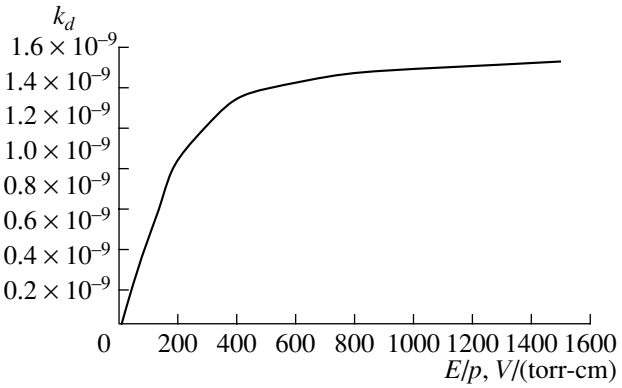


Fig. 1. Silane dissociation rate constant k_d vs. E/p.

where the coefficients $A = 1.06036$, $B = 0.15610$, $C = 0.19300$, $D = 47635$, $E = 1.03587$, $F = 1.52996$, $G = 1.76474$, and $H = 3.89411$. The parameters of the Lenard-Jones potential parameters [6, 7] are listed in Table 1.

The reactions involved in our model are listed in Table 2. We used the data from [6–8]. The rate constants for the reactions induced by electron impact were derived from the total cross section of silane dissociation by electron impact [6, 9]. We took into consideration that reactions (R1)–(R3) of total dissociation, account for 39, 48, and 13%, respectively, of the reaction products [7]. The rate constant for silane dissociation was calculated after the energy distribution of the electrons for the respective value of the applied field had been found [10, 11]. Figure 1 shows the rate constant for silane dissociation as a function of the applied field. Under the simulation conditions ($p = 0.125$ torr), the discharge is sustained by diffusion: $(n_i k_r \ll (1/\Lambda_D^2)D_a)$, where k_r is the recombination coefficient, D_a is the ambipolar diffusion coefficient, and Λ_D^2 is the diffusion length. Hence, the charged particles are lost on the walls of the chamber rather than in the bulk of the discharge.

In order to complete the statement of the problem, we should specify boundary conditions. To do this implies the consideration of the surface kinetics, which is difficult because of the lack of reliable data on the parameters of surface interaction between the components. That is why the boundary conditions from simple kinetic theory are used in the simulation. The relationship between the incident and reflected flows is given by

$$\Gamma_{out} = (1 - s)\Gamma_{in}, \qquad (5)$$

where $\Gamma_{in} = nv_t/4 + D/2(\partial n/\partial z)$ and $\Gamma_{out} = nv_t/4 - D/2(\partial n/\partial z)$ are the incident and the reflected flows, respectively; $v_t = 8kT/\pi m$ is the thermal velocity; m is the mass of the particle; and s is the deposition coefficient.

Then, the relationship for the gas-dynamic parameters on an absorbing surface has the form (diffuse

Table 2. Chemical reactions and rate constants

No.	Reaction	Rate constant*
	Electron-impact-induced reactions	
R1	$SiH_4 + e \longrightarrow SiH_2 + 2H + e$	3.0(–11)
R2	$SiH_4 + e \longrightarrow SiH_3 + H + e$	3.0(–11)
R3	$SiH_4 + e \longrightarrow SiH_2^+ + 2e + 2H$	2.0(–11)
R5	$Ar + e \longrightarrow Ar^m + e$	3.1(–11)
R6	$Ar^m + e \longrightarrow Ar^+ + 2e$	3.65(–3)
R4	$SiH_4 + e \longrightarrow SiH_3^+ + 2e + H$	2.0(–11)
R7	$Ar + e \longrightarrow Ar^+ + 2e$	1.0(–13)
R8	$SiH_4 + e \longrightarrow SiH + H_2 + H + e$	9.34(–12)
R9	$SiH_4 + e \longrightarrow SiH_2 + H_2 + e$	7.19(–12)
R10	$H_2 + e \longrightarrow 2H + e$	4.49(–12)
R11	$Si_2H_6 + e \longrightarrow SiH_4 + SiH_2 + e$	2.86(–10)
R12	$Si_2H_6 + e \longrightarrow Si_2H_4 + H_2 + e$	1.23(–11)
	Neutral–neutral reactions	
R13	$SiH_4 + H \longrightarrow SiH_3 + H_2$	$2.8 \times 10^{-11}\exp(-1250/T)$
R14	$SiH_4 + SiH \longrightarrow Si_2H_5^*$	1.0(–11)
R15	$SiH_4 + SiH \longrightarrow Si_2H_3 + H_2$	$4.4 \times 10^{-11}(1 - 1.4/(1 + \exp(P - 1.1)))$
R16	$SiH_4 + SiH \longrightarrow Si_2H_5$	2.5(–12)
R17	$SiH_4 + Si_2H_5 \longrightarrow SiH_3 + Si_2H_6$	5.0(–12)
R18	$SiH_4 + Si_2H_4 \longrightarrow Si_3H_8$	1.0(–10)
R19	$Si_2H_6 + H \longrightarrow SiH_4 + SiH_3$	$1.58 \times 10^{-11}\exp(-1250/T)$
R20	$Si_2H_6 + H \longrightarrow Si_2H_5 + H_2$	$2.4 \times 10^{-10}\exp(-1250/T)$
R21	$Si_2H_6^* \longrightarrow Si_2H_4 + H_2$	5.0(6) s^{-1}
R22	$Si_2H_6^* + Ar \longrightarrow Si_2H_6 + Ar$	5.00(–10)
R23	$Si_2H_6^{**} \longrightarrow SiH_4 + SiH_2$	2.3(7) s^{-1}
R24	$Si_2H_6^{**} \longrightarrow SiH_4 + H_2$	2.3(7) s^{-1}
	Reactions of radicals with radicals, silane, higher silanes, and hydrogen	
R25	$SiH_3 + SiH_3 \longrightarrow SiH_4 + SiH_2$	$4.1 \times 10^{-10}(1 + 3.2/(1 + \exp(0.4P)))$
R26	$SiH_3 + SiH_3 \longrightarrow Si_2H_6^{**}$	1.0(–11)
R27	$SiH_3 + Si_2H_5 \longrightarrow SiH_4 + Si_2H_4$	1.0(–10)
R28	$SiH_3 + Si_2H_5 \longrightarrow Si_3H_8$	1.0(–11)
R29	$SiH_3 + Si_2H_6 \longrightarrow SiH_4 + Si_2H_5$	$4.1 \times 10^{-10}\exp(-2500/T)$
R30	$SiH_2 + H \longrightarrow SiH + H_2$	3.30(–10)
R31	$SiH_2 + Si_2H_6 \longrightarrow Si_3H_8$	$3.9 \times 10^{-10}(1 - (1 + 0.55P)^{-1})$
R32	$Si_2H_3 + H_2 \longrightarrow Si_2H_5$	1.70(–10)
R33	$Si_2H_4 + H_2 \longrightarrow SiH_4 + SiH_2$	1.00(–10)
R34	$Si_2H_5 + H \longrightarrow Si_2H_4 + H_2$	1.00(–10)
R35	$Si_3H_8 + H \longrightarrow Si_2H_5 + SiH_4$	$2.4 \times 10^{-10}\exp(-1250/T)$
R36	$SiH_2 + Si_3H_8 \longrightarrow Si_4H_{10}$	1.00(–11)
R37	$Si_2H_5 + Si_2H_5 \longrightarrow Si_4H_{10}$	1.00(–11)
R38	$Si_2H_4 + Si_2H_6 \longrightarrow Si_4H_{10}$	1.00(–11)

Table 2. (Contd.)

No.	Reaction	Rate constant*
R39	$Si_3H_8 + SiH_3 \longrightarrow Si_4H_9 + H_2$	1.00(–11)
R40	$SiH + Si_2H_6 \longrightarrow Si_3H_7$	1.00(–11)
R41	$Si_3H_7 + H \longrightarrow Si_3H_8$	1.00(–11)
R42	$SiH + Si_3H_8 \longrightarrow Si_4H_9$	1.00(–11)
R43	$SiH_3 + Si_3H_8 \longrightarrow Si_4H_9 + H_2$	1.00(–11)
R44	$Si_4H_9 + H \longrightarrow Si_4H_{10}$	1.00(–11)
R45	$SiH + Si_4H_{10} \longrightarrow Si_5H_{11}$	1.00(–11)
R46	$SiH_2 + Si_4H_{10} \longrightarrow Si_5H_{12}$	1.00(–11)
R47	$SiH_3 + Si_4H_{10} \longrightarrow Si_5H_{11} + H_2$	1.00(–11)
R48	$Si_5H_{11} + H \longrightarrow Si_5H_{12}$	1.00(–11)
	Reactions with metastable particles	
R38	$Ar^m + SiH_4 \longrightarrow SiH_2 + 2H + Ar$	2.60(–10)
R39	$Ar^m + SiH_4 \longrightarrow SiH_3 + H + Ar$	1.40(–10)
R40	$Ar^m + H_2 \longrightarrow 2H + Ar$	7.00(–10)
R41	$Ar^m + SiH_3 \longrightarrow SiH_2 + H + Ar$	1.00(–10)
R42	$Ar^m + SiH_3 \longrightarrow SiH + H + Ar$	1.00(–10)
R43	$Ar^m + Si_2H_6 \longrightarrow Si_2H_4 + 2H + Ar$	6.60(–10)
R44	$Ar^m + Si_2H_4 \longrightarrow Si_2H_2 + 2H + Ar$	6.60(–10)
R45	$Ar^m + Ar^m \longrightarrow Ar^+ + Ar + e$	1.00(–9)
	Dust-forming reactions	
R46	$SiH_n + [Si_5H_{11}, Si_5H_{12}] \longrightarrow$ Dust	1.00(–11)
R47	$Si_2H_n + [Si_4H_9, Si_4H_{10}] \longrightarrow$ Dust	1.00(–11)
R48	$Si_2H_n + [Si_5H_{11}, Si_5H_{12}] \longrightarrow$ Dust	1.00(–11)
R49	$Si_3H_7 + [Si_3H_7, Si_3H_8] \longrightarrow$ Dust	1.00(–11)
R50	$Si_3H_7 + [Si_4H_9, Si_4H_{10}] \longrightarrow$ Dust	1.00(–11)
R51	$Si_3H_7 + [Si_5H_{11}, Si_5H_{12}] \longrightarrow$ Dust	1.00(–11)
R52	$Si_4H_9 + [Si_4H_9, Si_4H_{10}] \longrightarrow$ Dust	1.00(–11)
R53	$Si_4H_9 + [Si_5H_{11}, Si_5H_{12}] \longrightarrow$ Dust	1.00(–11)
R54	$Si_5H_{11} + [Si_5H_{11}, Si_5H_{12}] \longrightarrow$ Dust	1.00(–11)

* Rate constants are given for a pressure of 0.124 torr and a gas temperature of 500 K unless the associated relationship is specified.

reflection is assumed)

$$D\frac{\partial n}{\partial z} = -\frac{s v_t}{2(2-s)}n. \quad (6)$$

Under intense wall deposition, the supersaturated vapor pressure is considerable and the surface reevaporation from the surface can be neglected, as was done in Eq. (6).

COMPUTATIONAL METHOD

Let us consider the following equation written for 20 components in the cylindrical coordinate system:

$$\frac{\partial N_i}{\partial t} = D_i\left[\frac{1}{r}\frac{\partial}{\partial r}\left(r\frac{\partial N_i}{\partial r}\right) + \frac{\partial^2 N_i}{\partial z^2}\right] + F_i - A_i N_i. \quad (7)$$

The term $(F_i - A_iN_i)$ corresponds to the production and loss of the ith component in the chemical reactions; here, $A_i > 0$ and $\partial F_i/\partial N_i = 0$. Equations (2) were solved over the $[0, R] \times [0, L]$ domain within the $[0, \tau]$ time interval by using the Peaceman–Rachford (longitudinal–transverse) scheme:

$$\begin{aligned}\frac{\bar{n}-n}{0.5\Delta t} &= (\Lambda_z\bar{n} + \Lambda_r n) + F - A\bar{n},\\ \frac{\hat{n}-n}{0.5\Delta t} &= (\Lambda_r\bar{n} + \Lambda_z\bar{n}) + F - A\bar{n},\end{aligned} \quad (8)$$

where

$$\Lambda_r = \frac{1}{r}\frac{\partial}{\partial r}\left(D_i r\frac{\partial n_i}{\partial r}\right),$$

$$\Lambda_z = \frac{\partial}{\partial z}\left(D_i \frac{\partial n_i}{\partial z}\right),$$

$\bar{n}$ is the concentration on the time half-layer $\Delta t(m + 1/2)$, $\widehat{n}$ is the concentration on the time layer $\Delta t(m + 1)$, m is the number of a time layer, and Δt is the time step.

Taking the diffusion coefficient to be constant over the space of the problem and applying the integro-interpolation technique, we arrive at the following expressions for the operators Λ_r and Λ_z:

$$\Lambda_r n_k = \frac{1}{r_k}\left(r_k^- D \frac{n_k - n_{k-1}}{h_r}\right)_{r,k}$$

$$= \frac{D}{r_k h_r^2}(r_k^+ n_{k-1} - 2r_k n_k + r_k^- n_{k+1}), \quad 0 < r_i \le R,$$

$$\Lambda_r n_k = D\left(\frac{\partial^2 n}{\partial r^2}\right)_k = \frac{D}{h_r^2}(n_{k-1} - 2n_k + n_{k+1}), \quad r = 0, \tag{9}$$

$$\Lambda_z n_i = D\left(\frac{n_{i-1} - n_i + n_{i+1}}{h_z^2}\right),$$

where $r_k^- = (r_k - 0.5h)$, $r_k^+ = (r_k + 0.5h)$, $h_r = R/N_r$ is the step along the radius, $h_z = L/N_z$ is the longitudinal step,

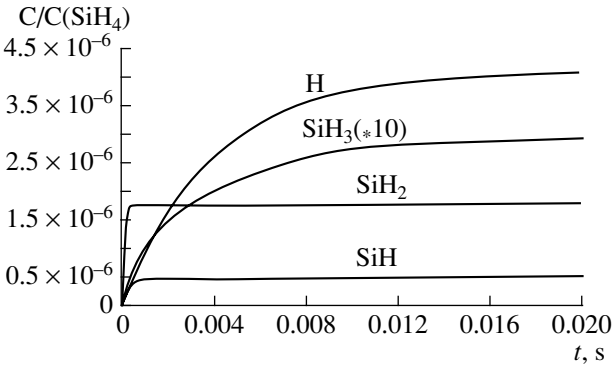


Fig. 2. Time dependences for concentrations of silane radicals and hydrogen.

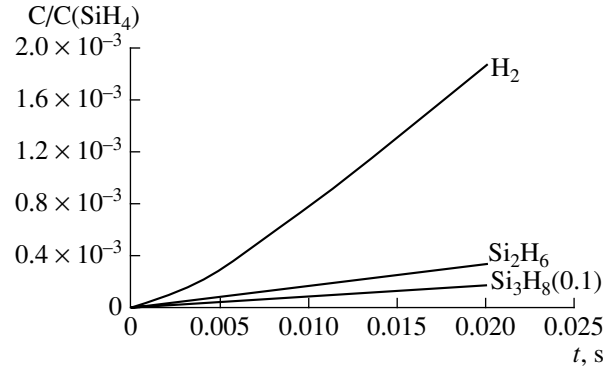


Fig. 3. Time dependences of the Si_2H_6, Si_3H_8, and H_2 concentrations.

$n_{i,k} = n(r_k, z_i)$, $r_k = kh_r$, and $z_i = ih_z$.

System (8) is closed by boundary conditions (similar to (6)) on the lateral surface and at the cylinder edges and is then rearranged to

$$\left.\frac{\partial n}{\partial r}\right|_{r=0} = 0, \tag{10a}$$

$$D\left.\frac{\partial n}{\partial r}\right|_{r=R} = -\frac{s v_t}{2(2-s)} n, \tag{10b}$$

$$D\left.\frac{\partial n}{\partial r}\right|_{z=0} = -\frac{s v_t}{2(2-s)} n, \tag{10c}$$

$$D\left.\frac{\partial n}{\partial r}\right|_{z=L} = -\frac{s v_t}{2(2-s)} n. \tag{10d}$$

The gas flow at the cylinder axis ($R = 0$) is taken to be zero. Condition (10c) can be approximated as

$$D\left.\frac{\partial n}{\partial z}\right|_{z=0} = -\beta n,$$
$$D\frac{n_{1,k} - n_{0,k}}{h_z} = -\beta n_{0,k} - 0.5h_z\mu_{0,k} + h_z\frac{\bar{n}_{0,k} - n_{0,k}}{\Delta t}. \tag{11}$$

Here, $\mu_{0,k} = (\Lambda_r n_{0,k} + F_{0,k} - \Lambda_{0,k} n_{0,k})$, and $\beta = (s v_t)/2(2 - s)$. Similar expressions can be obtained for conditions (10b) and (10c). In the difference form, condition (10a) appears as

$$D\frac{\widehat{n}_{i,1} - \widehat{n}_{i,0}}{h_r/4} = 0.5h_r\mu_{i,0} + 0.5h_r\frac{\widehat{n}_{i,0} - \bar{n}_{i,0}}{\Delta t/2}, \tag{12}$$

where $\mu_{0,k} = -(\Lambda_z \bar{n}_{i,0} + F_{i,0} - A_{i,0}\bar{n}_{i,0})$.

For calculations, it is convenient to represent expressions (11) and (12) in the form

$$\bar{n}_{0,k} = r_1 \bar{n}_{1,k} + p_1, \quad \widehat{n}_{i,0} = r_3 \widehat{n}_{i,0} + p_3, \tag{13}$$

where

$$r_1 = \frac{D\Delta t}{(h_z^2 + D\Delta t + \beta h_z \Delta t)},$$
$$p_1 = \frac{h_z n_{0,k} - \mu_{0,k}\Delta t}{(D\Delta t/h_z + \beta\Delta t + h_z)}, \tag{13a}$$

$$r_3 = \frac{4D\Delta t}{(4D\Delta t + h_r^2)}, \quad p_3 = \frac{\mu_{0,k} h_r \Delta t - h_r \bar{n}_{i,0}}{4D\Delta t + h_r^2}. \tag{13b}$$

In view of (9) and (13), system (8) takes the form (the subscript at the components is omitted)

$$\bar{a}_i \bar{n}_{i-1,k} - (\bar{a}_i + \bar{a}_{i+1} + 2 - A\Delta t)\bar{n}_{i,k} + \bar{a}_{i+1}\bar{n}_{i+1,k}$$
$$= -\Lambda_r n_{i,k}\Delta t - 2n_{i,k} - F\Delta t \quad (i = 1, N_z;\ k = 0, N_r),$$

$$\widehat{a}_k \widehat{n}_{i,k-1} - (\widehat{a}_k + \widehat{a}_{k+1} + 2 - A\Delta t)\widehat{n}_{i,k} + \widehat{a}_{k+1}\widehat{n}_{i,k+1}$$
$$= -\Lambda_z n_{i,k}\Delta t - 2\bar{n}_{i,k} - F\Delta t \quad (k = 1, N_r;\ i = 0, N),$$

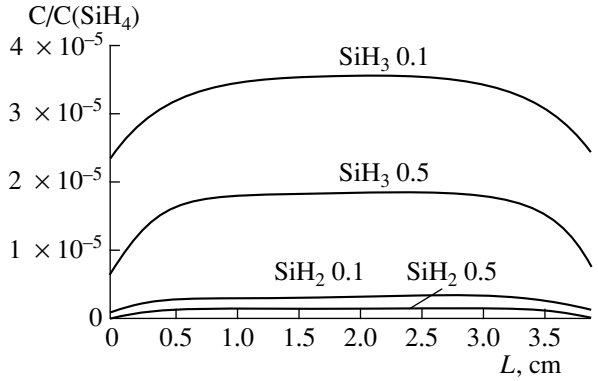


Fig. 4. Axial profiles of radical concentrations. The numbers at the curves indicate pressure p, torr.

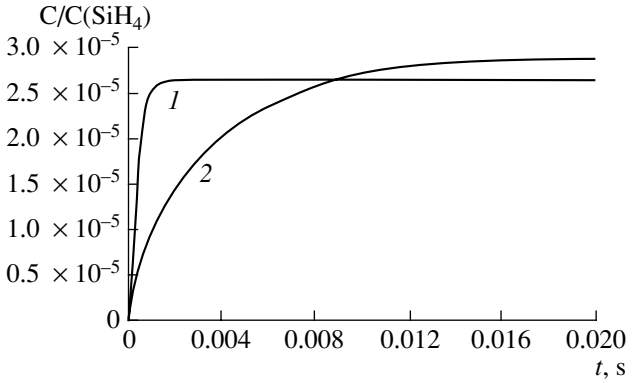


Fig. 5. Time dependence of the silyl concentration for (*1*) 100% SiH_4 and (*2*) 95% Ar + 5% SiH_4 mixture.

$$\bar{n}_{i,0} = r_1\bar{n}_{i,0} + p_1\bar{n}_{i,0} = r_2\bar{n}_{i,0} + p_2,$$
$$\hat{n}_{i,0} = r_3\hat{n}_{i,0} + p_3\hat{n}_{i,0} = r_4\hat{n}_{i,0} + p_4.$$

Here,

$$\bar{a}_i = \bar{a}_{i+1} = \frac{D\Delta t}{h_z^2}, \quad \hat{a}_k = \frac{D\Delta t r_k}{h_r r_k}, \quad \hat{a}_{k+1} = \frac{D\Delta t r_k^+}{h_r r_k}. \quad (14)$$

The coefficients r_2, p_2 and r_4, p_4 can readily be found from formulas (13a) by substituting corresponding h, μ, and $n_{i,k}(\bar{n}_{i,k})$. System (14) is solved for all of the components by scalar sweeping: first along the cylinder length for every fixed value of the radius r_k and then, along the radius for every fixed value of z_i. It is evident from the calculations that the solutions are stable for a time step $\Delta t \le 10^{-8}$ s.

RESULTS

We performed the numerical simulation of the physicochemical processes in a cylindrical reactor with the sizes L = 3.88 cm and R = 1.25 cm. Other parameters were as follows: the total gas pressure p = 0.125 torr, temperature T = 500 K, and the electric field frequency f = 13.56 MHz. The electron concentration was defined parametrically at a level $n_e = 10^8$ cm^{-3}.

Figure 2 shows the time dependences of the concentrations of silane radicals and atomic hydrogen. It is seen that SiH_2 and SiH rapidly (≈0.001 s) reach their equilibrium concentrations, whereas the time taken for the SiH_3 equilibrium concentration to set in is on the order of 0.01 s. Similar behavior is displayed by the atomic hydrogen. The considerable content of atomic hydrogen in the plasma correlates with the result [3] obtained for the diode-type PECVD reactor. Another important feature is the accumulation of Si_2H_6, Si_3H_8, and H_2 in the reactor volume (Fig. 3). The SiH_3 concentration profiles along the length of the discharge chamber at different pressures are plotted in Fig. 4. The SiH_3 concentration depends strongly on the addition of argon. This is because the $Ar^m + SiH_4 \longrightarrow SiH_3 + H + Ar$ reaction provides an additional channel of silane decomposition, thus increasing the silyl yield as compared to that in the pure silane plasma (Fig. 5). However, the presence of Ar extends the time taken for the SiH_3 concentration to reach its equilibrium level. Although the intense dust formation affects the kinetics of the basic components only slightly, it may deteriorate the film quality. Figure 6 demonstrates the dust concentration for different pressures of the discharge in pure silane and in the 95% Ar + 5% SiH_4 mixture. The relative content of the dust in the argon–silane plasma turns out to be lower: by a factor of 4 at high pressures (0.25–0.5 torr) and by a factor of 6–6.5 at a pressure of about 0.1 torr.

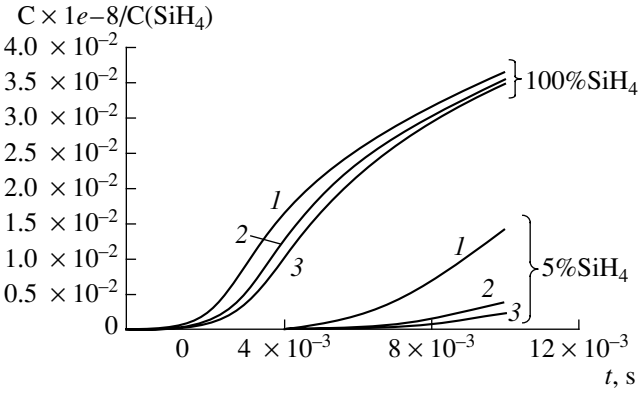


Fig. 6. Dust concentration at p = 0.25 (*1*), 0.125 (*2*), and 0.1 (*3*) torr.

REFERENCES

1. R. G. Sharafutdinov, M. R. Baklanov, B. M. Ayupov, *et al.*, Zh. Tekh. Fiz. **65** (1), 181 (1995) [Tech. Phys. **40**, 101 (1995)].

2. V. I. Strunin, G. Zh. Khudaĭbergenov, *et al.*, in *Proceedings of the Conference on Physics of Gas Discharge, Ryazan', 2000*, p. 182.
3. Yu. E. Gorbachev, M. A. Zatevakhin, and I. D. Kaganovich, Zh. Tekh. Fiz. **66** (12), 89 (1996) [Tech. Phys. **41**, 1247 (1996)].
4. R. B. Bird, W. E. Stewart, and E. N. Lighfoot, *Transport Phenomena* (Wiley, New York, 1960; Khimiya, Moscow, 1974).
5. S. P. Rudobashta, *Diffusion in Chemical Engineering Processes* (Khimiya, Moscow, 1993).
6. M. J. Kushner, J. Appl. Phys. **63**, 2532 (1988).
7. O. Leroy *et al.*, Plasma Sources Sci. Technol., No. 7, 348 (1998).
8. Yu. E. Gorbachev, M. A. Zatevakhin, V. V. Krzhizhanovskaya, and V. A. Shveĭgert, Zh. Tekh. Fiz. **70** (8), 77 (2000) [Tech. Phys. **45**, 1032 (2000)].
9. A. A. Deryugin, I. V. Kochetov, *et al.*, Teplofiz. Vys. Temp. **23**, 435 (1985).
10. V. I. Strunin, G. Zh. Khudaĭbergenov, and V. V. Shkurkin, in *Proceedings of the II All-Russia Scientific and Technical Conference "Computer Technology in Science, Design, and Production," 2000*, Part V, p. 29.
11. V. I. Strunin, G. Zh. Khudaĭbergenov, and V. V. Shkurkin, Vestn. Oms. Gos. Univ., No. 2, 27 (2000).

Translated by A. Sidorova-Biryukova

Technical Physics, Vol. 47, No. 6, 2002, pp. 767–771. Translated from Zhurnal Tekhnicheskoĭ Fiziki, Vol. 72, No. 5, 2002, pp. 115–119.

EXPERIMENTAL INSTRUMENTS AND TECHNIQUES

Charge and Size Distribution of Clusters Formed in Ion Sputtering of Metals

V. I. Matveev

Pomorsk M. V. Lomonosov University, Arkhangelsk, 163006 Russia

e-mail: matveev.victor@pomorsu.ru

Received September 24, 2001

Abstract—Description of the sputtering of metals by ion bombardment, in which large neutral or charged clusters are produced with the number of atoms $N \geq 5$, is further elaborated using simple physical assumptions. The results obtained, represented by simple formulas, are in agreement with experiment. By way of example, calculations of the mass spectra of neutral and singly charged clusters produced by ion sputtering of tantalum and the ionization coefficients of silver clusters are presented. © *2002 MAIK "Nauka/Interperiodica".*

INTRODUCTION

Considerable efforts are being devoted to research on ion sputtering of solids (see, for example, reviews [1–5] and references therein). Generally, the sputtering of solids can be either elastic (rebound) or inelastic (electronic), depending on the mechanism of energy transfer from the incident atoms to target atoms [1, 4]. In the inelastic sputtering process, which usually occurs in sputtering with fast ions [4] (traveling at a speed of $\sim 10^9$ cm/s), or multiply charged ions, the ion transfers its energy by exciting the electronic subsystem of the solid with subsequent transfer of the electronic excitation energy to atoms of the solid. In the elastic scattering process [1, 4] taking place under bombardment with ions carrying a small charge and possessing an energy from a few to tens of kiloelectronvolts, the energy is usually transferred from the impinging ion directly to a target atom, as a result of a (elastic) collision of the ion with a surface target atom with subsequent energy redistribution in the process of cascade collisions. We will consider the sputtering phenomenon of the latter, elastic type. Consistent calculations and the complete theoretical description of the sputtering process are extremely difficult, first of all, because of the many-particle nature of the problem both in the stage of penetration of the ions into the solid and in the formation stage of sputtering products consisting not only of single target atoms but also of polyatomic particles or clusters [6–8]. The first principle calculations using molecular dynamics methods [1] (see also [9–12]) are technically complicated in this case, especially for large numbers of atoms in the cluster, and poorly reproducible in calculations by different researchers. The difficulties increase considerably if the scheme of molecular dynamics is supplemented with formation processes of the charge states of sputtering products. Formation of the charge states of monoatomic particles, dislodged from or scattered by a metal surface, have been treated in a number of experimental and theoretical studies (see, for example [5], pp. 137–236). The mechanism by which charge states of polyatomic particles are formed has been studied much less extensively [13–16] both theoretically and experimentally.

In this study, a description is given of the sputtering of metal under ion bombardment for the case where sputtered particles are large (with the number of atoms $N \geq 5$) neutral and charged clusters. It is based on simple physical assumptions and is in agreement with experiment. Our treatment makes use of studies [17–20], but the result is represented by a simple formula for the probability of cluster ejection, as well as the size and charge state of the clusters, and includes the dependence on the target temperature. The proposed approach is absolutely inapplicable to the case of sputtering in the form of single atoms or small clusters. Comparison with experiment indicates that this approach [17, 21] can be applied starting with a certain number of atoms in the cluster ($N \geq 5$).

CALCULATION OF PROBABILITY

Let the bombarded solid consist of atoms residing in oscillating wells of depth Δ and having an intrinsic frequency ω. The characteristic period of oscillations is $T = 2\pi/\omega$. Let the speed of the bombarding ions be such that over a time $\tau \ll T$, the ions and fast recoil atoms in the metal experience multiple collisions with the result that the metal atoms receive momenta $\mathbf{q}_i$, where i is the number of atoms. Inequality $\tau \ll T$ allows a different formulation; i.e., the passing of an ion through the system of oscillators is equivalent to an instantaneous and simultaneous transfer to the oscillators of momenta $\mathbf{q}_i$ ($i = 1, 2, \ldots N$), where N is the number of oscillators. In what follows, we assume that all $\mathbf{q}_i$, $i = (1, 2, \ldots, N)$, are independent and all directions are equally probable. The objective of the calculation is to derive the proba-

1063-7842/02/4706-0767$22.00

bility of a system of N oscillators (which instantaneously received momenta $\mathbf{q}_i$) being found in finite bound states with the center of mass having a momentum $\mathbf{k}$. If the last condition is met, then the system of N oscillators moving as a whole will have the momentum $\mathbf{k}$. The easiest way of calculating the probability of such an event, which corresponds to the correlated ejection of a group of atoms, is to resort to quantum mechanics. Let us adopt Einstein's model and replace the group of N atoms by a system of N-independent identical oscillators with an intrinsic frequency ω.

The corresponding wave function has the form

$$\Psi_i = \Phi(\mathbf{R})\phi_1(\mathbf{r}_1)\phi_2(\mathbf{r}_2)\ldots\phi_N(\mathbf{r}_N), \quad (1)$$

where $\Phi(\mathbf{R})$ is the wave function of the center of mass of the group of N atoms; $\mathbf{R}$ is the coordinate of the center of mass; and ϕ_i is the wave function of the ith oscillator with coordinate $\mathbf{r}_i$.

Put differently, the result of the ion's journey through the system of oscillators amounts to an instantaneous and simultaneous transfer of the momentum $\mathbf{q}_i$ to each oscillator ($i = 1, 2, \ldots, N$) and of the momentum $\sum_{i=1}^{N}\mathbf{q}_i$ to the center of mass. This transforms wave function (1) into

$$\exp\left(\frac{i}{\hbar}\sum_{i=1}^{N}\mathbf{q}_i\cdot\mathbf{R}\right)\Phi(\mathbf{R})\exp\left(\frac{i}{\hbar}\mathbf{q}_1\cdot\mathbf{r}_1\right)\phi_1(\mathbf{r}_1) \times \exp\left(\frac{i}{\hbar}\mathbf{q}_2\cdot\mathbf{r}_2\right)\phi_2(\mathbf{r}_2)\times\ldots\times\exp\left(\frac{i}{\hbar}\mathbf{q}_N\cdot\mathbf{r}_N\right)\phi_N(\mathbf{r}_N). \quad (2)$$

We assume that prior to receiving the momenta $\mathbf{q}_i$, all oscillators are in their ground states, i.e.,

$$\phi_1 = \phi_2 = \ldots = \phi_N \equiv \phi_0, \quad \Phi \equiv \Phi_0.$$

Therefore, the amplitude of the probability of finding an arbitrary final state

$$\Psi_f = \Phi_\mathbf{k}(\mathbf{R})\phi_{n1}(\mathbf{r}_1)\phi_{n2}(\mathbf{r}_2)\ldots\phi_{nN}(\mathbf{r}_N)$$

with the center of mass in a state $\Phi_\mathbf{k}$ of the continuous spectrum and with a momentum $\mathbf{k}$ depends on the projection of the states Ψ_f onto states (2). After summation over quantum numbers n_i of the oscillators (subject to $n = \sum_{i=1}^{N} n_i$, where n has the meaning of the principal quantum number of a system of N oscillators), the square of the amplitude modulus is

$$W_\mathbf{k}^{(n)} = \frac{1}{n!}\left[\frac{1}{2\hbar^2\alpha^2}\sum_{i=1}^{N}q_i^2\right]^n\exp\left\{-\frac{1}{2\hbar^2\alpha^2}\sum_{i=1}^{N}q_i^2\right\} \times\left|\left\langle\Phi_\mathbf{k}(\mathbf{R})\left|\exp\left(\frac{1}{\hbar}\sum_{i=1}^{N}\mathbf{q}_i\cdot\mathbf{R}\right)\right|\Phi_0(\mathbf{R})\right\rangle\right|^2, \quad (3)$$

where $\alpha^2 = m\omega/\hbar$, m being the mass of an oscillator (atom).

In calculating the square of matrix element (3) and in summing, we used Fermi's results (paper #74 in [22]; see also [17–19, 23]). This is actually the probability that a cluster of N atoms will break away as a whole, with a momentum $\mathbf{k}$ and in an excited state n. We are interested in stable ejected clusters; therefore, the summation should be carried out over all excitation states with n lower than n_0, which is the state with enough energy stored in the excited oscillators to cause disintegration of a cluster. This state can be defined as $n_0 \approx \Delta/\hbar\omega$, or as the state in which the vibrational energy of all oscillators is large enough to eject one atom out of a well of depth Δ. Thus, it is necessary to calculate

$$W_\mathbf{k} = \sum_{n=0}^{n_0} W_\mathbf{k}^{(n)},$$

where $n_0 \gg 1$; therefore, the summation is performed as follows:

$$f(x) = e^{-x}\sum_{n=0}^{n_0}\frac{1}{n!}x^n \approx e^{-x}\sum_{n=0}^{\infty}\frac{1}{n!}x^n\exp\left\{-\frac{n}{n_0}\right\}$$

$$= \exp\left\{x\left[\exp\left(-\frac{1}{n_0}\right)-1\right]\right\}\approx\exp\left\{-\frac{x}{n_0}\right\}.$$

Hence,

$$W_\mathbf{k} = \exp\left(-\frac{1}{n_0}\frac{1}{2\hbar^2\alpha^2}\sum_{i=1}^{N}q_i^2\right) \times\left|\left\langle\Phi_\mathbf{k}(\mathbf{R})\left|\exp\left(\frac{1}{\hbar}\sum_{i=1}^{N}\mathbf{q}_i\cdot\mathbf{R}\right)\right|\Phi_0(\mathbf{R})\right\rangle\right|^2.$$

Now, the full probability W_N for the center of mass of a group of N atoms being found in the continuous spectrum can be obtained by integrating $W_\mathbf{k}$ over all possible $\mathbf{k}$, although this probability is easier to calculate by summing over all bound states $\Phi_n(\mathbf{R})$ of the center of mass and subtracting the obtained sum from unity. Summation using this method gives

$$\sum_{n=1}^{k_0}\left|\left\langle\Phi_\mathbf{k}(\mathbf{R})\left|\exp\left(\frac{1}{\hbar}\sum_{i=1}^{N}\mathbf{q}_i\cdot\mathbf{R}\right)\right|\Phi_0(\mathbf{R})\right\rangle\right|^2$$

$$= \exp\left(-\frac{1}{k_0}\frac{1}{2\hbar^2\alpha^2}\left(\sum_{i=1}^{N}\mathbf{q}_i\right)^2\right),$$

where summation is carried out from $n = 0$ through some value $n = k_0$ (its derivation is given below) and the notation $\beta^2 = mN\Omega/\hbar$ is used, Ω being the oscillation frequency of the center of mass.

The result is

$$W_N = \left[1 - \exp\left(-\frac{1}{k_0}\frac{1}{2\hbar^2\alpha^2}\left(\sum_{i=1}^{N}\mathbf{q}_i\right)^2\right)\right] \times \exp\left(-\frac{1}{n_0}\frac{1}{2\hbar^2\alpha^2}\sum_{i=1}^{N}q_i^2\right). \quad (4)$$

The procedure for deriving k_0 is as follows. It is assumed that the center of mass of a group of N atoms oscillates harmonically at a frequency Ω in a potential well of depth U_N, which represents the binding energy between the cluster and the metal. This binding energy is proportional to a contact area S_N between the group of N atoms and the rest of the metal. Suppose that this surface is a hemisphere with its center at the metal surface as it was prior to sputtering. Then $U_N = \sigma S_N = \delta N^{2/3}$ [17–20]. In this way, we discriminate between the depth Δ of the potential well, in which each atom of the solid is found, and the surface energy δ, binding the cluster to the bulk of the metal per one cluster atom. It us evident that in formula (4) $k_0 = U_N/(\hbar\Omega)$.

Then, probability (4) should be averaged over all possible values of $\mathbf{q}_i$, $i = (1, 2, \ldots, N)$. It is natural to assume that all quantities $\mathbf{q}_i$ are independent and their directions are equally probable, so that the average over angles $\Omega_{\mathbf{q}i}$ of the $\mathbf{q}_i$ vectors is

$$\overline{W}_N = \frac{1}{(4\pi)^N}\int\int\ldots\int d\Omega_{\mathbf{q}1}d\Omega_{\mathbf{q}2}\ldots d\Omega_{\mathbf{q}N}W_N. \quad (5)$$

Unlike the case considered in [17–20], the integrals in this formula cannot be factored; therefore, the average in question will be calculated using the following representation:

$$\exp\left(-\rho\left(\sum_{i=1}^{N}\mathbf{q}_i\right)^2\right) = \int d^3p\,\delta\left(\sum_{i=1}^{N}\mathbf{q}_i - \mathbf{p}\right)\exp(-\rho p^2)$$

$$= \frac{1}{(2\pi)^3}\int d^3p\int d^3r\exp\left(-i\left(\sum_{i=1}^{N}\mathbf{q}_i - \mathbf{p}\right)\cdot\mathbf{r}\right)$$

$$\times\exp(-\rho p^2) = \frac{1}{(2\pi)^3}\left(\sqrt{\frac{\pi}{\rho}}\right)^3\int d^3r$$

$$\times\exp\left(-i\left(\sum_{i=1}^{N}\mathbf{q}_i\right)\cdot\mathbf{r}\right)\exp\left(-\frac{r^2}{4\rho}\right),$$

where $\rho = 1/(k_0 2\beta^2) = 1/(2mNU_N) = 1/[N^{5/3}(k_{01})^2]$ and $k_{01} = (2m\delta)^{1/2}$. It is easy to arrive at $(4\pi)^{-1}\int d\Omega_{\mathbf{q}_i}\exp(-i\mathbf{q}_i\cdot\mathbf{r}) = (qr)^{-1}\sin(qr)$, and the average quantity (5) finally takes the form

$$\overline{W}_N = \exp\left(-\frac{1}{n_0}\frac{1}{2\hbar^2\alpha^2}\sum_{i=1}^{N}q_i^2\right)\left[1 - \frac{1}{2\sqrt{\pi}}(\rho)^{-3/2} \times\int_0^\infty r^2 dr\exp\left(-\frac{r^2}{4\rho}\right)\left(\frac{1}{qr}\sin(qr)\right)^N\right]. \quad (6)$$

The integral in this expression can be easily calculated using the formula

$$\left(\frac{1}{x}\sin x\right)^N \approx \exp\left(-\frac{Nx^2}{6}\right),$$

which is true for $N \gg 1$ and is easily obtainable recalling that the function

$$f(x) = \left(\frac{1}{x}\sin x\right)^N$$

at $N \gg 1$ satisfies the differential equation

$$\frac{d}{dx}f(x) \approx -N\frac{x}{3}f(x).$$

Indeed,

$$\frac{d}{dx}\left(\left(\frac{1}{x}\sin x\right)^N\right) = N\left(\frac{1}{x}\sin x\right)^{N-1}\left(\frac{\cos x}{x} - \frac{\sin x}{x^2}\right)$$

$$\approx N\left(\frac{1}{x}\sin x\right)^{N-1}\left(-\frac{x}{3}\right) \approx -N\frac{x}{3}\left(\frac{1}{x}\sin x\right)^N.$$

Therefore, calculation of the integral in (6) is simple. We ultimately obtain

$$\overline{W}_N = \left[1 - \frac{1}{(1 + Nq^2\rho 2/3)^{3/2}}\right] \times\exp\left(-\frac{1}{n_0}\frac{1}{2\hbar^2\alpha^2}\sum_{i=1}^{N}q_i^2\right). \quad (7)$$

We proceed further, assuming for simplicity as in [17–20], that though all $\mathbf{q}_i$ are directed absolutely chaotically, their absolute magnitudes are equal, i.e., $|\mathbf{q}_i| = q$ for any $i = 1, 2, \ldots, N$. Then, the probability takes the form

$$\overline{W}_N = \left[1 - \frac{1}{\left(1 + N^{-2/3}\frac{2}{3}\frac{q^2}{(k_{01})^2}\right)}\right]\exp\left(-N\frac{\delta}{\Delta}\frac{q^2}{(k_{01})^2}\right), \quad (8)$$

where $k_{01} = (2m\delta)^{1/2}$.

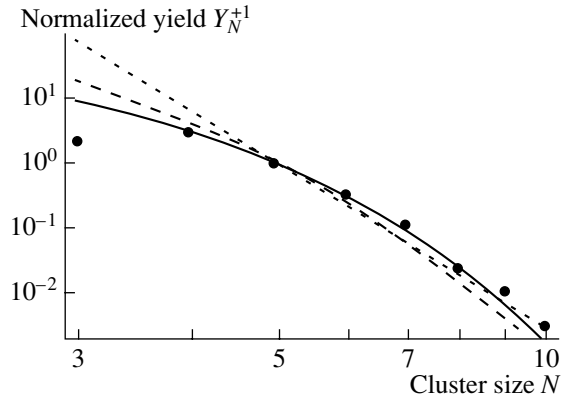


Fig. 1. Normalized yield Y_N^{+1} of singly charged clusters as a function of the number of atoms in the cluster. The solid curve is the calculation for a variable parameter value of q = 500 u (atomic units: $\hbar = m_e = e = 1$); solid circles is experiment [25]; the dashed curved is the calculated mass spectrum of neutral clusters (q = 500 u) for the same target temperature; the dotted curve is the power law [6] normalized by a cluster of 5 atoms, i.e., the function $N^{-8.5}/5^{-8.5}$.

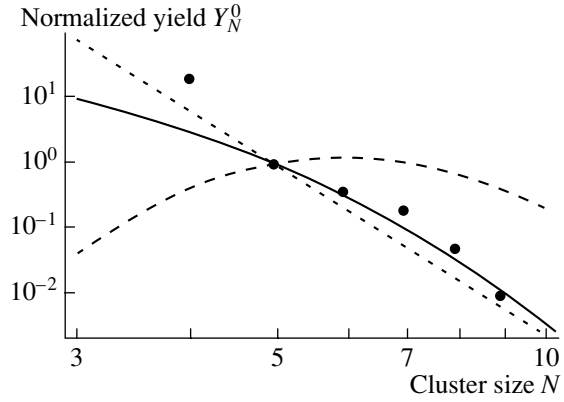


Fig. 2. Normalized yield Y_N^0 of neutral clusters as a function of the number of atoms in the cluster. Solid curve—calculation (q = 450 u); solid circles—experiment [6]; dashed curved—calculated mass spectrum of singly charged clusters (q = 450 u) for the same target temperature; dotted curve—the same as in Fig. 1.

CHARGE STATE CHARACTERISTICS

Thus, we calculated the possibility of a cluster breaking away as a whole (group) without its atoms changing their positions relative to each other. The charge-state formation process is part of the sputtering mechanism. We will proceed further relying substantially on the assumption that large clusters are dislodged as whole groups of atoms, in which case the charge state of a group of N atoms can be determined. Therefore, we consider that the exchange between the conduction band electrons in the metal and electrons of the atoms comprising the cluster is possible up to some (critical) distance ξ, as was done in the statistical derivation of the Saha–Langmuir formula [24]. At distances exceeding ξ, the electronic exchange ceases nonadiabatically. Then, according to [20], the possibility $P_N(Q)$ for an ejected cluster of N atoms to have a charge Qe (e being the electron charge) follows from the standard formula for the probability of fluctuations,

$$P_N(Q) = \frac{1}{D_N}\exp\left\{-\frac{1}{2}\frac{Q^2}{(\Delta Q_N)^2}\right\},$$
$$\overline{(\Delta Q_N)^2} = \frac{3^{1/2}}{\pi^{4/3}}\frac{m_e\Theta}{\hbar^2}\left(\frac{V}{N}\right)^{2/3}\gamma^{1/3}N, \quad (9)$$

where D_N is a normalization factor, which is determined by summation over all possible values of Q = 0, ±1, ±2, …; m_e is the conduction electron mass; V is the cluster volume; Θ is the target temperature; γ is the valence of the metal atoms; and $\overline{(\Delta Q_N)^2}$ is the average squared deviation of the cluster charge from the equilibrium values defined in terms of $\overline{\Delta N_e^2}$, the average squared deviation of the number of electrons in the cluster from the equilibrium value: $\overline{(\Delta Q_N)^2} = e^2\overline{\Delta N_e^2}$.

Thus, for obtaining the breakaway probability of a cluster of N atoms carrying a charge Qe, it is necessary to multiply the probability $\overline{W}_N$ given by (8) by $P_N(Q)$. This gives the final expression for the breakaway probability of a cluster of N atoms carrying the charge Qe:

$$\overline{W}_N^Q = \left[1 - \frac{1}{\left(1 + N^{-2/3}\frac{2}{3}\frac{q^2}{(k_{01})^2}\right)^{3/2}}\right]\exp\left(-N\frac{\delta}{\Delta}\frac{q^2}{(k_{01})^2}\right)$$
$$\times\frac{1}{D_N}\exp\left\{-\frac{1}{2}\frac{Q^2}{(\Delta Q_N)^2}\right\}. \quad (10)$$

COMPARISON WITH EXPERIMENT

A quantity usually measured in experiments is the relative probability Y_N^Q of the breaking away of clusters with a different number of atoms. Therefore, for comparison with experiment the probability expressed by (10) needs to be divided first by the value for a cluster with N = 5 (actually, for any cluster with $N \geq 5$, but in our case N = 5 is preferable). Thus, by definition, $Y_N^Q = \overline{W}_N^Q/\overline{W}_{N=5}^Q$. If required, any suitable units can then be used. Figures 1 and 2 show normalized yields of singly charged and neutral Ta_N clusters as a function of their size N for a tantalum target bombarded with singly charged ions Au^{-1} of energy 6 keV and with ions Ar^{+1} of energy 5 keV at target temperatures of Θ = 2273 and 300 K, respectively. For comparison, a power law curve

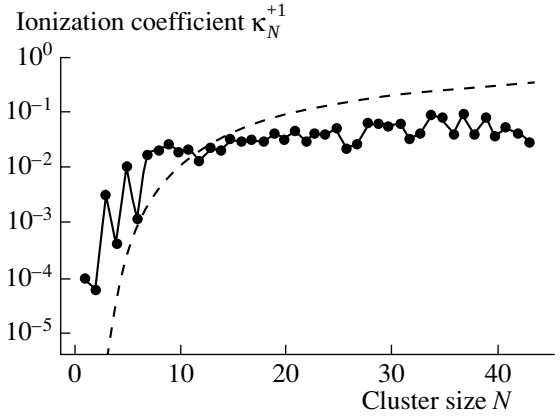


Fig. 3. Dependence of single ionization coefficient κ_N^{+1} on the number of atoms in clusters of Ag atoms. Dashed line—our calculation ($\delta = \Delta = 2.96$ eV, $q = 150$ u); solid circles—experiment [8].

for tantalum is shown normalized to a cluster of 5 atoms, i.e., the function $N^{-8.5}/5^{-8.5}$ [6]. In order to reduce the number of fitting parameters in the calculation, it is assumed that $\delta = \Delta = 8.1$ eV (sublimation energy [26]), so that the only remaining fitting parameter in (10) is the momentum q. It is noteworthy that mass spectra of the neutral clusters are only weakly dependent on the target temperature, while those of singly charged clusters change substantially with the target temperature (see also experimental data in [16]); but as the temperature is increased, these mass spectra become closer to those of neutral clusters. The charge state is usually characterized by an ionization coefficient equal to the ratio of the number of clusters with a charge Q to the number of neutral clusters (for a given cluster size N), which can be defined as follows:

$$\kappa_N^Q = \frac{\overline{W}_N^Q}{\overline{W}_N^{Q=0}}. \quad (11)$$

In Fig. 3, the dependence of the ionization coefficient κ_N^{+1} of silver clusters on the cluster size N at a target temperature of $\Theta = 400$ K is shown. An important feature of this dependence is the tendency to saturation (for $N > 5$) with increasing cluster size, so that further variation with increasing N is insignificant. Qualitatively similar behavior was observed experimentally [8, 13].

ACKNOWLEDGMENTS

The author thanks the Russian Ministry of Education for financial support (a grant for basic research in the natural sciences).

REFERENCES

1. *Fundamental and Applied Aspects of Solid Sputtering: Collection of Articles*, Ed. by E. S. Mashkova (Mir, Moscow, 1989).
2. H. H. Andersen and K. Dan, Vidensk. Selsk. Mat. Fys. Medd. **43**, 127 (1993).
3. H. M. Urbassek, W. O. Hofer, and K. Dan, Vidensk. Selsk. Mat. Fys. Medd. **43**, 97 (1993).
4. I. A. Baranov, Yu. V. Martynenko, S. O. Tsepelevich, and Yu. N. Yavlinskiĭ, Usp. Fiz. Nauk **156**, 478 (1988) [Sov. Phys. Usp. **31**, 1015 (1988)].
5. *Sputtering by Particle Bombardment*, Ed. by R. Behrisch and K. Vittmak (Springer-Verlag, New York, 1991; Mir, Moscow, 1998), Vol. 3.
6. A. Wucher and W. Wahl, Nucl. Instrum. Methods Phys. Res. B **115**, 581 (1996).
7. S. R. Coon, W. F. Calaway, and M. Y. Pellin, Nucl. Instrum. Methods Phys. Res. B **90**, 518 (1994).
8. C. Staudt, R. Heinrich, and A. Wucher, Nucl. Instrum. Methods Phys. Res. B **164–165**, 677 (2000).
9. A. Wucher and B. Y. Garrison, J. Chem. Phys. **105**, 5999 (1996).
10. A. Brunelle, S. Della-Negra, C. Deprun, *et al.*, Int. J. Mass Spectrom. Ion Processes **164**, 193 (1997).
11. Th. J. Colla, H. M. Urbassek, A. Wucher, *et al.*, Nucl. Instrum. Methods Phys. Res. B **143**, 284 (1998).
12. R. Kissel and H. M. Urbussek, Nucl. Instrum. Methods Phys. Res. B **180**, 293 (2001).
13. W. Wahl and A. Wucher, Nucl. Instrum. Methods Phys. Res. B **94**, 36 (1994).
14. V. Kh. Ferleger, M. B. Medvedeva, and I. A. Wojciechowski, Nucl. Instrum. Methods Phys. Res. B **125**, 214 (1997).
15. I. A. Wojciechowski, P. Bertrand, M. V. Medvedeva, and V. Kh. Ferleger, Nucl. Instrum. Methods Phys. Res. B **179**, 32 (2001).
16. C. Staudt, R. Heinrich, P. Mazarov, *et al.*, Nucl. Instrum. Methods Phys. Res. B **164–165**, 715 (2000).
17. V. I. Matveev and P. K. Khabibullaev, Dokl. Akad. Nauk **362**, 191 (1998) [Dokl. Phys. **43**, 544 (1998)].
18. V. I. Matveev, S. F. Belykh, and I. V. Verevkin, Zh. Tekh. Fiz. **69** (3), 64 (1999) [Tech. Phys. **44**, 323 (1999)].
19. S. G. Belykh, V. I. Matveev, I. V. Veryovkin, *et al.*, Nucl. Instrum. Methods Phys. Res. B **155**, 409 (1999).
20. V. I. Matveev, Zh. Tekh. Fiz. **70** (8), 108 (2000) [Tech. Phys. **45**, 1063 (2000)].
21. V. I. Matveev, Pis'ma Zh. Tekh. Fiz. **27** (18), 14 (2001) [Tech. Phys. Lett. **27**, 761 (2001)].
22. E. Fermi, *Scientific Works* (Nauka, Moscow, 1971), Vol. 1.
23. L. D. Landau and E. M. Lifshitz, *Course of Theoretical Physics*, Vol. 3: *Quantum Mechanics: Non-Relativistic Theory* (Nauka, Moscow, 1989, 4th ed.; Pergamon, New York, 1977, 3rd ed.).
24. L. N. Dobretsov and M. V. Gomoyunova, *Emission Electronics* (Nauka, Moscow, 1966).
25. S. G. Belykh, U. Kh. Rasulev, A. V. Samartsev, *et al.*, Nucl. Instrum. Methods Phys. Res. B **136–138**, 773 (1998).
26. C. Kittel, *Introduction to Solid State Physics* (Wiley, New York, 1976; Nauka, Moscow, 1978).

Translated by B. Kalinin

Technical Physics, Vol. 47, No. 6, 2002, pp. 772–776. Translated from Zhurnal Tekhnicheskoĭ Fiziki, Vol. 72, No. 5, 2002, pp. 120–124.

EXPERIMENTAL INSTRUMENTS AND TECHNIQUES

Crystal-Chemical Effect of the Thermodynamic History on the Composition–Temperature Phase Diagrams of (Na,Li)NbO_3 and (Na,K)NbO_3 Solid Solutions

I. P. Raevskiĭ, M. P. Ivliev, L. A. Reznichenko, M. N. Palatnikov, L. E. Balyunis, and M. A. Malitskaya

Research Institute of Physics, Rostov State University, pr. Stachki 194, Rostov-on-Don, 344090 Russia

e-mail: rip@ip.rsu.ru

Received October 16, 2001

Abstract—Composition (x)–temperature (T) phase diagrams obtained for the first time for single-crystalline (Na,Li)NbO_3 (system I) and (Na,K)NbO_3 (system II) niobates are compared with the known phase diagrams for ferroelectric niobate ceramics prepared by various techniques (standard firing and hot pressing). In system I, the thermodynamic history (process conditions) is shown to have a pronounced effect on the diagrams throughout the range of perovskite stability, while in system II, the effect is the most significant at low potassium concentration ($x \lesssim 0.15$). These data are associated with a small value of the tolerance factor t of the solid solutions, whose properties are sensitive to the thermodynamic history. Small ($\lesssim 0.9$–0.93) values of t indicate the "porosity" of the structure of solid solutions I and II in the above concentration ranges. Under these conditions, polarization- and rotation-related structure instabilities compete on equal terms and substitutional and interstitial solid solutions may coexist. © *2002 MAIK "Nauka/Interperiodica".*

INTRODUCTION

(Na,Li)NbO_3 and (Na,K)NbO_3 solid solutions (SSs) are promising piezoelectronic materials because they offer low permittivity ε, high velocity of sound, relatively high coefficient of electromechanical coupling, and a wide spectrum of the mechanical Q factor [1, 2]. Such a unique combination is not encountered in other piezoelectric materials. Moreover, these materials do not contain lead, which is of great importance in view of increasingly stringent environmental requirements. Sodium niobate (SN) $NaNbO_3$ exhibits multiple phase transitions (PTs) (six known to this day [1–3]), and the x–T phase diagrams of its solid solutions are very complex. Data for the PT number and temperatures are very contradictory [4–8] (Figs. 1a, 1b). The reason is that some transitions are hard to detect because associated changes in the structural and electrophysical parameters are very small. It should be noted that SN-based SSs have been studied mostly as ceramics. In this work, for the first time we construct x–T phase diagrams for SSs in systems I and II based on the dielectric and optical properties of the single crystals. The phase diagrams constructed are compared with the known x–T diagrams for ceramic materials obtained by various techniques (standard firing and hot pressing).

RESULTS AND DISCUSSION

(Na,Li)NbO_3 and (Na,K)NbO_3 transparent crystalline plates with $LiNbO_3$ and $KNbO_3$ concentrations up to 8 and 40 mol % were prepared by the method of spontaneous crystallization from solution in the $NaBO_2$ melt [9]. The composition of the SSs was determined by comparing their lattice parameters and PT temperatures with the published data for the ceramics. Ceramic I with a density of more than 95% of the theoretically predicted value was obtained by conventional ceramic technology (solid-phase synthesis, followed by firing without pressing). After grinding the samples, the powders were subjected to differential thermal analysis with a Paulik derivatograph. Optical studies of the 100- to 200-μm-thick plates were performed in a MIN-8 transmission polarization microscope provided with a heating stage. Dielectric studies of the crystals with Aquadag electrodes applied on the natural (001) faces were carried out in the interval 1–100 kHz with a P5083 ac bridge upon continuous heating or cooling at a rate of 2–3 K/min.

The x–T diagrams obtained are shown in Figs. 1 and 2. In system I, we succeeded in locating a phase transition at 220–330°C, which had not been discovered so far. One could anticipate this transition in system I based on the pyroelectric and optical properties of $Na_{1-x}Li_xNbO_3$ with $x \approx 0.02$ crystals [10, 11]. Unlike ceramic samples, crystals of both SSs, which, as follows from room-temperature X-ray data, contain orthorhombic antiferroelectric, P, and orthorhombic ferroelectric, Q, phases (the phases are designated according to [3, 8]), exhibit PTs in the interval 30–120°C. These PTs, being accompanied with anomalies in the optical and dielectric properties [9], have a wide

1063-7842/02/4706-0772$22.00

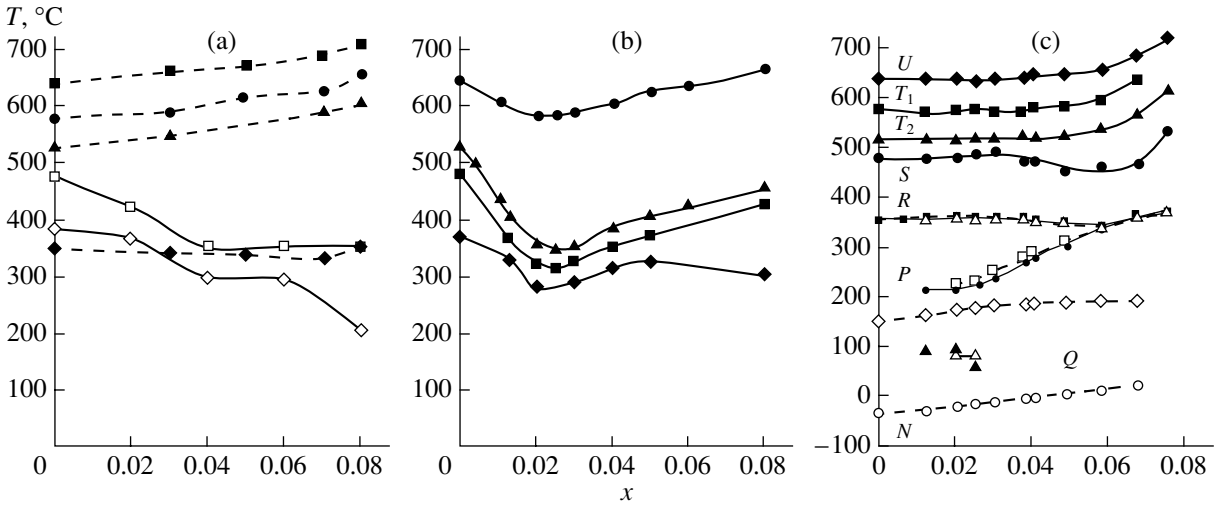


Fig. 1. *x–T* phase diagrams of the $Na_{1-x}Li_xNbO_3$ solid solutions in the range of small *x*. The phases are designated as in [3, 8]. (a) Ceramic obtained by conventional firing (empty symbols, X-ray diffraction data [4]; filled symbols, data of differential thermal analysis); (b) ceramic obtained by hot pressing, the diagram is constructed from the X-ray diffraction data [5]; and (c) single crystals obtained by mass crystallization from solution in melt (filled symbols, optical data in polarized light; empty symbols, dielectric measurements).

temperature hysteresis and are likely to be similar to the transitions between the *P* and *Q* phases at temperatures between 75 and 220°C in SN crystals [7]. In addition, in the *x–T* diagrams of systems I and II, shown are lines corresponding to the anomalies of the electrophysical and structural parameters at 150–190°C (Figs. 1c, 2), which were revealed by us earlier [9, 10]. For SN, these anomalies can be identified as PTs between two antiferroelectric orthorhombic phases, as was shown in [9, 12, 13].

Analysis of the phase states in systems I and II (Figs. 1, 2), as well as our experience in constructing the *x–T* diagrams for the ceramic SSs, shows the following. (i) Although the chemical compositions of the systems are close to each other (at small *x*), their *x–T* diagrams differ markedly by the number of PTs and by how the PT temperatures vary with the concentration of the second component. (ii) As the concentration of the second component grows, the manufacturability (i.e., the reproducible fabrication of high-density high-strength ceramics with a desired composition over a wide range of process parameters) in system I degrades, while in system II it improves. (iii) The form of the *x–T* diagram of system I depends considerably on the SS preparation conditions throughout the range of perovskite stability ($x \lesssim 0.15$).

In system II, a decrease in *x*, that is, the approach of the SS composition to SN, enhances the thermodynamic history effect. In particular, at $x \lesssim 0.5$, the physical properties of the SSs are virtually independent of the preparation conditions; at $x \lesssim 0.3$, the effect of the preparation conditions becomes more distinct; and at $x \lesssim 0.15$, the properties of the variously obtained SSs diverge considerably.

In [14, 15], we found correlation between the thermodynamic parameters of the PTs (temperatures, type of ordering, and phase sequence) and the crystal-chemical characteristics of perovskite-like compounds, which exhibit a large number of various, including rotational, transformations. The elucidation of the mechanism of their influence on the phase diagrams would allow researchers to understand changes in the SN thermodynamic properties when the crystal-chemical characteristics of compounds introduced as a second component of the solid solution vary. For analysis, it is convenient to use, along with ionic radii, the tolerance factor, which characterizes to what extent the sizes of atoms entering into a compound ABO_3 fit the perovskite structure. The tolerance factor is defined as $t = (R_A + R_B)/\sqrt{2}\,(R_B + R_O)$, where R_A, R_B, and R_O are the ionic radii of the cations A and B and the anion O taking into account their coordination number. Let us consider the properties of SN and its SSs in terms of such an approach in greater detail.

A specific feature of SN is the variety of dissimilar PTs, which is uncommon for oxides from the perovskite family. One is initiated by the rotation of oxygen octahedra (rotational PTs); the other, by the regular displacement of Nb^{5+} cations from the centers of the octahedra, i.e., by polarization (ferroelectric and antiferroelectric PTs). Rotational distortions of the SN structure occur either by rotating the octahedra in all layers orthogonal to the axis of rotation in the same direction (lattice vibration mode M_3) or by rotating the octahedra in adjacent layers in the opposite directions (R_{25} mode). The condensation of the modes causes orderings, which are treated in terms of order parameters (OPs). The modes M_3 and R_{25}, as well as the "polarization," are, in

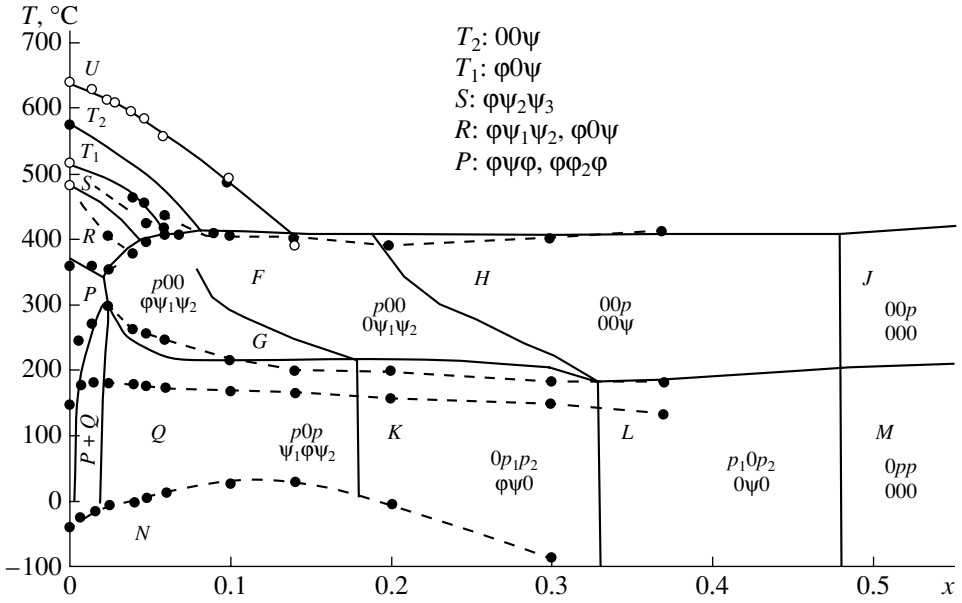


Fig. 2. *x–T* phase diagram of the $Na_{1-x}K_xNbO_3$ solid solutions. The phases are designated as in [3, 8, 16]. Filled symbols, dielectric measurements; continuous curves, *x–T* diagram of the ceramic constructed from X-ray diffraction data, specific heat data, and dielectric measurements [8, 16].

general, described by three OPs: ψ, φ, and *p*, respectively [14–17].

At $x \gtrsim 0.5$ in system II, the orderings are characterized by a single OP, polarization *p*; at $0.3 \lesssim x \lesssim 0.5$, by two OPs, rotational (ψ) and polarization (*p*). For *x* in the range between 0.15 and 0.3, the other rotational OP, φ, is added (Fig. 2). At $x > 0.15$, the prevailing OP is *p*, since the rotational OPs appear, if at all, only together with it. For $x < 0.15$, the rotational orderings arise independently of polarization; that is, the instabilities "have equal rights." As has been noted above, the effect of thermodynamic history on the PT temperatures and the physical characteristics of the SSs in system II is most significant precisely in the range $x < 0.15$. Thus, correlating changes in the phase-state picture in system II with the enhanced effect of preparation conditions on the properties of appropriate SSs, one notes that enhancement takes place with the appearance of each new OP, which characterizes the complete set of phase states for a given *x*.

Such a correlation may be explained as follows. A fabrication process implies a given sequence of external actions; hence, the correlation between the SS properties with the process conditions reflects different responses of the material to these actions. To some of the external actions, instabilities generating various OPs may respond oppositely. For example, hydrostatic pressure favors rotational PTs (i.e., raises their temperatures) and suppresses ferroelectric ones [18]. Under certain conditions, this may change the domains of existence of phase states, the phase states themselves, and, hence, the entire set of the SS physical properties. Thus, two or more dissimilar structure instabilities having comparable strengths cause a strong dependence of the SS properties on the process conditions. This shows up most vividly in system I, where the rotational and polarization instabilities coexist throughout the stability range of the perovskite phase. In SN, the domains of existence of the rotational and polarization orderings overlap heavily, while in perovskite-like titanates $ATiO_3$ (A = Ba, Sr, or Ca), they do not overlap at all [1, 16].

It should be borne in mind that the firing of the ceramics and the growth of the SS crystals take place at temperatures far exceeding those of the PTs. Therefore, the thermodynamic history influences the structural instabilities indirectly. For example, in the hot-pressed ceramics, high residual mechanical stresses arise [19]. They have been shown [20–22] to be the superposition of the hydrostatic pressure and the uniaxial compression along the axis of pressure application during hot pressing. According to the aforesaid, these stresses variously influence the properties of the SN-based ferroelectric and antiferroelectric SSs.

In SN, three high-temperature PTs are due to the rotation of the octahedra and the following three, to the combination of the rotation and polarization of the octahedra. In this case, two complex-ordered antiferroelectric phases (*R* and *P*) and one low-temperature ferroelectric (*N*) phase arise. From Figs. 1a–1c, it follows that the preparation conditions have a significant effect on the rotational PT temperatures. This means that the rotational instabilities are very sensitive to external actions. To realize the mechanism of this effect, it is necessary to consider how the rotational instabilities and the phase states induced by them form. The

sequence of purely rotational PTs in $NaNbO_3$ is the following:

$$U(Pm3m = O_h^1) \xrightarrow{640^\circ C} T_2(P4/mbm = D_{4h}^5)$$
$$\xrightarrow{575^\circ C} T_1(Ccmm = D_{2h}^{17}) \xrightarrow{520^\circ C} S(Pmmn = D_{2h}^{13}).$$

The tetragonal phase T_2 forms from the cubic phase U because of the condensation of the z component of the mode M_3 (the associated ordering is 00ψ). The phase T_1 forms from T_2 by the condensation of the y component of the mode R_{25} ($0\varphi\psi$). Finally, the phase S arises from T_1 through the condensation of one more component of the mode M_3 ($\varphi\psi_2\psi_3$). As has been shown [14, 15], this sequence of the orderings is favored by the specific features of the SN structure, which will be considered below.

In the symmetric phase, the size of the Na^+ cation (0.98 Å) is much less than that of the octahedral void where it is located ($t \approx 0.87 < 1$). In addition, the charge of Na^+ is small, so that the force attracting it to the void center is low. If $t < 1$, NbO_2 planes are compressed, while those of NaO are stretched. Compression causes a bend of Nb–O–Nb bonds, and stretching increases the rms displacements of the Na^+ cation from the void center. Eventually, the degree of Na^+ delocalization grows. The displacement also increases because, as was already mentioned, the force of Na^+ attraction to the void center is low (Na^+ has a small charge, +1). As the temperature is reduced, this effect is compensated for by bond bending due to the rotation of the octahedra. In the SN cubic phase, both rotational modes M_3 and R_{25} become "soft" as the temperature decreases but the former mode condenses first. Presumably, when the octahedra corresponding to M_3 rotate, a local quadrupole moment arises in the void and Na^+ cations delocalized from the void center interact with this moment, decreasing the energy of M_3 and, thereby, favoring its condensation. The OP φ_y appearing in the $T_2 \longrightarrow T_1$ transition is, to a great extent, associated with the triple interaction of type $M_3^z R_{25}^y$ [17, 23], where the rotational modes M_3 and R_{25} are coupled through the vibrational mode X of the Na^+ cations. Since the Na^+ cation is relatively free to move in the void, this "indirect" coupling turns out to be strong enough for φ_y to follow φ_z.

From this reasoning, it follows that the substitution of a larger K^+ cation (1.33 Å) for Na^+ (0.98 Å) will decrease the temperatures of all the rotational PTs, since the "free" space shrinks for both the anions (Nb–O–Nb bonds become difficult to bend) and the alkaline cations (the delocalization diminishes). In this case, t grows, the compression of the NbO_2 planes decreases, and Nb–O–Nb bonds straighten out. Accordingly, the conditions for ferroelectric ordering become more favorable and, at $x > 0.15$, this type of ordering prevails, reconfiguring rotational distortions. Thus, system II becomes more stable and denser (t increases) as x rises. As a result, the manufacturability improves.

The substitution of a smaller Li^+ cation (0.68 Å) for Na^+ must, as follows from the aforesaid, extend the free space and raise the temperatures of all the rotational PTs. This is the case both in the crystal and the ceramic of system I obtained by the various techniques but only for $x > 0.04$–0.06. At smaller x, the situation observed experimentally is more complicated.

In ceramic I, obtained by the conventional firing method without pressing, the temperatures of all the rotational PTs increase with x (Fig. 1a).

In hot-pressed ceramic I, the temperatures of all the rotational PTs decrease with increasing x for $0 < x < 0.02$, as in system II, and start rising at $x > 0.02$ (Fig. 1b). The decrease in the PT temperatures can be related to the increase in the structure density (since the lattice constant decreases [5]) rather than to the regular bond bending. High-temperature densification due to outer pressure can be associated with many factors. Among them are the isomorphic substitutional-to-interstitial transformation of the SSs (configurational PT) [24] (because of the high intercalation capacity of Li^+ ions [25]), octahedron deformation, lithium–defect interaction, etc. Today, it is difficult to draw any definite conclusions on this point without further investigation. The decrease in the S–R and R–P PT temperatures in hot-pressed ceramic I is likely to be associated with the fact that the antiferroelectric order arises together with the complex rotational order, and the suppression of the latter (because of the densification) causes the PT temperatures to diminish. With growing x (at $x > 0.02$), the Li $\longrightarrow$ Na substitution dominates, which "loosens" the structure (t drops) and raises the temperatures of the rotational PTs.

In single crystals I, the rotational PT temperatures vary insignificantly at $x < 0.04$. Then ($x > 0.04$), they grow with x, the growth being faster for $x > 0.06$ (Fig. 1c). It seems that, at small x, the probabilities of substitution and interstice occupation are close to each other, while substitution dominates at $x > 0.04$, which increases the temperatures of the rotational PTs.

CONCLUSION

Thus, the porosity of SN-based SSs is a primary factor responsible for the effect of the thermodynamic history (preparation conditions) on their physical properties. The porosity is associated with the misfit between the crystal-chemical parameters of the components and the perovskite structure (small tolerance factor t). This makes the structure unstable, specifically against rotational distortions. As the porosity grows (t decreases still further) when the SS composition changes, the material becomes even more sensitive to external actions and its response to them is more pronounced and diversified. The manufacturability of SN-based SSs diminishes appreciably at $t < 0.90$–0.93, since rota-

tional and polarization instabilities become comparable in strength in this interval of t. The manufacturability degradation is also typical of other oxides of the perovskite family for which $t \lesssim 0.93$ and which exhibit rotational PTs, such as $PbZrO_3$ [19] and $Pb_{1-x}Ca_xTiO_3$ at $x > 0.4$ [26]. Thus, the results reported can be considered as general.

ACKNOWLEDGMENTS

This work was partially supported by the Russian Foundation for Basic Research (grants no. 99-02-17575 and 00-03-32652a).

REFERENCES

1. G. A. Smolenskiĭ, V. A. Bokov, V. A. Isupov, *et al.*, The Physics of Ferroelectric Phenomena (Nauka, Leningrad, 1985).
2. A. Ya. Dantsiger, O. N. Razumovskaya, L. A. Reznichenko, and S. I. Dudkina, *High-Efficiency Piezoceramic Materials: Optimization of Search* (Paĭk, Rostov-on-Don, 1995).
3. H. D. Megaw, Ferroelectrics **7**, 87 (1974).
4. T. Nitta, J. Am. Ceram. Soc. **51**, 626 (1968).
5. L. A. Shilkina, L. A. Reznichenko, M. F. Kupriyanov, and E. G. Fesenko, Zh. Tekh. Fiz. **47**, 2173 (1977) [Sov. Phys. Tech. Phys. **22**, 1262 (1977)].
6. M. Palatnikov, O. Voloshina, Yu. Serebryakov, *et al.*, Ferroelectrics **131**, 227 (1992).
7. I. Lefkowitz, K. Lukaszewicz, and H. D. Megaw, Acta Crystallogr. **20**, 670 (1966).
8. M. Ahtee and A. Glazer, Acta Crystallogr. A **32**, 434 (1978).
9. I. P. Raevskiĭ, L. A. Reznichenko, V. G. Smotrakov, *et al.*, in *Proceedings of the International Workshop "Ordering in Minerals and Alloys," Rostov-on-Don, 2000*, p. 96.
10. A. Sadel, R. Von der Mühll, and J. Ravez, Mater. Res. Bull. **18**, 45 (1983).
11. A. Sadel, R. Von der Mühll, J. Ravez, and P. Hagenmuller, Ferroelectrics **47**, 169 (1983).
12. I. P. Raevskiĭ, L. A. Reznichenko, V. G. Smotrakov, *et al.*, Pis'ma Zh. Tekh. Fiz. **26** (16), 97 (2000) [Tech. Phys. Lett. **26**, 744 (2000)].
13. X. B. Wang, Z. X. Shen, Z. P. Hu, *et al.*, J. Mol. Struct. **385**, 1 (1996).
14. M. P. Ivliev, L. A. Reznichenko, and V. P. Sakhnenko, in *Proceedings of the International Scientific and Practical Conference "Piezotechnique-95," Rostov-on-Don, 1995*, p. 112.
15. M. P. Ivliev, in *Proceedings of the International Workshop "Order, Disorder, and Properties of Oxides," Sochi, 2001*, p. 141.
16. K. S. Aleksandrov, A. T. Anistratov, B. V. Beznosikov, *et al.*, *Phase Transitions in Crystals of ABX_3 Halogen Compounds* (Nauka, Novosibirsk, 1981).
17. C. N. W. Darlington, Phys. Status Solidi **76**, 231 (1976).
18. T. Hidaka, Phys. Rev. B **17**, 4363 (1978).
19. K. Okazaki, *Ceramic Engineering for Dielectrics* (Tokyo, 1969; Énergiya, Moscow, 1976).
20. I. P. Raevskiĭ and L. A. Reznichenko, Available form VINITI, No. 3251-78.
21. I. P. Raevskiĭ, L. A. Reznichenko, and A. N. Kalitvanskiĭ, Zh. Tekh. Fiz. **50** (9), 1983 (1980) [Sov. Phys. Tech. Phys. **25**, 1154 (1980)].
22. E. I. Bondarenko, Z. V. Bondarenko, M. V. Lomakov, and I. P. Raevskiĭ, Zh. Tekh. Fiz. **55**, 967 (1985) [Sov. Phys. Tech. Phys. **30**, 582 (1985)].
23. K. Ishida and G. Honjo, J. Phys. Soc. Jpn. **34**, 1279 (1973).
24. L. A. Reznichenko, L. A. Shilkina, O. N. Razumovskaya, *et al.*, in *Proceedings of the International Workshop "Phase Transformations in Solid Solutions and Alloys," Sochi, 2001*, p. 236.
25. C. N. R. Rao and J. Gopalakrishnan, *New Directions in Solid State Chemistry* (Cambridge Univ. Press, Cambridge, 1986; Nauka, Novosibirsk, 1990).
26. R. Ranjan, N. Singh, D. Pandey, *et al.*, Appl. Phys. Lett. **70**, 3221 (1997).

Translated by V. Isaakyan

Technical Physics, Vol. 47, No. 6, 2002, pp. 777–785. Translated from Zhurnal Tekhnicheskoĭ Fiziki, Vol. 72, No. 5, 2002, pp. 125–134.

EXPERIMENTAL INSTRUMENTS AND TECHNIQUES

IR Synchrotron Diagnostics. Part 1: Wide-Band Long-Focus Optics

A. A. Mal'tsev

Joint Institute for Nuclear Research, Dubna, Moscow oblast, 141980 Russia

e-mail: amaltsev@c.v.jinr.dubna.su

Received March 14, 2001; in final form, October 30, 2001

Abstract—An optical channel for extracting, transporting, and applying (to a detector) IR synchrotron radiation covering a wide wavelength range is developed. The design and study of optical systems with long-focus entrance objective and mirror lenses, as well as with a field condenser lens, are reported. Practical use of IR optics combined with integrated and position-sensitive detectors in accelerator experiments with synchrotron radiation is described. © *2002 MAIK "Nauka/Interperiodica".*

INTRODUCTION

Magneto-bremsstrahlung or synchrotron radiation (also referred to as fringe, undulator, coherent, or intersecting-beam radiation) is a well-known, very promising, and widely used effect. Synchrotron radiation opens the way to the nonintrusive (contactless) diagnostics of elementary particle beams and bunches without affecting their acceleration.

Accelerators currently available generate synchrotron radiation largely in the IR range [1–3], which makes it difficult or totally excludes the use of conventional techniques and tools of optical diagnostics. It is therefore necessary to devise IR diagnostics methods and equipment, as well as associated measuring, information, and detecting units.

At the Joint Institute for Nuclear Research (JINR), a compressor of toroidal (ring) high-current bunches of low-energy relativistic electrons (adgezator)[1] serves as a generator of IR synchrotron radiation [1]. The synchrotron radiation experiments being conducted with the JINR compressor are of great importance, since this radiation is the only source of direct objective (qualitative and quantitative) information on the formation, compression, and acceleration of an electron (electron–ion) bunch. Only IR synchrotron diagnostics provides the observation of a charged particle bunch without its collapse and without affecting fast processes under study.

Detecting, measuring, and information systems for nondestructive diagnostics of relativistic particle bunches in an accelerator or storage ring must detect, record, and analyze particle-generated synchrotron radiation, and eventually provide information on the basic (current, energy, and geometrical) parameters of the bunch, as well as on the fast processes. Because of the severe conditions of accelerator experiments, the creation of synchrotron diagnostics systems poses great technical difficulties and is very expensive.

Of special importance is the design and fabrication of high-sensitivity detecting equipment that allows for the detection of very weak light fluxes from a source rapidly moving within the accelerator chamber and rapidly varying its radiation spectrum. In designing such equipment, emphasis should be given on the optical channel, which provides object tracking. The channel must contain a window for extracting IR synchrotron radiation from the accelerator and an optical path for transporting and delivering the radiation to the detector. This work describes the optical channel designed in the JINR and its basic elements, which have been tested in related experiments.

EXPERIMENTAL CONDITIONS

The specific features of the optics developed and its use in the experiments with IR synchrotron radiation will be demonstrated by an example of the JINR adgezator diagnostics. Figure 1 shows the geometry of emitting and detecting the radiation. The object of diagnostics and study (i.e., an IR radiation source) is toroidal (ring) bunch *1* of low-energy relativistic electrons. Under the action of a pulsed time-growing magnetic field, the electrons spiral in the median plane of the cylindrical chamber of an accelerator–compressor, emitting narrow beam *2* of synchrotron radiation.

For the time of acceleration, the equilibrium (mean) radius R of the bunch decreases from 40 to 4 cm and the electron energy E grows from 2.5 to 20 MeV. For the same time, the maximum λ_m of the spectral distribution of the synchrotron radiation shifts from the rf range to the IR range, $\Delta\lambda \simeq 1$–8 μm. For the latter, there exist a variety of high-sensitivity fast-responding detectors,

[1] Transliteration of the Russian acronym адгезатор (адиабатический генератор заряженных тороидов—adiabatic generator of charged toroids).

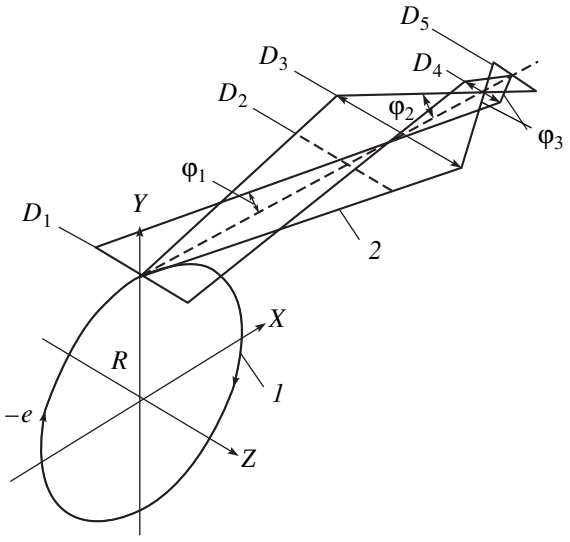


Fig. 1. Propagation and detection of synchrotron radiation. *1*, electron ring; *2*, synchrotron radiation; D_1, field of vision (object plane); D_2, window; φ_1–φ_3, aperture angles; D_3, entrance pupil of objective lens; D_4, condenser; and D_5, exit pupil.

which do not require sophisticated and expensive cryogenic facilities.

The requirements for the detecting part of the measuring system for synchrotron diagnostics depend on the parameters of the electron ring, radiation and technical experimental conditions, as well as the goal of the experiments. To provide the condition for collective acceleration (the electron bunch traps and confines heavy ions) at the final stage of ring compression, the parameters of the ring must be as follows: $R \simeq 4$ cm, linear sizes of the smaller cross section of the ring electron bunch a_R and a_Z = 2–4 mm, electron energy $E \simeq$ 20 MeV, and the number of electrons in the ring $N_e \gtrsim 10^{13}$. These parameters should be measured in the process of acceleration with the use of synchrotron radiation with an accuracy of no worse than 20%. The detectors must be placed as close to the accelerator as possible and be capable of operating under a high level of dc and pulse (with a pulse rate of up to 10 kHz) electromagnetic and radiation interference. The basic goals of the synchrotron radiation experiments are the following. (1) Extraction of the radiation from the high-vacuum chamber of the accelerator (the chamber may be heated) in the working spectral range $\Delta\lambda \simeq 1$–8 μm; (2) determination of the geometry of the toroidal bunch, namely, its radius (R = 4–6 cm), sizes of the smaller cross section (a_R and a_Z = 2–6 mm), as well as the elementary particle density function in the section with a_R and a_Z and the angular distribution of the elementary particles; and (3) determination of the total number of electrons (electron current) in the range $N_e \gtrsim 10^8$–10^{13} from the measured absolute intensity of the synchrotron radiation. The detecting part of the system must include the window, intermediate focusing optics, as well as integrated and multielement position-sensitive detectors. Bearing in mind the above goals and requirements and taking into account preliminary analysis of parameters of the electron bunches and its synchrotron radiation, we selected associated detectors and optical materials for the windows and lenses, and also designed the optical systems.

OPTICAL CHANNEL

Remote nondestructive (nonintrusive) diagnostics of ring bunch *1* by its IR radiation *2* includes the extraction of the radiation from the vacuum chamber of the accelerator through the window of diameter D_2, transport of the radiation along the optical path (system), and delivery of the radiation to the detectors. Intensity losses during the extraction and transport must be minimized throughout the working spectral range.

Detectors of the absolute intensity of the synchrotron radiation must collect the maximum possible amount of radiated energy irrespective of the position of the source (small cross section of diameter D_1 in a given space of the accelerating chamber). This can be provided with an optical system that is highly transparent in the spectral range of detector sensitivity. Also, this system must transmit the entire radiation arriving at the entrance pupil D_3 of the system from the source plane (which is bounded by the field of vision D_1) to the plane of the exit pupil D_5, where the detector is located, by means of the field condenser D_4. This can be done with an optical system based on a Fabry lens. The design procedure for such systems is well known [4, 5]. In the scheme with field condenser D_4 (Fig. 1), where the detector is placed at the exit pupil D_5 of the system, the image does not move over the detector sensitive surface and the exit pupil D_5 is constantly under irradiation even if the angle of beam incidence on the entrance pupil D_2 is varied.

The basic parameters of the system with a field condenser were selected in view of the Lagrange–Helmholtz invariant for image distortion-free transmission:

$$D_3 \sin\phi_1 = D_4 \sin\phi_2 = D_5 \sin\phi_5, \qquad (1)$$

where $\phi_{1\text{–}3}$ are the aperture angles of the objective lens, condenser, and detector, respectively.

For the normal operation of the optical system, it is necessary that the condition $\phi_3 \leq 1$ for the aperture angle of the detector be fulfilled.

Since appropriate IR windows and optics adequate for our problem are lacking, we developed special windows [6] and designed relatively simple optical systems with objective [5] and reflector lenses [7]. Either system can use both a multielement position-sensitive detector for measuring the geometrical parameters of the bunch (equilibrium radius R, cross-sectional sizes a_R and a_Z, electron distribution over the cross section,

etc.) and a precision single-element integrated detector for detecting the synchrotron radiation and determining the total number of electrons in the bunch from the absolute intensity of the radiation.

(1) The optical advantages of IR synchrotron diagnostics over UV and optical diagnostics are as follows. (i) The transparency of the window is not adversely affected by radiation-induced internal and external damage; (ii) the linear resolving power of the IR systems may be low (10 lines/mm) in comparison with that of visible-range optics (about 100 lines/mm); (iii) synchrotron radiation beams are detected in a wide IR wavelength range ($\Delta\lambda/\lambda \gg 1$, where $\Delta\lambda$ = 1–8 μm), unlike optical diagnostics, where $\Delta\lambda/\lambda \ll 1$; (iv) synchrotron radiation beams are narrow and polarized; (v) the source is of low intensity (~10^{-6} W) and small size (~2 mm); and (vi) the location and shape of the source in the field of vision are of no significance.

The use of the IR part of the electromagnetic spectrum for the diagnostics of accelerated beams improves considerably the accuracy of measurement of the absolute intensity of the synchrotron radiation in comparison with visible-range (optical) diagnostics. This is of special importance for precision measurements, e.g., in metrology. This advantage arises because radiation-induced dark spots (damages) on the window do not affect its transmissivity in the IR range, since the sizes of the damages on and inside the window are, as a rule, much less than the IR radiation wavelength. Therefore, they do not absorb or scatter the IR radiation. The visible and UV background from the window is of no concern, because either the detectors are insensitive to these wavelengths or special filters are used. As a result, with IR radiation used, the accelerator windows are easier to fabricate and they are less expensive.

In the IR range, the linear dimensions of the detector become comparable to the radiation wavelength, especially in the far-IR range. The minimal linear dimensions of the position-sensitive elements operating at $\lambda >$ 1 μm are usually no larger than $d_n \gtrsim 0.1$ mm. This circumstance makes it possible to radically (roughly ten times) loosen requirements for the resolving power of the optics compatible with position-sensitive IR detectors. In addition, this eliminates the need for carefully analyzing the optical system and calculating the lens curvature. Eventually, the fabrication of the lenses and the optical system as a whole is greatly simplified and becomes less expensive. In this respect, IR diagnostics is superior to UV and visible-range diagnostics.

(2) The problem of vacuum-tight windows has always been of vital importance in vacuum technology and cryogenics, including charged particle accelerators. The maintenance of a high vacuum is a necessary condition for their operation. Evacuation to a pressure as low as 10^{-9} torr sometimes implies heating of the units (including the radiation extracting window) mounted on the vacuum chamber to 350–400°C. Also, all vacuum seals must have a metallic gasket. Certainly, the window must remain air-tight after heating.

Windows used in magneto-bremsstrahlung diagnostic equipment must be transparent in a wide spectral range; offer high radiation hardness, thermal stability, plasticity, and mechanical strength; as well as provide extremely low gas leakage and gas release. Thus, the production of windows is an important, perhaps decisive, problem in designing the optical path. The cost of the optical material and its availability are also of great concern.

In light of the aforesaid, we tested windows made of various optical materials [6, 8]: fused quartz for the visible and near-IR (to 3.5 μm) ranges, CaF_2 for the visible and mid-IR (to 9 μm) ranges, MgF_2-based polycrystalline optical ceramic for λ = 1–8 μm, ZnS-based polycrystalline optical ceramic for λ = 1–14 μm, as well as ZnSe-based optical ceramic and GeSe-based oxygen-free optical glass for wavelengths up to 20 μm. These materials are highly transparent in the working ranges of the spectrum and are vacuum-tight. The diameter of the windows was varied between 60 and 115 mm.

For the heating of the chamber, the window made of MgF_2 was diffusion-welded to the stainless steel barrel [6]. Such windows are designed for multiple use in heated ultra-high-vacuum installations. The transparency region is $\Delta\lambda$ = 1–8 μm, and the optical diameter is 70 mm at a window thickness of 4 mm. The rate of gas release from the surface at 600°C and a pressure of ~10^{-9} torr does not exceed 7×10^{-7} torr/s.

(3) The lens system [5] was developed with the aim of being mounted directly on the extracting channel of the accelerator. It is intended to serve two functions: to determine the electron bunch geometry (source–objective lens–position-sensitive detector) and to measure the absolute intensity of radiation (source–objective lens–condenser–integrated detector). The optical part with transparent refractive elements (lenses) is made according to the scheme depicted in Fig. 1. Figure 2 shows the arrangement of the basic elements of the system with an objective lens and a field condenser. The optical channel has IR window *1* of diameter D_2 (Fig. 1), objective lens *2* (which simultaneously serves as the entrance pupil D_3 of the system), deflecting plane mirrors *3* and *4*, field condenser *5* in the plane of the image produced by the objective lens, and detector *6* in the plane of the exit pupil D_5.

All the lenses used in the optical system are spherical. The radii of curvature R_1 and R_2 depend on the focal length f, as well as on the thickness h and refractive index n of the lens material. These four parameters are related as [9]

$$\frac{1}{f} = (n-1)\left(\frac{1}{R_1} - \frac{1}{R_2}\right) + (n-1)^2 \frac{h}{n} \frac{1}{R_1 R_2}. \qquad (2)$$

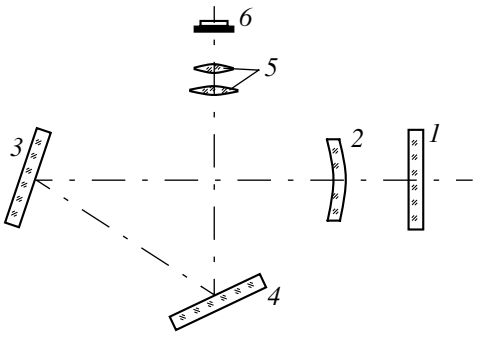


Fig. 2. Arrangement of the basic elements of the system with the objective lens.

The lens thickness can be neglected if the focal length-to-diameter ratio is much greater than unity. Then, formula (2) is recast as

$$1/f = (n-1)(1/R_1 - 1/R_2).$$

The radii of curvature were selected from the condition that the magnification due to spherical aberration be minimal. The refractive index value corresponded to the wavelength $\lambda = 5.5$ μm. For a simple lens, the spherical aberration is minimal when the radii of curvature of the lens surfaces are related as [4]

$$\frac{R_2}{R_1} = \frac{2n^2 + n}{2n^2 - n - 4}. \quad (3)$$

This relationship holds if

$$R_1 = \frac{2f(n-1)(n+2)}{n(2n+1)}. \quad (4)$$

The table lists the basic parameters of the lenses made of different IR materials when the spherical aberration is minimal. The minus sign at the radius R_2 means that the center of the circle is placed backward to the object of observation or study (i.e., on the right of the lens).

The lenses were made of specially developed optical ceramics. The materials of choice were MgF_2 and ZnS polycrystalline ceramics with high refractive indices, which define the spherical aberration. The use of the high-index optical materials made it possible to increase the optical diameter of the condenser lens to 19 and 23 mm, respectively. Because of this, the field-of-vision diameter in the plane of the smaller cross section of the torus becomes greater than 30 mm.

The long-focus lenses were designed largely for objective or intermediate lenses. The short-focus ones were used mainly as field condensers.

For all the lenses listed in the table and for the entire optical system (including the window, mirrors, objective lens, and condenser lens), we thoroughly studied the aberrations, resolving power, field of vision, and transmission spectrum.

The investigation of IR optics made of opaque optical materials is more complicated than similar investigations in the visible range. For this purpose, we elaborated a special technique [5] based on temperature quenching of photoluminescence. With this effect, the result of IR irradiation becomes visible. This technique includes several steps. First, the spatial distribution of the IR beam intensity is visualized with a flat luminescent screen. The phosphor is pre-excited by an UV source. When the IR beam passes through the absorbing surface, its temperature increases and the phosphor brightness drops. The spatial distribution of the radiation intensity, as well as the shape and geometry of the beam, can be judged from the degree of darkening of the luminescent screen. The temperature-quenchable luminescent screen, as such, is a nonselective detector sensitive to the range from microwaves to the near infrared. The working spectral range is specified by band-pass filters, the transmission or reflection coefficient of the optical elements in the path of the radiation, and the spectral responsiveness of the detector.

The resolving power of lenses depends on diffraction effects and aberrations. The effective diameter of a circle of confusion due to diffraction in the case of a point source is given by

$$d_{\text{dif}} = 1.22N\lambda, \quad (5)$$

where $N = f/D$ is the focus number (f, focus length; D, optical diameter of a lens).

Spherical aberration has the strongest effect on the magnification. For a simple lens with two spherical surfaces, the apparent image diameter (including an extra magnification due to spherical aberration) is approximated by the formula [4]

$$d_{\text{sph}} = 7.8 \times 10^{-3} N^{-3} A, \quad (6)$$

where the spherical aberration factor is given by

$$A = \frac{n+2}{n(n-1)^2}\sigma^2 - \frac{4(n+1)}{n(n-1)}\sigma + \frac{3n+2}{n} + \frac{n^2}{(n-1)^2},$$

$$\sigma = (R_2 + R_1)/(R_2 - R_1).$$

We studied the effect of various aberrations (spherical aberration, coma, and astigmatism) on the image size. It was found that diffraction and aberration are basic adverse effects in the spectral range $\Delta\lambda = 1$–8 μm. Moreover, diffraction limits the resolving power of the optical system. When calculating the radii of curvature of the lens surfaces, we took into consideration the spherical aberration alone. The influence of the other aberrations on the resolving power is negligibly small, while taking them into account makes the design and production of the lenses difficult. For the MgF_2 long-focus lens with the focus number $N = 5.3$, the diameter of a circle of confusion in the plane of the point source image is $d_{\text{cal}} = 0.43$ mm for a wavelength $\lambda = 5.5$ μm. This value, including the effect of diffraction and spher-

Parameters of the lenses made of different IR materials

Optical material	Lens parameters					
	$\Delta\lambda$, μm	$n_{\lambda = 5.5\ \mu m}$	f, mm	R_1, mm	R_2, mm	D, mm
MgF_2	1–8.5	1.3282	320.0	144.8	–394.0	60
ZnS	1–14.5	2.2416	320.0	288.0	1050.0	60
ZnS	1–14.5	2.2416	17.5	14.57	47.02	19
ZnS	1–14.5	2.2416	22.0	19.0	61.34	23
CaF_2	0.13–12	1.3952	40.0	21.0	–74.5	27

ical aberration, was found from formulas (5) and (6). The resolving power of the MgF_2 lens with $f = 320$ mm is presented in Fig. 3, where the point source size measured along the X and Y coordinates is shown. On the ordinate, the signal picked up from the detector as it moves in the point source image plane is plotted. The abscissa axis plots the real (apparent) point source size including the aberration and diffraction effects. The half-height diameter of the circle of confusion is $d_{real} \simeq 0.45$ mm. Thus, the resolving power of the MgF_2 long-focus lens with $f = 320$ mm completely corresponds to that of the multielement position-sensitive InSb photoresistors with a scale division of 0.4–1.2 mm.

One specific feature of the technique used to study the effect of diffraction and aberration should be noted. Usually, the resolving powers of individual lenses and an optical system as a whole are studied in a narrow wavelength range $\Delta\lambda/\lambda \ll 1$, where λ is the wavelength to which the refractive index n of the lens material corresponds. When included into calculations, the refractive index affects the curvature of the refracting surfaces. Moreover, it contributes to an extra magnification due to spherical aberration according to formulas (3), (4), and (6). Unlike the conventional approach, in this specific case we used the broad band $\Delta\lambda/\lambda = 1$–8 μm. The width of the range depends on the transmission spectrum of the optical elements (window and lenses) incorporated into a particular optical system and on the spectral responsiveness of the detectors.

In spite of the wide spectral range and the simplicity of the refracting surfaces of the lenses, our results (Fig. 3) indicate that diffraction and spherical aberration have a minor effect on the enlargement of the point source image; hence, they play an insignificant role in imaging in the plane of the position-sensitive elements of the detector.

To detect the synchrotron radiation, it was very important to know the linear size D_1 of the field of vision (Fig. 1) in the plane of object movement, i.e., where the median XY plane of the bunch crosses the YZ plane of its cross section. Associated studies were carried out, and the field of vision was measured. A point source of IR radiation moved in the focal plane of the objective lens, and the radiation was focused by the condenser in the plane of the photodetector, which converted the IR radiation to an electrical signal. Preliminary, the image plane of the entrance pupil was found with a temperature-quenchable luminescent screen. For the system with the CaF_2 condenser ($f = 40$ mm), the field-of-vision size is $D_1 = 24$ mm; with the ZnS condenser ($f = 17.5$ mm), $D_1 = 32$ mm. This totally meets the conditions for accelerator experiments. Thus, the parameters of the individual lenses and the optical systems as a whole that were measured experimentally with the temperature-quenchable luminescent screen agree well with the calculations.

We also studied the optical efficiency of the system with the MgF_2 condenser ($f = 320$ mm) and the ZnS condenser ($f = 17.5$ mm) in terms of window and lens losses, as well as reflection losses. It turned out that its working range (where the losses are the least) lies in the near and middle infrareds ($\Delta\lambda = 1$–8 μm), which fully meets the conditions for accelerator experiments. The transmission peak of the system corresponds to the spectral responsivenesses of the IR photodetectors used in the accelerator experiments.

Thus, the primary parameters of the base optical system with the objective lens (Fig. 2) are as follows: working spectral range $\Delta\lambda = 1$–8 μm, focal length of the MgF_2 objective lens $f_{obj} = 320$ mm, entrance pupil diameter $D_{ent} = 60$ mm, relative aperture of the objective lens $D_{ent}/f_{obj} = 1/5.3$, focal length of the ZnS condenser $f_{con} = 17.5$ mm, exit pupil diameter $D_{ex} = 2$ mm, field-of-vision size in the object plane $D_{f.v} = 32$ mm, and overall dimensions are $290 \times 300 \times 100$ mm.

(4) A mirror periscope is a specific optical channel [7] with an entrance objective lens composed of spherical mirrors. It provides the wide-band operation of the entire optical system and allows the detection of synchrotron radiation in the spectral range $\Delta\lambda = 0.3$–40 μm. A periscope is used when the detector must be placed beyond the range of the electromagnetic and radiation fields of the accelerator.

Multielement detectors must be reliably protected from the accelerator interference. Interference is caused largely by pulsed magnetic fields and gamma radiation. To remove the interference, it is necessary that the position-sensitive detector, on which the full-size image of the radiation source is focused, be at a distance of no closer than 2 m from the radiation

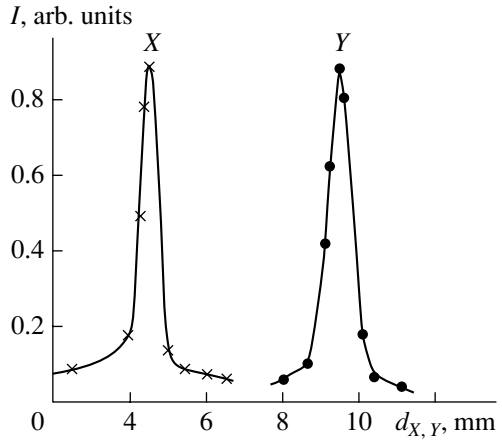


Fig. 3. Resolving power of the MgF_2 objective lens.

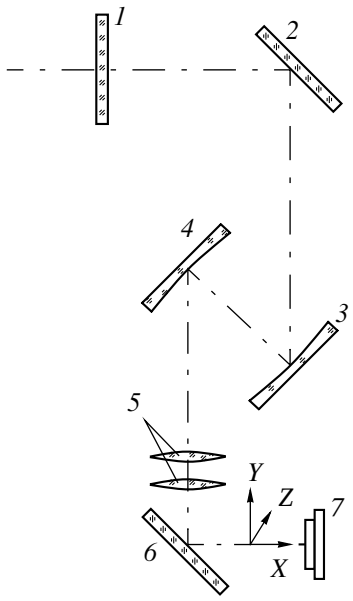


Fig. 4. Optical scheme of the periscope channel with the mirror objective.

source. This led us to develop an optical channel with a long-focus reflector entrance objective.

The wide-band operation of the channel is provided by the fact that it is composed only of reflecting elements (mirrors). The mirrors are made of optical glass and have a surface of a given curvature, on which a silver layer is applied by evaporation *in vacuo*. The channel is used in laboratory conditions under constant temperature and humidity; therefore, the metal layer is not additionally protected to minimize transmission losses. The short-wave edge of the spectral range depends on the reflecting surface finish and the coating material. The long-wave edge, being limited by diffraction, also depends on the relative aperture of the system and imaging wavelength.

The optical scheme of the periscope with the mirror objective lens is depicted in Fig. 4. Since the synchrotron radiation forms a narrow cone, only the section of the electron ring perpendicular to the optical axis of the system is seen. The radiation from the small cross section (≈1/60 of the entire section) of the electron ring is extracted from the vacuum chamber of the compressor through IR window *1*, in the neighborhood of which flat mirror *2* is placed. Mirror *2* deflects the divergent beam of the synchrotron radiation in the horizontal direction. First spherical mirror *3* is placed so that the object of observation is at its focus. With such an arrangement, the divergent beam becomes parallel to the optical axis and can travel any given distance. The torus section is imaged in the focal plane of second spherical mirror *4*, where detector *7* (its sensitive surface) or condenser *5* is located. Mirror *6* rotates the optical axis through 90° in the vertical plane. The focal length of both spherical mirrors is the same, 1850 mm. Diaphragms eliminating flare spots and stray radiations are placed along the path of the beam being focused.

Deflector *2* rotates the optical axis through 90°. Initially, it had a cylindrical surface to correct for the astigmatism of the spherical mirrors, which handle the inclined beam. To optimize the image quality, we studied the optical scheme of the system by the method of frequency–contrast characteristics. It was found that the best image is provided if the deflector has a flat, rather than cylindrical, surface. It is likely that the effect of astigmatism in this system is weaker than that of other aberrations (coma, spherical aberration, etc.). The frequency–contrast characteristics obtained are shown in Fig. 5. The spatial frequency ν (the number of lines per 1 mm of the image) is plotted on the abscissa axis; the change in the image contrast *K* relative to the object, on the ordinate axis. With a level of 0.02 (visual resolution) taken as the sign of contrast decrease, the resolving power of the optical system equals 5–6 lines/mm when the cylindrical mirror is placed in the meridional plane of the system (the plane of drawing). As follows from the calculations, this value is ultimate for the cylindrical mirror. With flat mirror *3* employed, the resolving power rises to 9–10 lines/mm. In Fig. 5, the upper curves (crosses) refer to the central part of the field of vision; the lower ones (circles), to the periphery of the field. The absence of chromatic aberration allows the adjustment of the device in the visible part of the spectrum.

An advantage of this scheme is low transmission losses (less than 4%). This is associated with the high reflection coefficient of silver at wavelengths between 0.3 and 40 μm and also with the absence of protective coatings on the reflecting surfaces. Losses of the synchrotron radiation intensity are inserted mainly by the IR window (up to 40% of the total losses) and detector window materials.

The channel is compatible with various cooled and uncooled IR detectors, particularly, with multielement

position-sensitive silicon detectors, InSb detectors (the working temperature T_w = 77 K), PbSe detectors (T_w = 250 K), pyroelectric detectors, etc. The wide-band operation of the mirrors allows them to be used in the UV and visible ranges. In addition, this optical channel can be used in the visible and near-IR ranges in combination with an SFR streak camera.

The optical channel with the mirror objective lens and condenser made it possible for the first time to detect synchrotron radiation in the JINR accelerator–compressor. In pioneering experiments, the radiation intensity (total number of electrons) was so low and the radiation spectrum was so uncertain that the radiation was impossible to detect without optical amplification and wide-band detectors.

The optical channel with the mirror objective lens is designed for the UV, visible, and IR ranges (the operating range is limited only by diffraction and mirror coatings) and has the following basic parameters: focal length of the spherical mirrors f = 1850 mm, relative aperture 1 : 21, magnification ×1, photographic resolution 7 mm^{-1}, and overall dimensions 2000 × 360 × 370 mm.

(5) The design of the optical systems enables them to be used as base devices for accelerator diagnostics and in experiments unrelated to accelerators. Structurally, the lens optical system (Fig. 2) is made as a separate unit. It can be mounted directly on the flange of a special branch pipe from which the radiation is extracted from the vacuum chamber of the accelerator.

The periscope channel (Fig. 4), provided with the rigid tubular support, is mounted as a separate unit on a concrete wall or cube. In other words, it must be attached to a vibration-free foundation.

All blocks and parts of the optical systems are made of nonmagnetic materials to meet conditions of physical accelerator experiments. The lens and mirror barrels are made of brass. The optical parts are mounted into the barrels with appropriate gaskets to avoid optical material–metal contact. The body of the lens system was made by turning and milling a D16t aluminum alloy slab to make the system monolithic and thereby provide a reliable attachment of all the units responsible for optical accuracy. Moreover, with such a construction, there is easy access to the basic optical elements and the system is easy to adjust and tune to a detector.

The lenses and mirrors can move along the optical axis. In addition, the inclination of the mirrors can be varied over wide limits in order to finely adjust the image plane, which is produced by the objective lens in the principal plane of the condenser or on the sensitive surface of the detector.

Once assembled and adjusted, all optical and mechanical units and parts of the system are mounted rigidly by means of locating pins. The body and fixings of the optical units are painted mat black to prevent flare spots and scattered light from penetrating to the image plane. The adjusting units of the optical systems are protected from environmental (optical and mechanical) effects by removable housings made of aluminum alloy sheets.

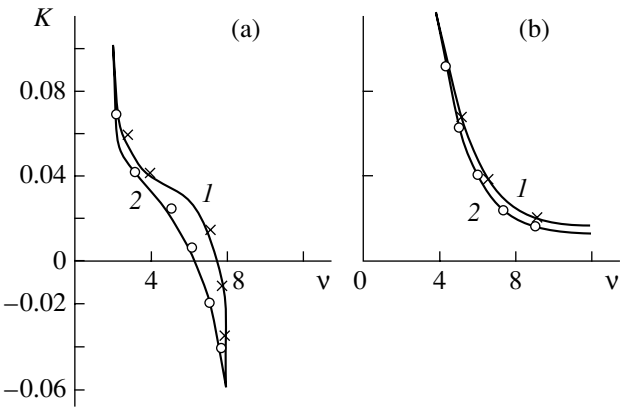


Fig. 5. Frequency–contrast characteristic of the optical channel with (a) cylindrical and (b) flat deflector. ν, the number of lines per 1 mm of the image; K, change in the image contrast relative to the object. (*1*) Central part and (*2*) periphery of the field of vision.

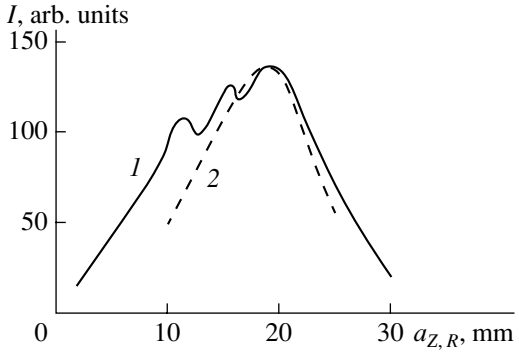


Fig. 6. Cross-sectional profile of the electron bunch and the distribution of the electron density over its cross section. (*1*) a_Z and (*2*) a_R.

APPLICATIONS

Thus, the IR optics described in this work has played a crucial role in studying the possibility of generating electron and electron–ion bunches with parameters optimal for the collective acceleration of electrons and heavy ions in various JINR compressors [1].

To solve the problem of synchrotron radiation detection, the author has designed the approach described above, which makes it possible to detect (with minimum losses) synchrotron radiation with detectors made of the appropriate materials and covering the entire spectrum of the radiation. Other advantages of this approach are as follows: (1) more than 20-fold amplification of the synchrotron radiation flux density on the sensitive element of the detector (hence, amplification

of detector signal) by applying the optical scheme with the field condenser, (2) use of precision detectors whose sensitivity is the highest in the working spectral range $\Delta\lambda$ = 1–8 μm, and (3) detection of the synchrotron radiation throughout the sensitivity range of these detectors ($\Delta\lambda/\lambda \gg 1$).

With this approach, synchrotron radiation was discovered for the first time and used to study relativistic electron bunches. The following are results illustrating the application of the long-focus wide-band mirror-lens optics in accelerator experiments.

From the maximal radiation signal, we located the ring bunch relative to the median plane *YZ* and the *Z* axis of the accelerator, measured the equilibrium radius *R*, found the bunch profile (a_R, a_Z), and studied the electron distribution over the cross section. It turned out that the median planes of the torus and accelerator coincide and the energy of the electrons and the radius of their orbit (accordingly, the radiation spectrum and intensity) are close to the calculated values. However, the size of the bunch, the number of electrons, and the distribution of the electrons over the cross section differ greatly from those required.

From the synchrotron radiation intensity, for the first time we determined the absolute number of electrons in an electron bunch formed in the JINR electron-ring compressor. The actual number of electrons in the bunch was found to be smaller ($N_e \sim 10^7$–10^8 electrons) than expected, and the bunch had the shape of a diffuse cylindrical helix rather than a ring.

The method of determining the instantaneous ($\tau <$ 0.1μs) geometrical parameter of the ring in the compressor relies on advances in IR optics and IR multichannel detecting technology [7]. Its essence is that, after it has fallen into the field of vision, the image of the smaller section of the ring is focused by the objective lens on the sensitive surface of a detector with a linear array of the elements. Figure 6 shows the instantaneous sizes of the smaller cross section of the electron bunch (scale division 1.2 mm) and the cross-sectional distribution of the electrons.

Thus, the difficulties associated with the detection of synchrotron radiation were due primarily to its low density on the sensitive surface of the detector. Therefore, the valid signal of the detector was hard to discriminate from intense electromagnetic and radiation fields of the accelerator. The problem has been solved by employing the diagnostic method developed, taking advantage of the unique properties of the IR optics, and using high-sensitivity detectors. Based on the information gathered, we upgraded the design of the accelerator units, which helped us to raise the ring current (number of electrons) from $N_e \sim 10^8$ to ~10^{12} electrons. Measurements of the absolute intensity of the synchrotron radiation allow the determination of the number of electrons in a bunch at the final stage of compression with an accuracy of no worse than 20%.

At the closing stages of the experiments on collective acceleration, a ring bunch with $N_e \sim 10^{12}$ electrons was obtained, the electron density distribution being fairly uniform. However, in spite of the considerable effort that went into upgrading the magnetic systems of the accelerator, we failed to generate an electron bunch with parameters optimal for collective acceleration. Today, in view of the results of IR synchrotron diagnostics, experiments on collective acceleration have been terminated, although the JINR compressor remains to be a unique source of synchrotron radiation in terms of its parameters [10].

CONCLUSION

IR diagnostics offers a number of advantages over visible-range diagnostics. The basic one is that requirements for the image size, i.e., for the resolving power of the system, are not so stringent.

Modern optical materials, specifically, optical ceramics, of which windows and lenses have been made, offer good radiation hardness, mechanical strength, and vacuum tightness. Moreover, they are highly transparent over a wide spectral range, which favors the extraction of IR synchrotron radiation from the vacuum chamber and its transport to the detector with minimal intensity losses.

A unique technique for window–barrel diffusion welding has been developed. The design of the barrel provides for the sealing of the vacuum window with the help of a metallic gasket.

In spite of the simplicity of the mirror and lens optics developed, we succeeded in discovering and detecting synchrotron radiation in the JINR (with the number of electrons in the ring about 10^8). The parameters of compressible electron bunches have been determined from their synchrotron radiation in the near and middle infrareds with a desired accuracy. The resolving power of the systems is adequate to that of the detectors used.

The optical channel totally meets the requirements for accelerator experiments and is compatible with various detectors for wide-band IR diagnostics in both pure scientific purposes and applications.

Also of importance are bench tests with IR optics, where the thermoluminescence effect was used. With such an approach, we succeeded in studying optical systems with lenses that are nontransparent in the visible range.

REFERENCES

1. A. A. Mal'tsev and M. A. Mal'tsev, Zh. Tekh. Fiz. **67** (4), 60 (1997) [Tech. Phys. **42**, 378 (1997)].

2. A. A. Mal'tsev and M. A. Mal'tsev, At. Énerg. **80** (3), 190 (1996).
3. É. A. Kuraev and A. A. Mal'tsev, Preprint No. R9-97-242, OIYaI (Joint Inst. for Nuclear Researches, Dubna, 1997).
4. E. M. Voronkova, B. N. Grechushnikov, *et al.*, *Optical Materials for Infrared Technology* (Nauka, Moscow, 1965).
5. A. A. Mal'tsev and M. A. Mal'tsev, Izmer. Tekh., No. 11, 17 (2000).
6. A. A. Mal'tsev, Prib. Tekh. Éksp., No. 3, 177 (1994).
7. A. A. Mal'tsev, Fiz. Plazmy **23**, 419 (1997) [Plasma Phys. Rep. **23**, 389 (1997)].
8. H. Ebert, *Physikalisches Taschenbuch* (F. Vieweg, Braunschweig, 1957; Fizmatgiz, Moscow, 1963).
9. K. M. Scott, Proc. IRE, No. 9, 1530 (1959).
10. A. A. Mal'tsev, Fiz. Élem. Chastits At. Yadra **27**, 797 (1996) [Phys. Part. Nucl. **27**, 330 (1996)].

Translated by V. Isaakyan

Technical Physics, Vol. 47, No. 6, 2002, pp. 786–787. Translated from Zhurnal Tekhnicheskoĭ Fiziki, Vol. 72, No. 5, 2002, pp. 135–136.

BRIEF COMMUNICATIONS

Effect of Magnesium Isovalent Impurity on the Temperature Stability of Blue Luminescence in Zinc Selenide

V. P. Makhniĭ, M. M. Slyotov, and Yu. Ya. Chaban

Fed'kovich National University, ul. Kotsyubinskogo 2, Chernovtsy, 58012 Ukraine

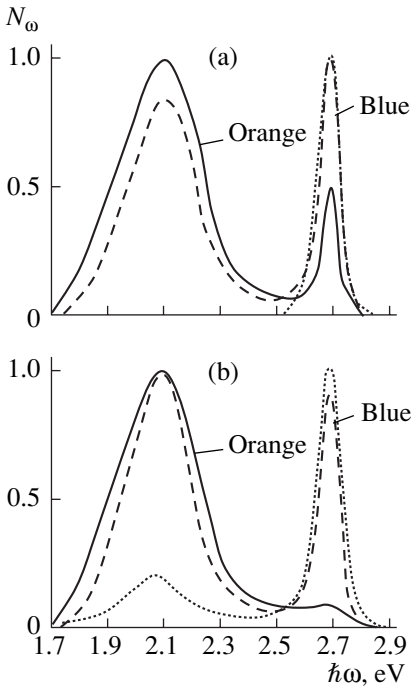
e-mail: oe-dpt@chnu.cv.ua

Received July 10, 2001

Abstract—The effect of Mg isovalent impurity on photoluminescence in undoped and Al(donor)-doped ZnSe crystals is investigated. The anomalously weak temperature quenching of the blue band in Mg-containing samples is detected. © *2002 MAIK "Nauka/Interperiodica".*

Doping by isovalent impurities may strongly improve those physical properties of semiconductors that are difficult or even impossible to improve by doping with conventional impurities [1]. In the first place, isovalent impurities increase the probability of radiative recombination and improve its temperature stability. Note that the discovery of intense edge luminescence in GaP⟨N⟩ [2] has provided the basis for producing commercial green LEDs. For blue (dark blue) radiation sources, it is necessary to use a wider gap material. Moreover, it is desirable that the material be direct-gap. Zinc selenide, having the appropriate gap width (E_g = 2.7 eV at 300 K [3]), meets these requirements. Therefore, most investigations are aimed at developing methods of obtaining low-resistivity ZnSe where edge luminescence prevails [4] for designing various electroluminescent structures. One way of solving this problem was suggested in [5], where doping of ZnSe by Mg isovalent impurity was shown to increase significantly the intensity of the blue luminescence band at 300 K. This paper studies the behavior of this band in diffusion layers of ZnSe⟨Mg⟩ at high temperatures.

Initial substrates were made of bulk ZnSe crystals of the two types that are used most frequently [3]. They were grown in a noble gas atmosphere from the stoichiometric melt (type 1), as well as from the melt containing ≈0.1 mol % of Al (type 2). The electron conductivity σ_n was (at 300 K) 10^{-10} and 10^{-3} Ω^{-1} cm^{-1} for type-1 and type-2 crystals, respectively. At room temperature, the photoluminescence spectrum contains two bands: yellow orange and blue (see figure). The lower energy emission is associated with donor–acceptor pairs consisting of deep acceptor ($E_a \approx 1.2$ eV) and shallow donor ($E_d \approx 0.02$–0.03 eV) levels [3, 4]. The former are caused by doubly-charged negative vacancies of zinc (V''_{Zn}); the latter, by singly charged positive vacancies of selenium ($V^{\bullet}_{Se}$) and/or by an Al substitutional impurity ($Al^{\bullet}_{Zn}$). $V^{\bullet}_{Se}$ vacancies that do not enter into the donor–acceptor pairs are responsible for the blue luminescence, resulting from the recombination of free holes with electrons localized on these centers [3]. The virtual absence of the blue band in the initial type-2 crystals indicates that the concentration of "free" Se vacancies is low. The number of these vacancies can be increased, in particular, by healing the V_{Zn} vacancies, which must change the relation between the orange and blue band intensities in favor of the latter. This is observed for crystals of both types (see figure). The



Photoluminescence spectra of the ZnSe samples at 300 K. Solid lines, initial samples; dashed lines, samples annealed in Zn vapor; and dotted lines, samples annealed in Zn and Mg vapors. (a) Type-1 and (b) type-2 crystals.

1063-7842/02/4706-0786$22.00

intensity of the blue radiation can be increased further by doping with Mg isovalent impurity. Note that, in Mg-doped type-1 crystals, the orange band is observed at none of the excitation levels L, which could be varied in the range of 10^{15}–10^{18} photons/s in our experiments. The incomplete quenching of the low-energy emission for the ZnSe⟨Al⟩ substrates indicates that doping by Mg does not heal all the V_{Zn} vacancies. Some of them remain to be incorporated into the donor–acceptor pairs, which form at temperatures (the temperature of crystal synthesis is $T_s \approx 1450$ K) higher than that of doping by the isovalent impurity (1070 K). Since the thermodynamic growth conditions are virtually the same for the crystals of both types, one should assume that the complexes with Al_{Zn} are more stable. Note that the ZnSe⟨Mg, Zn⟩ diffusion layers, as well as the substrates, have electron conductivity, but the value of σ_n in the former grows to 10^{-1} Ω^{-1} cm^{-1}. Thus, Mg doping of the ZnSe crystals with different ensembles of intrinsic and impurity defects results in a significant increase in the blue edge emission and suppresses the self-activated orange band.

Note also that the position of the blue band maximum remains virtually unchanged for all the samples: $\hbar\omega_m \approx 2.68$ eV at 300 K. However, its behavior in the layers containing magnesium and in the initial substrates, as well as the substrates doped by Zn alone, greatly differs. In the latter case, the position $\hbar\omega_m$ of the maximum is independent of the excitation level when it varies within three orders of magnitude. The difference between $\hbar\omega_m$ and the energy gap width E_g is equal to 0.02 eV, which correlates with the energy depth of $V_{Se}^{\bullet}$. Such a behavior is typical when free carriers pass to a discrete level by recombination [3]. In contrast, the maximum of the blue band for the ZnSe⟨Mg⟩ layers shifts toward lesser $\hbar\omega$ with a rise in L. In addition, the increase in the temperature and the excitation level smooth out the low-energy tail of the blue band. When combined, these factors testify to the annihilation of excitons when they are inelastically scattered by free charge carriers [6]. There is no doubt that the blue band is due to the isovalent Mg impurity; however, any certain conclusions about the exciton nature are beyond the scope of this work and requires special investigation, especially at low temperatures.

Indirect corroboration of Mg participation in the formation of the blue band is its high temperature stability. For example, with T increasing from 300 to 500 K, the blue emission intensity I_b in ZnSe⟨Mg⟩ decreases less than twice. At the same time, the doping of the crystals by Zn alone causes a noticeable temperature quenching of I_b, although this is somewhat smaller than for the initial substrates. The main results of the temperature investigations are listed in the table. Here, the symbol I_o denotes the intensity of the orange luminescence band.

Table

Dopant	I_b/I_o		I_b^{300}/I_b^{500}
	300 K	500 K	
Undoped	0.5	0.3	17
Zn	1.2	1.1	7
Zn, Mg	–	–	1.8
Al	0.07	0.05	13
Al, Zn	0.9	0.7	6
Al, Zn, Mg	5	6	2

Note that the position of the edge emission maximum is the same for all the samples at each particular T. The temperature dependence of $\hbar\omega_m$ is linear in the range from 300 to 500 K, and its slope, 6.5×10^{-4} eV/K, is close to the temperature coefficient gap width variation for ZnSe, $\gamma_{E_g} \approx 7 \times 10^{-4}$ eV/K [3, 4].

Thus, efficient temperature-stable blue luminescence can be excited by doping zinc selenide with Mg isovalent impurity.

REFERENCES

1. V. K. Bazhenov and V. I. Fistul', Fiz. Tekh. Poluprovodn. (Leningrad) **18**, 1345 (1984) [Sov. Phys. Semicond. **18**, 843 (1984)].
2. E. F. Gross and D. S. Nedzevitskiĭ, Dokl. Akad. Nauk SSSR **152**, 309 (1963) [Sov. Phys. Dokl. **8**, 896 (1964)].
3. D. D. Nedeoglo and A. V. Simashkevich, *Electrical Luminescent Properties of Zinc Selenide* (Shtiintsa, Chisinau, 1984).
4. *Physics of II–VI Compounds*, Ed. by A. N. Georgobiani and M. K. Sheĭnkman (Nauka, Moscow, 1986).
5. M. M. Slyotov, Pis'ma Zh. Tekh. Fiz. **27** (2), 48 (2001) [Tech. Phys. Lett. **27**, 63 (2001)].
6. E. Koh and D. V. Langer, J. Lumin. **1–2**, 514 (1970).

Translated by Yu. Vishnyakov

Technical Physics, Vol. 47, No. 6, 2002, pp. 788–791. Translated from Zhurnal Tekhnicheskoĭ Fiziki, Vol. 72, No. 5, 2002, pp. 137–140.

BRIEF COMMUNICATIONS

Analysis of Gain and Loss Anisotropy in the Guiding Structure of a Long-Wave Intervalley-Transfer Laser

Yu. A. Morozov*, I. S. Nefedov*, V. N. Gusyatnikov*, and V. Ya. Aleshkin**

** Institute of Radio Engineering and Electronics, Russian Academy of Sciences (Saratov Branch), Saratov, 410019 Russia*

e-mail: mor@ire.san.ru

*** Institute of Physics of Microstructures, Russian Academy of Sciences, Nizhni Novgorod, 603600 Russia*

e-mail: aleshkin@ipm.sci-nnov.ru

Received August 9, 2001

Abstract—Anisotropy of the gain and loss associated with different TM-wave components in a long-wave intervalley-transfer laser is analyzed numerically. The effect of field leakage into the substrate and of interference therein on the optical confinement coefficient and loss coefficient due to absorption by free charge carriers is studied. © *2002 MAIK "Nauka/Interperiodica".*

INTRODUCTION

At present, coherent radiation sources operating in the far- and mid-IR ranges are attracting great interest, hence, the considerable effort in developing these devices worldwide. This problem is topical because these lasers can be used in spectroscopy, environmental studies, and communication. The list of semiconductor lasers designed for the 10–15-μm range is actually limited to devices of two types: quantum cascade [1] and fountain [2] lasers. For a number of reasons (complex structure of the quantum wells, need for pumping, and failure to operate in the continuous-wave mode at room temperature), these lasers do not meet current requirements.

Several years ago, a new version of intraband-transition lasers, a hot-carrier quantum-well laser [3], was proposed by researchers at the Institute of Physics of Microstructures (Russian Academy of Sciences, Nizhni Novgorod). In these lasers, carriers are heated by an electric field directed along the layers of the quantum-well structure. The volume gain of the lasing medium in lasers based on hot electron intervalley transfer was estimated at 50–200 cm^{-1} for a wavelength of ≈10 μm [3], which indicates the feasibility of the device.

MODEL AND BASIC RELATIONSHIPS

Since the current in the laser under study passes along the layers of the structure, it can be built around a multilayer waveguide with confining undoped AlAs layers grown on an insulating GaAs substrate (Fig. 1). The lasing layer consists of alternating quantum-size GaAs films, which are wells for Γ valleys, and AlAs films, which are wells for *X* outer valleys. It has been shown [3] that, in a strong electric field, electrons are efficiently accumulated at those levels of *X* outer valleys located in the Brillouin zone along the growth direction. This causes a population inversion between these levels and the GaAs central valley [3]. The wave functions of these electrons overlap significantly. Therefore, direct transitions can occur at the heterojunctions, which provides amplification. The typical thickness of the GaAs films in each of the periods of the superlattice is 80 Å; that of the AlAs films, 20 Å. The overall thickness of the lasing layer, consisting of 100 superlattice periods, is about 1 μm.

It is significant that amplification in this structure is possible only for those waveguide eigenmodes that have the electric field component directed across the lasing region, i.e., for the TM modes. Our laser model

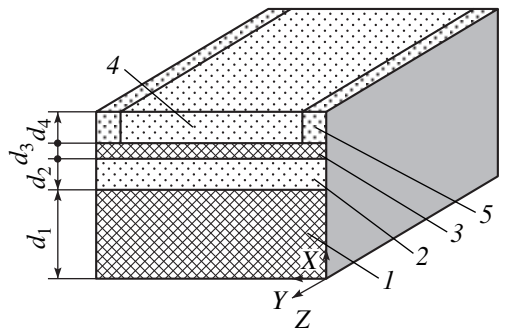


Fig. 1. Cross section of the laser: (*1*) substrate; (*2*, *4*) lower and upper confining layers, respectively; (*3*) active layer; and (*5*) contacts.

1063-7842/02/4706-0788$22.00

assumes that absorption by free carries (Drude absorption) is the basic reason for the attenuation of the TM modes propagating in the structure. It is known that the attenuation due to this mechanism is proportional to the square of the wavelength and may become noticeable for a long-wave laser if the carrier density is sufficiently high.

Since the electric field of the TM mode contains components both across and along the propagation direction (i.e., along the structure layers), it is clear that the gain and loss in the lasing layer must be estimated with allowance for its anisotropy. Specifically, the amplification of the lasing medium is associated with the transverse component of the TM mode field, while the loss is attributed to the longitudinal component because the carriers remain free in the plane of the quantum wells. The attenuation in other parts of the structure (confining layers and substrate) can be regarded as direction-independent, i.e., isotropic. Assume that the gain and the loss in the structure are low; i.e., $g/\beta \ll 1$ and $\alpha/\beta \ll 1$, where g is the volume gain for a plane wave in an infinite medium; α is the power loss factor; and β is the propagation constant, which characterizes the phase variation with distance.

In terms of the perturbation method [4, 5], the attenuation is given by

$$\alpha = \frac{\omega}{2}\epsilon_0 \frac{\int_{-\infty}^{+\infty} (\epsilon''_{xx}|E_x|^2 + \epsilon''_{zz}|E_z|^2)dx}{\int_{-\infty}^{+\infty} S_z dx}, \qquad (1)$$

where ϵ''_{xx} and ϵ''_{zz} are the imaginary parts of the permittivity tensor components for the layers, $S_z = 0.5\,\mathrm{Re}|[\mathbf{EH}^*]|_z$ is the time-averaged magnitude of the Poynting's vector in the z direction (brackets mean the vector product), and ω is the circular frequency. Integration is performed across the waveguide, i.e., in the direction of the layer growth.

As is usual in the perturbation method, the complex amplitudes of the electric, **E**, and magnetic, **H**, fields are calculated under the assumption that the waveguide structure is lossless. Let us rearrange expression (1) so that it contains only the field components that are continuous at the interfaces:

$$\alpha = \beta \frac{\int_{-\infty}^{+\infty} \left[\frac{\epsilon''_{xx}}{(\epsilon'_{xx})^2}|H_y|^2 + \left(\frac{k_0}{\beta\rho}\right)^2 \epsilon''_{zz}|E_z|^2\right] dx}{\int_{-\infty}^{+\infty} \frac{|H_y|^2}{\epsilon'_{xx}} dx}. \qquad (2)$$

Here, we used the relationship $E_x = \beta H_y/(\omega\epsilon_0\epsilon'_{xx})$, which is valid for TM waves, and introduced the designations $k_0 = 2\pi/\lambda$ for the wave number and $\rho = \sqrt{\mu_0/\epsilon_0} = 120\pi$ for the free-space wave impedance. As can be seen from expression (2), taking into account the anisotropy of all the layers except for the lasing one corrects the attenuation coefficient insignificantly. We will therefore assume that the real parts of the permittivity tensor components equal each other; i.e., $\epsilon'_{xx} = \epsilon'_{yy} = \epsilon'_{zz} = \epsilon'_i$, where i is the number of the layer. Obviously, the imaginary parts of the permittivity tensor components that enter into formula (2) can be expressed in terms of the volume gain α_0 and loss factor g as

$$\epsilon''_{xx} = -\sqrt{\epsilon'_a}\frac{g}{k_0}; \quad \epsilon''_{zz} = \sqrt{\epsilon'_a}\frac{\alpha_{0a}}{k_0}$$

for the active layer and

$$\epsilon''_{xx} = \epsilon''_{zz} = \sqrt{\epsilon'_i}\frac{\alpha_{0i}}{k_0}$$

for the other layers. In the lasing theory, of primary importance is the field confinement coefficient Γ. For TM waves, it is defined as

$$\Gamma = \frac{\int_{\text{active}} \frac{|H_y|^2}{\epsilon'_a} dx}{\int_{-\infty}^{+\infty} \frac{|H_y|^2}{\epsilon'} dx}. \qquad (3)$$

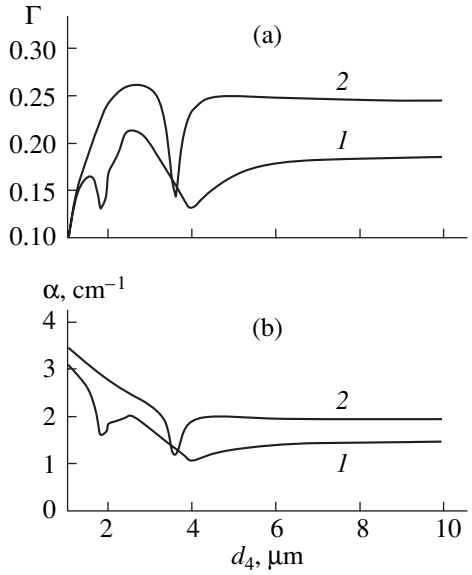


Fig. 2. (a) Optical confinement and (b) power loss factor for the TM wave in the intervalley-transfer laser vs. the upper confining layer thickness d_4. The lower confining layer thickness d_2 = 5 (*1*) and 10 μm (*2*).

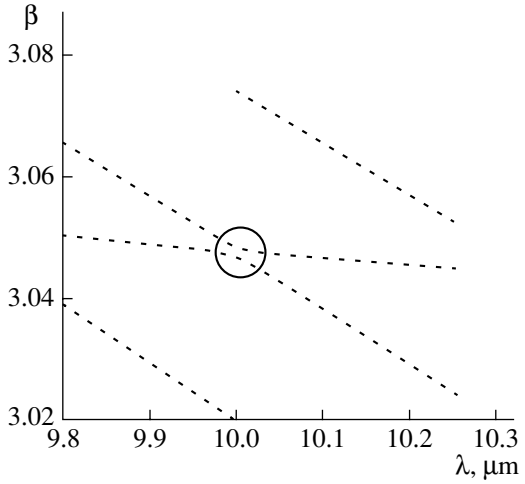


Fig. 3. Anticrossing of the dispersion characteristics as a result of interaction with substrate modes at $d_2 = 10$ μm and $d_4 = 3.6$ μm.

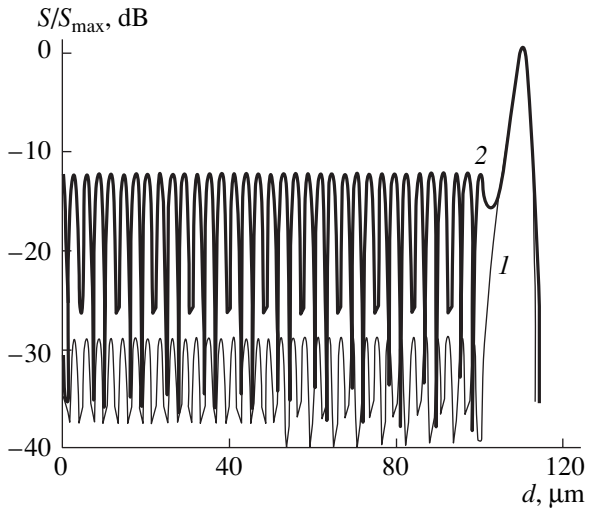


Fig. 4. TM-wave power density distribution over the cross section at $d_2 = 2.5$ (*1*) and 3.6 μm (*2*).

NUMERICAL RESULTS

To find the laser geometry optimal in terms of guiding characteristics, it is necessary to study the optical confinement and the loss factor as a function of the confining layer thickness (Fig. 2). The calculations were performed at a wavelength $\lambda = 10$ μm, $d_1 = 100$ μm, $d_3 = 1$ μm, $\epsilon_1' = 12.2$, $\epsilon_3' = 11.5$, and $\epsilon_2' = \epsilon_4' = 9$. The permittivity ϵ_3' of the lasing layer was calculated as an effective parameter dependent on the relative thicknesses of the GaAs (x) and AlAs ($1 - x$) layers of the superlattice:

$$\epsilon_3' = \epsilon_{GaAs}' x + \epsilon_{AlAs}'(1 - x).$$

The volume absorption α_{0a} of the lasing layer was estimated at about 500 cm^{-1} [4]. In the calculations, we neglected the loss due to scattering by free carriers in the undoped layers and the substrate.

Figure 2 shows that the optical confinement is maximum at a thickness of the upper confining layer $d_4 \approx 2.5$ μm for both thicknesses d_2 of the lower AlAs confining layer. The power loss factor α does not exceed 2.0–2.5 cm^{-1}. It is small primarily because the electric field component E_z in the propagation direction is small compared with the amplitude of the transverse component E_x. Thus, the volume gain factor $g \approx 10.5$ cm^{-1} is sufficient to compensate for this loss (the radiation loss through the end faces of the laser is neglected).

Notice the sharp dips in the curves in Fig. 2. It is likely that, under certain conditions, the lasing power flux resonantly tunnels into the substrate, causing interference effects therein. For example, for $d_4 \approx 3.6$ μm and $d_2 = 10$ μm, the dispersion characteristics of the lasing layer and the substrate intersect (see the outlined region in Fig. 3). As a result of the strong interaction between wave processes in the substrate and in the guiding layer, a significant portion of the lasing power penetrates into the substrate, sharply reducing the optical confinement coefficient. Wave interference in transparent substrates was studied earlier for a conventional semiconductor laser operating at 0.98 μm [6].

To illustrate the physical process described above, Fig. 4 shows the normalized distribution of the power flux over the laser cross section for the optimal (in terms of the confinement) structure ($d_4 = 2.5$ μm) and for the case when the radiation into the substrate is maximum ($d_4 = 3.6$ μm). It can be seen that the resonant interaction between the waves in the substrate and the lasing layer increases significantly (by 15–17 dB) the power peak density in the substrate.

CONCLUSION

Thus, our study of the electromagnetic properties of the long-wave intervalley-transfer laser are summarized as follows.

The optimal geometry of the structure that provides the maximum optical confinement for the TM wave ($\Gamma \approx 0.27$) is found.

Conditions for the intense penetration of the lasing power into the substrate are determined. The resonant interaction between the wave processes in the substrate and the optical waveguide may reduce significantly (to 0.13–0.15) the optical confinement coefficient.

Anisotropy of the loss and gain, which are associated with different field components of the TM wave in the laser structure, is analyzed. It is found that, with the optimal geometry of the optical waveguide at a volume

absorption of 500 cm^{-1} in the active layer, the gain $g \approx 10.5$ cm^{-1} is sufficient to compensate for the loss.

ACKNOWLEDGMENTS

This work was supported by the Russian Foundation for Basic Research (grant no. 00-02-81022).

REFERENCES

1. J. Faist, F. Capasso, D. Sivko, *et al.*, Science **264**, 553 (1994).
2. O. Ganthier-Lafayer, P. Boucaurd, F. Julien, *et al.*, Appl. Phys. Lett. **72**, 2388 (1998).
3. V. Ya. Aleshkin and A. A. Andronov, Pis'ma Zh. Éksp. Teor. Fiz. **68** (1), 73 (1998) [JETP Lett. **68**, 78 (1998)].
4. *Guided-Wave Optoelectronics*, Ed. by T. Tamira (Springer-Verlag, Berlin, 1990; Mir, Moscow, 1991).
5. T. D. Visser, H. Block, B. Demeulenaere, *et al.*, IEEE J. Quantum Electron. **33**, 1763 (1997).
6. I. A. Avrutsky, R. Gordon, R. Clayton, *et al.*, IEEE J. Quantum Electron. **33**, 1801 (1997).

Translated by A. Khzmalyan

Technical Physics, Vol. 47, No. 6, 2002, pp. 792–793. Translated from Zhurnal Tekhnicheskoĭ Fiziki, Vol. 72, No. 5, 2002, pp. 141–142.

BRIEF COMMUNICATIONS

Express Method of Designing Magnetic Mass Analyzers

O. N. Peregudov, V. F. Shkurdoda, and L. F. Sukhodub

Institute of Applied Physics, National Academy of Sciences of Ukraine, Sumy, ul. Petropavlovskaya 58, 40030 Ukraine

e-mail: iapuas@gluk.apc.org

Received November 20, 2001

Abstract—An original method for designing magnetic mass analyzers that is represented in the form of a computer program is proposed. The thorough optimization of the method makes it possible to rapidly and accurately simulate the focusing properties of a mass analyzer along its focal line, as well as to select designs appropriate for a given problem. *© 2002 MAIK "Nauka/Interperiodica".*

INTRODUCTION

The development of a magnetic mass analyzer, requiring a large body of computation and much experience, has always been a challenge for designers. To facilitate the design of new mass analyzers, as well as to tackle various comparison and optimization problems, we developed a new express design method.

EXPRESS SEARCH METHOD

Our approach is based on well-known formulas [1], which are extended by introducing higher order aberration coefficients into the classical dispersion relation. In addition, we took into account the results of computer simulations and optimized the mathematical model obtained. This has made it possible to simulate the focusing properties of a mass analyzer along its focal line over a wide range of masses with a much lesser body of computation.

Since the method is based on conventional formulas, we recommend applying it as the basis for designing magnetic mass analyzers. Further improvement of its accuracy can be achieved only by considering higher order effects of a stray edge field, as is done in more awkward models, for example, in the matriciant method [2]. Note that computational results obtained by this method and by the matriciant method [2] differ by approximately 2.5%. Because of the utmost simplicity of the method and, as a consequence, the high rate of computation, the application of our method for express analysis is well justified.

On the basis of our method, we devised a program (available from http://op.pochtamt.ru/) that makes it possible to evaluate the main properties of a magnetic mass analyzer along its focal line (resolution, positions of focal points, focusing parameters in the vertical and horizontal planes, etc.). In addition, the program enables an express search for designs according to given parameters. The optimization of the parameters is not performed. This is due to the fact that there are a number of optimization criteria, which do not form a continuous set. Frequently, improving one of the criteria results in the significant deterioration of the others. This problem has been considered carefully in [3], where it was noted that the result of any computer optimization cannot represent a single point in the space of initial parameters. One can only restrict the region of acceptable solutions, and the design of choice must meet certain nonformal criteria. As in [3], our program separates out a region of acceptable solutions according to rather loose formal criteria. The results of this separation form a database, which can readily be imported into "clearer" software tools (for example, into Microsoft Excel) for subsequent processing.

As was noted above, the basic feature of our method is its nonextensiveness, namely, it needs minimum computational resources for a rather high accuracy. The design search algorithm is based on the exhaustive search method (instead of the statistical sampling method [3]). As a result, the program does not require a large memory capacity and the search process and the formation of the database (i.e., the detailed simulation of the focusing properties of the mass analyzer along its focal line) in combination take several minutes depending on the search discreteness required. In addition, there is a possibility of simulating the process of electrostatic scanning over fixed collectors if a permanent magnet is applied [4] for rotations other than 180°.

RESULTS AND DISCUSSION

This method was successfully tested in the express search for designs of magnetic mass analyzers intended for the analysis of argon isotopes. The parameters simulated were resolving power R, mass dispersion D, tilt angle of the focal line Ψ, and sum of squared deviations of focal points from a straight line δ. Note that these parameters, except for the last one, are, in general, local; i.e., they are of interest only for a particular mass (or for a pair of masses, for example, in the case of the

1063-7842/02/4706-0792$22.00

Table

$\varepsilon' = \varepsilon''$	R	l'	D	ψ	δ
31	1014	0.8840	3.0608	28.7	1.243×10^{-4}
32	1016	0.9032	3.1772	28.2	1.415×10^{-4}
33	1017	0.9327	3.3024	27.7	1.640×10^{-4}
34	1019	0.9456	3.4376	27.1	1.860×10^{-4}
35	1021	0.9691	3.5840	26.6	2.137×10^{-4}
36	1023	0.9945	3.7428	26.1	2.464×10^{-3}
37	1025	1.0220	3.9160	25.6	2.864×10^{-3}
38	1028	1.0520	4.1040	25.1	3.362×10^{-4}
47	1063	1.5160	7.1720	20.3	1.867×10^{-3}
48	1069	1.6090	7.8040	19.7	2.401×10^{-3}

tilt angle) from a given range. The input parameters were angle of rotation in a magnetic field φ, radius of rotation in a magnetic field r, optical magnification M, and ion source parameters. The input arm (l') is an optional parameter (it is calculated unless specified). The output arm (l'') is found from direction focusing conditions. The entrance and exit boundaries of the magnet are straight. We were interested in designs with an optical magnification M in the range [0, 5, 1] and a resolution R of no less than 600 (the width of the source slit is $S_1 = 0.1$ mm, and the divergence angle of the beam is $2\alpha = 4 \times 10^{-2}$ rad).

The table lists comparative results for symmetric mass analyzers with the radius and angle of rotation 50 mm and 120°, respectively (the entrance and exit angles are integers).

Using the data from the table, one can construct a rather compact analyzer with good analytical characteristics. In addition, having a set of appropriate designs, one can choose that offering the best manufacturability and/or being easier to adjust. In our case, after simple interpolation (the correctness of which is confirmed by additional calculations), one can choose a version with the entrance/exit angles ($\varepsilon' = \varepsilon''$) 37°30′, $R = 1026$, $l' = 1.0370$, $D = 4.0080$, $\Psi = 25.2$, and $\delta = 3.088 \times 10^{-4}$.

CONCLUSIONS

Thus, the express search method proposed is adequate to available analogues. The high rate and efficiency of computation open up a wide spectrum of applications, one of which is presented here. The issue touched upon is rather topical nowadays, since the miniaturization of magnetic mass analyzers has become an important problem.

ACKNOWLEDGMENTS

Peregudov thanks S.N. Mordik, O.R. Savin, O.A. Sidora, and A.G. Ryabyshev for numerous useful discussions.

REFERENCES

1. A. S. Kuzema, O. R. Savin, and I. Ya. Chertkov, *Analyzers for Magnetic Mass Spectrometers* (Naukova Dumka, Kiev, 1987).
2. S. N. Mordik and A. G. Ponomarev, Zh. Tekh. Fiz. **71** (7), 98 (2001) [Tech. Phys. **46**, 883 (2001)].
3. S. B. Turtia, V. D. Satchenko, A. S. Berdnikov, *et al.*, Nucl. Instrum. Methods Phys. Res. A **427**, 382 (1999).
4. Yu. V. Brazhnik and V. F. Shkurdoda, Ukr. Fiz. Zh. **45**, 1015 (2000).

Translated by Yu. Vishnyakov