Spectral Characteristics of Interaction between a Short Intense Laser Pulse and Ionizable Gas

M. V. Chegotov

Institute of Thermal Physics of Extremal States, Institute of High Temperatures Scientific Association,
Russian Academy of Sciences, ul. Izhorskaya 13/19, Moscow, 127412 Russia

e-mail: chegotov@hedric.msk.su

Abstract—An expression for the mean-square frequency of a short intense laser pulse passing through an ionizable gas is derived for arbitrary 3D irradiation and observation geometries and pulse intensity. It is found that the resulting blue shift depends on the lasing intensity at the instant of ionization and on the ionization energy loss. Taking into account stimulated Raman backscattering increases the predicted value of the blue shift. © 2002 MAIK "Nauka/Interperiodica".

INTRODUCTION

Interaction of short intense laser pulses with various materials is generating considerable interest (see, e.g., [1]). Much attention to this subject is due mainly to the wide range of applications of short intense pulses. Neutral ionizable gases are promising for lasing sources [2], laser plasma acceleration of particles [3], up-conversion of electromagnetic waves by plasma methods [4–6], etc. Moreover, the blue shift resulting from the interaction of a laser pulse with an ionized medium can be employed for the diagnostics of this interaction [7].

The frequency characteristics of the interaction between an intense laser pulse and a medium being ionized resulting in plasma formation are usually described in the slowly varying amplitude approximation (because of the nonlinear interaction of laser radiation with the ionization front, this approximation was termed quasi-harmonic [5]). This approximation implies that the frequency depends on both the position on the pulse temporal profile and the spatial coordinates (see, e.g., [4–8]). This makes the conventional expression for the blue shift (see, e.g., [9]) $\delta\lambda \sim (zr_e\lambda_0^3/2\pi c)dn_e/dt$ (z is the interaction length, r_e is the classical electron radius, c is the velocity of light, $\lambda_0 = 2\pi c/\omega_0$, ω_0 is the laser radiation frequency, and n_e is the electron density) intrinsically contradictory, since $n_e = 0$ at the exit from the plasma and the outgoing radiation does not exhibit the blue shift. Moreover, the value of dn_e/dt is strongly dependent on the position on the temporal profile of the laser pulse and remains uncertain. This uncertainty is still more aggravated for polyelectron atoms, when different atomic shells are responsible for the ionization front. Therefore, the effect of ionization on the spectra of the radiation from the forming plasma remains quantitatively unclear.

In this study, we present a general approach that enables a strict quantitative description of the ionization effect on the radiation spectrum observed in experiments. With this method applied to moderate laser pulse intensities, we determined the blue shift of the mean-square frequency of the laser pulse passed through the ionizable gas. It is demonstrated that the shift is defined by the laser intensity at the instant of ionization, as well as by the amount of energy ionization losses and residual energy of electrons [10] (see formula (15)). With an increase in the laser pulse intensity, stimulated Raman scattering (SRS) may arise. Stimulated Raman backscattering (SRBS) is shown to enhance the blue shift of the ionizing radiation frequency.

BASIC RELATIONSHIPS

Let us start from the Maxwell equations

$$\nabla \times \mathbf{B} = \frac{4\pi}{c}\mathbf{J} + \frac{1}{c}\frac{\partial \mathbf{E}}{\partial t}, \quad \nabla \times \mathbf{E} = -\frac{1}{c}\frac{\partial \mathbf{B}}{\partial t}, \qquad (1)$$

where the electric current $\mathbf{J}$ is defined self-consistently through the electric $\mathbf{E}$ and the magnetic field $\mathbf{B}$. Differentiating (1) with respect to t; scalarly multiplying the resulting equations by $\partial\mathbf{E}/\partial t$ and $\partial\mathbf{B}/\partial t$, respectively; and subtracting the results, we get

$$\frac{\partial}{\partial t}\left\{\frac{1}{8\pi}\left[\left(\frac{\partial \mathbf{E}}{\partial t}\right)^2 + \left(\frac{\partial \mathbf{B}}{\partial t}\right)^2\right]\right\}$$
$$+ \text{div}\left\{\frac{c}{4\pi}\left[\frac{\partial \mathbf{E}}{\partial t} \times \frac{\partial \mathbf{B}}{\partial t}\right]\right\} = -\frac{\partial \mathbf{E}}{\partial t} \cdot \frac{\partial \mathbf{J}}{\partial t}. \qquad (2)$$

Assuming that there are zero fields at $t = \pm\infty$, Eq. (2) yields

$$\text{div}\left\{\frac{c}{4\pi}\int_{-\infty}^{+\infty}\left[\frac{\partial \mathbf{E}}{\partial t}\times\frac{\partial \mathbf{B}}{\partial t}\right]dt\right\} = -\int_{-\infty}^{+\infty}\frac{\partial \mathbf{E}}{\partial t}\cdot\frac{\partial \mathbf{J}}{\partial t}dt. \quad (3)$$

Integrating (3) over a certain volume V bounded by the surface S, we obtain

$$\frac{c}{4\pi}\int_{S}\int_{-\infty}^{+\infty}\left[\frac{\partial \mathbf{E}}{\partial t}\times\frac{\partial \mathbf{B}}{\partial t}\right]dtd\mathbf{s} = -\int_{V}\int_{-\infty}^{+\infty}\frac{\partial \mathbf{E}}{\partial t}\cdot\frac{\partial \mathbf{J}}{\partial t}dtd^3\mathbf{r}, \quad (4)$$

where $d\mathbf{s} = \mathbf{e}_S ds$ is the vector element of the surface S with the external normal $\mathbf{e}_S$.

For the further use of Eq. (4), one should specify the position of the surface S_{in} (a part of the surface S) through which the laser beam enters the gas and of the surface S_{out} ($S = S_{in} + S_{out}$) through which the radiation comes out of the volume V. The radiation may also leave the volume V through the surface S_{in}, for example, in the case of SRBS by plasma waves (see, e.g., [11]). We assume that the surfaces S_{in} and S_{out} are in the wave (Fresnel) zone with respect to the volume V_{pl} where the plasma is generated and that the permeability on these surfaces equals unity; hence, $\mathbf{B} = \mathbf{H}$, where $\mathbf{H}$ is the magnetic field strength.

In the wave zone, the fields $\mathbf{E}$ and $\mathbf{H}$ are related as $\mathbf{E} = [\mathbf{H} \times \mathbf{n}]$, where $\mathbf{n} = [\mathbf{E} \times \mathbf{H}]/|[\mathbf{E} \times \mathbf{H}]|$ is the unit Poynting vector at a certain point on the surface S [12] and $[\ldots]$ denotes the vector product. Substituting this relationship in (4) yields

$$\frac{c}{4\pi}\int_{S}\int_{-\infty}^{+\infty}\left(\frac{\partial \mathbf{H}}{\partial t}\right)^2 dt\mathbf{n}d\mathbf{s} = -\int_{V_{pl}}\int_{-\infty}^{+\infty}\frac{\partial \mathbf{E}}{\partial t}\cdot\frac{\partial \mathbf{J}}{\partial t}dtd^3\mathbf{r};$$

hence, the spectral energy density per unit area is given by

$$I(\omega, \mathbf{r}) = (c/4\pi^2)|\mathbf{H}(\omega, \mathbf{r})|^2,$$

where

$$\mathbf{H}(\omega, \mathbf{r}) = \int_{-\infty}^{+\infty}\mathbf{H}(t, \mathbf{r})\exp(i\omega t)dt \quad (\text{see } [12]),$$

$$\quad (5)$$

$$\int_{S}\int_{0}^{\infty}\omega^2 I(\omega, \mathbf{r})d\omega\mathbf{n}d\mathbf{s} = -\int_{V_{pl}}\int_{-\infty}^{+\infty}\frac{\partial \mathbf{E}}{\partial t}\cdot\frac{\partial \mathbf{J}}{\partial t}dtd^3\mathbf{r}.$$

Formulas (5) imply that the current $\mathbf{J}$ is concentrated within the plasma zone V_{pl}. In addition to (5), the energy conservation law in the form

$$\int_{S}\int_{0}^{\infty}I(\omega, \mathbf{r})d\omega\mathbf{n}d\mathbf{s} = -\int_{V_{pl}}\int_{-\infty}^{+\infty}\mathbf{E}\cdot\mathbf{J}dtd^3\mathbf{r} \quad (6)$$

is valid. Taking into consideration that the radiation may propagate through the surfaces S_{in} and S_{out}, we can rewrite Eq. (6) in the form

$$\mathscr{E}_{out} + \mathscr{E}_{back} - \mathscr{E}_{in} = -\int_{V_{pl}}\int_{-\infty}^{+\infty}\mathbf{E}\cdot\mathbf{J}dtd^3\mathbf{r}, \quad (7)$$

where $\mathscr{E}_{in} = -\int_{S_{in}}\int_{0}^{\infty}I_{in}(\omega, \mathbf{r})d\omega\mathbf{n}d\mathbf{s}$ is the radiation energy entering the volume V and $\mathscr{E}_{out} = \int_{S_{out}}\int_{0}^{\infty}I_{out}(\omega, \mathbf{r})d\omega\mathbf{n}d\mathbf{s}$ and $\mathscr{E}_{back} = \int_{S_{back}}\int_{0}^{\infty}I_{back}(\omega, \mathbf{r})d\omega\mathbf{n}d\mathbf{s}$ are the radiation energies leaving the volume V through the surfaces S_{out} and S_{in} (backward to the incoming radiation), respectively.

Introducing the mean-square frequencies

$$\langle\omega^2\rangle_{in} = \int_{S_{in}}\int_{0}^{\infty}\omega^2 I_{in}(\omega, \mathbf{r})d\omega\mathbf{n}d\mathbf{s}\left(\int_{S_{in}}\int_{0}^{\infty}I_{in}(\omega, \mathbf{r})d\omega\mathbf{n}d\mathbf{s}\right)^{-1},$$

$$\langle\omega^2\rangle_{out} = \int_{S_{out}}\int_{0}^{\infty}\omega^2 I_{out}(\omega, \mathbf{r})d\omega\mathbf{n}d\mathbf{s}$$
$$\times\left(\int_{S_{out}}\int_{0}^{\infty}I_{out}(\omega, \mathbf{r})d\omega\mathbf{n}d\mathbf{s}\right)^{-1},$$

$$\langle\omega^2\rangle_{back} = \int_{S_{back}}\int_{0}^{\infty}\omega^2 I_{back}(\omega, \mathbf{r})d\omega\mathbf{n}d\mathbf{s}$$
$$\times\left(\int_{S_{back}}\int_{0}^{\infty}I_{back}(\omega, \mathbf{r})d\omega\mathbf{n}d\mathbf{s}\right)^{-1},$$

and the backward reflection factor $R_{back} = \mathscr{E}_{back}/\mathscr{E}_{in}$, we reduce Eq. (5) to the formula that defines the mean-square frequency of the radiation outgoing through the S_{out} surface:

$$\langle\omega^2\rangle_{out}$$
$$= \frac{\langle\omega^2\rangle_{in} - \langle\omega^2\rangle_{back}R_{back} - \dfrac{1}{\mathscr{E}_{in}}\displaystyle\int_{V_{pl}}\int_{-\infty}^{+\infty}\dfrac{\partial \mathbf{E}}{\partial t}\cdot\dfrac{\partial \mathbf{J}}{\partial t}dtd^3\mathbf{r}}{1 - R_{back} - \dfrac{1}{\mathscr{E}_{in}}\displaystyle\int_{V_{pl}}\int_{-\infty}^{+\infty}\mathbf{E}\cdot\mathbf{J}dtd^3\mathbf{r}}, \quad (8)$$

where relationship (7) is included. It should be noted that formula (8) is valid for any intensity of the laser pulse and an arbitrary irradiation geometry. For the sake of definiteness, the following analysis will concern the spectral characteristics of the radiation that passed through the surface S_{out}.

IONIZATION-INDUCED FREQUENCY SHIFT

In order to separate out the effect of the ionization processes on the radiation spectrum, we consider an ionizing pulse of a relatively low intensity in the field of which electrons move with velocities much lower than the velocity of light c. With the ionization processes included, the electron current $\mathbf{J}$ has the form

$$\mathbf{J} = \mathbf{j} + \mathbf{j}_{ion}, \tag{9}$$

where $\mathbf{j}$ is the free electron current and

$$\mathbf{j}_{ion} = \frac{\mathbf{E}}{|\mathbf{E}|^2} \sum_{m=0}^{Z-1} W_m N_m U_m \tag{10}$$

is the ionization current, which takes into account the ionization loss of the laser field energy and momentum [13]. Here, N_m is the concentration of ions with a degree of ionization m (for neutral atoms, $m = 0$), U_m is the ionization potential required to take an ion of charge m to the $(m+1)$-charged state, W_m is the probability of m-to-$(m+1)$ ionization per unit time, and Z is the nucleus charge. The density of the free electron current $\mathbf{j}$ can be found from the equation

$$\frac{\partial \mathbf{j}}{\partial t} = \frac{e^2 n_e}{m_e} \mathbf{E}, \tag{11}$$

where n_e is the concentration of free electrons appearing as a result of ionization, and $m_e(e)$ is the electron mass (charge). The rate of change of n_e is defined by the ionization probabilities per unit time W_m:

$$\frac{\partial n_e}{\partial t} = \Gamma = \sum_{m=0}^{Z-1} \Gamma_m = \sum_{m=0}^{Z-1} W_m N_m, \tag{12}$$

$$\frac{\partial N_0}{\partial t} = -\Gamma_0, \quad \frac{\partial N_m}{\partial t} = -\Gamma_m + \Gamma_{m-1}, \tag{13}$$

$$m = 1, \dots, Z-1.$$

The ionization of atoms and ions by a sufficiently short laser pulse follows the tunneling mechanism (see, e.g., [10]). Therefore, the probability of ionization per unit time can be expressed by the Ammosov–Delone–Krainov formula [14]

$$W_m(|\mathbf{E}|) = \omega_{at} \frac{\exp(1)(m+1)^2}{2\pi \, n_*^4}$$

$$\times \left[4\exp(1)\frac{(m+1)^3}{n_*^4}\frac{E_{at}}{|\mathbf{E}|} \right]^{2n_* - 1} \exp\left(-\frac{2}{3}\frac{(m+1)^3}{n_*^3}\frac{E_{at}}{|\mathbf{E}|} \right), \tag{14}$$

where $n_* = (m+1)\sqrt{U_H/U_m}$, U_H is the potential of hydrogen ionization from the ground state, $E_{at} \approx 5.1 \times 10^9$ V/cm is the atomic field strength, and $\omega_{at} \approx 4.1 \times 10^{16}$ s^{-1} is the atomic frequency.

In the following analysis of expression (8), we will assume that the back reflection from the moving ionization front is insignificant for low concentrations of electrons $n_e \ll n_c = m_e\langle\omega^2\rangle_{in}/(4\pi e^2)$ [15]. Then, according to (9)–(12), formula (8) becomes

$$\langle\omega^2\rangle_{out} = \langle\omega^2\rangle_{in}$$

$$\times \left(1 + \frac{\mathscr{E}_{ion}^{fr}}{\mathscr{E}_{in} - \mathscr{E}_{ion}^p} \right) \left(1 - \frac{\mathscr{E}_{ion}^{res}}{\mathscr{E}_{in} - \mathscr{E}_{ion}^p} \right)^{-1}, \tag{15}$$

where

$$\mathscr{E}_{ion}^p = \int_{V_{pl}}\int_{-\infty}^{+\infty} \mathbf{E} \cdot \mathbf{j}_{ion}\, dt\, d^3\mathbf{r} = \sum_{m=0}^{Z-1} U_m \int_{V_{pl}}\int_{-\infty}^{+\infty} \Gamma_{in}\, dt\, d^3\mathbf{r} \tag{16}$$

is the laser pulse energy spent by electrons in the volume V_{pl} to overcome the ionization potential barriers U_m,

$$\mathscr{E}_{ion}^{res} = \int_{V_{pl}}\int_{-\infty}^{+\infty} \mathbf{E} \cdot \mathbf{j}\, dt\, d^3\mathbf{r} = \int_{V_{pl}}\int_{-\infty}^{+\infty} \Gamma\frac{m_e\mathbf{V}_E^2}{2}\, dt\, d^3\mathbf{r} \tag{17}$$

is the residual energy of electrons [10, 15, 16], $\mathbf{V}_E = \frac{e}{m_e}\int_{-\infty}^t \mathbf{E}\, dt'$, and

$$\mathscr{E}_{ion}^{fr} = \int_{V_{pl}}\int_{-\infty}^{+\infty} \Gamma\frac{e^2\mathbf{E}^2}{2m_e\langle\omega^2\rangle_{in}}\, dt\, d^3\mathbf{r}. \tag{18}$$

We derived expression (15) under the assumption that the duration τ_{ion} of the ionization front on the temporal profile of the laser pulse is much greater than the laser field period $2\pi/\sqrt{\langle\omega^2\rangle_{in}}$; therefore,

$$\int_{-\infty}^{+\infty} \frac{\partial \mathbf{E}}{\partial t} \cdot \frac{\partial \mathbf{j}_{ion}}{\partial t}\, dt = -\int_{-\infty}^{+\infty} \frac{\partial^2 \mathbf{E}}{\partial t^2} \cdot \mathbf{j}_{ion}\, dt$$

$$\approx \langle\omega^2\rangle_{in}\int_{V_{pl}}\int \mathbf{E} \cdot \mathbf{j}_{ion}\, dt\, d^3\mathbf{r}.$$

Expression (15) suggests an increase in the mean-square frequency of the laser radiation ionizing the gas medium through which it passes. The ratio between $\mathscr{E}_{ion}^{fr}$ (Eq. (18)) and $\mathscr{E}_{ion}^{res}$ (Eq. (17)), which specifies the blue shift, considerably depends on the polarization of the ionizing radiation because of the polarization dependence of the residual energy $\mathscr{E}_{ion}^{res}$ [16]. Ionization under the action of a linearly polarized beam occurs at time moments close to the maxima of the magnitude of the oscillating electric field, which correspond to a nearly zero velocity $\mathbf{V}_E$. Consequently, for the linear polarization of the laser radiation,

$$\mathscr{E}_{ion}^{fr} \gg \mathscr{E}_{ion}^{res}. \tag{19}$$

In case of the circular polarization and small changes in $\langle \omega^2 \rangle_{out}$ with respect to $\langle \omega^2 \rangle_{in}$, we can use the approximation $\mathbf{V}_E^2 \approx e^2 \mathbf{E}^2/(m_e^2 \langle \omega^2 \rangle_{in})$; then, from Eqs. (17) and (18), we get

$$\mathscr{E}_{ion}^{fr} \approx \mathscr{E}_{ion}^{res}. \tag{20}$$

When the polarization is linear, the magnitude of the field $\mathbf{E}$ in (17) and (18) is fast-oscillating. If the variation of the laser pulse envelope $\mathbf{E}_0(\mathbf{r}, t)$ is slow on the scale of the fast-oscillation period $2\pi/\sqrt{\langle \omega^2 \rangle_{in}}$ (which implies, in particular, that $\tau_{ion} \gg 2\pi/\sqrt{\langle \omega^2 \rangle_{in}}$ and the parameter $\alpha_m \equiv (U_H/U_m)^{3/2}|\mathbf{E}_0(\mathbf{r}, t)|/E_{at} \ll 1$ at the instants of ionization, which is true for light atoms with $Z < 10$), then Eqs. (17) and (18) can be simplified. Let us represent the field $\mathbf{E}$ in the form $\mathbf{E} = \mathbf{E}_0(\mathbf{r}, t)\cos(\omega_0 t + \varphi)$, where $\omega_0 = \sqrt{\langle \omega^2 \rangle_{in}}$. Residual energy (17) appears then as

$$\mathscr{E}_{ion}^{res} \approx \int\limits_{V_{pl}} \int\limits_{-\infty}^{+\infty} \sum_{m=0}^{Z-1} \overline{W}_m \overline{N}_m R_m dt d^3\mathbf{r}, \tag{21}$$
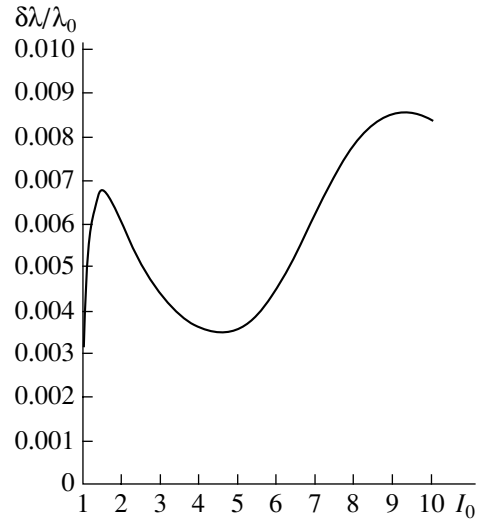
where (cf. [10, 16])

$$R_m(\mathbf{r}, t) \approx \frac{3\hbar\omega_0}{\sqrt{2}}\left(\frac{Q_p(\mathbf{r}, t)}{U_m}\right)^{3/2}, \tag{22}$$

and $Q_p(\mathbf{r}, t) = e^2|\mathbf{E}_0(\mathbf{r}, t)|^2/(4m_e\omega_0^2)$ is the ponderomotive potential of an electron. In this case,

$$\mathscr{E}_{ion}^{fr} = \int\limits_{V_{pl}} \int\limits_{-\infty}^{+\infty} \sum_{m=0}^{Z-1} \overline{W}_m \overline{N}_m F dt d^3\mathbf{r}, \tag{23}$$

where

$$F(\mathbf{r}, t) = 2Q_p(\mathbf{r}, t). \tag{24}$$

Dependence of $\delta\lambda/\lambda_0$ on the peak intensity I_0 of the Gaussian pulse with a half-width of 90 fs and a wavelength $\lambda_0 = 0.62$ μm. Helium pressure is 1 atm. I_0 is measured in 10^{15} W/cm^2.

Here,

$$\overline{W}_m(|\mathbf{E}_0|) = \omega_{at}\sqrt{3}\left(\frac{\exp(1)}{\pi}\right)^{3/2}\frac{(m+1)^2}{n_*^{4.5}}$$
$$\times \left[4\exp(1)\frac{(m+1)^3}{n_*^4}\frac{E_{at}}{|\mathbf{E}_0|}\right]^{2n_*-1.5} \tag{25}$$
$$\times \exp\left(-\frac{2}{3}\frac{(m+1)^3}{n_*^3}\frac{E_{at}}{|\mathbf{E}_0|}\right)$$

and $\overline{N}_m$ are ionization frequency (14) and the ion density, respectively, averaged over the laser field period $2\pi/\omega_0$. Since

$$R_m/F \approx \frac{3}{2}\alpha_m \ll 1, \tag{26}$$

the validity of inequality (19) ($\mathscr{E}_{ion}^{res}$ is much smaller than $\mathscr{E}_{ion}^{fr}$) under the conditions of linear polarization depends on the smallness of α_m.

When the frequency shift is small, a decrease in the wavelength $\delta\lambda$ can be readily obtained from Eq. (15):

$$\frac{\delta\lambda}{\lambda_0} = \frac{1}{2}\frac{\mathscr{E}_{ion}^{fr} + \mathscr{E}_{ion}^{res}}{\mathscr{E}_{ion} - \mathscr{E}_{ion}^p},$$

where λ_0 is the initial wavelength.

The figure presents the dependence of $\delta\lambda/\lambda_0$ on the peak intensity of a linearly polarized laser pulse in the approximation of one-dimensional propagation for the

parameters used in [7]. The interaction length is taken to be 50 μm, which is equal to the experimental focus length. The ratio $\delta\lambda/\lambda_0$ is close to that observed in [7].

IONIZATION-INDUCED FREQUENCY SHIFT IN THE PRESENCE OF SRBS

An increase in the intensity of short laser pulses gives rise to the SRS parametric process. In a uniform gas medium, the gain increment is maximal for three-wave SRBS, when the secondary wave propagates counter the initial laser wave and is red-shifted by the plasma wave frequency (Stokes component). As a laser pulse propagates in an ionizable gas, the Stokes component is excited at the leading edge of the pulse behind the ionization front (i.e., in the already existing plasma). Propagating backward to the laser pulse, the Stokes component is amplified by the laser field in the plasma independently of the ionization front. Such an amplification is effective if the plasma behind the ionization front, wherein the laser field is still large, is uniform over a considerable length and has an electron density $n_{e, \max}$.

If the SRBS amplification factor is sufficiently large, we can assume that the backreflected energy essentially exceeds the losses due to ionization and plasma wave excitation: i.e.,

$$R_{\text{back}} \gg \frac{1}{\mathscr{E}_{\text{in}}} \int\limits_{V_{\text{pl}}} \int\limits_{-\infty}^{+\infty} \mathbf{E} \cdot \mathbf{J} dt d^3\mathbf{r}.$$

With this assumption, Eq. (8) takes the form ($\omega_0 = \sqrt{\langle\omega^2\rangle_{\text{in}}}$)

$$\langle\omega^2\rangle_{\text{out}}$$
$$= \frac{\omega_0^2 - (\omega_0 - \omega_{\text{p, max}})^2 R_{\text{back}} - X_{\text{SRS}} + \omega_0^2 \mathscr{E}_{\text{ion}}^{\text{fr}}/\mathscr{E}_{\text{in}}}{1 - R_{\text{back}}}, \quad (27)$$

where $\omega_{\text{p, max}} = \sqrt{4\pi e^2 n_{e, \max}/m_e}$ is the plasma frequency, X_{SRS} is the SRS contribution to the expression

$$\frac{1}{\mathscr{E}_{\text{in}}} \int\limits_{V_{\text{pl}}} \int\limits_{-\infty}^{+\infty} \frac{\partial\mathbf{E}}{\partial t} \cdot \frac{\partial\mathbf{J}}{\partial t} dt d^3\mathbf{r},$$

and $\mathscr{E}_{\text{ion}}^{\text{fr}}$ is defined by formula (18). Here, we took into account that the contributions of ionization and SRS to the above expression are additive in view of their space–time separation. In the absence of ionization, the laser radiation passes through the plasma without a frequency shift: $\langle\omega^2\rangle_{\text{out}} = \omega_0^2$. Then, formula (27) yields the analytic dependence of X_{SRS} on R_{back} and $\omega_{\text{p, max}}$:

$$X_{\text{SRS}} = \omega_{\text{p, max}}(2\omega_0 - \omega_{\text{p, max}})R_{\text{back}}.$$

Substituting this expression into (27), we finally arrive at

$$\langle\omega^2\rangle_{\text{out}} = \omega_0^2\left(1 + \frac{\mathscr{E}_{\text{ion}}^{\text{fr}}/\mathscr{E}_{\text{in}}}{1 - R_{\text{back}}}\right). \quad (28)$$

Thus, the blue shift of the mean-square frequency increases with growing SRS reflection.

CONCLUSION

Our formula (15) for the frequency shift under conditions of the laser-induced ionization of a gas medium essentially differs from the widely used estimator discussed in the Introduction. Unlike the latter, expression (15) represents the frequency of an electromagnetic wave outside the generated plasma (in the wave zone). Moreover, according to (15), the frequency shift is defined by the ionizing radiation intensity I_{ion} at the ionization front. In particular, if the peak intensity I_0 of the laser pulse $\gg I_{\text{ion}}$, the blue shift turns out to be negligible. The blue shift is maximal when the time moment of ionization at the temporal profile of the pulse coincides with its maximal intensity. In gases with polyelectron atoms, the blue shift comprises the additive contributions from the ionization of several electron shells. The blue shift increases with peak intensity I_0 when the latter becomes close to the threshold intensity for the subsequent ionization of the ion. These regularities are in agreement with those observed in the experiment [7] at a relatively small gas pressure (1 atm).

ACKNOWLEDGMENTS

REFERENCES

1. *Laser Optics 2000: Ultrafast Optics and Superstrong Laser Fields*, Ed. by A. A. Andreev and V. E. Yashin, Proc. SPIE, Vol. 4352.

2. A. L'Huillier, L-A. Lompre, G. Mainfray, and C. Manus, *Atoms in Intense Laser Fields*, Ed. by M. Gavrila (Academic, Boston, 1992), pp. 139–201.

3. Z. Najmudin, A. E. Dangor, A. Modena, *et al.*, IEEE Trans. Plasma Sci. **28**, 1057 (2000).

4. S. C. Wilks, J. M. Dawson, W. B. Mori, *et al.*, Phys. Rev. Lett. **62**, 2600 (1989).

5. V. B. Gil'denburg, A. V. Kim, and A. M. Sergeev, Pis'ma Zh. Éksp. Teor. Fiz. **51**, 91 (1990) [JETP Lett. **51**, 104 (1990)].

6. E. Esarey, A. Ting, and P. Sprangle, Phys. Rev. A **42**, 3526 (1990).

7. W. M. Wood, C. W. Siders, and M. C. Downer, IEEE Trans. Plasma Sci. **21**, 20 (1993).

8. L. Oliveira e Silva and J. T. Mendonca, IEEE Trans. Plasma Sci. **24**, 316 (1996).

9. C. W. Siders, N. C. Turner, III, M. C. Downer, *et al.*, J. Opt. Soc. Am. B **13**, 330 (1996).

10. M. V. Chegotov, Fiz. Plazmy **26**, 940 (2000) [Plasma Phys. Rep. **26**, 881 (2000)].

11. Y. R. Shen, *The Principles of Nonlinear Optics* (Wiley, New York, 1984; Nauka, Moscow, 1989).

12. L. D. Landau and E. M. Lifshitz, *Course of Theoretical Physics*, Vol. 2: *The Classical Theory of Fields* (Nauka, Moscow, 1988; Pergamon, Oxford, 1975).

13. N. E. Andreev, M. E. Veĭsman, M. G. Kedzhyan, and M. V. Chegotov, Fiz. Plazmy **26**, 1010 (2000) [Plasma Phys. Rep. **26**, 947 (2000)].

14. M. V. Ammosov, N. B. Delone, and V. P. Krainov, Zh. Éksp. Teor. Fiz. **91**, 2008 (1986) [Sov. Phys. JETP **64**, 1191 (1986)].

15. M. V. Chegotov, Kvantovaya Élektron. (Moscow) **31**, 804 (2001).

16. N. A. Andreev, M. E. Veĭsman, S. P. Goreslavskiĭ, and M. V. Chegotov, Fiz. Plazmy **27**, 296 (2001) [Plasma Phys. Rep. **27**, 278 (2001)].

Translated by A. Sidorova

OPTICS,
QUANTUM ELECTRONICS

Optical Diode Based on a Highly Anisotropic Layer of a Helical Periodic Medium Subjected to a Magnetic Field

A. H. Gevorgyan

Yerevan State University, ul. Manukyana 1, Yerevan, 375025 Armenia
e-mail: agevorgyan@ysu.am
Received November 13, 2001; in final form, February 8, 2002

Abstract—The effect of an applied magnetic field on the optical properties of a layer of a helical periodic medium is studied in view of magnetooptic activity. The case when the radiation is normally incident on the layer and the magnetic field is aligned with the axis of the medium is considered. Irreversibility (nonreciprocity) effects in such a system are discussed. The situations with weak and high anisotropy are investigated. It is shown that the system can function as an optical shutter, optical diode, or one-side reflector. Reasons for the high irreversibility of transmission (reflection) are found. © 2002 MAIK "Nauka/Interperiodica".

INTRODUCTION

The effect of a magnetic field on the optical properties of helical periodic media, such as cholesterics, chiral smectics, and chiral ferronematics, has been the subject of much investigation. Emphasis has been on structure modifications in a helical periodic system (HPS) due to director reorientation [1–5]. However, it may happen that the magnetic field does not change the HPS structure. This may be observed in nonmagnetic HPSs with $\hat{\mu} = \hat{I}$ (where $\hat{I}$ is the unity matrix) and in HPSs with negative magnetic anisotropy if the magnetic field is aligned with the helix axis [1–5]. Irrespective of structure modifications, the magnetic field changes the local permittivity tensor of the structure and, in particular, causes the Faraday effect to occur. The coexistence of optical activity due to the helicity of the system and magnetooptic activity results in irreversibility effects [6–8]. The normal incidence of light on an HPS placed in a magnetic field aligned with the helix axis has been studied in [6, 7]. The Faraday effect in HPSs that arises when light propagates at an angle to the axis of the medium has been considered in terms of the dynamic diffraction theory in [8].

If the local anisotropy of the medium is weak, the magnetooptic activity is also relatively small. On the other hand, high (and even giant) anisotropy has been observed in HPSs, and some interesting features of such systems have been noted [9, 10]. In this work, we report an exact analytical solution to the boundary problem of normal transmission of light through a finite-thickness HPS layer subjected to a magnetic field aligned with the axis of the medium. The weak- and high-anisotropy cases are considered, and a number of intriguing applications are discussed.

If the anisotropy is significant, the permeability tensor of real systems is other than unity. Therefore, we will consider a more general case when the system has both dielectric and magnetic helicity. It should be noted that high anisotropy values are expected near individual absorption lines if one of the components of the permittivity (permeability) tensor is much larger than the other or if the components are of different sign. In the latter case, however, ordinary anisotropic media are known [11] to exhibit unusual features: specifically, the surface of wave vectors for such media is open (hyperboloid of revolution) rather than closed (closed ellipsoid). High values of gyroelectric and gyromagnetic activity parameters are expected near the resonances of the respective activities.

The possibility of creating artificial helical media [12–15], including those with preset parameters, and ferromagnetic helical structures simulating the properties of cholesterics at microwaves raises interest in their behavior and applications. In addition, it is known that in most modern theories regarding the optical properties of complex structures, the medium is represented as a multilayer system and any approach to associated problems is based on the exact solution of the problem of light transmission through a layer of finite thickness.

DISPERSION RELATION

In the presence of a magnetic field applied along the HRS axis (*z* axis), the permittivity, $\hat{\varepsilon}$, and permeability, $\hat{\mu}$, tensors have the form

$$\hat{\varepsilon}(z) = \varepsilon_m \begin{pmatrix} 1 + \delta_\varepsilon \cos 2az & \pm \delta_\varepsilon \sin 2az \pm i g_e/\varepsilon_m & 0 \\ \pm \delta_\varepsilon \sin 2az \mp i g_e/\varepsilon_m & 1 - \delta_\varepsilon \cos 2az & 0 \\ 0 & 0 & 1 - \delta_\varepsilon \end{pmatrix}, \quad (1)$$

$$\hat{\mu}(z) = \mu_m \begin{pmatrix} 1 + \delta_\mu \cos 2az & \pm \delta_\mu \sin 2az \pm i g_m/\mu_m & 0 \\ \pm \delta_\mu \sin 2az \mp i g_m/\mu_m & 1 - \delta_\mu \cos 2az & 0 \\ 0 & 0 & 1 - \delta_\mu \end{pmatrix}.$$

Here, $\varepsilon_m = (\varepsilon_1 + \varepsilon_2)/2$, $\varepsilon_a = (\varepsilon_1 - \varepsilon_2)/2$, $\delta_\varepsilon = \varepsilon_a/\varepsilon_m$, $\mu_m = (\mu_1 + \mu_2)/2$, $\mu_a = (\mu_1 - \mu_2)/2$, $\delta_\mu = \mu_a/\mu_m$, $\varepsilon_i = \varepsilon_{0i} + \Delta\varepsilon_i$ $(i = 1, 2)$, ε_{0i} are the principal values of the permittivity tensor in the absence of the external magnetic field, and $\Delta\varepsilon_i$ are disturbances of the principal values of the local permittivity tensor due to the external magnetic field [16–21], $\mu_i = \mu_{0i} + \Delta\mu_i$, μ_{0i} are the principal values of the permeability tensor in the absence of the external magnetic field, $\Delta\mu_i$ are disturbances of the principal values of the local permeability tensor due to the external magnetic field [16–21], $\mathbf{g}_e = \mathbf{g}_e(\mathbf{H}_{ext})$ and $\mathbf{g}_m = \mathbf{g}_m(\mathbf{H}_{ext})$ are the vectors of the gyroelectric and gyromagnetic magnetooptic activities, $a = 2\pi/\sigma$, $\sigma = \sigma_0 + \Delta\sigma$, σ_0 is the helix pitch in the absence of the magnetic field, and $\Delta\sigma$ is a disturbance of the helix pitch in the presence of the magnetic field [1–5]. We assume that the principal axes of the tensors $\hat{\varepsilon}$ and $\hat{\mu}$ coincide and one of them (z axis) is aligned with the external magnetic field.

In view of (1), for light propagating along the axis of the medium, a solution to the Maxwell equations $\mathrm{curl}\mathbf{H} = \frac{1}{c}\frac{\partial\mathbf{D}}{\partial t}$ and $\mathrm{curl}\mathbf{E} = -\frac{1}{c}\frac{\partial\mathbf{B}}{\partial t}$ is sought in the form

$$\mathbf{E}(z, t) = \sum_{j=1}^{4}\left\{ E_j^+ \mathbf{n}_+ \exp\left[i\frac{2}{\pi}\sqrt{\varepsilon_m\mu_m}(\chi + b)z \right]\right.$$

$$\left. + E_j^- \mathbf{n}_- \exp\left[i\frac{2\pi}{\lambda}\sqrt{\varepsilon_m\mu_m}(-\chi + b)z \right]\right\}\exp(-i\omega t),$$

(2)

where $\chi = \lambda/(\sigma\sqrt{\varepsilon_m\mu_m})$, λ is the wavelength in a vacuum, and $\mathbf{n}_\pm = (\mathbf{x} \pm \mathbf{y})/\sqrt{2}$ are the unit vectors of circular polarizations.

Substituting (2) into the Maxwell equations, we come to the dispersion relation for b:

$$b^4 + a_1 b^2 + a_2 b + a_3 = 0, \qquad (3)$$

where $a_1 = -2(1 + \chi^2 + G_e G_m - \delta_\varepsilon\delta_\mu)$, $a_2 = -4\chi(G_e + G_m)$, $a_3 = -2\chi^2(1 + \delta_\varepsilon\delta_\mu + G_e G_m) + (1 - \delta_\varepsilon^2 - G_e^2)(1 - \delta_\mu^2 - G_m^2) + \chi^4$, $G_e = g_e/\varepsilon_m$, and $G_m = g_m/\mu_m$.

Thus, unlike the case when the magnetic field is absent (the dispersion relation is biquadratic), the dispersion relation in the case under consideration is a complete fourth-degree equation. The solution to

Eq. (3) is given by

$$b_{1,2} = \sqrt{\frac{s}{2}} \pm \sqrt{\left(-\frac{a_1}{2} - \frac{s}{2} - \frac{a_2}{2\sqrt{2s}}\right)},$$

$$b_{3,4} = -\sqrt{\frac{s}{2}} \pm \sqrt{\left(-\frac{a_1}{2} - \frac{s}{2} + \frac{a_2}{2\sqrt{2s}}\right)},$$

(4)

where

$$s = w - \frac{p}{3w} - \frac{a_1}{3}, \quad w = \sqrt[3]{-\frac{q}{2} + v},$$

$$v = \sqrt{\left(\frac{q}{2}\right)^2 + \left(\frac{p}{3}\right)^3}, \quad p = -\frac{a_1}{12} - a_3,$$

$$q = a_1\left(a_3 - \frac{a_1^2}{36}\right) - \frac{a_2^2}{8}.$$

BOUNDARY-VALUE PROBLEM FOR A LAYER

Consider the boundary-value problem of light transmission through an HPS layer of finite thickness placed in a magnetic field. The field is aligned with the axis of the medium, and the axis is normal to the side surfaces. Light is incident normally to the layer surface. A solution to the problem is represented as

$$\mathbf{E}_r = \hat{R}\mathbf{E}_i, \quad \mathbf{E}_t = \hat{T}\mathbf{E}_i, \qquad (5)$$

where the subscripts i, r, t refer to the incident, reflected, and transmitted fields, respectively.

In (5),

$$\mathbf{E}_{i,r,t} = E_{i,r,t}^+ \mathbf{n}_+ + E_{i,r,t}^- \mathbf{n}_- = \begin{bmatrix} E_{i,r,t}^+ \\ E_{i,r,t}^- \end{bmatrix},$$

and $\hat{T}$ and $\hat{R}$ are 2×2 John matrices of the transmitted and reflected waves. Their components are given by

$$R_{11} = \frac{1}{2}\sum_{i,j,k,l} e_{ijkl}[-\gamma_i^-\alpha_j^- + \gamma_i^+\alpha_j^+$$

$$+ i(\gamma_j - \gamma_i + 2\alpha_i\beta_j)]\gamma_k^-\alpha_l^+ f_k f_l/\Delta,$$

$$R_{12} = \frac{1}{2}\sum_{i,j,k,l} e_{ijkl}[-\gamma_i^-\alpha_j^- - \gamma_i^+\alpha_j^+$$

$$+ i(\gamma_i - \gamma_j + 2\alpha_i\beta_j)]\gamma_k^-\alpha_l^+ f_k f_l/\Delta,$$

$$R_{21} = \frac{1}{2}\sum_{i,j,k,l} e_{ijkl}[-\gamma_i^-\alpha_j^- - \gamma_i^+\alpha_j^+$$

$$+ i(\gamma_j - \gamma_i - 2\alpha_i\beta_j)]\gamma_k^-\alpha_l^+ f_k f_l/\Delta,$$

$$R_{22} = \frac{1}{2}\sum_{i,j,k,l} e_{ijkl}[-\gamma_i^-\alpha_j^- + \gamma_i^+\alpha_j^+$$

$$-i(\gamma_j - \gamma_i + 2\alpha_i\beta_j)]\gamma_k^-\alpha_l^+ f_k f_l/\Delta,$$

(6)

and

$$T_{11} = \frac{1}{2}\sum_{i,j,k,l} e_{ijkl}(\alpha_i^- + i\gamma_i^+)$$

$$\times [\alpha_j^+(\gamma_k - \gamma_l) + 2\alpha_k\beta_l\gamma_j^-]f_k f_l f_j e^{-iad}/\Delta,$$

$$T_{12} = \frac{1}{2}\sum_{i,j,k,l} e_{ijkl}(\alpha_i^- - i\gamma_i^+)$$

$$\times [\alpha_j^+(\gamma_k - \gamma_l) + 2\alpha_k\beta_l\gamma_j^-]f_k f_l f_j e^{-iad}/\Delta,$$

(7)

$$T_{21} = \frac{1}{2}\sum_{i,j,k,l} e_{ijkl}(\alpha_i^- + i\gamma_i^+)$$

$$\times [\alpha_j^+(\gamma_k - \gamma_l) - 2\alpha_k\beta_l\gamma_j^-]f_k f_l f_j e^{-iad}/\Delta,$$

$$T_{22} = \frac{1}{2}\sum_{i,j,k,l} e_{ijkl}(\alpha_i^- - i\gamma_i^+)[\alpha_j^+(\gamma_k - \gamma_l)$$

$$+ 2\alpha_k\beta_l\gamma_j^-]f_k f_l f_j e^{-iad}/\Delta, \quad i,j,k,l = 1,2,3,4.$$

Here,

$$\Delta = \sum_{i,j,k,l} e_{ijkl}\gamma_i^+\alpha_j^-\gamma_k^-\alpha_l^+ f_k f_l, \quad \alpha_j^\pm = \alpha_j \pm \beta_j,$$

$$\gamma_j^\pm = 1 \pm \gamma_j, \quad f_j = \exp(i2\pi\sqrt{\varepsilon_m\mu_m}b_j d/\lambda),$$

$$\alpha_j = i\{b_j[b_j^2 - \chi^2 - G_e G_m - (1 - \delta_\mu)(1 + \delta_\varepsilon)]$$

$$+ \chi[G_m(1 + \delta_\varepsilon) + G_e(1 + \delta_\mu)]\}/\{\chi[b_j^2 - \chi^2 + G_e G_m$$

$$+ (1 + \delta_\mu)(1 - \delta_\varepsilon)] - b_j[G_m(1 - \delta_\varepsilon) + G_e(1 - \delta_\mu)]\},$$

$$\beta_j = \alpha\{b_j[iG_m + (1 - \delta_\mu)\alpha_j]$$

$$- \chi[\alpha_j G_m + i(1 - \delta_\mu)]\}/[G_m^2 - (1 - \delta_\mu^2)],$$

$$\gamma_j = \alpha\{b_j[i\alpha_j G_m - (1 + \delta_\mu)]$$

$$+ \chi[G_m + i\alpha_j(1 + \delta_\mu)]\}/[G_m^2 - (1 - \delta_\mu^2)],$$

$$\alpha = \sqrt{\varepsilon_m/\mu_m}\varepsilon,$$

ε is the permittivity of media adjacent to the layer on both sides (it is assumed that the parameter of magnetooptic activity of the adjacent media can be neglected in comparison with the activity parameters of the HPS), e_{ijkl} is the Levi–Civita symbol, and d is the thickness of the layer. With (5) and (6), one can calculate the coefficients of reflection, $R = |E_r|^2/|E_i|^2$, and transmission, $T = |E_t|^2/|E_i|^2$; the rotation of plane of polarization $\varphi = -\arg(\zeta)/2$; the polarization ellipticity $c = (|\zeta| - 1)/(|\zeta| + 1)$, where $\zeta = E_t^+/E_t^-$ is the polarization function of the transmitted wave; the absorption in the layer $Q = 1 - (R + T)$; and other optical characteristics of the system.

Consider first a simple case, namely, $\hat{\mu} = \hat{I}$, $g_m = 0$, and the local anisotropy of the medium $\delta_\varepsilon = \delta \ll 1$.

The numerical values of the reflection coefficients with and without the magnetic field indicate that the field shifts the range of selective (diffraction) reflection (SRR). This shift is associated only with the magnetooptic activity of the HPS (unlike the shift due to a change in the helix pitch), is directed toward shorter waves (effect quadratic in g), and is very small (15 Å for $g_e = g = \pm 0.1$). Near the SSR boundaries, where the width of the transmission coefficient peak and the shift are on the same order of magnitude at certain values of the parameter $\delta d/\sigma$, the external magnetic field changes the transmission coefficient by 100%. Specifically, near the SRR, the transmission minima ($T \approx 0$) may turn to maxima ($T \approx 1$) and vice versa at particular thicknesses. This means that the system can be used as a perfect optical shutter.

As follows from calculations, the transmission depends strongly on α in the presence of the field. Inside the SRR, the value $\Delta T_1 = T(g) - T(0)$ (where $T(0)$ is the transmission coefficient in the absence of the field and $T(g)$ is that in the presence of the field) depends on λ and g linearly. Outside the SRR, the function $\Delta T_1(\lambda)$ oscillates. The period of these oscillations, as well as of the transmission (T) and reflection (R) oscillations, is a function of the parameter $\delta d/\sigma$. These oscillations are modulated, the modulation period and amplitude varying with distance from the SRR boundary. At $\alpha \neq 1$, the amplitudes of the oscillations grow sharply and reach the same order of magnitude as the oscillations of T at certain wavelengths ($g = 10^{-4}$). If α is other than unity inside the SRR, ΔT_1 declines. Outside the SRR, the g dependence of ΔT_1 also oscillates. Such behavior of the HRS layer in the external magnetic field can be used for the intensity modulation of light.

As was noted above, HPSs subjected to a magnetic field exhibit the effect of wave irreversibility. Particularly, the transmitted intensities differ when the directions of light incidence and magnetic field coincide and when the light and the field are oppositely directed. It was shown numerically that inside the SRR the parameter $\Delta T_2 = T(g) - T(-g)$ (where $T(g)$ is the transmission when the directions coincide and $T(-g)$ is that when the
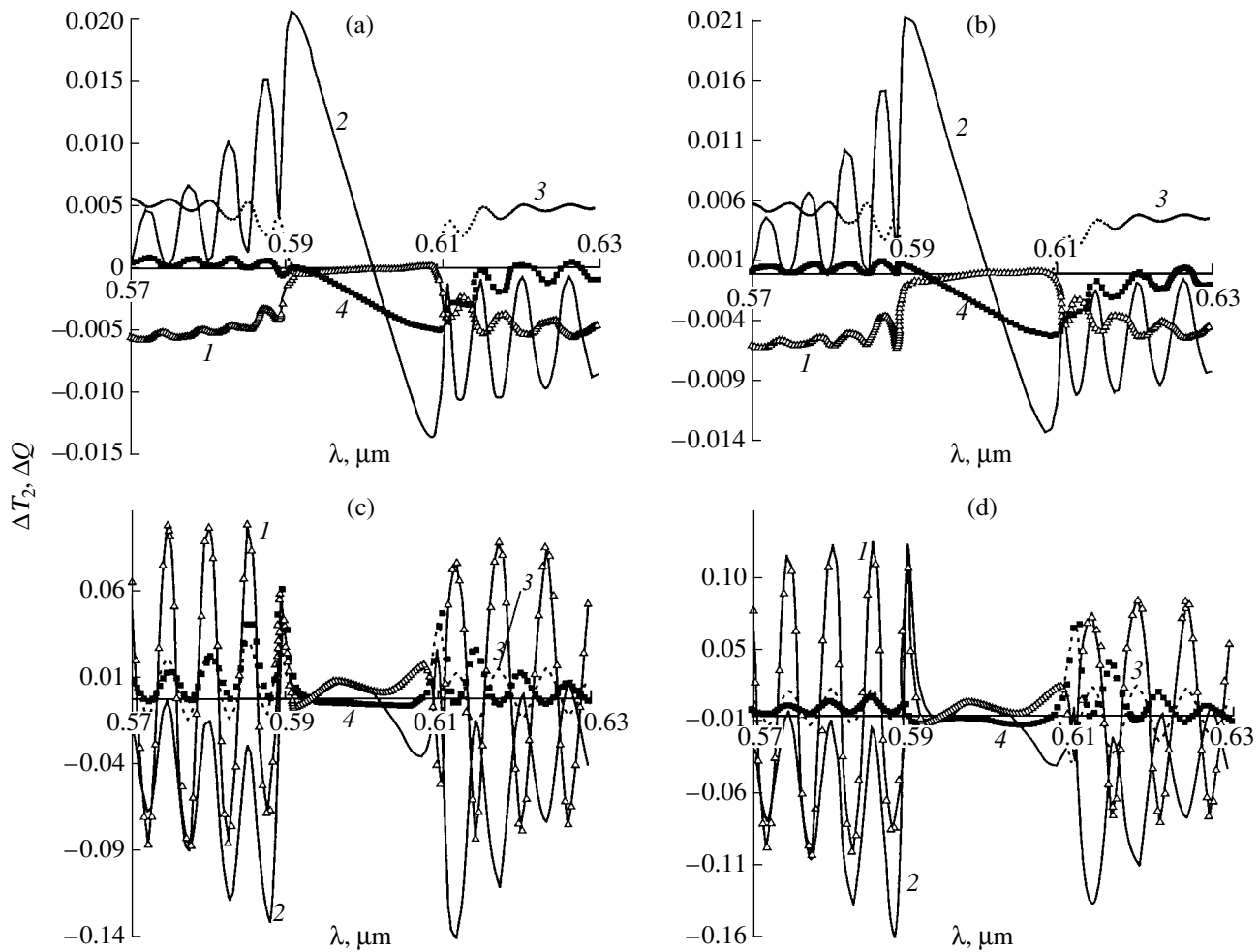
Fig. 1. ΔT_2 vs. wavelength for (a, c) isotropic ($\mathrm{Im}\varepsilon_m = 0.001$, $\mathrm{Im}\varepsilon_a = 0$) and (b, d) anisotropic ($\mathrm{Im}\varepsilon_m = 0.001$, $\mathrm{Im}\varepsilon_a = 0.001$) absorption in the (a, b) absence ($\alpha = 1$) and (c, d) presence ($\alpha = \sqrt{\varepsilon_m}$) of insulating boundaries. Right-hand helix, $\mathrm{Re}\varepsilon_m = 2.25$, $\mathrm{Re}\varepsilon_a = 0.0675$, $g = 0.1$, $\sigma = 0.4$ μm, and $d = 50\sigma$.

directions are opposite to each other) depends on λ and g linearly. Inside the SRR, both dependences oscillate. The following rules should be noted here.

(1) If outside the SRR α is other than unity, the transmission irreversibility rises (for $\alpha = 1$, the amplitude of ΔT_2 oscillations is roughly 100 times lower than for $\alpha \neq 1$). Because of this, the value of ΔT_2 becomes comparable to the value of T by the order of magnitude at certain wavelengths, and transmission irreversibility in the HPS is easy to observe in experiments. As α moves away from unity, the irreversibility does not grow. Conversely, starting from certain values ($\alpha > 1$ and $\alpha < 1$), the degree of irreversibility near the transmission peaks slightly decreases. At other frequencies, its decrease is more pronounced and α even approaches zero at particular frequencies. This is because Fresnel reflections are greatly enhanced: the intensity varies insignificantly within each of the HPS layers, and the

forward and backward propagation directions turn out to be almost equivalent.

(2) If α is other than unity inside the SRR, ΔT_2 decreases.

Note that the irreversibility effect is sometimes very significant and can be used for designing related elements in optical gyros, as well as for studying the Faraday effect and optical activity in HPS layers.

Numerical analysis shows that the polarization ellipticity e, rotation of plane of polarization φ, radiation absorption Q, and circular dichroism D_c are also nonreproducible.

Figure 1 shows the wavelength dependence of (*1, 2*) ΔT_2 and (*3, 4*) $\Delta Q = Q(g) - Q(-g)$ when the absorption is (a, c) isotropic and (b, d) anisotropic in the (a, b) absence and (c, d) presence of insulating boundaries. Curves *1* and *3* correspond to the case of circularly polarized light diffracted by the HPS structure; curves *2* and *4*, to light linearly polarized along the *x* axis. Fig-
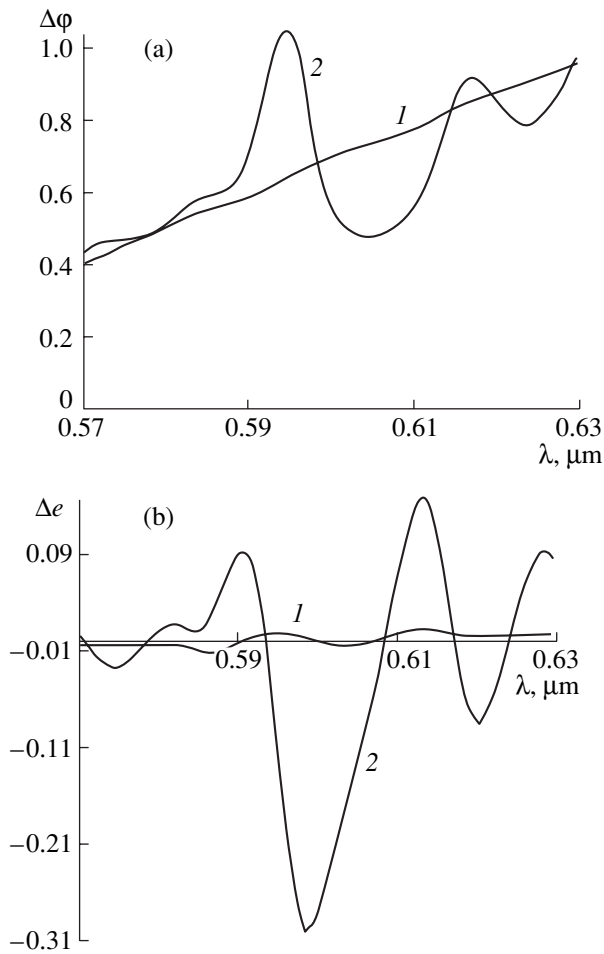
Fig. 2. Irreproducibilities of the (a) rotation of plane of polarization $\Delta\varphi$ and (b) polarization ellipticity Δe vs. wavelength. Absorption is absent. Light incident on the HPS layer is linearly polarized along the x axis. $d = 20\,\sigma$. The rest of the parameters are the same as in Fig. 1.

ure 2 demonstrates the irreversibility of (a) the rotation of plane of polarization $\Delta\varphi = \varphi(g) - \varphi(-g)$ and (b) ellipticity $\Delta e = e(g) - e(-g)$ vs. wavelength in the (1) absence and (2) presence of insulating boundaries.

Note that the irreversibility effects are weak if the local anisotropy is weak and are of no interest in spite of their uniqueness. When the anisotropy of the medium or the parameter of magnetooptic activity grows, they become appreciable and, hence, of practical value.

OPTICAL PROPERTIES OF THE HIGHLY ANISOTROPIC HPS LAYER IN A MAGNETIC FIELD

Figures 3 and 4 demonstrate the λ dependences of the irreproducibilities ΔT_2 and ΔT_1 when the anisotropy is high ($\delta \approx 1$).

From these dependences and also from the numerical results, the following conclusions can be drawn.

(1) As the local anisotropy δ (δ_ε or δ_μ) or the parameter of magnetooptic activity g (g_e or g_m) grows, the value of $|\Delta T_2|$ increases on average.

(2) If the anisotropy is high, $|\Delta T_2|$ is close to unity for specific wavelengths and gyration parameters g_e and g_m (i.e., for specific magnetic field strengths). This means that such a system can function as a perfect optical diode completely transmitting the radiation in one direction and cutting it in the opposite direction. This system can also operate as a one-side reflector.

(3) As the local anisotropy δ (δ_ε or δ_μ) or the parameter of magnetooptic activity g (g_e or g_m) grows, the value of $|\Delta T_1|$ also increases on average.

(4) If the anisotropy is high, $|\Delta T_1|$ is close to unity for specific wavelengths and gyration parameters g_e and g_m. This means that in this case, too, the system can function as a perfect optical diode.

DISCUSSION AND CONCLUSIONS

To shed light on a mechanism behind high transmission irreversibility, we studied the behavior of the wave numbers $k_j = 2\pi\sqrt{\varepsilon_m\mu_m}\,b_j/\lambda$ and self-polarizations (SPs). In the absence of an applied magnetic field, k_j curves are known to be symmetric about the frequency (or wavelength) axis. The application of the field breaks the symmetry: resonance k_j (i.e., those becoming complex in a certain frequency range in the absence of absorption) and nonresonance k_j shift in opposite directions along the λ axis. An unusual situation arises when the anisotropy of the medium or its magnetooptic activity is high. At particular values of these parameters, another (second) SRR appears in the long-wave part of the spectrum, with its boundary approaching infinity. The first SRR also tends to shorter waves in this case (i.e., at a high anisotropy), but the shift is now more significant. As the parameters mentioned above increase further, the short-wave boundary of the second SRR approaches the long-wave boundary of the first SRR and, at certain values of these parameters, the entire spectral range becomes a range of diffraction. Note that the second SRR also appears near the absorption line if the optical activity of cholesteric molecules is taken into account [22].

High values of $|\Delta T_2|$ reflect the fact that the diffraction reflection (and transmission) of light in the second SRR depends on the direction of the applied magnetic field.

It is known that SPs are two polarizations that do not change when light passes through a system. The SPs of an HPS layer are specific in that light with one SP experiences diffraction reflection (such SP will be referred to as diffracting SP), while light with the other does not. In the absence of absorption, the two SPs are orthogonal to each other. The wavelength dependence of the SP ellipticity indicates that the magnetic field reversal
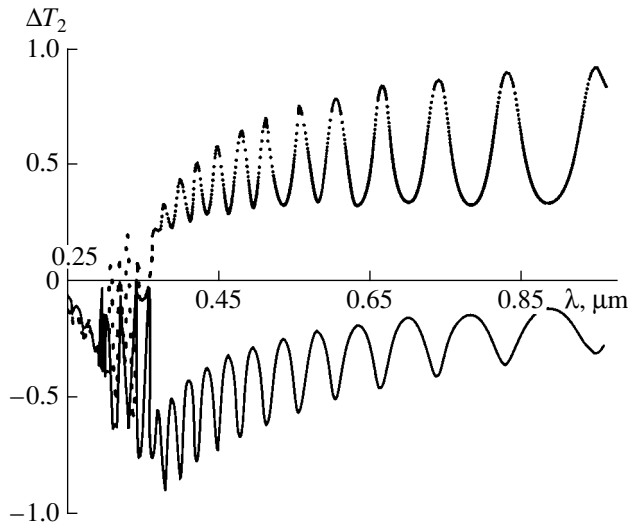
Fig. 3. Transmission irreversibility ΔT_2 vs. wavelength λ in the case of high anisotropy ($\delta_\varepsilon = 0.9$, $\varepsilon_{0m} = 0.25$, $\delta_\mu = 0.05$, $\mu_{0m} = 1.2$, $g_e = 0.275$, $g_m = 0.1$, $\Delta\varepsilon_1 = \Delta\varepsilon_2 = \Delta\mu_1 = \Delta\mu_2 = 0.00001$, $\Delta\sigma = 0.000001$ μm, and $d = 10$ σ). Continuous curve, incident light is clockwise polarized; dashed curve, light is linearly polarized along the x axis. Right-hand helix, $\sigma_0 = 0.4$ μm, absorption is absent. For the designation of the curves, see Fig. 1.
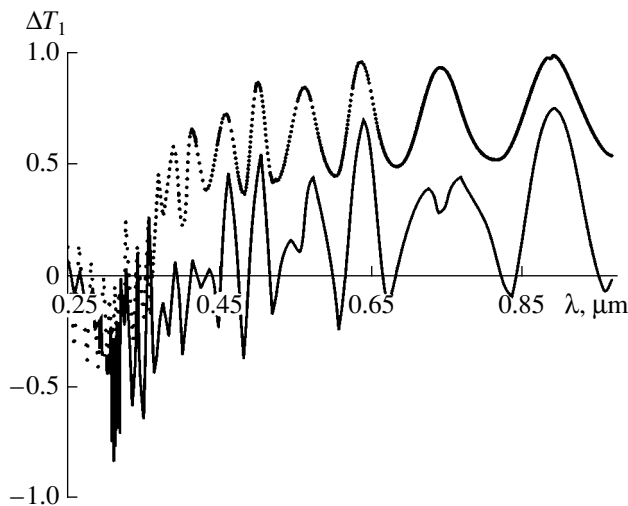

Fig. 4. Transmission irreversibility ΔT_1 vs. wavelength λ in the case of high anisotropy. The parameters and designations are the same as in Fig. 3.

changes the ellipticity of the diffracting (and the other) SP in the frequency range of the first SRR insignificantly. At the same time, the field reversal in the range of the second SRR changes the sign of the diffracting SP ellipticity. Thus, when the magnetic fields are oppositely directed, the ellipticities of the diffracting SP nearly coincide within the first SRR and are almost orthogonal to each other within the second one. Such an effect causes a high transmission irreversibility in the frequency range of the second SRR.

REFERENCES

1. P. de Gennes, *The Physics of Liquid Crystals* (Clarendon, Oxford, 1974; Mir, Moscow, 1977).
2. L. M. Blinov, *Electro- and Magnetooptics of Liquid Crystals* (Nauka, Moscow, 1978).
3. A. P. Kapustin, *Electrical and Acoustical Properties of Liquid Crystals* (Nauka, Moscow, 1973).
4. V. A. Belyakov, *Diffraction Optics of Periodic Media of Complex Structure* (Nauka, Moscow, 1988).
5. S. Chandrasekhar, *Liquid Crystals*, Ed. by A. A. Vedenov and I. G. Chistyakov (Cambridge Univ. Press, Cambridge, 1977; Mir, Moscow, 1980).
6. O. S. Eritsyan, Izv. Akad. Nauk Arm. SSR, Fiz. **13**, 347 (1978).
7. A. H. Gevorgyan, Uch. Zap. Erevan. Gos. Univ., No. 2, 66 (1987).
8. V. A. Kienya and I. V. Semchenko, Kristallografiya **39**, 514 (1994) [Crystallogr. Rep. **39**, 457 (1994)].
9. G. A. Vardanyan, A. H. Gevorgyan, O. S. Eritsyan, *et al.*, Kristallografiya **43**, 793 (1998) [Crystallogr. Rep. **43**, 740 (1998)].
10. O. M. Arakelyan, A. H. Gevorgyan, and O. S. Eritsyan, Izv. Akad. Nauk Arm., Fiz. **35** (5), 255 (2000).
11. L. B. Felsen and N. Marcuvitz, *Radiation and Scattering of Waves* (Prentice-Hall, Englewood Cliffs, 1973; Mir, Moscow, 1978).
12. G. Gerritsen and R. Yamaguchi, Usp. Fiz. Nauk **107**, 705 (1972).
13. P. D. Sunal, A. Lakhtakia, and R. Messier, Opt. Commun. **158**, 119 (1998).
14. P. I. Rovira, R. A. Yarussi, R. W. Collins, *et al.*, Appl. Phys. Lett. **71**, 1180 (1997).
15. I. Hodgkinson, Q. H. Wu, B. Knight, *et al.*, Appl. Opt. **39**, 642 (2000).
16. V. M. Agranovich and V. L. Ginzburg, *Crystal Optics with Spatial Dispersion, and Excitons* (Nauka, Moscow, 1979; Springer-Verlag, New York, 1984).
17. F. I. Fedorov, *Theory of Gyrotropy* (Nauka i Tekhnika, Minsk, 1976).
18. A. K. Zvezdin and V. A. Kotov, *Magnetooptics of Thin Films* (Nauka, Moscow, 1988).
19. G. S. Krinchik, *Physics of Magnetic Phenomena* (Mosk. Gos. Univ., Moscow, 1976).
20. V. V. Eremenko, N. F. Kharchenko, Yu. G. Litvinenko, and V. M. Naumenko, *Magnetooptics and Spectroscopy of Antiferromagnets* (Naukova Dumka, Kiev, 1989).
21. O. G. Vlokh, *Spatial Dispersion Phenomena in Parametric Crystal Optics* (Vishcha Shkola, L'vov, 1984).
22. V. G. Kamenskiĭ and E. I. Kats, Opt. Spektrosk. **45**, 1106 (1978) [Opt. Spectrosc. **45**, 877 (1978)].

Translated by V. Isaakyan

<div align="center">

OPTICS,
QUANTUM ELECTRONICS

</div>

Optical Data Recording with Vanadium Dioxide–Based Film Reversible Media

<div align="center">

A. S. Oleĭnik

Saratov State Technical University, Saratov, 410016 Russia
Received December 18, 2001; in final form, February 5, 2002

</div>

Abstract—Thermally induced semiconductor–metal phase transitions taking place in Al/VO_2/insulator and VO_2/insulator film systems upon recording optical data are investigated. Experiments show that the rate of the processes depends on thermophysical parameters of the films, as well as on the energy exposure. The dynamic range of the energy exposure sensitivity is defined by the hysteresis loop width. The shape and width of the hysteresis loop depends on the VO_2 film thickness and the nonstoichiometric ratio profile across the film. © 2002 MAIK "Nauka/Interperiodica".

The semiconductor–metal phase transition (SMPT) in VO_2 is used in optoelectronics for designing optical visualizers [1], holographic transparencies [2], and optical detectors [3]. The basic advantage of VO_2 films is their data storage capability because of SMPT-related hysteresis. In thin VO_2 films, the width of the hysteresis loop depends on fabrication technology and may reach 40 K [2].

In this work, we study the SMPT upon optical data recording on VO_2 films prepared by oxidizing evaporatively deposited V films in air. For the films thus prepared, the volume oxidation coefficient (the ratio of the oxide volume to the initial volume of a metal) is given by [4]

$$K = \frac{M\gamma_M}{nA\gamma_O},$$

where A is the atomic weight of the metal, n is the number of metal atoms incorporated into an oxide molecule, M is the molecular weight of the oxide, and γ_M and γ_O are the respective specific weights of the metal and oxide. For VO_2, $K = 2.25$; in this case, the oxide grows dense and offers protective properties. Internal stresses arise in the oxide film because of the difference between the metal and oxide volumes.

VO_2 layers with desired electrical and optical properties were prepared by simultaneously varying the V film initial thickness and time of oxidation in air at 753 ± 10 K [2]. In this case, the oxidation process in any temperature range obeys a particular parabolic law. It was shown [5] that when V is oxidized in air at 753 K, the temperature parabolic variation across the V film layer causes the parabolic distribution of the VO_2 phase across the oxide. The VO_2 phase thus formed is nonstoichiometric, with the nonstoichiometric ratio decreasing monotonically across the VO_2 layer within the homogeneity range. The slope of the hysteresis loop depends on the oxidation temperature. Note that the oxidation temperature is limited from above, since the oxidation takes an explosive character at $T > 753$ K.

Experiments showed [5] that the phase transition in VO_2 films exhibits hysteresis with the loop width depending on the VO_2 film thickness. Thus, the variation of the nonstoichiometric ratio across the VO_2 layer specifies the width of the temperature hysteresis. Figures 1 and 2 demonstrate, respectively, the temperature dependence of the sheet resistivity for the VO_2 films of various thickness and the temperature hysteresis for the reflection coefficient of the Al/VO_2 systems (with various thicknesses of VO_2 film) on a silicon single-crystal substrate at a wavelength of 1.06 μm.

In VO_2 polycrystalline films, the SMPT of the first kind occurs at temperatures between 317 and 359 K and is accompanied by heat absorption. Also, the monoclinic crystal lattice changes to tetragonal.

Al/VO_2/insulator (I) and VO_2/I systems were formed on optical-grade mica, quartz, and ceramic substrates. According to the topological theory of 2D crystallite growth [6], the local equilibrium condition (or the minimum condition for surface energy) requires the 2D space to be filled with regularly shaped crystallites in such a way that they are adjacent to each other along their sides, which in turn make an angle of 120° with each other. In other words, the 2D space must be filled with regular hexagons. With regard for the fact that a 2D set of polygons arises during growth from randomly located nuclei, the local equilibrium condition gives rise to boundaries between crystallites with the curved surface. Nevertheless, one can theoretically devise a stable set using a packing of regular polygons (polyhe-

dral crystallites) in this case. For the 3D space, such a construction is impossible.

Microscopic studies [2, 5] suggest that crystallites are regularly packed in the 2D layer of the VO_2 phase and that the crystallite sizes are well correlated. Today's technologies allow for the production of VO_2 films with given crystallite sizes. For example, if the film thickness is 80–100 nm, the crystallite sizes normal to the substrate equal the film thickness. In the plane of the substrate, the crystallite sizes vary between 50 and 100 nm according to the film thickness. The shape of the crystallites and its reproducibility depend on the crystal perfection of the substrate, since the oxide "inherits" lattice distortions in the vanadium film.

VO_2-based oxide films were shown [2, 5] to have a composite (layered) structure consisting of the thin surface V_2O_5 phase ($\approx 5\%$ of the total film thickness), basic VO_2 phase ($\approx 70\%$ of the total thickness), and successively arranged V_3O_5, V_2O_3, VO, and V layers accounting for the remaining 25% of the oxide thickness. Thus, we are dealing with a polycrystalline multiphase layered structure with a regular arrangement of the crystallites in the substrate plane. Since the SMPT in these film structures exhibit hysteresis, one can distinguish half-tones in optical data recording. Near the phase transition in VO_2 films, the semiconductor and metal phases were found to mix [7], their partial concentrations being a function of temperature. Since the composition of the VO_2 phase deviates from the stoichiometric one across the film, the phase transition occurs in succession. Because of this, the transition in the VO_2 film has a distinct boundary, which makes it possible to record complex images.

Upon irradiation, VO_2/I and Al/VO_2/I bistable structures experience several stages: absorption of light; energy transfer into the VO_2 layer, followed by heating the layer to the phase transition temperature; and absorption of the latent heat (phase transition of the first kind). It is known [8] that when the electron and ion masses in solids greatly differ (as in VO_2), energy exchange proceeds rather slowly. The times it takes for the electron and ionic subsystems to come to equilibrium are 10^{-14}–10^{-13} s and 10^{-13}–10^{-12} s, respectively. The time taken to set equilibrium between the subsystems is much greater, $\approx 10^{-10}$ s [8].

Hall effect measurements [9] showed that the carrier concentration in VO_2 films depends on temperature. At the phase transition temperature, the resistivity experiences a jump associated with a jump in the carrier concentration. One can expect that the early stage of absorption (intrinsic photoeffect), i.e., the growth of the electron concentration in VO_2, proceeds when the lattice is cold.

At the initial time instant, the thickness of the area heated is defined by the depth S of radiation penetration into the material: $S = \alpha^{-1}$, where α is the absorption coefficient. For a 100- to 120-nm-thick VO_2 layer, $\alpha \sim$
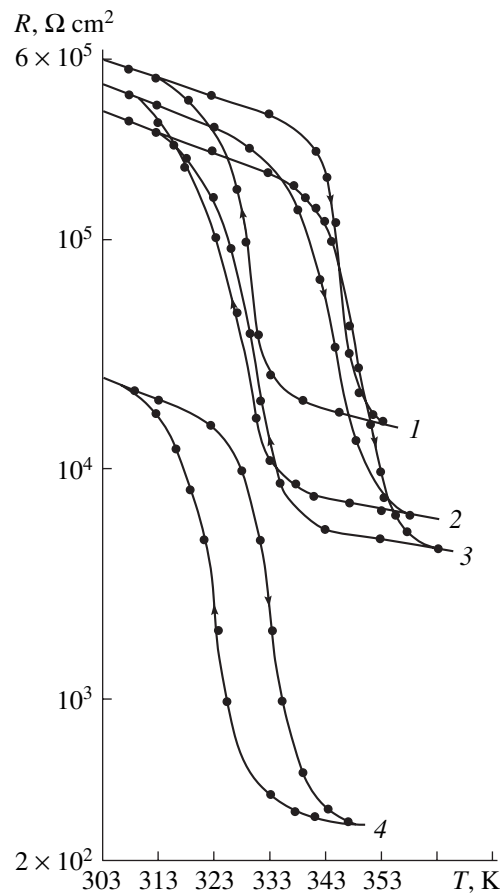

Fig. 1. Temperature dependence of the sheet resistivity of the VO_2 films with a thickness $d = 60$ (*1*), 80 (*2*), 100 (*3*), and 140 nm (*4*).

10^4 cm^{-1} for wavelengths between 0.4 and 1.06 μm. The thickness of heating increases with time because of heat conduction as $2.36\sqrt{a\tau}$ [10], where a is the thermal diffusivity and τ is the irradiation time (laser shot duration). At $\tau = 10^{-9}$ s, the thickness of the heated area in VO_2 is estimated at 133 nm. At $\tau = 10^{-9}$ s, the radiation power density causing the SMPT in a 100-nm-thick oxide layer was found to be 10^4 W/cm^2. Thus, the problem of heating a 100- to 120-nm-thick oxide layer (where VO_2 accounts for 70% of the total oxide thickness) by a shot of duration no longer than 7×10^{-10} s can be treated in terms of the conventional theory of heat conduction [11].

In [12], the SMPT in VO_2 was shown to be related to electron–phonon interaction, which is enhanced with temperature, and to the respective enhancement of electron delocalization. Because of this, the width W of the $3d$ gap increases, the energy U of intraatomic Coulomb repulsion decreases, and the SMPT occurs when the ratio U/W reaches a critical value. It is also known [6] that the formation of a new phase in a solid does not disturb the continuity of the crystal lattice: at the interface,
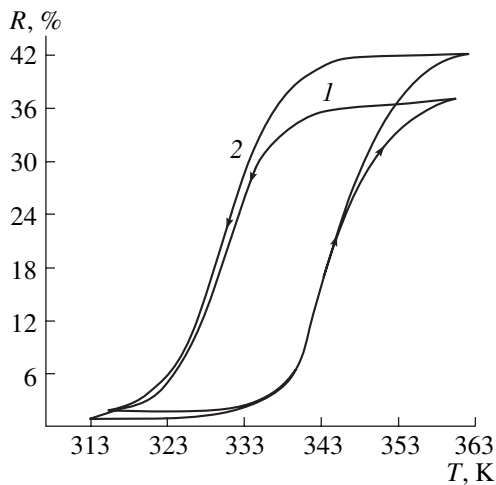
Fig. 2. Temperature hysteresis of the reflection coefficient of the Al/VO$_2$ system on the polysilicon substrate. The thickness d of the VO$_2$ film is 110 (*1*) and 112 (*2*) nm. The wavelength is 1.06 μm.
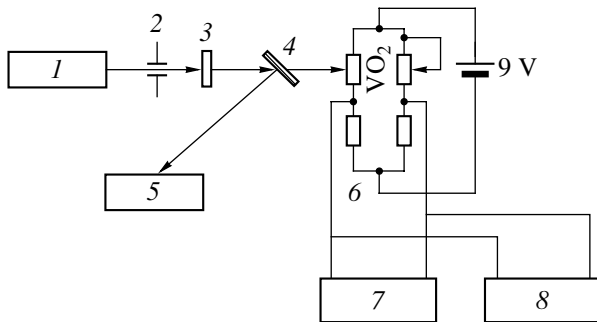

Fig. 3. Scheme for measuring the output of the VO$_2$-based thermal detector: *1*, LG-106M-1 laser; *2*, diaphragm; *3*, shutter; *4*, beam splitter; *5*, IMO-2-2 working standard of laser radiation; *6*, bridge with detector; *7*, Shch68003 voltmeter; and *8*, S1-70 oscilloscope.

one lattice smoothly passes into the other. In our case, the monoclinic lattice of VO$_2$ transforms into tetragonal. Phase transitions thermally activated can be considered as the result of a polymorphic transformation of the martensitic type, for which the growth of a phase takes place via the barrierless displacement (slip) of the interface. Martensitic transformation is initiated when a VO$_2$ film is heated up to the SMPT temperature. Process-related plastic strains in VO$_2$ films increase the degree of transformation and may also favor reverse transformation. Under martensitic transformation, the collective motion of many thousands of atoms proceeds with a speed approaching that of acoustic waves in a crystal. According to [13], the speed of sound in VO$_2$ is about 4500 m/s. Taking into account that the crystallite size in a VO$_2$ polycrystalline film is ≈100 nm, one can expect that the phase transition duration will not exceed ≈5×10^{-11} s; that is, laser heating does not retard the

growth of the VO$_2$ film. Thus, the speed of response of media under adiabatic exposure depends on the time of thermal diffusion of optical energy into the VO$_2$ film.

In VO$_2$/I and Al/VO$_2$/I systems, the protective layer is transparent to the radiation recorded; therefore, its heating is negligible. Under adiabatic exposure, a 100- to 120-nm-thick VO$_2$ film is heated for $(7-8) \times 10^{-10}$ s. It was shown [14] that when thin films, such as the VO$_2$ films, are heated by a source giving a thermal spot of radius $r \gg (at)^{1/2}$ (where a is the thermal diffusivity and t is the heating time), the temperature distribution in the film depends only on r and t for a given energy flux density. Thermal losses in a temperature-sensitive film on an insulating substrate were studied in [14]. The characteristic times of thermal losses were estimated by the following formulas: for radial losses, $\tau_a = r^2/4a$; for heat removal into the substrate, $\tau_s = d^2/4a_1^2$. Here, r is the radius of the thermal spot, a is the thermal diffusivity of the temperature-sensitive layer, d is the thickness of this layer, and a_1 is the thermal diffusivity of the substrate. The method used in [14] was applied to our VO$_2$/I and Al/VO$_2$/I systems to calculate the associated values of τ_a and τ_s. Specifically, they were calculated under adiabatic recording conditions for a thermal spot of radius 10^{-3} m on the VO$_2$/Al$_2$O$_3$ system (the respective thicknesses of the layers are 100 and 314 nm) applied on Polikor and mica substrates. The respective values are $\tau_a = 7.8 \times 10^{-2}$ s, $\tau_s = 0.25$ s, and $\tau_s = 1.5 \times 10^{-6}$ s. For the Al(100 nm)/VO$_2$(100 nm)/I(314 nm) system, $\tau_a = 7.8 \times 10^{-2}$ s and $\tau_s = 3.3 \times 10^{-7}$ s under the same recording conditions. The time of heat removal into the substrate specifies the time of data erasing. The recording and erasing times obtained are in good agreement with experimental data [15] found upon recording optical information on Al/VO$_2$ media under adiabatic exposure. In the thermal physical model considered, the effect of the substrate and protective layer on the optical sensitivity of the medium shows up in thermal losses due to contact heat transfer. Thus, when being absorbed, the energy of a laser shot forms a temperature profile depending on the thermal, physical, and optical properties of the thin films and substrate and also on the energy exposure.

In practical use of Al/VO$_2$/I and VO$_2$/I systems, one must be aware of activation processes taking place under the local or general heating of the surface. The former medium is used in optical visualizers and holographic transparencies, where the surface is heated locally. A metallic mirror enhances the contrast between the two (semiconductor and metal) phases. The latter medium is employed as temperature-sensitive layers of thermal detectors, where the entire surface is heated. Experiments have shown that the phase transitions thermally induced in Al/VO$_2$/I and VO$_2$/I systems provide high-efficiency optical data recording. Under adiabatic exposure, when recording is accom-

plished with a laser (the duration of a laser shot is no more than 4×10^{-8} s and the minimal area of laser action is limited by the wavelength), the threshold sensitivity of the medium is the least [16]. At longer shot durations, radiation losses and radial heat removal increase, which affects the threshold sensitivity, as well as the times of response and relaxation. Thus, when recording optical data on these media, one should take into account the activation processes, which depend on the thermophysical properties of the films and substrate, as well as on the energy exposure of the radiation source.

Essentially, the dynamic range of the energy exposure sensitivity of the VO_2 films is defined by the hysteresis width. $Al/VO_2/I$ and VO_2/I systems are used in RAMs, so that they are thermostatically controlled. The temperature at which the systems are maintained is selected inside the loop (as a rule, near the quasi-linear segment of the optical or electrical hysteresis).

Of interest are VO_2/I-based thermal detectors of optical radiation operating in the writing/erasing mode when the shot duration is 1 s (the boundary between pulsed and continuous radiation) and its power is low. The detector with a 100-nm-thick VO_2 film and a sensitive area of 2×2 mm was made on a 40-μm-thick mica substrate with aluminum electrodes. The response time of the detector was estimated by the technique described in [17]. According to this technique, the transmission of a temperature wave through the temperature-sensitive layer is assumed to be instantaneous (since the thickness of the layer is much smaller than that of the substrate and the thermal diffusivity of the layer is much higher than that of the substrate) and the wavelength of the temperature wave is found from the relationship $1/\lambda = \sqrt{\dfrac{\omega}{2a}}$, where λ is the wavelength of the temperature wave in the substrate, $\omega = 2\pi f$ is the circular frequency, and a is the thermal diffusivity of the substrate. For $f = 1$ and 4 Hz, $\lambda = 178$ and 44 μm, respectively; hence, the response time of the detector is ≈0.25 s.

The detector was tested with the working standard no. 30 of pulsed laser radiation (wavelength 0.5 μm) [18]. The measurement scheme is depicted in Fig. 3. The working standard no. 30 consists of an LG-106-1M stabilized gas laser, optical system, mechanical shutter (modulator), device measuring the average power, vibration-absorbing table, and recording system. The modulator converts the laser cw radiation with a given average power to a shot of duration 1 s, which is applied to the input of the detector connected in a bridge circuit. The output of the bridge is measured with a Shch68003 digital voltmeter and displayed on an S1-70 oscilloscope.

The conversion efficiency of the detector at wavelengths between 0.3 and 3.39 μm equals 5.55 mV/mW. The radiation power density 1 mW/mm² heats the VO_2
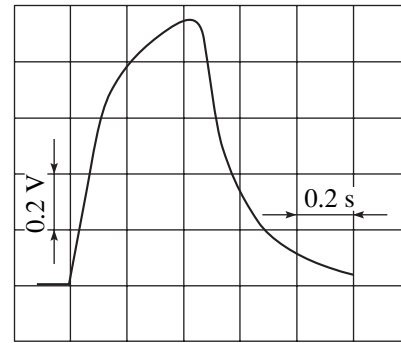

Fig. 4. Detector output waveform.

layer by 1 K. The range where the sheet resistivity of the 100-nm-thick VO_2 layer varies with temperature quasi-linearly is 13 K. For a power density of 13.2 mW/mm² and a shot duration of 1 s, the detector output is maximal in the dynamic range. Figure 4 shows the detector output waveform. At the half-peak level, the leading edge of the pulse has a tilt ≈0.2 s. The experimental data for the speed of response of the detector are in good agreement with the analysis.

Thus, $Al/VO_2/I$ and VO_2/I film media seem to be promising for optical data writing, storage, and reading systems. The writing speed and the resolution of the media meet holography requirements. When writing optical data, one must select an appropriate film system and radiation source for each application.

REFERENCES

1. A. S. Oleĭnik, V. F. Smolyakov, V. M. Stepanov, and N. M. Rudenko, Élektron. Prom-st', Nos. 5–6, 111 (1982).

2. A. S. Oleĭnik, Zh. Tekh. Fiz. **63** (1), 97 (1993).

3. A. S. Oleĭnik, in *Proceedings of the International Conference "Urgent Problems in Electronic Instrument Making"* (Saratov. Gos. Tekh. Univ., Saratov, 1998), p. 69.

4. M. P. Bogoroditskiĭ and V. V. Pasynkov, *Materials in Radioelectronics* (Gosénergoizdat, Moscow, 1961).

5. A. S. Oleĭnik, Neorg. Mater. **27**, 534 (1991).

6. J. W. Christian, *Theory of Transformations in Metals and Alloys* (Pergamon, Oxford, 1975; Mir, Moscow, 1978), Part 1.

7. Yu. M. Gerbshteĭn, T. V. Smirnova, E. I. Terukov, and F. A. Chudnovskiĭ, Fiz. Tverd. Tela (Leningrad) **18**, 503 (1976) [Sov. Phys. Solid State **18**, 290 (1976)].

8. *Physical Encyclopedia* (Sov. Éntsiklopediya, Moscow, 1994), Vol. 4, pp. 72–74.

9. E. I. Terukov, K.-D. Ufert, and F. A. Chudnovskiĭ, Fiz. Tverd. Tela (Leningrad) **18**, 2479 (1976) [Sov. Phys. Solid State **18**, 1450 (1976)].

10. A. V. Lykov, *Theory of Heat Conduction* (Vysshaya Shkola, Moscow, 1967).

11. N. N. Rykalin, A. A. Uglov, I. V. Zuev, and A. N. Kokora, *Laser and Electron-Beam Treatment of Materials: A Handbook* (Mashinostroenie, Moscow, 1985).

12. V. G. Mokerov and V. V. Saraĭkin, Fiz. Tverd. Tela (Leningrad) **18**, 1801 (1976) [Sov. Phys. Solid State **18**, 1049 (1976)].

13. *Physicochemical Properties of Oxides: A Handbook*, Ed. by G. V. Samsonov (Metallurgiya, Moscow, 1978).

14. A. B. Granovskiĭ, G. I. Bukman, B. M. Stepanov, and E. B. Shelemin, in *Thin Magnetic Films for Recording Optical Information: Scientific Works of All-Union Research Institute of Physicotechnical and Radio Engineering Measurements* (Moscow, 1979), pp. 44–53.

15. A. A. Bugaev, B. P. Zakharchenya, and F. A. Chudnovskiĭ, *Metal–Semiconductor Phase Transition and Its Application* (Nauka, Leningrad, 1979).

16. D. Dayton and D. Eden, Opt. Eng. **20**, 377 (1981).

17. O. V. Roĭtsina, Zh. Prikl. Spektrosk. **111**, 403 (1965).

18. *Measurement Assurance of Energy Photometry: A Handbook*, Ed. by B. M. Stepanov (Atomizdat, Moscow, 1979).

Translated by V. Isaakyan

OPTICS,
QUANTUM ELECTRONICS

Formation of the Spectra of Multiply Charged Ions from the Plasma of Nd-Doped Glass Irradiated with a Neodymium Laser

M. R. Bedilov, Kh. B. Beĭsembaeva, and I. Yu. Davletov

Research Institute of Applied Physics, Ulugbek National University of Uzbekistan, Tashkent, 700174 Uzbekistan
e-mail: bedilov@iaph.post.tps.uz
Received January 18, 2000; in final form, February 11, 2002

Abstract—The effect of neodymium ions on laser-induced glass destruction and the formation of the spectra of laser plasma ions is studied by the mass-spectrometric and optical-microscopy methods. As the Nd ion concentration increases, the degree of laser-induced damage in silicate glass increases, the threshold intensity for the glass destruction decreases, the maximum ion charge number Z_{max} and the maximum ion energy E_{max} decrease, and the energy spectra of target ions change. All this is related to both the coincidence between the lasing lines and the absorption bands of Nd ions and the intensification of recombination processes in the multicomponent laser plasma. © 2002 MAIK "Nauka/Interperiodica".

In recent years, physical processes related to the interaction of laser radiation with matter has attracted growing interest. Under irradiation of a solid with laser radiation with an intensity of $q \geq 0.1$ GW/cm^2, a plasma plume is produced, which acts as a source of broadband radiation, electrons, multiply charged ions, and neutrons. The processes of plasma formation under intense laser irradiation of solids is closely related to the destruction of these solids. Hence, various static and dynamic defects of the target can substantially affect the process of destruction and, finally, the spectra of laser plasma ions. Experimental studies of the physical processes (ionization, recombination, etc.) in multicomponent laser plasmas at $q = 0.1–100$ GW/cm^2 [1–6] revealed specific effects related to the formation and expansion of some heavy ion species doped in the target. It is found that the charge number, energy, and fluxes of plasma ions depend not only on q but also on the target composition. Thus, it was found that, at low concentrations of gold [1] and silver [2] ($\sim 10^{-2}–10^{-3}$%) in a multicomponent target, the concentrations of the single-charged Au and Ag ions in a multicomponent plasma are significantly lower than those in a monocomponent plasma. It should be noted that studies of the effect of small additives in the target on the formation of the spectra of laser plasma ions have begun only recently.

In this study, the effect of Nd ions on the formation of the mass, charge, and energy spectra of multiply charged ions in a laser plasma is investigated. The target was made of GLS silicate laser glass with and without Nd doping. The Nd concentration in glass was varied in the range $C = 0–5.6 \times 10^{20}$ cm^{-3}. The targets were 1.0-cm-diameter ~0.5-cm-thick pellets with a known

Nd concentration. The laser pulses were directed onto different places of the target, which allowed us to acquire the data on the destruction of the sample surface. The experimental data were acquired under identical initial conditions: the target chamber was loaded with more than ten samples with and without Nd doping. The measurements were carried out with a time-of-flight mass-spectrometer and an electrostatic analyzer [3]. In experiments, we used a neodymium laser with a wavelength of $\lambda = 1.06 \times 10^{-4}$ cm, pulse duration of 50 ns, and pulse energy of 3.0 J. The angle of incidence of laser radiation with respect to the normal to the target surface was 18° and the radiation intensity on the target surface was varied in the range 1–100 GW/cm^2 with the help of neutral light filters. To determine the threshold intensity for the destruction of glass under study, the laser radiation intensity was gradually increased from pulse to pulse until destruction occurred, which was accompanied by a glow and plasma generation. From the recorded ion signals, we could determine not only the threshold intensity for the target destruction, but also the composition of the absorbing additives, which played a significant role in the glass destruction and plasma production. The development of damages on the glass surface was monitored by using an optical microscope.

We measured the threshold intensity for the laser destruction of glass and the formation of the mass, charge, and energy spectra of plasma ions within the Nd concentration range $C = 0–5.6 \times 10^{20}$ cm^{-3} for laser intensities of $q = 0.1–100$ GW/cm^2. The threshold intensity for the laser destruction of glass and the formation of the spectra of multiply charged ions in laser
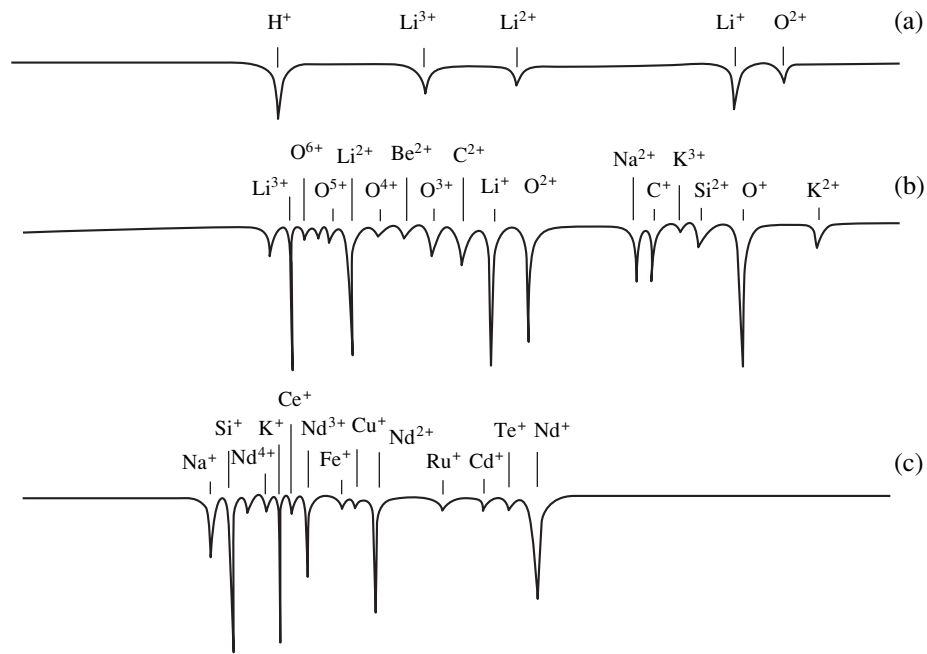
Fig. 1. Typical mass and charge spectra of the ions from an Nd-doped glass for the laser intensity $q = 90$ GW/cm^2 at $E/Z =$ (a) 50, (b) 100, and (c) 400 eV.

plasma was determined as a function of the Nd concentration in glass and the laser intensity q.

Figure 1 presents typical mass and charge spectra of multiply charged ions in laser plasma. The spectra were obtained at $q = 90$ GW/cm^2 for a glass with the Nd concentration $C = 1.9 \times 10^{20}$ cm^{-3}. An analysis of the mass and charge spectra of laser plasma ions shows that the element and charge compositions of the plasma depend not only on the laser intensity q, but also on the Nd concentration in glass. At $q = 9$ GW/cm^2, the mass and charge spectra of laser plasma ions mainly consist of the ion signals from the constituents of the glass matrix and some impurities (Be, B, and C). As the laser intensity q increases, peaks from other ionized atoms arise in the mass and charge spectra. At $q = 90$ GW/cm^2, together with the peaks from the ions of glass matrix elements (Li, O, Na, Si, K, and Ca) and Nd dopant, extra peaks from the ions of some impurities (H, Be, B, C, Fe, Cu, and Ce) and Si isotopes are observed. When q is close to the threshold value, the mass and charge spectra of plasma ions contain ion peaks from the matrix elements, as well as from the B and C impurities. This fact indicates that the presence of boron and carbon in glass somewhat facilitates laser destruction of the surface. The microscopic studies show that damages on the surface of an undoped target look like a crater with a fused rim. Inside the crater, there are petty dents with dimensions of several tenths of micron or less. Presumably, optical inhomogeneities and impurity inclusions in a sample, whose presence increases the absorption of laser radiation in local sites, are responsible for these microcraters. Under our experimental conditions, the threshold intensity for the laser destruction of the GLS glass surface is ~7 GW/cm^2. At this intensity, the crater diameter is equal to $d \sim 5.0 \times 10^{-3}$ cm. The crater diameter increases with laser intensity and, at $q = 1000$ GW/cm^2, is ~3.0×10^{-2} cm. The increase in the Nd content in glass from 0 to 5.6×10^{20} cm^{-3} leads to a decrease in the destruction threshold to 2 GW/cm^2, i.e., by a factor of 3.5, whereas the corresponding crater diameter amounts to 1.0×10^{-2} cm. Figure 2 shows the threshold intensity for the laser destruction of the glass surface as a function of the Nd concentration, which was varied in the range $C = 0$–5.6×10^{20} cm^{-3}. It is seen that the threshold for laser destruction depends strongly on the Nd concentration in the target. The increase in the Nd concentration from 0 to 5.6×10^{20} cm^{-3} decreases the threshold intensity by a factor of 3.5. Moreover, in the case of undoped glass, the mass and charge spectra show that the yield of matrix ions (Li$^+$, O$^+$, and Si$^+$) decreases by a factor of 1.5–2. The revealed decrease in the threshold intensity for glass destruction with increasing Nd concentration correlates with the data on the absorption spectra of glasses under study. It was found that, in the case of Nd-doped glass, the background absorption at wavelengths close to 1.06×10^{-4} cm is about seven times higher than that in the case of undoped glass. The increase in the background absorption near the lasing line of neodymium laser seems to be related to various defects (dislocations, inhomogeneities, etc.) occurring in glass doped with Nd. The presence of these defects inside glass and on its surface increases light absorption.

An analysis of the measured charge spectra of the plasma ions produced in the laser–glass interaction at Nd concentrations in the glass in the range $C = 0$–5.6×10^{20} cm^{-3} shows that the plasma charge composition depends strongly on the Nd concentration. Thus, it was shown that the Nd doping of glass significantly changes the charge spectra of plasma ions because of the redistribution of the charge number Z among the ions of target atoms. Typical mass and charge spectra of the plasma ions from an undoped glass irradiated with laser at $q = 9$–90 GW/cm^2 and $E/Z = 400$ eV are shown in Fig. 3. It can be seen that the charge number Z of the target ions depends not only on the laser intensity q, but also on the element species (specifically, neodymium) doped in the glass under study. For example, at $q = 9$ GW/cm^2 and different E/Z values, there are ions of oxygen and silicon with $Z_{max} = 3$ and ions of K, Ca, and Na with $Z_{max} = 1.2$. As the laser intensity q increases, Z_{max} of the plasma ions increases and the signals from the ions of H, Al, Li, and Si isotopes, as well as the ions of other impurities, appear in the spectra. At $q = 100$ GW/cm^2, we recorded the following ions of the elements present in an undoped glass: H$^+$, Li$^+$–Li^{3+}, O$^+$–O^{6+}, Na$^+$–Na^{5+}, Al$^+$, Al^{2+}, Si$^+$–Si^{4+}, K$^+$–K^{4+}, Ca$^+$–Ca^{2+}, Be$^+$, B$^+$, C$^+$, Fe$^+$, and Cu$^+$.

It is established that, in the charge spectra of the plasma ions produced when irradiating an Nd-doped glass, the charge number Z of the plasma ions is redistributed. For example, at $q = 100$ GW/cm^2, Z_{max} for oxygen and sodium ions decreases to 5 and 4, respectively, whereas for neodymium ions we have $Z_{max} = 4$. As the laser intensity q decreases, the charge number of neodymium ions also decreases. The decrease in Z_{max} observed for the ions of some constituent elements of an Nd-doped glass stems from the fact that the presence of heavy neodymium ions in the plasma significantly slows down recombination processes, because Nd doping substantially increases the mass of the ionized plume, which significantly decreases the speed at which it expands.

The study of the ion energy spectra from glass with the Nd concentration $C = 0$–5.6×10^{20} cm^{-3} at laser intensities of $q = 1$–100 GW/cm^2 revealed the following features:

(i) The energy spectra of the laser-produced ions of glass constituents and doped elements are broad; their width and intensity, as well as the location of the spectrum maximum, depend on the concentration of Nd ions in the plasma.

(ii) The energies of the ions of light elements are lower than the energy of heavy neodymium ions.

(iii) At the plasma expansion stage, the energy spectra of the ions from an Nd-doped glass change greatly because of the interplay between the ions of light and heavy elements.
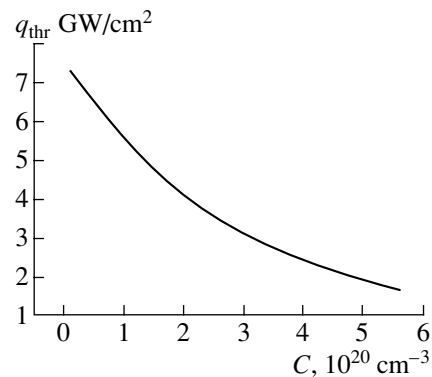
Fig. 2. Threshold intensity q_{thr} for the laser destruction of glass vs. Nd concentration in the target.

(iv) The presence of the ions of heavy elements (Nd) in the expanding plasma decreases both the maximum and minimum energies (E_{max} and E_{min}) of the ions of the lighter glass constituents, although E_{max} and E_{min} of the ions of heavy elements (Nd) increase substantially.

The energy spectra of Si ions produced under laser irradiation of Nd-doped and undoped glasses at laser intensities of $q = 100$ GW/cm^2 are shown in Fig. 4. It can be seen that the character of the energy distribution does not depend on whether the glass has been doped or not; in both cases, the ion energy spectra are broad and have one maximum. The higher the ion charge number, the larger the shift of the maxima toward higher energies. However, in the case of an Nd-doped glass, the increase in q decreases the width of the energy spectra of Si ions with a given charge number, shifts the spectrum maxima toward lower energies, and decreases the E_{max} and E_{min} values. The higher the value of q, the larger the change in the energy spectra. Similar changes in the energy spectra are observed for ions of the other glass constituents (Li, O, Na, Al, K, and Cu). The influence of Nd ions on the formation of the spectra of the glass constituents (Li, O, Na, Si, and K) is explicitly seen when comparing the spectra from Nd-doped and undoped glasses. At $q = 9$ GW/cm^2, in either glass, the energy spectra of all these ions have one maximum. In an undoped glass, the maxima in the energy spectra of the above ions occur at energies of 155 ± 20 eV, whereas in an Nd-doped glass, these energies decrease (approximately by a factor of 2) and amount to 80 ± 10 eV. This shift in the maxima of the energy spectrum toward lower energies is accompanied by a decrease in the maximum and minimum ion energies, E_{max} and E_{min}. Figure 5 presents the energy spectra of singly charged ions obtained by irradiating Nd-doped and undoped glasses with laser radiation with an intensity of $q = 90$ GW/cm^2. It can be seen that Nd doping changes the energy spectra of Li$^+$, O$^+$, Na$^+$, Si$^+$, and K$^+$ ions. Thus, for Li$^+$, O$^+$, and Si$^+$ ions, the maxima in the spectra increase by a factor of about 1.5–2, whereas for
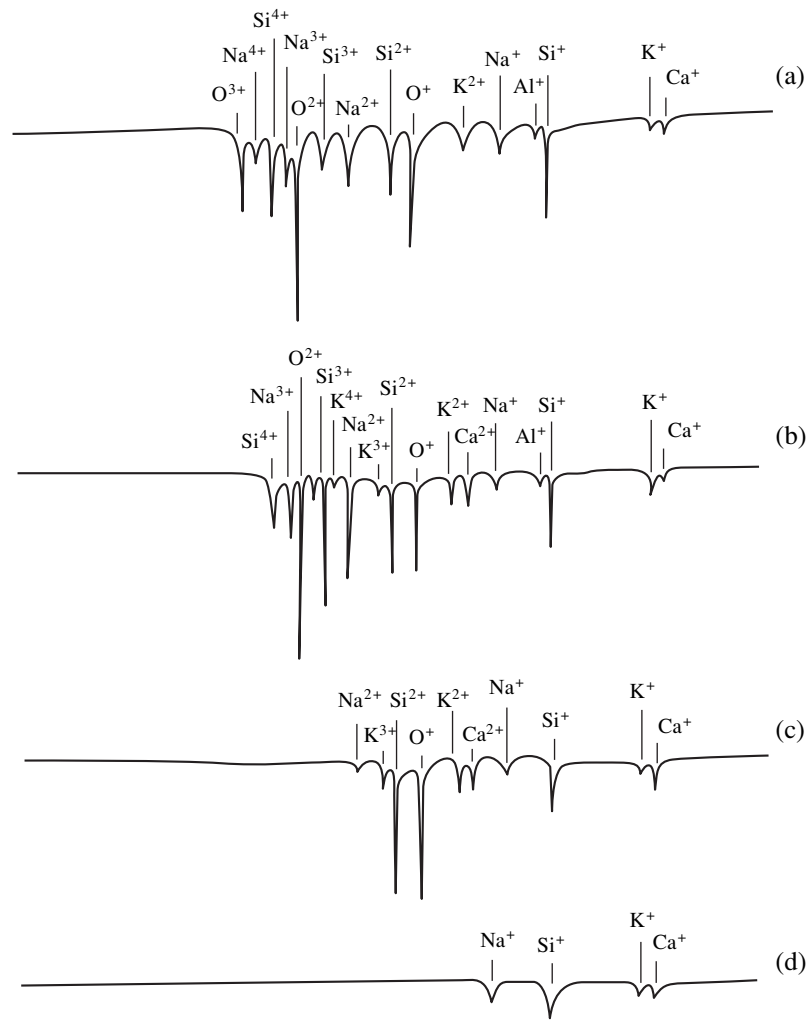
Fig. 3. Typical mass and charge spectra of the ions from undoped glass for q = (a) 90, (b) 60, (c) 20, and (d) 9 GW/cm^2 at E/Z = 400 eV.
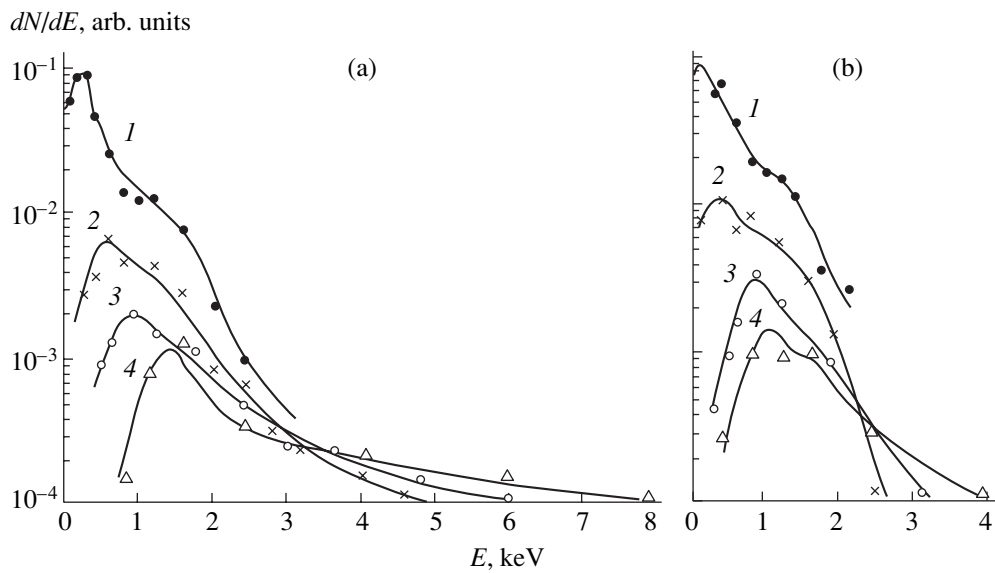

Fig. 4. Energy spectra of silicon ions with charge numbers Z = (*1*) 1, (*2*) 2, (*3*) 3, and (*4*) 4 produced under laser irradiation of (a) undoped and (b) Nd-doped glasses at q = 90 GW/cm^2.
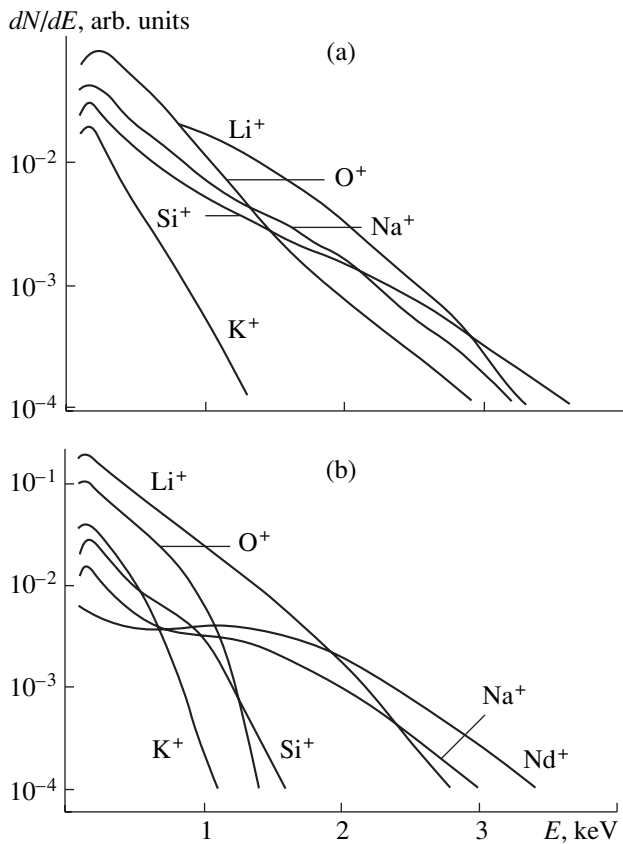
dN/dE, arb. units

Fig. 5. Energy spectra of the Li$^+$, O$^+$, Na$^+$, Si$^+$, and K$^+$ single-charged ions from (a) undoped glass and (b) Nd-doped glass at $q = 90$ GW/cm^2.
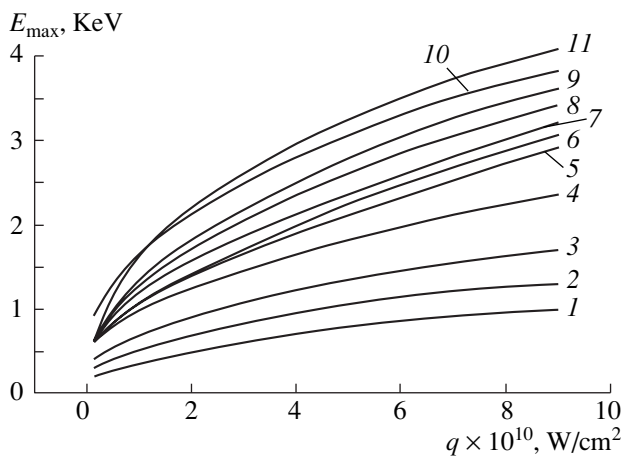
E_{max}, KeV

Fig. 6. Maximum energy E_{max} of (*1, 6*) Li$^+$, (*2, 7*) O$^+$, (*3, 8*) Na$^+$, (*4, 9*) Si$^+$, (*5, 10*) K$^+$, and (*11*) Nd$^+$ ions from (*1–5, 11*) Nd-doped and (*6–10*) undoped glasses vs. laser intensity q.

Na$^+$ and K$^+$ ions the maxima decrease by nearly the same factor. Figure 6 presents the dependences of E_{max} of the singly charged ions of the glass matrix elements from Nd-doped (curves *1–5*) and undoped glasses (curves *6–10*) on the laser intensity q. It is seen that Nd doping reduces E_{max} of the ions. Note that during the expansion of the plasma that contains ions of different masses, the ions of heavy additive (Nd) undergo strong acceleration (Fig. 6, curve *11*); the higher the value of q, the stronger the acceleration. Apparently, the strong acceleration of the ions of heavy elements stems from the increasing role of collision processes occurring in a multicomponent plasma. As a result, the ions of light elements transfer their energy to the heavy ions. The increase in the laser intensity q intensifies the production of the multiply charged ions, which, in turn, stimulates collision processes among the ions. As a result, the ions of heavy elements (particularly, Nd) are accelerated even more strongly.

Thus, we have found that Nd ions affect the threshold intensity for laser-induced silicate glass destruction, the degree to which the glass is destroyed, and the spectra of multicomponent laser plasma ions. It is revealed that, as the Nd ion concentration increases, the degree of laser-induced destruction of silicate glass increases, the threshold intensity for destruction decreases, the maximum ion charge number Z_{max} and the maximum ion energy E_{max} decrease, and the energy spectra of target ions change. An analysis of experimental data obtained with Nd-doped and undoped glasses shows that the observed effects are related to the presence of the absorption bands that are characteristic of Nd ions and that possibly overlap with the lasing lines of the neodymium laser.

REFERENCES

1. M. R. Bedilov, Yu. A. Bykovskiĭ, and D. Kuramatov, Kvantovaya Élektron. (Moscow) **14**, 1675 (1987).
2. M. R. Bedilov, Yu. A. Bykovskiĭ, and D. Kuramatov, Kvantovaya Élektron. (Moscow) **21**, 79 (1991).
3. M. R. Bedilov and K. Khaitbaev, Prib. Tekh. Éksp., No. 6, 139 (1996).
4. Yu. A. Bykovskiĭ and V. N. Nevolin, *Laser Mass-Spectrometry* (Énergoatomizdat, Moscow, 1985).
5. M. R. Bedilov, Kh. B. Beĭsembaeva, and M. S. Sabitov, Kvantovaya Élektron. (Moscow) **30**, 48 (2000).
6. N. G. Basov, Yu. A. Zakharenkov, A. A. Rupasov, *et al.*, *Diagnostics of Dense Plasma* (Nauka, Moscow, 1989).

Translated by N. Ustinovskiĭ

ACOUSTICS,
ACOUSTOELECTRONICS

Synthesis of Ladder-Type Filters Based on SAW Resonators

V. F. Dmitriev

Avangard-Élionika Close Joint-Stock Company, St. Petersburg, 195271 Russia
e-mail: elionica@cityline.spb.ru
Received September 19, 2001; in final form, January 9, 2002

Abstract—A synthesis method for ladder-type filters based on SAW resonators is described. SAW devices are designed with modified equations for coupled waves. Nomograms for choosing the geometry of the filters on *ST*, *X*-cut quartz and on 36° *Y*, *X*-cut LiTaO$_3$ substrates are presented. Theoretical results are compared with experimental data. © 2002 MAIK "Nauka/Interperiodica".

INTRODUCTION

In recent decades, designers of SAW filters have concentrated on low-loss band-pass filters. One of the best designs is ladder-type filters on SAW resonators (SAW-R filters) [1]. They not only offer inherently high band-pass performance, but also extend the high-frequency range of SAW devices.

Ladder-type filters with lumped inductances and capacitances (*LC* elements) have proved to be efficient at comparatively low frequencies [2]. Since the *Q* factor of *LC* elements is low, ladder-type filters use quartz bulk-acoustic-wave resonators at high frequencies (in the megahertz range) [3]. However, such filters feature good band-pass performance only up to several tens of megahertz. At frequencies above 100 MHz, SAW resonators are the most appropriate. A number of manufacturers have recently developed and produced SAW-R filters operating at up to 2.4 GHz, while some pilot devices have reached operating frequencies as high as 5 GHz [4].

At the same time, although the number of publications on the problem is large (see [1] and refs. therein), design methods for these filters have so far been inadequately studied. The design of SAW-R ladder-type filters is somewhat different from that of standard SAW filters. To take advantage of the properties of the former (low insertion loss, flat frequency response in the pass band, and high off-band rejection), the geometry of the resonators must be chosen very carefully.

This paper describes a method for synthesizing SAW-R ladder-type filters and provides nomograms for choosing the geometry of these filters fabricated on *ST*, *X*-cut quartz substrates and on 36° *Y*, *X*-cut LiTaO$_3$ substrates. Modified coupled-mode equations (COM equations) for SAW devices are proposed, and experimental frequency responses of SAW-R ladder-type filters synthesized by this method are provided. The theo-retical frequency responses are compared with those found experimentally.

1. SYNTHESIS OF SAW RESONATORS

A ladder-type filter uses one-port resonators, each having particular static and dynamic parameters. The equivalent circuit of a resonator with an input admittance *Y* near the resonance frequency (Fig. 1) includes a static capacitance C_0, dynamic capacitance C_D, dynamic inductance L_D, and dynamic (active) resistance R_D. The equivalent circuit in Fig. 1 describes both SAW-based and *LC* resonators and thus allows us to pass from the parameters of an *LC* resonator to those of an SAW resonator.

A building block of a ladder-type filter is the Γ-connected resonators shown in Fig. 2, where Y_P and Y_S are the admittances of the parallel and series arms, respectively. More complex filters are constructed by series connecting the building blocks (sections) in Figs. 2a and 2b in any combination.

The parameters of the resonators inserted in the series and parallel arms of the section must be selected so that its input and output reactances be close to zero in the filter's passband. The active component R_D of the impedance must be as small as possible, because it specifies the *Q* factor of the resonator and insertion loss of the filter.
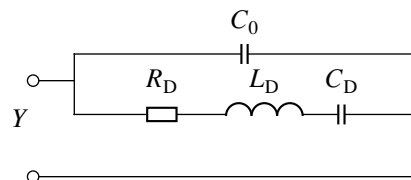

Fig. 1. Equivalent circuit of an SAW resonator.

The theory of ladder-type filters on LC resonators is well developed. Given filter parameters, one can find the inductances and capacitances, as well as the number of sections in the filter (see, e.g., [2]). Let us know center frequency F_0; frequency band $\Delta F_C = F_{2C} - F_{1C}$, where F_{2C} and F_{1C} are the upper and lower cutoff frequencies; and bandwidth ratio $K_B = (F_{2\infty} - F_{1\infty})/\Delta F_C$, where $F_{2\infty}$ and $F_{1\infty}$ are, respectively, the upper and lower frequencies at which the insertion loss is infinite. Then, from the theory of ladder-type LC filters, we can find the equivalent parameters of an LC prototype for the parallel section (C_{0P}, C_{DP}, and L_{DP}) and series section (C_{0S}, C_{DS}, and L_{DS}). Figures 3a and 4a show theoretical nomograms obtained with data in [2] that relate C_D/C_0 to ΔF_C for several values of K_B.

Synthesis of an SAW-R ladder-type filter thus consists in establishing a relationship between the geometrical parameters of the SAW resonator and the equivalent parameters of the LC prototype.

Using the known expression for the static capacitance of an interdigital transducer and for the dynamic capacitance of an SAW resonator near the resonance frequency [3], we can write the ratio between the dynamic and static SAW resonator capacitances as

$$(C_D/C_0)_{P0} = 8k^2(h_1)Nr_0(h_2)/[\pi^2(Nr_0(h_2) + 1)], \quad (1)$$

where $k^2(h_1)$ is the electromechanical coupling factor squared, h_1 is the electrode thickness, N is the total number of electrodes in the transducer, $r_0(h_2)$ is the absolute value of the reflection coefficient from an inhomogeneity, and h_2 is the height of the inhomogeneity in the reflecting structure.

Note that an inhomogeneity in the reflecting structure can have the form of metal strips (in this case, $h_1 = h_2$) or grooves (protrusions) etched on the piezoelectric surface. The coupling factor and the reflection coefficient depend on the thickness of interdigital transducer electrodes and on the height of the inhomogeneity. The dependences $k^2(h_1)$ and $r_0(h_2)$ are usually empirically approximated as

$$k^2(h_1) = k^2(h_1 = 0) + \zeta_1(h_1/\lambda) + \zeta_2(h_1/\lambda)^2 + \ldots, \quad (2)$$

$$r_0(h_2) = r_0(h_2 = 0) + r_{10}(h_2/\lambda) + r_{20}(h_2/\lambda)^2 + \ldots, \quad (3)$$

where $k^2(h_1 = 0)$ and $r_0(h_2 = 0)$ are the values of the respective parameters at the zero thicknesses of the metal film h_1 and the discontinuity h_2; λ is the SAW wavelength; and ζ_1, ζ_2, r_{10}, and r_{20} are empirical coefficients.

Our method for synthesizing an SAW resonator for a ladder-type filter relies on formulas (1)–(3). It is necessary to find the geometrical parameters of the resonator elements from empirical [5] or theoretical dependences $k^2(h_1)$ and $r_0(h_2)$ for a particular piezoelectric material. In certain cases, Eq. (1) in view of relation-
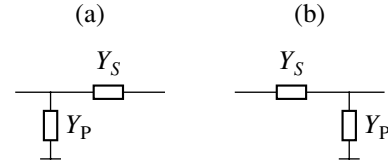

Fig. 2. Elementary section of a ladder-type filter: (a) type-*a* section and (b) type-*b* section.

ships (2) and (3) can be solved analytically. For the sake of generality, this equation should however be solved numerically with $h_1 = 0$ and $h_2 = 0$ as the initial approximation. For a particular piezoelectric material, Eq. (1) does not necessarily have a solution for arbitrary $\Delta F_C/F_0$ and K_B. Therefore, for Eq. (1) to have a solution for filters with both narrow and wide pass bands $\Delta F_C/F_0$, it is necessary to use piezoelectric materials with low and high k^2, respectively.

In a ladder-type filter, the resonance frequency of the series-arm resonator is somewhat lower than that of the parallel-arm resonator. The values of h_1 and h_2 given by the solution to Eq. (1) are also lower for the series-arm resonator. Therefore, one should first solve Eq. (1) for the parallel-arm resonator and then use the resulting value of h to calculate the parameters of the series-arm resonator. Before solving Eq. (1), it is necessary to specify an initial approximation. The number of electrodes N will subsequently be refined by a more rigorous analytical method (the COM technique).

As follows from the theory of ladder-type filters, the resonance frequency of the parallel resonator must correspond to the frequency $F_{1\infty}$, while the antiresonance frequency of the series resonator must correspond to the frequency $F_{2\infty}$. Also, the antiresonance frequency of the parallel resonator must equal the resonance frequency of the series one. This requirement specifies the period of interdigital transducers used in the filter.

Let the filter employ a reflecting structure consisting of metal strips and a solution to Eq. (1) yield the thickness h_1 for the parallel resonator and the corresponding number N_{P0} of electrodes. Then, the number of electrodes in the series resonator is given by the expression following from (1):

$$N_{S0} = \pi^2(C_D/C_0)_{S0}/[8k^2(h_1)]/\{r_0(h_1) \\ \times [1 - \pi^2(C_D/C_0)_{S0}]/[8k^2(h_1)]\}. \quad (4)$$

With the numbers N_{S0} and N_{P0} of electrodes for both resonators determined and the piezoelectric material chosen, the apertures of the interdigital transducers can be calculated as

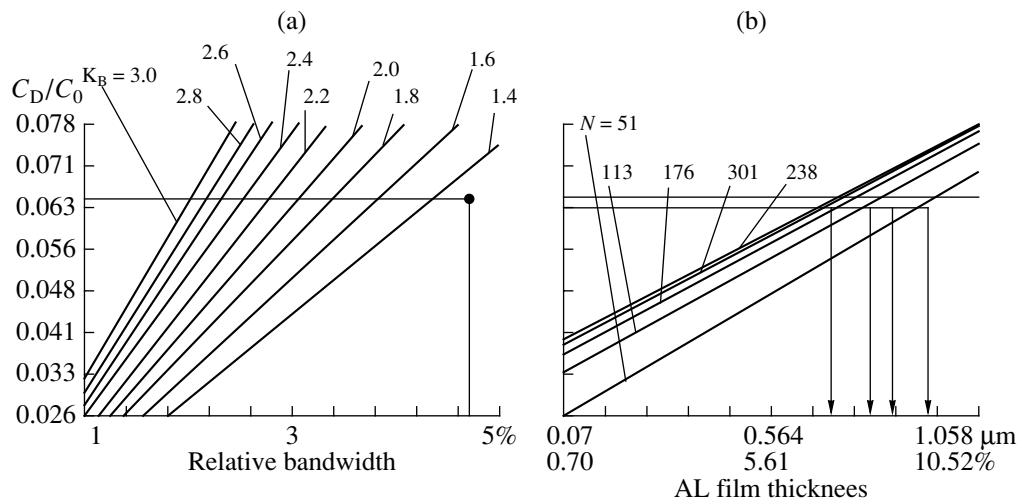$$W_P = C_0/(2C_2N_{P0}), \quad (5)$$

$$W_S = C_0/(2C_2N_{S0}), \quad (6)$$

(a) (b)

Fig. 3. Nomogram for determining the parameters of the resonator on a 36° Y, X-cut LiTaO$_3$ piezoelectric.
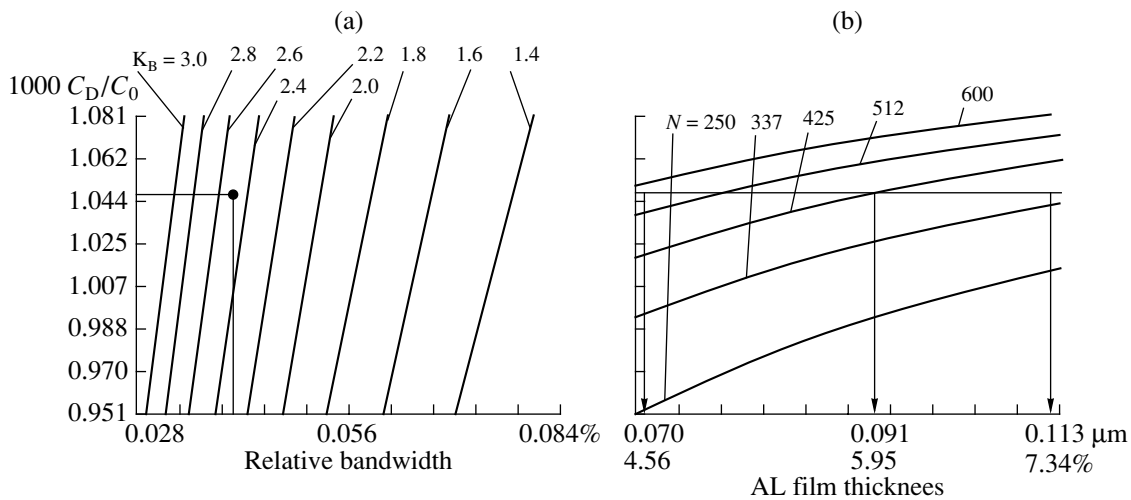
(a) (b)

Fig. 4. Nomogram for determining the parameters of the resonator on a ST, X-cut quartz piezoelectric substrate. $\Delta F_C/F_0 = 0.042\%$ and $K_B = 2.5$ is illustrated.

where C_2 is the capacitance of a pair of electrodes per unit length.

Figures 3b and 4b show solutions to Eq. (1) in the form of nomograms for two conventional piezoelectrics: 36° Y, X-cut LiTaO$_3$ (Fig. 3) and ST, X-cut quartz (Fig. 4). These nomograms help in selecting the geometry of resonators incorporated into ladder-type filters. The curves in Figs. 3b and 4b plot the resonator's dynamic-to-static capacitance ratio versus thickness of the metal (in this case, aluminum) electrodes in the transducer and reflecting structure with the number of transducer electrodes used as a parameter.

The maximum possible width of the relative passband for a ladder-type filter is defined by the maximum value of $k^2(h)$ of a piezoelectric material. For 36° Y, X-cut LiTaO$_3$ substrates, the maximum feasible value, $k^2(h) \approx 12\%$ [5], is limited by a rapid increase in losses

due to bulk wave excitation, starting with a thickness of the aluminum film for which $h/\lambda \approx 8\%$ [1, 6]. The minimum relative passband width of a ladder-type filter depends on the maximum Q factor of the piezoelectric resonator. For 36° Y, X-cut LiTaO$_3$, the feasible passband widths are approximately 1 to 5%. The terms "higher boundary" and "lower boundary" are a matter of convention, because a ladder-type filter can be designed with a narrower or wider passband but at a sacrifice in the filter performance as described in Section 4.

2. SELECTION OF SAW RESONATOR PARAMETERS

Once the filter's relative passband width and bandwidth ratio have been determined, it is necessary to find

the dynamic-to-static capacitance ratio for each of the resonators. One can solve this problem similarly to that for conventional *LC* ladder-type filters [2] or use nomograms (Figs. 3a, 4a). Note that, when the number of stages and the passband width are given, the bandwidth ratio defines the off-band rejection.

The nomograms should be used as follows. From given $\Delta F_C/F_0$ and K_B, one determines the ratio C_D/C_0 for the parallel resonator (Fig. 3a or 4a). The values chosen must be within the area of the nomogram. Otherwise, the SAW equivalent of the filter is unfeasible or, at least, the filter response may be inadequate. The feasibility of a filter with parameters outside the area of the nomogram will be discussed later.

Now, we find the thickness of the aluminum film as the abscissa of a point of intersection between the line of particular C_D/C_0 and a line which refers to a total number of electrodes in the SAW structure (Figs. 3b, 4b). As an example, Fig. 3 illustrates the choice of the parameters for a filter with a center frequency of 392 MHz, bandwidth $\Delta F_C = 18.5$ MHz ($\Delta F_C/F_0 = 4.7\%$), and bandwidth ratio of 1.3. Feasible numbers of electrodes in the interdigital transducer of the parallel resonator found from Fig. 3a for $C_D/C_0 = 0.0643$ are plotted in Fig. 5 versus aluminum film thickness calculated from Eq. (1) or found from Fig. 3b.

The hatched region in Fig. 5 corresponds to $h > 0.85$ μm ($h/\lambda > 8\%$). The values of h_0, N_{P0}, and W_{P0} found for the parallel resonator and N_{S0} and W_{S0} found for the series resonator do not suffice to design the filter. Moreover, since these values are calculated within the simplified resonator model, which is valid near F_0, they should be refined under the constraints

$$N_{P0}W_{P0} = \text{const}, \qquad (7)$$

$$N_{S0}W_{S0} = \text{const}. \qquad (8)$$

To refine the h, N, and W values and to subsequently choose the remaining parameters of the resonator, one must use a more rigorous model that takes into account the distributed nature of the resonator, for example, the COM method.

3. MODIFIED COM METHOD

The conventional COM theory (see, e.g., [7]), which implies the derivation of inhomogeneous differential equations, greatly complicates the solution of the problem. In the framework of this theory, it is also difficult to include such factors as the variable period of the structure, apodization, and nonuniform distribution of the surface charge over the electrodes. All these factors are easy to take into account by the modified COM method, which deals with an element of the structure (an electrode of the interdigital transducer or the reflecting structure). This method is also more promising for constructing more complex models of the structure. The parameters of the SAW structure as a whole
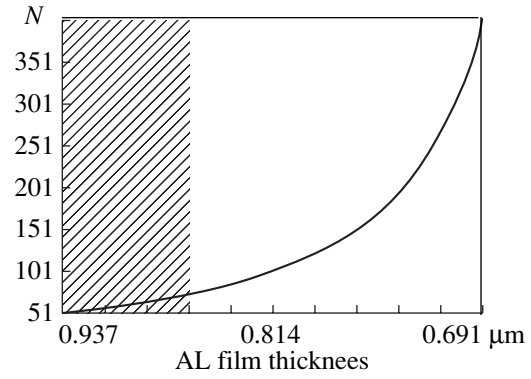

Fig. 5. Number of transducer electrodes in the parallel resonator versus thickness of the aluminum film.

(interdigital transducer, reflecting structure, or their combinations) are found by multiplying the *P* matrices of individual electrodes (as is done in the standard theory of two-port networks in terms of scattering matrices). The time taken to design an SAW structure with as many as several thousands of electrodes with a modern personal computer equals several seconds.

Let an SAW structure have the form of arbitrarily spaced alternating-polarity electrodes that overlap. We also assume that a source of a signal with an amplitude U_0 is placed on the left. Consider the Kth electrode of the interdigital transducer (Fig. 6a). Let $R(x, \omega)$ and $S(x, \omega)$ be coupled inhomogeneous plane waves that propagate in the electrode structure of the transducer in the positive and negative X directions, respectively. Let these waves be expressed as

$$R(x, \omega) = R(\omega)\exp(-jkx), \qquad (9)$$

$$S(x, \omega) = S(\omega)\exp(+jkx), \qquad (10)$$

where $R(\omega)$ and $S(\omega)$ are the complex amplitudes of the corresponding waves and k is the wave number.

Let the wave $R_K(x, \omega)$ impinge on the Kth electrode from the left and the wave $S_{K+1}(x, \omega)$, from the right. Then, with allowance for reflection, transmission, and transformation with a coefficient ξ_K, the complex amplitudes of the transmitted waves can be written as

$$S_K(\omega) = r_K\eta_{1K}\exp[-j(k_{eff}-k_0)p_K]R_K(\omega)$$

$$\times \eta_{1K}(1-|r_K|^2)^{1/2}\exp[-j(k_{eff}-k_0)p_K]S_{K+1}(\omega) \quad (11)$$

$$+ \xi_K\eta_{2K}\exp[-j(k_{eff}-k_0)p_K/2]U_0,$$

$$R_{K+1}(\omega) = \eta_{1K}(1-|r_K|^2)^{1/2}\exp[-j(k_{eff}-k_0)p_K]R_K(\omega)$$

$$+ r_K\eta_{1K}\exp[-j(k_{eff}-k_0)p_K]S_{K+1}(\omega) \quad (12)$$

$$+ \xi_K\eta_{2K}\exp[-j(k_{eff}-k_0)p_K/2]U_0,$$

where r_K is the complex coefficient of reflection from the Kth electrode, k_{eff} is the effective wave number of the SAW, $k_0 = 2\pi/p_K$, $p_K = x_{K+1} - x_K$, ξ_K is the SAW
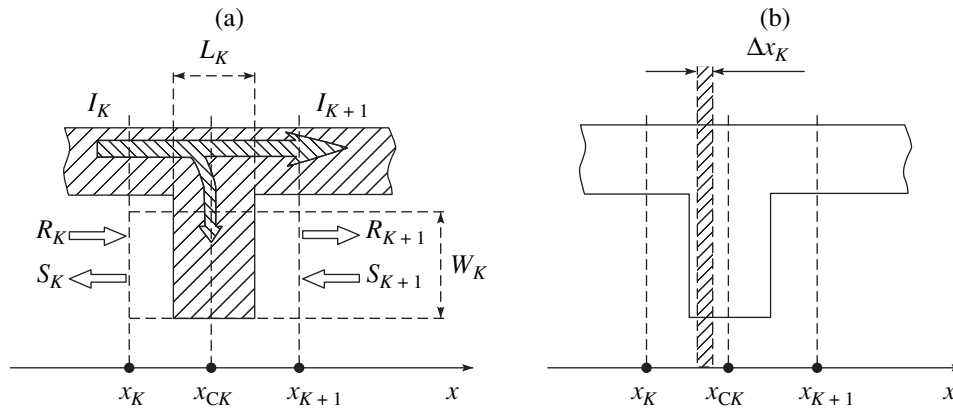
Fig. 6. *K*th element of the interdigital transducer.

transformation coefficient at the *K*th electrode, $\eta_{1K} = W_{1K}/W_0$, $\eta_{2K} = W_{2K}/W_0$, W_0 is the widest aperture, and W_{1K} is the electrode overlap. $W_{2K} = W_0$ when idle electrodes are used; otherwise, $W_{2K} = W_{1K}$.

The phase factors at the terms associated with the wave reflection (transformation) define the phase advance from the center of SAW reflection (transformation) to the corresponding boundary (x_K for $S_K(\omega)$ and x_{K+1} for $R_K(\omega)$). The center of reflection (transformation) is assumed to be at the center of the electrode. Let the effective wave number be expressed as $k_{\mathrm{eff}} = 2\pi/\lambda_{\mathrm{eff}} = \omega/[V_0 + L_K(V_M - V_0)/p_K] - j\alpha$, where V_0 is the SAW velocity on the free surface, V_M is the SAW velocity under the metallized surface, and α is the total propagation loss of the SAW in the electrode structure per unit length.

The current in the interdigital transducer varies because of the transformation of forward and backward waves and a voltage drop across the electrode capacitance:

$$\Delta I_K(\omega) = I_K(\omega) - I_{K+1}(\omega)$$
$$= +2\xi_K \exp[-j(k_{\mathrm{eff}} - k_0)p_K/2]R_K(\omega)$$
$$+ 2\xi_K \exp[-j(k_{\mathrm{eff}} - k_0)p_K/2]S_K(\omega) + j\omega(C_2/2)U_0.$$

Let us consider the terms associated with the SAW transformation as the wave passes through the transducer electrode (Figs. 6a, 6b) and take into account the fact that the excitation is distributed. Assume that efficiencies of direct and inverse SAW transformations at the electrodes are equal, i.e., that the transformation is reciprocal. Let the surface current distribution over the electrodes, $J(x)$, be given and let a small area Δx_K of the electrode's surface current transform the SAW in the same manner as the electrode as a whole. Then, we sum the partial contributions to the SAW transformation over the electrode width and pass to the limit $\Delta X_K \longrightarrow 0$

to obtain

$$\xi_K = G_{\mathrm{a}} \int_{-L_K/2}^{L_K/2} J(x)\exp[-j(\omega/V_M - k_0)x]dx, \quad (13)$$

where G_{a} is the acoustic admittance of the radiation at the synchronism frequency (see, e.g., [7]).

The distribution $J(x)$ of the surface current over the electrodes can be calculated by the technique reported in [8], which allows for edge effects and the finite length of the transducer, as well as disregards the back reaction of a piezoelectric. Formulas (11)–(13) can be written in matrix form:

$$\begin{vmatrix} S_K(\omega) \\ R_{K+1}(\omega) \\ \Delta I_K(\omega) \end{vmatrix} = \begin{vmatrix} P_{11} & P_{12} & P_{13} \\ P_{21} & P_{22} & P_{23} \\ P_{31} & P_{32} & P_{33} \end{vmatrix} \begin{vmatrix} R_K(\omega) \\ S_{K+1}(\omega) \\ U_0 \end{vmatrix}. \quad (14)$$

Then, the *P* matrix of the interdigital transducer as a whole can be calculated by multiplying the partial *P* matrices for each of the electrodes. Using Eqs. (11)–(13) with arbitrary coefficients for two SAW structures connected in series, we easily calculate the components of the complete *P* matrix:

$$(P_{11})_{\mathrm{S}} = (P_{11})_1 + (P_{12})_1(P_{11})_2(P_{21})_1/P_0, \quad (15)$$

$$(P_{12})_{\mathrm{S}} = (P_{12})_1(P_{12})_2/P_0, \quad (16)$$

$$(P_{13})_{\mathrm{S}} = (P_{13})_1$$
$$+ (P_{12})_1[(P_{13})_2 + (P_{11})_2(P_{23})_1]/P_0, \quad (17)$$

$$(P_{21})_{\mathrm{S}} = (P_{21})_1(P_{21})_2/P_0, \quad (18)$$

$$(P_{22})_{\mathrm{S}} = (P_{22})_2 + (P_{21})_2(P_{22})_1(P_{12})_2/P_0, \quad (19)$$

$$(P_{23})_{\mathrm{S}} = (P_{23})_2$$
$$+ (P_{21})_2[(P_{23})_1 + (P_{13})_2(P_{22})_1]/P_0, \quad (20)$$

$$(P_{31})_S = (P_{31})_1$$
$$+ (P_{21})_1[(P_{31})_1 + (P_{11})_2(P_{32})_1]/P_0, \quad (21)$$

$$(P_{32})_S = (P_{32})_2$$
$$+ (P_{12})_2[(P_{32})_1 + (P_{22})_1(P_{31})_2]/P_0, \quad (22)$$

$$(P_{33})_S = (P_{33})_1 + (P_{33})_2$$
$$+ \{(P_{32})_1[(P_{13})_2 + (P_{11})_2(P_{23})_1] \quad (23)$$
$$+ (P_{31})_2[(P_{23})_1 + (P_{22})_1(P_{13})_2]\}/P_0,$$

where $P_0 = 1 - (P_{11})_2(P_{22})_1$. The subscripts S, 1, and 2 at the parentheses refer to the complete P matrix, P matrix of the SAW structure on the left, and P matrix of the SAW structure on the right, respectively. The SAW structure can be viewed as an individual electrode or a group of electrodes for which the complete P matrix is known.

The P matrix of the reflecting structure can be obtained from (11)–(13) under the additional conditions $\xi_K = 0$, $U_0 = 0$, and $C_2 = 0$. Note that the elements P_{13}, P_{23}, P_{31}, P_{32}, and P_{33} are zero.

When designing filters and resonators, we need to know the P matrix for the gap between the transducer and the reflecting structure or between two transducers. The form of this matrix is well known. Therefore, we only note that it can also be obtained from (11)–(13) under the conditions $\xi_K = 0$, $U_0 = 0$, $C_2 = 0$, $r_K = 0$, and $W_K = 0$. In this case, only the P_{12} and P_{21} elements of the P matrix are zero.

The above expressions can be used for calculating the input admittance of the interdigital transducer in a filter or resonator where the electrode spacing and electrode aperture vary arbitrarily along the transducer and the reflecting structure with the surface current (charge) distribution over the transducer electrodes is real. Note that the input admittance of the transducer depends on the P_{33} element of the P matrix for the whole SAW structure.

4. ANALYSIS OF SAW RESONATORS BY THE COM METHOD

A still better optimization of a set of h, N, and W obtained by solving Eq. (1) (or from the nomograms in Figs. 3 and 4) under the conditions $N_{P0}W_{P0} = $ const and $N_{S0}W_{S0} = $ const can be made by analyzing the frequency responses of the resonators and the filter as a whole with the COM method. We used the modified COM method to refine the geometry of the resonator elements. The frequency responses of the resonators and filter were calculated for 3 or 4 possible pairs of h and N that lie on the curve in Fig. 5 with the corresponding aperture sizes W computed from formulas (4) and (5) to find the best combination of h, N, and W for each of the resonators of the filter. The combination of choice must satisfy the maximum condition for the Q factor and the
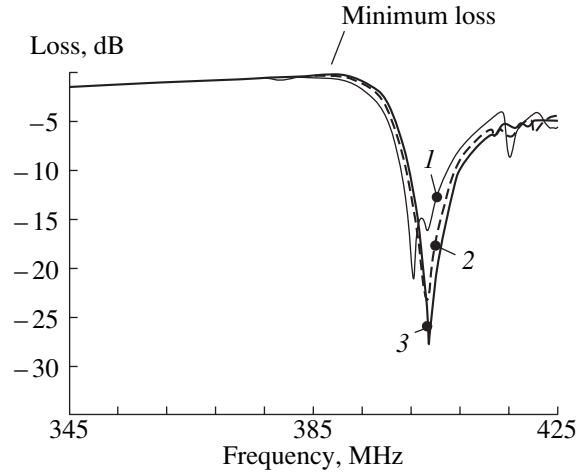
Fig. 7. Frequency response of SAW resonators (modified COM method).

minimum condition for the loss of the resonators, with individual resonators connected as in the filter. Figure 7 shows the frequency responses of the series resonator on the 36° Y, X-cut LiTaO$_3$ substrate for three values of N_S (the condition $N_{S0}W_{S0} = $ const is met). The table summarizes the Q factors and minimal losses corresponding to the curves in Fig. 7.

Since the filter must provide good electrical performance with the piezoelectric substrate size not too small, we took $N_S = 127$, $W_S = 17.8\lambda$, and $h = 0.85$ μm. Other parameters of the resonator that must be chosen as a result of the rigorous analysis are the gap between the transducer and reflecting structure and the ratio of their periods. These parameters are selected so as to make the frequency response as flat as possible, e.g., without spikes or dips. The number of reflectors in the reflecting structure depends on the total loss per unit length α in the transducer. For an SAW traveling through a reflecting structure, the most significant loss factors are wave leakage [1, 6], reradiation into bulk waves (for metal strips [1, 6] and grooves [9]), and ohmic loss in metal strips. Also, with the number of reflectors chosen properly, the frequency response of the filter is free of regular ripple.

5. LADDER-TYPE FILTER WITH A WIDER-THAN-OPTIMAL PASSBAND

The maximal bandwidth of a ladder-type filter depends on the maximum possible coupling coefficient. For materials with pronounced piezoelectric properties like lithium niobate and lithium tantalate, $k^2(h)$ depends on the maximum permissible thickness of the metal film at which a decrease in the insertion loss due to an increase in $k^2(h)$ is not accompanied by an increase in the loss due to higher leakage and efficiency of transformation into bulk waves [1, 6]. For instance, for 36° Y, X-cut LiTaO$_3$, the maximum permissible thickness of
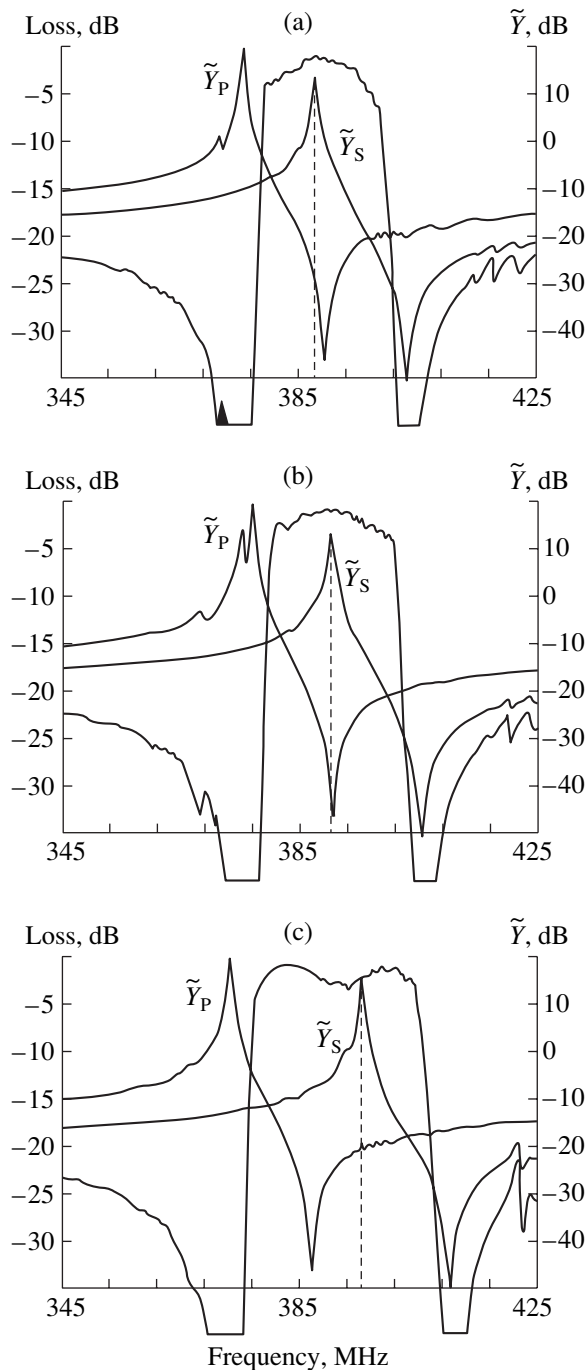
Fig. 8. Frequency response of ladder-type filters and normalized input admittances of their resonators, $\tilde{Y} = 20\{\log_{10}[|Y|/G_0]$, for $\Delta F_C/F_0 =$ (a) 4.7, (b) 5.7, and (c) 8.6%.

Fig. 9. (*1*) Experimental frequency response of a five-section SAW ladder-type filter and (*2*) theoretical response calculated by the modified COM method.

to have the maximum bandwidth corresponding to the maximum permissible thickness of the metal film, and the resonance frequencies of the series and parallel resonators are subsequently separated by an appropriate amount. If the shift in the resonance frequencies is too large, a dip appears at the center of the frequency response.

By way of example, Fig. 8 shows theoretical frequency responses of a ladder-type filter with the above electrical parameters. Five sections of the filter are connected in the order *a–b–a–b–a* (see Fig. 2). Here, the normalized input admittances $\tilde{Y}$ of the filter's resonators as a function of frequency are also shown. These plots illustrate the possibility of changing the passband width by separating the resonance frequencies. The normalized input admittance is defined as $\tilde{Y} = 20\{\log_{10}[|Y|/G_0]\}$, where $G_0 = (4/\pi)k^2(h_1)\omega_0 N^2 C_2$. For the given bandwidth $\Delta F_C = 4.7\%$, the theoretical –3-dB bandwidth is $\Delta F = 4.4\%$ (Fig. 8a); for $\Delta F_C = 5.7\%$, $\Delta F = 5.1\%$ (Fig. 8b); and for $\Delta F_C = 8.7\%$, $\Delta F = 6.9\%$ (Fig. 8c). The parallel transducer contains 81 electrodes with an aperture size of 37.3λ, the series transducer contains 127 electrodes with an aperture size of 17.8λ, the thickness of the aluminum film is 0.85 µm, each reflecting structure contains 120 reflectors, and the loss $\alpha' = 10\log_{10}[\exp(\alpha\lambda)]$ is taken to be 0.02 dB/λ [1]. Functions $k^2(h)$ and $r^0(h)$ were obtained from the data presented in [5].

Note that increasing the bandwidth to ~6% does not produce a dip at the center of the response.

6. EXPERIMENTAL RESULTS

With the synthesis method proposed above and supplemented by the modified COM technique, we designed and fabricated a ladder-type filter consisting of single-port SAW resonators with $F_0 = 392$ MHz,

the aluminum film is $h \approx 8\%$ [1]. Accordingly, the maximum feasible bandwidth is ≈5%. However, the resonance frequency of the parallel resonator is equal, in this case, to the resonance frequency of the series resonator. Separating the resonance frequencies, one can extend the filter's bandwidth without noticeably degrading its performance. Then, the filter is designed
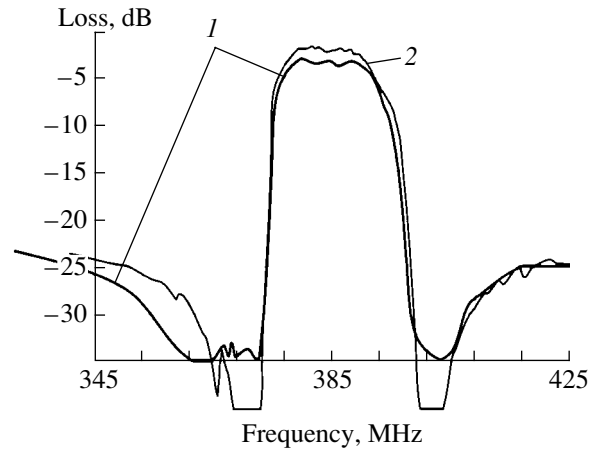
Table

Number of electrodes, N_S	Aperture W in wavelengths	Q factor	Minimum loss, dB	Curve number in Fig. 7
63	35.6	279	0.29	1
127	17.8	359	0.23	2
253	8.9	389	0.22	3

$\Delta F_C = 18.5$ MHz, and $K_B = 1.3$. The filter had five sections connected in series in the order a–b–a–b–a (see Fig. 2). The piezoelectric material used was $36°$ Y, X-cut LiTaO$_3$. For the parallel resonator, $C_D/C_0 = 0.0643$; for the series resonator, $C_D/C_0 = 0.0623$. The transducers of the parallel and series resonators contained, respectively, $N_{P0} = 81$ and $N_{S0} = 127$ 0.85-μm-thick aluminum electrodes with apertures $W_P = 37.3\lambda$ and $W_S = 17.8\lambda$. The ratio between the period of the reflecting structure and the transducer half-period was 1.03 for the parallel resonator and 0.97 for the series one. The analytical frequency response of the filter (the modified COM method) and the experimental one are shown in Fig. 9. The slightly higher insertion loss at the center of the passband is due to the inappropriate choice of the gap between the transducer and reflecting structure in the resonators. The small ripple in the passband is attributed to the interaction between the SAW and bulk waves reflected from the back surface of the substrate [7]. The gap width was equal to the half-period of the transducer. The discrepancy in the shapes of the analytical and experimental curves near $F_{1\infty}$ stems from the inaccurately selected phase of the coefficient of reflection from the transducer electrodes.

CONCLUSION

The method described in this paper can be used for synthesizing SAW-R ladder-type filters based on low-k^2 materials (such as quartz or langasite), as well as on materials with pronounced piezoelectric properties (lithium niobate or lithium tantalate).

REFERENCES

1. K. Hashimoto, M. Yamaguchi, S. Mineyoshi, *et al.*, Proc. IEEE Ultrason. Symp., 245 (1997).
2. O. Przesmycki, *Filtry elektryczne* (Wydawn. Komunikacji i Laczno Sci., Warsaw, 1962; Svyaz', Moscow, 1968).
3. J. Zelenka, *Piezoelectric Resonators and Their Applications* (Elsevier, Amsterdam, 1986; Mir, Moscow, 1990).
4. S. Lehtonen, V. Plessky, M. T. Honkanen, *et al.*, Appl. Phys. Lett. **75**, 142 (1999).
5. U. Rosler, D. Cohrs, and A. Dietz, Proc. IEEE Ultrason. Symp., 247 (1995).
6. S. N. Kondratiev, T. Thorvaldson, *et al.*, Proc. IEEE Ultrason. Symp. (1999).
7. S. V. Birykov, G. Martin, V. G. Polevoi, *et al.*, IEEE Trans. Ultrason. Ferroelectr. Freq. Control UFFC-42, 612 (1995).
8. V. F. Dmitriev and B. A. Kalinikos, Zh. Tekh. Fiz. **59**, 197 (1989) [Sov. Phys. Tech. Phys. **34**, 118 (1989)].
9. V. F. Dmitriev, Zh. Tekh. Fiz. **65** (8), 111 (1995) [Tech. Phys. **40**, 809 (1995)].

Translated by A. Khzmalyan

New Optimal Orientations for Surface Acoustic Waves in Langasite, Langanite, and Langatate Piezoelectric Crystals

M. Yu. Dvoesherstov, S. G. Petrov, V. I. Cherednik, and A. P. Chirimanov

Lobachevsky State University, pr. Gagarina 23, Nizhni Novgorod, 603600 Russia
e-mail: dvoesh@rf.unn.runnet.ru
Received December 11, 2001

Abstract—Spatial regions and optimal orientations for surface acoustic waves in langasite, langanite, and langatate piezoelectric crystals are analyzed. An objective function for the search for optimal orientations in the piezocrystals is suggested. New optimal spatial regions and orientations that are promising for SAW-based devices are found. © 2002 MAIK "Nauka/Interperiodica".

INTRODUCTION

Langasite ($La_3Ga_5SiO_{14}$, LGS), langanite ($La_3Ga_{5.5}Nb_{0.5}O_{14}$, LGN), and langatate ($La_3Ga_{5.5}Ta_{0.5}O_{14}$, LGT) piezocrystals, which have the trigonal structure (class 3 : 2), are promising substrate materials for SAW-based acoustooptic devices. To date, spatial regions where parameters of propagating SAWs are optimal (specifically, high electromechanical coupling coefficient K^2, zero angle of energy flux (AEF), appropriate anisotropy parameter γ, and zero first-order temperature coefficient of delay $\alpha_\tau^{(1)}$) have been found in these crystals [1–4].

Any spatial orientation in a crystal can be described with three Euler angles (ϕ, θ, and Ψ). The optimal space regions found in [1–3] are the following: $-15° < \phi < +10°$, $120° < \theta < 165°$, and $20° < \Psi < 45°$ for LGS; $-5° < \phi < +5°$, $130° < \theta < 150°$, and $15° < \Psi < 35°$ for LGN and LGT. These regions are thermally stable in terms of SAW propagation ($\alpha_\tau^{(1)} < 10^{-6} 1/°C$) only near room temperature ($t_0 = 25°C$).

The aim of this work was to seek new optimal thermally stable orientations for SAW propagation in LGS, LGN, and LGT in a temperature range of $-60 < t < +60°C$. Mathematically, this means that $\alpha_\tau^{(1)} \approx 0$ and the second-order temperature coefficient of SAW delay $\alpha_\tau^{(2)}$ ($\alpha_\tau^{(2)} \sim 10^{-9} 1/°C^2$) is as small as possible.

Orientations that are optimal for SAW propagation can be found only with numerical experiments. The accuracy of an optimal orientation found by calculation depends on the proper choice of the material constants of a piezocrystal (tensors of elastic constants, C_{ijkl}; piezoelectric constants, p_{ijk}; dielectric constants, ε_{ij}; linear expansion coefficients, β_{ij}; as well as the densities ρ at various temperatures t), which specify the crystal behavior and are the input in the problem stated. The

published data for the material constants of the piezocrystals considered in this work diverge [5–8]. The existing discrepancies affect the SAW velocity, K^2, *AEF*, and γ insignificantly; however, the first- and second-order temperature coefficients of delay change considerably [9]. This is because the disagreements refer primarily to the thermal constants (TC_{ijkl}, Tp_{ijk}, $T\varepsilon_{ij}$, and $T\rho$) of LGS, LGN, and LGT crystals. When calculating the SAW parameters in the spatial regions mentioned above, we use the data reported in [1–3]. Namely, the ranges of the Euler angles were $\phi = 0°–30°$, $\theta = 0°–180°$, and $\Psi = 0°–180°$. The search for new optimal orientations was accomplished by the technique described in [10]. Having carried out the numerical experiments, we found thermally stable orientations that are optimal for SAW propagation in a wide temperature range. Other SAW parameters also remained optimal.

SEARCH METHOD FOR OPTIMAL ORIENTATIONS

The most efficient way of searching for spatial orientations that are optimal in any SAW parameters or in a linear combination of the parameters is global optimization [10–12]. It is only necessary that solutions for SAWs exist throughout the search domain [12]. The SAW phase velocity can be directly calculated by the well-known Farnell–Jones [4] and Adler [13] procedures.

When using the method of searching for a global extremum of a multivariable function [12], one should write this function in the form of an objective function (OF) and analyze the objective function to find an orientation optimal for SAW propagation in a piezocrystal. The objective function can be represented as a linear combination of the wave basic parameters (K^2, *AEF*, $\alpha_\tau^{(1)}$, $\alpha_\tau^{(2)}$, etc.) with expert coefficients a_i, which make

it possible to vary the contribution of each of the terms to the OF:

$$OF(\phi, \theta, \psi) = \min\{a_1 AEF + a_1 \alpha_\tau^{(1)} + a_3 \alpha_\tau^{(2)} \\ + a_4(\gamma + 1) + \dots \}. \quad (1)$$

The search for orientations at which all the wave parameters are optimal simultaneously is very problematic, since one cannot be sure that such an orientation will be found. It is easier to find the optimum of a parameter that is of importance for a specific application or some trade-off between several of the parameters. For example, setting all the expert coefficients in (1), except for those at $\alpha_\tau^{(1)}$ and $\alpha_\tau^{(2)}$, equal to zero, one can seek a thermally stable direction in a wide temperature range by varying all the three Euler angles between $0°$ and $360°$ with a preset step. Alternatively, setting the expert coefficients, except for those at the AEF, $\alpha_\tau^{(1)}$, and $\alpha_\tau^{(2)}$, equal to zero, one can find a thermally stable direction for SAW propagation at which the AEF is small, etc. The time of searching for an optimal orientation in the crystal is naturally proportional to the number of parameters to be optimized (velocity V, K^2, AEF, $\alpha_\tau^{(1)}$, $\alpha_\tau^{(2)}$, γ, etc.). The Euler angles ϕ, θ, and Ψ in (1) should be varied with a step as small as possible to find the minimum of the OF. In this case, the orientation found will be optimal in terms of the optimal Euler angles ϕ_0, θ_0, and Ψ_0. The step of variation of the Euler angles should also be optimized. If the step is large, one may miss a desired optimization during the search; at a very small step, the time of search increases sharply. In this work, the Euler angles were varied with a step of $2°$.

In searching for an optimal orientation (ϕ_0, θ_0, Ψ_0), one must find the phase velocity of the wave on the open, V_o, and metallized, V_m, crystal surfaces for each current orientation (ϕ, θ, and Ψ). The other SAW parameters (K^2, AEF, $\alpha_\tau^{(1)}$, $\alpha_\tau^{(2)}$, and γ) are calculated if the values of V_o and V_m of the SAW are known in the vicinity $\pm\Delta\Psi$ of a direction chosen and in the interval $\pm\Delta t$ of a temperature selected. According to the Farnell–Jones technique, the SAW phase velocity is also a parameter in the search algorithm; therefore, the method of global optimization [12] is also applied for searching the SAW true velocity in a given direction in the crystal.

Thus, for each specific set of Euler angles (ϕ, θ, Ψ), one must first find the minimum of the function $F(V)$ of boundary conditions [12], which is represented either in the form of the boundary condition determinant (Farnell–Jones technique) or as a function of effective permittivity (Adler technique), and then find the minimum of function (1).

SAW BASIC PARAMETERS

It is known that the phase velocity of an SAW propagating in some direction in a piezocrystal can be determined by solving a set of coupled equations from the elasticity theory as applied to an anisotropic piezoelectric medium [4]. Then, the other SAW parameters are calculated. The electromechanical coupling coefficient for an SAW is given by [4]

$$K^2 = \frac{2(V_o - V_m)}{V_o}, \quad (2)$$

where V_o and V_m are the SAW phase velocities on the open and metallized (shorted) surfaces of a piezocrystal, respectively.

The anisotropy parameter γ, which relates the SAW velocity and the angle Ψ, $\left(V(\Psi) \approx V(0)\left(1 + \gamma\frac{\Psi^2}{2}\right)\right)$, that is, diffraction losses of the wave due to crystal anisotropy, is calculated from the formula [4]

$$\gamma = \frac{d(AEF)}{d\Psi}, \quad (3)$$

where AEF is the angle between the energy flux direction and the wave normal. According to [4],

$$AEF = \arctan\frac{1}{V}\frac{dV}{d\Psi}. \quad (4)$$

In an optimal direction, AEF must be close to zero and $\gamma = -1$ [4].

The sensitivity to temperature variations is a most important characteristic of an SAW. For a given orientation, SAW thermal properties are described by the first- and second-order temperature coefficients of delay, which are given, respectively, by

$$\alpha_\tau^{(1)} = \frac{1}{\tau}\frac{d\tau}{dt} = \beta^{(1)} - \frac{1}{V}\frac{dV}{dt} = \beta^{(1)} - \alpha_v^{(1)}, \quad (5)$$

$$\alpha_\tau^{(2)} = \frac{1}{2}\frac{1}{\tau}\frac{d^2\tau}{dt^2}. \quad (6)$$

Here, $\tau = l/V$ is the delay time, l is the SAW travel, $\beta^{(1)} = 1/l\,dl/dt$ is the first-order linear expansion coefficient in the wave direction, and $\alpha_v^{(1)}$ is the first-order temperature coefficient of SAW velocity.

From (5), it follows that $\alpha_\tau^{(1)}$ is easy to calculate if β and the phase velocity of the wave in the temperature interval $\pm\Delta t$ are known. The mathematics that follow is necessary to derive an expression from which $\alpha_\tau^{(2)}$ can be calculated. Let us differentiate the time delay τ with respect to temperature t once more:

$$\frac{d^2\tau}{dt^2} = \frac{l}{V}\left[\frac{1}{l}\frac{d^2l}{dt^2} - \frac{1}{V}\frac{d^2V}{dt^2} + \frac{2}{V^2}\left(\frac{dV}{dt}\right)^2 - \frac{2}{lV}\frac{dV}{dt}\frac{dl}{dt}\right],$$

then, the second-order temperature coefficient of SAW delay is

$$\frac{1}{2}\frac{1}{\tau}\frac{d^2\tau}{dt^2} = \frac{1}{2}\frac{1}{l}\frac{d^2l}{dt^2} - \frac{1}{2}\frac{1}{V}\frac{d^2V}{dt^2} + \frac{1}{V^2}\left(\frac{dV}{dt}\right)^2 - \frac{1}{lV}\frac{dV}{dt}\frac{dl}{dt}$$

or

$$\alpha_\tau^{(2)} = \beta^{(2)} - \alpha_v^{(2)} + \alpha_v^{(2)} - \beta^{(1)},$$
$$\alpha_v = \beta_l^{(2)} - \alpha_v^{(2)} - \alpha_\tau^{(1)}\alpha_v^{(1)},$$

(7)

where $\beta^{(2)}$ and $\alpha_v^{(2)}$ are the second-order temperature coefficients of linear expansion and velocity, respectively.

On the other hand,

$$\frac{d\alpha_\tau^{(1)}}{dt} = -\frac{1}{\tau^2}\frac{d\tau}{dt}\frac{d\tau}{dt} + \frac{1}{\tau}\frac{d^2\tau}{dt^2},$$

hence,

$$\frac{1}{\tau}\frac{d^2\tau}{dt^2} = \frac{d\alpha_\tau}{dt} + \alpha_\tau^{(1)^2}.$$

Then,

$$\alpha_\tau^{(2)} = \frac{1}{2}\frac{1}{\tau}\frac{d^2\tau}{dt^2} = \frac{1}{2}\left(\frac{d\alpha_\tau^{(1)}}{dt} + \alpha_\tau^{(1)^2}\right).$$

Since $\alpha_\tau^{(1)} \sim 10^{-6} 1/°C$ and $\alpha_\tau^{(2)} \sim 10^{-9} 1/°C^2$, we always have $\alpha_\tau^{(1)^2} \ll \alpha_\tau^{(2)}$ and can use the simpler expression for $\alpha_\tau^{(2)}$

$$\alpha_\tau^{(2)} = \frac{1}{2}\frac{d\alpha_\tau^{(1)}}{dt}.$$

(8)

Similarly, the second-order temperature coefficient of SAW frequency, $\alpha_f^{(2)}$, can be calculated by the formula

$$\alpha_f^{(2)} = \frac{1}{2}\frac{1}{f}\frac{d^2f}{dt^2} = -\frac{1}{2}\frac{1}{\tau}\frac{d^2\tau}{dt^2} + \alpha_\tau^{(1)^2} = -\alpha_\tau^{(2)} + \alpha_\tau^{(1)^2}.$$

(9)

In the abridged matrix form, the material constants of a piezoelectric depend on temperature t as [4]

$$C_{ij}(t) = C_{ij}(t_0)(1 + TC_{ij}^{(1)}(t - t_0) + TC_{ij}^{(2)}(t - t_0)^2 + \ldots),$$
$$e_{ij}(t) = e_{ij}(t_0)(1 + Te_{ij}^{(1)}(t - t_0) + Te_{ij}^{(2)}(t - t_0)^2 + \ldots),$$
$$\varepsilon_{ij}(t) = \varepsilon_{ij}(t_0)(1 + T\varepsilon_{ij}^{(1)}(t - t_0) + T\varepsilon_{ij}^{(2)}(t - t_0)^2 + \ldots),$$
$$\rho(t) = \rho(t_0)(1 + T\rho^{(1)}(t - t_0) + T\rho^{(2)}(t - t_0)^2 + \ldots),$$

(10)

where $TC_{ij}^{(1)}$, $Te_{ij}^{(1)}$, $T\varepsilon_{ij}^{(1)}$, $TC_{ij}^{(2)}$, $Te_{ij}^{(2)}$, $T\varepsilon_{ij}^{(2)}$, $T\rho^{(1)}$, and $T\rho^{(2)}$ are, respectively, the first- and second-order temperature coefficients of the elastic, piezoelectric,

and dielectric constants of a crystal ($t_0 = 25°C$ is the room temperature). The temperature coefficient of density $T\rho^{(1)}$ is calculated by the formula [4] $T\rho^{(1)} = -(2\beta_{11}^{(1)} + \beta_{33}^{(1)})$, where $\beta_{11}^{(1)}$ and $\beta_{33}^{(1)}$ are the first-order linear expansion coefficients on the X_1 and X_3 directions, respectively.

The thermal stability of SAWs in a wide temperature range is usually estimated through the relative variation of the time delay (or the frequency) $\Delta\tau/\tau_0$ ($\Delta f/f_0$) in a given direction of a piezocrystal:

$$\frac{\Delta\tau}{\tau_0} = \frac{\tau - \tau_0}{\tau_0} = \alpha_\tau^{(1)}(t - t_0) + \alpha_\tau^{(2)}(t - t_0)^2 + \ldots \quad (11)$$

It can be shown that $\Delta\tau/\tau_0 = -\Delta f/f_0$.

When calculating $\alpha_\tau^{(1)}$ near the room temperature t_0, one usually uses only the first-order values of the temperature elastic, piezoelectric, and dielectric constants; density; and thermal expansion coefficient [4]. However, as follows from (10), if the temperature differs considerably from the room value t_0, the factor $(t - t_0)$ may be sufficiently large. Moreover, the second-order temperature material constants can also be very high. Therefore, the material temperature constants of both orders should be taken into account in order to more accurately evaluate $\alpha_\tau^{(1)}$ in a wide temperature range. Below, we perform the necessary calculations to show how the value of $\Delta\tau/\tau_0$ can be evaluated with a higher accuracy. From (5), it follows that $(d\tau)/\tau = \alpha_\tau dt$. Integrating this expression from t_0 to t yields

$$\ln\frac{\tau}{\tau_0} = \int_{t_0}^{t}\alpha_\tau dt$$

or

$$\frac{\tau}{\tau_0} = \exp^{\int_{t_0}^{t}\alpha_\tau dt} ;$$

hence,

$$\frac{\tau - \tau_0}{\tau_0} = \exp^{\int_{t_0}^{t}\alpha_\tau dt} - 1.$$

In practice, we always have $\int_{t_0}^{t}\alpha_\tau dt \ll 1$; hence,

$$\exp^{\int_{t_0}^{t}\alpha_\tau dt} \approx 1 + \int_{t_0}^{t}\alpha_\tau dt,$$

and eventually

$$\frac{\tau - \tau_0}{\tau_0} = \int_{t_0}^{t} \alpha_\tau dt. \tag{12}$$

Expression (12) is more general than (11), because the former includes all orders of α_τ, rather than just the first and second ones. In calculating $\Delta\tau/\tau$ by formula (12), it is sufficient to know only the dependence $\alpha_\tau(t)$, which is calculated for all the temperature material constants of the first, second, etc. order [see (10)].

Thus, with the formulas given above, one can evaluate the SAW basic parameters in a wide temperature range and in any direction of a piezocrystal.

NUMERICAL EXPERIMENT

Using the above techniques, we analyzed various spatial regions on LGS, LGN, and LGT piezocrystals for the purpose of optimizing the SAW parameters. To find a particular thermally stable direction of SAW propagation, a family of isolines for the SAW parameters in the space of the Euler angles θ and Ψ (with ϕ fixed) was calculated and constructed. Figures 1–3 show the respective isolines for $\alpha_\tau^{(1)}$ and *AEF* in the ranges $0° < \theta < 180°$ and $0° < \Psi < 180°$ ($\phi = 10°$) for the LGS, LGN, and LGT crystals.

As follows from Fig. 1, the LGS crystal contains previously discovered region *2* [1] and region *1* ($\phi = 10°$, $20° < \theta < 40°$, $70° < \Psi < 85°$), where the isolines of zero $\alpha_\tau^{(1)}$ are observed (Fig. 1a). Also, region *1* contains isolines of zero *AEF* (Fig. 1b). As for the value of K^2 in this region, it is relatively high (0.1–0.2%). In region *1*, we looked for an orientation for which both $\alpha_\tau^{(1)}$ and *AEF* are zero. To do this, all the expert coefficients a_i in objective function (1), except for a_1 and a_2, were set equal to zero. After the calculations, we found a particular orientation in the LGS crystal (10°, 21.3°, 78.9°) for which the SAW parameters are as follows: $V = 2.935$ km/s, $K^2 = 0.2\%$, $AEF = 0.1°$, $\alpha_\tau^{(1)} = -0.07 \times 10^{-6}$ 1/°C, and $\alpha_\tau^{(2)} = 95.6 \times 10^{-9}$ 1/°C^2. For comparison, the SAW parameters in region *2* of the LGS crystal with the orientation (0°, 141°, 25°) are $V_o = 2.743$ km/s, $K^2 = 0.35\%$, $AEF = -0.8°$, $\alpha_\tau^{(1)} = -1.75 \times 10^{-6}$ 1/°C, and $\alpha_\tau^{(2)} = 76.1 \times 10^{-9}$ 1/°C^2.

The region ($\phi = 10°$, $15° < \theta < 25°$, $70° < \Psi < 85°$) of the LGN crystal (Fig. 2) also contains the isolines of zero $\alpha_\tau^{(1)}$ (Fig. 2a) and zero *AEF* (Fig. 2b). In this angular range, the electromechanical coupling coefficient K^2 lies in the range 0.1–0.3%. The numerical search for a specific orientation where both $\alpha_\tau^{(1)}$ and *AEF* are zero gave the following angles: $\phi = 10°$, $\theta = 19.2°$, and $\Psi =$
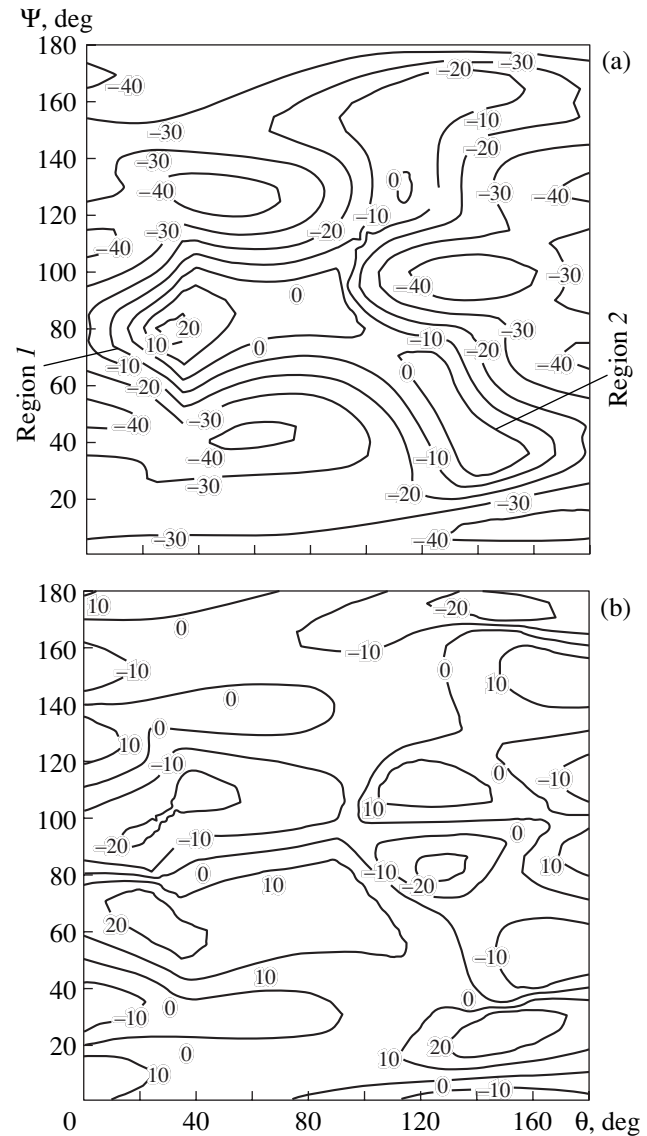

Fig. 1. Isolines for (a) α_τ ($\times 10^{-6}$ 1/°C) and (b) *AEF* in the LGS crystal (10°, θ, Ψ).

79.4°. The associated SAW parameters were found to be $V_m = 2.799$ km/s, $V_o = 2.804$ km/s, $K^2 = 0.278\%$, $AEF = -0.28°$, $\alpha_\tau^{(1)} = -0.26 \times 10^{-6}$ 1/°C, and $\alpha_\tau^{(2)} = -264 \times 10^{-9}$ 1/°C^2.

In the angular range ($\phi = 10°$, $5° < \theta < 20°$, $70° < \Psi < 85°$), the LGT crystal (Fig. 3) also has the isolines of zero $\alpha_\tau^{(1)}$ (Fig. 3a) and *AEF* (Fig. 3b). The Euler angles specifying a new optimal orientation for SAW propagation in the LGT crystal were found to be $\phi = 10°$, $\theta = 10°$, and $\Psi = 79.5°$. In this direction, the SAW parameters are the following: $V_m = 2.609$ km/s, $V_o = 2.612$ km/s, $K^2 = 0.267\%$, $AEF = 0.9°$, $\alpha^{(1)} = -0.239 \times 10^{-6}$ 1/°C, and $\alpha_\tau^{(2)} = 576 \times 10^{-9}$ 1/°C^2. In the angular range ($\phi = 10°$, $5° < \theta < 20°$, $70° < \Psi < 85°$), the search

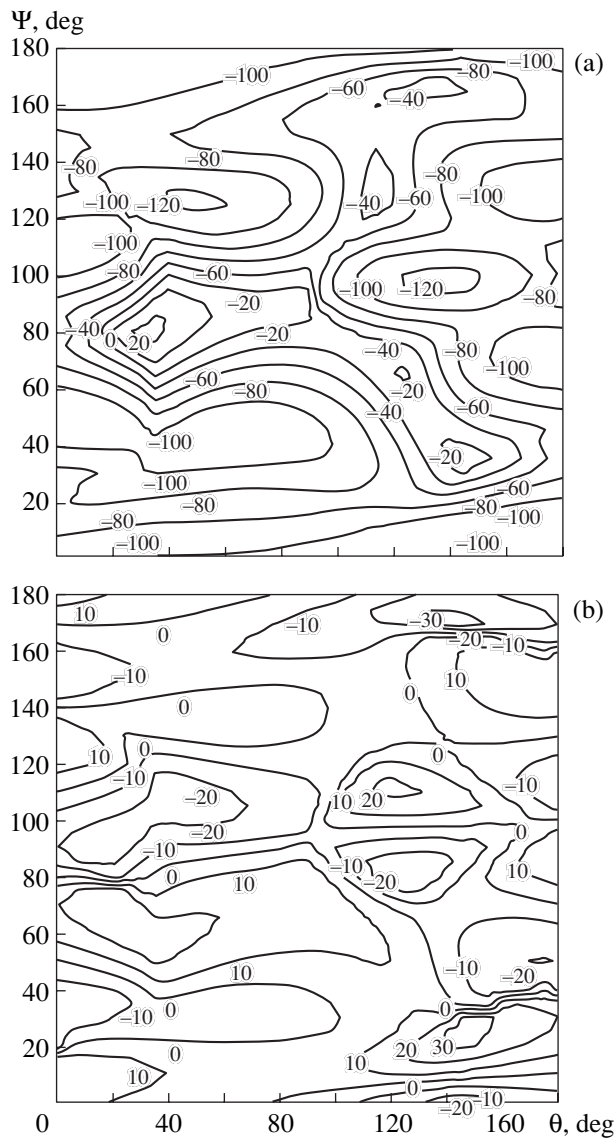**Fig. 2.** Isolines for (a) α_τ ($\times 10^{-6}$ 1/°C) and (b) *AEF* in the LGN crystal (10°, θ, Ψ).

Fig. 3. Isolines for (a) α_τ ($\times 10^{-6}$ 1/°C) and (b) *AEF* in the LGT crystal (10°, θ, Ψ).

for a direction with high K^2 and zero $\alpha_\tau^{(1)}$ failed. Moreover, for the LGT crystal, the absolute values of $\alpha_\tau^{(2)}$ in thermally stable ($\alpha_\tau^{(1)} \approx 0$) directions are several times higher than for the LGS and LGN crystals. This means that the directions found in this crystal are thermally stable in a narrow temperature range (near room temperature, $t_0 = 25$°C).

The calculations showed that optimal orientations in the crystals are observed not only at fixed $\phi = 10$° but also in the interval $0 < \phi < 30$°. For example, the thermally stable orientations and associated wave parameters in the LGT crystal are as follows: (1) $\phi = 0$°, $\theta = 9.85$°, $\Psi = 89.8$°; *AEF* = 0.75°; $\alpha_\tau^{(1)} = -0.27 \times 10^{-6}$ 1/°C; and $K^2 = 0.26$%; (2) $\phi = 20$°, $\theta = 10.6$°, $\Psi = 69.5$°; *AEF* =

-0.54°; $\alpha_\tau^{(1)} = -0.06 \times 10^{-6}$ 1/°C, and $K^2 = 0.33$%; and (3) $\phi = 30$°, $\theta = 11.5$°, $\Psi = 59$°; *AEF* = -0.58°; $\alpha_\tau^{(1)} = -0.17 \times 10^{-6}$ 1/°C, and $K^2 = 0.39$%.

When searching for a thermally stable direction for SAW propagation in the LGN crystal in a wide temperature range (all the expert coefficients, except for a_2 and a_3, were set equal to zero), we found the orientation $\phi = 10$°, $\theta = 23.48$°, $\Psi = 86.1$°, which has the near-zero value of $\alpha_\tau^{(2)}$. The associated SAW parameters were as follows: $V_m = 2.767$ km/s, $V_o = 2.771$ km/s, $K^2 = 0.295$%, *AEF* = -18°, $\alpha_\tau^{(1)} = 0.012 \times 10^{-6}$ 1/°C, $\alpha_\tau^{(2)} = 0.105 \times 10^{-9}$ 1/°C², and $\gamma = -0.64$. For this orientation
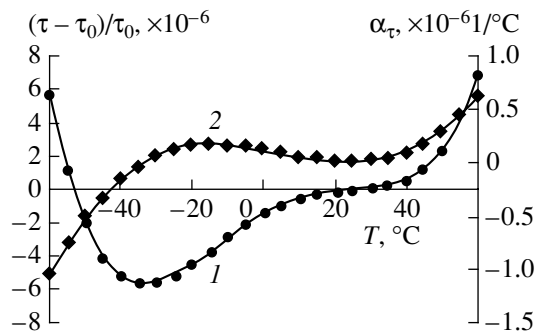
Fig. 4. (*1*) $\Delta\tau/\tau_0$ and (*2*) α_τ vs. temperature in the LGN crystal with the orientation (10°, 23.48°, 86.1°).

in the LGN crystal, the temperature dependences of $\Delta\tau/\tau_0$ (curve *1*) and $\alpha_\tau^{(1)}$ (curve *2*) in the interval from –60 to +60°C are plotted in Fig. 4. This direction is seen to be thermally stable in a wide range. Unfortunately, the value of *AEF* is very large, indicating a high anisotropy in this direction.

CONCLUSION

The objective function for searching for optimal orientations for SAW propagation in piezocrystals is suggested. New spatial regions in LGS, LGN, and LGT piezoelectrics where the SAW parameters are optimal are found. In the LGS crystal, a new spatial region optimal for SAW propagation is bounded by the ranges $0° < \phi < 30°$, $20° < \theta < 40°$, $70° < \Psi < 85°$; in the LGN crystal, $0° < \phi < 30°$, $15° < \theta < 25°$, $70° < \Psi < 85°$; and in the LGT crystal, $0° < \phi < 30°$, $5° < \theta < 20°$, $70° < \Psi < 85°$. In the LGN crystal, an orientation thermally stable in a wide range is found.

REFERENCES

1. N. F. Naumenko, *High Frequency SAW Device*, International Patent. PCN/US 96/17906. H03H 9/17.9/02. WO 97/25776 (1997).
2. N. F. Naumenko, *Optimal Cut for SAW Devices on Langanite*, International Patent. PCT/US5905325. H01L 41/08 (1999).
3. N. F. Naumenko, *An Optimal Cut for SAW Devices on Langatate*, International Patent. PCT/US99/05871. H03H 9/02. WO 99/48200 (1999).
4. *Acoustic Surface Waves*, Ed. by A. A. Oliner (Springer-Verlag, New York, 1978; Mir, Moscow, 1981).
5. A. Bungo, C. Jian, K. Yamaguchi, and Y. Sawada, Proc. IEEE Ultrason. Symp., 1 (1999).
6. T. Sato, M. Murota, and Y. Shimizu, Jpn. J. Appl. Phys. **37**, 2914 (1998).
7. S. Sakharov, P. Senushencov, A. Medvedev, and Yu. Pisarevsky, in *Proceedings of the IEEE Frequency Control Symposium, 1995*, p. 647.
8. D. C. Malocha, M. P. Cunha, and E. Adler, in *Proceedings of the IEEE Frequency Control Symposium, 2000*, p. 200.
9. M. Yu. Dvoesherstov, V. I. Cherednik, A. P. Chirimanov, and S. G. Petrov, Zh. Tekh. Fiz. **71** (4), 89 (2001) [Tech. Phys. **46**, 450 (2001)].
10. M. Yu. Dvoesherstov, V. I. Cherednik, A. P. Chirimanov, and S. G. Petrov, Proc. SPIE **3900**, 283 (1999).
11. D. J. Wilde, *Optimum Seeking Methods* (Prentice-Hall, Englewood Cliffs, 1964; Nauka, Moscow, 1967).
12. M. Yu. Dvoesherstov, V. I. Cherednik, and A. P. Chirimanov, Izv. Vyssh. Uchebn. Zaved., Radiofiz. **43**, 801 (2000).
13. E. L. Adler, IEEE Trans. Ultrason. Ferroelectr. Freq. Control **41**, 876 (1994).

Translated by V. Isaakyan

SURFACES, ELECTRON AND ION EMISSION

High-Temperature Field Evaporation of Rhenium

O. L. Golubev and V. N. Shrednik

Ioffe Physicotechnical Institute, Russian Academy of Sciences,
ul. Politekhnicheskaya 26, St. Petersburg, 194021 Russia
e-mail: V.Shrednik@pop.ioffe.rssi.ru
Received January 21, 2002

Abstract—High-temperature field emission of Re, Pt, Ta, and W is studied by field-emission methods. Metal ions are found to evaporate mainly from the tops of thermal–field microprotrusions produced by high electric fields and temperatures on the emitter surface. For field intensities of up to $F = 1$–2 V/Å and temperatures of 1500–2000 K, the ion currents i are recorded from the entire emitter surface. They range from several tenths of nanoamperes to several nanoamperes. The activation energies of field evaporation determined from the Arrhenius plots $\log i = f(1/T)$ are found to be appreciably lower than those calculated within the charge exchange model for known parameters of the process and the metals evaporated. Reasons for such a difference in the activation energies and mechanisms of ion evaporation at high F and T are discussed. © 2002 MAIK "Nauka/Interperiodica".

INTRODUCTION

No more than twenty works have studied high-temperature field evaporation, while field evaporation at cryogenic temperatures has been the subject of investigation in more than a thousand papers. The list of materials studied is still narrower; these are Ir, W, Pt, and Au metals and several alloys (W–Hf, Mo–Hf, and W–Si), although this phenomenon is of great scientific and practical interest. High-temperature field evaporation occurs under the simultaneous action of a strong electric field (the field intensity F for refractory metals is usually 1–2 V/Å) and a high temperature, causing intense surface metal self-diffusion. This is accompanied by nontrivial physical phenomena on the metallic tip surface. For example, under these conditions, the effect of field crystalline growth has been discovered. This effect consists in the formation of large crystalline outgrowths (macrooutgrowths) on close-packed faces of the tip [1] and small microprotrusions on the vertices and edges of a faceted tip. Specific features of the dynamic growth and evaporation of macrooutgrowths give rise to the effect of ring collapse [2] in the tip surface image formed by evaporating ions. From the practical point of view, ion evaporation at high T and F allows the design of continuous-operation ion and electron point sources offering high reliability and long service time. Such sources can be successfully used in nanotechnology, surface examination, and in designing radically new electronic devices.

EXPERIMENTAL TECHNIQUES

The results presented in this paper were obtained largely for Re; some experiments were also carried out for Ta, W, and Pt. Tips with a radius of several fractions of a micron were prepared by standard electrolytic etching [3]. The experiments were carried out both in sealed diode microscopes containing the metal tips to be studied and luminescent collecting screens and with an instrument that combines the operating conditions of electron, ion, and field desorption microscopes and is equipped with a microchannel light amplifier. The tip temperature T was determined by optical pyrometry with an accuracy of 20°. The voltages and currents were measured by digital voltmeters accurate to 0.5–1%. The electric field intensity F was determined from the slope of the Fowler–Nordheim characteristics [4] under the assumption that the work function $\varphi = 5.0$ eV for Re, 4.4 eV for W, 5.32 eV for Pt, and 4.1 eV for Ta [5]. Complicated processes, such as surface self-diffusion and field crystalline growth, occurring on the emitter surface during ion evaporation under strong fields and high temperatures result in the complex dependence of the field on applied voltage U and T. Here, two values of F characterizing changes in the tip shape are of importance: the initial field, which should be applied to the emitter in order to change its shape at a given T for a fixed time, and the final field F, which results when the emitter takes any steady-state shape. The former will be referred to as the processing field F_{pr} and the latter, as the steady-state field F_{st}. The steady-state field does not necessarily cause field evaporation. The steady-state field F_{st} at which field evaporation takes place will be called the evaporation field F_{ev}. Static electron images were photographed, while field-desorption images, rapidly varying with time, were recorded by a streak camera. In the sealed devices, the residual gas pressure was 10^{-10}–10^{-1} torr; in the setup with the light amplifier, it was 10^{-8} torr.

EXPERIMENTAL RESULTS

The technique used in the experiments allowed us to simultaneously measure the ion current from the emitting tip, determine the values of F_{ev}, and observe the emitter surface with a near-atomic resolution. The dependences of the ion current i from the entire Re tip surface on the emitter temperature T are shown in Fig. 1 for different values of the voltage U applied between the tip and the screen. The temperature dependences of F_{ev} are shown in Fig. 2, where each curve $F_{ev}(T)$ corresponds to different values of i but to the same U. As follows from Fig. 1, the rise in the emitter temperature T by $\approx 15\%$ results in an increase in the ion current by a factor of 10–15. The dependences $i(T)$ can be divided into two parts: the slow rise at low T and the steep rise at high T. It should also be mentioned that i fluctuates appreciably at constant T and U: from tens of percent to several times. Therefore, the values of i measured are time-average values.

The dependence $F_{ev}(T)$ is shown in Fig. 2 for different U. The values of F varies with T only slightly and nonmonotonically. As a rule, F varies noticeably in the middle of the T range rather than at its extremes. If we take F_{ev} as the mean between the maximum and minimum values of F throughout the range, for $U = 10$ kV, we have $F_{ev} = 1.15 \pm 0.18$ V/Å; for $U = 11$ kV, $F_{ev} = 1.52 \pm 0.12$ V/Å; and for $U = 13$ kV, $F_{ev} = 1.92 \pm 0.08$ V/Å. It is seen that the growth F_{ev} is not proportional to the applied voltage U.

The reasons for such dependences of i and F_{ev} become clear from field-emission images of the Re tip surface subjected to high F and T (Fig. 3). The initial surface of the Re tip oriented in the $[11\bar{2}0]$ direction is shown in Fig. 3a, where the $(11\bar{2}0)$, $(10\bar{1}0)$, $(10\bar{1}1)$, and $(11\bar{2}2)$ faces are revealed on the surface. The most close-packed (0001) faces are on the top and bottom of the image. In the case of refractory metals, the tip of such a shape cannot emit ions at $F_{pr} = 0.3$–0.6 V/Å usually used in the experiments. However, the simultaneous action of high T and F on the tip leads to the well-known phenomenon of the thermal–field reconfiguration of the tip [6]. Figure 3b shows the tip reconfigured in the field; Fig. 3c, the fully reconfigured tip, where the surface is made up of close-packed faces separated by edges and vertices. For the fully reconfigured tip, ion emission can already be observed because the acting field increases from initial $F_{pr} = 0.41$ V/Å to $F_{ev} = 0.95$ V/Å for the fully reconfigured tip. For such values of F, Re ions may evaporate at $T \approx 1800$ K. The early portion of the dependence $i(T)$ corresponds to the emission from the fully (or almost fully) reconfigured tip or from the fully reconfigured tip with several microprotrusions (Fig. 3d). For high values of i, the next stage of thermal–field reconfiguration is observed, i.e., the formation of macrooutgrowths on the close-packed faces, with their vertices and edges covered by microprotru-
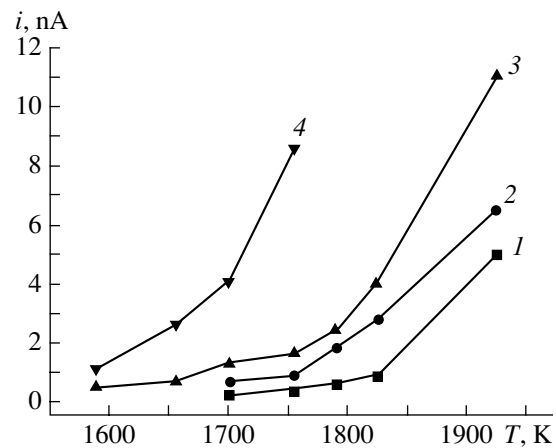

Fig. 1. Ion current from the entire surface of the Re tip as a function of the emitter temperature. $U = (1)$ 10, (2) 11, (3) 12, and (4) 13 kV.
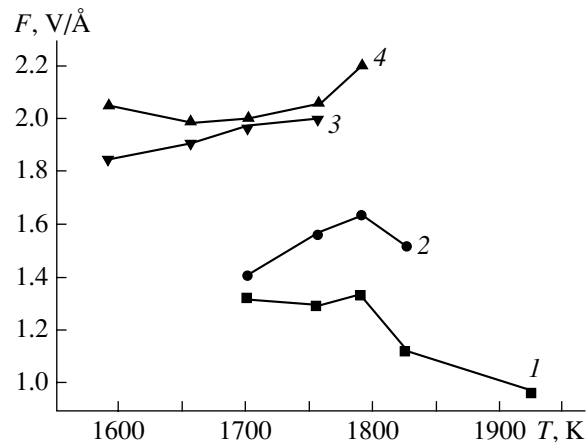

Fig. 2. Temperature dependence of the evaporating field. The values of U are the same as in Fig. 1.

sions. This shape of the tip is shown in Fig. 3e; the last figure in this series, Fig. 3f, shows the tip surface after the smoothing of the microprotrusions by short-term heating at moderate temperatures, $T \approx 1000$ K. The macrooutgrowths are seen on all the close-packed faces. As was shown in [2], in spite of the macroscopic sizes, the macrooutgrowths are unstable: they evaporate though ion emission from the tops of the microprotrusions and then appear again. These alternating processes of growth and evaporation of the macrooutgrowths covered by the microprotrusions lead to appreciable fluctuations in the ion current. Thus, the field-emission images suggest that the ions are emitted mainly from the tip surface covered by the macrooutgrowths and microprotrusions. These shapes obtained at different T and the same U differ slightly. This fact explains the minor change in F_{ev} with T in Fig. 2. The sharp growth in F with increasing T at constant U takes place only when the initially smooth annealed tip reconfigures its shape.
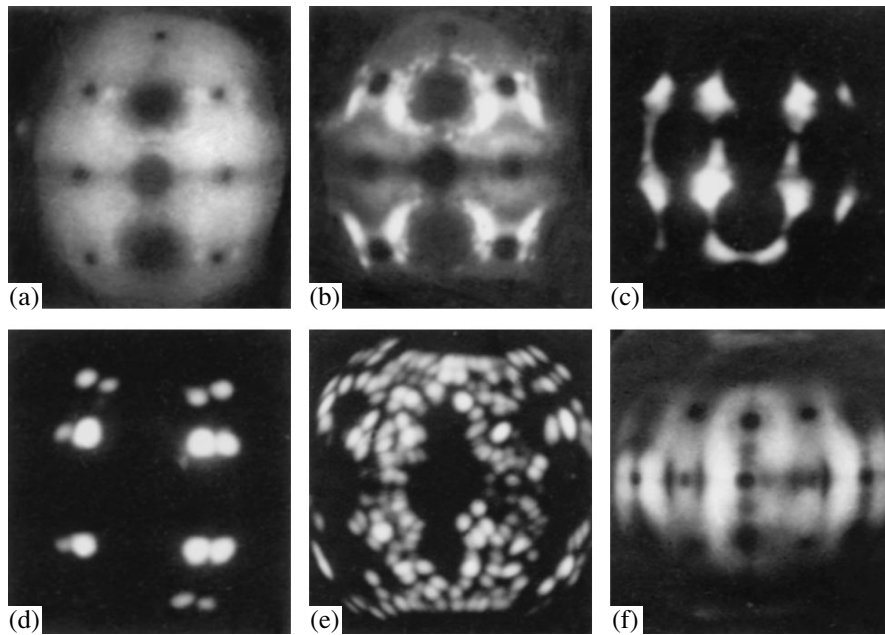
Fig. 3. Field electron images of the Re tip surface subjected to high F and T simultaneously. (a) Initial annealed shape, (b) tip reconfigured in the field, (c) fully reconfigured tip, (d) fully reconfigured tip with several microprotrusions, (e) tip surface with many microprotrusions grown on the vertices and edges of macrooutgrowths, and (f) tip surface after smoothing microprotrusions at $T \approx$ 1000 K (macrooutgrowths are seen on the close-packed faces of Re).

The same behavior is typical of the other metals studied. The initial shape of the annealed Ta tip is shown in Fig. 4a, and the shape with the microprotrusions covering the macrooutgrowths is illustrated in Fig. 4b. The microprotrusions generate the Ta ion current $i = 0.5$ nA at $T = 1800$ K and $F_{ev} = 1.27$ V/Å. The initial field factor of the tip $\beta = F/U$ increases from $\beta = 4995$ 1/cm for the annealed shape (Fig. 4a) to $\beta = 12\,656$ 1/cm (Fig. 4b). Accordingly, F increases from the initial value for the annealed shape, $F_{pr} = 0.51 \pm 0.02$ V/Å, to $F_{ev} = (1.24–1.27) \pm 0.02$ V/Å for the shape shown in Fig. 4b. In the experiments with W, we managed to measure i from a single microasperity. For $T =$
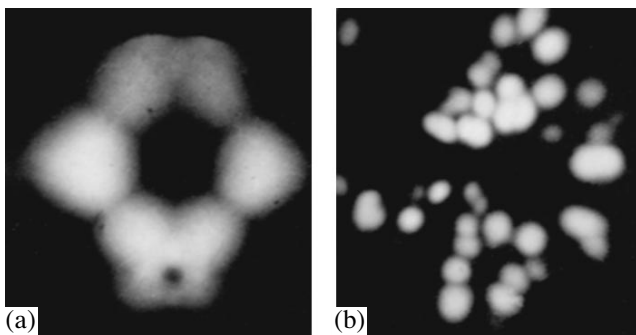

Fig. 4. Field electron images of the Ta tip surface subjected to high F and T simultaneously. (a) Initial annealed shape and (b) tip surface with microprotrusions on the tops of macrooutgrowths.

1600 K and $F \approx 1.5$ V/Å, the ion current usually varied between 10^{-13} and 10^{-12} A.

Measuring i at different T and the same U, one can determine kinetic parameters of high-temperature field evaporation, such as the energy of activation Q_n and the preexponential ν_0. According to Forbes [7], the ion current is given by

$$i = n_{hr}\nu_0\exp(-Q_n/kT), \tag{1}$$

where n_{hr} is the number of surface atoms in the "high-risk" position: in other words, the number of active emission centers on the surface.

In the case of low-temperature field emission, n_{hr} is the number of atoms near kinks, while in our case, it is the number of microprotrusions on the surface, because the ions evaporate from the tops of the microprotrusions. The slope of the Arrhenius plot $\log i = f(1/T)$ defines the value of Q_n; the current intercept, $n_{hr}\nu_0$. To construct the plot, the values of i were measured at different T and the same U, since the value of i is a measure of the rate of evaporation from the tip surface. The Arrhenius plot for the Re tip is shown in Fig. 5 at $U = 12$ kV. Here, F_{ev} varies from 1.9 to 2.1 V/Å; i.e., $F_{ev} = 2.0 \pm 0.1$ V/Å throughout the range of T, which is within the experimental error for appreciable fluctuations of i. The slope of the plot yields $Q_n = 2.6 \pm 0.34$ eV and the preexponential $n_{hr}\nu_0 = 7 \times 10^{17}$ 1/s. Here, the frequency factor is $\nu_0 = 7 \times 10^{15}$ 1/s since the number of microprotrusions on the emitter surface n_{hr} is usually about 10^2.

The values of $\nu_0 = 10^{15}$–10^{16} l/s are typical of field evap-oration, as pointed out by Muller long ago [8]. The Arrhenius plot for Re at $U = 13$ kV and $F_{ev} = 1.92 \pm 0.08$ V/Å gives $Q_n = 2.6 \pm 0.18$ eV, $n_{hr}\nu_0 = 6 \times 10^{20}$ l/s, $\nu_0 \sim 10^{18}$ l/s, while at $U = 10$ kV, we have $Q_n = 3.7 \pm 0.75$ eV and $\nu_0 \sim 10^{18}$ l/s.

The Arrhenius plot constructed for field evaporation from the Pt tip in the range $T = 1400$–1700 K at $F = 1.1 \pm 0.14$ V/Å gave $Q_n = 3.05 \pm 0.1$ eV and the frequency factor $\nu_0 \sim 3 \times 10^{15}$ l/s.

DISCUSSION

The relatively low experimental values of Q_n are of prime interest. According to the generally accepted charge exchange model [9], Q_n is given by

$$Q_n = \lambda_0 + \Sigma_n I_n - n\varphi - 3.8\sqrt{nF} + 0.0345\alpha F^2, \quad (2)$$

where λ_0 is the heat of evaporation, $\Sigma_n I_n$ is the total ion-ization potential of a multiply charged ion, n is the ion charge, and α is the atom polarizability.

When expressed in the form of (2), Q_n is found in eV if F and α are defined in units of V/Å and Å^3, respec-tively. Since the value of α for Re atoms is unknown, we take its value for an Ir atom on the (110) face of W, $\alpha = (8 \pm 1)$ Å^3 [10], and use the tabulated values of λ_0, I_1, and φ for Re to find $Q_n = 7.64$ eV for $F_{ev} = 1$ V/Å and $Q_n = 7.05$ eV for $F_{ev} = 2.0$ V/Å, i.e., values much higher than those found experimentally (2.6–3.6 eV).

The value of Q_n was calculated on the assumption that only singly charged Re ions evaporate; i.e., $n = 1$. Generally speaking, the charge state of ions evaporat-ing at high temperatures is poorly understood: it is only known from experiments that the ion charge rapidly decreases with increasing T [11]. One can reason as fol-lows. The current of differently charged ions depends on Q_n. The temperature does not influence the charge state of an ion because it does not influence the param-eters entering into Q_n if the temperature dependence of φ is neglected. This dependence is first weak and, sec-ond, the same for differently charged ions, since it char-acterizes not an ion but a surface from which an ion evaporates. A growth of T necessitates a decrease in F for the evaporation rate to remain constant. At the same time, a decrease in F reduces the ion charge according to Eq. (2). In the absence of an external field (purely surface ionization), singly charged ions alone evaporate as a rule, because $Q_1 = \lambda_0 + I_1 - \varphi$, $Q_2 = \lambda_0 + I_1 + I_2 - 2\varphi$, and $Q_2 - Q_1 = I_2 - \varphi$ in this case. The values of Q_1 are always less than Q_2 for all elements since $I_2 > 10$ eV, while φ, on the contrary, is always less than 10 eV. Thus, the presence of an external field and the Schottky term in Eq. (2) are responsible for multiply charged ions, and the higher F, the higher the charge of an evap-orating ion. At the field evaporation of Pt at $T = 78$ K
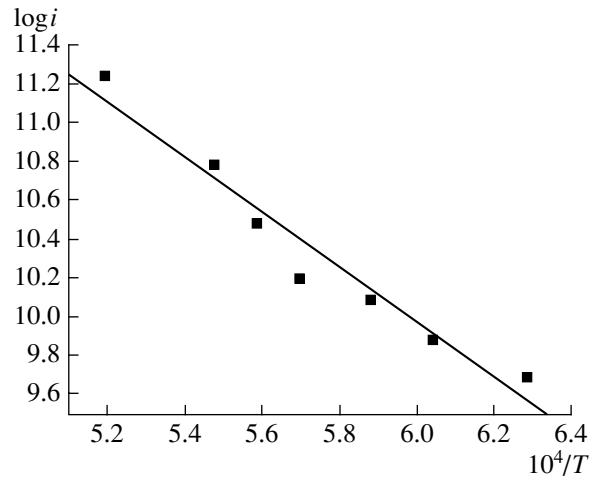

Fig. 5. Arrhenius plot $\log i = f(1/T)$ for high-temperature field evaporation of Re at $U = 12$ kV. Ion current i is given in the number of ions evaporating in a second; temperature T is given in K.

and $F = 4.8$ V/Å, we are dealing with Pt^{+2}, Pt^{+3}, and Pt$^+$ ions (in order of decreasing intensity [3]), while at the field evaporation of Pt at $T > 1000$ K and $F_{ev} = 1.2$ V/Å, singly charged Pt ions dominate [12]. Note, however, that multiply charged ions can also be produced by the postionization mechanism. Calculations carried out for this case [13] show that doubly charged Re ions account for several tenths of a percent for a field $F \approx 1.0$ V/Å. Even for $F \approx 2.0$ V/Å, their contribution to the total current is less than 10%. For $F \approx 1$ V/Å, doubly charged ions that contribute appreciably to the current are those of Ba, Ca, V, La, and some rare earths. More-over, under the assumption that doubly charged Re ions evaporate at the second ionization potential value $I_2 = 24.5$ eV, the activation energy becomes unrealistically high, $Q_n = 25$–26 eV.

With $Q_n = 7.64$ and 7.04 eV substituted into Eq. (1), the corresponding currents at $T = 1700$ K must equal $i \sim 10^{-24}$ and $\sim 10^{-22}$ A instead of the measured values $i \sim 10^{-9}$ A for the same preexponential; i.e., they must be less by many orders of magnitude. What is the rea-son for such a great discrepancy between the calculated and measured values of i and Q_n? In [14], where such a difference was observed for W evaporation, it was sug-gested that the ionization potentials I_n may be smaller than the tabulated data, since I_n for small clusters may be smaller than for a single atom. In fact, in [15], where the ionization of small Sb clusters was studied, the value of I_n was shown to decrease from 7.6 to 7.1 eV as the number of atoms in a cluster increased from 4 to 36, while for a single Sb atom, $I_n = 8.7$ eV. If Re and other metals evaporated clusterwise, the question of cluster ionization potential and the still unclear question of the heat of cluster evaporation would actually arise. The pulsed evaporation of W$_2^{+3}$, Re$_2^{+3}$, and Mo$_2^{+3}$ dimers at

$T \approx 1000$ K and $F \approx 1.8$ V/Å was observed in [16]. The dimers were only triply charged, while atomic ions were mainly singly charged and the number of the dimers was much less than that of singly charged ions. For Pt, Ni, and Rh, cluster ions were absent. A detectable number of Si cluster ions alone was observed. Therefore, it seems unlikely that cluster evaporation may contribute appreciably to the ion current in this case too, and microprotrusions cannot be considered as small clusters.

A change in the value of φ due to a strong field may also be a factor. As was pointed out in [9], a positive field, when penetrating deep into a metal, causes a decrease in the electron density near the surface, thus increasing φ. According to [9],

$$\varphi_F = \varphi_0 + neFk, \tag{3}$$

where k is the field penetration depth.

However, if Eq. (3) is valid the increase in φ is 0.5–1.0 eV for $F = 1.0$–2.0 V/Å and $k \approx 0.5$ Å, which is much less than the difference between the calculated and measured values of Q_n.

Let us also consider a possible error in determining the field F. In fact, we determine this value based on the Fowler–Nordheim theory, which implies that the field at the emitter surface is uniform. However, it was indicated (see, e.g., [17]) that an electron tunneling from the emitter surface "sees" this surface as flat starting with tip radii $r > 500$ Å. For much lesser r, which is the case for microprotrusions, such an approach fails. Therefore, applying the Fowler–Nordheim theory to microprotrusions, we must substitute some effective emitter with a uniform field for a real emitter with a nonuniform field. In this case, the values of F and current density i for the effective emitter are always less than for the real one [18]. It was shown [17] that j of the real emitter with $r = 100$ Å is 10–20 times higher than j of the effective emitter. In view of the exponential dependence of j on F, this means that the actual value of F is higher by 15–20% (at $\varphi = 5.0$ eV for Re), which changes Q_n insignificantly.

It remains to consider the last quantity in Eq. (2), the heat of evaporation λ_0, which always corresponds to atom evaporation from a kink. Evaluating λ_0 from Eq. (2) with $Q_n = 3.42$, 2.64, and 3.70 eV for $U = 12$, 13, and 10 kV, respectively, we find that $\lambda_0 = 4.75$, 4.35, and 4.8 eV, relatively; these values are much lower than the heat of evaporation for Re, $\lambda_0 = 8.3$ eV. What is the reason for the twofold decrease in λ_0? This cannot be associated with the effect of the strong field, as was suggested in [9], where even an approximate expression for the field dependence was derived:

$$\lambda_F = \lambda_0 - bF \tag{4}$$

(b is a constant).

Such a dependence is attributed [9] to field penetration into a metal and a decrease in the electron density near the surface, which loosens atom–surface bonds. However, this assumption has not been corroborated experimentally. There is only one work [19] confirming the dependence of λ_0 on F, where Rh evaporated in the field $F = 4.8$ V/Å at $T = 78$ K. However, the value of λ_F turned out to be appreciably higher than the heat of evaporation of rhodium (by 0.7–1.2 eV for different Rh faces). Subsequently, high-resolution atomic probe studies into the cryogenic field evaporation of Pt, Rh, Fe, Ni, Co, and W at $F = 3.5$–5.7 V/Å did not reveal any F dependence: λ_0 was always equal to λ_F. In our case, the fields were much lower: $F = 1.0$–2.0 V/Å; therefore, it would be unreasonable to expect a strong dependence of λ_0 on F.

As was mentioned above, the heat of evaporation λ_0 always refers to atom evaporation from a kink, where a Re atom in the hcp structure has six short-range-order neighbors and three long-range-order neighbors [22]. For the case of high-temperature field evaporation, the atom evaporates not from a kink but from the tops of the microprotrusions, where it has three short-range-order neighbors and three long-range-order neighbors; i.e., the binding energy of the atom on the microasperity top must be roughly half as large as λ_0. During the steady-state evaporation, atoms are continuously delivered to the microasperity top via the mechanism of surface migration. Thus, instead of λ_0, the values of Q_n measured include the quantity λ_0^*, which is considerably smaller than that corresponding to the bonding energy at a kink. The evaporation from the microprotrusions can also be limited by the energy of migration activation in the electric field Q_M^F (strictly speaking, this energy may be associated not only with migration but also with the formation of a microasperity as a crystalline body; for simplicity, we will consider Q_M^F alone). If $\lambda_0^* < Q_M^F < \lambda_0$, the evaporation is limited by the value of Q_M^F. The value of Q_M^F for Re is unknown; for metals close to Re in properties, Q_M^F equals 2.36 eV for W [23], 2.0 eV for Mo [24], and 1.9 eV for Ta [25] at field strengths $F = 0.1$–0.8 V/Å. In our case, the value of Q_M^F is apparently also below λ_0^*; hence, the evaporation from the microprotrusions is a "bottleneck" for the process of high-temperature field evaporation and defines the evaporation rate, i.e., the ion current. Therefore, not λ_0 but much smaller λ_0^* has to enter into the energy of activation for this process Q_n, which is determined from the Arrhenius plot. Substituting λ_0^* for λ_0 in Eq. (2), we will eliminate the differences between the calculated and measured values of Q_n and, correspondingly, the ion currents. In the case of Pt evaporation, where the experimentally found value $\lambda_F = 2.90$ eV is also much lower than the tabulated value of the heat of

evaporation $\lambda_0 = 5.3$ eV, the differences can be explained in a similar way.

Thus, high-temperature field evaporation involves the evaporation of atoms from the tops of thermal–field microprotrusions. The atoms are delivered to these tops by surface migration and leave the surface mainly in the form of singly charged ions. The energy of activation Q_n determined from the Arrhenius plot $\log i = f(1/T)$ depends not on the heat of evaporation λ_0 but on the binding energy of an atom on the microasperity top, i.e., on the heat of evaporation from a microasperity λ_0^*.

CONCLUSIONS

High-temperature field evaporation of Re, Pt, Ta, and W was studied in the range $T = 1500$–2000 K at $F = 1$–2 V/Å. For such values of T and F, the ion current i from the entire emitter surface range from several tenths of a nanoampere to several nanoamperes.

Ions evaporated mainly from the tops of microprotrusions growing on the emitter surface under the simultaneous action of high T and F.

The evaporating field F_{ev} is a complicated function of T and the applied voltage U, which reflects the complex reconfiguration of the emitter during high-temperature evaporation.

The energy of activation of high-temperature field evaporation Q_n was determined from the temperature dependence $i = f(T)$. The experimental values of Q_n are much lower than those calculated within the charge exchange model.

The difference between the calculated and experimental values of Q_n can be explained as follows. For high-temperature field evaporation, the expression for Q_n has to include not the total heat of metal evaporation λ_0 but the considerably lower value λ_0^*, which corresponds to the binding energy of an atom on the top of a microasperity from which ions evaporate.

REFERENCES

1. V. N. Shrednik, V. G. Pavlov, A. A. Rabinovich, and B. M. Shaĭkhin, Izv. Akad. Nauk SSSR, Ser. Fiz. **38**, 296 (1974).

2. V. G. Butenko, O. L. Golubev, E. L. Kontorovich, and V. N. Shrednik, Pis'ma Zh. Tekh. Fiz. **18** (8), 86 (1992) [Sov. Tech. Phys. Lett. **18**, 275 (1992)].

3. M. K. Miller and G. D. W. Smith, *Atom Probe Microanalysis: Principles and Applications to Material Problems* (Materials Research Society, Pittsburgh, 1989; Mir, Moscow, 1993).

4. *Nonincandescent Cathodes*, Ed. by M. I. Elinson (Sov. Radio, Moscow, 1974), Chap. 6, pp. 165–169.

5. V. S. Fomenko, *Electronic Properties of Materials: A Handbook* (Naukova Dumka, Kiev, 1981).

6. V. N. Shrednik, *Problems of Modern Crystallography* (Nauka, Moscow, 1975), pp. 150–171.

7. R. G. Forbes, Appl. Surf. Sci. **87/88**, 1 (1995).

8. E. W. Muller, Usp. Fiz. Nauk **77**, 481 (1962).

9. R. Gomer and L. W. Swanson, J. Chem. Phys. **39**, 2813 (1963).

10. J. Bardon and M. Audiffren, J. Phys. Colloq. **45**, C9-245 (1984).

11. E. W. Muller and T. T. Tsong, *Field Ion Microscopy: An Introduction to Principles, Experiments, and Applications* (Elsevier, New York, 1969; Metallurgiya, Moscow, 1972).

12. R. Vanselov and W. A. Schmidt, Z. Naturforsch. A **21**, 1690 (1966).

13. D. R. Kingham, Surf. Sci. **116**, 273 (1982).

14. Vu Thien Binh and N. Garcia, Ultramicroscopy **42–44**, 80 (1992).

15. D. Rayane, P. Milinon, B. Trobollet, *et al.*, J. Chem. Phys. **91**, 3100 (1989).

16. G. L. Kellog, Surf. Sci. **120**, 319 (1982).

17. H. Jun, P. H. Cutler, and N. N. Miskovsky, J. Appl. Phys. **52**, 5320 (1981).

18. A. A. Rabinovich, Surf. Sci. **70**, 181 (1978).

19. N. Ernst and Th. Jentsch, Phys. Rev. B **24**, 6234 (1981).

20. J. Lui, C. Wu, and T. T. Tsong, Surf. Sci. **246**, 157 (1991).

21. J. Lui, C. Wu, and T. T. Tsong, Phys. Rev. B **43**, 11595 (1991).

22. A. A. Chernov, *Modern Crystallography* (Nauka, Moscow, 1980), Chap. 1, pp. 8–232.

23. P. C. Bettler and F. M. Sharbonnier, Phys. Rev. **119**, 85 (1960).

24. I. L. Sokol'skaya, H. Neumann, and E. Klose, Fiz. Tverd. Tela (Leningrad) **6**, 1439 (1964) [Sov. Phys. Solid State **6**, 1126 (1964)].

25. D. M. Pautov and I. L. Sokol'skaya, Fiz. Tverd. Tela (Leningrad) **10**, 2473 (1968).

Translated by M. Astrov

Speckle Velocimeter for a Self-Powered Vehicle

A. Aliverdiev[1, 2], M. Caponero[1], and C. Moriconi[3]

[1]*Robotics and Informatics Division, ENEA Casaccia, v. Anguillarese 301, 00060 Roma, Italy*
[2]*Institute of Physics, Dagestan Scientific Center, Russian Academy of Sciences,*
ul. 26 Bakinskikh komissarov 94, Makhachkala, 367003 Dagestan, Russia
e-mail: aliverdi@mail.ru
[3]*Applied Physics Division, ENEA Frascati, v. E. Fermi 455, 00044 Frascati, Italy*
Received November 12, 2001

Abstract—Laboratory tests of a laser speckle velocimetry technique are reported. The use of integral statistics in the space–time plane is discussed. Preliminary experimental results are compared with data obtained by other speckle velocimeter methods. © 2002 MAIK "Nauka/Interperiodica".

INTRODUCTION

The purpose of this work was to develop a precision velocimeter for a self-powered robotic vehicle designed for operation in the Antarctic Regions.

This problem seems to be topical because velocity is difficult to measure with appropriate accuracy under slip conditions and in the absence of reliable reference points over vast snow or ice fields. Speckle velocimetry is expected to provide fairly accurate measurements under Antarctic conditions.

The results of early theoretical and experimental studies on the problem are very encouraging.

THE HISTORY OF THE PROBLEM

The use of the speckle effect in velocimetry has been repeatedly discussed by the scientific community. Several approaches that rely on both time and space statistics have been suggested. Differential and integral (with respect to time) intensity functions have been used.

In 1980, Fercher [1] suggested using the temporal differential statistics of the differential (with respect to time) intensity function. He succeeded in gaining fairly accurate results for the velocity of motion of an object made of plexiglass. The velocities were varied from 1 μm/s to 1 mm/s. Fercher used simple and descriptive formulas that relate the velocity measured and statistical values found experimentally:

$$v^2 = \frac{\lambda d}{8} \frac{R_i(0)}{R_I(0)} \qquad (1)$$

without the coherent reference beam and

$$v^2 = \frac{\lambda d}{8} \frac{R_i(0)}{R_I(0)} - \frac{\pi^2 I_r}{\langle I \rangle + 2I_r} (\Delta v)^2 \qquad (2)$$

with the coherent reference beam. Here, $R_I(\tau) = \langle I(t)I(t + \tau) \rangle - \langle I(t) \rangle \langle I(t + \tau) \rangle$, $R_i(\tau) = \langle \dot{I}(t)\dot{I}(t + \tau) \rangle - \langle \dot{I}(t) \rangle \langle \dot{I}(t + \tau) \rangle$, $I(t)$ is the intensity, $\dot{I}(t)$ is the derivative of the intensity with time, I_r is the intensity of the reference beam, λ is the wavelength, d is the laser beam diameter in the direction of motion, and Δv is the Doppler broadening of the scattered light.

However, the quantities appearing in formulas (1) and (2) are very inconvenient for calculation; therefore, Fercher suggested using analog computers.

In works devoted to vibration analysis (e.g., [2]), the temporal statistics of the differential function was used in a somewhat different manner. However, the statement of the problem in those works differed from ours; therefore, we will not enlarge upon it.

The use of the spatial statistics of the integral (with respect to time) intensity function was described in [3]. A simple expression was derived that relates the rms spatial deviation of the speckle intensity function within the time of averaging, the time of averaging, and the characteristic correlation time:

$$\frac{\sigma_s}{\langle I \rangle} = \sqrt{\left(\frac{\tau_c}{2T}\right)\left(1 - e^{-\frac{2T}{\tau_c}}\right)}, \qquad (3)$$

where σ_s is the rms spatial deviation of the speckle intensity function within the time of averaging, T is the time of averaging, and τ_c is the correlation time. The correlation time here is the time that directly defines the velocity of motion of speckle-producing diffusers.

Subsequently, this method was refined both mathematically and experimentally [4–7]. Its basic advantage is its simplicity and clearness of information gained. With this method, one can easily estimate the spatial distribution of diffusers and find regions where the velocity of motion is higher or lower. In addition, the
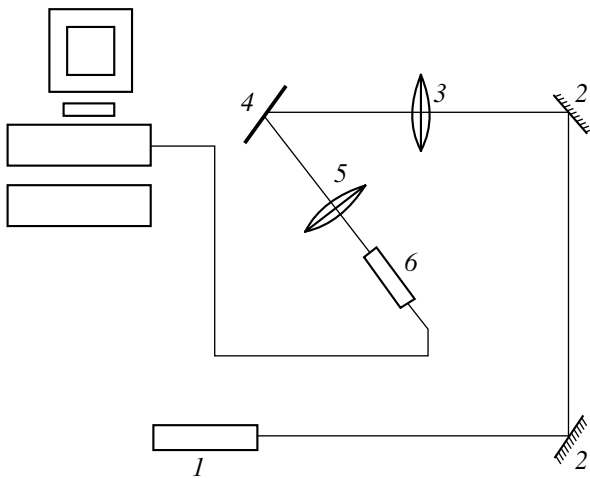
Fig. 1. Schematic diagram of the laboratory setup.

mathematical apparatus developed provides a relatively high accuracy.

A disadvantage of the method is that it is impossible to determine the direction of object motion. Moreover, one cannot establish whether translational motion is observed or vibration. However, this disadvantage naturally becomes an advantage when the velocity of motion of a liquid is found (in this case, diffusers are arranged randomly instead of being regularly spaced). This method is successfully used in medical diagnostics; however, it can hardly be considered optimal as applied to our purpose.

STATEMENT OF THE PROBLEM

From the aforesaid, it follows that speckle velocimetry has proved itself a promising technique but calls for additional investigation in order to be applied in determining the velocity of a self-powered vehicle. This work will hopefully shed more light on the potentiality of speckle velocimetry.

Speckle velocimetry offers a sufficiently high accuracy and can be applied when a reference point is extremely difficult (if possible) to select. The vehicle is supposed to have two, speckle and Doppler, velocimeters, so that comparative analysis of the two measuring techniques can be carried out.

In this work, we developed software tools for flat-screen speckle velocimetry and performed tentative laboratory experiments (Fig. 1).

SOLUTION OF THE PROBLEM WITH THE USE OF INTEGRAL STATISTICS

Figure 2 shows a speckled image of an object made of foamed plastic, whose surface closely simulates the surface of ice.

Consider now the time-resolved image, where only the space coordinate is left (Fig. 3). Clearly, the velocity of the object is determined from the slope of time traces for each speckle. This gives a chance to find the velocity directly without statistical calculations. However, the distinct separation of the traces of individual speckles for subsequently determining the slope is a nontrivial procedure even in Fig. 3, which was obtained under nearly perfect laboratory conditions. Moreover, even
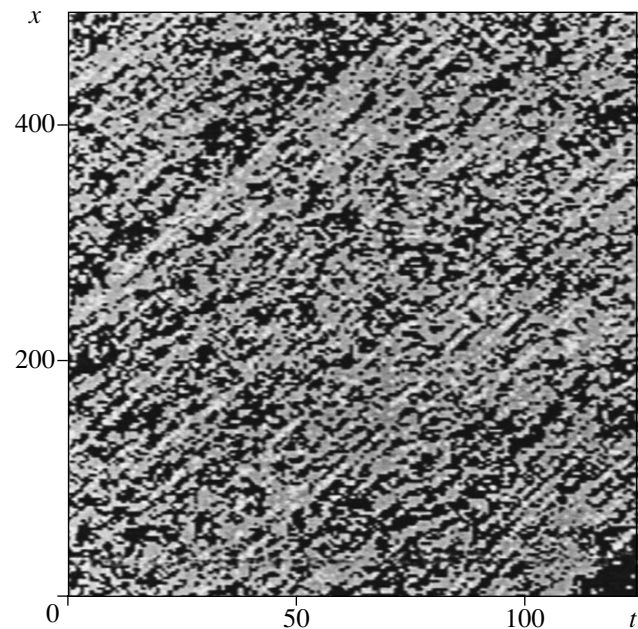

Fig. 2. Speckle photograph. Resolution 512×512 pixels. The vertical scale division is 1 mm.

Fig. 3. Time-resolved speckle picture. On the axes, the number of pixels are plotted. $v_0 = 0.732$ mm/s.
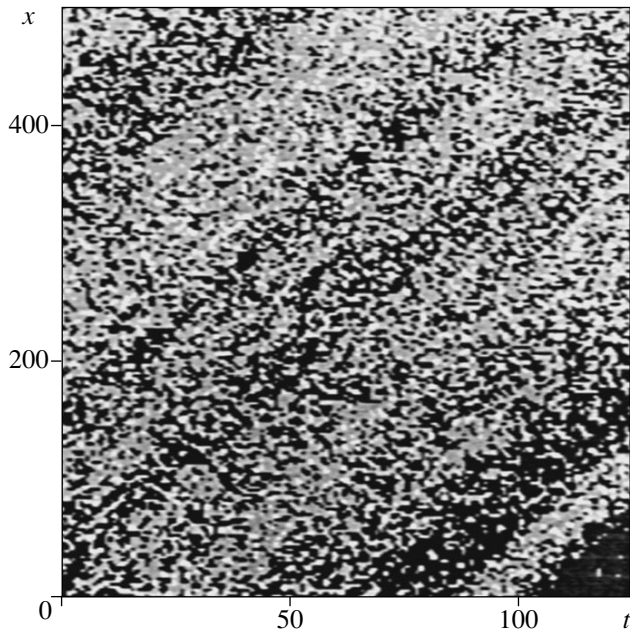
Fig. 4. Time-resolved speckle picture. On the axes, the number of pixels are plotted. $v_0 = 0.732$ mm/s. Vibration in the orthogonal direction.
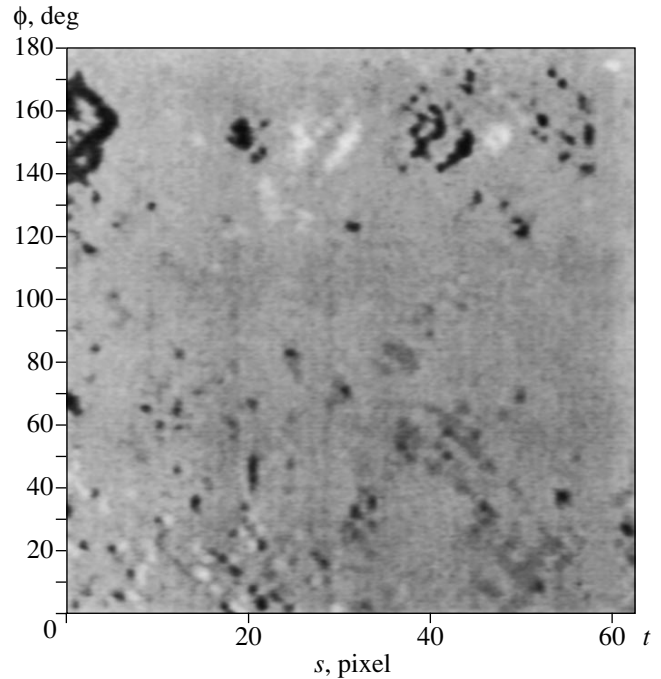

Fig. 5. Radon transform of time-resolved speckle picture. $v_0 = 0.732$ mm/s. Vibration in the orthogonal direction.

insignificant noise, inevitable under vibration (Fig. 4), may become catastrophic for such a "direct" method. That is why this simplest approach has not received wide recognition.

Let us apply the statistical method to our problem. Preference was given to the analysis of the integral speckle intensity function because it is more stable against intense noise, which may arise in our case. The data represented in Fig. 3 (with a full set along the *y* coordinate in the general case) suffice to perform Briers analysis, i.e., to find a set of rms deviations of the speckle intensity function for a set of times of averaging and then, having calculated τ_c, find the desired velocity. In so doing, however, we lose some data used in the direct method. Therefore, we will slightly modify the analysis of the integral statistics. Namely, instead of several averagings over different periods *T*, it is proposed that the averaging be done over one period and over a set of angles in the space–time plane [8]. Such a procedure is similar to the Radon transformation [9] with the fixed limit in the quasi-uniform plane (Fig. 5). Each straight line along which we perform averaging will have normal coordinates (s, ϕ); that is, we find the function

$$g(s, \phi) = \int_{-Tv_0}^{Tv_0} I\left(s\cos(\phi) - p\sin(\phi),\right.$$

$$\left.\frac{s}{v_0}\sin(\phi) + \frac{p}{v_0}\cos(\phi)\right)dp. \quad (4)$$

Here, the parameter *s* is an analog of a space parameter over which one can calculate the spatial rms deviation of the speckle intensity function:

$$\sigma_s^2(\phi) = \langle(g(s, \phi) - \langle g(s, \phi)\rangle_s)^2\rangle_s. \quad (5)$$

The rms deviation is maximal for the angle that corresponds to the object velocity being measured:

$$v = v_0\tan\left(\phi((\sigma_s)_{max}) - \frac{\pi}{2}\right), \quad (6)$$

where v_0 is the time-to-spatial resolution ratio of a TV camera ($v_0 = 0.732$ mm/s in our case).

The value of v_0 specifies a measuring range of velocity with a fairly good accuracy. Roughly, this range spans the interval from $0.1v_0$ to $10v_0$. At high *v*, the tangent tends to infinity and the accuracy drops. At very low velocities, the relative error increases because of a decrease in the absolute value appearing in the denominator.

Thus, to improve the accuracy at very high or very low velocities, one should either decrease the spatial resolution (it is bounded from below by the size of a speckle, which must exceed that of a pixel) or increase the time resolution (it is limited by the parameters of the TV camera and real-time computing systems).

EXPERIMENTAL SETUP

The schematic diagram of the experimental setup is shown in Fig. 1. A beam from Siemens LGK 7621

Table 1

Number of averagings N	1 Without averaging	2	4
V	0.877	0.868	0.849
ε, %	2.7%	1.7%	0.5%

Table 2

Number of averagings N	1 Without averaging	2	4
V	0.371	0.882	0.846
ε, %	56%	3.3%	0.8%

Table 3

Number of averagings N	1 Without averaging		2		4	
	(x, t)	(t, x)	(x, t)	(t, x)	(x, t)	(t, x)
V	−0.51	1.253	0.88	0.883	0.844	0.848
ε, %	160%	47%	3.1%	3.4%	1.0%	0.7%
$|\langle \phi_{(x, t)} + \phi_{(t, x)} - 270 \rangle|$, deg	68.5		0.1		0.2	

(Q-4001-K7 654) He–Ne laser *1* reflects from mirrors *2*, passes through lens *3*, and strikes moving object *4*. The object, a screen made of foamed plastic, is mounted on a support, which is capable of displacing the object in three directions with a micrometer accuracy.

The light scattered by the screen is incident on lens *5*, which projects it onto JVC-TK-S350EG video camera *6*, coupled with an IBM-compatible computer via an IMAQ PCI-1408 grabber. The acquisition, storage, and preliminary data processing are performed with a dedicated C++-based software package using IMAQ Lab VIEW libraries (National Instruments).

The experiments simulated motion; that is, all displacements were made in steps and were recorded separately. It was assumed that the frame frequency meets the TV standard: 25 frames per second. Since speckles are defined by their position and do not depend on time and velocity, such a simulation is adequate to the situation and the velocity can be estimated from the parameters given very accurately.

The total experimental error in this case is the sum of the relative accuracy of displacement (associated with the displacing tools) and the relative accuracy of displacement measurement by the TV camera. It varied from 0.1 to 0.5%.

RESULTS AND DISCUSSION

In the experiments, we varied the object velocities, screens, and experimental conditions (the amplitude and frequency of orthogonal vibration, etc.).

The general results are as follows: for velocities up to 2.5 mm/s (0.1-mm displacement per frame), the relative measurement error was less than 1% without vibration and less than 2% when severe orthogonal vibration (exceeding really possible values for the vehicle by several orders of magnitude) was simulated; for a velocity of 12.5 mm/s (0.5-mm displacement per frame), the error was less than 5%.

The error for a velocity of 12.5 mm/s (and for greater velocities, 5–10 cm/s) can be decreased in real experiments (this relatively large error is due largely to the scarcity of temporal data gathered in this laboratory experiment because of the restrictions associated with the micrometer displacing means). A further increase in the velocity to be measured is related to a rise in the time resolution. In practice, good real-time results are expected for velocities of up to several meters per second.

Consider a number of experimental findings in greater detail.

(i) The real velocity measured was $v_{real} = 2.5$ mm/s; spatial resolution, 512 pixels/15 mm; time resolution, 25 pixels/s. Thus, $v_0 = 0.732$ mm/s and $v_{real}/v_0 = 3.413$.

In addition, we averaged the data over 32 straight lines orthogonal to the y axis. After processing the data with the above technique, we found $v_{exp}/v_0 = 3.508$.

The accuracy of the measurements was 2.7%. This value exceeds the experimental error but remains quite appropriate for applications.

To improve the measurement accuracy, we modified the data processing algorithm by introducing spatial averaging along the x axis and thereby shrinking the spatial resolution. Averaging was performed over two and four points. The results are summarized in Table 1.

Upon averaging, the spatial resolution drops and v_0 grows. As was mentioned before, the accuracy is the highest at $v/v_0 \approx 1$; therefore, the velocities in Table 1 are given in terms of unity for the number of averagings $N = 4$ (it was assumed that $N = 4$ will provide the best accuracy), and all the relative velocities measured are given in the same units ($V = v/v_0 N/4$). The real relative velocity was 0.853. It can be seen from Table 1 that the error after the fourth averaging is no more than 0.5%, i.e., does not exceed the experimental error.

(ii) In another experiment, the velocity of another part of the screen was measured (Table 2). The real

Table 4

| Number of averagings N	1 Without averaging	2	4
V	0.371 (0.769)	0.692 (0.849)	0.87 (0.869)
ε, %	48% (8.4%)	16% (0.4%)	1.7% (1.6%)

velocity was the same as in experiment (i). Without averaging, the measurement accuracy drops drastically. The reason becomes clear if we construct the dependence of σ_s on angle ϕ for each y (Fig. 6). For a number of y, spurious maxima at zero ϕ are higher than maxima due to the measured velocity.

In this case, the measurement accuracy can be improved by modifying the maximum search procedure. It is necessary to perform the same integral transformation but with the mirror-symmetric choice of the coordinates x and t and compare the results. It is easy to check that maxima on the associated diagrams in the (x, t) and (t, x) coordinates obey the condition

$$\phi(x, t) + \phi(t, x) = \frac{3}{2}\pi. \tag{7}$$

Maxima that do not meet this relationship (spurious maxima) must be rejected.

In Tables 1 and 2, the averages of the velocities calculated in the coordinates (t, x) and (x, t) are given. Table 3 lists the velocities obtained in the coordinates (x, t) and (t, x) separately (their averages are listed in Table 2).
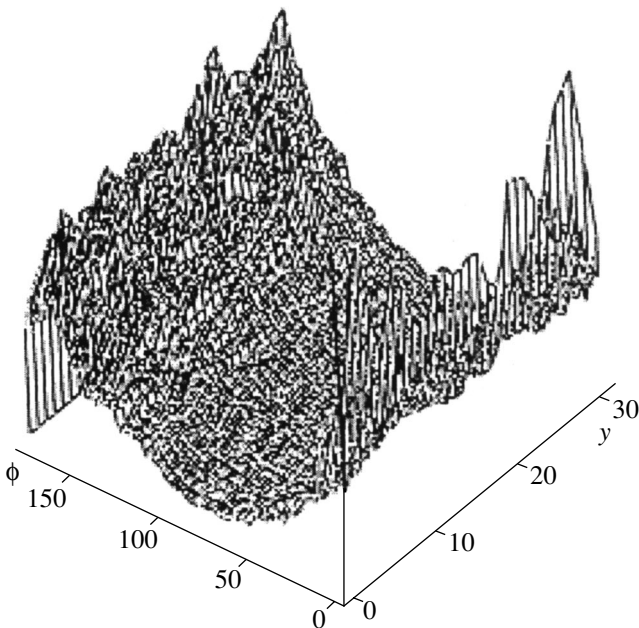
Fig. 6. σ_s vs. ϕ (degrees) and y (pixels) for example (ii).

Since the calculation procedures for the coordinates (x, t) and (t, x) are nearly the same, they can (and must) be combined into a single calculation cycle. This will improve the accuracy of final results and reject artifacts, without requiring much computational resources.

(iii) Consider one more experiment illustrated in Fig. 3.

The velocity was the same as in the two previous experiments, but the moving screen vibrated in the direction orthogonal to its plane with a frequency of 0.20 ± 0.05 relative frequency units and an amplitude of 1 relative amplitude unit (in other words, the velocity in the orthogonal direction was five times higher than the velocity measured). The associated data are summarized in Table 4 (results obtained after rejecting spurious maxima according to (4) are given in parentheses).

Table 4 indicates than even severe vibration deteriorates the accuracy insignificantly. The value of 1.6% is quite appropriate for applications.

CONCLUSION

Thus, our experimental data proved to be promising for practical use. We are going to design a field model of the velocimeter and perform a series of field tests.

Some remarks concerning the measurement of the angular velocity: it can be determined by the time averaging of the spatial contrast on the full (x, y) plane (Briers method). In the case of translational motion, the spatial contrast on the full image is uniform (does not depend on spatial coordinates). In the presence of rotation, the velocity of different object points is different. One can first estimate the angular velocity ω and then calculate it with a great accuracy by comparing with the translational velocity.

REFERENCES

1. A. F. Fercher, Opt. Commun. **33**, 129 (1980).
2. N. Takai, Opt. Commun. **25**, 31 (1978).
3. A. F. Fercher and J. D. Briers, Opt. Commun. **37**, 326 (1981).
4. A. F. Fercher, M. Peukert, and E. Roth, Opt. Eng. **25**, 731 (1986).
5. J. D. Briers and S. Webster, Opt. Commun. **116**, 36 (1995).
6. G. Richards and J. D. Briers, Proc. SPIE **2981**, 160 (1997).
7. J. D. Briers, G. Richards, and X. W. He, J. Biomed. Opt. **4**, 164 (1999).
8. M. Caponero, C. Moriconi, and A. Aliverdiev, in *Technical Digest of the Forth Italian–Russian Laser Symposium, St. Petersburg, 2001*, p. 122.
9. A. Aliverdiev, Zh. Tekh. Fiz. **67** (9), 132 (1997) [Tech. Phys. **42**, 1102 (1997)].

Translated by V. Isaakyan

BRIEF
COMMUNICATIONS

Solid-Phase Reaction of Aluminum with the Hexagonal and Cubic Phases of Cobalt in Film Systems

V. G. Miagkov, L. E. Bykova, G. N. Bondarenko, and F. V. Miagkov

*Kirenskiĭ Institute of Physics, Siberian Division, Russian Academy of Sciences,
Akademgorodok, Krasnoyarsk, 660036 Russia*

e-mail: miagkov@iph.krasn.ru

Abstract—Solid-phase reactions taking place in Al/β-Co/MgO(001) and Al/α-Co/MgO(001) film systems under conditions of self-propagating high-temperature synthesis (SHS) are investigated. In both systems, SHS products exhibit the single CoAl superstructure, which epitaxially grows on the MgO(001) surface in the Al/β-Co/MgO(001) structures and has a fine-dispersed disordered structure in the Al/α-Co/MgO(001) films. It appears that the difference in the reagent structure has an effect on the energy of activation but does not change the SHS initiation temperature and the temperature at which the initial phase involved in the reaction products nucleates. It is shown that the SHS initiation temperature in the Al/β-Co/MgO(001) and Al/α-Co/MgO(001) systems coincides with the temperature of CoAl superstructure ordering. © 2002 MAIK "Nauka/Interperiodica".

Thin-film components are a building block of today's microelectronic devices. During fabrication and operation, these components are subjected to temperature effects, which cause solid-phase transformations and thus change the performance of the devices. That is why solid-phase reactions in thin films have been the subject of extensive investigation [1–3]. It has been repeatedly shown that upon heating bi- and multilayer films, the initial (first) phase of reaction products appears at some temperature. As the temperature is increased further, new phases form sequentially (so-called phase sequence) [1–3]. Grain-boundary diffusion is believed to be the basic mechanism of solid-phase reactions. Therefore, the energy of activation and the rate of a solid-phase reaction must depend on the film structure and crystallographic orientation. For example, early in $NiAl_2O_4$ spinel formation at the NiO–single-crystal Al_2O_3 interface, the orientation of the Al_2O_3 substrate specifies the energy of activation, rate, and even kinetic law of the process [4, 5].

At high heating rates, the initial stage of many solid-phase reactions in thin films has been shown [6, 7] to proceed under SHS conditions. In thin films, SHS may be considered as a surface combustion wave and is characterized by the initiation temperature T_0. The effect of reagents with different polymorphic modifications on the reaction product structure, the formation of the first phase, and the SHS initiation temperature is of great scientific and applied interest.

In this work, we study solid-phase reactions under the SHS conditions in Al/β-Co/MgO(001) and Al/α-Co/MgO(001) film systems. The total thickness of the films was no more than 100 nm. It is known that the low-temperature hexagonal structure of bulk α-Co is stable up to 420°C. Above this temperature, it transforms into the fcc phase of β-Co by the martensitic mechanism. However, in cobalt thin films and nanocrystals, the β-Co phase is also stable at room temperature under various process conditions [8]. The structure and modification of Co films on the MgO(001) surface depend strongly on the substrate temperature T_s. For $T_s = 220–240°C$, the high-temperature fcc cubic phase of β-Co grows epitaxially on the MgO(001) surface. The first constant of magnetic anisotropy was found to be $K_1 = -(6–7) \times 10^5$ erg/cm^3. Diffraction reflections (Fig. 1a) and the arrangement of the easy axes show that the β-Co film and the MgO substrate form an epitaxial system [100](001)β-Co∥[100](001)MgO. The situation changes radically when the substrate temperature rises to $T_s = 370–400°C$. As follows from diffraction reflections (Fig. 2), an α-Co layer grows in this case so that its (110) plane is parallel to the (001)MgO surface. The reflections (110)α-Co and (110)β-Co coincide. Electron microscopy studies [9] indicate, however, that Co films grown at high temperatures consist of α-Co grains with their c axis aligned with the [100] and [010] directions in MgO. The effective constant of bilayer magnetic anisotropy for these films is $K_{eff} = (1.1–1.3) \times 10^6$ erg/cm^3. The magnetic anisotropy energy per unit volume E_{an} of the hexagonal crystal (with the in-plane anisotropy neglected) is given by [10] $E_{an} = K_1\sin^2\varphi + K_2\sin^4\varphi + \ldots$, where $K_1 = 4.3 \times 10^6$ erg/cm^3 and $K_2 = 1.2 \times 10^6$ erg/cm^3 for hexagonal cobalt (φ is the angle between the c axis and the magnetization vector).
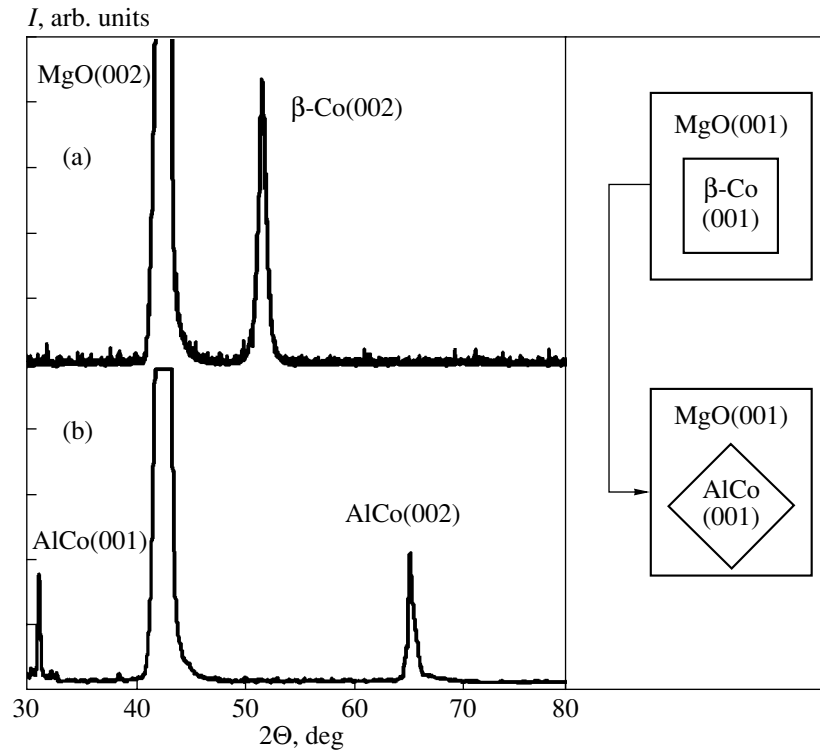
Fig. 1. Diffraction patterns from the Al/β-Co/MgO(001) film structure (a) before and (b) after the passage of the SHS front. Crystallographic epitaxial relations of the MgO(001) substrate and the initial β-Co phase and the CoAl phase of the reaction product are shown on the right of the figure.

Assuming that the α-Co grains whose *c* axes are aligned with the [100] and [010] directions of MgO experience exchange coupling and their volumes are the same, we calculated the effective bilayer magnetic anisotropy of such a film system and found that $K_{eff} = K_2$. The coincidence between K_{eff} and K_2 found experimentally also confirms the formation of the α-Co grains on the MgO(001) surface. An aluminum layer with a thickness equal to that of the Co film was deposited on the initial β-Co/MgO(001) and α-Co/MgO(001) film systems. The absence of aluminum reflections from the as-prepared Al/Co/MgO(001) structures (Figs. 1a, 2) implies that the aluminum film is amorphous or fine-grained. The Al/β-Co/MgO(001) and Al/α-Co/MgO(001) samples were heated to the SHS initiation temperature T_0 with a rate of 20 K/s at a pressure of 10^{-4} Pa. The appearance of the SHS front was observed visually. The initiation temperature T_0 lay in the 750–780 K interval. The discrepancy in the T_0 values is associated with heat removal into the substrate, which depends on the substrate thickness and heating rate. However, the values of T_0 for the Al/β-Co/MgO(001) and Al/α-Co/MgO(001) systems coincide within the experimental error. To exclude the Headwall effect, we experimentally checked that solid-phase transformations and α $\rightleftarrows$ β transitions in the initial β-Co and α-Co films do not occur before the initiation temperature. The reaction products from the Al/β-Co/MgO(001) system contain the ordered AlCo

phase that grows epitaxially on the MgO(001) surface (Fig. 1b). This phase has a cubic lattice with spacing $a = 0.286$ nm. The lattice mismatch between MgO and AlCo is the smallest if the orientation relation is (001)[100]AlCo‖(001)[110]MgO (Fig. 1b). After SHS in the Al/β-Co/MgO(001) system, the AlCo phase may grow with its (110) plane parallel to the (001)MgO surface, although such cases are untypical. In the Al/α-Co/MgO(001) system, diffraction reflections after SHS are absent. However, phase analysis of bi- and multilayer polycrystalline Al/Co films indicates the presence of the AlCo phase alone in the reaction products. From the aforesaid and the equality of the initiation temperatures, it follows that the SHS products in the Al/α-Co/MgO(001) system also contain the AlCo phase that grows randomly on the (001)MgO surface and is amorphous or fine-grained. The different crystal structure of the SHS products in the Al/α-Co/MgO(001) and Al/β-Co/MgO(001) systems implies different reaction mechanisms, which are governed by the crystallographic planes containing the reagents.

The results obtained lead us to conclude that polytypic forms of the reagents play an insignificant role in the formation of the first phase during solid-phase SHS. The energy of activation E_a of the solid-phase reaction must include the binding energy and the energy of reagent transfer. In the reactions considered, the transfer energies are the same. Hence, the difference in the energies of activation of the solid-phase SHS reactions
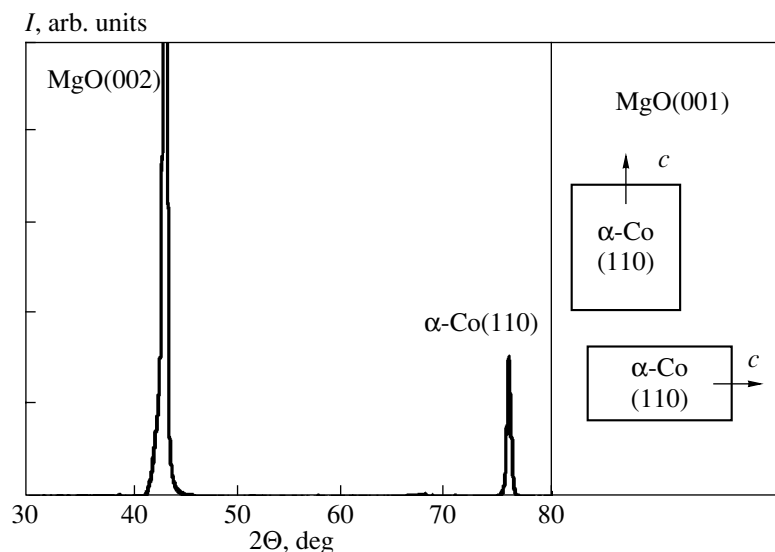
Fig. 2. Diffraction patterns from the Al/α-Co/MgO(001) film structure before the SHS reaction. Crystallographic epitaxial relations of the MgO(001) substrate and the initial α-Co phase are shown on the right of the figure.

taking place in the Al/β-Co/MgO(001) and Al/α-Co/MgO(001) systems equals that in the binding energies of cobalt in the hexagonal and cubic modifications. This difference is insignificant because the enthalpy $\Delta H^{\beta \to \alpha} = -220$ cal/mol is small [11]. The SHS front velocity V_f depends on the temperature T_s by the near-Arrhenius law: $V_f \sim \exp(-E_a/kT_s)$; therefore, a change in the energy of activation will insignificantly influence only the temperature dependence of the front velocity. This means that the low-energy difference in the reagent structures affect the SHS basic parameters in thin films (initiation temperature and the formation of the first phase) only slightly. Specifically, the SHS parameters in amorphous films may not differ from those in polycrystalline and single-crystal layers, since the enthalpy of amorphous–crystalline transition is low (≈ 1000 cal/mol for most alloys). These conclusions agree with those drawn in [12], where the solid-phase reaction at the titanium–aluminum interface was shown to form a single $TiAl_3$ phase irrespective of whether fine- or coarse-grained aluminum was used as a substrate. The formation of the $TiAl_3$ phase was also insensitive to the sequence of layer deposition in the Ti/Al bilayer system.

The prediction of the onset of the first phase (or of the SHS initiation temperature T_0) is of great importance for applications. At the initiation temperature, solid-phase transformations are accompanied by bond breaking and atomic rearrangements with the formation of new bonds. Therefore, as the annealing temperature rises, a phase having the lowest temperature of solid-phase transformation must appear first at the interface between reagents in a bilayer system. For example, in a S/Fe bilayer system, the temperature T_0 of initiating the solid-phase reaction producing FeS iron monosulfide and the temperature of metal–insulator phase transition

were shown to coincide [13]. In Cu/Au thin bilayer films (the Cu–Au system is a model system for studying ordering in alloys), the SHS initiation temperature T_0 was shown [14] to coincide with the Kurnakov temperature T_K of the CuAu ordered alloy, which is a reaction product: $T_0(Cu/Au) = T_K(CuAu)$. Later [15], this equality was confirmed for Al/β-Co/MgO(001) and other bilayer systems having superlattices in their reaction products.

Thus, the study of SHS in Al/hexagonal Co and Al/cubic Co bilayer structures revealed the following.

(1) In the thin films, the polytypic modifications of the reagents and the low-energy difference in their structures do not influence the SHS initiation temperature and the first phase of the reaction products.

(2) In the film reagents, the SHS initiation temperature correlates with the Kurnakov temperature of the phase forming in the reaction products provided that there are no solid-phase transformations preceding the order–disorder transition.

(3) Synthesis mechanisms, which are responsible for the reaction product microstructure, depend on the crystal orientations of reacting pairs.

ACKNOWLEDGMENTS

REFERENCES

1. *Thin Films: Interdiffusion and Reactions*, Ed. by J. M. Poate, K. Tu, and J. Meier (Wiley, New York, 1978; Mir, Moscow, 1982).

2. U. A. Clevenger, B. Arcort, W. Ziegler, *et al.*, J. Appl. Phys. **83**, 90 (1998).

3. M. Zhang, W. Yu, W. Wang, *et al.*, J. Appl. Phys. **80**, 1422 (1996).

4. P. G. Kotula and C. B. Carter, Phys. Rev. Lett. **77**, 3367 (1996).

5. P. G. Kotula, M. T. Johnson, and C. B. Carter, Z. Phys. Chem. **207**, 39 (1998).

6. V. G. Myagkov and L. E. Bykova, Dokl. Akad. Nauk **354**, 777 (1997).

7. V. G. Myagkov, V. S. Zhigalov, L. E. Bykova, *et al.*, Zh. Tekh. Fiz. **68** (10), 58 (1998) [Tech. Phys. **43**, 1189 (1998)].

8. H. Sato, O. Kitakami, T. Sakurai, *et al.*, J. Appl. Phys. **81**, 1858 (1997).

9. A. S. Komalov, Candidate's Dissertation (Inst. Fiz., Krasnoyarsk, 1970).

10. S. V. Vonsovskiĭ, *Magnetism* (Nauka, Moscow, 1971; Wiley, New York, 1974).

11. *Thermodynamic Constants of Substances* (Nauka, Moscow, 1972), No. 6.

12. X.-A. Zhao, F. C. T. So, and M.-A. Nicolet, J. Appl. Phys. **63**, 2800 (1988).

13. V. G. Myagkov, L. E. Bykova, G. N. Bondarenko, *et al.*, Dokl. Akad. Nauk **371**, 763 (2000) [Dokl. Phys. **45**, 157 (2000)].

14. V. G. Myagkov, L. E. Bykova, G. N. Bondarenko, *et al.*, Pis'ma Zh. Éksp. Teor. Fiz. **71**, 268 (2000) [JETP Lett. **71**, 183 (2000)].

15. V. G. Myagkov, L. E. Bykova, G. N. Bondarenko, *et al.*, Zh. Tekh. Fiz. **71** (6), 104 (2001) [Tech. Phys. **46**, 743 (2001)].

Translated by V. Isaakyan

BRIEF
COMMUNICATIONS

Generation of Electrohydrodynamic Waves
at the Liquid–Vacuum Interface

A. A. Shutov

Karpov Institute of Physical Chemistry (Obninsk Branch),
Obninsk, Kaluga oblast, 249030 Russia

e-mail: fci@meteo.ru

Received October 26, 2001; in final form, January 31, 2002

Abstract—The linear electrodynamic problem of the stability of a viscous liquid surface at the liquid–air interface is studied in view of the surface charge kinetics. The electric field critical intensities which break the stability of the plane interface are determined. Conditions for generating sustained waves at the interface are found. © *2002 MAIK "Nauka/Interperiodica".*

INTRODUCTION

The electrostatic approximation describing the behavior of perfect conductors and insulators is also applicable to a nonviscous liquid [1, 2]. In contrast, the problem of the motion of a finite-conductivity liquid should be considered in the viscous statement. Charges generated at the interface cause interfacial tangent electric forces, which must be compensated for by viscous tensions. The surface charge has to match convective fluxes and ohmic currents. In this case, not only the values of the currents may be of great importance, but also the nature of the ohmic coupling between the interface and charge sources. In particular, experiments [3, 4] yield different critical intensities of plane-surface destabilization depending on the way in which the voltage is applied (an abrupt or gradual increase in the field intensity). Here, the instability is treated as the response of a liquid system to a small perturbation of the interface with the resulting spontaneous growth of the perturbation and emission of some amount of the liquid from the free surface. The analysis of the interface between liquids with equal kinematic viscosities revealed the presence of unstable and decaying perturbations [5]. It turns out that under certain conditions a viscous system may excite sustained surface waves. For instance, in an electric field tangential to the plane interface, a standing wave is excited at an intensity much smaller than critical [6].

We consider the problem of the stability and generation of sustained waves in a field perpendicular to the interface. In contrast to [3, 5], the interface is assumed to be initially charged.

THEORY

The kinematic equation for a surface charge of density σ on a surface $q_3 = f(q_1, q_2, t)$ has the form [7]

$$\frac{\partial}{\partial t}(\sigma\sqrt{g_1 g_2}) + \frac{\partial}{\partial q_1}(\sigma V_1 \sqrt{g_2}) + \frac{\partial}{\partial q_2}(\sigma V_2 \sqrt{g_1}) \tag{1}$$
$$+ \sqrt{g_1 g_2}(\lambda_2 E_{2n} - \lambda_1 E_{1n}) = 0,$$

where the orthogonal curvilinear coordinate system (q_1, q_2, q_3) on the interface is chosen such that the unit vector of the coordinate q_3 is directed along the normal to the surface from medium 1 to medium 2, and q_1 and q_2 form an orthogonal grid.

The charge density is defined in the conventional electrodynamic form through the volume charge density ρ_e:

$$\sigma = \int_{f-\delta}^{f+\delta} \rho_e h_3 dq_3.$$

Here, the metric coefficients of the surface are related to the Lamé coefficients h_i of Cartesian-to-curvilinear coordinate conversion as $g_i(q_1, q_2) = h_i^2(q_1, q_2, q_3 = f)$, $i = 1$ and 2. In this relationship, V_1 and V_2 are the q_1 and q_2 components of the liquid velocity at the interface; λ_1 and λ_2, the conductivities; and E_{1n} and E_{2n}, the normal components of the electric field intensity on both sides of the interface. Consider the plane motion of the liquid occupying the lower half-space in the Cartesian coordinates x and z in a uniform electric field. The free surface is described by the equation $z = f(x, t)$. The gravity force **g**

is directed oppositely to the z axis. The field intensity over the liquid is E_{20}. We assume that charge transfer occurs in the absence of the volume charge. The problem is stated as

$$\partial \mathbf{V}/\partial t + (\mathbf{V} \cdot \nabla)\mathbf{V} = -\nabla p/\rho + \nu\Delta\mathbf{V} + \mathbf{g}, \quad \nabla \cdot \mathbf{V} = 0, \tag{2}$$

$$\Delta\varphi_{1,2} = 0 \tag{3}$$

(Δ is the Laplacian operator) with the conditions on the boundary $z = f$

$$\partial f/\partial t - w + u\partial f/\partial x = 0, \tag{4}$$

$$\partial(\sigma\sqrt{1 + f'^2})/\partial t + \partial(\sigma u)/\partial x - \lambda E_{n1}\sqrt{1 + f'^2} = 0, \tag{5}$$

$$\varphi_1 = \varphi_2, \quad \varepsilon\mathbf{n} \cdot \nabla\varphi_1 - \mathbf{n} \cdot \nabla\varphi_2 = \sigma/\varepsilon_0, \tag{6}$$

$$\mathbf{p}_{n1} = \mathbf{n}Tf''/(1 + f'^2)^{3/2} + \mathbf{F}_e, \tag{7}$$

$$\mathbf{E}_e = (D_{2n}\mathbf{E}_2 - D_{1n}\mathbf{E}_1) - (D_2E_2 - D_1E_1)\mathbf{n}/2, \tag{8}$$

$$z \longrightarrow \infty: E_{2z} \longrightarrow E_{20}.$$

The components of the velocity $\mathbf{V}$ are denoted in view of the correspondence $(u, w) \longrightarrow (x, z)$. The density, kinematic viscosity, surface tension, and permittivity are designated by ρ, ν, T, and ε, respectively. The viscosity, density, pressure, and conductivity of the medium above the liquid are zero. The electric potential is denoted by φ; $\mathbf{E} = -\nabla\varphi$; and $D = \varepsilon\varepsilon_0 E$, where ε_0 is the absolute dielectric constant. Equation (5) for the surface charge results from Eq. (1) and takes into account both convective charge transfer over the surface and the ohmic current of the liquid. Subscripts 1 and 2 at the potential correspond to the values inside and outside the liquid, respectively. The vector $\mathbf{n}$ is the outer normal to the liquid, and prime means differentiation with respect to x. The vector $\mathbf{p}_{n1}$ in boundary condition (7) is the viscous stress vector. Relation (8) is the surface density of the electric force [8].

Equations (2)–(8) admit the following solution for the plane interface $z = 0$:

$$\mathbf{V}_0 = 0, \quad f_0 = 0, \quad \varphi_{20} = -E_{20}z, \tag{9}$$

$$\varphi_{10} = zE_{20}\exp(-t/\tau)/\varepsilon,$$

$$p_0 = -\rho gz - E_{20}^2(1 - \exp(-2t/\tau)/\varepsilon)/2, \tag{10}$$

$$\sigma_0 = \varepsilon_0 E_{20}(1 - \exp(-t/\tau)),$$

where $\tau = \varepsilon\varepsilon_0/\lambda$.

We assume that the charge relaxation time far exceeds the characteristic time of the hydrodynamic instability development. On such a time scale, the potential and surface charge density are also slowly varying functions. The unperturbed intensities outside, E_{20}, and inside, E_{10}, the liquid are assumed to be constant in the statement of the perturbed problem. The potential inside the liquid, surface charge density, and pressure are given by

$$\varphi_{10} = -E_{10}z, \quad \sigma_0 = \varepsilon_0(E_{20} - \varepsilon E_{10}),$$

$$p_0 = -\rho gz - (E_{20}^2 - \varepsilon E_{10}^2)/2. \tag{11}$$

Consider problem (2)–(8) in the case of very viscous liquids, when the inertial terms in (2) can be neglected. The perturbations are defined as follows:

$$A = A_0 + \overline{A},$$

where A_0 is any quantity in (9) and (11) and $\overline{A}$ is a perturbation.

The linearized problem has the form

$$\nabla p = \mu\Delta\mathbf{V}, \quad \text{div}\mathbf{V} = 0, \quad \Delta\varphi_{1,2} = 0$$

with the boundary conditions at $z = 0$

$$w = \partial f/\partial t, \quad \partial\sigma/\partial t + \sigma_0\partial u/\partial x - \lambda E_{n1} = 0,$$

$$\varphi_2 - \varphi_1 = (E_{20} - E_{10})f, \quad \varepsilon\mathbf{n}\nabla\varphi_1 - \mathbf{n}\nabla\varphi_2 = \sigma/\varepsilon_0,$$

$$-p + 2\mu\partial w/\partial z + \rho gf - Tf'' = \varepsilon_0 E_{20}E_{2z} - \varepsilon\varepsilon_0 E_{10}E_{1z},$$

$$\mu(\partial u/\partial z + \partial w/\partial x) = \varepsilon_0(E_{20}E_{2x} - \varepsilon E_{10}E_{1x})$$

$$+ f'\varepsilon_0(E_{20}^2 - \varepsilon E_{10}^2)/2.$$

Here, μ is the liquid viscosity and the bar at the perturbations is omitted. Assuming that the x and t dependences of each of the variables have the factor $\exp(st + ikx)$, we find the following dispersion relation:

$$\Omega^2(\varepsilon + 1) + a\Omega + b = 0, \tag{12}$$

$$a = (\varepsilon + 1)\xi + C^2 + H - (\varepsilon - 1 + C)^2/\varepsilon,$$

$$b = (C^2 + H)\xi \tag{13}$$

$$- (2(\varepsilon - 1 + C)H + (\varepsilon - 1)(3C^2 - 2C) + C^4)/2\varepsilon,$$

$$C = \sigma_0/\varepsilon_0 E_{20}, \quad \Omega = 2\mu s/\varepsilon_0 E_{20}^2, \quad H = 2\mu\lambda/\varepsilon_0^2 E_{20}^2,$$

$$\xi = (\alpha^2 H^2/4 + y^2)/y, \quad \alpha = (\rho gT\varepsilon_0^2/\mu^2\lambda^2)^{1/2},$$

$$y = Tk/\varepsilon_0 E_{20}^2. \tag{14}$$

Here, H is the electrical analogue of the Hartmann number. We choose $\gamma = \alpha H = (4\rho gT)^{1/2}/\varepsilon E_{20}^2$ as an instability parameter. The inequalities $a \geq 0$ nd $b \geq 0$ are the stability conditions for quadratic equation (12). In view of (13) and (14), we find

$$\gamma \geq \gamma_1 = \frac{\alpha(\varepsilon - 1)(\varepsilon - 1 + 2C - C^2)}{\varepsilon(\alpha(\varepsilon + 1) + 1)}, \tag{15}$$

$$\gamma \geq \gamma_2 = \frac{\sqrt{A^2 + 2\alpha\varepsilon((\varepsilon - 1)(3C^2 - 2C) + C^4)} - A}{2\varepsilon}, \tag{16}$$

where $A = \alpha\varepsilon C^2 - \varepsilon + 1 - C$.

Table 1

Liquid	$\rho \times 10^{-3}$, g/m^3	$T \times 10^3$, N/m	ε	λ, Ω^{-1} m^{-1}	$\mu \times 10^3$, Pa s
1	0.88	28.9	2.28	10^{-16}	0.65
2	1.59	27	2.24	4×10^{-16}	0.97
3	0.88	29.8	2.25	10^{-12}	17.3
4	0.88	24.8	4.87	5×10^{-7}	0.73
5	1.25	32.2	10.4	3×10^{-8}	0.9
6	0.95	36.8	36.7	1.8×10^{-6}	1.0

In the limiting cases of an uncharged perfect insulator ($C = 0$, $H \longrightarrow 0$, and $\alpha \longrightarrow \infty$) and a perfect conductor ($C = 1$, $\alpha \longrightarrow 0$, and $H \longrightarrow \infty$), Eq. (12) yields

$$\Omega = -\frac{\rho g + Tk^2}{\varepsilon_0 E_{20}^2 k} + \frac{(\varepsilon - 1)^2}{\varepsilon(\varepsilon + 1)}, \tag{17}$$

$$\Omega = -\frac{\rho g + Tk^2}{\varepsilon_0 E_{20}^2 k} + 1. \tag{18}$$

Periodic solutions of Eqs. (17) and (18) exist only in the form of standing waves at $\Omega = 0$. In the range $\Omega < 0$, the solutions are damped aperiodic functions.

The values $\alpha \ll 1$ ($H \gg 1$) and $\alpha \gg 1$ ($H \ll 1$) correspond to highly and poorly conducting liquids, respectively. The onset of instability is defined as the minimum field intensity violating either of inequalities (15) and (16). As is seen from (10), C varies in the range [0, 1]. The threshold intensities breaking the stability of the interface depend on the liquid properties and means of potential increase.

Let us consider two cases of instability excitation. At an abrupt increase in the intensity to the threshold, C can take all the values from the range [0, 1]. The minimum intensity is specified by the maximum of γ from (15) and (16):

$$\gamma_s = \max_C(\gamma_1, \gamma_2). \tag{19}$$

At a gradual increase in the intensity, the surface charge density is close to its maximum value $C = 1$. In this case, we have

$$\gamma_g = \max(\gamma_1(C = 1), \gamma_2(C = 1)). \tag{20}$$

As is seen from (9) and (10), the electrohydrodynamic system under consideration has a single equilibrium state corresponding to the maximum charge of the interface. The gradual increase in the voltage keeps the system in the near-equilibrium state. When the voltage grows abruptly, the charge can vary from zero to the equilibrium value. Here, it is assumed that the charge relaxation process is quasi-equilibrium. This means that in the nonequilibrium state, a rapid hydrodynamic response to any slow variation of the surface charge

takes place at any time. In terms of the characteristic times of the process, the quasi-equilibrium condition has the form $s\tau > 1$.

The figure shows the critical field intensity as a function of the conductivity according to (19) and (20). The threshold intensity for real liquids is always higher than that for a conductor (see (18)). At the same time, the critical field of a perfect insulator, Eq. (17), is unachievable even for a very poor conductor. Criteria (19) and (20) yield different instability excitation curves. The threshold intensity in the case of abrupt increase in the voltage is always lower than that in the gradual case. It is worth noting that in the former case, the instability takes place at $C < 1$.

We will study the wave propagation in the case of gradual increase ($C = 1$). As is seen from (12), these waves can propagate under the conditions $a = 0$ and $b > 0$:

$$\frac{\rho g + Tk^2}{\varepsilon_0 E_{20}^2 k} = \frac{\varepsilon - 1 - H}{\varepsilon + 1}, \tag{21}$$

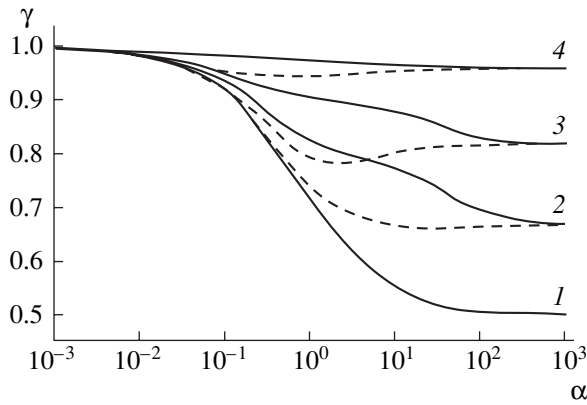$$\frac{\rho g + Tk^2}{\varepsilon_0 E_{20}^2 k} > \frac{2H + 1}{2(H + 1)}. \tag{22}$$

The allowable values of the wave numbers are given by Eq. (21), which can be recast in the form

$$\left(k - \frac{\varepsilon - 1 - H}{\varepsilon + 1}\frac{\varepsilon_0 E_{20}^2}{2T}\right) + \frac{\rho g}{T}$$
$$- \left(\frac{\varepsilon - 1 - H}{\varepsilon + 1}\frac{\varepsilon_0 E_{20}^2}{2T}\right)^2 = 0. \tag{23}$$

This equation has a solution if

$$\gamma \leq \frac{\alpha(\varepsilon - 1)}{1 + (\varepsilon + 1)\alpha}. \tag{24}$$

Comparing (24) with stability condition (20), we conclude that the only case when the waves can propagate is that when (24) is an equality. Then, the wave

Threshold intensity as a function of the liquid conductivity. The dashed line corresponds to the gradual increase in the voltage. $\varepsilon = $ (*1*) 2, (*2*) 5, (*3*) 10, and (*4*) 40.

number is determined from (23):

$$k = \sqrt{\frac{\rho g}{T}}. \tag{25}$$

This wave number is independent of the field and is expressed through the capillary constant of a liquid. Substituting (25) into (22), we find the condition for wave propagation

$$H < \frac{\sqrt{3 + 2\varepsilon} - 3}{2}. \tag{26}$$

This relation shows that the generation of sustained waves is possible only in liquids with $\varepsilon > 3$.

RESULTS AND DISCUSSION

The destabilization of the liquid surface under various conditions was studied in a cylindrical glass cell. The level of the liquid was adjusted by the movable conducting grounded bottom. The potential was applied to the planar electrode placed at a distance of 0.5–2 cm from the liquid surface. The typical instability field strength was ≈ 20 kV/cm. Sinusoidal ripples from the source voltage were about 5% of the mean value. The total error in the dimensionless quantities measured was within 10%. In the experiments, the voltage at which a liquid jet was emitted from the surface was measured. Under conditions of gradual increase, the potential grew stepwise with the delay time equal to the charge relaxation time τ. At an abrupt increase, the potential was applied instantly and was kept for the time τ. If the emission did not occur, the experiment was repeated at a higher voltage after the liquid had been discharged.

The lower instability intensity was observed in the latter regime. In the former case, the intensities were 5–20% higher. The properties of the liquids studied are listed in Table 1.

The liquids were 1, benzene; 2, carbon tetrachloride; 3, transformer oil; 4, butyl acetate; 5, dichloroethane; and 6, dimethylformamide. The measured parameters of these liquids are presented in Table 2. The parameters γ and H were calculated at the critical experimental value of E_{20} for an abrupt increase in the voltage. The values of $\gamma(\alpha)$ from Table 2 agree with the calculated data shown in the figure within the experimental error. For liquids 4–6, the charge relaxation time is comparable to the time of setting the source voltage. Therefore, these liquids destabilize in the intermediate (between abrupt and gradual) regime. Nevertheless, instability appears at the near-threshold intensity for the regime of abrupt increase. The measured intensity turned out to be lower than the threshold. This is probably due to the fact that the system experiences the peak field intensity, which is higher than the measured effective value.

The liquids in Table 1 with low $\varepsilon \approx 2$ are customarily regarded as good insulators. However, our results indicate that their behavior differs essentially from that of a perfect insulator. The interface is charged in the course of measurement, and even a poor conductor does not achieve the dielectric limit

$$\gamma = \frac{(\varepsilon - 1)^2}{\varepsilon(\varepsilon + 1)},$$

which corresponds to dispersion relation (17). The surface of such liquids becomes unstable at $\gamma \approx 0.5$.

Of the liquids presented, only dichloroethane can generate sustained waves. In the experiment, a stable two-dimensional system of surface standing waves is generated at field intensities somewhat lower than the critical one. This effect depends drastically on the liquid purity. A small amount of impurities raising the conductivity results in the violation of (26), and the waves disappear. Although butyl acetate and dimethylformamide have $\varepsilon > 3$, their parameters do not satisfy (26), and sustained waves are not generated.

Table 2

Liquid	$E_{20} \times 10^{-6}$, V/m	α	H	γ
1	2.5	2.2×10^9	2.6×10^{-10}	0.57
2	2.9	4.7×10^8	1.2×10^{-9}	0.56
3	2.5	8.3×10^3	7.1×10^{-5}	0.58
4	1.9	0.36	2.6	0.92
5	2.2	6.6	0.14	0.93
6	2.0	9.2×10^{-2}	12	1.1

REFERENCES

1. Ya. I. Frenkel', Zh. Éksp. Teor. Fiz. **6**, 347 (1936).
2. É. K. Kopeĭkina, Magn. Gidrodin., No. 3, 142 (1970).
3. J. R. Melcher and C. V. Smith, Phys. Fluids **12**, 778 (1969).
4. J. R. Melcher, *Field-Coupled Surface Waves* (MIT Press, Cambridge, 1963).
5. D. F. Belonozhko, A. I. Grigor'ev, and S. O. Shiryaeva, Izv. Akad. Nauk, Mekh. Zhidk. Gaza, No. 6, 116 (1998).
6. J. R. Melcher and G. I. Taylor, in *Annual Review of Fluid Mechanics*, Ed. by W. R. Sears and M. Van Dyke (Palo Alto, California, 1969), Vol. 1, pp. 111–146.
7. A. A. Shutov, N. M. Alontseva, and V. Ya. Shkadov, in *Proceedings of the Regional Contest of Projects in Natural Sciences* (Éĭdos, Kaluga, 2000), Vol. 1, pp. 67–88 (ISBN 5-938810-006-2).
8. I. E. Tamm, *The Principles of Theory of Electricity* (Nauka, Moscow, 1976).

Translated by M. Fofanov

Kinetics of Pore Healing and Copper Hardening upon Uniform Compression

A. I. Petrov and M. V. Razuvaeva

Ioffe Physicotechnical Institute, Russian Academy of Sciences,
Politekhnicheskaya ul. 26, St. Petersburg, 194021 Russia
e-mail: An.Petrov@pop.ioffe.rssi.ru

Received November 29, 2001; in final form, January 31, 2002

Abstract—X-ray diffraction is employed to determine the dependence of the mobile dislocation density on the strain arising upon healing porous copper subjected to a hydrostatic pressure of up to 1500 MPa. The mean free path of dislocations is estimated as a function of the compression strain, strain rate under pressure, rate of dislocation generation, and mean velocity of their motion. © 2002 MAIK "Nauka/Interperiodica".

According to [1], the healing of pores in crystals under hydrostatic pressure at moderate temperatures is associated with a stationary dislocation ensemble forming near the pores. The pore size stabilizes when shear stresses produced by an applied pressure are balanced by oppositely directed stresses due to dislocations generated previously. In [1], the pore size was related to an applied pressure and the number of dislocation loops emitted by a source in the vicinity of pores. Experimental data (mostly for alkali halide crystals) support the basic ideas of the dislocation mechanism of healing under pressure [1]. However, the pressure-induced change in the dislocation density near pores has not been estimated.

In this study, we investigate the effect of a hydrostatic pressure on the dislocation density in porous metals and elucidate the role of dislocations surrounding pores on pressure hardening observed in [2].

Porous copper samples were obtained by evaporating out zinc from brass *in vacuo* at 1073 K for 8 h. The samples resulted measured $30 \times 3 \times 1$ mm and had an initial porosity of 12%. A high-pressure camera was used to subject the samples to a hydrostatic pressure of up to 1500 MPa. Dilute silicone oil (one part petrol per two parts oil) served as a pressure-transmitting medium. The porosity before (W_0) and after (W) the action of a pressure up to 1500 MPa was determined by density measurements and optical microscopy. The longitudinal compression strain due to the pore contraction was estimated as $\varepsilon = 1/3(W_0 - W)$. The dislocation density ρ was calculated from the physical broadening of X-ray diffraction lines β according to the expression [2]

$$\rho = \beta^2 \cot^2\theta / 2b^2, \tag{1}$$

where b is the Burgers vector and θ is the diffraction angle. In order to find the contribution of physical factors to the broadening of the lines in the porous copper, we performed a preliminary analysis of the broadening

for the (111) and (222) reflection lines. Applying the approximation technique [3], we established that the broadening is due largely to dislocation-induced microstrains; the broadening associated with the size of blocks turned out to be insignificant.

Now, let us consider the results obtained in more detail. Figure 1 (curve *1*) shows the longitudinal strain dependence of the density of mobile dislocations, which provide for plastic flow of the material into pores. The nonlinear character of the dependence under hydrostatic pressure is evident. According to the well-known expression

$$\varepsilon = \rho b \bar{L}, \tag{2}$$

this suggests that the mean free path $\bar{L}$ of dislocations varies as the samples are strained under pressure. Figure 2 indicates the linear decrease in $\bar{L}$ with the plastic strain.

Let us estimate the influence of increasing the dislocation density under pressure on the hardening of porosity copper. As follows from the microhardness measurements of the healed samples (Fig. 3, curve *1*), an increase in the density of mobile dislocations hardens the porous copper samples.

It is interesting to compare our results with those published previously. For example, the data obtained in [4, 5] indicate that the mean free path of dislocations in fcc metals is sensitive to the strain rate. In particular, the free path is essentially smaller when dynamic, rather than static, loading is used (Fig. 4). Comparing the data for fcc metals presented in Fig. 4 with the values of $\bar{L}$ obtained in this study, we see that the rate of plastic flow near pores under hydrostatic pressure is no less than $\dot{\varepsilon} \approx 10^1$ s^{-1}. From this value of $\dot{\varepsilon}$ and the expression for the mean rate of plastic strain $\dot{\varepsilon} = \varepsilon/t_p$, one can estimate the time of pressure action at 10^{-3} s or less;
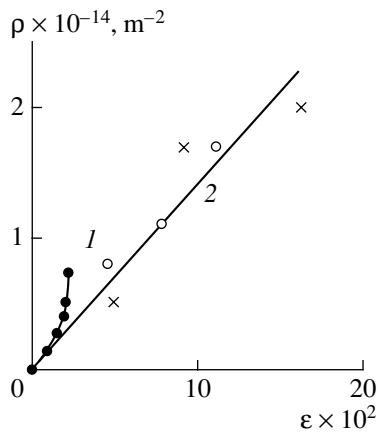
Fig. 1. Strain dependence of the mobile dislocations density for (●) porous copper, (○) pore-free copper, and (×) brass [5].
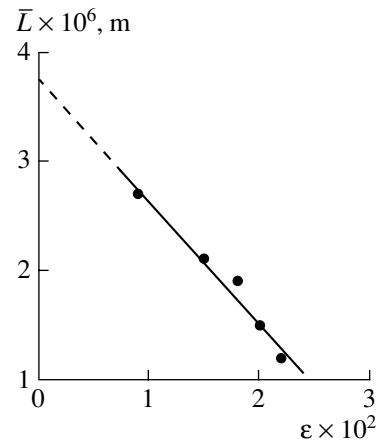

Fig. 2. Strain dependence of the dislocation mean free path upon pore healing in copper.
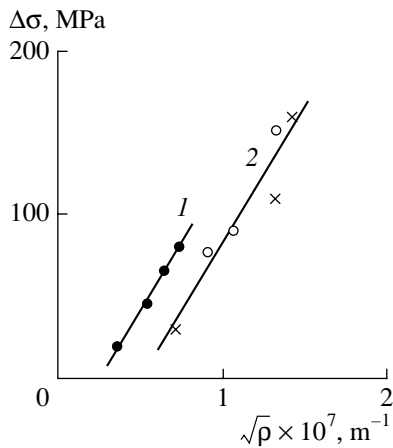

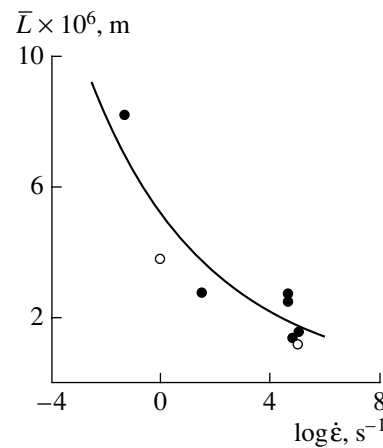**Fig. 3.** Hardening increment versus the mobile dislocations density. Notation is the same as in Fig. 1.

Fig. 4. Mean free path of dislocations in fcc metals versus the strain rate: (●) copper and (○) nickel [4, 5].

then, the mean velocity of dislocations determined from the Orovan equation $\vartheta = \dot{\varepsilon}/b\rho$ amounts to approximately 1×10^{-3} m/s and the rate of their generation $\dot{\rho} = \rho/t_{\mathrm{p}} \approx 5 \times 10^{16}$ m^{-2}/s.

The above estimates show that the plastic strain built up near pores under the pressure corresponds to dynamic, rather than static, loading conditions. Bearing this in mind, let us compare our data obtained under the hydrostatic pressure with the results of testing the samples subjected to pulsed dynamic compression. Figure 1 (curve 2) and Fig. 3 (curve 2) [5] show, respectively, the strain dependence of the mobile dislocation density and the ρ dependence of the hardening increment for pore-free copper and brass (70% Cu + 30% Zn) samples subjected to shock loading with a pulse duration of 2 μs and at pressures of up to 20 GPa. It is indicated by Fig. 1 that the dependences of ρ on ε are very close to each other at small ε. The discrepancy at large strain values may be associated with the pressure-induced decrease

in $\bar{L}$ (under pulsed loading, the value of $\bar{L}$ is strain-independent and equals 2.8×10^{-6} m). The effect of the initial (before the tensile test) dislocation density is practically the same for both loading methods (Fig. 3). The minor discrepancy between curves 1 and 2 and may be due to the difference in the dislocation distribution, which is uniform over the sample in the case of pulsed loading but has a peak near pores when hydrostatic pressure is applied.

A possible reason why $\bar{L}$ decreases with increasing ε is the following. The optical microscopy study of metallographic sections made on the porous copper revealed one distinct dislocation slip system in the vicinity of the pores at low pressures. However, as the pressure rises, other slip systems are involved in the deformation process. The interaction of dislocations lying in different planes must result in the formation of Lomer–Cottrell sessile dislocations. Based on the Seeger concept [6], one can infer that barriers forming in

slip planes as a result of dislocation interactions may decrease the mean free path of dislocations when the degree of pore healing increases.

CONCLUSIONS

(1) The degree of pore healing under hydrostatic pressure is related to the density of dislocations generated in the vicinity of pores.

(2) The dislocation mean free path decreases with pressure.

(3) The copper hardening depends on a pressure-induced rise in the dislocation density near pores.

REFERENCES

1. Ya. E. Geguzin and V. G. Kononenko, Fiz. Khim. Obrab. Mater., No. 2, 60 (1982).

2. V. I. Betekhtin, A. I. Petrov, N. K. Ormanov, *et al.*, Fiz. Met. Metalloved. **67**, 318 (1989).

3. Ya. S. Umanskiĭ, Yu. A. Skakov, A. N. Ivanov, *et al.*, *Crystallography, X-ray Diffraction, and Electron Microscopy* (Metallurgiya, Moscow, 1982).

4. G. N. Épshteĭn, *Structure of Explosion-Deformed Metals* (Metallurgiya, Moscow, 1980).

5. *Shock Waves and High-Strain-Rate Phenomena in Metals*, Ed. by M. A. Meyers and L. E. Murr (Plenum, New York, 1981; Metallurgiya, Moscow, 1984).

6. A. Seeger, in *Dislocations and Mechanical Properties of Crystals: Proceedings of the International Conference, Lake Placid, 1956,* Ed. by J. C. Fisher *et al.* (Wiley, New York, 1959; Inostrannaya Literatura, Moscow, 1960).

Translated by A. Sidorova-Biruykova

BRIEF
COMMUNICATIONS

Effect of Superconductivity–Magnetism Interaction on the Differential Conductivity in Ho(NiB)₂C/Ag Point Contacts

I. N. Askerzade[1, 2] and B. Tanatar[3]

[1] *Institute of Physics, Academy of Sciences of Azerbaijan,*
ul. Dzhavida 33, Baku, 370143 Azerbaijan

e-mail: solstphs@lan.ab.az

[2] *Physics Department, Ankara University, Tandogan, 06100 Ankara, Turkey*

[3] *Physics Department, Bilkent University, 06533 Bilkent, Ankara, Turkey*

Received January 15, 2002

Abstract—In terms of the Blonder–Tinkham–Klapwjik theory, the differential conductivity of Ho(NiB)₂C/Ag point contacts is explained by the coexistence of magnetic ordering and superconductivity in holmium boron carbides. © *2002 MAIK "Nauka/Interperiodica".*

The measurement of the differential conductivity in normal metal–superconductor junctions is a sensitive method for studying superconductor properties [1–3]. It was applied to study the order parameter symmetry in cuprate semiconductors [4, 5], as well as the properties of recently discovered semiconductor MgB_2 [6]. Investigation into the superconducting properties of exotic boron carbide–(boron nitride–) based semiconductors of the RTBC(N) class, where R is a rare-earth element and T = Ni, Pd or Pt, is also of great importance for the elucidation of a microscopic mechanism of superconductivity in these compounds [7]. Point contact spectroscopy of boron carbides is needed for the detailed study of the order parameter symmetry and is also dictated by the coexistence of superconductivity and magnetism in magnetic boron carbides. Andreev spectroscopy of Y(La)(NiB)₂C nonmagnetic boron carbides gives peak values at the gap value of the order parameter [3]. Similar behavior was observed in Dy(Er)(NiB)₂C magnetic boron carbides. The compound with Dy exhibits superconductivity in the presence of antiferromagnetic ordering at a temperature $T_N = 10.5$ K and is the only boron carbide–based compound with the Néel temperature exceeding the superconducting transition critical temperature, $T_N > T_c = 6$ K. In the Er-containing compound (the critical temperature $T_c = 10.8$ K), antiferromagnetic ordering takes place below $T_N = 5.9$ K. Ho(NiB)₂C compounds have a more complex magnetic structure. Here, antiferromagnetic ordering arises below the Néel temperature $T_N \approx 5$ K and is associated with a commensurate magnetic structure modulated along the *z* axis with a wave vector $\mathbf{Q}_{AF} = \mathbf{c}^* = 2\pi/c$. Other magnetic structures were discovered in the temperature range $T_N < T < T_m = 6$ K. In this interval, an incommensurate phase helical along the *z* axis with a wave vector $\mathbf{Q}_C = 0.91\mathbf{c}^*$ and an incommensurate phase modulated along the *x* axis with a vector $\mathbf{Q}_a = 0.55\mathbf{c}^*$ form. The reentrant or almost reentrant behavior of superconductivity was detected in the magnetic ordering range [8]. Ho(NiB)₂C/Ag point contacts were experimentally studied in [9], but the suppression of Andreev features remained unexplained. In this work, we concentrate on the effect of the helical structure on the differential conduction in Ho(NiB)₂C/Ag point junctions, invoking the Blonder–Tinkham–Klapwjik (BTK) formalism [10].

First, let us turn to the effect of the helical structure on the superconductivity. For the first time, this point was considered by Morozov in [11]. Recently, he has extended his approach for holmium boron carbides [12]. Applying the Bogoliubov transformation, one can show that the gap parameter in the quasi-particle spectrum becomes highly anisotropic and disappears at the line of intersection between the Fermi plane and Bragg planes, which are generated by magnetic ordering. The conventional BTK theory for isotropic semiconductors can be generalized for the anisotropic case by introducing the dependence of the gap on the momentum $\Delta(\mathbf{k})$ into the expressions for the Andreev reflection probability $A(\varepsilon, \Delta(\mathbf{k}))$ and normal tunneling probability $B(\varepsilon, \Delta(\mathbf{k}))$. Then, the normalized zero-temperature con-

ductance of a point junction is given by

$$\frac{G_{NS}}{G_{NN}} = \frac{\partial I_{NS}/\partial V}{\partial I_{NN}/\partial V}$$

$$= \frac{\partial/\partial V \int d^3 k \, v_z \{1 + A(\varepsilon, \Delta(\mathbf{k})) - B(\varepsilon, \Delta(\mathbf{k}))\}}{\partial/\partial V \int d^3 k \, v_z \{1 - Z^2/1 + Z^2\}}, \quad (1)$$

where v_z is the velocity positive component normal to the *NS* interface and Z is the potential barrier height at the interface. In this approximation, the proximity effect can be neglected although the effect of the surface on the order parameter is significant in purely *d*-wave and *p*-wave superconductors.

The differential conductivity for *N*/*d*-wave superconductors was calculated in [13] with Eq. (1). Similar calculations for ferromagnet/*d*-wave structures were carried out in [14]. The dependence of the intragap structure on the orientation of the *d* wave with respect to the interface was found. Of interest also are calculations performed within the same approach for Sr_2RuO_4 in view of the *p*-wave symmetry of the order parameter [15]. For the UPt_3 heavy-fermion system, the oddness of the order parameter was taken into account [16]. In all the cases, the anisotropy of the order parameter causes the plateau in the $(-\Delta, \Delta)$ interval to transform into a triangular peak of the conductance inside the gap.

As was shown by Morozov [11, 12], the order parameter in the presence of the helical structure is given by

$$\Delta(\mathbf{k}, T) = (u_{\mathbf{k}}^2 - v_{\mathbf{k}}^2)\Delta(T),$$

where

$$u_{\mathbf{k}}^2 - v_{\mathbf{k}}^2 = \left\{ \frac{(\varepsilon_{\mathbf{k}} - \varepsilon_{\mathbf{k}+\mathbf{Q}})^2}{(\varepsilon_{\mathbf{k}} - \varepsilon_{\mathbf{k}+\mathbf{Q}})^2 + I^2 S^2} \right\}^{1/2}, \quad (2)$$

I is the exchange integral, *S* is the averaged spin, $\varepsilon_{\mathbf{k}}$ is the dispersion relation in the paramagnetic phase,

$$\Delta(T) = \int_0^\omega d\varepsilon \frac{\Delta(T)(1 - 2n_{\mathbf{k}})}{\varepsilon^2 + \Delta^2(T)}$$

$$\times \left(\int_{MFS} \frac{dS'}{(2\pi)^3} \frac{(u_{\mathbf{k'}}^2 - v_{\mathbf{k'}}^2)^2}{|\nabla_{\mathbf{k'}}\tilde{\varepsilon}_{\mathbf{k'}}|} \right), \quad (3)$$

$\tilde{\varepsilon}_{\mathbf{k}}$ is a new dispersion relation, and n_k is the Fermi distribution function. Equation (4) corresponds to the conventional BCS equation with the effective parameter of interaction in parentheses. The dependence $\lambda_{eff}(T)$ depends on the Bogoliubov coefficients and on the Fermi surface slope. Since anomalies of the vector dependence are observed near the intersections of the Fermi surface with Bragg planes, the difference between the actual interaction constant and its effective value $\Delta\lambda = \lambda - \lambda_{eff}(T)$ can be expanded in units of IS/E_F. The difference $\Delta\lambda$ was estimated at $\Delta\lambda/\lambda = 0.12$ [17] from the data for the band structure of boron carbides. This value was used to explain the reentrant behavior of the upper critical field in holmium boron carbides.

Experimental data [9] indicate the insensitivity of the curve $G_{NS}(V)$ to the orientation of the contact plane with respect to crystallographic axes. This fact supports the isotropy of the electronic structure of these compounds. Thus, the possibility of *d*-wave or *p*-wave pairing in helicoidal superconductors should be excluded; otherwise, the current–voltage characteristic would be sensitive to the contact plane orientation.

The evolution of the $G_{NS}(V)$ shape with the barrier height is analyzed in the BTK theory [10]. Clearly, the intragap plateau in the absence of the barrier changes to peaks for $\pm\Delta$ as the barrier grows. For $Ho(NiB)_2C/Ag$ at $T < 5$ K, where the helicoidal structure transforms into the antiferromagnetic phase, the current–voltage characteristic has two peaks.

However, in the temperature interval $5 < T < 8.1$ K, which is equivalent to $\Delta T/T \approx 3/4 = 0.4$, the gapless behavior is observed (note that this parameter equals 0.2 for $Er(NiB)_2C$ compounds). In our opinion, the broadening and the gapless behavior are due to the decline of the order parameter when the helical structure is present in the system. As was noted [17], the decrease in the interaction constant to $\Delta\lambda/\lambda \approx 0.12$ is insufficient for the order parameter to be suppressed completely. On the other hand, when calculating the differential conductivity, we must include the additional factor $u_{\mathbf{k}}^2 - v_{\mathbf{k}}^2$. Since this factor is less than unity, we have one more channel for suppressing the order parameter.

Thus, the broadened gapless behavior of $G_{NS}(V)$ for $Ho(NiB)_2C/Ag$ point contacts at near-critical temperatures can be explained by the order parameter suppression. In experiment [9], the suppression is incomplete because $Ho(NiB)_2C$ samples were extrapure. It was shown [18] that nonmagnetic impurities heavily destroy superconductivity in helicoidal systems. We can thus conclude that helicity "retards" the emergence of two peaks in the $G_{NS}(V)$ curve.

REFERENCES

1. A. F. Andreev, Zh. Éksp. Teor. Fiz. **46**, 1823 (1964) [Sov. Phys. JETP **19**, 1228 (1964)].

2. S. N. Artemenko, A. F. Volkov, and A. V. Zaĭtsev, Zh. Éksp. Teor. Fiz. **76**, 1816 (1979) [Sov. Phys. JETP **49**, 924 (1979)].

3. I. K. Yanson, in *Symmetry and Pairing in Superconductors*, Ed. by M. Auslos and S. Kruchinin (Academic, Kluwer, 1999), pp. 271–285.

4. M. Belogolovski, M. Grajcar, P. Kus, *et al.*, Phys. Rev. B **59**, 9617 (1999).

5. R. A. Clemm, in *High Temperature Superconductivity—Ten Years after Discovery*, Ed. by K. B. Garg and S. M. Bose (Narosa, New Delhi, 1998), pp. 179–190.

6. A. Kohen and G. Deutscher, Phys. Rev. B **64**, 060506 (2001).

7. S.-L. Drechsler, S. V. Shulga, K.-H. Muller, *et al.*, Physica C (Amsterdam) **317–318**, 117 (1999).

8. M. S. Lin, J. H. Shich, B. Yon, *et al.*, Phys. Rev. B **52**, 1181 (1995).

9. L. F. Rybaltchenko, A. G. M. Jansen, P. Wider, *et al.*, Physica C (Amsterdam) **319**, 189 (1999).

10. G. E. Blonder, M. Tinkham, and T. M. Klapwjik, Phys. Rev. B **25**, 4515 (1982).

11. A. I. Morozov, Fiz. Tverd. Tela (Leningrad) **22**, 3372 (1980).

12. A. I. Morozov, Zh. Éksp. Teor. Fiz. **110**, 1903 (1996) [JETP **83**, 1048 (1996)].

13. Y. Tanaka and S. Kashiwaya, Phys. Rev. Lett. **74**, 3451 (1995).

14. J.-X. Zu, B. Freidman, and C. S. Ting, Phys. Rev. B **59**, 9558 (1999).

15. C. Honercamp and M. Sigrist, J. Low Temp. Phys. **111**, 895 (1998).

16. G. Goll, C. Bruder, and H. V. Lohneyesen, Phys. Rev. B **52**, 6801 (1995).

17. A. Amici, P. Thalmeier, and P. Fulde, Phys. Rev. Lett. **84**, 1800 (2000).

18. K. Krug, H. Heinecke, and K. Winzer, Physica C (Amsterdam) **267**, 321 (1996).

Translated by V. Isaakyan

BRIEF
COMMUNICATIONS

Suppression of Interference by Quantum Fluctuations in a Josephson Interferometer with a Single Small Junction

I. N. Askerzade[1, 2]

[1]Institute of Physics, Academy of Sciences of Azerbaijan,
ul. Dzhavida 33, Baku, 370143 Azerbaijan
e-mail: solstphs@lan.ab.az
[2] Physics Department, Ankara University, Tandogan, 06100 Ankara, Turkey
Received January 15, 2002

Abstract—The geometric inductance of a single-junction interferometer at which quantum fluctuations suppress interference and the Coulomb blockade of Cooper pairs is observed is found. © *2002 MAIK "Nauka/Interperiodica".*

INTRODUCTION

Quantum effects in small Josephson junctions have recently become the subject of great interest [1, 2]. Quantum fluctuations basically show up in so-called macroscopic quantum phase tunneling through the energy barrier of a Josephson junction with current [3]. An increase in quantum fluctuations at low temperatures causes another effect—Coulomb blockade in small Josephson junctions [4, 5]. The condition for this effect to occur in a single small Josephson junction has the form [6]

$$\min[\hbar\omega_p, \hbar\omega_{ch}] \ll E_J, \tag{1}$$

where $\omega_p = (2eI_c/\hbar C)^{1/2}$ is the plasma frequency; $\omega_{ch} = 2eI_cR_N/\hbar$ is the characteristic electron frequency of a Josephson junction with critical current I_c, normal resistance R_N, and capacitance C; and $E_J = \hbar I_c/2e$ is the Josephson energy. In terms of resistance, condition (1) is recast as

$$R_N < R_Q; \quad R_Q = \hbar/4e^2, \tag{2}$$

where the constant $R_Q = 1$ KΩ is the quantum unit of resistance. Conditions for Coulomb blockade to take place depend on the environment of a Josephson junction and junction parameters. In this work, we analyze these conditions as applied to a single-junction interferometer.

BASIC EQUATIONS

For small Josephson junctions with current that are subjected to a small voltage ($eV < \Delta$), where Δ is the energy gap of a superconductor), the Hamiltonian has

the form

$$H = \hat{Q}^2/2C + E_J(1 - \cos\phi)$$
$$- \hbar/2e(I(t) - I_N(x)) + H_e(x), \tag{3}$$

where H_e and x are the electronic Hamiltonian and the set of internal degrees of freedom of a thermostat with which the junction is coupled through the "normal" current I_N of quasi-particles. The general theory of the Josephson effect applies if

$$E_Q \ll E_J, \quad E_Q = Q^2/2C. \tag{4}$$

In this case, electrical charge Q and Josephson phase ϕ can be considered as classical variables. In the inverse limit ($E_Q \gg E_J$), such an approach fails and Q and ϕ should be viewed as noncommuting operators [7], with ϕ meeting the condition

$$\hat{Q} = -2ei(\partial/\partial\phi). \tag{5}$$

If the external current is small, the junction has been shown to be described with a simple equation for quasi-charge in terms of the resistive model [8]:

$$\dot{Q}' = I(t) - V(Q')/R_N + I_F(t), \tag{6}$$

where $V(Q') = dE/dQ'$ and $I_F(t)$ is the fluctuation current. For $I < e/R_NC$, Eq. (6) may have a quasi-static solution. In the inverse limit, the energy, voltage, and current oscillate (so-called Bloch oscillations) with a frequency

$$\omega_B = \pi/e(I - \langle V \rangle/R_N). \tag{7}$$

Thus, with conditions (2) and (4) met, quantum fluctuations not only merely grow but give rise to new phenomena.

It has also been shown [9] that the properties of the small junction depend on the impedance of its environment. Only if the resistance of the environment (leads)

is high, $R_s \gg R_Q$, can Coulomb blockade not be removed by charge fluctuations in the leads. In a single junction, these effects are difficult to detect because of interference. Since a high-frequency SQUID and a single-junction interferometer as its basic component are well shielded from external fields and interference [6], it is of interest to see how the smallness of the junction influences the properties of a single-junction interferometer. To this end, we consider the Coulomb blockade conditions in a Josephson junction shaped into a superconducting ring. In addition, it is of interest to trace the effect of the junction environment.

It is well known that the state of a single-junction interferometer is described by the equation

$$\phi + l\sin\phi = \phi_e, \tag{8}$$

where $\phi_e = 2\pi\Phi_e/\Phi_0$ is the external magnetic flux normalized to the magnetic flux quantum, and $l = 2\pi LI_c/\Phi_0$ is the normalized geometrical inductance of the ring. Quantum fluctuations render ϕ a quantum variable, which satisfies the Heisenberg uncertainty relation

$$\Delta Q \Delta \phi > 2e. \tag{9}$$

To analyze the operation of a single-junction interferometer under Coulomb blockade conditions, it is necessary to consider the Hamiltonian

$$H = \hat{Q}^2/2C + E_J\left(1 - \cos\phi + \frac{(\phi - \phi_e)^2}{2l}\right). \tag{10}$$

A solution to the problem in terms of the Schrödinger equation can be found only in limiting cases. Obviously, at large inductances, when $\frac{(\phi - \phi_e)^2}{2l} \ll (1 - \cos\phi)$, the interferometer behaves as a single junction. In this limit, the quadratic potential is small compared with the periodic potential, and Coulomb blockade shows up in the same way as in a single junction. To quantitatively estimate the geometrical inductance, we will reason as follows. The current in the interferometer is a classical quantity; that is, it can be determined exactly. To be more precise, its uncertainty is extremely small:

$$\frac{\Delta I}{I_t} \ll 1, \tag{11}$$

where $I_t = \dfrac{e}{R_N C}$ is the current at the edge of Coulomb blockade in the I–V characteristic of the small Josephson junction. At a fixed time instant, the phase uncertainty $\Delta\phi$ can be expressed from equation (8) of state of an interferometer as

$$\Delta\phi = 2\pi L \Delta I / \Phi_0. \tag{12}$$

In view of (12), inequality (11) can be recast in the form

$$L/CR_N \gg R_Q. \tag{13}$$

It is assumed in (13) that the phase uncertainty in Eq. (11) is on the order of 2π.

DISCUSSION

In (13), $\tau_N = R_N C$ is roughly equal to the period of Bloch oscillations. It is known that, within this time interval, the mere discharge of a capacitor produced by the electrodes takes place. When the quasi-charge on the capacitor plates approaches an odd number of the elementary electron charge, the superconductor current arises and a Cooper pair is transferred from one electron to another. The transfer occurs very rapidly, so that τ_N virtually defines the period of this process, that is, the frequency of Bloch oscillations. The left-hand side of (13) describes the impedance of the ring at the Bloch oscillation frequency. If the inductance of the interferometer is large, condition (13) is readily fulfilled, because the junction behaves as a single junction, as follows from equation (8) of state of an interferometer. Hence, the condition of high environmental impedance for Coulomb blockade to occur is also valid for a high-inductance interferometer.

For low-impedance interferometers, quantum fluctuations may also be of prime importance. If the inductance is low [6, 10], it shunts the Josephson junction, and thermal fluctuations are thus removed from the junction. Then, at low temperatures, quantum fluctuations begin to play a decisive role and suppress interference in the ring if the inequality [6]

$$L/CR_N \gg R_Q^2/R_N \tag{14}$$

is met.

Thus, as in a single Josephson junction, quantum interference is suppressed by quantum fluctuations and the Coulomb blockade of Cooper pairs arises if the impedance of the interferometer ring is high.

REFERENCES

1. K. K. Likharev, Mikroélektronika **16**, 195 (1987).

2. B. Ruggerio, M. G. Castelliano, G. Torrioli, *et al.*, Phys. Rev. B **59**, 177 (1999).

3. M. H. Devoret, D. Esteve, G. Urbina, J. Martinis, A. Cleland, and J. Clark, *Quantum Tunneling in Condensed Media*, Ed. by Y. Kagan and A. J. Leggett (Elsevier, Amsterdam, 1992), p. 313.

4. D. V. Averin, K. K. Likharev, and A. B. Zorin, Zh. Éksp. Teor. Fiz. **88**, 692 (1985) [Sov. Phys. JETP **61**, 407 (1985)].

5. D. V. Averin and K. K. Likharev, in *Mesoscopic Phenomena in Solids*, Ed. by B. L. Altshuler, P. A. Lee, and R. A. Webb (Elsevier, Amsterdam, 1991).

6. K. K. Likharev, *Introduction to the Dynamics of Josephson Junctions* (Nauka, Moscow, 1985).

7. P. W. Anderson, in *Lectures on Many Body Problems*, Ed. by E. R. Caianiello (1962), p. 113.

8. K. K. Likharev and A. B. Zorin, J. Low Temp. Phys. **62**, 345 (1986).

9. L. S. Kuzmin, Yu. Pashkin, D. S. Golubev, and A. D. Zaikin, Phys. Rev. B **54**, 10074 (1996).

10. I. N. Askerzade, Zh. Tekh. Fiz. **71** (12), 88 (2001) [Tech. Phys. **46**, 1575 (2001)].

Translated by V. Isaakyan

BRIEF
COMMUNICATIONS

Magnetic Properties of Thermally Deposited Fe/GaAs(100) Thin Films

A. A. Veselov, A. G. Veselov, S. L. Vysotsky, A. S. Dzhumaliev, and Yu. A. Filimonov

Institute of Radio Engineering and Electronics, Saratov Branch,
Russian Academy of Sciences, Saratov, 410019 Russia
e-mail: fil@sfire.san.ru
Received January 21, 2002

Abstract—Thin (50–200 Å) films of iron have been prepared on (100) gallium arsenide substrates by thermal evaporation at a deposition rate of 3–30 Å/s and a pressure of ~10^{-5} torr. Dependences of the saturation magnetization, cubic and uniaxial planar anisotropy constants, and the ferromagnetic resonance linewidth on the film thickness were studied by ferromagnetic resonance at 9.8 GHz. It has been found that the parameters of thermally deposited Fe/GaAs (001) films are comparable to those achieved with molecular-beam epitaxy. © 2002 MAIK "Nauka/Interperiodica".

INTRODUCTION

The magnetic properties of thin films of iron grown on gallium arsenide are being extensively studied because of possible applications of such structures in integrated microelectronic devices [1, 2] and spin-dependent nanoelectronics [3], on the one hand, and because of interest in fundamental problems of magnetism [4], on the other. Lattice constants of these materials differ by less than 5%, making possible the oriented growth of Fe films. Using molecular-beam epitaxy (MBE), it is possible to grow on GaAs substrates Fe films of orientations (110) [5] and (100) [6–8]. At thicknesses of $t \sim 30$–200 Å, these films have a high crystalline and magnetic quality that shows up, for example, in the ferromagnetic resonance (FMR) linewidth of $\Delta H \approx 20$ Oe, which is as narrow as in the bulk material. Also, epitaxial Fe/GaAs films grown by ion-beam sputtering were reported [9], which had $\Delta H \approx 20$ Oe only at rather large thicknesses, in excess of 700 Å. Properties of thermally deposited Fe/GaAs have not been investigated as yet. In this work, the possibility has been demonstrated of preparing by thermal vaporization Fe/GaAs(001) films of a thickness of $t \approx 50$–200 Å, which have properties comparable to films of the same thickness grown by MBE.

Note that, to our knowledge, the properties of thermally deposited iron films have been discussed in only two publications. In [10], an MgO substrate was used, and the FMR linewidth of a film 700 Å thick, the thinnest of those studied, was about 40 Oe. In [11], deposition onto a cleaved surface of a lithium fluoride single crystal was carried out; the best result for a film 900 Å thick was 85 Oe. Properties of Fe films of lesser thickness have not been reported at all. Meanwhile, thin films are of interest not only with regard to their applications, but also as an object in which the effect of the

film surface and the interface on the formation of magnetic properties can be studied [6, 8]. In Fe/GaAs films produced by MBE, this effect is most clearly seen in the dependence of the constant of uniaxial planar anisotropy K_u and parameter ΔH on film thickness t. The problem of interest here appears to be identifying the combination of thermal sputtering parameters permitting detection of the effect of the interface on the formation of the properties of evaporated Fe/GaAs (001) at pressures of about 10^{-5} torr in the presence of residual oil vapors.

EXPERIMENTAL

Fe films were prepared in a VUP-5 general-purpose vacuum installation providing a vacuum not worse than 10^{-5} torr. Preliminarily degassed Fe powder was uniformly spread over a tantalum boat positioned in such a manner as to ensure normal incidence of the Fe atoms onto the GaAs substrate. The boat temperature in the evaporation process did not exceed the iron melting point (1536°C) to maintain a sublimation regime. The 4×3 mm substrates were cut out of commercial semi-insulating GaAs (100) wafers with a class 14 surface finish. The samples were cleaned of organic matter by washing in an alkaline solution and distilled water following a procedure described in [7] and clamped in a holder, that could be heated by Joule heat for degassing or maintaining the operating temperature. Close to the holder, a quartz plate was placed whose resonance frequency was measured before and after the sputtering and, using a known method of thickness measurements [12], the film deposition rate was verified and its thickness determined.

The sputtering process was carried out as follows. The substrate was degassed for 5 min at a temperature
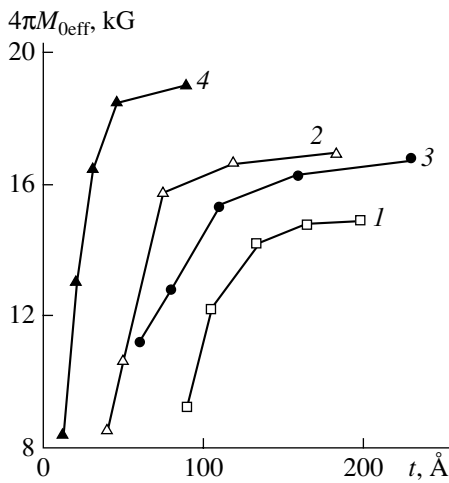
$4\pi M_{0eff}$, kG

Fig. 1. Dependence of the saturation magnetization $4\pi M_{0eff}$ on film thickness.

of 550°C and then left to cool down. Meanwhile, the boat was heated to a working temperature, its attainment being judged both by the power consumed and visually, by observing through a sight glass the appearance of an evaporated layer on the polished shutter surface. At the selected substrate temperature (reached either in the cooling process or by heating the holder), which was controlled by a thermocouple, the shutter was opened for a certain period of time so that the sputtered iron atoms could reach both the substrate and the quartz plate. Magnetic parameters of the films were determined by FMR at a frequency of ≈9.8 GHz with the magnetic field modulated at a frequency of 30 kHz. The FMR linewidth was determined as the distance between the extrema of the observed dependence of the derivative of reflected power on the magnetic field, the midpoint between the extrema giving the value of the resonance field H_r. From comparison of the experimen-

tal and calculated dependences of H_r on the angle θ between the direction of the external constant magnetic field H_0, applied tangentially relative to the structure surface and the [100] axis, values of the effective saturation magnetization $4\pi M_{0eff}$ and constants of the cubic K_1 and uniaxial K_u anisotropy were determined as in [7].

RESULTS AND DISCUSSION

In Figs. 1–4, dependences of $4\pi M_{0eff}$, K_1, K_u, and ΔH on the film thickness obtained under different conditions are shown. Curves *1* and *2* were obtained for evaporations on substrates that were allowed to cool after degassing and then, before opening the shutter, heated to ~50 or ~150°C; we refer to these procedures as regimes 1 and 2, respectively. Curves *3* were obtained in the case where the shutter was opened as the substrate cooled to a temperature of ~150°C after degassing (regime 3). For comparison, results obtained with MBE growth [7] are given (curves *4*). Note that the variation of magnetic parameters with film thickness corresponds in general to data in [5–7].

It is seen in the figure that with the technological regime 2, the best quality films are obtained. Saturation of the $4\pi M_{0eff}(t)$ (Fig. 1) dependence takes place already at thicknesses of about 80 Å, comparable to values for the MBE growth (curve *4*) though with $4\pi M_{0eff} \approx 16$ kG lower by 10%. The reason for this is, apparently, the greater content of foreign inclusions in the lattice because of a worse vacuum. In regime 1, the substrate temperature rises from the beginning of sputtering due to thermal radiation from the boat. Because no special measures were taken to stabilize the temperature in the sputtering process, the temperature rise stopped only when a thermal equilibrium had been reached with the surrounding armature at ≈150°C (Fig. 5). Probably, this is the reason why films less than 90 Å thick, which grew in the constantly changing con-
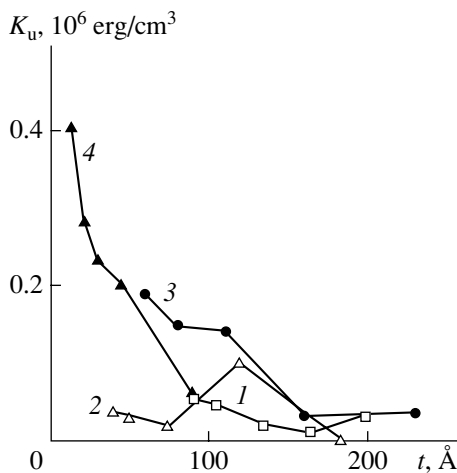
K_u, 10^6 erg/cm^3

Fig. 2. Dependence of the uniaxial anisotropy constants K_u on film thickness.

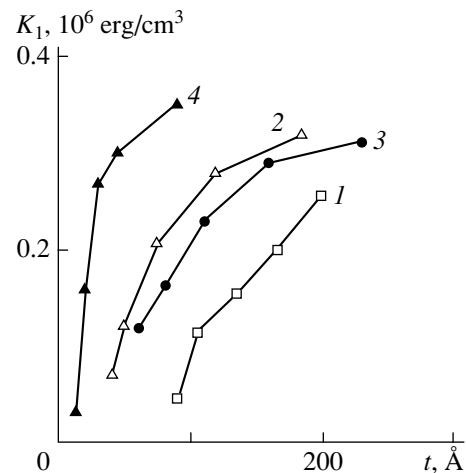K_1, 10^6 erg/cm^3

Fig. 3. Dependence of the cubic anisotropy constants K_1 on film thickness.
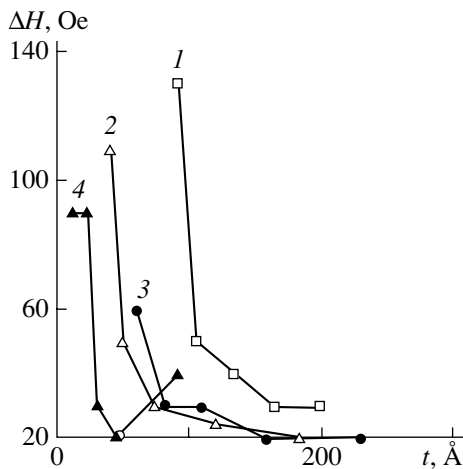
Fig. 4. Dependence of the FMR linewidth ΔH on film thickness.

Fig. 5. Variation with time of the substrate temperature for sputtering in regime 1.

ditions, had such an imperfect structure that they could not be studied by the FMR method. Note that in regime 2 the substrate temperature stayed practically constant.

The results obtained are remarkable in that the film quality was found to be dependent not only on the stability of the substrate temperature but also on its thermal prehistory. It is seen in Fig. 2 that although the substrate temperatures for sputtering in regimes 2 and 3 are the same, the $K_u(t)$ curves differ considerably. Curve 2, except for a local maximum at $t \approx 100$ Å, in general demonstrats a stable and low value of this constant, which is evidence of the high degree of cubic symmetry, whereas K_u for regime 3 increases appreciably with decreasing thickness as it does at $t < 100$ Å in MBE technology (Fig. 3, curves 3 and 4). At the same time, the cubic anisotropy curves (curves 2 and 3 in Fig. 3) have similar shapes, although at equal thicknesses, curve 2 corresponds to higher K_1 values.

Let us discuss possible causes of the different behavior of the $K_u(t)$ curves for films evaporated in regimes 2 and 3. In [6], the increase of uniaxial anisotropy with decreasing film thickness was attributed to the effect of the nonequivalent crystallographic directions [110] and [1$\bar{1}$0] in the GaAs (100) substrate, which, in turn, determined the preferred growth direction of islands at the initial stages of film growth. The authors indicated that this effect can be the largest for Ga- or As-stabilized surfaces of GaAs (100). It is possible that this was precisely the case in our experiments: in regime 3 as a result of evaporation (during degassing) of the more volatile arsenic, formation of a layer rich in gallium could take place, leading to the emergence of uniaxial anisotropy, whereas in a substrate that cooled before the evaporation, the equilibrium between gallium and arsenic could be restored due to diffusion
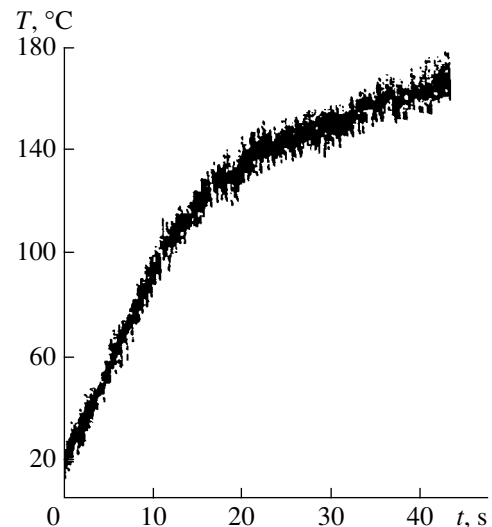
processes. So, the advantage of regime 3, which eliminates readsorption of residual gases at the substrate surface, turns out to be a hindrance to the growth of cubic anisotropic crystals but can be useful where films with strong uniaxial anisotropy are required.

Figure 4 shows the dependence on thickness of the FMR linewidth. It is seen that, regarding this parameter, the regimes corresponding to curves 2 and 3 also yield better results than regime 1; ΔH values in a thickness range of $70 < t < 200$ Å do not exceed those achieved with MBE while displaying a monotonic rather than resonant character of the dependence on thickness.

CONCLUSION

Thus, it has been demonstrated that thin (50–200 Å) films of iron with a narrow (20–60 Oe) ferromagnetic resonance linewidth can be grown on gallium arsenide (100) substrates by thermal sputtering. Technological regimes of sputtering have been identified, which yield dependences of the saturation magnetization and cubic and uniaxial anisotropy constants on thickness close to those of films grown by MBE. It was shown that the emerging anisotropic properties of iron films are affected not only by the substrate temperature, but also by its thermal prehistory.

ACKNOWLEDGMENTS

REFERENCES

1. E. Schloemann, R. Tuistison, J. Weissman, *et al.*, J. Appl. Phys. **63**, 3140 (1988).

2. C. S. Tsai, Jun Su, and C. C. Lee, IEEE Trans. Magn. **35**, 3178 (1999).

3. S. A. Wolf and D. Treger, IEEE Trans. Magn. **36**, 2748 (2000).

4. Hans J. Hag, B. Stiefel, A. Moser, *et al.*, J. Appl. Phys. **79**, 5609 (1996).

5. G. A. Prinz and J. J. Krebs, Appl. Phys. Lett. **39**, 397 (1981).

6. J. J. Krebs, B. T. Jonker, and G. A. Prinz, J. Appl. Phys. **61**, 2596 (1987).

7. S. L. Vysotskiĭ, A. S. Dzhumaliev, Yu. A. Filimonov, *et al.*, Radiotekh. Élektron. (Moscow) **45**, 209 (2000).

8. S. A. Oliver, C. Vittoria, E. Schloemann, *et al.*, J. Appl. Phys. **63**, 3802 (1988).

9. Yu. V. Goryunov, N. N. Garif'yanov, G. G. Khaliullin, *et al.*, Phys. Rev. B **52**, 13450 (1995).

10. L. Thomas, S. S. P. Parkin, Jun Yu, *et al.*, Appl. Phys. Lett. **76**, 766 (2000).

11. Y. Gondo, S. Usami, K. Itoh, *et al.*, J. Appl. Phys. **34**, 1081 (1963).

12. A. B. Vaganov and V. V. Vasilevskiĭ, Fiz. Tverd. Tela (Leningrad) **16**, 2044 (1974).

13. *Handbook of Thin-Film Technology*, Ed. by L. I. Maissel and R. Glang (McGraw-Hill, New York, 1970; Sov. Radio, Moscow, 1977).

Translated by B. Kalinin

GAS DISCHARGES, PLASMA

Structure of Luminous Clouds Near a Sublimated Hydrogen Pellet

B. V. Kuteev, A. Yu. Kostryukov, and O. A. Bakhareva

St. Petersburg State Technical University, St. Petersburg, 195251 Russia
e-mail: kuteev@stu.neva.ru
Received November 28, 2001; in final form, February 7, 2002

Abstract—The spatial structure of a luminous cloud near a hydrogen pellet sublimated in a tokamak plasma is investigated experimentally. The luminous cloud has the form of a constant-intensity plateau extended along the magnetic field and is surrounded by an exponentially fading halo. The cross size of the plateau is determined by the boundary of the ionization region of atoms leaving the pellet surface. The plateau length is approximately determined by the balance between the pressures in the luminous cloud and the bulk plasma. Taking into account the structure described is important for the correct interpretation of the signals from alpha-particle diagnostics by charge-exchange neutral fluxes, because these signals depend on the characteristic dimensions of the cloud. © 2002 MAIK "Nauka/Interperiodica".

A method for injecting pellets consisting of cooled, solid-state hydrogen isotopes into a plasma [1] is developed in order to create an efficient system for delivering the fuel to the plasma core in tokamaks and stellarators. A neutral plasma cloud produced due to the sublimation of the pellet acts as a screen for thermal fluxes from the plasma onto the pellet and, thus, determines the sublimation rate. The structure of the radiating cloud was studied in a number of papers. Photographs of the cloud were obtained over a wide range of plasma and pellet parameters. From the results of spectroscopic measurements of an optically thin cloud, it was found that the luminous region was in local thermal equilibrium, the electron density was about 10^{17}–10^{18} cm^{-3}, and the electron temperature was 1–4 eV [2–5]. At the same time, only a few attempts were made to experimentally study the spatial distribution of the cloud radiation intensity [6].

An important parameter that influences the results of calculations of the sublimation rate in the plasma screening models [7, 8] is the ionization radius of atoms in the luminous cloud. The authors usually compare the calculated ionization radius with the full cross size of the luminous cloud observed experimentally. At the same time, the spatial distribution of the cloud radiation intensity is not considered.

In this paper, photographs of luminous clouds obtained during the injection of hydrogen pellets into the T-10 tokamak [6] are analyzed and an attempt is made to explain physical mechanisms for the formation of the cloud structure.

Figure 1 shows a typical intensity distribution of optical radiation of a cloud produced from a 0.05-cm-radius hydrogen pellet moving toward the center of the plasma with a velocity of $V_p = 710$ m/s. The plasma electron temperature at the point of the pellet location was 500 eV, and the plasma density was 3.5×10^{13} cm^{-3}. The figure shows the contour lines of the intensity on the logarithmic scale. During an exposure time of 5×10^{-6} s, the pellet moving at an angle of 15° to the observation axis was displaced by 0.9 mm in the image plane of the photographic camera, which was much less than the cloud size. Consequently, the particle displacement during the exposure time can be ignored.

In photographs, a narrow cigar-shaped plateau region in the center of the cloud is seen. The intensity of this plateau is nearly constant. The cross size of the plateau $2r_{pl}$ is about 2–4 mm, which is larger than the distance 0.9 mm, by which the cloud is displaced during the exposure time. The longitudinal size of the plateau L_{pl} is about 4 cm. The plateau is surrounded by a halo in which the intensity falls exponentially away from the plateau both along and across the magnetic field.

The cloud structure calculated by the model [9] for typical experimental conditions of the T-10 tokamak is shown in Fig 2. The density and the temperature are determined by the gas-dynamic expansion of the neutral cloud. The degree of ionization and the intensity of emission in the H_α line are calculated in the local thermodynamical equilibrium model. In the experiment, the total visible radiation from the cloud was recorded on a film. The H_α line comprises a considerable fraction of the visible radiation from the cloud. That is why, for estimates, we restricted ourselves to calculating the intensity of the H_α line.

According to calculations, the luminous cloud forms a hollow shell nearly 4 mm in diameter. Since the light is integrated along the line of sight, this shell around the pellet is recorded on the photograph as a pla-
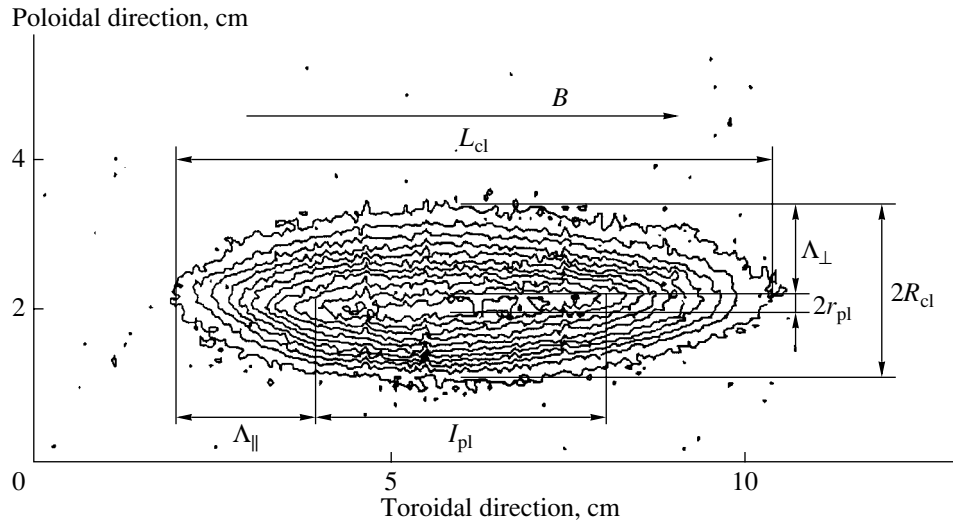
Poloidal direction, cm

Fig. 1. Intensity distribution of optical radiation from a pellet cloud: L_{pl} is the length of a plateau where the intensity is nearly constant, r_{pl} is the plateau radius, L_{cl} is the full length of the cloud observed, R_{cl} is the radius of the cloud observed, and $\Lambda_\perp$, $\Lambda_\parallel$ are the transverse and longitudinal widths of the exponential halo.

teau (with a nearly constant intensity) surrounded by a fading halo. The 4-mm diameter of the spherical luminous shell predicted by the model [9] nearly coincides with the plateau cross size observed in the experiment.

In Fig. 2, it is seen that, because of the large temperature gradient in the cloud, there is a sharp ionization boundary, which lies near the maximum-luminosity region. Therefore, the position of the ionization boundary can be determined from the photograph as a region where the plateau transforms into the exponential halo.

A simple estimate for the ionization radius r_{ion} can be obtained from the balance between the energy absorbed by the cloud and the energy required for the heating and acceleration of the cloud and the dissociation and ionization of neutrals [7],

$$E_{loss}dN/dt = 2\pi r_{ion(1)}^2 \frac{1}{4}n_e\sqrt{\frac{8T_e}{\pi m_e}}2T_e. \qquad (1)$$

Here, dN/dt is the sublimation rate and $E_{loss} \approx 0.011$ eV + 1eV + 1eV + 2 eV + 13.6 eV is the energy expended on sublimation, the heating of the cloud, its acceleration, dissociation, and ionization; the temperature T_e is in energy units. Based on the results of measurements, which showed that the temperature in the maximum-luminosity region is $T_{cl} \approx 1.0$–1.5 eV [4, 5] and the cloud expansion velocity is $v_{el} \approx 10^6$ cm/s [2] (which corresponds to an energy of 1 eV), the energy of the directed and thermal motion of the cloud atoms is estimated at 1 eV + 1 eV. The model [9] also gives $T_{cl} \approx$ 1.0–1.5 eV (Fig. 2) and $v_{cl} = 10^6$ cm/s. We ignore the radiation losses because, in the local thermodynamical equilibrium, the rates of radiative processes are lower than those of collisional ionization (recombination) and excitation (de-excitation) [10].

To estimate the ionization radius by formula (1), it is necessary to independently calculate the sublimation rate. In what follows, Parks scaling [11] is used for this purpose.

Figure 3 shows the ionization radius $r_{ion(1)}$ calculated by formula (1) for the pellet sublimation in the T-10 shot no. 49968 and the plateau radius r_{pl} measured from the photographs as functions of the minor radius r. The figure also shows the profile of the calculated sublimation rate dN/dt in the T-10 tokamak. As is seen, the estimates for the sublimation boundary agree well with the measured plateau size. For minor radii of $12 < r < 16$ cm, we observed the stratification of the cloud, namely, the separation of the cloud in the poloidal direction. For this reason, we obtained an overestimated value of the cloud radius in this region.

Calculation of the ionization boundary from the power balance provide only a lower estimate, because the energy lost before the plasma flow reaches the most luminous region is ignored and it is assumed that the plasma heat flux arriving at this region is completely absorbed there. However, the fact that calculations by formula (1) describe well the size of the ionization boundary allows us to conclude that it is the most luminous region where the energy is mainly lost. The spherical model [9] allows us to adequately describe the expansion of the cloud up to the ionization boundary. The model can also be used to calculate the neutral screening, because this screening occurs in a thin spherical shell near the pellet surface. It turns out that the model predicts a reasonable transverse size of the cloud, but it cannot adequately describe the cloud expansion after ionization.

Let us consider the longitudinal expansion of the cloud after ionization. From photographs and measure-

ments carried out in [2, 5, 10], it is evident that the intensity of cloud radiation changes only slightly over several centimeters along the plateau. We assume that the electron density and temperature are equalized along the plateau surface. This is because the main source of the plateau radiation is the sharp ionization boundary extended along the magnetic field. The deceleration of the transverse expansion of the ionized cloud results in the elongation of the cloud along the magnetic field in such a way that the plasma parameters (temperature, density, and expansion velocity) at the plateau boundary nearly coincide with the values predicted by the spherical model for the ionization boundary.

Let us estimate the pressure inside a homogeneous cylindrical cloud of length L_{pl} and radius r_{pl},

$$p_{cl} = N_{cl}T_{cl} = \left(\frac{dN/dt}{2\pi r_{pl}L_{pl}v_{cl}}\right)T_{cl}. \quad (2)$$

Here, the values of r_{pl} and L_{pl} are determined from photographs, the sublimation rate is calculated by Parks scaling [11], and the temperature $T_{cl} = 1$ eV and the expansion velocity $v_{cl} = 10^6$ cm/s are assumed to be constant during the entire sublimation process (their values were found experimentally in [2]). The results of calculations of different phases of the pellet sublimation are shown in Fig. 4. It turns out that, under T-10 conditions, the pressure p_{cl} in the luminous cloud is close to the pressure in the external plasma $p_{ext} = n_iT_i + n_eT_e$ in all the phases of pellet sublimation, i.e., over a wide range of pellet sublimation rates and pressures p_{cl} and p_{ext}.

The presence of a cigar-shaped plateau surrounded by a halo in which the intensity falls exponentially is typical not only of radiation from hydrogen clouds, but also of various ionization states of impurity clouds, e.g., for carbon clouds in plasmas of the T-10 tokamak [6] and the Wandelstein-AS stellarator [12]. For the application of pellet injection to, e.g., the pellet diagnostic of fast plasma particles [13] (which is based on the charge exchange of fast particles in the sublimation cloud with the following detection of these particles by a neutral particle analyzer), the characteristic parameters of the cloud are the characteristic lengths on which the radiation intensity falls, rather than the visible cloud dimensions. Below, we will show that, taking into account this fact, it is possible to refine the results of the measurements of the energy spectrum of high-energy alpha-particles [14].

To determine the energy spectrum of high-energy alpha-particles from the measured energy dependence of the count rate of charge exchange neutrals, it is necessary to know the cross size of the pellet cloud. In [14], the cross size of a Li^+ cloud (in which, presumably, most charge-exchange events occur) was estimated at several centimeters. Apparently, this estimate was obtained from the width of the visible part of the cloud (Fig. 5). The figure is taken from [14] and shows an
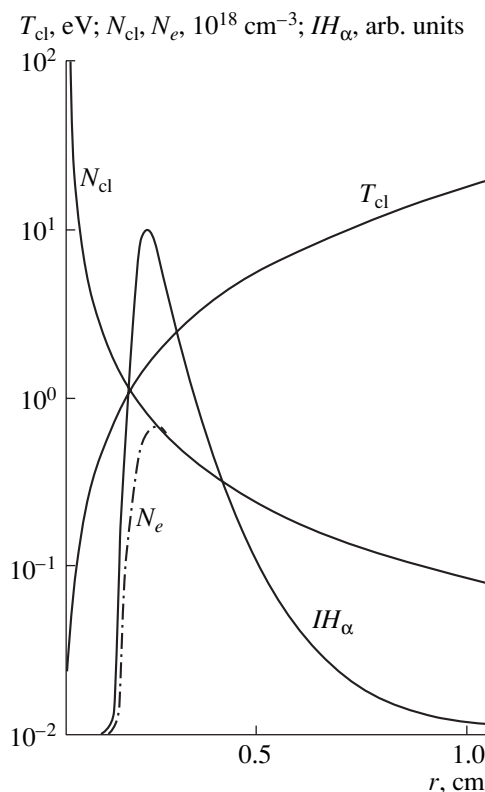
Fig. 2. Profiles of the cloud temperature T_{cl} and the total atom and ion density N_{cl} calculated by the model [9] for the following sublimation parameters: an H_2 pellet, $r_p = 0.05$ cm, $T_e = 500$ eV, $n_e = 3.5 \times 10^{13}$ cm^{-3}, $dN/dt = 1.7 \times 10^{23}$ atom/s (r is the distance to the pellet, N_e is the electron density, and IH_α is the intensity of the H_α line).
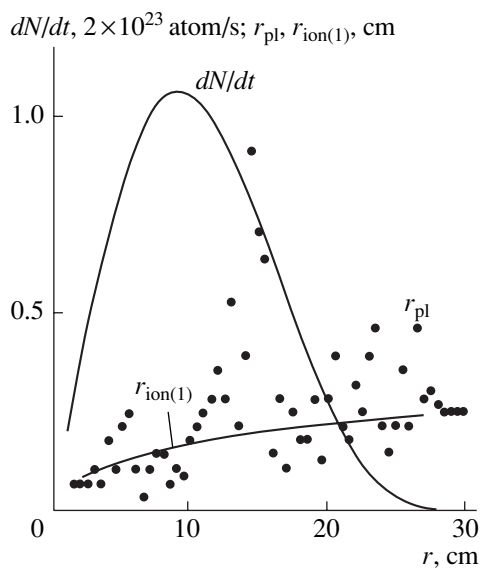
Fig. 3. Sublimation rate and the cloud cross size vs. minor radius in the T-10 tokamak (shot no. 49 968): an H_2 pellet, $r_p = 0.056$ cm, $V_p = 710$ m/s, $T_e(0) = 770$ eV, $n_e(0) = 4.7 \times 10^{13}$ cm^{-3}, and max$(dN/dt) = 2.2 \times 10^{23}$ atom/s.
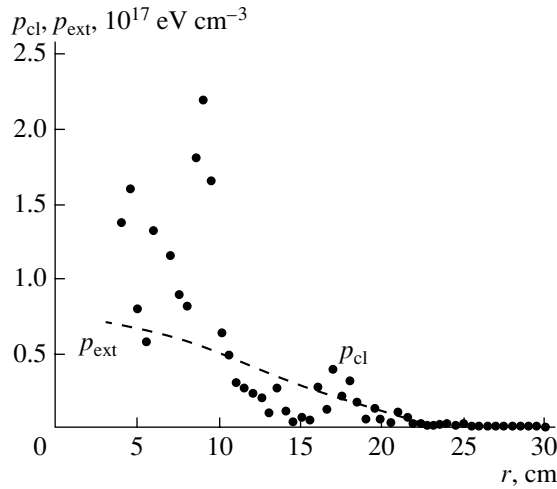
$p_{cl}, p_{ext}, 10^{17} eV cm^{-3}$

Fig. 4. Pressures in the luminous cloud and in the external plasma, p_{cl} (circles) and p_{ext} (dashed curve), vs. minor radius r in the T-10 tokamak (shot no. 49 968).

example of the singly ionized part of a lithium pellet cloud in TFTR (shot no. 52 197, $n_e(0) \approx 6.8 \times 10^{13}$ cm^{-3}, $T_e(0) \approx 2.3$ keV). For reasons listed below, we believe that in [14] the cross size r_{cl} of the Li$^+$ cloud was overestimated and that r_{cl} was really on the order of several millimeters.

(i) In the photographs of the pellet cloud [6] obtained during the sublimation of carbon pellets in the T-10 plasma, it is seen that both characteristic cross sizes of the carbon cloud (the plateau size and the ionization length) lie in the millimeter range.

(ii) In order to estimate the cross size of the lithium cloud for TFTR (shot no. 52 197), let us calculate the length over which Li^0 atoms are singly ionized,

$$r_{ion(2)} = c_s \tau_{ion}, \qquad (3)$$

where

$$c_s = \sqrt{\frac{5}{3}\frac{T_{cl}}{m_a}}$$

is the expansion velocity of sublimated atoms, $T_{cl} \sim 1$ eV is the cloud temperature near the surface of the Li pellet, m_a is the mass of an Li^0 atom,

$$\tau_{ion} = \frac{1}{n_e R_1(T_e)},$$

n_e (cm^{-3}) and T_e (eV) are the electron density and temperature in the ambient plasma,

$$R_1(T_e) = \sqrt{\frac{I_1}{T_e}}\left(4.5456 \times 10^{-8} \ln\left(\frac{T_e}{I_1}\right)\right.$$
$$\left. + \sum_{n=0}^{2} \beta 1_n \left(\frac{I_1}{T_e}\right)^n\right)[cm^3/s]$$

is the single-ionization rate, $I_1 = 5.39$ eV, $\beta 1_0 = 2.78 \times 10^{-7}$, $\beta 1_1 = -1.583 \times 10^{-6}$, and $\beta 1_2 = 5.465 \times 10^{-6}$ [15].

In Fig. 6, the cloud cross size $r_{ion(2)}$ estimated as an ionization length is shown as a function of the effective minor radius in TFTR.

(iii) The cloud cross size in TFTR (shot no. 52 197) was estimated also from energy balance equation (1), where the energy expended per lithium atom was assumed to be $E_{loss} \approx 2.089$ eV + 1 eV + 1 eV + 0 eV + 5.39 eV and the sublimation velocity dN/dt was calculated by the neutral and electrostatic screening model [16]. As $r_{ion(1)}$, we chose the largest of two values—the value defined by formula (1) and the pellet radius. The resulting profile $r_{ion(1)}$ is shown by the solid line in Fig. 6. It can be seen that, for shot no. 52 197, the cloud cross size estimated from either the energy balance or the length of ionization is on the order of one millimeter and is one order of magnitude higher than the estimate obtained in [14].

In [14], the alpha-particle energy distribution function in TFTR (shot no. 86 297) was measured with the help of a pellet diagnostics of fast particles. The result is shown by dark squares in Fig. 7 (Fig. 15 in [14]). The measured magnitude of the distribution function was one order of magnitude lower than that calculated theoretically by the FPPT code (Fig. 7, solid line); how-
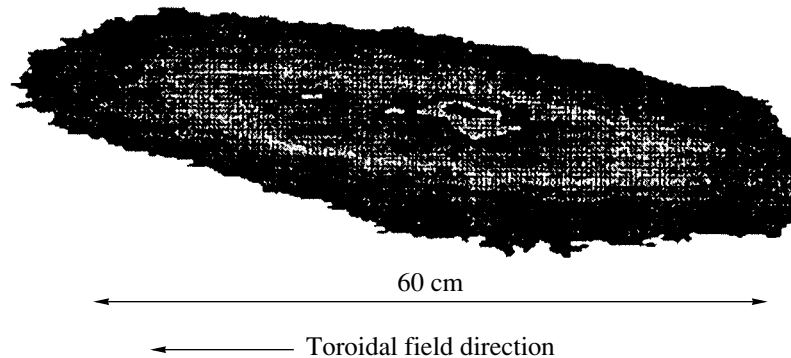

60 cm

Toroidal field direction

Fig. 5. Image of a Li$^+$ cloud in the 548.5-nm line in TFTR (shot no. 52 197) [14].
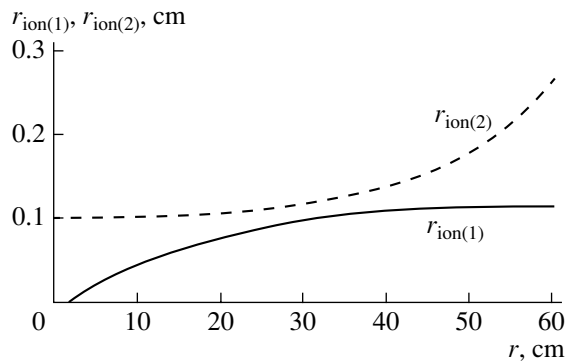
Fig. 6. Cross size of a Li^+ cloud in TFTR (shot no. 52 197) vs. minor radius r.
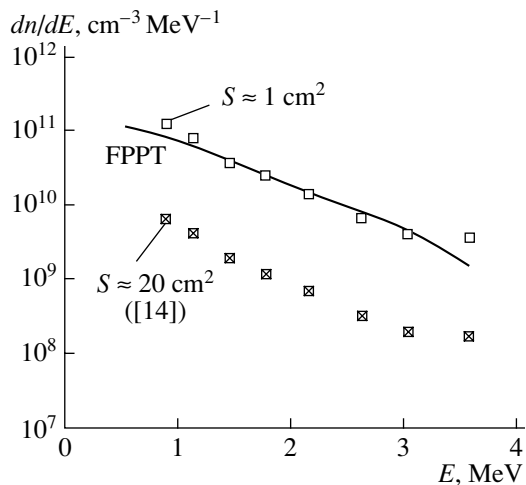

Fig. 7. Comparison of the measured and calculated energy spectra of fast alpha-particles in TFTR (shot no. 86 291).

ever, the shapes of the experimental and theoretical energy spectra were the same. We suggest that one of the causes of this discrepancy is the overestimated value of the area S of the cloud region from which charge-exchange atoms detected by a neutral particle analyzer (NPA) are emitted. In the TFTR experiment analyzed in [14], the radius of the spatial region viewed by the analyzer around the pellet was $r_{NPA} \sim 2.5$ cm. In [14], it was assumed that the cross size of the cloud surrounding a boron pellet exceeded 2.5 cm and the area of the visible region of the cloud was estimated as $S = \pi(r_{NPA})^2 \approx 20$ cm^2. From estimates similar to those above, the cross size of the cloud region contributing mainly to the NPA signal ($r_{cl} \approx 1$ mm) is much less than r_{NPA}. Therefore, the effective cloud area from which neutrals are collected depends on the cloud cross size as $S = 4r_{cl}r_{NPA} \approx 1$ cm^2. The count rate of particles in a certain energy range is proportional to the visible cloud area. Hence, if $S \approx 1$ cm^2 is taken instead of $S \approx 20$ cm^2 in calculating the energy distribution from the neutral count rate, then, for shot no. 86 291, we obtain the curve shown

in Fig. 7. This energy spectrum almost coincides with the theoretically calculated spectrum (Fig. 7, solid line).

Therefore, the structural analysis of hydrogen clouds in the T-10 tokamak has shown that the ionization boundary in the transverse direction is well defined and nearly coincides with the region where the radiation intensity is maximum. The size of this region can be measured from the cloud photographs. These measurements allow one to check the results of calculations by the codes that account for the transverse deceleration of the cloud by the $\mathbf{j} \times \mathbf{B}$ force, arising when the ionized cloud moves across the magnetic field.

The ionization boundary in the longitudinal direction approximately corresponds to the balance of the pressures in the expanding cloud and the bulk plasma.

The characteristic cloud cross size, which is on the order of 1 mm, is typical not only of hydrogen, but also of light-impurity pellets. When the characteristic cloud structure is taken into account, a better correspondence can be achieved between the energy spectra of fast alpha-particles measured with the pellet diagnostics in the TFTR tokamak [14] and those calculated theoretically.

REFERENCES

1. S. L. Milora, W. A. Houlberg, L. L. Lenguel, *et al.*, Nucl. Fusion **35**, 657 (1995).
2. R. D. Durst, W. L. Rowan, R. A. Austin, *et al.*, Nucl. Fusion **30**, 3 (1990).
3. V. Yu. Sergeev, A. Yu. Kostrukov, and S. A. Shibaev, Fusion Eng. Des. **34–35**, 323 (1997).
4. TFR Group, Europhys. Lett. **2**, 267 (1986).
5. D. H. McNeill, G. J. Greene, J. D. Newburger, *et al.*, Phys. Fluids B **3**, 1994 (1991).
6. S. M. Egorov, V. A. Galkin, V. G. Kapralov, *et al.*, in *Proceedings of the 13th International Conference on Plasma Physics and Controlled Nuclear Fusion Research, Washington, 1990*, Vol. 1, p. 599.
7. W. A. Houlberg, S. E. Milora, and S. E. Attenberger, Nucl. Fusion **28**, 595 (1988).
8. L. Garzotti, B. Pegourie, A. Geraud, *et al.*, Nucl. Fusion **37**, 1167 (1997).
9. B. V. Kuteev and L. D. Tsendin, Fusion Technol. **26**, 642 (1994).
10. D. H. McNeill, G. J. Greene, and D. D. Schuresco, Phys. Rev. Lett. **55**, 1398 (1985).
11. P. B. Parks and R. J. Turnbull, Phys. Fluids **21**, 1735 (1978).
12. V. Yu. Sergeev, V. M. Timokhin, O. A. Bakhareva, *et al.*, Europhys. Conf. Abstr. **25A**, 1953 (2001); http://www.cfn.ist.utl.pt/EPS2001/fin/pdf/P4.023.pdf.
13. S. S. Medley, D. K. Mansfield, A. L. Roquemore, *et al.*, Rev. Sci. Instrum. **67**, 3122 (1996).
14. J. M. McChesney, P. B. Parks, R. K. Fisher, *et al.*, Phys. Plasmas **4**, 381 (1997).
15. K. L. Bell, H. B. Gilbody, G. U. Hujhes, *et al.*, J. Phys. Chem. Ref. Data **12**, 891 (1983).
16. B. V. Kuteev, Nucl. Fusion **35**, 431 (1995).

Translated by N. Larionova

Determination of the Electron Energy Distribution in a Plasma from the First and Second Derivatives of the Probe Current

N. A. Gorbunov, A. N. Kopytov, and F. E. Latyshev

Research Institute of Physics, St. Petersburg State University,
Ul'yanovskaya ul. 1, Petrodvorets, St. Petersburg, 198504 Russia
e-mail: gorbunov@paloma.spbu.ru
Received January 23, 2002

Abstract—An approach to the reconstruction of the electron energy distribution function (EEDF) as the sum of the second derivative of the electron current to the probe and some value proportional to its first derivative is proposed (hereafter we mean derivatives with respect to voltage). Solutions to model problems for typical electron distributions in a plasma show that this approach lowers the systematic error of EEDF reconstruction by several times in comparison with the conventional techniques using Langmuir and diffusion probes. The approach is applicable in a wide range of the ratio of the probe radius to the free path of electrons. It can be used for the determination of fast nonequilibrium electron distribution in neon discharge afterglow and also for the evaluation of the Maxwellian electron temperature in the low-energy range of nitrogen discharge afterglow. © 2002 MAIK "Nauka/Interperiodica".

INTRODUCTION

At present, probe techniques [1] are widely used for the determination of the electron energy distribution function in a plasma. Those based on the evaluation of the probe current derivatives as a function of the probe potential V have enjoyed the widest application. The method that employs the second derivative is used at low pressures of a neutral gas when the probe radius a is much shorter than the free path λ of electrons (Langmuir probe). The one using the first derivative [2] is valid at high pressures of a neutral gas when $a \gg \lambda$ (the diffusion method).

To date, various approaches to finding the derivatives have been developed. Among them are methods of numerical differentiation, probe current modulation, and differentiation with amplifiers [1]. These offer a high time resolution, accumulate the desired signal, and separate it out of noise. The determination of the probe current derivatives provides information on the high-energy part of the EEDF at potentials exceeding the potential V_p of an isolated probe. This is because the derivative of the probe I–V characteristic is governed by the electron current in a wide energy range $V > V_p$, where the current of positive ions virtually saturates and, therefore, depends weakly on V. This is an obvious advantage of the differential method for EEDF determination over the integral method in which the EEDF is found immediately from the integral equation relating the EEDF with the electron current to the probe. In the latter case, the EEDF determination at $V > V_p$ is practically impossible.

However, the methods of EEDF determination from the first and second derivatives of the probe current suf-

fer from unavoidable systematic errors associated with the finite ratio a/λ in a real experiment. The errors were analyzed for different forms of the EEDF and values of a/λ [3]. It was shown that the greatest distortions of the EEDF arise near the plasma potential and in the nonequilibrium high-energy part of the EEDF.

To diminish the systematic error, it was proposed that the EEDF from integral equations relating the EEDF and the probe electron current be found [4]. Such a procedure belongs to the class of ill-posed problems that can be solved by regularization techniques. In this case, it is difficult to relate the solution and the random error of a measurement.

Another approach to the error problem is to introduce corrections to the second derivative [5, 6]. In many cases, this way allows one to find the EEDF without considerable distortions and to avoid the need for regularization algorithms. However, such an approach is valid only for $a/\lambda < 1$ and for EEDFs decreasing monotonically with increasing energy.

In this paper, we suggest a method for finding the EEDF from the first and second derivatives of the probe electron current. Its efficiency is demonstrated by solving model problems for various forms of the EEDF in a wide range of a/λ. The method can be used for the determination of the EEDF in neon and nitrogen afterglow, where conventional techniques distort the EEDF considerably.

METHOD FOR FINDING THE EEDF

In [3], a formula that relates the probe electron current to the EEDF was derived on the basis of the kinetic

Table

Model EEDF	ψ	ξ		
		$f_1 = -\dfrac{\psi}{V}\dfrac{dJ}{dV}$	$f_2 = \dfrac{d^2 J}{d(V)^2}$	$f_3 = f_1 + f_2$
Maxwell EEDF	0.2	0.871	0.142	0.047
	1	0.676	0.440	0.133
$\psi = $ const	5	0.388	0.782	0.162
Step	2	0.715	0.711	0.307
Maximum	10	1.347	1.571	0.402
Maxwell EEDF in nitrogen, $T_e = 0.6$ eV	$\psi(T_e) = 0.5$	0.682	0.388	0.108

approach. It was obtained for the condition $\lambda_\varepsilon > a + d$, where λ_ε is the energy relaxation length, a is the probe radius, and d is the space charge thickness. In the practically important case of thin space charge, $a \gg d$, this formula has the form

$$j_e(eV) = C_0 \int\limits_{eV}^{\infty} \frac{(\varepsilon - eV) f(\varepsilon) d\varepsilon}{1 + \psi(\varepsilon)\left(1 - \dfrac{eV}{\varepsilon}\right)}, \tag{1}$$

where $C_0 = (8\pi e)/(3m^2\gamma_0)$ is the normalization factor; e and m are the electron charge and mass, respectively; and V is the probe potential with respect to the plasma.

The value of $\gamma_0(a/\lambda)$ decreases monotonically as the argument increases and varies in the range $4/3 \geq \gamma_0 \geq 0.71$. This quantity takes the asymptotic values $\gamma_0 = 4/3$ at $a \ll \lambda$ (Langmuir probe) and $\gamma_0 = 0.71$ at $a \gg \lambda$ (diffusion probe). The probe diffusion parameter is $\psi = ac_i/\gamma_0\lambda$, where $c_i = 1$ for a sphere and $c_i = \ln\left(\dfrac{\pi l}{4a}\right)$ for a cylindrical probe of length l. It was also shown [3] that formula (1) includes both the Langmuir and the diffusion limits.

A formula analogous to (1) was derived in [7] from the qualitative analysis of a probe-induced plasma disturbance. A similar expression was obtained in [8] by solving the moment equation for the EEDF. These formulas differ from each other only in values of the coefficient γ_0.

Adding the first and second derivatives of Eq. (1) yields

$$f(V) = \frac{1}{C_0}\left[j''(V) - j'(V)\frac{\psi}{V}\right]$$

$$+ \int\limits_{V}^{\infty} \frac{2\psi\varepsilon^2 f(\varepsilon) d\varepsilon}{[\varepsilon(1 + \psi) - \psi V]^3} - \frac{\psi}{V}\int\limits_{V}^{\infty} \frac{\varepsilon^2 f(\varepsilon) d\varepsilon}{[\varepsilon(1 + \psi) - \psi V]^2}, \tag{2}$$

where eV is replaced by V for brevity.

The first two terms in (2) are a combination of the Langmuir and diffusion limits, which relate the EEDF and the derivatives of the probe current. In fact, following the qualitative reasoning in [7], one can split the space around the probe into zones of free-transit and diffusion motions of electrons. The boundary between them lies approximately at a distance $r \approx \lambda$ from the probe surface. In the first approximation, the contributions of both zones to the probe current are additive. If $\psi \ll 1$ (Langmuir probe), free-transit electrons reaching the probe from the region $r < \lambda$ play a dominant role. The contribution of these electrons is described by the first term in (2). Electrons coming from distances $r > \lambda$ "see" the probe at a small space angle; therefore, the contribution of the diffusion component, described by the second term in (2), is small. As the pressure of the neutral gas grows, the ratio a/λ becomes greater than unity; then, from a distance $r \approx \lambda$, the probe is seen at a sufficiently large space angle. In this case, the diffusion component increases, while the free-transit component decreases. This change is reflected by a growth of $\psi > 1$ with increasing pressure. Thus, one can conclude that the difference between the integral terms in (2) refers to electrons reaching the probe from the layer $r \approx \lambda$. These electrons experience single collisions and are neither free-transit nor diffusion. In many cases, their contribution is minor and we have

$$f(V) \approx \frac{1}{C_0}\left[j''(V) - j'(V)\frac{\psi}{V}\right]. \tag{3}$$

Consider first the simplest case $\psi = $ const. To some extent, this situation occurs for example, in neon at $\varepsilon \leq 20$ eV and in helium at $\varepsilon \leq 4$ eV. In this case, for the Maxwell EEDF, it is possible to derive analytical expressions for j'' and j' [9]. Substituting them into (3), one can easily see that the right-hand side of the inequality tends to $f(V)$ at $\psi \ll 1$ and $\psi \gg 1$. Maximum discrepancies from the actual EEDF take place at $V < T_e$, where T_e is the temperature of Maxwellian electrons. To quantitatively compare the EEDFs reconstructed with the help of the conventional differential techniques and with formula (3), we used the mean relative reconstruc-
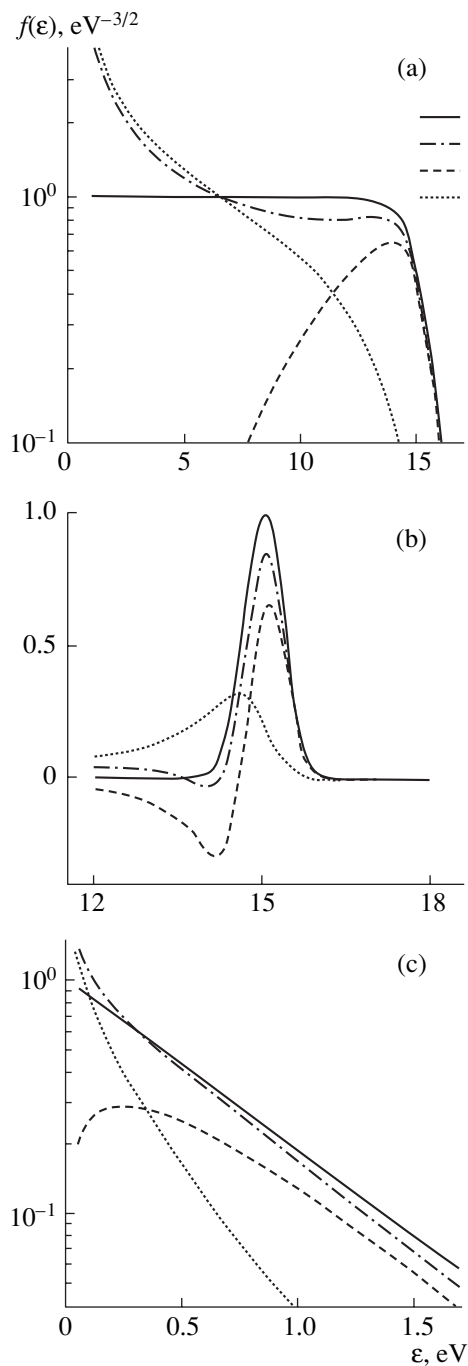
Fig. 1. Solutions to model problems of EEDF reconstruction. (a) Step EEDF ($\psi = 2$), (b) EEDF with maximum ($\psi = 10$); and (c) Maxwell EEDF for a plasma of molecular nitrogen ($T_e = 0.6$ eV, $\psi(T_e) = 0.5$). (*1*) f_{mod}, (*2*) new technique, (*3*) (j''), and (*4*) ($-\frac{\psi}{V}j'$).

tion error

$$\xi(\psi) = \frac{1}{\varepsilon_2 - \varepsilon_1}\int_{\varepsilon_1}^{\varepsilon_2} \frac{|f_{mod}(\varepsilon) - f_i(\varepsilon)|}{f_{mod}(\varepsilon)}d\varepsilon, \qquad (4)$$

where f_{mod} is a model EEDF and f_i is the EEDF reconstructed with different techniques.

The reconstruction errors for the Maxwell EEDF in the range from $0.1T_e$ to $5T_e$ are listed in the table. The function f_1 was reconstructed in accordance with the diffusion theory (the second term in (3)); f_2, from the second derivative; and f_3, using formula (3). It is seen that the technique proposed reconstructs the EEDF with smaller errors in a wide energy range and at any value of ψ. It is especially effective in the intermediate situation between the Langmuir and diffusion limits at $\psi \approx 1$.

Now, consider nonequlibrium distributions at $\psi = $ const. It was shown [3] that a step EEDF is distorted most strongly by the conventional techniques in the range $0.5 < \psi < 5$; an EEDF with a maximum, at $2 < \psi < 20$.

Figure 1a demonstrates a solution for a model step EEDF ($\psi = 2$). The second derivative is seen to underestimate the EEDF at low energies and becomes negative at $\varepsilon < 3$ eV. Agreement is observed only in the range of EEDF fast decrease. The first derivative, conversely, overestimates the function at low energies and underestimates at high energies. The combination of the derivatives by formula (3) reconstructs the EEDF with much lesser distortions. Similar results are observed in Fig. 1b for the EEDF with a maximum ($\psi = 10$).

The simulation carried out in the wide range of ψ shows that our technique decreases the error in EEDF reconstruction by several times in comparison with the conventional approaches. For the Langmuir probe ($\psi \ll 1$), the correction due to the first derivative diminishes appreciably the reconstruction error at low energies. For the diffusion probe ($\psi \gg 1$), the correction due to the second derivative decreases the reconstruction error at both low and high energies.

The simulation also shows that our technique is effective in the case when ψ depends on energy. We will illustrate this by reconstructing the Maxwell EEDF in a plasma of molecular nitrogen (Fig. 1c), where the transport collision cross section has a resonance character with a maximum at $\varepsilon \approx 2$ eV [10]. We simulated the situation for the pressure $P = 0.3$ torr and $T_e = 0.6$ eV. In this case, $\psi(T_e) = 0.5$ [11]. As follows from Fig. 1c, in the range where electron–electron collisions prevail ($\varepsilon < 1.7$ eV), the second derivative greatly distorts the EEDF. The characteristic scale of decrease in the interval from 0.6 to 1.7 eV, as determined from the second derivative, is 1.5 times greater than for f_{mod}. At higher electron temperatures, measurements based on the second derivative make no sense. However, the new approach allows one to reconstruct the EEDF even under these conditions (Fig. 1c). The systematic error of T_e determination in the range from 0.6 to 1.7 eV does not exceed 7%.

APPLICATION OF THE NEW APPROACH IN EXPERIMENT

Consider the reconstruction technique using a combination of the first and second derivatives for reconstructing the nonequilibrium part of the EEDF in neon afterglow and the Maxwellian part of the EEDF in the thermal energy range in nitrogen afterglow.

Recently [12, 13], the electron temperature relaxation in a diffusion plasma of neon afterglow has been measured and associated theoretical calculations have been carried out. In the experiment, T_e differed from the atom temperature T_a under quasi-steady-state conditions of EEDF formation. Under these conditions, T_e does not depend on the initial discharge temperature and thermal electrons are heated by fast electrons generated in the chemical ionization reaction involving two metastable neon atoms. When comparing the calculated and experimental values of T_e, Brauer *et al.* [12] hypothesized that the chemical ionization reaction generates low-energy electrons, which is contrary to fact [14]. In [13], the heating of Maxwellian electrons was considered in terms of the nonlocal formation of the EEDF fast part. In particular, it was noted that the wall potential should be taken into account for the correct description of the interaction between Maxwellian and nonequilibrium fast electrons.

It is noteworthy that in [12, 13] the fast part of the EEDF, $f_{ef}(\varepsilon)$, was not measured and its formation remained unclear. However, the evaluation of the second derivative of the probe current by numerical differentiation [13] was made in the energy range of the nonequilibrium EEDF. Under these conditions, $\psi = 2$ and the second derivative reconstructs the EEDF with large systematic distortions. Using our approach, we reconstructed the EEDF based on data from [13]. The result is shown in Fig. 2. The measurements were performed at a delay $\tau = 100$ μs after the current pulse had been switched off. The total electron concentration was $n_e = 2 \times 10^{10}$ cm^{-3}, and the electron concentrations on the most populated excited states were $Ne(^3P_2) = 4 \times 10^{11}$ cm^{-3} and $Ne(^3P_1) = 10^{11}$ cm^{-3}. Figure 2 also shows the theoretical results for $f_{ef}(\varepsilon)$ in different approximations described in detail in [13]. Here, we give only qualitative explanations.

To describe $f_{ef}(\varepsilon)$, it is necessary to know the spectrum of fast electrons generated and an energy relaxation mechanism. The spectrum of electrons generated in reactions with the participation of metastable states in inert gases was studied by the method of plasma electron spectroscopy [14]. It was shown that the characteristic width of the energy spectrum of arising electrons is $D_\varepsilon \approx 1$ eV, which is much less than the energy giving rise to fast electrons. Therefore, in calculations, a source of electrons can be represented in the form of the delta function at an energy $\varepsilon_0 = 11.6$ eV. Under the conditions considered, $f_{ef}(\varepsilon)$ forms in the nonlocal regime [15]. However, under free diffusion conditions,
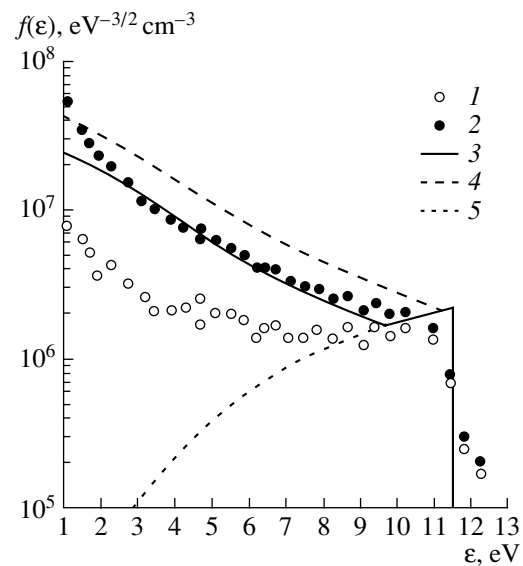

$f(\varepsilon)$, eV$^{-3/2}$ cm^{-3}

Fig. 2. EFDE in neon. (*1*) Experimental data (*j''*) and (*2*) EEDF reconstructed with the new technique. Calculation: (*3*) a near-wall discontinuity in the potential $\varphi_w = 9.5$ eV, (*4*) local calculation, and (*5*) EEDF under free diffusion conditions.

only fast electrons with an energy exceeding the wall potential can reach the walls. The rest of the electrons will be confined in the potential well, and their energy relaxation will occur in the volume via collisions. In its turn, the steady-state wall potential is set when the fast electron current toward the wall and that of positive ions equal each other (the latter current depends on the rate of ambipolar diffusion). If the number of fast electrons is sufficiently large, a discontinuity in the wall potential may arise in afterglow [16]. This means that a potential drop in the electron sheath greatly exceeds the value (3–5) T_e, which is typical of the Maxwell distribution. The wall potential under these conditions is shown to be $\varphi_w = 9.5$ eV at a temperature $T_e \approx 0.05$ eV of the majority of the electrons. The analytical result for $f_{ef}(\varepsilon)$ is shown by curve 3. It is in good agreement with the EEDF obtained with the new approach at $\varepsilon > 3$ eV. The discrepancy at low energies stems from the fact that the measurement systematically overestimates the EEDF at $\varepsilon < 3$ eV.

Figure 2 also depicts $f_{ef}(\varepsilon)$ calculated in the local approximation (curve *4*), which overestimates the EEDF. The assumption of free diffusion of fast electrons (curve *5*) strongly underestimates the EEDF at low energies. Figure 2 clearly demonstrates how different approaches to $f_{ef}(\varepsilon)$ calculation affect the heating of Maxwell electrons. In [17], it was shown that the effective energy transferred to them from fast electrons depends on the integral of the fast part of the EEDF. The difference in heating may reach two orders of magnitude. Only the calculation allowing for a discontinuity in the wall potential provides agreement with the tem-
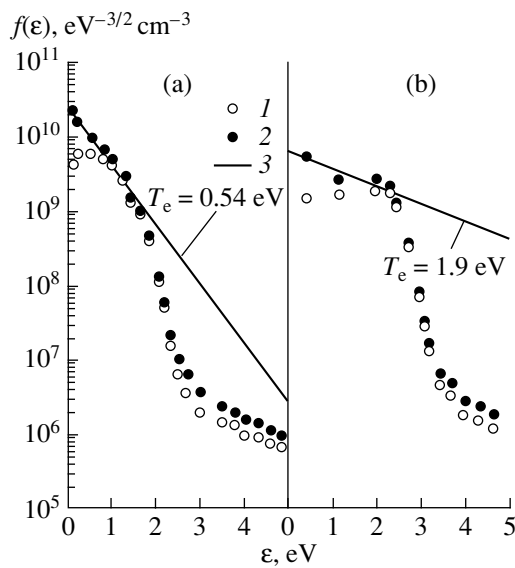
Fig. 3. EEDF in nitrogen. $E/N = 0.4$ (a) and 1.7 Td (b). (*1*) Experimental data (f''), (*2*) EEDF reconstructed with the new technique; and (*3*) approximation of the EEDFs reconstructed with the new technique at low energies.

perature T_e measured experimentally [13]. Data for the concentrations of electrons and excited atoms [12] imply that a potential jump in afterglow may exist. Ignoring this factor in the theoretical calculation probably led to the disagreement between the theory and the experiment.

Let us apply the new approach to the determination of the low-energy electron temperature in nitrogen afterglow with the electron distribution close to the Maxwellian one in the range $\varepsilon < 1.5$ eV, where electron–electron collisions prevail. In [11], large systematic distortions in the EEDF determined from the second derivative made it impossible to find the temperature of thermal Maxwellian electrons heated by an electric field pulse. In Fig. 3, we show the EEDF reconstructed with the new approach at $\varepsilon < 5$ eV. Obviously, the new method allows the determination of T_e. Namely, in a heating field $E/N = 0.4$ Td, the temperature T_e equals 0.54 ± 0.11 eV; i.e., it is twice as high as the temperature of vibrational distribution T_v, to which the scale of EEDF decrease in the energy range $1.5 < \varepsilon < 3$ eV corresponds. This contradicts the theoretical calculations in [11, 18], which gave $T_e \approx T_v$ for the same value of E/N. It should be noted that the calculations in [11, 18] were carried out assuming the Boltzmann distribution of molecules over vibrational levels of the ground state. At $T_e \approx T_v$, the rates of electron impact excitation and de-excitation of vibrational levels are equal to each other. Then, if the rate of vibrational–vibrational exchange (*VV* processes) for higher levels exceeds the rates of vibrational–translational relaxation (*VT* processes) and diffusion loss at the tube walls, the Treanor distribution over vibrational levels sets in. This

distribution has a denser population of higher vibrationally excited states than the Boltzmann distribution. For the first eight vibrationally excited levels, the value of T_e calculated with the technique described in [11] for the Treanor distribution ($T_v = 0.26$ eV, the gas temperature $T_g = 300$ K) was found to be $T_e = 0.64$ eV, which agrees with the experiment within the measurement error.

With the Treanor distribution over vibrational levels, the variation of the field strength up to $E/N = 1.7$ Td has a minor effect on the electron energy balance on the assumption that the EEDF is Maxwellian. This assumption is valid because the electron gas is heated by collisions with vibrationally excited molecules and not by the electric field. However, the experiment (see Fig. 3b) reveals a considerable transformation of the EEDF in the thermal energy range. In our opinion, this is due to the nonlinear properties of the electron–electron collision integral I_{ee}. Using the known approximations [20] $I_{ee} = A_0(\varepsilon/T_e)\nu_{ee}$, where $A_0(\varepsilon/T_e) = 0.385\varepsilon/T_e$ at $\varepsilon/T_e \leq 2.6$ and $A_0(\varepsilon/T_e) = 1$ at $\varepsilon/T_e \geq 2.6$, one can see that a growth of T_e reduces the effect of electron–electron collisions at low energies ($\varepsilon < 1.5$ eV). A characteristic plateau in the thermal energy range, observed in the experiment, is in qualitative agreement with the EEDF calculation neglecting electron–electron collisions and assuming the Treanor distribution of molecules over vibrational states [21].

CONCLUSION

We presented a technique for EEDF reconstruction using a combination of the first and second derivatives of the probe *I–V* characteristics. The technique does not require regularization algorithms to find the EEDF from the measured derivatives. This simplifies appreciably the interpretation of the measurements as compared with the method of EEDF reconstruction from the integral equations relating the EEDF to the derivatives. Our technique reduces the systematic error in EEDF measurements by several times in comparison with the traditional Langmuir approach and the diffusion probe method. Except for the range of low energies (as compared with the electron mean energy), the error does not exceed 20% for any ratio of the probe radius to the free path of electrons. The reconstruction of nonequlibrium EEDFs from probe measurements in heavy inert gases calls for special consideration.

REFERENCES

1. V. I. Demidov, N. B. Kolokolov, and A. A. Kudryavtsev, *Probe Methods for Low-Temperature Plasma Study* (Énergoatomizdat, Moscow, 1996).

2. Yu. B. Golubovskiĭ, V. M. Zakharova, V. N. Pasunkin, and L. D. Tsendin, Fiz. Plazmy **7**, 620 (1981) [Sov. J. Plasma Phys. **7**, 340 (1981)].

3. N. A. Gorbunov, N. B. Kolokolov, and A. A. Kudryavtsev, Fiz. Plazmy **15**, 1513 (1989) [Sov. J. Plasma Phys. **15**, 881 (1989)].

4. R. R. Arslanbekov, A. A. Kudryavtsev, and N. A. Khromov, Fiz. Plazmy **17**, 855 (1991) [Sov. J. Plasma Phys. **17**, 500 (1991)].

5. A. I. Lukovnikov and M. Z. Novgorodov, Kratk. Soobshch. Fiz., No. 1, 27 (1971).

6. N. L. Bashalov, Le Van Kh'eu, V. M. Milenin, *et al.*, Zh. Tekh. Fiz. **64** (10), 66 (1994) [Tech. Phys. **39**, 1012 (1994)].

7. J. D. Swift, Proc. Phys. Soc. London **79**, 697 (1962).

8. L. M. Friedland and Yu. M. Kagan, J. Phys. D **12**, 739 (1979).

9. A. M. Devyatov and M. A. Mal'kov, Izv. Vyssh. Uchebn. Zaved., Fiz., No. 3, 34 (1984).

10. L. G. H. Huxley and R. W. Crompton, *The Diffusion and Drift of Electrons in Gases* (Wiley, New York, 1974; Mir, Moscow, 1977).

11. N. A. Gorbunov, N. B. Kolokolov, and F. E. Latyshev, Fiz. Plazmy **27**, 1143 (2001) [Plasma Phys. Rep. **27**, 1079 (2001)].

12. T. Brauer, S. Gortchakov, D. Loffhagen, *et al.*, J. Phys. D **30**, 3223 (1997).

13. N. A. Gorbunov, N. B. Kolokolov, and F. E. Latyshev, Zh. Tekh. Fiz. **71** (4), 28 (2001) [Tech. Phys. **46**, 391 (2001)].

14. N. B. Kolokolov and A. B. Blagoev, Usp. Fiz. Nauk **163** (3), 55 (1993) [Phys. Usp. **36**, 152 (1993)].

15. L. D. Tsendin, Plasma Sources Sci. Technol. **4**, 200 (1995).

16. V. I. Demidov, N. B. Kolokolov, and O. G. Toronov, Fiz. Plazmy **12**, 702 (1986) [Sov. J. Plasma Phys. **12**, 402 (1986)].

17. A. B. Blagoev, Yu. M. Kagan, N. B. Kolokolov, and R. I. Lyagushchenko, Zh. Tekh. Fiz. **44**, 339 (1974) [Sov. Phys. Tech. Phys. **19**, 215 (1974)].

18. N. A. Dyatko, I. V. Kochetov, and A. P. Napartovich, Fiz. Plazmy **18**, 888 (1992) [Sov. J. Plasma Phys. **18**, 462 (1992)].

19. *Nonequilibrium Vibrational Kinetics*, Ed. by M. Capitalli (Springer-Verlag, Berlin, 1986; Mir, Moscow, 1989).

20. Yu. B. Golubovskiĭ, Yu. M. Kagan, and R. I. Lyagushchenko, Zh. Éksp. Teor. Fiz. **57**, 2222 (1969) [Sov. Phys. JETP **30**, 1204 (1970)].

21. F. Paniccia, C. Gorse, J. Bretagne, *et al.*, J. Appl. Phys. **59**, 4004 (1986).

Translated by N. Mende

Scaling Laws for Oxygen Discharge Plasmas

E. A. Bogdanov*, V. I. Kolobov*, A. A. Kudryavtsev**, and L. D. Tsendin**

* CFD Research Corporation, 215 Wynn Drive, Huntsville, AL, USA
** St. Petersburg State Technical University,
Politekhnicheskaya ul. 29, St. Petersburg, 195251 Russia
e-mail: akud@ak2138.spb.edu
Received February 8, 2002

Abstract—A fluid model is used to simulate ICP discharges in oxygen for a wide range of conditions under which commercial plasma-chemical reactors typically operate. Simple scaling laws are constructed with which different parameters of discharge plasmas in electronegative gases can be readily estimated from the given external parameters—the specific input power W and the product pL of the gas pressure and the characteristic plasma dimension. © 2002 MAIK "Nauka/Interperiodica".

Increased interest in studying discharges in electronegative gases is associated with their wide use in present-day plasma technology [1]. In this context, an important practical task is to construct scaling laws that would make it possible to readily estimate different plasma parameters and to predict how they will change depending on the external conditions. It has been established that, in the presence of negatively charged particles, diffusion processes, which determine the profiles of the plasma density and other plasma parameters, have some important special features [1, 2]. For this reason, the scaling laws known for electropositive plasma [3] do not apply directly to electronegative plasma and thus require significant modification.

In recent years, several papers have been written [4–12] that used a volume-averaged global model [4], in which the density profiles for different plasma components are assumed to be given (a flat profile for electrons and parabolic profiles for ions). The assumption of equilibrium Boltzmann densities not only for electrons but also for negative ions, which is an underlying assumption of the global model, was severely and outspokenly criticized in [13, 14]. In addition, in [15, 16], it was pointed out that taking into account exclusively ion–ion recombination while neglecting the loss of negative ions due to electron detachment limits the applicability of the global model [4–10] to only the range of low pressures ($pL < 0.1$ cm torr). In turn, the results that were obtained in some papers [4–12] in studying discharges in the same gas with the help of the global model differ significantly from each other. Thus, Gudmindsson *et al.* [12] arrived at the following result: in a chamber with the radius $R = 15.2$ cm and height $H = 7.6$ cm, ion–ion recombination is the main loss mechanism for negative ions only at pressures of $p \leq 5$ mtorr, whereas the loss of negative ions at higher pressures is dominated by electron detachment. The conclusion that electron detachment predominates over ion–ion recom-

bination at such low pressures raises doubts about the results obtained with the help of the global model in [4–10], in which only the loss of negative ions due to recombination was taken into account, while electron detachment was neglected. The version of the global model used in [4–12] also has a significant drawback in that it yields complicated formulas; moreover, in [4–12], these formulas are sometimes written with an unreasonably high accuracy (see below). In our opinion, the approach developed in [2] on the basis of the results of the earlier paper [17] is more systematic and attractive. The essence of this approach consists in determining the spatial profiles of the densities of charged particles from the analysis of the corresponding density balance equations under different conditions, in which case the ionization rate K_1 is to be analyzed as an eigenvalue of the problem and the rate constants of the remaining processes involving electrons are assumed to be given. In [17], it was established that the distinctive feature of discharges in electronegative gases is the presence of an external electron–ion plasma region (shell), which is essentially free of negative ions. As it is usually thin, the shell plays a fundamental role in the discharge because it traps negative ions in the plasma volume. As a result, at the boundary between the plasma and the wall sheath, the Bohm velocity of the ions is equal to that in the conventional electropositive plasma:

$$V_B = \sqrt{T_e/M}. \tag{1}$$

In turn, if we take into account the effect of negative ions at the boundary between ion–ion and electron–ion regions in the plasma volume, we arrive at the modified Bohm velocity of the ions,

$$V_{Bn}/V_s = \sqrt{n_n T_n/(n_e T_n + n_n T_e)}, \tag{2}$$

which is substantially lower than the Bohm velocity V_B (1) even when the degree of electronegativity $\alpha = n_n/n_e$ of the plasma is about several units. Since there are essentially no negative ions at the boundary between the plasma and the wall sheath, the appeal to expression (2) seems to be unnecessary.

Recall that the scaling laws for the main parameters of an electropositive plasma can be readily obtained from the known discharge characteristics, namely, the pressure, the geometry of the discharge volume, and the sort of gas (see, e.g., [3]).

The main parameters of the corresponding problem are the product pL (where p is the gas pressure and L is the characteristic dimension of the plasma region) and the specific power W fed into a unit volume of an electron gas ($W = P/V$). Recall that, in diffusion problems, introducing the effective plasma dimension $L = V/S$ allows the discharges to be treated approximately in plane-parallel geometry ($x = 0, L$). Thus, for a cylinder of radius R and height H, we have $L = V/S = RH/2(R + H)$ [1].

It is well known (see, e.g., [3]) that the scaling laws for discharges in electropositive gases follow from the balance relations for charged particles and electron energy. From the condition for the discharge to be steady, it follows that the ionization rate and the diffusive loss rate at the walls are the same:

$$K_1(T_e)N\tau_s(T_e) = 1, \qquad (3)$$

where K_1 is the ionization rate constant and N is the gas density.

The time τ_s of electron and ion losses to the wall can be estimated from the interpolation formula

$$\tau_s = \tau_a + \tau_B, \qquad (4)$$

where $\tau_a = L^2/\pi^2 D_a$ is the characteristic ambipolar diffusion time, during which the electrons escape to the wall in the fluid model in the limiting case $\lambda_p T_e/LT_a > 1$. In the opposite limiting case $\lambda_p T_e/LT_a < 1$, expression (4) for τ_s passes over to

$$\tau_B = L/V_B, \qquad (5)$$

where V_b is Bohm velocity (1).

In the case of a Maxwellian electron distribution function (EDF) in an arbitrary gas, the quantities K_1 and τ_s are functions of the electron temperature T_e, so that condition (3) makes it possible to determine T_e as a function of pL. In the general case, both of these quantities are determined by the heating electric field, namely, the longitudinal electric field in the positive column and the RF electric field in RF discharges. In what follows, we will restrict ourselves to the simplest case of a Maxwellian EDF.

If the specific input power W and the electron temperature T_e are known, then we can determine the elec-

tron density n_e from the balance relation for the electron gas energy,

$$W = n_e \sum_j \varepsilon_j \nu_j, \qquad (6)$$

where ε_j is the electron energy loss in the jth elementary process with the rate constant $\nu_j(T_e)$. The electron energy balance is dominated by all possible inelastic processes, in which case the sum on the right-hand side of the balance relation (6) is usually only weakly dependent on T_e. Consequently, as was mentioned in [16], the electron energy balance equation (6) is practically inapplicable for determining the electron temperature. It is more convenient to determine T_e from condition (3), because the dependence $K_1(T_e)$ of the ionization rate on the electron temperature is strong (usually exponential). As for balance relation (6), it serves to determine the electron density from the specific power fed into the discharge (or from the external current in the case of DC discharges). In [1, 4–10], the right-hand side of balance relation (6) was normalized to the ionization rate constant K_1 through the introduction of the energy cost $\varepsilon_c(T_e)$ of the production of one electron–ion pair in the form $W = n_e\varepsilon_c\nu_i$. However, keeping in mind the strong dependence of K_1 on T_e, we think that this way is not very promising because the quantity K_1 is not known *a priori*.

The goal of our study is to construct scaling laws that are as simple as possible in order to provide simple and rapid estimates of the parameters of electronegative plasmas. Since discharges in electronegative gases have an additional degree of freedom (the degree of electronegativity), to construct the desired scaling laws requires the specification of one more parameter in addition to W and pL. The role of this additional parameter is played by the ratio of the electron attachment time to the time of the diffusion of electronegative ions at electron temperature [2, 17]; however, this ratio reduces to the parameter pL. We restrict ourselves to the case that is of primary interest from a practical standpoint, specifically, the case of low pressures ($pL < 1$ cm torr) and/or low degrees of ionization ($n_e/N < 10^{-4}$), corresponding to classical condition (3) for the Schottky column, in which the diffusion mechanism for the losses of positive ions predominates over the recombination mechanism.

We consider a cylindrical ICP reactor with the dimensions $H = 7.62$ cm and $R = 15.42$ cm and assume that the working gas is oxygen. This is a typical test reactor for most studies with the global model [4–12]. Following [4–16], we assume that the EDF is Maxwellian. Unfortunately, this assumption is in no way justified for the hydrodynamic (collisional) regime of the electrons, in which the electron mean free path is $\lambda_e < L$. The EDF can be expected to be nearly Maxwellian only under the conditions of the Langmuir paradox, i.e., only in the free-fall (collisionless) regime of
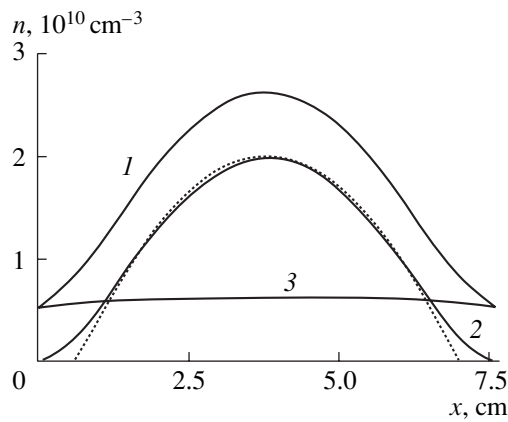
n, 10^{10} cm^{-3}

Fig. 1. Spatial profiles of the charged-particle densities (*1*) n_p, (*2*) n_n, and (*3*) n_e for $p = 20$ mtorr and $W = 15$ W. The dotted curve shows the results of calculations from expressions (17) and (21).
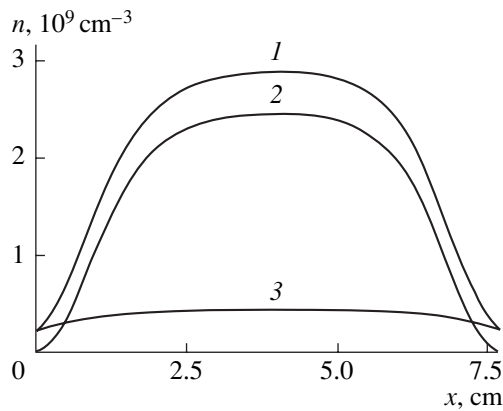
n, 10^9 cm^{-3}

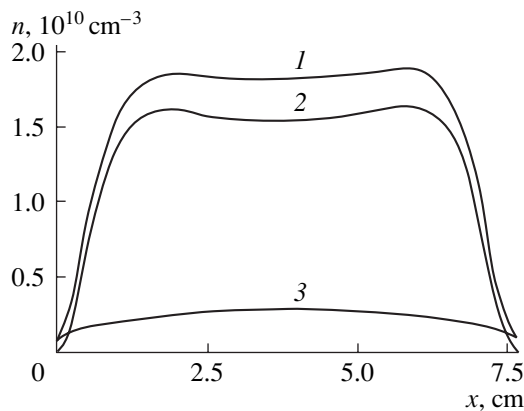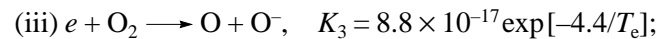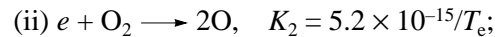**Fig. 2.** Same as in Fig. 1, but for $p = 200$ mtorr and $W = 25$ W.

n, 10^{10} cm^{-3}

Fig. 3. Same as in Fig. 1, but for $p = 600$ mtorr and $W = 15$ W.

the electrons, in which $\lambda_e > L$ (see [18] for details). The set of the ion balance equations, written in the drift–diffusion approximation, was solved numerically with the help of commercial software developed at the CFD Research Corporation (Huntsville, AL, USA) [19–21].
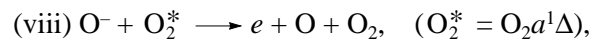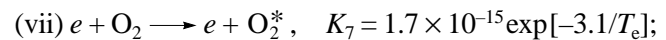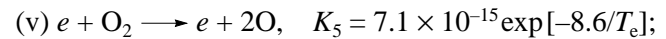
The ion mobility was calculated from the Langevin formula, and the corresponding diffusion coefficient was calculated from the Einstein relationship. As in most simulations [4–12], we assumed that the mobilities of the positive and negative ions are the same. Note that, at low pressures ($p < 200$ mtorr), the drift velocity of the ions is higher than their thermal velocity. As a result, the ion diffusion coefficients and ion drift rates obtained in our simulations are higher than the corresponding tabulated values calculated from their thermal velocities (see, e.g., [2, 22] for details). Thus, in our simulations, the ion diffusion coefficient was $D_p = 291/p$ cm^2/s (where the gas pressure is in mtorr). The processes and their rate constants incorporated into our model correspond to those used in [4–12]:

$$(i)\ e + O_2 \longrightarrow 2e + O_2^+,$$

$$K_1 = 9 \times 10^{-16} T_e^2 \exp[-12.6/T_e];$$

$$(ii)\ e + O_2 \longrightarrow 2O, \quad K_2 = 5.2 \times 10^{-15}/T_e;$$

$$(iii)\ e + O_2 \longrightarrow O + O^-, \quad K_3 = 8.8 \times 10^{-17} \exp[-4.4/T_e];$$

$$(iv)\ O^- + O_2^+ \longrightarrow O + O_2,$$

$$K_4 = 1.5 \times 10^{-13}(300/T_g)^{1/2};$$

$$(v)\ e + O_2 \longrightarrow e + 2O, \quad K_5 = 7.1 \times 10^{-15} \exp[-8.6/T_e];$$

$$(vi)\ O + O^- \longrightarrow O_2 + e, \quad K_6 = 2 \times 10^{-16}(300/T_g)^{1/2};$$

$$(vii)\ e + O_2 \longrightarrow e + O_2^*, \quad K_7 = 1.7 \times 10^{-15} \exp[-3.1/T_e];$$

$$(viii)\ O^- + O_2^* \longrightarrow e + O + O_2, \quad (O_2^* = O_2 a^1\Delta),$$
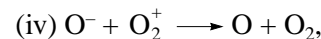
$$K_8 = 3 \times 10^{-16}(300/T_g)^{1/2}.$$

In the energy balance equation for the electron gas, we took into account electron energy transport by thermal conduction and thermal diffusion. The electron density was determined from the quasineutrality condition. The numerical scheme for solving the set of hydrodynamic equations was described in detail in [19], and examples of how this scheme is used to solve some plasma problems can be found in [20, 21].

We simulated discharges in oxygen for wide ranges of pressures, $p = 20$–1000 mtorr, and specific powers, $W = 10^3$–3.6×10^4 W/m^3. As in [10–12], the gas temperature was set equal to 600 K. The representative results and profiles that were obtained in our simulations are illustrated in Figs. 1–11.

In a plasma with a single ion species, the electron and ion density profiles coincide. In contrast, in a plasma with negative ions, the spatial density profiles for electrons and for positive and negative ions are different (see, e.g., [1, 2]). In simulations, we obtained three main types of density profiles for charged plasma particles (Figs. 1–3). At low pressures (Fig. 1), the ion density profiles are parabolic and the electron density profile is flat (it is the flat profile that is used in the global model [4, 5]). At higher gas pressures, the ion den-

sity profiles become flatter (Fig. 2), which corresponds to the flat-top profiles of the ion and electron densities [13, 14]. Along with these two well-known types of profiles, we obtained interesting profiles of the third type: the density of negative ions is minimum at the center and increases toward the periphery (Fig. 3). Earlier, analogous nonmonotonic profiles were observed experimentally in [23, 24].

As the pressure increases, the electron temperature T_e decreases for any W values, all other conditions being the same (Fig. 4). In turn, the dependence $T_e(W)$ is weak (Fig. 5). This weak dependence, which was pointed out in all of the papers [4–16], agrees qualitatively with the well-known profiles for electropositive gases.

The profiles $n_e(W)$ are nearly linear (Fig. 6), and the dependences $n_e(p)$ are far weaker (Fig. 7), which also agrees with the previously known results [4–16]. In Figs. 4 and 7, the triangles are for the experimental results obtained in [11] for discharges at low pressures in the reactor under consideration. One can see that there is good agreement between modeling and experiment.

Notably, as a function of pressure, the degree of electronegativity n_n/n_e has a maximum (Fig. 9). An increase in n_n/n_e at low pressures corresponds to the data obtained in [12], in which the degree of electronegativity n_n/n_e was found to increase monotonically with decreasing pressure over the entire pressure range ($p < 100$ mtorr) under investigation. Our simulations show that, at higher pressures, the degree of electronegativity n_n/n_e decreases. When the pressure is fixed, the degree of electronegativity n_n/n_e decreases with increasing specific input power (Fig. 8).

In order to explain the profiles obtained in our simulations and the scaling laws of interest to us, we turn to the formulation of the problem used in [4–16]. In plane-parallel geometry ($x = 0, L$), the equations of the drift–diffusion approximation and the Boltzmann electron distribution $E = -T_e \nabla n_e/n_e$ yield the following basic set of equations ($dn/dx \equiv n'$):

$$-D_p(n_p' + k n_p n_e'/n_e)' = \nu_i n_e - K_4 n_n n_p, \qquad (7)$$

$$-D_n(n_n' - k n_n n_e'/n_e)' = \nu_a n_e - \nu_d n_n - K_4 n_n n_p, \qquad (8)$$

$$n_p = n_n + n_e. \qquad (9)$$

Here, $\nu_a = K_3 N$ is the attachment rate in process (iii), $\nu_d = K_6 N_0 + K_8 N_0^*$ is the detachment rate in processes (vi) and (viii), and K_4 is the rate constant of ion–ion recombination (iv). In our approach, these three rate constants are assumed to be given. The ionization rate $\nu_i = K_1 N$ in process (i) is not known *a priori*: it is the eigenvalue of the boundary-value problem (7)–(9). Since, in the case of a Maxwellian EDF, all rate constants depend on T_e, the electron temperature can be found from the condition for the discharge to be steady.
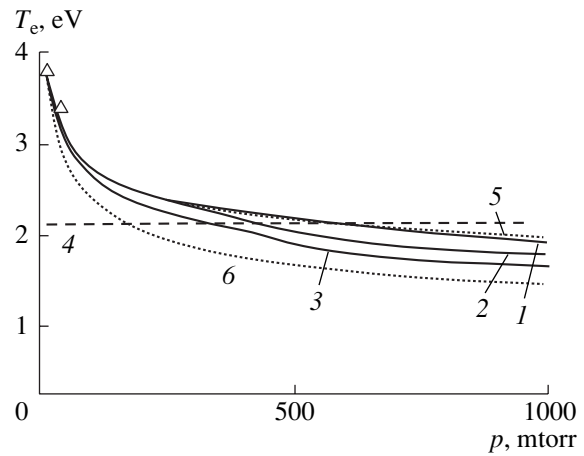
Fig. 4. Electron temperature T_e vs. pressure at different specific input powers: $W =$ (*1*) 430, (*2*) 860, and (*3*) 2153 W/m^3. Curves *4* and *5*, computation by (21), (22); curve *6*, computation by (17).
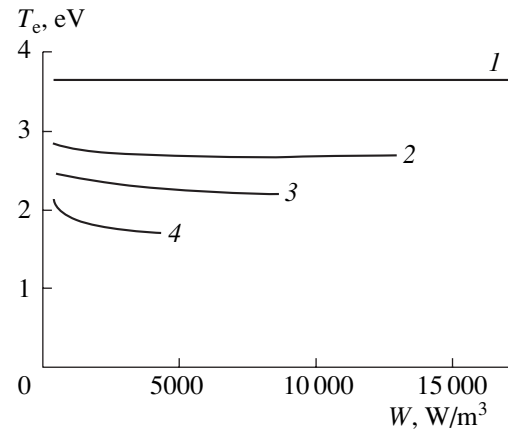

Fig. 5. Electron temperature T_e vs. specific power deposited in the discharge at different pressures: $p =$ (*1*) 20, (*2*) 80, (*3*) 200, and (*4*) 600 mtorr.
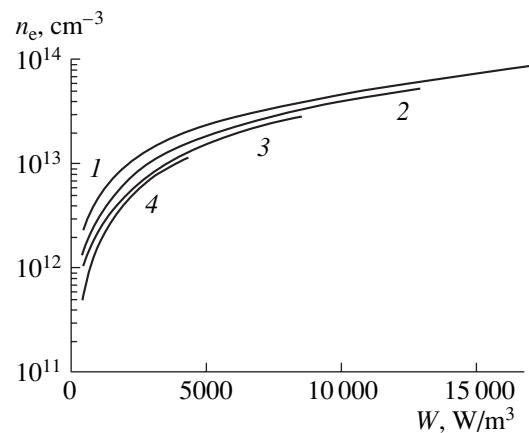

Fig. 6. Electron density n_e vs. specific power deposited in the discharge at different pressures: $p =$ (*1*) 20, (*2*) 80, (*3*) 200, and (*4*) 600 mtorr.
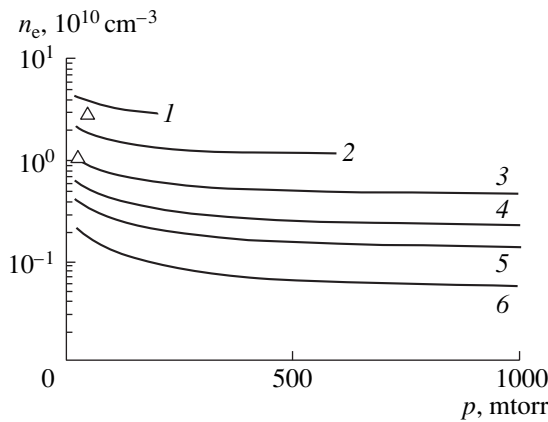
n_e, 10^{10} cm^{-3}

Fig. 7. Electron density n_e vs. pressure at different specific input powers: $W = $ (*1*) 8600, (*2*) 4300, (*3*) 1720, (*4*) 860, (*5*) 603, and (*6*) 430 W/m^3.
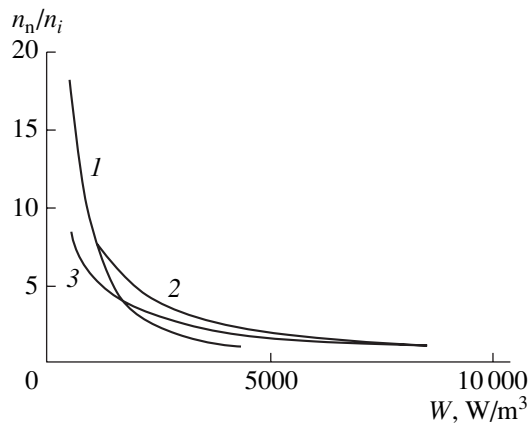
n_n/n_i

Fig. 8. Degree of electronegativity n_n/n_e vs. specific power fed into the discharge at different pressures: $p = $ (*1*) 20, (*2*) 80, (*3*) 200.
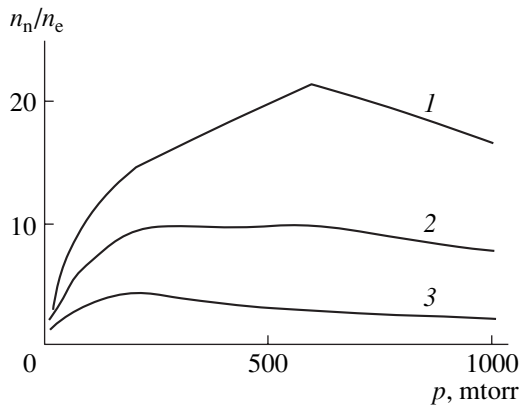
n_n/n_e

Fig. 9. Degree of electronegativity n_n/n_e vs. pressure at different specific input powers: $W = $ (*1*) 430, (*2*) 860, and (*3*) 2153 W/m^3.

The boundary conditions to Eqs. (7) and (8) have the form [1, 2]

$$n_n = n_p = dn_n/dx = 0 \quad \text{for} \quad x = \pm L,$$
$$dn_n/dx = dn_p/dx = 0 \quad \text{for} \quad x = 0. \tag{10}$$

Since there are no fluxes of negative ions to the walls, balance relation (6) yields the following relationship for the densities ($\bar{n}$), averaged over the cross section of the discharge [1, 2]:

$$\nu_a \bar{n}_e = \nu_d \bar{n}_n + K_4 \bar{n}_n \bar{n}_p. \tag{11}$$

The set of equations (7)–(9) was investigated in [1–16] under different simplifying assumptions. These studies established the formation of different spatial distributions of the charged-particle densities over the discharge cross section (Figs. 1–3). Essentially all negative ions occur in the inner ion–ion plasma region (the thickness of this region, which will be denoted by the index 0, is $x = x_0$). The outer plasma region (which will be denoted by the index 1) consists of electrons and positive ions and contains practically no negative ions.

Historically, in the late 1950s, Oskam [25] and Tompson [26] virtually simultaneously proposed two simple relationships between the charged-particle densities and their gradients in steady-state discharge plasmas in electronegative gases. Oskam [25] postulated that the electron density profile and the negative-ion density profile are similar:

$$\nabla n_e/n_e = \nabla n_n/n_n, \quad n_e(x)/n_n(x) = \text{const}, \tag{12}$$

Tompson [26] asserted that not only n_e but also n_n should obey a Boltzmann distribution:

$$T_e \nabla n_e/n_e = T \nabla n_n/n_n. \tag{13}$$

Condition (13) yields the relationship $n_e(x)/n_e(0) = n_n(x)/n_n(0)^{1/k}$, which is highly sensitive to the temperature ratio $k = T_e/T$ and coincides with profile (12) only in the particular case $k = 1$. Since in discharges the temperature ratio is $k \gg 1$, condition (13) implies that the electron density profile is nearly flat, $n_e(x) \approx n_{e0} \approx \text{const}$. Conditions (12) and (13) are meaningful only for the inner plasma region, which is denoted by 0. The criteria that allow these conditions to be satisfied in real discharges and are based on the approach developed in [2, 17] were discussed in detail in [27] and will be considered below.

The simplest way to determine the structure of the solution to Eqs. (7)–(9) is to turn to the equation introduced in [2, 17]. Specifically, summing up Eqs. (7) and (8) divided by the corresponding diffusion coefficients yields the equation

$$-2n_n''/k - n_e'' = n_e/l_e^2 - 2n_n/kl_n^2, \tag{14}$$

which is valid for $k \gg 1$ and has the following charac-

teristic spatial scales l_e and l_n:

$$1/l_e^2 = 1/l_{ion}^2 + 1/l_a^2 \tag{15}$$

$$= \nu_i/D_{ap} + \nu_a/D_{an} = \tau_{ap}\nu_i/L^2 + \tau_{an}\nu_a/L^2,$$

$$1/l_n^2 = 1/l_{nd}^2 + 1/l_{nr}^2 = \nu_d/2D_n + n_p K_r/D_{np} \tag{16}$$

$$= \tau_n \nu_d/L^2 + \tau_{np}K_r n_p/L^2,$$

where $D_{an,\,ap} = D_{n,\,p}(k + 1)$ and $D_{np} = 2D_n D_p/(D_n + D_p)$ are the electron–ion and ion–ion ambipolar diffusion coefficients and the corresponding characteristic time scales are denoted by $\tau = L^2/D_j$.

The scale l_e determines the maximum possible thickness of the outer region 1 [17]. In this region, the density of negative ions is low, $n_n(x) \approx 0$, because they are expelled by the ambipolar field to the inner plasma region, so that we have $n_e(x) \approx n_p(x)$.

The distribution of charged particle densities in the inner plasma region depends substantially on the relationship between the characteristic dimension L and the distance l_n over which negative ions diffuse during their lifetime before they disappear in internal processes [2, 17]. When $l_n > L$ (which is the case at low gas pressures), a negative ion can diffuse across the entire discharge volume. In order for negative ions to be trapped in the discharge, their density in the inner region should obey Boltzmann distribution (13) (as is the case with electrons). From Eqs. (7)–(9), it follows (see, e.g., [28] for details) that the diffusion of negative ions usually dominates ($l_n \gg L$) when the product of the rate of their ambipolar diffusion (ion diffusion at electron temperature) and the attachment rate is much smaller than unity, $\tau_{an}\nu_a \ll 1$. Since, in this approximation, Schottky condition (3) is satisfied, the last inequality passes over to the condition $K_1 \geq K_3$, which indicates that the ionization rate exceeds the attachment rate and, for a Maxwellian EDF, leads to the condition $T_e > T_{ec}$ for the electron temperature. The temperature on the right-hand side of this condition was introduced in [15] and is to be determined from the equality $K_1(T_{ec}) = K_3(T_{ec})$. For the rate constants of processes (i)–(viii), we have $T_{ec} = 2.2$ eV (Fig. 4, curve 4). In [15], it was rightly noted that T_{ec} is the lowest possible temperature in the global model [1, 4–10]; this indicates that the global model is inapplicable to such discharge conditions for which simulations based on it give $T_e < T_{ec}$.

At low pressures ($\tau_{an}\nu_a \ll 1$, which corresponds to electron temperatures $T_e > T_{ec}$), the electrons and negative ions both obey Boltzmann distribution (13). Recall that, under these conditions, the electron density profile in the inner region at $k \gg 1$ is flat, $n_e(x) \approx n_{e0}$. The field-driven flux of negative ions and their diffusive flux are of roughly equal intensity and are oppositely directed. Consequently, the two terms on the right-hand side of Eq. (7) for $n_p(x)$ are approximately equal and should be
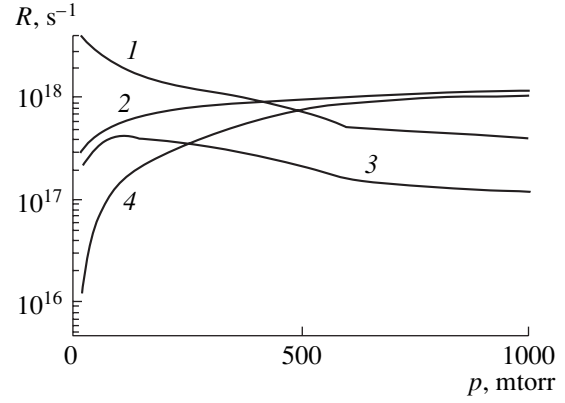
Fig. 10. Rate constants of reactions in the balance of n_e: (1) R_1, (2) R_3, (3) R$_4$, and (4) $R_6 + R_7$.
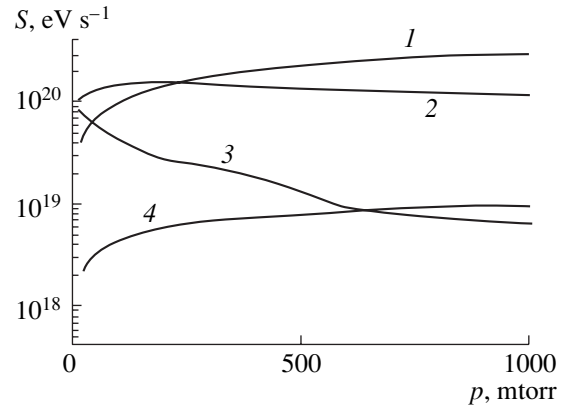

Fig. 11. Rate constants of reactions in the balance of T_e: (1) $\varepsilon_7 R_7$, (2) $\varepsilon_5 R_5$, (3) $\varepsilon_1 R_1$, and (4) $\varepsilon_3 R_3$.

summed, so that, in the case $n_n(0) > n_e(0)$, in which we are interested here, Eq. (7) becomes $-2D_p n_n'' = \nu_i n_{e0}$. This gives a parabolic profile of the ion densities [17] (for details, see also [1, 4–7]):

$$n_n(x) = n_{n0}(1 - x^2/L^2),$$

$$n_{n0}/n_{e0} = \nu_i \tau_{sn}, \quad \tau_{sn} = \tau_p + \tau_{bn}. \tag{17}$$

By analogy with expression (4), profile (17) was derived using the interpolation formula for the characteristic loss times for charged particles. The only difference is that the expression for the diffusion time τ_p contains the ion diffusion coefficient $2D_p$, while the expression for τ_{bn} contains the Bohm velocity of the ions (including the negative ions); i.e., we have $\tau_p = L^2/4D_p$ and $\tau_{bn} = L/V_{bn}$ [cf. expression (4)]. Since the sell (region 1) is thin and the sell thickness is highly sensitive to changes in the parameters (see [28] for details), we simplified the approximate analytic approach to the problem by setting $x_0 \approx L$ in expressions (17). Note that, in order to achieve profiles (17) in practice, it is sufficient to satisfy the condition $\tau_{an}\nu_a \ll 1$ (or $T_e > T_{ec}$). In

this case, the internal losses of negative ions [see the right-hand side of relationship (11)] may occur via either the recombination ($\nu_d < K_4 n_p$) or the attachment ($\nu_d > K_4 n_0$) mechanism.

In the opposite limiting case $\tau_{an}\nu_a \gg 1$ (i.e., at high gas pressures), we have $l_n < l_e < L$ for the typical values $\bar{n}_n/\bar{n}_e < k \approx 100$. In this case, in the basic set of equations (7)–(9), we can neglect the terms describing ion diffusion (as was done in [13–16]) and approximately set $x_0 \approx l_e$. Since the ion fluxes are dominated primarily by the electric field, the fluxes of positive and negative ions with nearly equal densities ($n \approx p$) in the inner region $l_e < x < L - l_e$ have the same intensity and are oppositely directed; i.e., we have [2, 17]

$$\Gamma_n/b_n \approx -\Gamma_p/b_p. \tag{18}$$

Because of the equality of the flux divergences (18), the left-hand side of Eq. (14) in the central region $l_e < x < L - l_e$ is small, reflecting the local balance between the plasma-chemical processes of production and loss of the ions, $n_e/l_e^2 = 2n_n/k l_n^2$, which corresponds to the condition [2, 17]

$$(\nu_i/D_p + \nu_a/D_n)n_e = \nu_d n_n/D_n \\ + K_4 n_n(n_n + n_e)/(1/D_p + 1/D_n). \tag{19}$$

In contrast to the case of low gas pressures, the charged-particle density profiles at high pressures may be different depending on the mechanism that dominates the loss of negative ions [i.e., on the relationship between the two terms on the right-hand side of condition (19)]. This leads to different relationships between the ionization rate and the rates of other process (and, accordingly, to different electron temperatures T_e). When the recombination mechanism is dominant ($\nu_d < K_4 n_p$), condition (19) yields the relationship $\nabla n_e/n_e = \nabla n_n/n_n + \nabla n_p/n_p \approx 2\nabla n_n/n_n$ [2, 17], in which case, in contrast to profiles satisfying condition (12), the ion profiles are flatter than the electron profile (see [28] for details). From Eq. (7), relationships (11) and (18), and condition (19), we obtain the estimate

$$K_1 = K_3 + (1 + D_p/D_n)K_3 N_{e1}/N_{e0} \\ \approx K_3 + (1 + D_p/D_n)(L - x_0)/x_0 \approx K_3, \tag{20}$$

i.e., $K_1 \approx K_3$. Here, $N_{e1} \approx n_{e0}(L - x_0)$ and $N_{e0} \approx n_{e0}x_0$ are the averaged (over the discharge cross section) electron densities in regions 1 and 0, respectively. Hence, in the recombination regime, the electron temperature T_e is close to the constant value T_{ec} (as was pointed out in [15]).

In the detachment regime ($\nu_d > K_4 n_p$), for flat spatial profiles that were adopted in [2, 13–18] for oxygen atoms O and metastable oxygen molecules O_2^* participating [detachment processes (vi) and (viii), respectively], condition (19) immediately yields condition

(12), which corresponds to flattened density profiles for electrons and ions. For the electron density $n_e(x)$ in the inner region, we obtain from Eq. (7), relationships (11) and (18), and condition (19) the following profile, which decreases slightly toward the periphery (see [2, 17] for details):

$$n_e(x) = n_{e0}\sin(\sqrt{\nu_i\tau_a\nu_d/\nu_a}\,x). \tag{21}$$

Since the loss of O atoms and O_2^* metastables is dominated by diffusion, the profiles of their densities are not flat but decrease toward the walls; in plane geometry, they are proportional to $\sim\sin(\pi x/L)$. According to condition (19), the resulting decrease in the rate of electron detachment from the negative ions toward the periphery in reactions (vi) and (viii) leads to a local increase in their densities.

From relationship (18), we obtain the following relationship for the ionization rate in the detachment regime [2, 17]:

$$K_1 = (D_p/D_n)K_3 N_{e1}/N_{e0} \approx (D_p/D_n)K_3(L - x_0)/x_0; \tag{22}$$

i.e., the number of electrons attached in the outer region is equal to the number of atoms ionized in the inner region. Relationship (22) is inconvenient for estimates, because it requires knowledge of and is sensitive to the shell thickness. If we use electron density profile (21) in relationship (22) (see [2, 17] for details) and turn to relationships (11) and (18)–(21), then we can obtain the following more convenient expressions for the ionization rate:

$$\nu_i\tau_{an} = \nu_a/\nu_d \quad \text{for} \quad \sqrt{\tau_{an}\nu_a} > \nu_a/\nu_d, \tag{23}$$

$$\nu_i\tau_{an} = \sqrt{\nu_a\tau_{an}} \quad \text{for} \quad \sqrt{\tau_{an}\nu_a} < \nu_a/\nu_d. \tag{24}$$

Expression (24) was also obtained earlier in [28, 29].

According to expressions (23) and (24), we have $\nu_i\tau_{an} > 1$ in the detachment regime at hand; this indicates that the ionization rate is higher than that determined from Schottky condition (3) for a conventional plasma. Accordingly, the electron temperature is also higher than the temperature obtained from condition (3).

The results of our analysis show that the parameters of the plasma in an electronegative gas depend substantially on a new parameter—the quantity $\tau_{an}\nu_a$, which, however, cannot serve as a new scaling parameter, because it is determined by the product pL.

Hence, considering regimes dominated by one or another process makes it possible to obtain simple theoretical relationships for the calculation of the plasma parameters of discharges in electronegative gases. This approach is illustrated by Fig. 10, which shows the rates of different processes incorporated into the balance equations for the charged-particle densities in oxygen discharge plasmas in the ICP reactor under investigation. From Fig. 10, we can see that, at low pressures ($p < 200$ mtorr), the most intense process is ionization,

i.e., reaction (i). This corresponds to the case $\tau_{an}\nu_a \ll 1$, in which the condition $T_e > T_{ec}$ holds and the Boltzmann equilibrium condition for negative ions is satisfied. In this case, the ion density profiles are parabolic, and the electron density profile is flat. Parabolic profile (17), which is plotted by the dotted curve in Fig. 1, is seen to agree well with the calculated results. At low pressures, the negative ions disappear mainly in recombination process (iv).

At higher pressures ($p > 200$ mtorr), the condition $\tau_{an}\nu_a \gg 1$ is satisfied and the ionization rate is lower than the attachment rate (Fig. 10). This situation is governed to a large extent by the mechanism for the internal losses of negative ions. For $\nu_d < K_4 n_p$ (when negative ions disappear due to the recombination mechanism), the condition $K_1 = K_3$ should hold (this indicates that the ionization rate is equal to the attachment rate and that $T_e \approx T_{ec}$). In oxygen, this is not the case, because, at high pressures, the inequality $\tau_{an}\nu_d > 1$ implies a transition to the detachment regime of the losses of negative ions ($\nu_d > K_4 n_p$). In this case, the ion and electron densities obey a kind of flat profiles: the ion densities either slightly decrease or slightly increase toward the periphery (Figs. 2, 3). The dotted curves in Fig. 4 show the electron density profiles $n_e(x)$ calculated from expression (21) with allowance for condition (19) in which the negative-ion density profile $n_n(x)$ is given by expression (21). We can see that the dotted curves agree well with the numerical results.

The electron temperature T_e is determined by the relationship between the ionization rate and the rates of other processes that govern the steady plasma state. At low pressures (such that $\tau_{an}\nu_a < 1$), the T_e values agree well with those obtained from expressions (17) (Fig. 4, curve 6). At higher pressures ($\tau_{an}\tau_a > 1$), the numerical values of the electron temperature agree well with those obtained from expressions (21) and (22) (Fig. 4, curve 5). Since expressions (17), (21), and (22) do not contain the specific input power W, the dependence $T_e(W)$ is weak (Fig. 5).

By analogy with Fig. 10, Fig. 11 illustrates the relationships that determine the energy balance of an electron gas. We can see that the rates of the governing processes, specifically, dissociation process (v) and process (vii) of the excitation of metastable states, are substantially higher than the ionization rate. Recall that the electron energy balance equation is practically inapplicable for determining the electron temperature. However, this balance equation yields simple formulas for estimating the electron density n_e from the specific power fed into the discharge. From energy balance equation (6) and from Fig. 11, we obtain the relationship $W \approx (\varepsilon_5 K_5 + \varepsilon_7 K_7)n_e$; for the weak dependence $T_e(W)$, this relationship yields the direct proportionality $W \sim n_e$, indicating that the electron density $n_e(W)$ as a function of specific input power increases essentially linearly (Fig. 7).

The degree of electronegativity n_n/n_e can be estimated from global balance relationship (11) for negative ions. At low pressures ($\tau_{an}/\nu_a < 1$), when the negative ion density n_n decreases because of the ion–ion recombination, the degree of electronegativity n_n/n_e increases with pressure. At higher pressures ($\tau_{an}\nu_a > 1$), i.e., in the detachment regime, the degree of electronegativity n_n/n_e decreases with pressure (Fig. 8).

Hence, we have simulated ICP discharges in oxygen with the help of commercial software developed at the CFD Research Corporation. We have obtained simple scaling laws that make it possible to readily estimate different parameters of discharge plasmas in electronegative gases from the given external parameters pL and W.

ACKNOWLEDGMENTS

REFERENCES

1. M. Lieberman and A. Lichtenberg, *Principles of Plasma Discharges and Materials Processing* (Wiley, New York, 1994).

2. A. V. Rozhansky and L. D. Tsendin, *Transport Phenomena in Partially Ionized Plasma* (Taylor & Francis, London, 2001).

3. G. Francis, *Ionization Phenomena in Gases* (Butterworths, London, 1960).

4. C. Lee and M. A. Lieberman, J. Vac. Sci. Technol. A **13**, 368 (1995).

5. A. J. Lichtenberg, V. Vahedi, M. A. Lieberman, *et al.*, J. Appl. Phys. **75**, 2339 (1994).

6. M. A. Lieberman and S. Ashida, Plasma Sources Sci. Technol. **5**, 145 (1996).

7. A. J. Lichtenberg, M. A. Lieberman, I. G. Kouznetsov, *et al.*, Plasma Sources Sci. Technol. **9**, 45 (2000).

8. T. H. Chung, J. Korean Phys. Soc. **34**, 24 (1999).

9. T. H. Chung, H. T. Yoon, and D. C. Seo, J. Appl. Phys. **86**, 3536 (1999).

10. D. C. Seo and T. H. Chung, J. Phys. D **34**, 2854 (2001).

11. J. T. Gudmindsson, A. M. Marakhtanov, K. K. Patel, *et al.*, J. Phys. D **33**, 1323 (2000).

12. J. T. Gudmindsson, I. G. Kouznetsov, K. K. Patel, *et al.*, J. Phys. D **34**, 1100 (2001).

13. R. N. Franklin and J. Snell, J. Phys. D **32**, 2190 (1999).

14. R. N. Franklin, Plasma Sources Sci. Technol. **10**, 174 (2001).

15. R. N. Franklin, J. Phys. D **34**, 1243 (2001).

16. R. N. Franklin, J. Phys. D **34**, 1834 (2001).

17. L. D. Tsendin, Zh. Tekh. Fiz. **59** (1), 21 (1989) [Sov. Phys. Tech. Phys. **34**, 11 (1989)].

18. A. A. Kudryavtsev and L. D. Tsendin, Zh. Tekh. Fiz. **69** (11), 34 (1999) [Tech. Phys. **44**, 1290 (1999)].

19. *CFD Research Corporation: CFD-PLASMA User Manual* (CFD Research Corporation, Huntsville, 1999).

20. S. Yun, V. I. Kolobov, and G. R. Tynan, Phys. Plasmas **8**, 3069 (2001).

21. http://www.cfdrc.com/.cfdplasma.

22. V. V. Ivanov, K. S. Klopovsky, D. V. Lopaev, *et al.*, IEEE Trans. Plasma Sci. **27**, 1279 (1999).

23. U. Buddemeier, PhD Thesis (Ruhr University, 1997).

24. S. V. Berezhnoy, C. B. Shin, U. Buddemeier, *et al.*, Appl. Phys. Lett. **77**, 800 (2000).

25. H. J. Oskam, Philips Res. Rep. **13**, 335 (1958).

26. J. B. Tompson, Proc. Phys. Soc. London **73**, 818 (1959).

27. P. R. Daniels and R. N. Franklin, J. Phys. D **22**, 780 (1989).

28. E. A. Bogdanov and A. A. Kudryavtsev, Pis'ma Zh. Tekh. Fiz. **27** (21), 36 (2001) [Tech. Phys. Lett. **27**, 905 (2001)].

29. C. M. Ferreira, G. Gousset, and M. Touzeau, J. Phys. D **21**, 1403 (1988).

Translated by G. Shepekina

Numerical Investigation of Subcritical Microwave Discharges in a High-Pressure Gas

O. I. Voskoboĭnikova, S. L. Ginzburg, V. F. D'yachenko, and K. V. Khodataev

*Keldysh Institute of Applied Mathematics, Russian Academy of Sciences,
Miusskaya pl. 4, Moscow, 125047 Russia*

e-mail: ginz@kiam.ru

Received June 19, 2001; in final form, January 21, 2002

Abstract—Subcritical microwave streamer discharges are investigated using a two-dimensional model that describes gas-dynamic processes in the ideal gas approximation and a self-consistent electromagnetic field in the wave approximation and takes into account the minimum required number of kinetic processes (such as ionization, attachment, recombination, diffusion, and electric conduction). The initial conditions imitate the initiation of a discharge from a small cavity with a reduced gas density and an arbitrarily small degree of gas ionization. The possibility of describing streamer discharges without reference to ionizing hard radiation is confirmed. © 2002 MAIK "Nauka/Interperiodica".

INTRODUCTION

In traditional physics, the phenomenon of breakdown followed by the onset and development of a gas discharge has been studied for a long time because it has a broad range of applications. Being driven by an electromagnetic field of an arbitrary frequency in different ranges (from the frequency range corresponding to constant currents to the laser frequency range), the gas discharge develops in very different ways. The development of the discharge initiated by an electromagnetic field with a wavelength ranging from millimeters to decimeters is most peculiar. Progress in the technique for generating high-power microwave radiation made it possible to achieve fairly high microwave energy densities in the focal region. Spark gas discharges were initiated by breakdown in the focus of a centimeter-wavelength radiation beam at moderate pressures (up to a pressure of 0.3 atm) [1, 2] and at high pressures (up to several atmospheres) in an open microwave cavity [3].

A discharge in the beam focus occurs in the form of a complicated system of very thin plasma filaments, which occupy the entire focal region and propagate toward the source [1, 4]. Such discharges are rather difficult to investigate in detail because of their complicated spatial filamentary structure. In contrast, discharges in the focal region of a cavity occur as a regular solitary formation—a bright thin filament that is stretched out along the electric field vector and whose length is close to half of the radiation wavelength [5]. Such a simple geometric structure offers expanded opportunities for studying the physics of the phenomenon in question.

Since microwave discharges can develop in regions in which the unperturbed electric field strength is below the threshold for breakdown, it was supposed that the streamer mechanism plays a leading role in the discharge process. In [6] and in some subsequent papers, attempts were made to construct an elementary model of a high-frequency streamer in an overcritical field in the quasisteady (rather than wave) approximation. Since under conditions of an overcritical field the streamer develops at a significantly faster rate than the gas-dynamic processes, the latter can be neglected in describing the initial stage of the streamer evolution. However, in a subcritical field, a time scale for the discharge evolution is comparable to that of the gas-dynamic processes, which ensure the conditions required for the development of the streamer—namely, the reduction of the gas density in the streamer channel [7]. For the first time, this situation was described theoretically by Khodataev [8], who constructed a relatively simple numerical model of a developing streamer with a length comparable to the radiation wavelength and took into account the reduction in the gas density during gas heating. He showed that, under the electrodynamic resonance conditions (when the streamer length becomes equal to half of the radiation wavelength), the field energy is resonantly absorbed by the streamer plasma. He also demonstrated that, in contrast to the common practice [1], the propagation of the streamer can be described without reference to the ionizing ultraviolet radiation from the plasma in the streamer channel. In this way, it is sufficient to take into account only electron diffusion, which is free in the range of low electron densities. In the region in front of the streamer head, where the field amplitude exceeds both its unperturbed value and the threshold for breakdown, the electron density decreases abruptly away from the head and becomes so low that the ambipolarity effect turns out to be insignificant. In this case, the

growth rate of the streamer is determined by the sum of the ionization rate in the electric field and the attachment rate and by the coefficient of free electron diffusion [9]. The results of our investigations confirm this conclusion.

Our efforts to construct a more adequate theoretical model of a streamer ran into significant difficulties. Of course, the model should be capable of describing electron-balance processes such as electron impact ionization, recombination, attachment, and electron diffusion. The model to be constructed is intended to describe spark streamer discharges observed in a microwave field at gas densities higher than about 0.1 of normal density. Under such conditions, the electron transport collision frequency ν_{tr} exceeds the angular frequency ω of the field, and the electrical conductivity σ of a partially ionized gas is dominated by electron collisions and is thus determined by the transport collision frequency ν_{tr} and electron density n_e. The current density and the electric field are related by Ohm's law:

$$\mathbf{j} = \sigma\mathbf{E}. \tag{1}$$

The microwave-driven electric current rapidly heats the streamer plasma and thus gives rise to intense gas-dynamic processes. In this case, the ionization and dissociation coefficients are assumed to be small, which makes it possible to use the model of an ideal gas with the adiabatic index γ.

The development of an electrically conducting plasma formation greatly perturbs the initial electromagnetic field distribution, so the streamer effect comes into play. Since the length of a developed streamer is comparable to the radiation wavelength, the self-consistent electrodynamic problem is to be solved in the wave approximation. In actual experiments, the initial microwave field is generated by a monochromatic source and the characteristic time scale of the process is much longer than the field period; consequently, it is natural to formulate the problem in terms of the Helmholtz equation with the Sommerfeld condition at a remote boundary.

MATHEMATICAL MODEL

The development of a streamer discharge in a microwave field is described by the set of equations consisting of the gas-dynamic equations for an ideal gas,

$$\frac{\partial n}{\partial t} + \frac{\partial n V_\alpha}{\partial x_\alpha} = 0, \tag{2}$$

$$\frac{\partial n V_\beta}{\partial t} + \nabla p/M + \frac{\partial n V_\beta V_\alpha}{\partial x_\alpha} = 0, \quad p = (\gamma - 1)Mn\varepsilon, \tag{3}$$

$$\frac{\partial nw}{\partial t} + \frac{\partial}{\partial x_\alpha} V_\alpha(p/M + nw) = q\sigma|E|^2, \quad w = \varepsilon + \frac{V^2}{2}, \tag{4}$$

the ionization balance equation for the electron density n_e,

$$\frac{\partial n_e}{\partial t} + \frac{\partial}{\partial x_\alpha}\left(n_e V_\alpha - D_{ef}(n, |E|, n_e)\frac{\partial n_e}{\partial x_\alpha}\right) \tag{5}$$

$$= n_e[\nu_i(n, |E|) - \nu_a(n, |E|) - \beta_r n_e]$$

(where α, $\beta = 1, 2, 3$); and the Helmholtz equation for the microwave field amplitude $\mathbf{E}$,

$$\nabla \times \nabla \times \mathbf{E} - k^2\left(1 + i\frac{4\pi\sigma}{\omega}\right)\mathbf{E} = 0, \tag{6}$$

$$\sigma = \frac{n_e e^2}{m\nu_{tr}}, \quad k = \frac{\omega}{c}, \quad i = \sqrt{-1}.$$

Here, n is the gas density, V is the velocity, p is the pressure, ε is the internal energy per unit mass of the gas, M is the mean mass of a molecule, γ is the adiabatic index, ν_i and ν_a are the ionization and attachment rates, β_r is the recombination coefficient, n_e is the electron density, σ is the electrical conductivity, E is the complex amplitude of the microwave field, D_{ef} is the coefficient of free electron diffusion (in a plasma region with a high electron density, it passes over to the ambipolar diffusion coefficient [9]), and m and e are the mass and charge of an electron.

As the measurement units for the gas density and gas pressure, we adopt their values $[n]$ and $[p]$ in an unperturbed gas. We also express the velocity in the units of $[V] = \{[p]/M[n]\}^{1/2}$. Since the length of a fully developed streamer is comparable to the radiation wavelength, it is convenient to choose $[x] = 1/k$ and $[\sigma] = \omega/4\pi$ as the units of length and conductivity, in which case the units of time and electron density are $[t] = [x]/[V]$ and $[n_e] = [\sigma m\nu_{tr}/e^2]$. It is also expedient to normalize the electric field to its critical value $[E] = E_{cr}$ in an unperturbed state; in turn, the critical value of the field is determined from the condition $\nu_i([n], [E]) - \nu_\alpha([n], [E]) = 0$.

In the dimensionless variables, Eqs. (2)–(6) take the form

$$\frac{\partial n}{\partial t} + \frac{\partial n\nu_\alpha}{\partial x_\alpha} = 0, \tag{7}$$

$$\frac{\partial n\nu_\beta}{\partial t} + \nabla p + \frac{\partial n\nu_\beta\nu_\alpha}{\partial x_\alpha} = 0, \quad p = (\gamma - 1)n\varepsilon, \tag{8}$$

$$\frac{\partial nw}{\partial t} + \frac{\partial}{\partial x_\alpha}\nu_\alpha(p + nw) = q\sigma|E|^2, \quad w = \varepsilon + \frac{\nu^2}{2}, \tag{9}$$

$$\frac{\partial n_e}{\partial t} + \frac{\partial}{\partial x_\alpha}\left(n_e\nu_\alpha - D(n, E, n_e)\frac{\partial n_e}{\partial x_\alpha}\right) \tag{10}$$

$$= n_e[nF(n, |E|) - fn_e],$$

$$\nabla \times \nabla \times \mathbf{E} - (1 + i\sigma)\mathbf{E} = 0, \quad \sigma = \frac{n_e}{n}, \tag{11}$$

where

$$q = [\sigma][E]^2[t]/[p],$$
$$f = \beta_r[n_e][t], \tag{12}$$

$$F(n, |E|) = F_0[\exp(b(1 - n/|E|)) - A(n, |E|)],$$

$$F_0 = \nu_a([n], [E])[t], \tag{13}$$

$$A(n, |E|) = \frac{\nu_a(n, |E|)}{\nu_a([n], [E])}, \tag{14}$$

$$D(n, E, n_e) = \frac{D_0}{n}\left(\frac{F^2(n, |E|)}{F^2(n, |E|) + \left(\dfrac{n_e}{n}\right)^2} + \sqrt{\frac{m}{M}}\right), \tag{15}$$

$$D_0 = D_{ef}([n], [E], 0)[t]k^2.$$

Dependence (12) approximates the experimental data on ionization and electron attachment (see, e.g., [10]). The function $A(n, |E|)$, which is used to model dissociative attachment, has a slightly pronounced maximum at the point $(n = 1, |E| = 1)$, at which it is equal to unity. Model expression (15) relates the free electron diffusion coefficient to the ambipolar diffusion coefficient at the ionization front [9]. In model (7)–(15), an important role is played by the function $nF(n, |E|)$, which is the total dimensionless rate of ionization and attachment. Note that this function is positive for $n < E$ and is peaked about $n = |E|/b$.

Equations (7)–(11) were solved in cylindrical coordinates (r, z) under the assumption that the discharge configuration possesses both axial symmetry $(\partial/\partial\varphi = 0)$ and symmetry with respect to the $z = 0$ plane. The computation region was $(0 < r < R, 0 < z < Z)$, with the obvious symmetry conditions at the axes. The sides of the computation region, R and Z, were chosen to be large enough for the gas-dynamic perturbations not to reach the boundaries of the computation region during the run time of the code. Helmholtz equation (11) was solved with the Sommerfeld asymptotic condition at infinity:

$$\mathbf{E} \sim \mathbf{E}^{ext} + \mathbf{C}(\mathbf{r}/|\mathbf{r}|)\exp(i|\mathbf{r}|)/|\mathbf{r}|. \tag{16}$$

The algorithm used to solve Helmholtz equation (11) with regard to the diffraction by a thin object was described in [11].

Our numerical simulations were carried out for conditions prevailing in the laboratory experiments described in [2–5]. Specifically, we simulated discharges driven by microwave radiation with a wavelength of 8.9 cm in air at a pressure of 152 torr at room temperature. Under such conditions, the measurement units and the constants are equal to

$$[r, z] = 1.42 \text{ cm}, \quad [n] = 4.8 \times 10^{18} \text{ cm}^{-3},$$

$$[p] = 2 \times 10^5 \text{ erg/cm}^3, \quad [V] = 2.88 \times 10^4 \text{ cm/s},$$
$$[t] = 49 \text{ μs},$$
$$[E] = 20 \text{ esu (6 kV/cm)}, \quad F_0 = 800, \quad f = 25,$$
$$D_0 = 0.16, \quad q = 166, \quad \gamma = 1.4, \quad b = 5.3.$$

The initial conditions require a separate discussion.

In studying discharges in an overcritical field, it is natural to use initial conditions corresponding to an immobile homogeneous gas medium with $n = 1$, $v = 0$, $w = 1$, and $\sigma = 0$. The external field is specified as a standing wave in which the nonzero amplitude of the axial field component is expressed in terms of the zero-order Bessel function, $E^{ext} = E_0 J_0(r)$. In order to initiate the discharge when $E^{ext} > 1$ (i.e., $|E| > n$), it is sufficient to create an arbitrarily small region with an arbitrarily low degree of ionization. In this case, the condition $nF > 0$ holds everywhere, so that nothing can prevent avalanche ionization followed by the development of the discharge. However, we are interested here in the opposite situation, namely, in discharges in a subcritical field, $E_0 < 1$. In this case, we have $|E| < n$ and $nF < 0$, so that the electron density in the initial stage should decrease. However, the discharge generally continues to develop because the heating source (described by the term $q\sigma|E|^2$) creates an excess pressure, which expels the gas from the central region. Since the gas density in this region can significantly decrease (i.e., fall below $|E|$), the function F may become positive and the number of electrons may start to increase. The higher the initial electron density, the earlier the situation changes. However, when the initial electron density in a local region is low, the time interval after which the situation changes is so long that the avalanche ionization practically does not come into play. That is why, in order to initiate a subcritical discharge, it is necessary either to increase the electric field to an overcritical level by inserting a metal initiator into a local region or to create (e.g., by laser breakdown) a local cavity with a reduced gas density at the minimum of which the breakdown condition $n < E_0$ is satisfied.
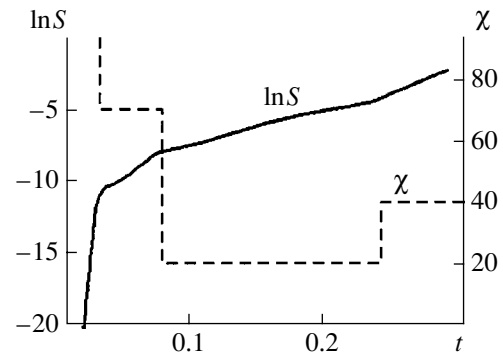

Fig. 1. Time evolutions of both the logarithm of the total number of electrons and the growth rate of this number ($[t] = 49$ μs).

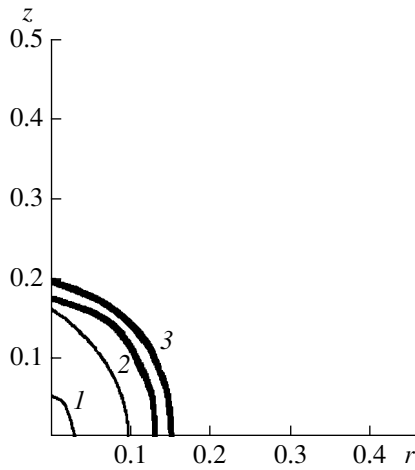**Fig. 2.** Contours of constant gas density at $t = 0.15$: $n =$ (*1*) 0.07, (*2*) 0.2, and (*3*) 1.5 ([r, z] = 1.42 cm, [n] = 4.8 × 10^{18} cm^{-3}).
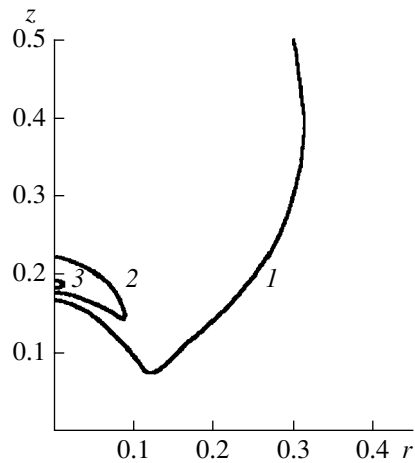
Fig. 3. Contours of constant field amplitude at $t = 0.15$: $|E| =$ (*1*) 0.5, (*2*) 1, and (*3*) 1.5 ([r, z] = 1.42 cm, [E] = 20 esu, [t] = 49 μs); the values of [r, z] and [t] also refer to Figs. 4–8.
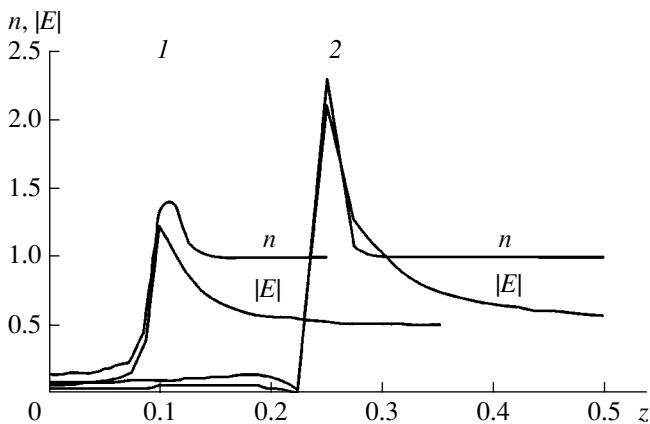
Fig. 4. Profiles of the gas density and field amplitude along the z-axis ($r = 0$) at $t =$ (*1*) 0.1 and (*2*) 0.2 ([n] = 4.8 × 10^{18} cm^{-3}, [E] = 20 esu).

Here, we consider precisely the latter way of initiating discharges and adopt the initial conditions corresponding to a local cavity (a local region with a reduced gas density) in the vicinity of the origin of coordinates. The time during which the gas density in the cavity is low is restricted by conventional heat conduction and, thus, can be sufficiently long. On the other hand, recombination quickly reduces the electron density to a low level. Because of this, the initial electron density in the cavity is assumed to be low.

The avalanche ionization condition is equivalent to the condition for the zero solution to Eq. (10) to be unstable. A rough estimate based on the solution to this equation for constant values of D and nF yields the inequality

$$D < nFL^2, \qquad (17)$$

where L is the characteristic dimension of the region with $F > 0$. We thus arrive at the familiar breakdown condition [10].

If, in the region of a reduced gas density, inequality (17) is satisfied and σ is small, then, the only process in the initial discharge stage is electron production. The gas remains immobile and the field is unperturbed. As the electrical conductivity begins to increase, the gas and the field both start to evolve.

The results of simulations described in the next section were obtained for the following initial conditions, which satisfy inequality (17): the amplitude of the external field is two times smaller than the critical value ($E_0 = 0.5$), and, in a cavity with dimensions of about $r, z < 10^{-2}$, the gas density and electrical conductivity are $n < 0.5$ and $\sigma \sim 10^{-10}$.

CALCULATED RESULTS

In presenting the numerical results, we start with the analysis of the integral parameter $S = \iint n_e\, 2\pi r\, dr\, dz$, which is the total number of electrons. The time evolution of this parameter is illustrated in Fig. 1. A linear increase in the function $\ln S(t)$ indicates an exponential increase in the function $S(t)$, the characteristic exponential index (the growth rate) being $\chi = d(\ln S)/dt$. The values of the growth rate (which is also shown in Fig. 1) allow us to see that the discharge process proceeds in the following four stages.

In the first stage ($0 < t < 0.02$), we have $\chi = 1000$ (this value, however, is beyond the frame of the figure). The electrons are produced in the way described above, and the gas state changes insignificantly.

The second stage ($0.02 < t < 0.07$) is in a sense a transition stage. As a result of the initial heating, the gas starts to expand, giving rise to a shock wave. The characteristic growth rate drops abruptly to $\chi = 70$.

In the third, fairly long, stage ($0.07 < t < 0.23$), the processes of the formation of the ionization front, rar-
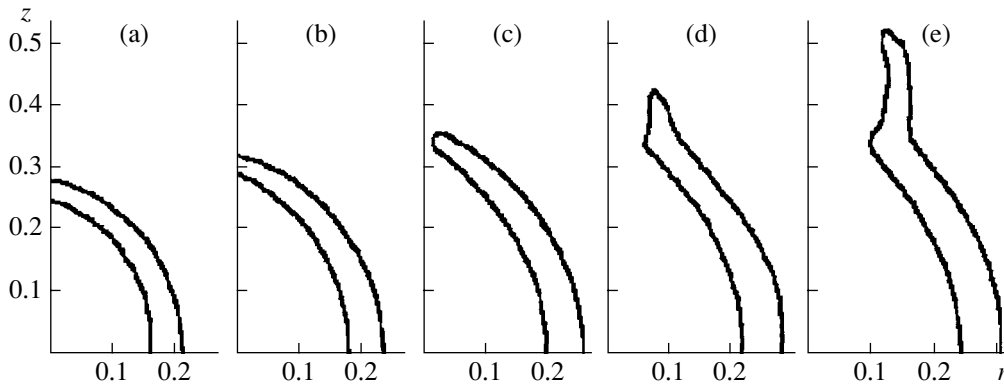
Fig. 5. Dynamics of the compressed gas layer ($n > 1.3$) at the times $t =$ (a) 0.19, (b) 0.21, (c) 0.23, (d) 0.25, and (e) 0.27 from the time interval $t = 0.19$–0.27 ($[n] = 4.8 \times 10^{18}$ cm^{-3}).

efaction wave, and shock wave all come to an end. The shock wave slows the discharge evolution: the growth rate χ reduces to 20. The representative distributions of the gas density and electric field in this stage are shown in Figs. 2 and 3, respectively. The gas density distribution, represented by the contours $n = 0.07$, $n = 0.2$, and $n = 1.5$, as well as by two heavy curves corresponding to the boundaries of the compressed gas layer, is still seen to keep its original elliptically symmetric shape.

The electric field distribution represented in Fig. 3 by the contours $|E| = 0.5$ and $|E| = 1$ differs radically from the gas density distribution. The electric field is seen to have the maximum $|E| = 1.5$ at the z-axis. Due to the presence of this maximum, the conducting plasma cloud extends slowly in the z direction.

Figure 4 shows profiles of the gas density and electric field along the z-axis at the times $t = 0.1$ and 0.2. It may be said that the n profile "follows" (approaches from below) the $|E|$ profile in such a way that the parameter $|E|/n$ tends to unity and the ionization rate nF falls to zero. Note also that, by the time $t = 0.2$, the field perturbation overtakes the shock wave.

The successive positions of a compressed ($n > 1.3$) gas layer during the time interval from $t = 0.19$ to 0.27 are shown in Fig. 5, which illustrates the dynamics of a transition from the third to the fourth discharge stage. The fourth stage begins at the time $t = 0.23$, at which the gas in the layer compressed by the shock wave breaks through the layer boundary near the z-axis to redistribute the discharge configuration. During the same time interval, the propagation velocity of the ionization front (shown by the contours $\sigma = 1$ in Fig. 6) increases abruptly.

The distribution of the current flowing through a cross section $z = $ const is illustrated in Fig. 7, which shows profiles of the real and imaginary parts of the current amplitude $I_z(z) = \int \sigma E_z 2\pi r dr$ at two times. The real parts are everywhere positive, and the imaginary parts are almost everywhere negative. Time evolutions of the real and imaginary parts of the amplitude of the

current flowing through the cross section $z = 0$ are given in Fig. 8, which also shows how the ionization front $Z(t)$ ($\sigma = 1$) propagates along the z-axis. From this figure, we can again see that the discharge process changes radically at the time $t = 0.23$: the propagation velocity of the ionization front increases by one order of magnitude (from 1 to 10).

We may say that only after reaching this stage did the discharge and breakdown actually happen. That the discharge process intensifies after breakdown is also
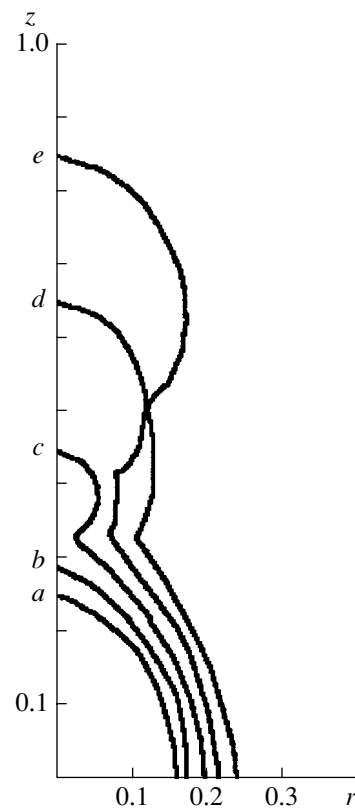

Fig. 6. Positions of the ionization front ($\sigma = 1$) during the time interval $t = 0.19$–0.27 at the same times as in Fig. 5 ($[\sigma] = \omega/4\pi$).
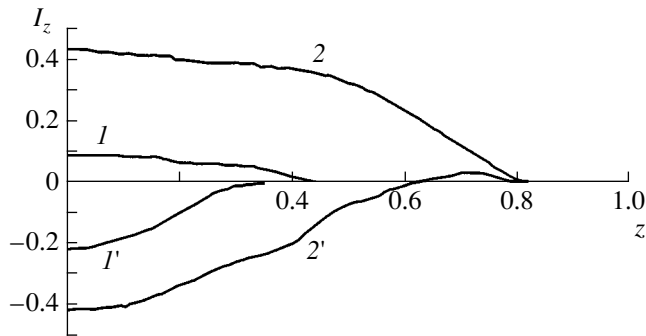
Fig. 7. Profiles of the real (curves *1*, *2*) and imaginary (curves *1'*, *2'*) parts of the current amplitude $I_z(z) = \int \sigma E_z 2\pi r dr$ ([σ] = ω/4π, [E] = 20 esu). Curves *1* and *1'* were calculated for $t = 0.23$, and curves *2* and *2'* were calculated for $t = 0.27$.
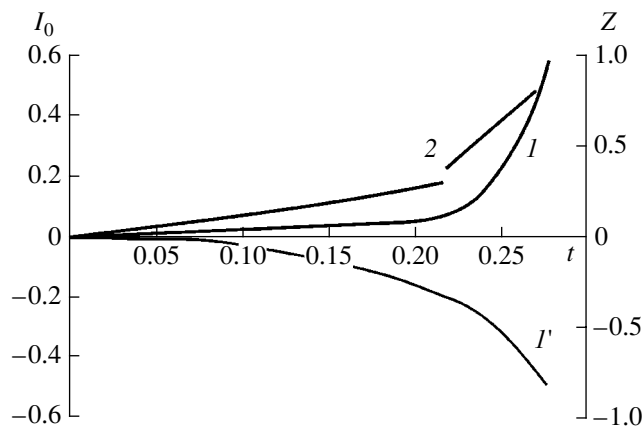

Fig. 8. Time evolutions of the real (curve *1*) and imaginary (curve *1'*) parts of the amplitude $I_z(z) = j\sigma E_z 2\pi r dr$ of the current flowing through the cross section $z = 0$ and change in the position $Z(t)$ of the ionization front (curve *2*) with time ([σ] = ω/4π, [E] = 20 esu)

clear from Figs. 1 and 8, which show the time evolutions of the total number of electrons $S(t)$ and of the current amplitude $I_z(z, t)$ in the cross section $z = 0$. In particular, during the last (fourth) time interval, the growth rate χ sharply increases to 40.

As the total length of the streamer approaches the resonant value (equal to half of the radiation wavelength), all discharge processes become faster and the induced current increases in a jumplike manner, so that the streamer will soon explode as a result of rapid heating. However, this discharge stage is beyond the scope of our study.

CONCLUSION

We have developed a mathematical model that takes into account the minimum possible number of required factors. The model not only confirms to the results obtained thus far but also promotes a better understanding of discharges in a subcritical field (in particular, at a qualitative level).

From our investigations, it is possible to draw the following conclusions.

The results obtained confirm the streamer nature of microwave discharges in gases at comparatively high pressures and prove that microwave discharges can develop in a subcritical field (i.e., in a field whose strength is below the threshold for breakdown).

Our results also confirm that, in describing the development of a subcritical streamer discharge, it is possible to dispense with ionizing hard radiation from the streamer channel, which is not regarded now as a necessary factor.

It is shown that, in order to initiate a microwave discharge in a subcritical field, it is sufficient to create a local cavity with a radius larger than a certain radius and with a reduced gas density at the minimum of which the condition of positive electron balance in an unperturbed field should be satisfied.

The wings of the streamer develop mainly at the expense of an increase in the field at the streamer ends and also, but only later, because of the plasma heating in the streamer channel. The hot region of the streamer is markedly shorter than its conducting shoulders.

As the streamer wings develop, the spherical shock wave originating in the initial stage evolves into a cylindrical shock wave. In the region in front of the streamer head, there is no shock wave.

REFERENCES

1. G. M. Batanov, S. I. Gritsinin, I. A. Kossyĭ, *et al.*, Tr. Fiz. Inst. Akad. Nauk SSSR **160**, 174 (1985).
2. L. P. Grachev, I. I. Esakov, G. I. Mishin, *et al.*, Zh. Tekh. Fiz. **55**, 389 (1985) [Sov. Phys. Tech. Phys. **30**, 228 (1985)].
3. L. P. Grachev, I. I. Esakov, G. I. Mishin, and K. V. Khodataev, Zh. Tekh. Fiz. **64** (2), 26 (1994) [Tech. Phys. **39**, 130 (1994)].
4. L. P. Grachev, I. I. Esakov, G. I. Mishin, and K. V. Khodataev, Zh. Tekh. Fiz. **66** (7), 32 (1996) [Tech. Phys. **41**, 652 (1996)].
5. L. P. Grachev, I. I. Esakov, and K. V. Khodataev, Zh. Tekh. Fiz. **68** (4), 33 (1998) [Tech. Phys. **43**, 378 (1998)].
6. V. B. Gil'denburg, I. S. Gushchin, S. A. Dvinin, and A. V. Kim, Zh. Éksp. Teor. Fiz. **97**, 1151 (1990) [Sov. Phys. JETP **70**, 645 (1990)].
7. K. V. Khodataev, Khim. Fiz. **12**, 303 (1993).
8. K. V. Khodataev, in *Proceedings of the XXIII International Conference on Phenomena in Ionized Gases, Toulouse, 1997*, Ed. by M. C. Bordage and A. Gleizes (EDP Sciences, Les Ullis, 1997), Contributed Papers, IV-24.
9. K. V. Khodataev and B. R. Gorelik, Fiz. Plazmy **23**, 236 (1997) [Plasma Phys. Rep. **23**, 215 (1997)].
10. A. D. MacDonald, *Microwave Breakdown in Gases* (Wiley, New York, 1966; Mir, Moscow, 1969).
11. O. I. Voskoboĭnikova, Preprint No. 38 (Keldysh Institute of Applied Mathematics, Russian Academy of Sciences, Moscow, 1997).

Translated by O. Khadin

====================== **SOLIDS** ======================

Potential Distribution in Ferroelectric Ceramic during Pulsed Exoelectron Emission

V. N. Ivanov and L. M. Rabkin

Rostov State University, pr. Stachki 194, Rostov-on-Don, 344090 Russia
e-mail: ivanov@phys.rnd.runnet.ru
Received July 11, 2001; in final form, January 18, 2002

Abstract—The potential and polarization distributions in a planar emitting layer of PLZT-9/65/35 ferroelectric ceramic with a set of conductive strip electrodes on the emitting side and a continuous electrode on the opposite side are studied by numerical methods. The state arising immediately after polarization switching at the leading edge of an applied voltage pulse (i.e., before the polarization charges are screened by free charges) is considered. When the pulsed field strength far exceeds the double coercive field, regions with alternating polarization are found to form in the surface layer between the strips. The normal component of the polarization at its maxima is close to saturation. The electric field on both sides of the surface varies as the polarization vector and reaches 200 kV/cm. At surface microirregularities, the electric field strength is much higher. This means that field emission is responsible for electron escape from the ferroelectric ceramic during pulsed polarization switching. © 2002 MAIK "Nauka/Interperiodica".

INTRODUCTION

Although electron emission from ferroelectric ceramics has been studied for a long time, its nature still remains unclear. It has been established [1] that this phenomenon is related to spontaneous polarization switching under the action of an electric field pulse. Clearly, to gain a better understanding of the emission mechanism, one has to study the distribution of the potential near the surface of the ferroelectric ceramic emitter at the instant the pulsed emission starts. The pulse is observed immediately after the leading edge of an applied field pulse, i.e., until the free charges of the ceramic neutralize the polarization charges.

The aim of this work is a self-consistent calculation of this potential. The field-pulse-switched polarization is assumed to be dependent on the field strength in the form adopted in [2] (the Preisach statistical model [3]).

MODEL AND CALCULATION OF POTENTIAL

We assume that the half-space $y > 0$ (hereafter, domain I) is occupied by the ferroelectric ceramic and the half-space $y < 0$ (domain II) is a vacuum. In the boundary plane $y = 0$, an array of thin conducting strips that are infinitely long in the z direction and have a width b in the x direction is arranged. Let the period of the array in the x direction be equal to $2b$ for simplicity. We also assume that the potential of the strips equals zero and that domain I is subjected to a pulsed electric field, which can be viewed as uniform away from the boundary:

$$E_y = E_0, \quad y \gg b. \tag{1}$$

A sufficiently high field strength E_0 switches the polarization, causing volume and surface (in the plane $y = 0$) polarization charges to appear. The field responsible for the polarization is the sum of the applied field and the self-field of the nonuniform polarization. When calculating the fields, one must take into account the effect of the grounded conducting strips. The field strength–polarization relationship is nonlinear [2]; therefore, a self-consistent solution to the problem of potential distribution can be found only numerically within our model.

We will use the method of successive approximations. At each step, the potential is calculated by the finite-difference approach combined with the Fourier transformation [4–6]. Then, the electric field strength is found by numerical differentiation and the polarization vector is determined by the technique reported in [2]. The normal component of the polarization vector between the strips specifies the surface polarization charges. The volume polarization charges are calculated by numerically differentiating the polarization components. The field sources thus found are subsequently used to calculate the potential in the next approximation.

POTENTIAL DUE TO THE APPLIED FIELD STRENGTH

The origin is placed in the middle of the strip spacing. In the plane $x0y$, we choose a square mesh with a step $d = b/M$ so that the edge of the strip corresponds to $x = dM/2$, where M is even. The symmetry of the problem allows us to calculate the potential only in the region $0 \le x \le b$.

For the potential at the domain boundary that is specified at points $x_m = md$ ($m = 0, 1, \ldots, M$), we write the discrete Fourier transform

$$v_m = \hat{v}_0/2 + \sum_{k=1}^{M-1} \hat{v}_k \cos\frac{\pi km}{M} + \hat{v}_M/2\cos\pi M, \quad (2)$$

$$m = 0, 1, 2, \ldots M,$$

$$\hat{v}_k = 2/M\left(v_0/2 + \sum_{m=1}^{M/2-1} v_m \cos\frac{\pi km}{M}\right), \quad (3)$$

$$m = 0, 1, 2, \ldots M.$$

Expression (3) for the Fourier coefficients takes into account that $v_m = 0$ at $m = M/2, M/2 + 1, \ldots, M$ (i.e., at points on the grounded strip).

Using the method of separation of variables [4], we write the solution to the Laplace difference equation for the potential at the points $x_m = md$ ($m = 0, 1, 2, \ldots, M$) and $y_n = nd$ ($n = 0, 1, 2, \ldots$) of domain I:

$$V_{nm}^{I} = \hat{v}_0/2 + \sum_{k=1}^{M-1} \hat{v}_k \cos\frac{\pi km}{M} e^{-\beta_k n}$$

$$+ \hat{v}_M/2\cos\pi M e^{-\beta_M n} - E_0 nd, \quad (4)$$

and at the points $x_m = md$ ($m = 0, 1, 2, \ldots, M$) and $y_n = -nd$ ($n = 0, 1, 2, \ldots$) of domain II:

$$V_{nm}^{II} = \hat{v}_0/2 + \sum_{k=1}^{M-1} \hat{v}_k \cos\frac{\pi km}{M} e^{-\beta_k n}$$

$$+ \hat{v}_M/2\cos\pi M e^{-\beta_M n}, \quad (5)$$

$$\beta_k = \text{arccosh}\left(2 - \cos\frac{\pi k}{M}\right). \quad (6)$$

The potential given by (4) and (5) meets the continuity condition at the boundary between domains I and II. To satisfy the continuity condition for the normal component of the electric induction between the strips, we write the expression for the y component of the field:

$$E_{ynm}^{VI} = E_0 + d^{-1}\sum_{k=1}^{M-1} \hat{v}_k \beta_k \cos\frac{\pi km}{M} e^{-\beta_k n}$$

$$+ d^{-1}\hat{v}_M/2\beta_M \cos\pi M e^{-\beta_M n}, \quad (7)$$

$$E_{ynm}^{VII} = -d^{-1}\sum_{k=1}^{M-1} \hat{v}_k \beta_k \cos\frac{\pi km}{M} e^{-\beta_k n}$$

$$- d^{-1}\hat{v}_M/2\beta_M \cos\pi M e^{-\beta_M n}. \quad (8)$$

The boundary condition for the normal component of the induction,

$$\varepsilon E_{ynm}^{VI} - E_{ynm}^{VII} = 0,$$

$$m = 0, 1, 2, \ldots, M/2 - 1 \quad (9)$$

(where ε is the ceramic permittivity unrelated to the spontaneous polarization), gives rise to the set of equations for the Fourier coefficients of the potential:

$$\sum_{k=1}^{M-1} \hat{v}_k \beta_k \cos\frac{\pi km}{M} + \hat{v}_M/2\beta_M \cos\pi m = -\varepsilon/(\varepsilon+1)dE_0, \quad (10)$$

$$m = 0, 1, 2, \ldots, M/2 - 1.$$

Set (10) should be supplemented by the set

$$\hat{v}_0/2 + \sum_{k=1}^{M-1} \hat{v}_k \cos\frac{\pi km}{M} + \hat{v}_M/2\cos\pi M = 0, \quad (11)$$

$$m = M/2, M/2 + 1, \ldots, M,$$

which is derived by equating the potential across the conducting strips to zero. A solution to sets (10) and (11), together with expressions (4) and (5), defines the potential in both domains in the absence of the spontaneous polarization.

The numerical solution thus constructed is a solution of the simple problem. It can also be solved analytically by conformal mapping. Then, the potential across the boundary plane is given by

$$v = 2\pi^{-1} E_0 b \, \text{arccosh}[2^{1/2}\cos(\pi x/2b)]. \quad (12)$$

However, the problem of potential due to polarization charges is solved only by the difference method. It is therefore of interest to evaluate its accuracy by contrasting with the result obtained analytically. In the model of indefinitely thin strips, the electric field strength obtained analytically grows infinitely at the strip edge, which is easy to check by differentiating expression (12). The solution found by the difference method does not possess such a property.

The solutions at the strip edges are inaccurate in our model also because of the fact that real electrodes have a finite thickness. From the following, however, it will be seen that the model describes adequately the typical behavior of the potential and polarization between the strips and, hence, sheds light on a possible emission mechanism.

The potentials in the middle of the strip spacing that are found numerically and analytically differ by 7% for $M = 8$. This is quite sufficient for qualitative estimates. The discrepancy decreases fast with increasing M.

Note also that the factors β_k in finite-difference expressions (7) and (8) for field strength should be replaced by $\sinh\beta_k$. However, the factors β_k, which appear when continuous Fourier series are used, provide a better accuracy, as follows from a comparison

between the results obtained by the difference and analytical methods.

POTENTIAL DUE TO POLARIZATION CHARGES

To solve the difference Poisson equation, from which the potential in domain I is found, one usually uses double discrete Fourier transformation [5]. In this case, the potential and its Fourier coefficients are related as

$$U_{nm}^{I*} = \sum_{l=1}^{N-1} \sum_{k=0}^{M}{}' \hat{u}_{lk}^* \sin\frac{\pi ln}{M} \cos\frac{\pi km}{M}, \qquad (13)$$

$$\hat{u}_{lk}^* = 4M^{-1}N^{-1} \sum_{n=1}^{N-1} \sum_{m=0}^{M}{}' U_{nm}^{I*} \sin\frac{\pi ln}{N} \cos\frac{\pi km}{M}, \qquad (14)$$

where primes at the signs of summation over k and m mean that the weighting coefficient of the terms with the numbers 0 and M is 1/2. The volume density ρ_{nm}^* of the polarization space charge and its Fourier coefficient $\hat{\rho}_{lk}^*$ are related by expressions similar to (13) and (14).

When calculating the potential and the space charge in the ceramic, we, according to the procedure adopted here, will specify the potential at $(N + 1)$ points in the y direction. However, it is known from the electrodynamic theory [7] that the array appreciably disturbs a uniform electric field normal to its surface only at a distance on the order of its period. One can therefore expect that effects related to the polarization nonuniformity show up also within this distance; that is, we need only take $N = 2M$. The validity of such a selection of N is supported by subsequent results.

The Fourier coefficients of the potential and charge density relate as [5]

$$\hat{u}_{lk}^* = \hat{\rho}_{lk}^* d^2/(\varepsilon\varepsilon_0\lambda_{lk}), \qquad (15)$$

where

$$\lambda_{lk} = 4 - 2\cos\frac{\pi l}{N} - 2\cos\frac{\pi k}{M} \qquad (16)$$

are the eigenvalues of the difference Laplacian.

Equalities (13)–(15) define the potential vanishing at the boundary between the domains. To construct a solution satisfying the nonuniform boundary conditions

$$U_{0m}^{I} = \begin{cases} u_m, & m = 0, 1, ..., M/2 - 1 \\ 0, & m = M/2, M/2 + 1, ..., M, \end{cases} \qquad (17)$$

we take advantage of the equivalent charge method [6]. To this end, we relate the equivalent charge ρ_{nm}^* and the real charge ρ_{nm} by the equality

$$\rho_{nm}^* = \begin{cases} \rho_{1m} + \varepsilon\varepsilon_0 U_{0m}^{I}, & n = 1 \\ \rho_{nm}, & n \neq 1, \end{cases} \qquad (18)$$

which provides the fulfillment of the nonuniform boundary conditions for the potential U_{0m}^{I} we are interested in.

It is convenient to introduce the Green difference function

$$G(n, m; n', m') = \frac{4d^2}{MN} \sum_{l=1}^{N-1} \sum_{k=0}^{M}{}' \lambda_{lk}^{-1} \sin\frac{\pi ln}{N}$$
$$\times \cos\frac{\pi km}{M} \sin\frac{\pi ln'}{N} \cos\frac{\pi km'}{M}, \qquad (19)$$

which relates the potential in domain I and its sources:

$$U_{nm}^{I} = (\varepsilon\varepsilon_0)^{-1} \sum_{n'=1}^{N-1} \sum_{m'=0}^{M}{}' G(n, m; n', m')\rho_{n'm'}$$
$$+ d^{-2} \sum_{m'=0}^{M/2-1}{}' G(n, m; 1, m')u_{m'}. \qquad (20)$$

The potential in domain II that takes boundary conditions (17) is defined by (2)–(4), where u_m and $\hat{u}_k$ are substituted for u_m and $\hat{u}_k$, respectively.

The y components of the electric field are represented through the difference derivatives of associated (volume for domain I and surface for domain II) Green functions with respect to n:

$$G_n(n, m; n', m') = \frac{4d}{MN} \sum_{l=1}^{N-1} \sum_{k=0}^{M}{}' \lambda_{lk}^{-1} \sin\frac{\pi l}{N} \cos\frac{\pi ln}{N}$$
$$\times \cos\frac{\pi km}{M} \sin\frac{\pi ln'}{N} \cos\frac{\pi km'}{M}, \qquad (21)$$

$$g_n(m; m') = \frac{2}{dM} \sum_{k=1}^{M}{}' \beta_k \cos\frac{\pi km}{M} \cos\frac{\pi km'}{M}, \qquad (22)$$

$$E_{ynm}^{U\text{II}} = -(\varepsilon\varepsilon_0)^{-1} \sum_{n'=1}^{N-1} \sum_{m'=0}^{M}{}' G_n(n, m; n', m')\rho_{n'm'}$$
$$- d^{-2} \sum_{m'=0}^{M/2-1}{}' G_n(n, m; 1, m')u_{m'}, \qquad (23)$$

$$E_{ynm}^{U\text{II}} = -d^{-1} \sum_{m'=1}^{M}{}' g_n(m; m')u_{m'}. \qquad (24)$$

The continuity condition for the normal component of the electric induction, including the switchable part

of the polarization, between the strips,

$$\varepsilon\varepsilon_0 E_{y0m}^{U\mathrm{I}} - \varepsilon_0 E_{y0m}^{U\mathrm{II}} = -P_{y0m},$$
$$m = 0, 1, 2, \ldots, M/2 - 1 \tag{25}$$

gives rise to a set of equations for the potential u_m at the boundary:

$$\sum_{m'=0}^{M/2-1}{}' \{ d^{-1} g_n(m; m') - \varepsilon d^{-2} G_n(0, m; 1, m') \}$$

$$= -\varepsilon_0^{-1} P_{y0m} + \varepsilon_0^{-1} \sum_{n'=1}^{N-1} \sum_{m'=0}^{M}{}' G_n(0, m; n', m') \rho_{n'm'}, \tag{26}$$

$$m = 0, 1, \ldots, M/2 - 1.$$

In (25) and (26), the polarization component $-P_y$ is the surface density of charges related to the switchable part of the polarization.

Once set of equations (26) has been solved, the potential in domain I is calculated with expression (20), where the potential matrix is supplemented by the row U_{0m}^{I} according to (17) and the row of zeros, $U_{Nm}^{\mathrm{I}} = 0$. In domain II, the potential U_{nm}^{II} is calculated with expressions similar to (3) and (5).

POLARIZATION AND POLARIZATION CHARGES

The potential produced by an applied field and the switchable polarization is represented as the sums $V_{nm}^{\mathrm{I}} + U_{nm}^{\mathrm{I}}$ and $V_{nm}^{\mathrm{II}} + U_{nm}^{\mathrm{II}}$. The total field strength, which is necessary for calculating the switchable part of the polarization, is determined by numerically finding the potential gradient. We use the central differences except for the boundaries $n = 0$ and $n = M$, where ordinary differences are used. At the boundaries $m = 0$ and $m = M$, the array symmetry is taken into account.

According to [2], the polarization vector is given by

$$P = \begin{cases} 0, & E < 2E_\mathrm{c} \\ (\mathbf{E}/E)(P_\mathrm{s}/3)\{ \mathrm{erf}[(E - E_\mathrm{c})/(2^{1/2}\sigma_\mathrm{i})] \\ \quad - \mathrm{erf}[E_\mathrm{c}/(2^{1/2}\sigma_\mathrm{i})] \}, & E > 2E_\mathrm{c}, \end{cases} \tag{27}$$

where $\mathrm{erf}(x)$ is the error function [8], P_s is the spontaneous polarization, E_c is the coercive field, and σ_i^2 is the internal field variance.

The volume density of the polarization charges is calculated by the formula

$$\rho = -\mathrm{div}\mathbf{P}. \tag{28}$$

Numerical differentiation here is accomplished similarly to the potential differentiation. In calculating the

Fourier coefficients of the space charge $\hat{\rho}_{lk}$ with an expression like (14), the space charge densities at the boundaries $n = 0$ and $n = N$ are ignored. This means that the total polarization charge, responsible for the potential, is other than zero. To eliminate the arising error, the matrix elements ρ_{1m} are replaced by $\rho_{1m} + \rho_{0m}/2$ and the matrix elements $\rho_{N-1, m}$, by $\rho_{N-1, m} + \rho_{Nm}/2$ prior to subsequent calculations. Then, as may be easily checked, the condition of charge conservation is fulfilled if the above numerical differentiation formulas are used.

METHOD OF SUCCESSIVE APPROXIMATIONS

The calculation sequence described above is one step of the method of successive approximations. The potential V_{nm} in the absence of polarization charges is taken as the zero approximation. The mere repetition of steps does not provide the convergence of the method; therefore, after each ith step, the refined approximation $\tilde{\rho}_{nm}^{(i)}$ of the charge density matrix is calculated through the matrix $\rho_{nm}^{(i)}$ calculated at this step by the formula

$$\tilde{\rho}_{nm}^{(i)} = \alpha\rho_{nm}^{(i)} + (1 - \alpha)\tilde{\rho}_{nm}^{(i-1)}, \tag{29}$$

where the constant parameter α is taken so that it provides the highest rate of convergence; in other words, the modified Newton–Kantorovich method [9] is applied.

The tentative value of the parameter α is determined from the first several steps by the formula for convergence acceleration [9]:

$$\alpha = \frac{l_{i+1} - l_i}{l_{i+2} - 2l_{i+1} + l_i}, \tag{30}$$

where lower harmonics of the space charge matrix, for example, $\hat{\rho}_{12}^{(i)}$, are taken as the linear functional l_i.

RESULTS OF CALCULATION

The potential and polarization were calculated for PLZT-9/65/35 ferroelectric ceramic. Its permittivity $\varepsilon = 4650$ [10]. The parameters entering into expression (27) were calculated in [2]: at 20°C, $P_\mathrm{s} = 74\ \mu\mathrm{C/cm^2}$, $E_\mathrm{c} = 3.73$ kV/cm, and $\sigma_\mathrm{i} = 11.4$ kV/cm. The width b of the conducting strips is taken to be 0.2 mm. As the applied electric field, we took the field at which intense electron emission was observed [11]: $E_0 = 10$ kV/cm.

Electrons are emitted from the ceramic surface between the strips; therefore, in what follows the field characteristics are given for this part of the ceramic. Recall that the origin is in the middle of the strip spacing, the x axis lies in the plane normal to the strips, and the y axis is directed inward to the ceramic.

Figure 1 shows the calculated potential U and the field components in a vacuum, E_{yv}, and in the ceramic,

E_{yc}, on the surface $y = 0$ in the gap. The normal components E_{yc} and $-E_{yv}$ differ insignificantly; hence, they are produced largely by surface polarization charges with a density $-P_y$. The x dependence of the tangential field component E_x on the surface (omitted in Fig. 1) also has an oscillatory character. Its extrema are close to and lie between those of the normal components.

Such dependence of the field is related to the distribution of the switchable part of the polarization, as follows from Fig. 2, which demonstrates the directions of this vector near the surface in the middle part of the gap. In the immediate vicinity of the surface, the polarization directions alternate. On the surface, the polarization reaches a maximum: $P_s/3\{1 - \text{erf}[E_c/(2^{1/2}\sigma_i)]\} \cong 18 \ \mu\text{C/cm}^2$, as follows from (27). At a distance equal to one variation of x from the surface along the y direction, the polarization drops by one order of magnitude. The field of the vector $\mathbf{E}$ in the ceramic is similar to that of $\mathbf{P}$ but drops significantly with y at a distance half as large as that for $\mathbf{P}$.

The reason for the polarization alternation is that in a uniform external field somewhat exceeding $2E_c$, the near-saturation polarization produces a depolarizing field far exceeding $2E_c$; therefore, the polarization cannot be uniform. In other words, the reason is akin to that causing domain nucleation in a single crystal. However, in the ceramic, a large number of domains is supposed to exist in either of the regions (I and II) according to the Preisach model [2], within which expression (27) has been derived in [2].

The nonuniform distribution of the polarization over the ceramic surface is likely to be due to the nonuniform boundary conditions because of the presence of the strip array. To check this assumption, we studied the polarization in a uniform structure. Let a semi-infinite ceramic sample be subjected to a uniform external field produced by unlike charges uniformly distributed over the free surface placed at infinity. With the uniform field taken as the initial approximation, layers that are parallel to the surface and have an alternating-direction polarization normal to the surface appear. If the initial approximation is disturbed by a periodic harmonic perturbation of the surface potential, successive approximations converge to a solution where regions similar to those existing in the ceramic with the strips and having the associated period will be observed in the surface layer. Such a regular structure apparently has a number of equilibrium states, each with its own local energy minimum, and iterations converge to one of them depending on the initial approximation chosen.

Let us now discuss the admissible number M of partitions (mesh steps). The number M defines the number of Fourier components involved in the solution and the maximal number of the variations of the solution in the x direction.

It was shown that, as M grows, the number of variations increases in proportion while the surface layer
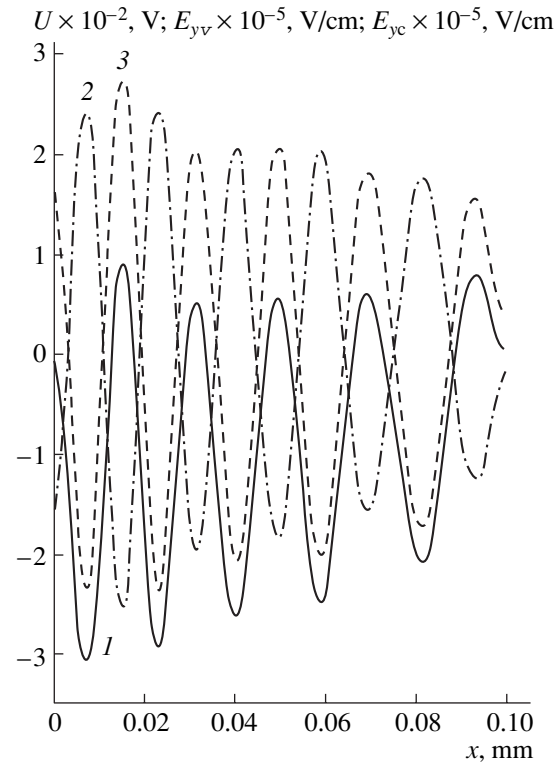

$U \times 10^{-2}$, V; $E_{yv} \times 10^{-5}$, V/cm; $E_{yc} \times 10^{-5}$, V/cm

Fig. 1. Distributions of (*1*) the potential U and electric field components (*2*) in vacuum E_{yv} and (*3*) in ceramic E_{yc} over the ceramic surface.

thickness decreases in inverse proportion. The field, polarization, and stored energy density in this layer are high but their maximal values remain virtually unchanged. It appears that the solution corresponds to the least possible electrostatic energy for a given M. However, such an assumption is valid if the ceramic is highly uniform. Actually, it consists of crystallites separated by the nonpolar glassy phase. One can expect that crystallites (up to 10 μm in size) are highly polarized regions. Therefore, in the calculations (Figs. 1, 2), we put $M = 32$, although the numerical method retains stability at M twice as large. With such M, the applicability condition for the Preisach model and, accordingly, expression (27) are fulfilled: the domain size (1 μm in the ceramic) is much smaller than the characteristic length of the electrostatic problem.

RESULTS AND DISCUSSION

Thus, if a pulsed electric field of strength far exceeding the double coercive force is applied to a PLZT-9/65/35 ferroelectric ceramic emitter, regions of alternating-direction polarization arise in the near-surface layer between the strips. The polarization is normal to the surface and approaches the saturation value in the maximum. Near the surface (both inside and outside the ceramic), an electric field of strength reaching 200 kV/cm appears. Its direction varies consistently
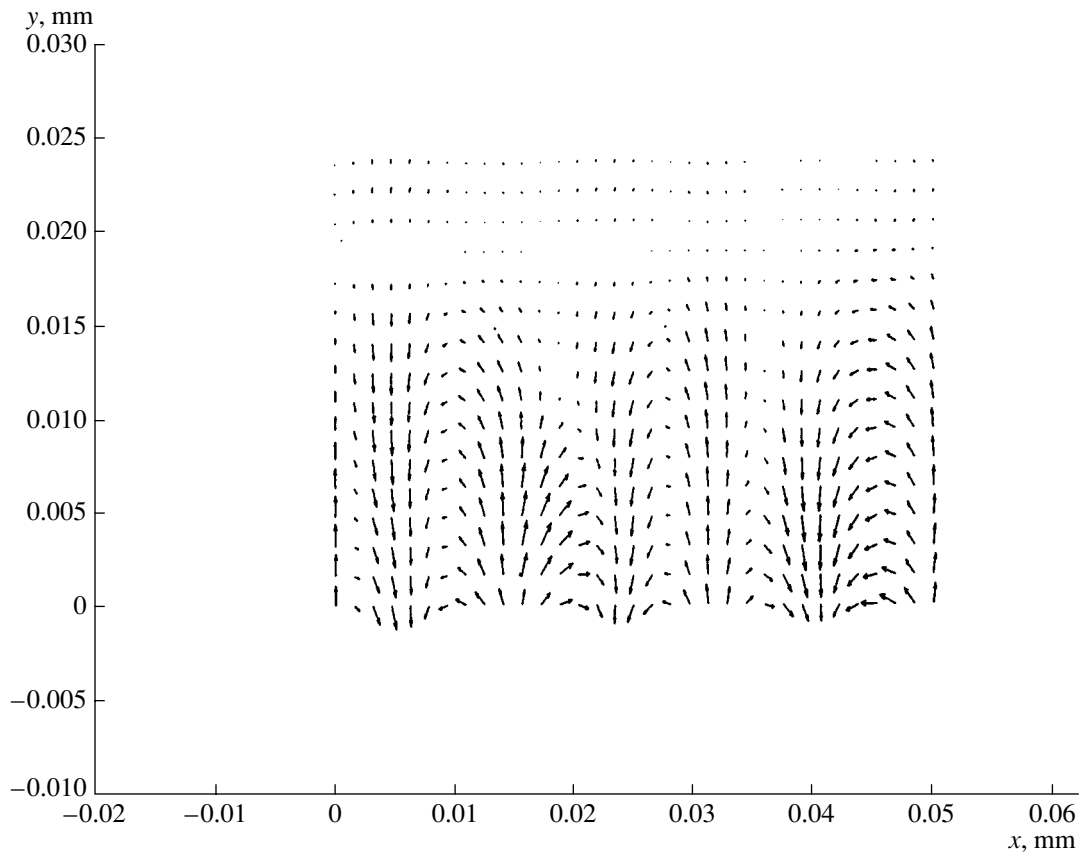
Fig. 2. Polarization vector field near the surface in the middle of the strip spacing.

with the direction of the polarization vector. Note that this field in the ceramic is comparable to the breakdown value but, being concentrated in the thin surface layer, only bends the bands and causes carrier avalanche multiplication by impact ionization. In a layer of thickness on the order of 10^{-7} cm, the carrier concentration was estimated as 10^{21} cm^{-3} [12, 13]. It can be expected that, as in barium titanate [12], carrier sources here are impurity levels (such as F centers of oxygen vacancies) or surface levels.

The calculated value of the field strength in the surface layer sheds additional light on electron emission from ferroelectric ceramics observed in experiments and suggests that there is a fundamental similarity between this emission and the emission from insulators due to the injection of electrons with energies ranging from 0.2 to 0.5 MeV [14, 15]. Emission from insulators, called critical or explosive, arises when the charge accumulated in the surface layer produces a surface field on the order of 100 kV/cm. An emission pulse with a duration of several to several tens of nanoseconds and a current density of several tens to several thousands of amperes/cm^2 ejects the injected charge and then the process of charge accumulation/ejection may repeat many times. The similarity with emission from ferroelectric ceramics is also in that a plasma resulting in

vacuum discharge under certain conditions is generated on the insulator surface.

Taking into account the qualitative and quantitative similarity of these two types of electron emission, we believe that approaches to the development of a unified emission theory put forward in [15] apply to emission from ferroelectric ceramics. Specifically, it can be argued that microirregularities on the ceramic surface (surface roughness and surface structure nonuniformity) are responsible for the high current density. It seems likely that a local field enhancement by one or two orders of magnitude at a surface irregularity causes intense field emission.

To further develop the quantitative theory of emission from ferroelectric ceramics that arises during polarization switching, it is necessary to elucidate changes in the electron subsystem of an imperfect ceramic during polarization switching.

ACKNOWLEDGMENTS

REFERENCES

1. W. Zhang and W. Huebner, Ferroelectrics **221**, 199 (1999).
2. V. N. Ivanov and L. M. Rabkin, Zh. Tekh. Fiz. **70** (12), 43 (2000) [Tech. Phys. **45**, 1553 (2000)].
3. A. V. Turik, Fiz. Tverd. Tela (Leningrad) **5**, 1213 (1963) [Sov. Phys. Solid State **5**, 885 (1963)].
4. P. M. Morse and H. Feshbach, *Methods of Theoretical Physics* (McGraw-Hill, New York, 1953; Inostrannaya Literatura, Moscow, 1958), Vol. 1.
5. D. E. Potter, *Computational Physics* (Wiley, New York, 1973; Mir, Moscow, 1975).
6. R. Hockney and J. Eastwood, *Computer Simulation Using Particles* (McGraw-Hill, New York, 1984; Mir, Moscow, 1987).
7. L. A. Vaĭnshteĭn, in *High-Power Electronics* (Akad. Nauk SSSR, Moscow, 1963), Vol. 2, pp. 25–56.
8. *Handbook of Mathematical Functions*, Ed. by M. Abramowitz and I. A. Stegun (National Bureau of Standards, Washington, 1964; Nauka, Moscow, 1979).
9. M. A. Krasnosel'skiĭ, G. M. Vaĭnikko, P. P. Zabreĭko, *et al.*, *Approximate Solutions of Operator Equations* (Nauka, Moscow, 1969).
10. G. H. Haertling and C. E. Land, J. Am. Ceram. Soc. **50**, 1 (1971).
11. H. Hundel, J. Hadrek, H. Riege, *et al.*, Ferroelectrics **110**, 183 (1990).
12. G. M. Guro, I. I. Ivanchik, and N. F. Kovtonyuk, in *Barium Titanate* (Nauka, Moscow, 1973), pp. 71–77.
13. M. E. Lines and A. M. Glass, *Principles and Applications of Ferroelectrics and Related Materials* (Oxford Univ. Press, Oxford, 1977; Mir, Moscow, 1981).
14. A. Watson and J. Dow, J. Appl. Phys. **39**, 5935 (1968).
15. D. I. Vaĭsburd, S. I. Tverdokhlebov, and T. A. Tukhfatullin, Izv. Vyssh. Uchebn. Zaved., Fiz. **40** (11), 45 (1997).

Translated by V. Isaakyan

───────────────────────── **SOLIDS** ─────────────────────────

On the Mechanism of Cratering on Solid Surfaces Exposed to an Intense Charged Particle Beam

N. B. Volkov, A. E. Maĭer, and A. P. Yalovets

Institute of Electrophysics, Ural Division, Russian Academy of Sciences,
ul. Amundsena 106, Yekaterinburg, 620016 Russia
e-mail: nbv@ami.uran.ru
Received July 17, 2001

Abstract—A physical mechanism of cratering on the surface of solid targets exposed to intense charged particle fluxes is suggested. According to this mechanism, the craters form due to surface gravitational waves and Richtmyer–Meshkov instability of the plasma torch free surface [1–4]. The crater sizes and shapes predicted theoretically agree well with experimental observations. It is shown that the stresses arising in the target are the highest under the crater, which explains the localization of structure modification observed in the experiments. © 2002 MAIK "Nauka/Interperiodica".

INTRODUCTION

Material irradiation by intense charged particle beams is widely used in various technologies for improving the properties of the surface layer. Resulting defects and surface films improve the strength, as well as wear and erosion resistances, of the material surface. The disadvantage of such a surface processing is the formation of craters and other surface features, making the surface rough. Cratering is a complex physical process. Its study is of interest both for applications (i.e., for controlling the beam processing of structural materials) and from the scientific point of view. In this article, our goal is (1) to construct physical and mathematical models of cratering and (2) to perform a numerical experiment to derive basic conditions for cratering on the surface and relationships for stress fields in the target.

PHYSICAL MODEL OF CRATERING

Before constructing a model of cratering, we will recall some experimental data for this phenomenon found in [5, 6]. In the works cited, the beam ion energy was $E = 250$ keV; shot duration τ, 50 ns; and beam current density j_b, 50–200 A/cm^2. Carbon ions and protons accounted for 70 and 30% of the beam. The number of shots N was between 1 and 10. Targets were thin (0.3 mm) plates made of different materials. The surface to be irradiated was electrochemically polished.

In [5], craters were observed on all the samples irrespective of the current density even if the surface temperature was below the melting point. Their sizes varied over a wide range. The craters were always centrosymmetric and had an annular structure. Central symmetry may be somewhat violated, i.e., by the nearest neighbor. The ring-like structure may be both well developed

and lack peripheral rings. Because of this, the craters apparently exhibit a wide variety of forms. The central part of the crater may contain both a recess and a protrusion.

In general, the craters are randomly arranged on the surface. However, one can distinguish regions where small craters ($d = 0.1$–2.0 μm) of density $n \approx 5 \times 10^{10}$ m^{-2} dominate, regions with small and medium-size craters ($d \approx 3$–10 μm, $n \approx 10^9$ m^{-2}), and those with large craters ($d > 10$ μm, $n \leq 10^7$–10^8 m^{-2}). Large craters concentrate, as a rule, near the beam center, where the radiation energy density is the highest. Small ones prevail over the periphery of the irradiation area. From the crater size distribution for nickel, it follows that small craters are associated with a low current density. A fraction of large craters increases grows as the current density rises. For molybdenum, the situation is similar [5]. The positions of the craters do not correlate with grain boundaries and shear tracers, that is, to the material structure under the surface. As the number of shots increases from 1 to 10, the typical sizes of the craters rise from several microns to several tens of microns and their density drops from $n = 1.5 \times 10^8$ to 4×10^7 m^{-2}. The depth of the craters was no larger than 1.5–2.0 μm.

Images of the target surface [5, 6] show well-defined high-contrast craters and many diffuse craters of much lower contrast. It was assumed [5, 6] that the distinct craters were produced by the last shot, while the diffuse ones were produced by preceding shots and then were smoothed out by surface melting, evaporation, and plasma cloud condensation.

From the experiments, it also follows that protective films and structure defects form most intensely near the craters. The surface inside and near the craters is covered by a film about 100 nm thick. Away from the craters, the films either do not form at all or their thickness

is no more than 10–20 nm. At low current densities ($j = 5 \times 10^5$ A/m^2), the dislocation density in the surface layer is high (10^{14} m^{-2}) near the craters and low away from them. At $j = 2 \times 10^6$ A/m^2, the dislocations are uniformly distributed with a density of $(2–5) \times 10^{14}$ m^{-2}; however, the concentration of structure defects near the craters is even higher.

Bearing in mind the experimental results [5, 6], we suggest the following mechanism of cratering. When striking a solid target, intense charged particle beams with a power density of more than 10^{11} W/m^2 heat it up heavily, producing a plasma torch with a temperature $T = (1–5) \times 10^3$ K and a strongly nonuniform concentration. Experiments and calculations (see, e.g., [7]) have shown that the plasma front velocity becomes as high as ~10^3 m/s for the shot duration 100 ns and the plasma front acceleration reaches ~10^{10} m/s^2. With such acceleration values, gravitational waves may be excited on the surface irradiated [1]. Nuclei (emitters) of the gravitational waves are local microprotrusions or microrecesses on the surface. Their height (depth) may be very small (from several tenths to several hundredths of a micron).

It will be shown below that gravitational waves originating at a localized perturbation are annular waves. The size and number of their rings depend on the size and shape of the perturbation, as well as on the value and time period of acceleration. By the time the shot is off, a gravitational wave produces a protocrater whose amplitude is smaller than the amplitude of the initial perturbation. Such a protocrater, however, may possess a considerable kinetic energy.

When the beam current is switched off, the plasma spreads without energy delivery. Its acceleration decreases very quickly and may even change sign. In the absence of acceleration, the kinetic energy stored by a protocrater causes Richtmyer–Meshkov instability to arise [3, 4, 8]. The second stage of cratering sets in, during which the protocrater deepens and transforms into a full crater several microns deep. The resulting crater has the same radius and annular structure as the protocrater.

When the gravitational waves propagate and the instability develops, not only the surface but also near-surface layers of the material become disturbed. As a result, the processes on the plasma surface deform the liquid phase–plasma interface. Further solidification of the liquid phase freezes the deformation, giving rise to the relief observed on the surface with the craters. The depth of each crater depends on the kinetic energy stored in the associated protocrater (i.e., on the phase of the gravitational wave when the acceleration drops abruptly) and on the time period from the onset of instability to liquid phase crystallization.

MATHEMATICAL MODEL OF CRATERING

In the general case, one can distinguish three phases in the target irradiated (Fig. 1): the plasma phase, from which the material leaves the surface; the liquid phase, which crystallizes on the target surface upon cooling; and the solid phase. Note that the plasma phase may be absent if the irradiation intensity is low. Gravitational waves arising on the free surface deform areas adjacent to the surface. However, we neglect the disturbance of the liquid–solid interface, assuming that the gravitational waves and the instability occupy a layer of finite depth. Flows associated with the free surface deformation will be described under the assumption that the plasma and liquid phases of the material constitute a unified plasma-like medium [9, 10].

The problem will be solved for the axisymmetric cylindrical geometry (Fig. 1). The axis $0z$ is directed along the normal to the undisturbed free surface. The medium is assumed to be incompressible in the first approximation. Consider a layer of an incompressible liquid bounded by the flat boundary $z = -H$. The liquid moves with a variable acceleration, i.e., is in an external potential field $G(r, z, t) = -g(t)z$. If the inertial forces are directed inward to the medium ($g < 0$), gravitational waves appear on the free surface [1]; otherwise, Rayleigh–Taylor instability develops [2, 3].

Let us write a set of equations that describe the dynamics of the free surface $z = Z(r, t)$, that of the plasma–liquid interface $z = Z_c(r, t)$, and the pressure at the solid–liquid interface ($z = -H$). For a potential liquid flow, we write the incompressibility condition

$$\Delta\varphi(r, z, t) = 0, \tag{1}$$

where $\varphi(r, z, t)$ is the velocity potential; Bernoulli equation,

$$\frac{\partial\varphi(r, Z, t)}{\partial t} + \frac{1}{2}\left[\left(\frac{\partial\varphi}{\partial r}\right)^2_{Z(r,t)} + \left(\frac{\partial\varphi}{\partial z}\right)^2_{Z(r,t)}\right] \tag{2}$$
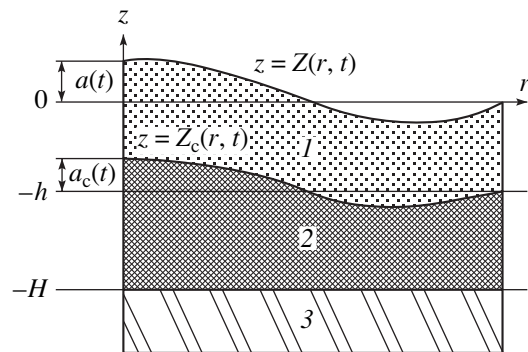$$- g(t)Z(r, t) = 0;$$

Fig. 1. Geometry of the problem: *1*, plasma; *2*, liquid phase; and *3*, elastic medium.

kinematic condition for the free surface,

$$\frac{\partial Z(r, t)}{\partial t} = \left(\frac{\partial \varphi}{\partial z}\right)_{Z(r, t)} - \left(\frac{\partial \varphi}{\partial r}\right)_{Z(r, t)} \frac{\partial Z(r, t)}{\partial r}; \quad (3)$$

and the impermeability condition for the lower boundary,

$$\frac{\partial \varphi}{\partial z}(r, -H, t) = 0. \quad (4)$$

In the general case, nonlinear set (1)–(4) is difficult to solve. It has been shown [8] that this set can be reduced to a closed set of equations for the free surface by eliminating flows in the volume. Unfortunately, one must know strain fields and pressures in the volume of the material when describing cratering.

The linear theory has been demonstrated to be valid in many practical cases. If the velocity and displacement of the liquid are assumed to be small, one can neglect the second-order terms in $Z(r, t)$ and $\nabla\varphi$ in Eqs. (2) and (3) to arrive at a linearized set that is solved by the method of separation of variables.

For a layer in the cylindrical coordinates, an elementary solution to Laplace equation (1) that satisfies condition (4) is

$$\varphi(r, z, t) = C(t)(e^{kz} + e^{-kz - 2kH})J_0(kr), \quad (5)$$

where $J_0(kr)$ is the zero-order Bessel function.

This solution corresponds to the elementary profile of the free boundary $Z(r, t) = a(t)J_0(kr)$. For the amplitude of such a perturbation, the Taylor formula is given by [1–3]

$$\frac{d^2 a(t)}{dt^2} = g(t)k\tanh(kH)a(t). \quad (6)$$

For $g(t) = -g_0$, where g_0 is a positive constant, the perturbation amplitude experiences harmonic oscillations with a frequency $\omega = \sqrt{g_0 k\tanh(kH)}$; that is, gravitational waves set in the system. In this case, the inertial forces tend to stabilize the system, i.e., to return it to the equilibrium (plane) shape. In a traveling gravitational wave, the perturbations propagate with a group velocity [1]

$$V_{\mathrm{g}} = \frac{\partial \omega(k)}{\partial k}$$
$$= \frac{1}{2}\sqrt{\frac{g_0}{k\tanh(kH)}}\left(\tanh(kH) + \frac{kH}{\cosh^2(kH)}\right). \quad (7)$$

If the acceleration of the system (and hence, the inertial forces) vanishes at a time instant $t = t_0$, the perturbation amplitude starts growing with a constant rate da/dt ($t = t_0$), as in the case of Richtmyer–Meshkov instability [3, 4], and the perturbations do not propagate over the surface. Actually, the dependence $g(t)$ is rather complicated, so that it is difficult to establish a strict

line of demarcation between the conditions of gravitational wave propagation and instability conditions. Therefore, in real situations, the solution can be found only numerically.

Integration of the velocity field $\nabla\varphi$ with respect to time yields the displacement of the liquid in the volume. The Z component of the displacement at a depth $z = -h$ defines the crater shape $Z_{\mathrm{c}}(r, t)$ (Fig. 1). If the plasma–liquid interface is perfectly flat at $t = 0$, then

$$Z_{\mathrm{c}}(r, t) = (a(t) - a(0))\frac{e^{-kh} - e^{kh - 2kH}}{1 - e^{-2kH}}J_0(kr). \quad (8)$$

Let us find the pressure field due to the processes on the free surface. The linearized equation of motion in the layer has the form

$$\frac{\partial \nabla\varphi(r, z, t)}{\partial t} = -\frac{\nabla P(r, z, t)}{\rho} + g(t). \quad (9)$$

Let $P(r, z, t) = P_0(z, t) + P_1(r, z, t)$, where $P_0(z, t) = \rho g(t)z$ and $P_1(r, z, t)$ is defined by the free surface deformation. Taking the constant of integration such that $P_0(Z(r, t), t) + P_1(r, Z(r, t), t) = 0$, we find in view of (3)

$$P_1(r, z, t) = -\rho g(t)a(t)\frac{e^{kz} + e^{-kz - 2kH}}{1 + e^{-2kH}}J_0(kr). \quad (10)$$

In the linear range, it is easy to describe the evolution of a perturbation of arbitrary shape. In a circle of radius R, an arbitrary function $Z(r, t)$ is expanded in Bessel functions:

$$Z(r, t) = \sum_{n=1}^{\infty} a_n(t)J_0(k_n r), \quad (11)$$

where $k_n = \mu_n/R$ and μ_n is the nth zero of the Bessel function.

Each of the elementary perturbations $a_n J_0(k_n r)$ will develop independently. Adding the elementary perturbation, we obtain from (8)

$$Z_{\mathrm{c}}(r, t)$$
$$= \sum_{n=1}^{\infty} (a_n(t) - a_n(0))\frac{e^{-k_n h} - e^{k_n h - 2k_n H}}{1 - e^{-2k_n H}}J_0(k_n r). \quad (12)$$

Similarly, from (10), we derive the pressure on the rigid wall:

$$P_1(r, -H, t) = \sum_{n=1}^{\infty} P_n^0 J_0(k_n r),$$
$$P_n^0 = -\rho g(t)a_n(t)\frac{2e^{-k_n H}}{1 + e^{-2k_n H}}. \quad (13)$$

Let us find the applicability domain for the linear theory. In a layer of finite thickness, two criteria of smallness must be fulfilled simultaneously: $a(t)/H \ll 1$

and $a(t)k \ll 1$. Nonlinear effects are essential when either of the left-hand sides approaches unity. It has been found experimentally [6] that the crater depth does not exceed 1.5–2.0 μm. The depth H of the molten surface layer depends on the particle range in the target and the thermal conductivity of the target. For the beam used in [6], $H \geq 3$ μm. The value of H can be obtained from the temperature fields in the material, which are found by numerically solving the equations of continuum mechanics. Thus, the criterion $a(t)/H < 1$ is met for most craters. The second criterion, $a(t)k < 1$, imposes restrictions on the wave number range: $k < 10^6$ m^{-1}. Thus, nonlinear effects are essential only for very deep and very shallow craters.

In the nonlinear range $a(t)/H \sim 1$ or $a(t)k \sim 1$, the solution to Eq. (1) can be represented in the form

$$\varphi(r, z, t) = C(t)(e^{z/L(r, z, t)} + e^{-(z + 2H)/L(r, z, t)})J_0(kr),(14)$$

where L is some effective depth of perturbation. The substitution of (14) into (1) yields an equation for L. From (14) and (5), we find that $L = 1/k$ in the linear range.

To approximately calculate the effective depth of perturbation $L(r, t)$ and the evolution of the plasma surface profile, we used the nonlinear approach described in [8]. The crater shape and the pressure fields in the nonlinear range were evaluated with Eqs. (8) and (10), where k in the exponents was replaced by $1/L(r, t)$ calculated with the nonlinear approach.

To describe the evolution of a perturbation of arbitrary shape in the nonlinear range, k_n in the exponents in (12) and (13) should be replaced by $1/L_n$, where L_n is the effective depth of perturbation for the nth component.

The generalization of the theory [8] for the evolution of an arbitrary perturbation yields $L_n = L(r, t)k_1/k_n$. In the linear range, $L = 1/k_1$ and accordingly $L_n = 1/k_n$.

ACCELERATION OF TARGET SURFACE EXPOSED TO INTENSE CHARGED PARTICLE FLUXES

The interaction between an intense charged particle beam and a target can be described in terms of a set of equations including the kinetic equation for fast particles, equations of continuum mechanics, and an equation of state that provides a stable solution in a wide range of thermodynamic parameters.

The off-diagonal components of the target stress tensor were found within the elastoplastic model [11]. Beyond the solid, the shear modulus was set equal to zero. The diagonal components of the stress tensor were calculated with the wide-range equation of state suggested in [12]. Equations of continuum mechanics were solved with the technique suggested in [13], which yields a stable solution in a very wide range of thermodynamic parameters of a medium. This technique also applies to liquid–vapor equilibrium states, for which approaches involving artificial viscosity are associated with certain difficulties. The essence of this technique is that an analytical solution found for a small volume element and a short time interval is applied for the description of the behavior of the whole system.

For the beams with power densities $<10^{13}$ W/cm^2, the radiative thermal conductivity is much smaller than the electron thermal conductivity and can be neglected [14]. The electron thermal conductivity was calculated within the plasma model of metal [9, 10] over a wide range of material parameters. The heat evolution due to the beam was taken into account by adding the energy release function to the right-hand side of the equation for internal energy.

The energy release function, which equals the dose rate per unit mass of the material, was found by solving the problem of charged particle transfer with the semi-analytic method suggested in [15]. In evaluating the electron transfer, we took into consideration energy loss fluctuations, multiple electron scattering, and the production of secondary electrons. The ion transfer was evaluated in the continuous deceleration approximation with multiple scattering ignored. The energy losses for the ions were calculated according to [16].

The acceleration of the surface irradiated was calculated by the formula

$$g(t) = -\frac{1}{\int\limits_{z_a} \varrho(z, t)dz} \frac{d}{dt}\int\limits_{z_i}^{z_t} \varrho(z, t)v_z(z, t)dz. \quad (15)$$

Here, z_t is the coordinate of the plasma torch surface, z_i is the coordinate of the plasma–solid interface, and z_a is the coordinate of the surface inside the solid that has been reached by the perturbation by the time instant t. Results given below were obtained for a flat target irradiated by the flux moving in the positive direction along the $0z$ axis. Therefore, the target plasma spreads in the negative z direction.

The properties of a plasma generated by intense charged particle beams striking a target depend primarily on the energy release function $D(z)$ [17]. Our numerical experiment implies that the width of the energy release region is specified by the particle range in the target. Under electron irradiation, this region is two orders of magnitude larger than that under irradiation by protons of the same energy. Since the range of carbon ions in a material is much smaller than that of protons, the presence of ions in a carbon beam increases the energy release in surface layers of the target. A rise in the particle energy increases largely the volume irradiated but changes insignificantly the energy being released.

Consider first the dynamics of the surface accelerated by a powerful electron beam with parameters used
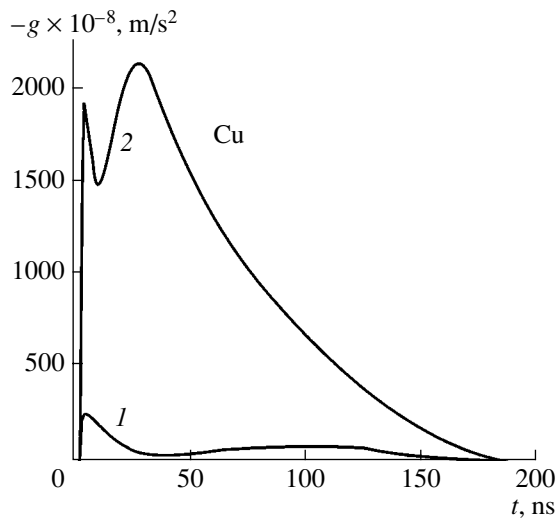
Fig. 2. Time variation of the target surface acceleration under electron acceleration. $j_b = 10$ (*1*) and 100 kA/cm^2 (*2*). $E_e = 0.3$ MeV, $\tau = 100$ ns.

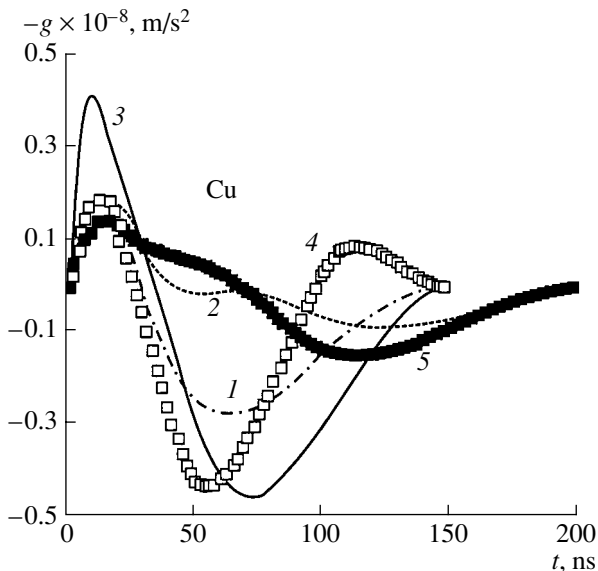**Fig. 3.** The same as in Fig. 2 for irradiation by low-energy ion beams containing protons (P) and carbon ions (C). $E_b = 0.5$ MeV; $j_b = 100$ (*1, 2, 4, 5*) and 200 A/cm^2 (*3*). $\tau = 50$ (*1, 3, 4*) and 100 ns (*2, 5*). (*1–3*) P and (*4, 5*) 0.7C + 0.3P.

in experiments carried out in [18]. Figure 2 shows the time variation of the surface acceleration. The surface starts to move noticeably within several nanoseconds after the beginning of irradiation. Then, within 10–15 ns, the acceleration declines because the energy delivered is spent on target melting and evaporation. The further increase in the acceleration is associated with a pressure rise in the plasma due to heating by the beam.

For the target irradiated by the powerful ion beam, the surface dynamics is more complex (Fig. 3). At a low

ion current density ($\sim 10^6$ A/m^2) and a shot duration of 50 ns, the acceleration of the target surface changes to its fast deceleration when the irradiation ends (curves *1, 3, 4*). This is explained by evaporation energy losses when the medium expands (the volume irradiated is in liquid–vapor equilibrium). As the shot duration grows, the negative acceleration diminishes considerably. Note that when the target is irradiated by a low-energy electron shot of nanosecond duration, the pulsed acceleration response is similar to curve *1* in Fig. 3 but has a lesser amplitude and accordingly is longer. For example, when a copper target is irradiated by a 40-keV electron shot with a current density of 5×10^6 A/m^2 and a duration of 10^{-6} s, the maximal acceleration reaches $\approx 5 \times 10^6$ m/s^2. In this case, the durations of the accelerating and decelerating phases are roughly the same, $\approx 4 \times 10^{-7}$ s.

For low irradiation intensities, the distributions of the density and temperature of the spreading plasma show that the plasma density remains high ($\approx 0.7\rho_0$, where ρ_0 is the initial density of the medium), and the temperature drops to the crystallization temperature within 100 ns after the beginning of irradiation. Such a high rate of temperature drop is related primarily to the high electron thermal conductivity: a ≈ 18-μm-thick layer is heated for the time interval mentioned above.

A rise in the irradiation intensity or an increase in the energy delivery by changing the beam composition causes intense sputtering of the irradiated part of the target. The acceleration of the surface grows considerably (Fig. 4), and the decelerating phase is absent. The density of the spreading plasma decreases by several orders of magnitude compared with the initial value. The temperature of the medium also rapidly drops with time. In this case, this is due to both the high thermal conductivity and internal-to-kinetic energy conversion.

RULES OF CRATERING: NUMERICAL EXPERIMENT

To elucidate general rules of cratering, we performed calculations for a model acceleration response: $g(t) = -g_0$ for $t \leq \tau_g$ and $g(t) = 0$ for $\tau_g < t \leq \tau_g + \tau_c$. Here, τ_g is the duration of the pulsed acceleration response (which can be found from Figs. 2–4), and τ_c is the time to liquid phase crystallization after the shot has been switched off (which can be estimated from the variation of the temperature field).

In simulating the effect of N shots, the acceleration cycle was repeated N times. The profile of the liquid boundary $Z_c(r, t = \tau_g + \tau_c)$ obtained in the preceding cycle was taken as the initial perturbation of the plasma surface $Z(r, t = 0)$ in the next cycle. With such an approach, the smoothening of the crater due to the removal of the material evaporated is taken into account. We used the weakly nonlinear approximation; that is, the plasma surface profile was determined in
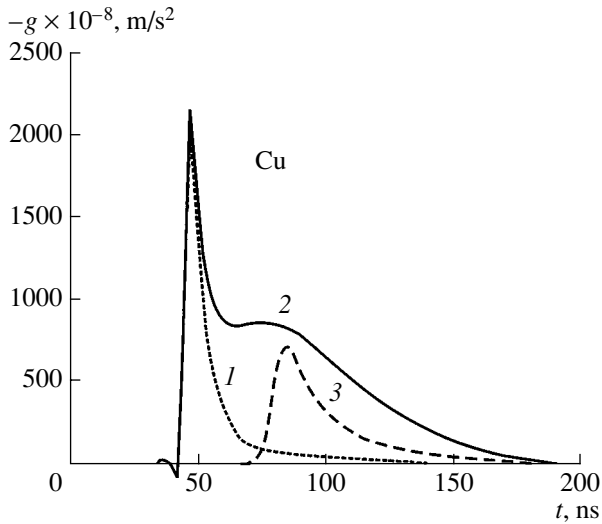
Fig. 4. The same as in Figs. 2 and 3 for irradiation by high-energy ion beams containing protons and carbon ions. $E_b = 0.5$ MeV. $j_b = 200$ A/cm^2 (3). $\tau = 50$ (1) and 100 ns (2, 3). (1, 2) 0.7C + 0.3P and (3) P.

terms of the nonlinear theory [8], while the liquid boundary profile was calculated by formula (12). Figure 5a demonstrates various initial perturbations of the surface; Fig. 5b, resulting craters after $N = 6$ shots (curve 1 refers to smoothed microprotrusion; curve 3, to rings; and curves 2, 4, and 5 refer to steps of various radius). The irradiation conditions are $g = 3 \times 10^{10}$ m/s^2, $\tau_g = 50$ ns, $\tau_c = 100$ ns, $h = 2$ μm, and $H = 4$ μm. At the maximum of the pulsed acceleration response (≈ 60 ns), the plasma is still near the surface, which justifies the order of magnitude for h and H. The mechanism considered is seen to provide the evolution of a small localized perturbation into a crater whose parameters and shape (structure) meet those observed in experiments.

Of special note is the weak dependence of the crater shape (structure) on the shape and size of an initial perturbation. Initial perturbations with different shapes and sizes transformed into virtually identical annular structures, which are specified by the narrow spectrum in the expansion in Bessel functions. For various initial perturbations, the typical value of k is $k = (3.3–3.8) \times 10^5$ m^{-1}. Since $kH > 1$, expression (7) for the group velocity of the wave is simplified. In this case, the crater radius can be estimated as

$$R \approx R_0 + 0.5\sqrt{g_0/k}\,\tau_g N, \qquad (16)$$

where R_0 is the radius of the initial perturbation.

In (16), it is taken into account that the perturbation propagates along the radius only as a gravitational wave.

If the second term in (16) much exceeds the first one, the crater radius depends on irradiation conditions, i.e., on the duration and amplitude of the pulsed acceleration response. Otherwise, the crater size depends on
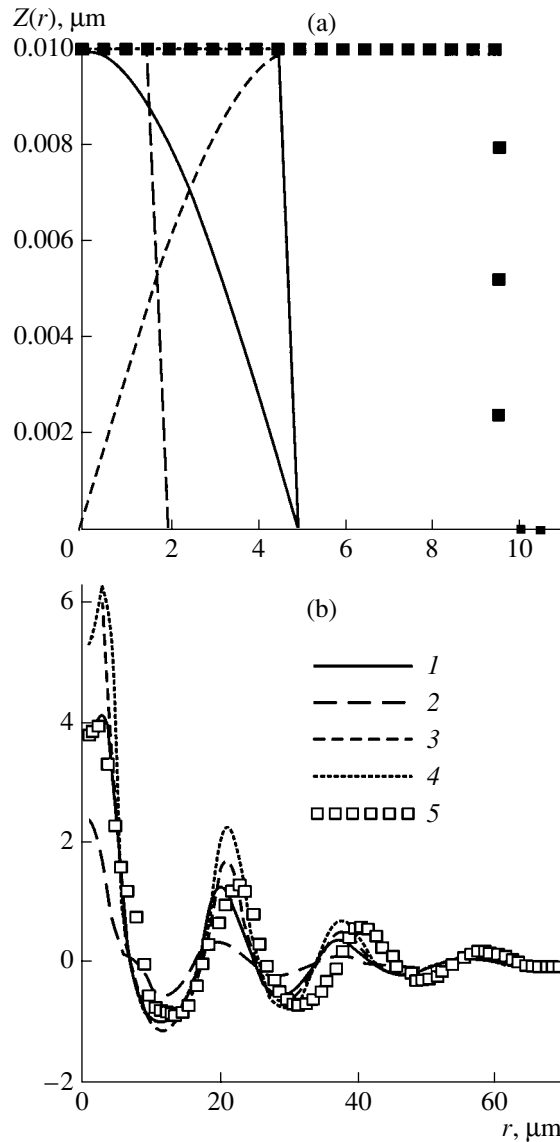

Fig. 5. Initial perturbations and associated craters formed under the action of model acceleration.

the initial perturbation. In our case, the radius of the perturbed region grows to 40–45 μm through the propagation of gravitational waves; therefore, the crater radius depends on the initial perturbation only slightly. Thus, the annular structure and the crater radius in this case depend on the irradiation conditions and conditions at the plasma surface.

Figure 5 depicts the crater shapes for the case where the initial perturbations are protrusions. The craters formed also have a protrusion at their center. If the initial perturbation is a recess, the associated craters will also have recesses at the center with the annular structure retained.

It should be noted that the cratering mechanism suggested in this work includes also the mechanism of filtering of the initial perturbation spectrum. Indeed, for-
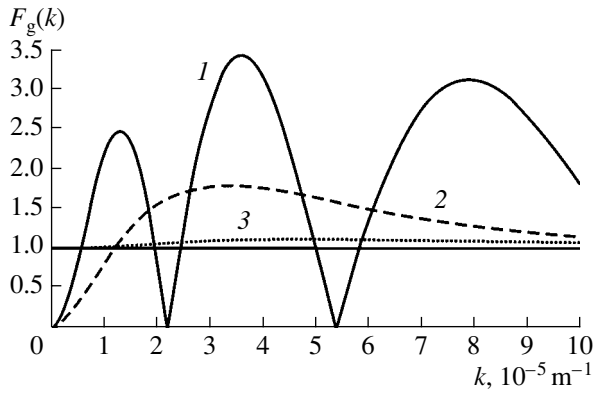
Fig. 6. Crater growth function vs. *k*.

mulas (7) and (12) imply that short-wave perturbations make a smaller contribution to cratering than long-wave ones, because the former propagate with a lower velocity and decay with depth faster. However, according to the Taylor formula [3], long-wave perturbations have a greater time to onset of instability. Thus, for a specific beam–target system, there is a certain range of wave numbers that specifies the crater profile. Quantitatively, this fact can be described with the crater growth function $F_g(k)$. It is defined as the ratio of the crater amplitude (i.e., the amplitude of the solidified boundary of the liquid) after the action of one shot to the initial amplitude of an elementary perturbation for a given *k*. Within our model of acceleration, the crater growth function is found analytically by integrating Taylor formula (6) in view of (8):

$$F_g(k) = \frac{e^{-kh} - e^{-2kH + kh}}{1 - e^{-2kH}}$$

$$\times [\cos(\xi_g) + \tau_k\sqrt{g_0 k \tanh(kH)}\sin(\xi_g) - 1], \quad (17)$$

$$\xi_g = \sqrt{g_0 k \tanh(kH)}\tau_g,$$

where ξ_g is the phase of the gravitational wave at the instant the pulse acceleration response is completed.

The crater growth function is uniquely specified by irradiation conditions. For our model, the absolute value of the growth function vs. *k* (for $k \le 10^6$ m^{-1}) is shown in Fig. 6 (curves *1–3*). Under irradiation, only those harmonics of the initial perturbation spectrum for which $|F_g(k)| > 1$ grow. Harmonics with the largest $|F_g(k)|$ (i.e., for $k \approx 3.5 \times 10^5$ m^{-1} in our case) grow most rapidly. The filtering efficiency of one shot is low: local peaks (at $k \approx 1.5 \times 10^5$ m^{-1} and $k \approx 8 \times 10^5$ m^{-1}) with roughly equal heights are observed; however, the dependence $[F_g(k)]^N$ for $N > 1$ has a distinct peak in a rather narrow range of wave numbers. Even at $N = 6$, the growth function peaks sharply at $k \approx (3–4) \times 10^5$ m^{-1}. Because of this, the annular structure of the crater

depends on the initial shape of the perturbation insignificantly (Fig. 6, curves *2* and *3*).

Now consider the crater growth under the action of an actual pulsed acceleration response. The function $g(t)$ is rather complicated; therefore, the growth function is found by integrating Taylor function (6) for the plasma surface profile jointly with the equation

$$\frac{da_c(t)}{dt} = \frac{da(t)}{dt}\frac{e^{-kh} - e^{kh - 2kH}}{1 - e^{-2kH}}, \quad (18)$$

$$a_c(0) = 0,$$

for the boundary of the liquid. Here, a_c is the amplitude of the perturbation at the liquid–plasma interface. Equation (18) follows from (8). Because of the plasma spread, the depths *h* and *H* continuously vary. Within the model of incompressible liquid, these variations for the liquid–plasma system are impossible to consistently take into account. However, one can refine the model by formally substituting into Eqs. (6) and (18) the time-varying functions $H(t)$ and $h(t)$ for the constants. The model refined includes the compressibility of the medium as a whole in the course of the basic process, which causes gravitational waves and instability.

As has been shown above, the dependence of the irradiation-induced acceleration of the medium on the energy delivered has a strongly nonlinear character. We consider two irradiation regimes with greatly differing functions $g(t)$ (Figs. 3, 4). Figure 7 shows the absolute values of the growth function for two actual acceleration regimes. Curve *2* reflects the case of an ion beam with a large energy (the associated function $g(t)$ is depicted by curve *1* in Fig. 4), and curve *3* refers to a small-energy ion beam (the associated function $g(t)$ is depicted by curve *3* in Fig. 3). The total time from the beginning of the shot action to crystallization is $\tau_g + \tau_c = 150$ ns. In evaluating these crater growth functions, the above time dependences were taken into account: $h(t)$ was derived from one-dimensional calculations (the coordinate of the front surface of the target) and $H(t)$ was set equal to $h(t) + h_{cons}$, where h_{cons} is the constant thickness of the liquid layer.

It follows from the plots that conditions for crater growth exist in both cases. For the low-energy beam, crater smoothening due to the plasma spread is absent. This partially compensates for a small acceleration. However, the crater growth functions are small: only relatively large initial perturbations (0.1–0.3 μm in size) can transform into a crater. For this reason, low-energy beams generate a smaller number of craters than high-energy ones (in the latter case, a perturbation even ≈0.01 μm deep may give rise to a crater).

For the high-energy beam, the peak acceleration is one order of magnitude higher than g_0 in the above model case. However, the sharp but not instantaneous drop of the acceleration and the inclusion of the time dependence of the depths result in a much smaller crater growth function.

For the high-energy beam, the absolute value of the growth function is maximal at $k \approx 3.3 \times 10^5$ m^{-1}. However, this peak is not so sharp as for the model acceleration. Therefore, the sensitivity to an initial perturbation in this case is higher. For the low-energy beam, the crater growth function almost saturates at $k \geq 3 \times 10^5$ m^{-1} and the crater shape and radius are completely defined by an initial perturbation. Such irradiation conditions deepen the crater nucleus without regularly structuring it.

If the growth function is known, the final expression for the crater profile after the action of N shots in the linear range is given by

$$Z(r, t = N(\tau + \tau_c)) = \sum_{n=1}^{\infty} a_n(0)(F_g(k_n))^N J_0(k_n r), \quad (19)$$

where $a_n(0)$ is the initial perturbation spectrum.

Thus, we found the profiles of the actual craters (Fig. 7) that were formed under the action of the high-energy shots ($N = 10$). The initial perturbations had the form of recesses 0.01 μm deep with a radius $R_0 = 10$ (curve 1), 30 (curve 2), and 1 μm (curve 3).

All the craters have nearly the same parameters of the annular structure, $k = (3.0–3.5) \times 10^5$ m^{-1} (the values close to the maximum of the absolute value of the growth function). Vertical lines indicate the position of the rings of the crater described in [5], where the irradiation conditions were the same as in this work but the targets were made of molybdenum alloys. Nevertheless, our calculations are seen to be in good agreement with the experimental data.

Thus, under the irradiation conditions used in [5], the crater radius depends strongly on that of the initial perturbation. The increase in the crater radius because of the propagation of gravitational waves is as small as 10–30 μm. Large craters ($R \geq 10$ μm) with a developed annular structure may arise only from sufficiently large initial perturbations; for this reason, a fraction of large craters is low if the surface was prepolished. As for local perturbations with $R_0 \leq 1$ μm, they give rise to a large number of small craters with $R \leq 10$ μm.

STRESS FIELDS IN A SOLID DUE TO CRATERING

According to formula (13), gravitational waves originating on the plasma surface exert pressure on the solid surface. The higher the instantaneous acceleration and the crater amplitude, the higher the pressure. The pressure deforms near-surface layers of the metal and produces a stress field. This field is the subject of the subsequent discussion.

We assume that the deformations are elastic and quasi-static. Since their range is typically $\approx k^{-1} \sim 3$ μm and the velocity of sound is $\approx 3.3 \times 10^3$ m/s, the relaxation times are on the order of 10^{-9} s. The geometry of the problem is the same as in Fig. 1 with the only dif-
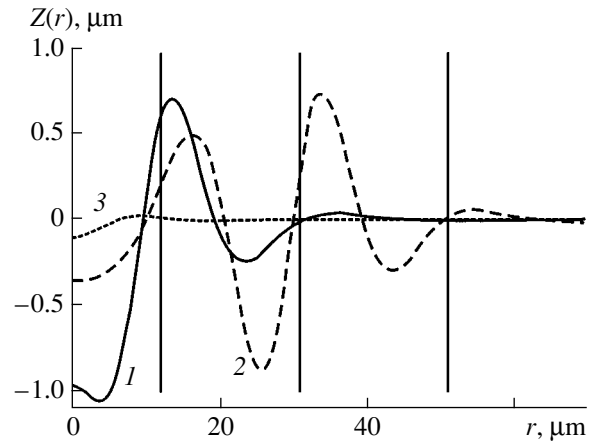

Fig. 7. Crater profiles under actual acceleration conditions.

ference that the origin is shifted downward by H so that $z = 0$ corresponds to the solid surface. Balance equations for elastic medium in the cylindrical geometry have the form

$$c_l^2 \left(\frac{\partial^2 u_r}{\partial r^2} + \frac{1}{r}\frac{\partial u_r}{\partial r} + \frac{\partial^2 u_r}{\partial z^2} - \frac{u_r}{r^2} \right)$$
$$- (c_l^2 - c_{tr}^2)\left(\frac{\partial^2 u_r}{\partial z^2} - \frac{\partial^2 u_z}{\partial r \partial z} \right) = 0,$$
$$c_{tr}^2 \left(\frac{\partial^2 u_z}{\partial r^2} + \frac{1}{r}\frac{\partial u_z}{\partial r} + \frac{\partial^2 u_z}{\partial z^2} \right) \quad (20)$$
$$+ (c_l^2 - c_{tr}^2)\left(\frac{\partial^2 u_r}{\partial r \partial z} + \frac{1}{r}\frac{\partial u_r}{\partial z} - \frac{\partial^2 u_z}{\partial z^2} \right) = 0,$$

where c_l and c_{tr} are the longitudinal and transverse velocities of sound, respectively, and $u_r(r, z)$ and $u_z(r, z)$ are the displacement vector components.

Boundary conditions for the stresses are

$$\sigma_{zz}(r, z = 0) = -\sum_{n=1}^{\infty} P_n^0 J_0(k_n r),$$
$$\sigma_{rz}(r, z = 0) = 0, \quad \sigma_{zz}(r, z = -\infty) = 0, \quad (21)$$
$$\sigma_{rz}(r, z = -\infty) = 0.$$

Jointly solving (20) and (21), we find the displacement field:

$$u_r(r, z) = -\sum_{n=1}^{\infty} \frac{P_n^0}{2k_n \rho c_{tr}^2} e^{k_n z}((1 - 2\sigma) + k_n z)J_1(k_n r),$$
$$(22)$$
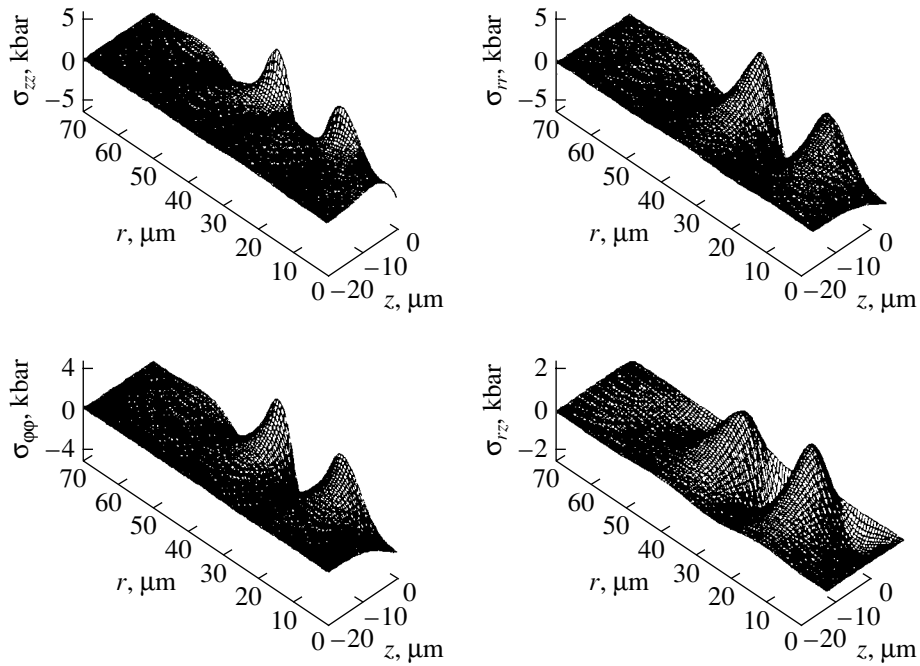$$u_z(r, z) = -\sum_{n=1}^{\infty} \frac{P_n^0}{2k_n \rho c_{tr}^2} e^{k_n z}((2 - 2\sigma) - k_n z)J_0(k_n r),$$

Fig. 8. Stress fields under the crater in the solid part of the target.

where σ is Poisson's ratio and $J_1(k_n r)$ is the first-order Bessel function.

The components of the desired stress field are

$$\sigma_{zz}(r, z) = -\sum_{n=1}^{\infty} P_n^0 e^{k_n z}(1 - k_n z)J_0(k_n r),$$

$$\sigma_{rr}(r, z) = -\sum_{n=1}^{\infty} P_n^0 \left[e^{k_n z}(1 + k_n z)J_0(k_n r) \right.$$

$$\left. - e^{k_n z}((1 - 2\sigma) + k_n z)\frac{J_1(k_n r)}{k_n r} \right],$$

$$\sigma_{\varphi\varphi}(r, z) = -\sum_{n=1}^{\infty} P_n^0 \left[e^{k_n z}2\sigma J_0(k_n r) \right.$$

$$\left. + e^{k_n z}((1 - 2\sigma) + k_n z)\frac{J_1(k_n r)}{k_n r} \right],$$

$$\sigma_{rz}(r, z) = -\sum_{n=1}^{\infty} P_n^0 k_n z e^{k_n z} J_1(k_n r).$$

$$(23)$$

Taking into account the time dependence of the perturbation spectrum $P_n^0(t)$, one can trace the stress dynamics in the quasi-stationary approximation. Figure 8 shows the stress fields for the peak acceleration after the tenth high-energy shot (the situation corresponds to curve *2* in Fig. 6). It is seen that the stresses are localized by depth and by radius (they concentrate in a thin surface layer about 10 μm thick and under the crater, respectively). The stress amplitude is one order of magnitude higher than the yield point Σ (for copper, $\Sigma = 6.85 \times 10^7$ Pa). However, this value is the highest throughout the irradiation time.

The high stresses in the metal under the crater explain an experimentally observed high concentration of structure defects near craters.

As follows from (13), when the acceleration disappears, so does the pressure on the surface and, hence, the quasi-stationary deformations. The energy of deformation is converted to standing cylindrical Rayleigh waves. They are also localized near the surface. Eventually, the total energy is accumulated in the surface layer and is spent on the formation of structure defects.

CONCLUSION

We studied the physical processes responsible for cratering on a solid surface exposed to intense charged particle beams (shots). Physical and mathematical cratering models were constructed. The reason for cratering is gravitational waves and the instability of the plasma torch free surface. The crater sizes and shape found in the numerical experiment are in good agreement with experimental findings. During cratering, the stress fields arise under craters, which are responsible for structural changes in the target.

REFERENCES

1. L. D. Landau and E. M. Lifshitz, *Course of Theoretical Physics*, Vol. 6: *Fluid Mechanics* (Nauka, Moscow, 1988; Pergamon, New York, 1987).

2. S. Chandrasekhar, *Hydrodynamic and Hydromagnetic Stability* (Clarendon, Oxford, 1961).

3. R. D. Richtmyer, Commun. Pure Appl. Math. **72**, 297 (1960).

4. E. E. Meshkov, Izv. Akad. Nauk SSSR, Mekh. Zhidk. Gaza **5**, 151 (1969).

5. A. D. Korotaev, S. V. Ovchinnikov, Yu. I. Pochivalov, *et al.*, Surf. Coat. Technol. **105**, 84 (1998).

6. A. D. Korotaev, A. N. Tyumentsev, M. V. Tretjak, *et al.*, Structure, Phase Transformations and Diffusion **89**, 54 (2000).

7. S. A. Chistyakov, S. V. Khalikov, and A. P. Yalovets, Zh. Tekh. Fiz. **63** (1), 31 (1993) [Tech. Phys. **38**, 5 (1993)].

8. N. B. Volkov, A. E. Maĭer, and A. P. Yalovets, Pis'ma Zh. Tekh. Fiz. **27** (1), 47 (2001) [Tech. Phys. Lett. **27**, 20 (2001)].

9. N. B. Volkov and A. Z. Nemirovsky, J. Phys. D **24**, 693 (1991).

10. N. B. Volkov, Doctoral Dissertation (Inst. of Electrophysics, Ural Division, Russian Academy of Sciences, Yekaterinburg, 1999).

11. M. L. Wilkins, in *Fundamental Methods in Hydrodynamics*, Ed. by B. Alder, S. Fernbach, and M. Rotenberg (New York, 1964; Mir, Moscow, 1967), Vol. 3.

12. S. N. Kolgatin and A. V. Khachatur'yants, Teplofiz. Vys. Temp. **20** (3), 90 (1982).

13. A. P. Yalovets, Prikl. Mekh. Tekh. Fiz., No. 1, 151 (1997).

14. B. A. Demidov, M. V. Ivkin, V. V. Obukhov, and Yu. F. Timoshchuk, Zh. Tekh. Fiz. **50**, 2209 (1980) [Sov. Phys. Tech. Phys. **25**, 1288 (1980)].

15. V. V. Val'chuk, S. V. Khalikov, and A. P. Yalovets, Mat. Model. **4** (10), 111 (1992).

16. F. Zeigler, *Stopping Cross Section for Energetic Ions in All Elements* (Pergamon, New York, 1977).

17. N. B. Volkov and A. P. Yalovets, in *Proceedings of 1st Congress on Radiation Physics, High Current Electronics, and Modification of Materials*, Ed. by G. Mesyats, S. Bugaev, and A. Ryabchikov (TPU, HCEI, Tomsk, 2000), Vol. 3, p. 90.

18. A. F. Akkerman, A. V. Bushman, B. A. Demidov, *et al.*, Zh. Éksp. Teor. Fiz. **89**, 852 (1985) [Sov. Phys. JETP **62**, 489 (1985)].

Translated by V. Isaakyan

Chemically Produced ZnS : Cu Films: Structure, Properties, and Mechanism of Electroluminescence

V. S. Khomchenko, L. V. Zavyalova, N. N. Roshchina, G. S. Svechnikov, I. V. Prokopenko, V. E. Rodionov, P. M. Lytvyn, O. S. Lytvyn, and Yu. A. Tsyrkunov

Institute of Semiconductor Physics, National Academy of Sciences of Ukraine, Kiev, 03028 Ukraine
Received September 6, 2001; in final form, February 26, 2002

Abstract—ZnS : Cu films are prepared with a chemical nonvacuum technique by joint pyrolysis of zinc and copper dithiocarbamates at a substrate temperature between 260 and 300°C. It is shown that the films have the hexagonal lattice and are polycrystalline, with grains oriented mostly in the $\langle 0001 \rangle$ direction. The luminance–voltage characteristics, charge–voltage characteristics, brightness waveforms, electroluminescence spectrum, and degradation characteristics of the phosphors are investigated. A mechanism of electroluminescence and the degradation type are discussed. It is concluded that the MOCVD technique is promising for fabricating ZnS : Cu electroluminescent films. © *2002 MAIK "Nauka/Interperiodica".*

INTRODUCTION

ZnS : Cu is a widely used blue- and green-emitting powder electroluminescent (EL) phosphor. At the same time, thin-film EL phosphors of this composition prepared by conventional vacuum technology have been until recently regarded as inappropriate for practical use because of their fast degradation. The poor performance of these thin-film EL structures is due to copper ion migration and the recombination mechanism of emission. However, copper-doped high-luminance stable SrS and ZnS phosphors prepared by advanced vacuum technology [1, 2] have proved to be promising phosphor materials. Importantly, in these structures, electroluminescence is excited by the impact mechanism, which is the most efficient for films. This has given impetus to the search for new thin-film materials and methods.

Over the past years, MOCVD has become common in preparing EL phosphors based on ZnS : Mn films [3]. The most simple yet efficient version of this technique is chelate CVD [4–6]. From these precursors, many thin-film components and devices, including high-luminance EL ZnS : Mn-based phosphors, have been made [5, 6]. The films were deposited on a substrate kept at a temperature $T_s = 240$–300°C by the joint pyrolysis of zinc and manganese dithiocarbamates taken in a given proportion.

In [7], we reported our pioneering experiments on depositing blue-, green-, and yellow-emitting ZnS : Cu films. They were obtained by the same technique as the ZnS : Mn films, that is, by the joint pyrolysis of zinc and copper dithiocarbamates. The films thus produced have a more ordered structure and smoother surface than those prepared by the conventional electron-beam evaporation (EBE) *in vacuo*. The good results obtained in [7] have stimulated us to further refine the chemical

technique and study the properties of the films under various preparation conditions. The deposition of ZnS : Cu films from the dithiocarbamates is very different from their fabrication by the EBE method: copper is introduced without a coactivator and a large amount of free sulfur and oxygen is present. Because of this, the probability of the Cu_2S (or ZnO) phase and Cu_2S–ZnS heterojunctions occurring rises. In this case, the undesired injection mechanism of electroluminescence cannot be excluded.

In the light of the aforesaid, the elucidation of an electroluminescence mechanism in these films seems to be topical, because this mechanism specifies the quality of ZnS : Cu films and their suitability for practical use. In this work, we study the structure, as well as the EL and electrical performance, of ZnS : Cu films prepared by the joint pyrolysis of zinc and copper dithiocarbamates in order to establish an electroluminescence mechanism and estimate their promise for EL phosphors.

EXPERIMENTAL

The objects of the investigation were metal–insulator–semiconductor–insulator–metal (MISIM) structures on glass and ceramic substrates. These phosphor structures had various materials of their auxiliary layers and various directions of emission extraction. In the direct-type structures (glass substrate), the emission was extracted through the substrate; in the inverse-type ones (ceramic substrate), through the upper transparent electrode. Samples were multilayer thin-film EL structures, which are usually used in EL studies. Here, the emitting ZnS : Cu layer is sandwiched in two electrodes, one of which is transparent and serves to extract the emission. In the direct-type structure, the electrodes were made of In_2O_3 and Al, while the insulators were

SiO_2, Al_2O_3, and $SiO_2 + Al_2O_3$ films. The total thickness of the insulators was roughly equal to the thickness of the EL layer (≈ 0.5 μm). The conducting and insulating films were deposited by EBE or thermal evaporation *in vacuo*. In the inverse-type structures, one of the insulators was a thick (≈ 40 μm) layer of $BaTiO_3$ ferroelectric ceramic applied by the special technique [8] on the metallic electrodes of the ceramic substrate. A semi-transparent gold film was used as the second electrode.

The ZnS : Cu films were obtained by spraying a 0.1 M pyridine solution of zinc and copper dithiocarbamates onto substrates kept at 260–300°C as in [7]. This method provides relatively high film growth rates, 40–120 nm/min, and allows for the introduction of dopants in any proportions immediately during the growth.

The crystal structure of the films was studied by X-ray diffraction analysis with a DRON-3M diffractometer ($CuK\alpha$ radiation). The surface morphology of the films was examined with a NanoScope IIIa atomic force microscope (AFM) (Digital Instruments). Electroluminescence was excited by a sine-wave voltage of frequency 2 kHz. The spectra were recorded photoelectrically by means of a KSVU spectral complex. The emission intensity was measured with an FP4-BPU photometer. Charge–voltage curves were taken with the Sawyer–Tower scheme [3].

RESULTS

Figure 1 shows a typical diffraction pattern for ZnS : Cu films prepared from zinc and copper dithiocarbamates. The films are polycrystalline and have the hexagonal lattice. Diffraction maxima indicating the presence of the ZnS cubic phase, Cu_2S phase, or ZnO phase were absent.

The surface morphology was studied for the ZnS : Cu films simultaneously deposited on the glass and ceramic substrates. The grain size distribution for the films of both types is illustrated in Fig. 2. The insets in Fig. 2 are AFM images of the low-temperature (260°C) ZnS : Cu films. Different feature heights correspond to different colors; the highest points, 200 nm over the surface, are colored white.

The photoluminescence and EL spectra were taken from the ZnS : Cu films deposited on the different substrates under various process conditions.

The luminance (B)–voltage dependences for the EL structures on the glass and ceramic are depicted in Fig. 3 (curves *1*, *2*). Both have a threshold, a region of steep rise, and a saturation region. The variations of the luminance–voltage curves with operating time for the ZnS : Cu films on the glass substrate are shown in Fig. 3 (curves *2–4*). The exciting voltage frequency is 2 kHz, i.e., more than 30 times higher than the usual operating frequency (60 Hz). The heavy-duty operating condi-
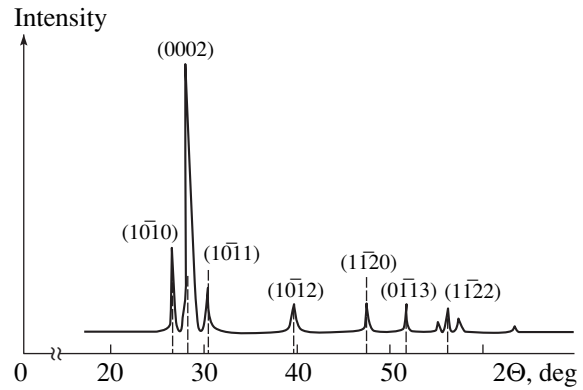

Fig. 1. Diffraction patterns for ZnS : Cu films on glass at $T_s = 280$°C (continuous line) and hexagonal ZnS (dashed line).
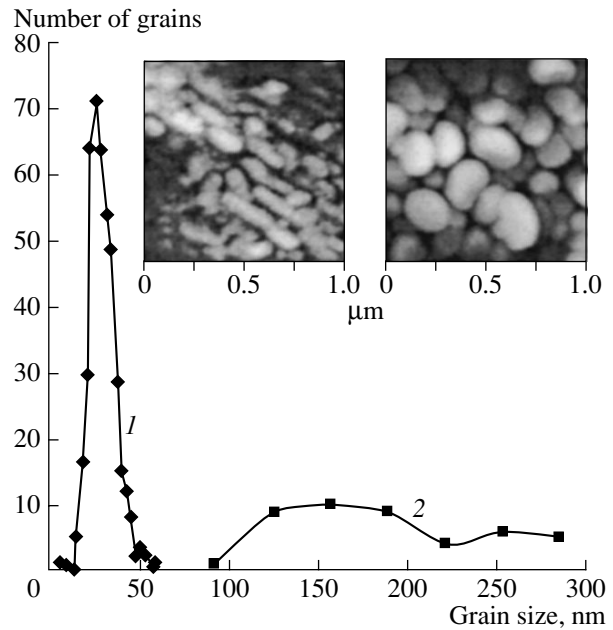

Fig. 2. Grain size distribution for ZnS : Cu films on (*1*) $BaTiO_3$ and (*2*) glass and AFM images of the film surface.

tions, which exploit the fact that the degradation rate increases with frequency superlinearly [9], were taken to minimize the effect of environmental moisture on the unpackaged samples by cutting the exposure time.

The charge–voltage curves $Q(U)$, or the dependence of the carrier concentration on the exciting voltage, are shown in Figs. 4a and 4b for both EL structures for two cases: before the onset of electroluminescence (voltage U_0, $B = 0$) and under the operating conditions.

Brightness waveforms for the ZnS : Cu EL structures excited by a sine-wave voltage U are depicted in Fig. 5.

DISCUSSION

From the X-ray diffraction spectra taken from the films, it follows that they are polycrystalline and have the hexagonal lattice. In the diffraction patterns for both the ZnS : Cu films and polycrystalline ZnS powder, the ratio between the (0002) reflection intensity and that of other reflections is very high. This means that grains in the films are oriented mostly in the ⟨0001⟩ direction. Such a crystal structure of the ZnS films prepared by the method suggested was also observed in [10–12], where the process conditions were similar. Additional phases, such as Cu_2S or ZnO, were not detected. Therefore, the films can be assumed to be homogeneous in composition, so that electric field nonuniformities, causing high-rate degradation of the EL properties, are absent. One could thus expect that degradation processes in these homogeneous films are slow.

The surface morphologies of the films deposited on the glass and ceramic substrates (Fig. 2) at 260 and 300°C are different. Unlike those obtained at T_s = 300°C [7], the lower temperature (260°C) films consist of unfaceted grains loose against each other. The AFM images of the films deposited on the glass and $BaTiO_3$ show differences in their structure.

The surface of the ZnS : Cu films on the ceramic contains nanocrystallites about 25 nm in size, which form conglomerates of size ranging from 70 to 400 nm (hereafter, by grain size we mean that of the major diameter of an ellipse closely approximating the grain shape). It was found that the number and shape of the conglomerates to a great extent depend on the surface relief of the $BaTiO_3$ substrate, which consists of microcrystallites 7–9 μm in size. On the nearly horizontal surface of individual ceramic crystallites, nanocrystallites of ZnS : Cu films form small symmetric conglomerates about 70 nm in size. On the slopes of ceramic grains, ZnS nanocrystallites produce conglomerates

extended across the slope. This is apparently due to the creeping-down of precursor molecules (and hence the film) along the slopes of substrate (ceramic) grains. The surface of the ZnS film covering the flat surface of a substrate grain has a roughness R_q = 8–13 nm, which was estimated by the well-known formula [13].

The polycrystalline surface of the ZnS : Cu films applied on the glass substrate consists of 80- to 300-nm grains and is fairly uniform. The grain size distribution shows that 150 and 250 nm are the most typical sizes. The surface roughness here is roughly twice as large as that of the films on the ceramic (23 nm on average). This difference in surface morphology stems from different thermodynamic parameters on the surfaces of the substrates. Accordingly, mass transfer and growth processes in the plane of a growing film proceed in a different manner.

The photoluminescence and EL spectra also depend strongly on the type of the substrate (for the same activator concentrations and substrate temperatures). The yellow band (λ_{max} = 570 nm) dominates for the ceramic substrate, while the green band (λ_{max} = 520 nm) prevails for the glass substrate. The presence of blue (420, 445, and 485 nm), green (520 nm), and yellow (570 nm) emission bands specifies the nearly white (yellowish green) color of the emission from the thin-film EL structures.

The structure and emission properties of the EL films correlate. The ZnS : Cu films consisting of fine crystallites are largely yellow-emitting phosphors. In those with more perfect coarser grains, conditions for higher energy green and blue emissions are provided. The different ratios of the band intensities in the emission spectra may also be associated with the difference in crystal structure of the films deposited on the ceramic and glass. The difference in grain size and shape for the films deposited on the different substrates results in different conditions for copper incorporation, the appearance of EL centers (structural defects) that may neutralize the charge of a copper ion (Cu^+), and various excitation conditions.

The shape of the luminance–voltage curves in Fig. 3 (the presence of a threshold voltage, as well as sharply rising and saturation regions) is typical of conventional EL structures with the impact excitation mechanism [3]. For our structures on the ceramic and glass, the threshold voltage is 20 and 110 V, respectively. Their electric strength margins, which characterize the stability of the EL structures and show up in the extent of the maximal brightness region, also differ. As follows from curves *1* and *2*, the EL films on the ceramic, unlike those on the glass, have a large electric strength margin (≈150 V). For the as-prepared structures on the glass, the brightness saturation region is absent. It appears only after the structures have been trained under heavy-duty conditions for several hours (curves *3*, *4*).

From curves *3* and *4* (Fig. 3), it is also seen that the luminance–voltage curves shift toward higher voltages
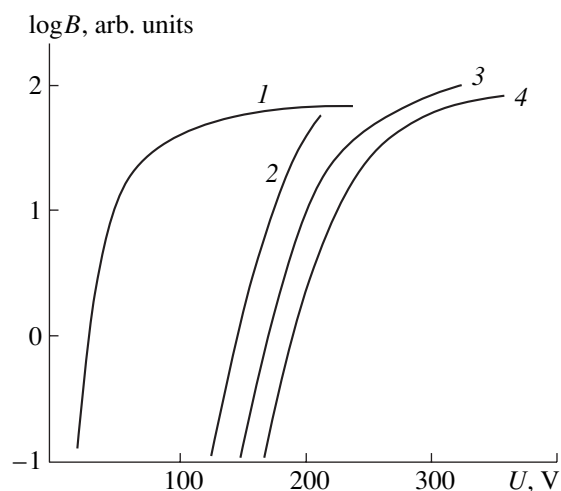
Fig. 3. Luminance–voltage characteristics of the films on (*1*) ferroelectric and (*2–4*) glass: (*2*) as-prepared structure, (*3*) in 1 h, and (*4*) in 4 h.

within the first several hours of operation. Subsequently, this trend becomes less pronounced. The curves taken of our ZnS : Cu films are similar to those obtained for widely used ZnS : Mn and ZnS : Tb EL phosphors [3] and indicate that the training does not cause the degradation of the structures; on the contrary, it stabilizes their characteristics. Thus, the feasibility of ZnS : Cu-based stable EL phosphors is obvious.

The charge–voltage characteristics $Q(U)$ of the ZnS : Cu films deposited on the ceramic and glass (Figs. 4a and 4b, respectively) change drastically after the onset of electroluminescence. An extended ellipse for the ceramic and a straight line for the glass change to parallelograms. Note that the slope of the $Q(U)$ curve defines the total capacitance of the structure in the subthreshold range and the capacitance of the insulators in the above-threshold range (the slope of the curve is higher). The $Q(U)$ curve for the structures on the ceramic (the presence of charge before the onset of luminescence) suggests the existence of leakage charge. This may be related to the use of the ferroelectric as an insulator or to defects in the EL structure.

The two different voltage dependences of the charge (in the subthreshold and above-threshold ranges for the luminance curve) can be explained in terms of equivalent circuits for these two states. In the subthreshold range, the voltage dependence of the charge passing through the structure is linear and its slope yields the integral capacitance of the structure. The associated equivalent circuit can be represented as the capacitance of the insulators, C_i, series-connected with that of the ZnS film, C_{ZnS}. Because of the difference in permittivity ε, $C_i > C_{ZnS}$. This difference is especially large for the films on the ceramic, since ε for ferroelectrics is as high as 1600 (for traditional insulators, $\varepsilon = 8$–10). In this case, the equivalent circuit is akin to a capacitive voltage divider. Therefore, the voltage across the EL layer is much higher for the films on the ceramic substrate if the applied voltage is the same. This causes the threshold voltage of this structure to shift toward lower voltages although the field strengths for both structures remain roughly equal to each other, $(4$–$7) \times 10^5$ V/cm. From the capacitance measurements, it follows that the EL layer does not contain field concentrators due to the presence of a barrier (e.g, Mott–Schottky) layer or conductive phase (Cu_2S or ZnO) inclusions.

The luminance–voltage and charge–voltage curves correlate: both the luminance and the charge increase with voltage. However, the former curve is steeper.

The brightness waveforms (Fig. 5) for the films on the glass suggest that the emission is almost in phase with the voltage. The constant brightness component ΔB_0 results from the overlap of the brightness peaks because of their slow decay within the voltage half-period. Such brightness waveforms are typical of conventional ZnS : Mn EL film phosphors, where the emission is excited by the impact mechanism [3]. In ZnS : Cu powder EL phosphors, where electroluminescence is
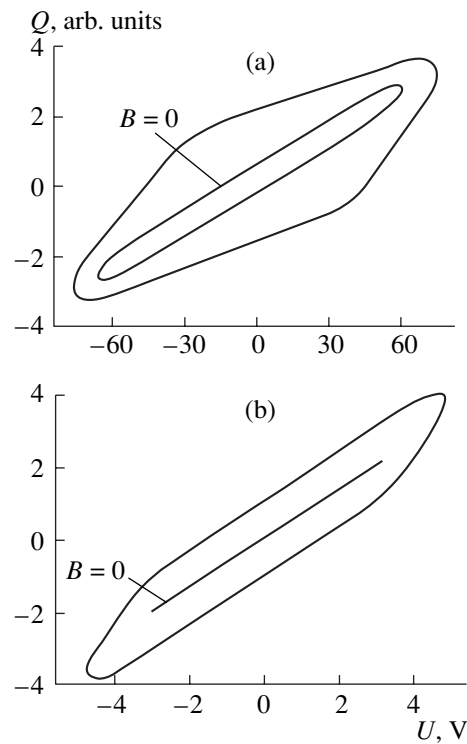

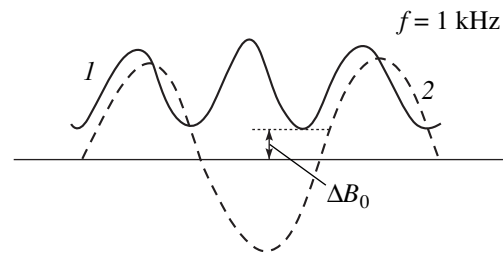**Fig. 4.** Charge–voltage characteristics for ZnS : Cu EL films on (a) ferroelectric and (b) glass.

Fig. 5. (*1*) Brightness waveform for the EL structure on glass and (*2*) exciting voltage waveform.

initiated by the recombination mechanism, two, instead of one, brightness peaks are observed within the voltage half-period: one is in phase with the voltage peak, while the other is in phase with the zero value of the voltage.

To summarize, the EL characteristics of the thin-film structures strengthen the supposition that Cu^+ ions are excited by the hot-electron impact mechanism.

CONCLUSION

Thus, the study of the structure and also the electrical and EL performance of ZnS : Cu films prepared by the chemical method (that is, by the joint pyrolysis of zinc and copper dithiocarbamates) shows that they are homogeneous in composition and do not contain field concentrators. As follows from the EL characteristics,

the mechanism of electroluminescence in these films is hot-electron impact excitation of emission centers, because the emission appears at electric fields as high as 5×10^5 V/cm. The luminance–voltage characteristic has a threshold and a region of steep rise. The luminance and the carrier concentration vary with voltage in a similar manner. The emission starts immediately after the EL layer has been excited. The training of the structures causes the EL characteristics to shift toward higher voltages, which is also typical of conventional ZnS : Mn and ZnS : Tb EL structures made by vacuum techniques. Thus, one can conclude that ZnS : Cu films prepared by the chemical method studied are promising for EL displays.

REFERENCES

1. Sey-Shing Sun, E. Dickey, J. Kane, and P. Niel Yocom, JDRC, Vol. 301 (Toronto, 1997).

2. L. J. Berezhinsky, V. S. Khomchenko, V. E. Rodionov, and Yu. A. Tzyrkunov, in *Book of Extended Abstracts of the 5th International Conference on the Science and Technology of Display Phosphors, San Diego, California, 1999.*

3. Y. A. Ono, *Electroluminescent Displays* (World Sci., London, 1995).

4. L. F. Zharovsky, L. V. Zavyalova, and G. S. Svechnikov, Thin Solid Films **128**, 241 (1985).

5. L. V. Zavyalova and G. S. Svechnikov, Displays **18**, 73 (1997).

6. L. V. Zavyalova, A. I. Beletski, and G. S. Svechnikov, Semicond. Sci. Technol. **14**, 446 (1999).

7. S. V. Svechnikov, L. V. Zavyalova, N. N. Roshchina, *et al.*, Fiz. Tekh. Poluprovodn. (St. Petersburg) **34**, 1178 (2000) [Semiconductors **34**, 1128 (2000)].

8. V. E. Rodionov, Ukr. Patent 15 210 (1997).

9. V. S. Khomchenko, Ja. F. Kononec, N. A. Vlasenko, *et al.*, J. Cryst. Growth **101**, 994 (1990).

10. P. H. Holloway, J.-E. Yu, P. Rack, *et al.*, Mater. Res. Soc. Symp. Proc. **345**, 289 (1994).

11. T. Yamaduchi, Y. Yamamoto, T. Tanaka, *et al.*, Thin Solid Films **281–282**, 375 (1996).

12. V. G. Bessergenev, V. I. Belyi, A. A. Rastorguev, *et al.*, Thin Solid Films **279**, 135 (1996).

13. P. A. Arutyunov and A. L. Tolstikhina, Kristallografiya **43**, 524 (1998) [Crystallogr. Rep. **43**, 483 (1998)].

Translated by V. Isaakyan

SOLID-STATE ELECTRONICS

Growth Kinetics of CdSe₁₋ₓSₓ Microcrystallites in Silicate Glass

Growth Kinetics of $CdSe_{1-x}S_x$ Microcrystallites in Silicate Glass

P. G. Petrosyan and L. N. Grigoryan

Yerevan State University, ul. Manukyana 1, Yerevan, 375025 Armenia
e-mail: ppetros@www.physdep.r.am
Received December 5, 2001

Abstract—The absorption spectra of silicate glass $CdSe_{1-x}S_x$ composite as a function of temperature are studied. At the early stage of thermal treatment, a shift in the absorption edge toward longer waves is associated with an increase in the density of critical nuclei. A further shift of the spectra is due to the growth of semiconductor nanocrystallites. © 2002 MAIK "Nauka/Interperiodica".

The optoelectronic performance of semiconducting composites (SCs) (semiconductor inclusions in an insulating matrix) have recently become the subject of much interest. In these materials, the size quantization of the electron and hole energy spectrum may occur depending on inclusion dimensions. The optical behavior of SCs containing elementary and binary semiconductor inclusions has been investigated elsewhere [1–3].

In this work, we report the growth kinetics of $CdSe_{1-x}S_x$ microcrystallites in silicate glass. The transmission and absorption spectra of the SCs were measured at various stages of microcrystallite formation, and kinetic features were elucidated from a shift in the optical absorption edge.

$CdSe_{1-x}S_x$–glass composites were prepared by the conventional glass foundry process. CdSe and CdS semiconductor phases with concentrations of about 1 wt % were added to the initial charge. The foundry temperature was between 1350 and 1400°C. The concentration of the semiconductor phase depends on the limiting solubility of the semiconductor in the glass at the foundry temperature and was found to be several tenths of percent [1]. Upon fast cooling of the melt, a supersaturated solid solution of the semiconductor in the glass forms. At room temperature, this metastable phase can exist infinitely long. At higher temperatures, however, it decays and semiconductor microcrystallites nucleate by the fluctuation mechanism. Subsequently, they grow via the diffusion of the semiconductor from the supersaturated solid solution.

Samples used were 1-mm-thick plane-parallel plates that were polished on both sides. The transmission spectra were taken at wavelengths ranging from 200 to 700 nm with a KSVU-23 general-purpose optical computing complex.

As-prepared $CdSe_{1-x}S_x$–silicate glass composites were heat-treated in air. It was found that even long-term treatment below 400°C changes the transmission spectra insignificantly. The heat treatment at temperatures above 400°C shifts the absorption edge toward longer wavelengths. As the heat treatment time increases at a given temperature, the shift slows down and eventually ceases. The spectrum can be shifted further only by increasing the treatment temperature.

To elucidate the role of the glass matrix in the spectrum transformations, we recorded the transmission spectra of the silicate glass free of semiconductor inclusions. Figure 1 shows the transmission spectra for the glass (curves *1–3*) and the $CdSe_{1-x}S_x$–glass composite (curves *4–9*) vs. thermal treatment time.

It was found that the long-term treatment of the glass at 570°C may cause the transmission edge to shift to 340–350 nm. In the range 350–700 nm, the glass matrix is virtually transparent. Therefore, it can be concluded that the shift of the absorption edge at wavelengths longer than 350 nm is due to the presence of the semiconductor phase in the glass. At treatment durations shorter than 40 min, the transmission spectrum of the composite exhibits a long tail (Fig. 1). At durations greater than 40 min, the tail shortens and the absorption edge becomes distinct (curves *6–9*).

We believe that the early stage of thermal treatment produces $CdSe_{1-x}S_x$ solid solution inclusions of size such that the band structure of the semiconductor affects the optical properties of the material.

To determine the absorption edge, we found the spectral dependence of the absorption coefficient $\alpha = \alpha(\hbar\omega)$. It is known that $\alpha \sim (\hbar\omega - E_g)^{1/2}$, which corresponds to allowed direct transitions. With this approximation, we determined the absorption edge related to the energy gap of the semiconductor solid solution. The shift of the absorption edge of the SC is associated with the presence of the semiconductor microcrystallites in the matrix, which form when the supersaturated solid solution decays. Here, three steps of the phase decay are distinguished: nucleation, diffusion growth, and coalescence of semiconductor crystallites [1].
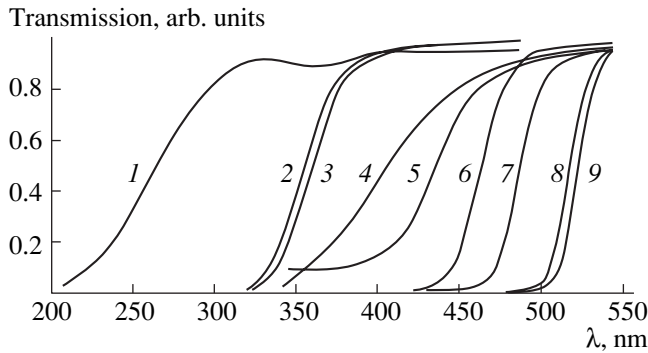
Transmission, arb. units

Fig. 1. Transmission spectra of (*1–3*) silicate glass [heat treatment temperature $T = 570°C$; heat treatment time t is 0 (*1*), 2 (*2*), and 60 min (*3*)] and (*4–9*) CdSe$_{1-x}$S$_x$–silicate glass composite [$T = 450°C$; $t = 0$ (*4*), 15 (*5*), 40 (*6*), 60 (*7*), 120 (*8*), and 180 min (*9*)].
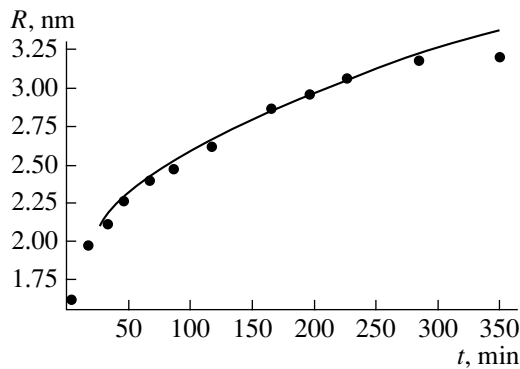

Fig. 2. Mean radius of CdSe$_{1-x}$S$_x$ nanocrystallites vs. heat treatment time.

We assume that the shift of the transmission spectrum at the early stage of thermal treatment (<40 min) is related to nucleation, while the subsequent shift is associated with the diffusion growth of nuclei [4].

The presence of the nucleation step is also supported by the fact that the absorption edge of the composite heat-treated at 550°C shifts from 380 to 600 nm within several minutes. Such a high rate is untypical of diffusion processes.

According to [5], the crystallite radius R grows during diffusion by the law

$$R(t) = \left[\frac{2Dt(C_0 - C_e)}{(C_{cr} - C_e)}\right]^{1/2},$$

where C_0 is the initial semiconductor concentration in the matrix, C_e is the equilibrium semiconductor concentration in the matrix, C_{cr} is the semiconductor concentration in a crystallite, D is the effective diffusion coefficient, and t is time.

In semiconductor nanocrystallites, the size quantization of the electron and hole energy spectra is known to take place, which shifts the absorption edge toward

shorter waves relative to that in a bulk crystal. Assuming that size quantization is also responsible for the shift of the optical transmission spectrum observed in our experiments, we can find the mean radius of the nanocrystallites by the formula [6]

$$R = \frac{\hbar\pi}{\sqrt{2\mu(\hbar\omega_0 - E_g)}},$$

where m_e and m_h are the electron and hole mass, respectively; $\mu = m_e m_h/(m_e + m_h)$ is the reduced mass; E_g is the energy gap; and $\hbar$ is the Planck constant. Here, we assume that the hole and electron bands are parabolic and ignore electron–hole interaction.

Now, we can derive the dependence of the mean radius of CdSe$_{1-x}$S$_x$ crystallites on heat treatment time (Fig. 2). The energy gap of bulk CdSe$_{1-x}$S$_x$ was found from the transmission spectra taken from similar samples subjected to long-term thermal treatment at higher temperatures (e.g., at 570°C). In this case, it can be assumed that the semiconductor inclusions grow to the point where their characteristic size exceeds the de Broglie wavelength for CdSe$_{1-x}$S$_x$, so that the size quantization effect can be neglected. The associated energy gap of bulk CdSe$_{1-x}$S$_x$ is $E_g = 2$ eV. In [7], the energy gap of CdSe$_{1-x}$S$_x$ solid solution vs. composition dependence was found to be

$$E_g = 1.89 + 0.16x + 0.41x^2. \tag{1}$$

From expression (1), we have $x = 0.36$ for $E_g = 2$ eV. Early in the crystallization, the inclusion radius depends linearly on the square root of the heat treatment time: $R = R(t)$ can be described by the empirical relationship $R = (16 + 0.9t^{1/2})$ Å, where t is measured in minutes. This means that the initial size of a nucleus is 16 Å and then it grows by the diffusion mechanism. The deviation from the radical dependence is observed after a lapse of much time (more than 4 h). The reason for the deviation may be that the heat-treated matrix becomes depleted and can no longer be considered as the solid solution, for which the radical dependence $R = R(t)$ is valid. After the long heat treatment at 450°C, the nucleus size was estimated at 32 Å.

As was shown [1], at the initial stage of the phase decay, the mean size of the stable particles virtually equals the critical one and does not change, but their number increases.

The shift of the transmission edge toward longer wavelengths at small treatment times can be explained in terms of the Maxwell–Garnett theory. Within this theory, the long-wave shift of the transmission curve upon annealing is explained by increasing absorption as the volume fraction of the microcrystallites grows [8].

In the Maxwell–Garnett theory, the dependence of the absorption coefficient $\alpha(t)$ on the annealing time is

given by

$$\alpha \;=\; \frac{18\pi\varepsilon_{m}^{3/2}f}{\lambda}\frac{\varepsilon_{2}}{(\varepsilon_{1}+2\varepsilon_{m})^{2}+\varepsilon_{2}^{2}}, \qquad (2)$$

where ε_1 and ε_2 are the real and imaginary parts of the permittivity of semiconductor microcrystallites, ε_{m} is the permittivity of the matrix, $f = (4/3)\pi N(t)R_{0}^{3}$ is the volume fraction of the microcrystallites, $N(t)$ is the density of the nuclei, and R_0 is the mean radius of critical nuclei.

In [7], the dispersion curves for the refractive index n and absorption coefficient α in the range 350–600 nm for $CdSe_{1-x}S_x$ solid solution were found. At $x = 0.36$, these curves can be closely approximated with the empirical relationships

$$n \;=\; 2.49 + 0.29e^{-102(\lambda-0.57)^{2}}, \quad \alpha \;=\; 71.2 - 85.6\sqrt{\lambda}.$$

From the expressions $\varepsilon_1 = n^2 - \alpha^2$ and $\varepsilon_2 = 2n\alpha$, we determined the spectral dependences of the real and imaginary parts of the permittivity (Fig. 3).

Under the assumption that the shift of the transmission spectrum is initially associated with the growing number of critical nuclei (with their size remaining unchanged), the experimental curves were approximated by the expression $I = I_0(1 - R)e^{-ad}$ (in experiments, $R \ll 1$), where α is defined by (2). From this approximation, the density of the nuclei as a function of heat treatment time was determined (see table).

At the stage of diffusion growth of the nuclei (i.e., immediately after the nucleation stage), the density of critical nuclei does not change but their size increases. If it is assumed that an increase in the volume fraction of the microcrystallites (with the concentration of the nuclei unchanged, $N_n = 6.8 \times 10^{15}$ cm^{-3}) is responsible also for the shift of the absorption edge to the long-wave part of the spectrum at the second step of heat treatment, then the radius of the microcrystallites vs. treatment time can be approximated by a similar relationship (the values of R are also listed in the table). Since we suppose that the crystallites grow by the diffusion mechanism at the second stage of heat treatment, the radical dependence of the inclusion radius on heat treatment time must be observed. Figure 4 shows the dependence $R = R(t)$, which is best approximated by the empirical expression

$$R \;=\; (16.74 + 0.6\sqrt{t})\ \text{Å}.$$

Thus, the dependences of the nucleus radius on heat treatment time derived above are well approximated by the radical curve, which is typical of the diffusion growth of the microcrystallites.

The shift of the transmission curves toward longer waves upon annealing can be explained by both size quantization in nanocrystallites and the enhancement of the absorption as the volume fraction of the microcrys-
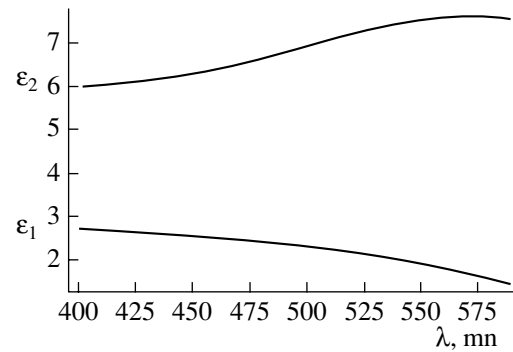

Fig. 3. Spectral dependences of the real, ε_1, and imaginary, ε_2, parts of the permittivity for $CdSe_{1-x}S_x$ ($x = 0.36$).

tallites grows. We believe that both mechanisms are responsible for the shift. At the nucleation stage, where the crystal structure of the semiconductor has not yet formed, clustering changes the effective permittivity, shifting the transmission edge to the long-wave part. After the initial stage, where the crystal structure of the clusters has already formed, the spectral shift is due to the absorption in coarse crystallites.

The relative contributions of these two mechanisms were evaluated by comparing the dependence of the absorption coefficient on heat treatment time obtained experimentally for $\lambda = 450$ nm with analytical expression (2) (Fig. 5). At the initial stage of the treatment, the curves are close to each other. This means that, at the first stage, the shift of the transmission edge can be explained by an increase in the volume fraction of the crystallites in the matrix. At times exceeding 40 min, the curves diverge; hence, the absorption in the crystallites should also be taken into consideration.

The sizes of the $CdSe_{1-x}S_x$ microcrystallites were determined by the low-angle X-ray scattering method. The typical size of the microcrystallites in the samples heat-treated at 450°C for 4 h was found to be $R_0 = 42$ Å.

As was shown in [6], the X-ray sizes of the microcrystallites are always larger than the mean value, because microcrystallites of size smaller than average take a minor part in scattering. According to [6], the actual average size of the microcrystallites is $\bar{R} =$

Table

t, min	N, cm^{-3}	R, Å
0	3.2×10^{15}	16
15	5.2×10^{15}	16
40	6.8×10^{15}	16
60	6.8×10^{15}	21
120	6.8×10^{15}	23
180	6.8×10^{15}	24
280	6.8×10^{15}	25.5

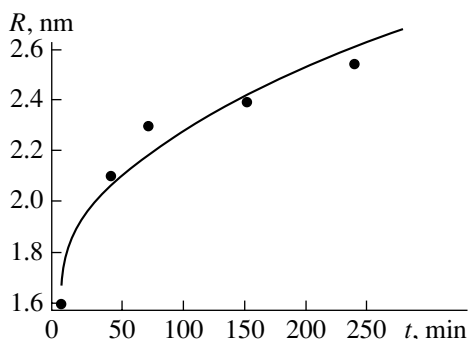**Fig. 4.** Mean radius of $CdSe_{1-x}S_x$ nanocrystallites vs. heat treatment time (Maxwell–Garnett theory).
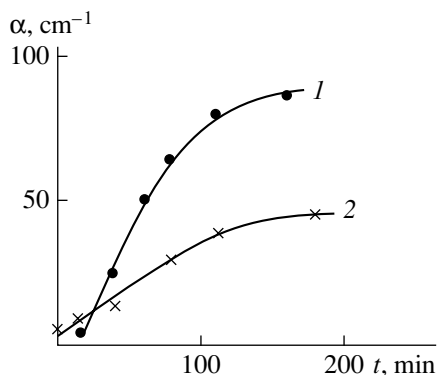

Fig. 5. Absorption coefficient vs. heat treatment time for $\lambda = 450$ nm: (*1*) experiment and (*2*) theory.

$0.86\bar{R}_0$, where $\bar{R}_0$ is the value obtained by X-ray measurements in the monodisperse approximation. In our case, $\bar{R} = 36$ Å, which is in close agreement with the value given by empirical expression (1).

The results obtained indicate that, at the second stage of the heat treatment, the shift of the transmission spectrum is basically due to size quantization. If, however, it is assumed that the shift depends totally on the size quantization effect, the average size of the microcrystallites must be smaller than the actual value that was observed in our measurements.

Thus, the shift of the absorption edge for $CdSe_{1-x}S_x$–silicate glass semiconductor composites at the initial stage of heat treatment (<40 min, $T = 450°C$) is associated with the increase in the number of critical nuclei 16 Å in size. Subsequently, the spectrum shifts because of the growth of semiconductor nanocrystallites.

ACKNOWLEDGMENTS

REFERENCES

1. S. A. Gurevich, A. I. Ekimov, I. A. Kudryavtsev, *et al.*, Fiz. Tekh. Poluprovodn. (St. Petersburg) **28**, 830 (1994) [Semiconductors **28**, 486 (1994)].

2. K. M. Hassan, A. K. Sharma, and J. Narayan, Appl. Phys. Lett. **75**, 1222 (1999).

3. R. Thielsh, T. Bohme, R. Reiche, and D. Schlafer, Nanostruct. Mater. **10**, 131 (1998).

4. P. G. Petrosyan and L. N. Grigoryan, in *Proceedings of the Second National Conference "Semiconductors Microelectronics"* (Erevan. Gos. Univ., Yerevan, 1977), p. 212.

5. R. Kampmann and R. Wagner, in *Decomposition of Alloys, the Early Stages: Proceedings of the 2nd Acta-Scripta Metallurgica Conference, Sonnenberg, 1983*, Ed. by P. Haasen *et al.* (Pergamon, Oxford, 1983), p. 91.

6. A. I. Ekimov and Al. A. Éffros, in *Proceedings of the XII Winter School on Physics of Semiconductors, Leningrad, 1986,* p. 65.

7. M. M. El-Nahass, O. Jamjoum, and A. M. Al Howaity, J. Mater. Sci. Lett. **9**, 79 (1990).

8. C. F. Bohren and D. R. Haffman, *Absorption and Scattering of Light by Small Particles* (Wiley, New York, 1983; Mir, Moscow, 1986).

Translated by V. Isaakyan

SOLID-STATE
ELECTRONICS

Spin-Valve Magnetoresistive Structures Based on Co/Tb Multilayer Films

A. V. Svalov[1], P. A. Savin[1], G. V. Kurlyandskaya[2, 3], J. Gutiérrez[2], and V. O. Vas'kovskiy[1]

[1] *Gor'kiĭ State University, pr. Lenina 51, Yekaterinburg, 620083 Russia*
e-mail: andrey.svalov@usu.ru
[2] *University of Basque Provinces, Bilbao, Spain*
[3] *Institute of Metal Physics, Siberian Division, Russian Academy of Sciences,
ul. S. Kovalevskoĭ 18, Yekaterinburg, Russia*
Received December 25, 2001

Abstract—The effect of the layer thickness on the magnetic properties of $\{Co/Tb\}_n$, $\{Co/Tb\}_n/Co$, and $\{Co/Tb\}_n/Co/Cu/Co$ multilayer films is studied. The dependence of the hysteresis and magnetoresistive properties of $\{Co(1\ nm)/Tb(1\ nm)\}_n/Co(5\ nm)/Cu(L_{Cu})/Co(5\ nm)$ structures on the thickness of the $\{Co/Tb\}_n$ layer and copper spacing are obtained. The feasibility of spin-valve structures based on $\{Co/Tb\}_n$ multilayer films with in-plane anisotropy is demonstrated. © *2002 MAIK "Nauka/Interperiodica".*

INTRODUCTION

A change in the mutual orientation of the magnetizations of ferromagnetic layers separated by a conductive nonmagnetic spacer under the action of a magnetic field may change the conductivity of the layered structure. This effect is called giant magnetoresistance (GMR) [1]. To date, GMR-based devices, such as reading heads for magnetic recording systems or position- and/or velocity-sensitive detectors, have been designed and implemented. By convention, multilayer systems exhibiting GMR can be subdivided into two groups: Fe/Cr superlattices, where magnetic layers are antiferromagnetically ordered through a nonmagnetic spacer [2], and so-called spin-valve structures, where a difference in the coercive forces of the layers or the presence of unidirectional anisotropy (magnetic biasing) in one of the layers causes layer-by-layer magnetization reversal [3, 4]. Unidirectional anisotropy is provided by the use of additional (pinning) layers of antiferromagnetic or magnetically hard materials [1]. Of the former, NiO insulator seems to be the most promising for spin-valve structures. It ensures a high blocking temperature, i.e., the temperature at which exchange interaction between the ferromagnetic and antiferromagnetic layers being in contact disappears. Of the high-coercivity materials used for the pinning layer, amorphous TbCo ferrimagnetic films have shown great promise. Along with the high coercive force H_c, they provide a high blocking temperature in FeNi/Cu/FeNi/TbCo and FeNi/Cu/Co/TbCo spin-valve structures [5, 6]. The high coercive force observed in [5, 6] was achieved by using magnetic compositions exhibiting magnetic compensation at room temperature. However, such films have perpendicular magnetic anisotropy. Moreover, in the films with strictly perpendicular anisotropy, adjacent ferro-

magnetic layers did not show magnetic biasing. For the bias to exist, at least a low magnetization in the plane of the TbCo layer is necessary [5]. Tb/Co multilayer films, which also exhibit magnetic compensation and offer high H_c [7], are an alternative to TbCo amorphous films. In this work, we show the feasibility of producing spin-valve structures based on Co/Tb multilayer films with in-plane anisotropy and study the dependence of the hysteresis and magnetoresistive properties of $\{Co/Tb\}_n/Co/Cu/Co$ structures on the thickness of the $\{Co/Tb\}_n$ layer and copper spacing.

EXPERIMENTAL

The multilayer films were prepared by successively depositing layers on glass or silicon substrates in argon with rf ion–plasma sputtering. The deposition rate was predetermined on thicker uniform films with the Tolansky method and low-angle X-ray scattering. It was found to be 2.5 nm/min for cobalt layers and 5 nm/min for terbium and copper layers. The films were deposited in a constant magnetic field of 100 Oe parallel to the substrate plane. This field is necessary for the easy magnetic axis (EMA) to lie in the film plane. In the $\{Co/Tb\}_n/Co/Cu/Co$ spin-valve structures, the thicknesses of the "free" (separated by the Cu spacing) and "pinned" (being in contact with the Tb/Co structure) cobalt layers were fixed (5 nm). The thicknesses of the cobalt and terbium layers in the $\{Co/Tb\}_n$ multilayer film were also kept constant (1 nm). The number n of periods in the multilayer film was varied between 8 and 60, and the thickness L_{Cu} of the copper spacing ranged from 1 to 5 nm.

As a free ferromagnetic layer, we used a Co film of thickness $L_{Co} = 5$ nm. Its coercive force varied from

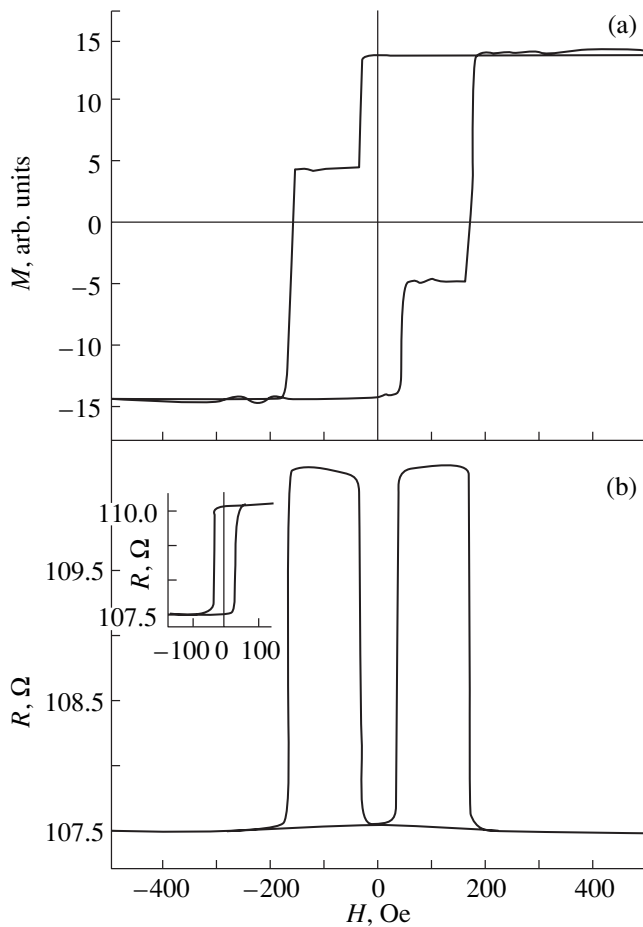**Fig. 1.** (a) Magnetometric and (b) magnetoresistive hysteresis loops for the {Co(1 nm)/Tb(1 nm)}$_{16}$/Co(5 nm)/Cu(2.5 nm)/Co(5 nm) structure. The inset shows the partial magnetoresistive loop measured between –150 and 150 Oe.

sample to sample within 35 and 40 Oe. Magnetization reversal fields for cobalt films are usually higher than those for Permalloy films; however, the former provide a higher value of the magnetoresistive effect [8]. The thickness L_{Co} = 5 nm was also taken so as to provide the maximum possible GMR [1, 9].

A {Co(1 nm)/Tb(1 nm)}$_n$ multilayer film composition was used as a pinning composition. As follows from preliminary experiments, these films have in-plane anisotropy. The spontaneous magnetization increases monotonically in the temperature range 90–370 K, which suggests the ferrimagnetic nature of this layered structure. The coercive force of the films exceeds 200 Oe at room temperature. It should also be noted that if the thickness of the layers is less than 1 nm, the multilayer films have a normal magnetization component. For much thicker layers incorporated into the films, the magnetization varies with temperature non-monotonically.

The magnetic properties of the films were studied with the Kerr magnetooptic effect, a torsional magnito-

meter, and a SQUID magnetometer (Quantum Design Co). The magnetoresistance was determined with a bridge circuit made of 12×2-mm strips. The current passing through the sample was 15 mA. The GMR value $\Delta R/R$ is defined as $(R_a - R_p)/R_p$, where R_a and R_p are the respective resistances of the spin-valve structure with the antiparallel and parallel mutual orientation of the magnetizations in the free and pinned Co layers.

RESULTS AND DISCUSSION

The magnetic pinning of one of the layers is an important stage in producing spin-valve structures. In our case, the pinning was accomplished through exchange coupling between the relatively thin Co layer and the Tb layer completing the multilayer structure {Co(1 nm)/Tb(1 nm)}$_n$. In essence, the Co layer and the {Co(1 nm)/Tb(1 nm)}$_n$ structure can be viewed as two films with different coercive forces that interact via direct exchange. Such an approach allows us to estimate the critical thickness t_{cr} below which magnetization reversal in the coupled films follows a single hysteresis loop. The critical thickness is found from the relationship [10]

$$t_{cr} \approx \pi \sqrt{\frac{A}{M_s(H_{c2} - H_{c1})}},$$

where A, M_s, H_{c1} are the exchange interaction constant, saturation magnetization, and coercive force of the magnetically soft (Co) film, respectively, and H_{c2} is the coercive force of the magnetically hard ({Co(1 nm)/Tb(1 nm)}$_n$) film.

Putting $A = 1 \times 10^{-6}$ erg/cm, M_s = 1400 G, and (H_{c2} – H_{c1}) $\approx$ 200 Oe, we have $t_{cr} \approx$ 60 nm. Thus, the thickness of the Co film selected (5 nm) must provide the formation of the {Co(1 nm)/Tb(1 nm)}$_n$/Co(5 nm) substructure that behaves as a unit in a magnetic field. The coercive force H_c^{sub} of such a substructure must lie between the coercive force of the Co(5 nm) film and that of the {Co(1 nm)/Tb(1 nm)}$_n$ multilayer film. The experiments confirmed these estimates.

Figure 1a shows a hysteresis loop obtained with the SQUID magnetometer for the fully formed {Co/Tb}$_{16}$/Co/Cu(2.5 nm)/Co. The magnetic field was aligned with the EMA. The sharp changes in the magnetizations that are observed on the loop correspond to magnetization reversal in the free and pinned Co layers, which takes place at fields of 40 and 170 Oe, respectively. The magnetoresistive loop measured on the same film with the magnetic field and the current aligned with the EMA is depicted in Fig. 1b. From Figs. 1a and 1b, one can conclude that when the mutual orientation of the magnetizations in the free and pinned layers changes, so does the structure resistance; i.e., the GMR effect is realized.

The inset in Fig. 1b demonstrates the partial magnetoresistive loop measured at fields lower than the reversal field for the pinned layer. The loop is somewhat shifted along the field axis relative to $H = 0$, reflecting the coupling between the free and pinned Co layers through the copper spacing. A similar shifted magnetization reversal loop for the free layer was observed when the Kerr magnetooptic effect was employed in the measurements. The reversal field for the free layer is the sum of the coercive force and the bias field H_b. The coupling of the magnetic layers through the copper spacing can be accomplished both via microholes in the spacing and magnetostatically as a result of the layer roughness. The bias H_b is a measure of such coupling. In the presence of this interaction, magnetization reversal in the layers cannot take place separately, and the magnetizations in adjacent layers are not strictly antiparallel; hence, the GMR diminishes. It would be natural to expect that the degree of coupling will increase with decreasing copper spacing thickness L_{Cu}. Figure 2a shows the dependence $H_b(L_{Cu})$ for the $\{Co/Tb\}_{60}/Co/Cu(L_{Cu})/Co$. At $L_{Cu} = 5$ nm, H_b is low and increases as L_{Cu} decreases. For $L_{Cu} = 1$ nm, the coupling between the Co layers becomes so strong that the spin-valve structure behaves as a unit and the magnetization is reversed in the entire structure under high fields. The change in the magnetic properties of the structure affects its magnetoresistive properties. Figure 2b shows the dependence of $\Delta R/R$ on the copper spacing thickness for samples with differing numbers of periods in the $\{Co(1\ nm)/Tb(1\ nm)\}_n$ structure. When L_{Cu} is small, the GMR is absent. As L_{Cu} rises, this effect shows up only if the magnetizations in the free and pinned layers are reversed separately. The further increase in L_{Cu} decreases $\Delta R/R$ (curve 1) because of the shunting effect of the spacing and also because of electron scattering in the spacing. In combination, these observations indicate that a significant increase in L_{Cu} with the aim to decrease H_b is inappropriate, since the GMR drops in this case. In our structures, the optimal thickness of the copper spacing was between 1.7 and 2.5 nm.

The $\{Co(1\ nm)Tb(1\ nm)\}_n$ pinning film also has a shunting effect and, thereby, affects the value of $\Delta R/R$. Obviously, the number of periods n should be decreased in order to weaken the shunting effect. On the other hand, when n is small, the coercive force of the $\{Co(1\ nm)Tb(1\ nm)\}_n/Co(5\ nm)$ substructure diminishes, as is shown in Fig. 3 (curve 1). The decrease in H_c^{sub} reduces the difference in the reversal fields for the free and pinned layers. As a result, magnetization reversal taking place in the free layer does not provide the strictly antiparallel orientation of the magnetizations in these layers; consequently, the shape of the magnetoresistive loop and the value of $\Delta R/R$ change. Specifically, the plateau on the magnetoresistive loop (Fig. 4a) at $n = 60$ convincingly indicates that
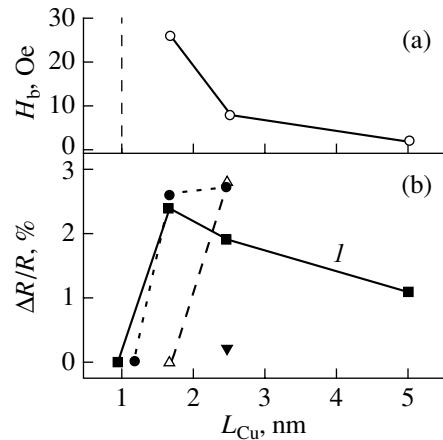

Fig. 2. (a) H_b for $\{Co/Tb\}_{60}/Co/Cu(L_{Cu})/Co$ and (b) $\Delta R/R$ for $\{Co/Tb\}_n/Co/Cu(L_{Cu})/Co$ structures vs. copper spacing thickness L_{Cu}. $n = 60$ (■), 30 (●), 16 (△), and 8 (▼).
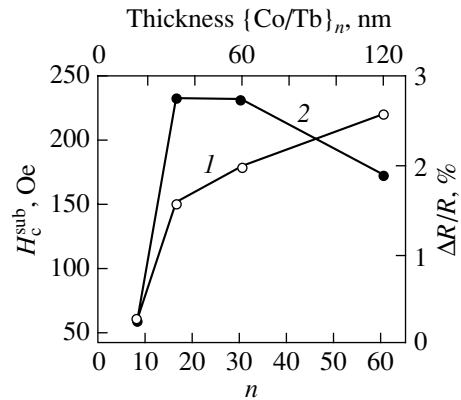

Fig. 3. Coercive force H_c^{sub} for the $\{Co(1\ nm)Tb(1\ nm)\}_n/Co(5\ nm)$ structure and $\Delta R/R$ for the $\{Co(1\ nm)/Tb(1\ nm)\}_n/Co(5\ nm)/Cu(2.5\ nm)/Co(5\ nm)$ spin-valve structure vs. the thickness of the $\{Co(1\ nm)/Tb(1\ nm)\}_n$ film.

magnetization reversal proceeds separately. At $n = 8$, the magnetoresistive loop (Fig. 4b) has a very different shape. The sharp peak implies the absence of a distinct antiparallel arrangement of the magnetizations in adjacent layers. Because of the two competing trends in the GMR variation with the thickness of the $\{Co(1\ nm)/Tb(1\ nm)\}_n$ pinning layer, the dependence of $\Delta R/R$ on n is nonmonotonic (Fig. 2, curve 2).

The difference between the reversal fields, which affects $\Delta R/R$, depends not only on H_c^{sub} but also on the efficiency of coupling through the copper layer, as follows from the $\Delta R/R$ vs. L_{Cu} curves taken for various n (Fig. 2b). The smaller the value of n, the thicker must be the spacing for magnetization reversal in the free and pinned layers to proceed separately and the GMR effect to occur. Eventually, we can achieve the same values of
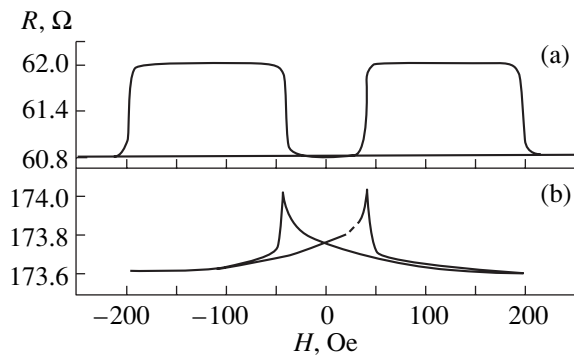
Fig. 4. Magnetoresistive hysteresis loops for the {Co(1 nm)/Tb(1 nm)}$_n$/Co(5 nm)/Cu(2.5 nm)/Co(5 nm) spin-valve structure. n = (a) 60 and (b) 8.

$\Delta R/R$ for different combinations of {Co(1 nm)/Tb(1 nm)}$_n$ and Cu spacer thicknesses.

CONCLUSION

Thus, we demonstrated the feasibility of {Co(1 nm)Tb(1 nm)}$_n$ multilayer ferrimagnetic films with in-plane anisotropy and spin-valve structures on their basis. It was found that the thicknesses of the {Co(1 nm)Tb(1 nm)}$_n$ layer and copper spacing have an effect on the reversal field in the magnetic layers and on the magnetoresistive properties of the {Co/Tb}$_n$/Co/Cu/Co spin-valve structure. The highest value of the magnetoresistive effect (2.8%) was obtained with the {Co(1 nm)/Tb(1 nm)}$_{16}$/Co(5 nm)/Cu(2.5 nm)/Co(5 nm) multilayer film.

ACKNOWLEDGMENTS

REFERENCES

1. *Magnetic Multilayers and Giant Magnetoresistance: Fundamentals and Industrial Applications*, Ed. by U. Hartman (Springer-Verlag, Berlin, 2000).

2. M. N. Baibich, J. M. Broto, A. Fert, *et al.*, Phys. Rev. Lett. **61**, 2472 (1988).

3. B. Dieny, V. S. Speriosu, B. A. Gurney, *et al.*, J. Magn. Magn. Mater. **93**, 101 (1991).

4. V. O. Vas'kovskiy and A. V. Svalov, J. Magn. Magn. Mater. **157–158**, 285 (1996).

5. P. P. Freitas, J. L. Leal, T. S. Plaskett, *et al.*, J. Appl. Phys. **75**, 6480 (1994).

6. N. J. Oliveira, J. L. Ferreira, J. Pinheiro, *et al.*, J. Appl. Phys. **81**, 4903 (1997).

7. Z. S. Shan and D. J. Sellmyer, in *Handbook on the Physics and Chemistry of Rare Earths*, Ed. by K. A. Gschneidner, Jr. and LeRoy Eyring (North-Holland, Amsterdam, 1996), Vol. 22.

8. S. S. P. Parkin, Appl. Phys. Lett. **61**, 1358 (1992).

9. S. Tumanski, *Thin Film Magnetoresistive Sensors* (Institute of Physics Publ., Bristol, 2001).

10. A. Yelon, in *Physics of Thin Films: Advances in Research and Development*, Ed. by G. Hass, M. Francombe, and R. Hoffman (Academic, New York, 1971; Mir, Moscow, 1973), Vol. 4.

Translated by V. Isaakyan

<div align="center">

OPTICS,
QUANTUM ELECTRONICS

Nonlinear Optical Parameters and Optical Limitation in Cobalt-Doped Polyvinylpyrrolidone Solutions

</div>

<div align="center">

R. A. Ganeev*, A. I. Ryasnyanskiĭ**, M. K. Kodirov**, Sh. R. Kamalov*,
V. A. Li***, R. I. Tugushev*, and T. Usmanov*

* Akadempribor Research and Production Association, Tashkent, 700143 Uzbekistan
** Navoi State University, Samarkand, 703004 Uzbekistan
e-mail: ryasn2000@yahoo.com
*** Institute of Physics and Chemistry of Polymers, Academy of Sciences of Uzbekistan,
Tashkent, 700128 Uzbekistan
Received August 13, 2001; in final form, March 1, 2002

</div>

Abstract—Nonlinear optical parameters (nonlinear refractive indices and nonlinear absorption coefficients) of solutions of polyvinylpyrrolidone doped by cobalt to various concentrations are measured at the lasing and second-harmonic wavelengths of a picosecond Nd:YAG laser (λ = 1064 and 532 nm, respectively, and τ = 35 ps). Data on optical limitation in these solutions are presented. The absence of nonlinear absorption in the IR spectral range and its significant effect in the visible range are demonstrated. Optical limitation at a wavelength of 1064 nm is related to defocusing, whereas at 532 nm, this effect is caused by two-photon absorption and partially by inverse saturated absorption and defocusing. Nonlinear optical parameters of metal–polymer complexes are reported. © 2002 MAIK "Nauka/Interperiodica".

INTRODUCTION

The search for new media for optical limitation and laser frequency conversion necessitates extensive studies of nonlinear optoelectron materials. Recently, polymers, especially those with conjugated π bonds (in particular, polyphenylvinylene and polyvinylpyrrolidone) and delocalized electrons, have become the subject of much investigation. Molecules with delocalized electrons exhibit large transition dipole moments and an increased nonlinear optical response. Nonlinear absorption with a substantial deviation from the Lambert law is typical of films made of such materials. Materials with optical limitation properties that combine pronounced optical nonlinearity and stability against high and medium levels of laser irradiation are of prime importance. It has been demonstrated that the nonlinear susceptibilities of a number of organic compounds with conjugated double bonds are comparable to the resonance nonlinear susceptibilities of atoms [1, 2]. Note that the delocalization of π electrons in fullerenes (C_{60}) is the reason for their high nonlinear optical susceptibility observed in the processes of harmonic generation [3–5] and phase conjugation [6].

Systematic studies of metallorganic molecules containing metals such as ruthenium, gold, and nickel, and metal clusters have been dictated by high second- and third-order nonlinearities in these substances. Some of these systems exhibit inverse saturable absorption (ISA), which also makes them promising for optical limiters. Being incorporated into polymer molecules, metals change both optical and nonlinear optical parameters of these systems [7, 8]. Pan *et al.* [7] used metal–polymer complexes for optical switching and the optical limitation of nanosecond pulses. Qureshi *et al.* [8] studied porphyrin polymers doped by zinc as an optical-limitation medium. It was demonstrated that these polymers effectively limit laser pulses with a duration τ = 500 ps and a wavelength λ = 532 nm. Note that polymer–metal complexes have been employed for optical limitation primarily in the visible spectral range.

Below, we report data for the optical limitation of picosecond IR (λ = 1064 nm) and visible (λ = 532 nm) radiations in solutions of polyvinylpyrrolidone (PVP) doped by cobalt at various concentrations of the metal. Nonlinear optical parameters of the polymer (nonlinear refractive indices, Kerr nonlinear susceptibilities, and nonlinear absorption coefficients) were measured.

EXPERIMENTAL

In the absence of a solvent, PVP can incorporate various molecules and atoms without deteriorating its originally high optical properties (in particular, a low level of scattering). PVP can also be used as a stabilizer in metal nanoclusters. This enables one to employ the polymer for preparing stable suspensions used in nonlinear optical studies [9]. At the same time, the thermal response of the compound to long (nanosecond or microsecond) radiation pulses may lead to effects detrimental to its nonlinear optical parameters. In particular, the interface between cold and heated areas may cause stresses in the polymer matrix. In this case, the
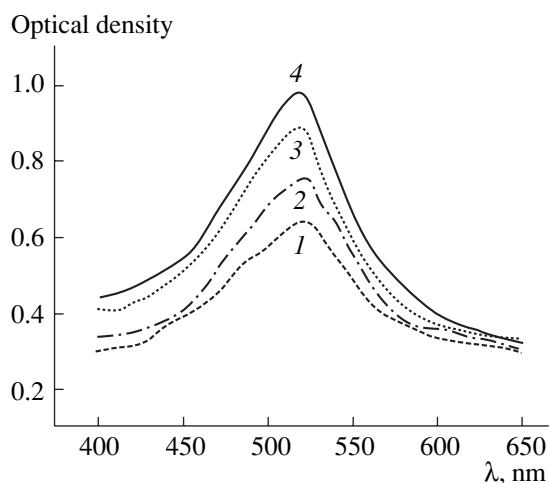
Optical density

Fig. 1. Absorption spectra of PVP aqueous solutions with cobalt concentrations of (*1*) 2.0, (*2*) 5.3, (*3*) 6.2, and (*4*) 13.5%.
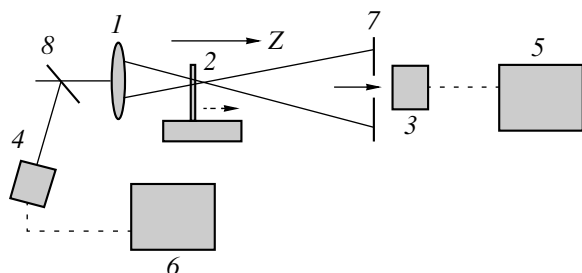

Fig. 2. Experimental setup for *Z*-scan experiments: *1*, focusing lens; *2*, quartz cuvette with the solution of the metallorganic polymer; *3* and *4*, FD-24K photodiodes; *5* and *6*, V4-17 digital voltmeters; *7*, iris; and *8*, deflector.

refractive index will change both because of variations in the density of the medium and because of birefringence due to stresses arising in the area of interaction between the laser beam and the polymer (elastooptical effect). By dissolving metallorganic systems, we eliminate this effect and can thus take into account only thermal expansion effects in the medium (for long pulses) and high-frequency Kerr effect (for short pulses).

In experiments, we used aqueous solutions of PVP doped by cobalt at concentrations of 2.0, 5.3, 6.2, and 13.5%. The solutions were prepared by dissolving 1 g of the polymer in 100 ml of distilled water at room temperature. The solution had a typical red color. With an SF-26 spectrophotometer, we measured the spectral parameters of PVP. Figure 1 shows the absorption spectra of the solutions near the absorption peak at 525 nm.

The conventional *Z*-scan method [10] was used to measure the nonlinear optical parameters of the solutions at wavelengths of 532 and 1064 nm (for details, see [11]). A Nd:YAG laser generated trains of picosecond pulses of duration $t = 35$ ps and energy $E = 2$ mJ. A

pulse selected from the train was focused by lens *1* with a focal length of 25 cm (Fig. 2). The diameters of the beam waists were 150 and 100 μm for the fundamental and second harmonics, respectively. The radiation maximum intensity at $\lambda = 1064$ nm reached 8×10^{11} W/cm^2. The energy of the laser pulses was measured by FD-24K calibrated photodiode *4* and detected by V4-17 digital voltmeter *5*. Calibrated neutral density filters attenuated the laser radiation. Micropositioner *8* moved 1-mm-thick quartz cuvette *2* with the polymer along the optical (*Z*) axis through the focal area.

An iris with a limiting diameter of 1 mm transmitting about 1% of the incident intensity was placed at a distance of 150 cm from the focal spot. The signal from FD-24K photodiode *3* was applied to V4-17 digital voltmeter *6*. To eliminate the effect of fluctuations in the laser output, we normalized the signal detected by photodiode *3* to the signal detected by photodiode *4*. The scheme with the limiting iris enables one to determine both the sign and the value of the nonlinear refractive index n_2 and Kerr nonlinear susceptibilities $\chi^{(3)}$ of the solutions. To determine the nonlinear absorption coefficient β, we took away diaphragm *7* and measured the transmittance of the solution vs. the position of the cuvette relative to the focal point (the open-iris scheme). To measure the transmittance vs. laser intensity, we placed the detector at a distance from the cuvette such that the entire transmitted radiation could be detected. In the open-iris experiments, the decrease in the transmittance was thus related to the nonlinear absorption of the metallorganic polymer. Particular attention was given to avoiding optical breakdown. The optical breakdown intensities for the polymer structures were found to be 4×10^{11} and 10^{11} W/cm^2 for the fundamental and second harmonics, respectively, whereas the working intensities were no more than 8×10^{10} and 10^{10} W/cm^2, respectively.

RESULTS AND DISCUSSION

Below, we present the experimental results for the nonlinear optical properties of cobalt-doped PVP obtained by the *Z*-scan method and in the process of optical limitation. Figure 3a shows data points for the normalized transmittance at a wavelength of 1064 nm for the variously doped metallorganic polymers. Each point is the result of averaging over five measurements. Figure 3b shows associated data obtained at 532 nm. The experiments were carried out at laser intensities of 4×10^{10} W/cm^2 ($\lambda = 1064$ nm) and 3×10^9 W/cm^2 ($\lambda = 532$ nm). The *Z*-scan technique allows for the immediate determination of the sign of the nonlinear refractive index and the character of the nonlinear process. In the given case, the sign is negative and self-defocusing is observed. Note the constancy of the sign and of the concentration dependence of the nonlinearity at the two wavelengths. It can be seen that the transmittance $T(Z)$ increases with cobalt concentration. This may be

related to stronger defocusing at higher concentrations of cobalt. The data for pure water and undoped PVP did not yield the typical nonlinear dependence $T(Z)$.

We neglected the thermal lens effect. It is known that this effect is associated with the propagation of an elastic wave. On the order of magnitude, the characteristic time taken to establish the steady-state density distribution is given by the ratio between the radius of the beam cross section (d) and the speed of sound in an insulator (ϑ_{sound}):

$$\tau = \frac{d}{\vartheta_{\text{sound}}}. \quad (1)$$

For our experimental conditions ($d = 75$ and 50 μm for the fundamental and second harmonics, respectively, and $\vartheta_{\text{sound}} = 1500$ m/s), the characteristic times are 50 and 33 ns, respectively. The pulse duration was shorter by three orders of magnitude (35 ps). Note also a rather low repetition rate of laser pulses (0.4 Hz) and, accordingly, the absence of the thermal lens, which is related to the cumulative effect. Hence, the effect observed depends on the Kerr nonlinearity.

To determine the nonlinear optical parameters of PVP ($\chi^{(3)}$ and n_2), we used the relationship [10]

$$\Delta T_{\text{p}-\text{v}} = 0.404(1 - S)^{0.25}\left|\frac{2\pi n_2 I[1 - \exp(-\alpha L)]}{\alpha\lambda}\right|, \quad (2)$$

where $\Delta T_{\text{p}-\text{v}}$ is the normalized peak-to-valley difference in the transmittances $T(Z)$ for the limiting-iris scheme, I_0 is the intensity at the focal spot, S is the transmittance of the iris (a fraction of the intensity detected by the photodiode), λ is the wavelength, L is the sample length, and α is the linear absorption coefficient.

The Kerr nonlinear susceptibility of the cobalt-doped PVP solution is represented as [12]

$$\chi^{(3)} = \frac{n_0 n_2}{3\pi}, \quad (3)$$

where n_0 is the linear refractive index of the aqueous solution of the metallorganic polymer. The nonlinear refractive index and the Kerr nonlinear susceptibility for the solution of PVP doped by cobalt with a concentration of 13.5% were found to be -3×10^{-12} and -4.7×10^{-13} esu, respectively.

In the open-iris scheme, all the samples exhibited nonlinear absorption at a wavelength of 532 nm (Fig. 4). No detectable nonlinear absorption was observed at 1064 nm in spite of the high intensity levels.

In the open-iris scheme, the normalized transmittance is given by [13]
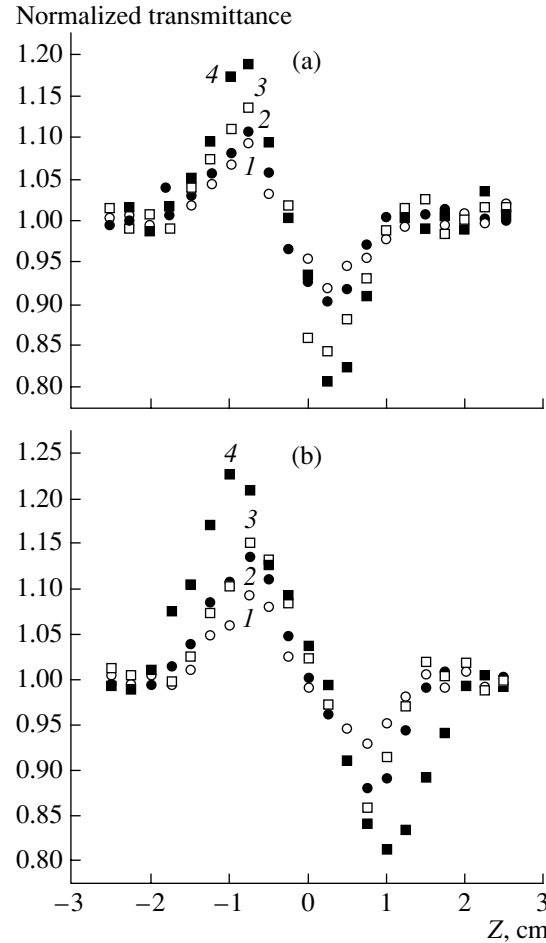
$$T(z) = 1 - \frac{q_0}{2\sqrt{2}} \quad (4)$$

Normalized transmittance

Fig. 3. Normalized transmittance at (a) $\lambda = 1064$ nm and $I_0 = 4 \times 10^{10}$ W/cm^2 and (b) $\lambda = 532$ nm and $I_0 = 3 \times 10^9$ W/cm^2 versus the position of the cuvette in the limiting-iris scheme for cobalt concentrations of (*1*) 2.0, (*2*) 5.3, (*3*) 6.2, and (*4*) 13.5%.

for $|q_0| < 1$, where

$$q_0 = \beta\frac{I_0[1 - \exp(-\alpha L)]}{\left(1 + \left(\frac{z}{z_0}\right)^2\right)\alpha}. \quad (5)$$

Here, $z_0 = k\omega_0^2/2$ is the diffraction length of the laser beam, $k = 2\pi/\lambda$ is the wave number, and ω_0 is the radius of the beam waist. Having determined the minimum normalized transmittance at the focal spot ($z = 0$) for the open-iris scheme and using expressions (4) and (5), we can find the nonlinear absorption coefficient. In each of the experiments, the position $z = 0$ was determined from the lowest normalized transmittance with allowance for the symmetry of the curve $T(Z)$ in this scheme. The decrease in the transmittance at 532 nm is likely to be related to inverse saturated absorption, which is typical of such metallorganic and polymer structures in the visible spectral range [14]. The nonlinear absorption in
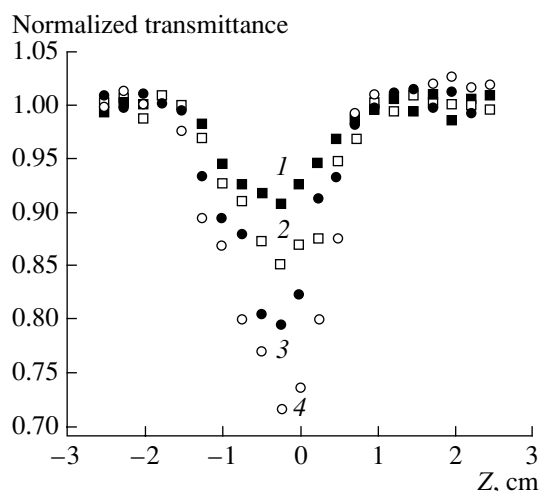
Normalized transmittance

Fig. 4. The same as in Fig. 3 at 532 nm for the open-iris scheme.
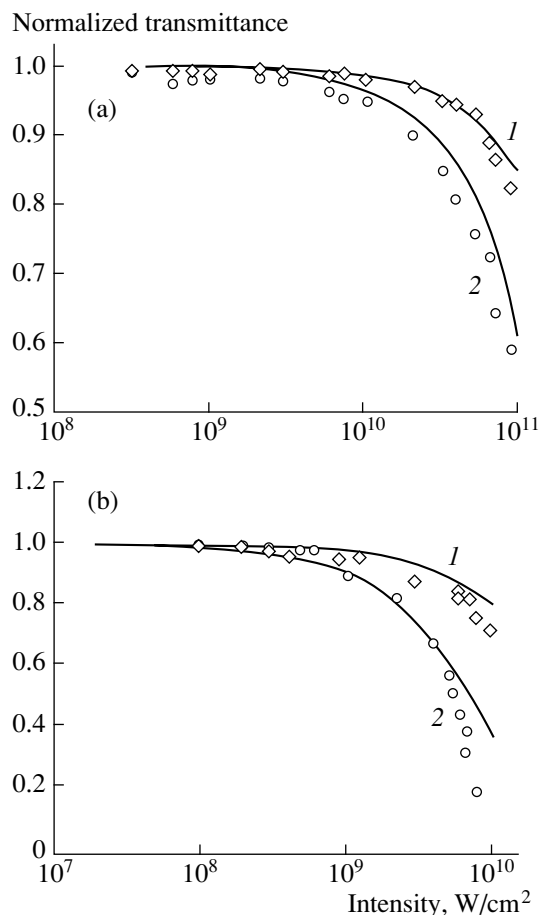
Normalized transmittance

Fig. 5. Normalized transmittance at (a) 1064 and (b) 532 nm versus the incident intensity for the solutions of the metal–PVP complexes with (*1*) 2.0 and (*2*) 13.5% of cobalt. Solid lines are theoretical curves.

the PVP solution with a cobalt concentration of 13.5% was 9.4×10^{-10} W/cm. Apparently, the wavelength 1064 nm does not fall into the two-photon absorption band,

which explains the absence of nonlinear (in particular, two-photon) absorption at this wavelength. The level of shot-to-shot fluctuations in the normalized transmittance was no greater than 3%. Hence, the upper limit for the nonlinear absorption at 1064 nm did not exceed 10^{-11} W/cm.

The table lists the nonlinear refractive indices, Kerr nonlinear susceptibilities, and nonlinear absorption coefficients for the solutions of the metallorganic polymers investigated.

Consider the data on optical limitation in PVP with various concentrations of cobalt. The effect of optical limitation is very promising for protecting eyes and sensitive detectors against high-intensity radiation. The mechanisms of optical limitation are varied. ISA due to high cross sections of absorption from excited levels is responsible for optical limitation in metal colloids [15], fullerenes, and phthalocyanines [16]. Intense nonlinear refraction gives rise to optical limitation in inorganic clusters [17]. Two-photon absorption is the reason for optical limitation in semiconductor structures [10]. As was mentioned, ISA is a highly effective nonlinear process. Tutt and McCahon [18] treated ISA in terms of a five-level model. A medium exhibits ISA if (i) the cross section of absorption from an excited level is greater than that from the ground state and (ii) the lifetime of the excited level is long (as compared with the excited pulse duration). Recently, it has been demonstrated that some metallorganic and organic substances, such as metallorganic clusters [19], porphyrins [8], metal phthalocyanines [20, 21], and fullerenes [11, 22, 23], meet these criteria. It would be natural to expect that the structures studied in this work also exhibit ISA (along with the other mechanisms of optical limitation).

We studied optical limitation at two wavelengths: 532 and 1064 nm. In spite of the absence of nonlinear absorption in the IR range, metal–PVP complexes demonstrate the pronounced defocusing effect. This suggests that optical limitation in the IR range is due to self-action (in particular, self-defocusing).

Figure 5a shows the normalized transmittance of the PVP solutions with various concentrations of cobalt versus the incident intensity for the limiting-iris scheme ($\lambda = 1064$ nm). The cuvettes were located in the area of the lowest transmittance, i.e., in the region corresponding to the valley in Fig. 3. Note good agreement with the theoretical curves constructed from the measurements of $\chi^{(3)}$.

At a wavelength of 532 nm, the situation is quite different (Fig. 5b). In this case, the optical limitation depends both on self-defocusing (which played the decisive role for the IR radiation) and, to a greater extent, on nonlinear absorption. The nonlinear absorption can be related to two-photon processes and ISA (at high intensities). The lack of data for the spectral characteristics and lifetimes of the excited singlet and triplet states of the metal complexes investigated gives no way of establishing a relationship between the contributions

Nonlinear optical parameters of metal–PVP complexes

Material	λ, nm	n_2, esu	$\chi^{(3)}$, esu	β, cm^{-1} W^{-1}
PVP (2% Co)	1064	-4.5×10^{-13}	-6.78×10^{-14}	–
PVP (5.3% Co)	1064	-7.3×10^{-13}	-1.1×10^{-13}	–
PVP (6.2% Co)	1064	-9.1×10^{-13}	-1.38×10^{-13}	–
PVP (13.5% Co)	1064	-1.3×10^{-12}	-1.97×10^{-13}	–
PVP (2.0% Co)	532	-7.6×10^{-13}	-1.18×10^{-13}	2.1×10^{-10}
PVP (5.3% Co)	532	-1.7×10^{-12}	-2.65×10^{-13}	4.2×10^{-10}
PVP (6.2% Co)	532	-2.3×10^{-12}	-3.59×10^{-13}	6.1×10^{-10}
PVP (13.5% Co)	532	-3.0×10^{-12}	-4.7×10^{-13}	9.4×10^{-10}

of these two processes. The theoretical curves, including only two-photon absorption, to some extent fit the experimental data. However, the experimental curves imply stronger optical limitation than the theoretical ones. We believe that this discrepancy is associated with the additional effect of ISA and, partially, by self-defocusing.

CONCLUSION

In experiments with picosecond laser pulses, we demonstrated that the nonlinear optical response of aqueous solutions of cobalt–polyvinylpyrrolidone complexes with various concentrations of the metal is related to the Kerr effect. Nonlinear absorption is absent in the IR range but plays a significant role in the visible part. The cobalt concentration in the complexes was shown to be the key parameter governing the nonlinear interaction of the picosecond radiation with the polymer. The reason for optical limitation at a wavelength of 1064 nm is self-defocusing. The mechanism behind optical limitation at 532 nm involves two-photon absorption and partially ISA and defocusing. The nonlinear optical parameters of the cobalt–polyvinylpyrrolidone complexes were measured.

REFERENCES

1. V. F. Lukinykh, S. A. Myslivets, A. K. Popov, *et al.*, Appl. Phys. B **38**, 143 (1985).
2. K. S. Aleksandrov, A. S. Aleksandrovskiĭ, S. V. Karpov, *et al.*, Dokl. Akad. Nauk SSSR **296**, 85 (1987).
3. J. S. Meth, H. Vanherzeele, and Y. Wang, Chem. Phys. Lett. **197**, 26 (1992).
4. D. Neher, G. I. Stegeman, and F. A. Tinker, Opt. Lett. **17**, 1491 (1992).
5. R. A. Ganeev, A. I. Ryasnyansky, N. V. Kamanina, *et al.*, J. Opt. B **3**, 88 (2001).
6. R. J. Knize and J. P. Partanen, Phys. Rev. Lett. **68**, 2704 (1992).
7. G. S. Pan, R. Kesavamoorthy, and S. A. Asher, J. Am. Chem. Soc. **120**, 6525 (1998).
8. F. M. Qureshi, S. J. Martin, X. Long, *et al.*, Chem. Phys. **231**, 87 (1998).
9. Y.-P. Sun, J. E. Riggs, H. W. Rollins, *et al.*, J. Phys. Chem. B **103**, 77 (1999).
10. M. Sheik-Bahae, A. A. Said, and E. W. Van Stryland, Opt. Lett. **14**, 955 (1989).
11. R. A. Ganeev, A. I. Ryasnyansky, M. K. Kodirov, *et al.*, Opt. Commun. **185**, 473 (2000).
12. J. F. Reintjes, *Nonlinear Optical Parametric Processes in Liquids and Gases* (Academic, New York, 1984).
13. M. Sheik-Bahae, A. A. Said, T. Wei, *et al.*, IEEE J. Quantum Electron. **26**, 760 (1990).
14. J. M. Perry, K. Mansour, S. R. Marder, *et al.*, Opt. Lett. **19**, 625 (1994).
15. W. Ji, H. J. Du, S. H. Tang, *et al.*, J. Opt. Soc. Am. B **12**, 876 (1995).
16. L. W. Tutt and A. Kost, Nature (London) **356**, 224 (1992).
17. Z. R. Chen, H. W. Hou, X. Q. Xin, *et al.*, Phys. Chem. **99**, 8717 (1995).
18. L. W. Tutt and S. W. McCahon, Opt. Lett. **15**, 700 (1990).
19. M. Pittman, P. Plaza, M. M. Martin, *et al.*, Opt. Commun. **158**, 201 (1998).
20. J. S. Shirk, R. G. S. Pong, F. J. Bartoli, *et al.*, Appl. Phys. Lett. **63**, 1880 (1993).
21. P. Zhu, P. Wang, W. Qiu, *et al.*, Appl. Phys. Lett. **78**, 1319 (2001).
22. M. Cha, N. S. Sariciftci, A. J. Heeger, *et al.*, Appl. Phys. Lett. **67**, 3850 (1995).
23. N. V. Kamanina, N. K. Kaporskii, and B. V. Kotov, Opt. Commun. **152**, 280 (1998).

Translated by A. Chikishev

OPTICS,
QUANTUM ELECTRONICS

Signal Recording in Acoustooptical Imaging
of Scattering Media

A. P. Solov'ev, M. I. Perchenko, Yu. P. Sinichkin, and O. V. Zyuryukina

Research Institute of Mechanics and Physics, Saratov State University,
Moskovskaya ul. 155, Saratov, 410026 Russia
e-mail: yasin@optics.sgu.ru
Received October 30, 2001

Abstract—An expression for a signal at the ultrasonic frequency from a photodetector arising when diffraction waves at the cathode mix is obtained for the case when the optical and acoustic beams are space-limited and Bragg angles are small. The alternating current generated when the scattering medium is crossed by a focused 3-MHz ultrasonic beam and illuminated by a cw He–Ne laser is measured. Satisfactory agreement between experimental results and those calculated from the formulas obtained indicates that our model treating acoustooptical interaction in the medium in terms of Raman–Nath diffraction is valid. Conditions for measuring the alternating current (which is a parameter of acoustooptical imaging) that are optimal from the viewpoint of maximizing the signal and signal-to-noise ratio are predicted theoretically and corroborated experimentally. © 2002 MAIK "Nauka/Interperiodica".

INTRODUCTION

Optical tomography relies on changes in the characteristics of light passed through a scattering medium. High-contrast and high-resolution optical imaging of inhomogeneities present in scattering media has been the subject of much investigation [1]. Acoustooptical imaging, a version of coherent optical imaging, is based on analysis of light passed through a medium crossed by an ultrasonic beam [2, 3]. In this case, the transmitted optical radiation is frequency-modulated at the ultrasonic frequency. The higher the ultrasonic frequency, the higher the spatial resolution of this method. However, a rise in the frequency enhances the attenuation of the acoustic wave in the medium. Therefore, in the deep (up to 10 cm) sounding of aqueous, including biological, media, ultrasonic frequencies of 1 to 10 MHz are used [4, 5] in order to avoid strong attenuation and, at the same time, provide a sufficiently high spatial resolution. The properties of the scattering medium sounded by crossed optical and acoustic beams can be analyzed by measuring the spatial distribution of the modulated transmitted light over the sounded area and by subsequently recovering the characteristics of the medium from this distribution.

As a cause of the modulation, this paper considers the diffraction of light by an elastic acoustic wave. Since the diffracted field is produced in the region where the optical and acoustic beams intersect, this field carries information on the state of the medium in this region and, therefore, on the presence and properties of optical inhomogeneities. At low acoustic frequencies, the diffraction angles are small and the diffracted beams overlap, as a result of which the unmodulated and diffracted signals are hard to spatially separate. In this case, to study the modulation characteristics of the light leaving the medium, it is convenient to use an alternating current with the ultrasonic frequency that arises from mixing diffracted waves in the photodetector. The phase difference between these fields becomes important, because the alternating current depends on experimental conditions.

In this paper, we show that the diffracted beams that are mixed on the photocathode produce an alternating current with an ultrasonic frequency as a result of heterodyning with the zero-order diffraction field, which is the most intense when the acoustic beam is weak, plays the role of the heterodyne. An expression for the alternating component of the photoelectric current at ultrasonic frequency is derived for the situation when the optical and acoustic beams are space-limited and the diffraction angles are small. Spatial amplitude–phase characteristics of the alternating component are studied theoretically and experimentally as a function of the photodetector position relative to the point of intersection between the acoustic and laser beams. Optimal conditions for observing the alternating component obtained theoretically are compared with the experiment at different acoustic frequencies. The possibility of using an external heterodyne is studied.

THEORY

Let an optical radiation propagating in the positive x direction (for the sake of definiteness) in a medium with a refractive index n_0 (ω and k_o are the frequency and wave number of the optical wave in the medium) that is crossed by an ultrasonic beam (Ω and K are the circular

frequency and wave number of the acoustic wave, respectively, and w is the beam width) directed in the positive z axis generate Raman–Nath diffraction waves when the condition $Q = wK^2/k_o \ll \pi/2$ [4] is satisfied. All diffraction waves whose orders differ by unity can contribute to the current at the ultrasonic frequency Ω as a result of their mixing at the photocathode of the detector. When the parameter of Raman–Nath diffraction is small, $v = n_0^2 pS_0 k_o w/2 < 1$ [4], the contribution of higher- order diffraction waves can be neglected. Here, p is the piezoelectric constant and S_0 is the strain amplitude caused by the acoustic wave. According to, for example, [4], at a small Bragg angle $\theta = K/2k_o \ll 1$, the zero- and first-order diffraction fields in a homogeneous optically transparent medium can be written as

$$E_n \sim E_L(-j)^n J_n(v)\exp j[(\omega + n\Omega)t$$
$$- (k_o x + nKz - |n|Kx\theta) + \varphi_o + n\varphi_a], \quad n = 0, \pm 1. \tag{1}$$

Here, E_L is the amplitude of the plane optical wave at the entrance to the acoustic beam, and φ_a and φ_o are the initial phases of the acoustic and optical fields, respectively. When the zero- and first-order waves mix, an elementary area $dzdy$ of the photocathode located perpendicularly to the x axis produces elementary photoelectric currents di_n at the ultrasonic frequency. They are proportional to the product of the fields $E_0 E_{+1}^* + E_0^* E_{+1}$ and $E_0 E_{-1}^* + E_0^* E_{-1}$ at the photocathode. Using expression (1) for a symmetric acoustic beam whose axis is aligned with the z axis, one can write the elementary currents as

$$di_n \sim -n J_0(v) J_1(v) E_L^2 \sin[\Omega t - Kz_0$$
$$- nK(L_0 + n_0 L)\theta + \varphi_a]dydz, \quad n = \pm 1. \tag{2}$$

Here, L_0 and L are the distances from the boundary of the medium to the acoustic beam axis in the medium and to the photodetector, respectively (Fig. 1). As can easily be seen from (2), the phase components $nK\theta(L_0 + n_0L)$ of the elementary currents have opposite signs. Therefore, the total elementary current $di = di_1 + di_{-1}$ depends significantly on the positions of the receiver, L, and the acoustic beam axis in the medium, L_0. In particular, when $K\theta(L_0 + n_0L) = (m-1)\pi$, where $m = 1, 2, \ldots$, the elementary currents are in antiphase and the total current is zero, while at $K\theta(L_0 + n_0L) = (2m-1)\pi/2$, the total current is maximal, because the elementary currents are in phase

Expressions (1) and (2) have been obtained for the optical field E_L with a constant amplitude in the plane perpendicular to the propagation direction. Actually, the optical field is space-limited and its amplitude varies in the cross section. Assume that the optical field is the field of the TEM_{00} mode of the laser radiation whose axis coincides with the x axis and a $d_Y \times d_Z$ rectangular aperture diaphragm is installed in front of the
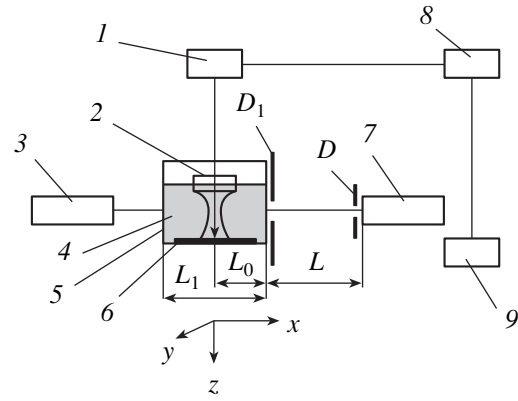

Fig. 1. Experimental setup: *1*, ultrasonic generator; *2*, ultrasonic transducer with a focusing lens; *3*, He–Ne laser; *4*, medium under study; *5*, cell; *6*, ultrasonic absorber; *7*, PMT photodetector; *8*, double-beam oscilloscope; and *9*, frequency-selective voltmeter.

photocathode. Then, the acoustic-frequency total current from the photodetector can be written as

$$i_\Omega \sim J_0(v) J_1(v) E_L^2 (a_0^2/a_L)(2/K)(\pi(A^2 + B^2))^{0.5}$$
$$\times \sin(Kd_Z/2)\exp[-(z_0^2 + y_0^2 + H^2/2)/a_L^2] \tag{3}$$
$$\times \cos[\Omega t - Kz_0 - \varphi(L, L_0, z_0) + \varphi_a],$$

where

$$A = 2\sinh(z_0 H/a_L^2)\cos\Phi; \quad B = 2\cosh(z_0 H/a_L^2)\sin\Phi;$$

$$\Phi = K\theta(Ln + L_0); \quad H = 2\theta[n_0 L + L_0];$$

$$\varphi = \arctan[\tan(z_0 H/a_L^2)\cot\Phi] \tag{4}$$

$$\text{when } 2\pi n \le \Phi \le (2n+1)\pi,$$

$$\varphi = \arctan[\tan(z_0 H/a_L^2)\cot\Phi] + \pi$$

$$\text{when } (2n+1)\pi \le \Phi \le 2(n+1)\pi, \quad n = 0, 1, 2, \ldots. \tag{5}$$

Here, $x_0 = L_0 + L$, y_0, and z_0 are the coordinates of the center of the aperture diaphragm; E_L and a_0 are the amplitude of the optical field at the axis and the beam radius near the plane exit mirror of the laser; and a_L is the radius of the laser beam at the distance L. The radius of the laser beam is determined at the point where the field amplitude decrease e times, L_a is the distance from the laser's exit mirror to the acoustic axis, and H is the distance between the centers of the zero- and first-order diffraction beams at the distance L.

Equation (3) is written for the case when the acoustic wave power is constant over the region of intersection between the acoustic wave and laser beam, the dimensions of the diaphragm meet the conditions $d_Z \ll a_L$ and $d_Y \ge a_L$, and the amplitudes of the diffracted beams vary with z much more slowly than Kz_0. There-
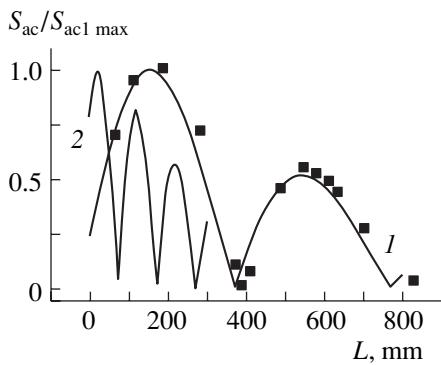
$S_{ac}/S_{ac1\,max}$

Fig. 2. Experimental (symbols) and calculated (solid lines) normalized amplitude of the ultrasonic-frequency alternating current, S_{ac}, versus position of the photodetector on the axis of the laser beam at $\Omega = (1)$ 3 and (2) 6 MHz.
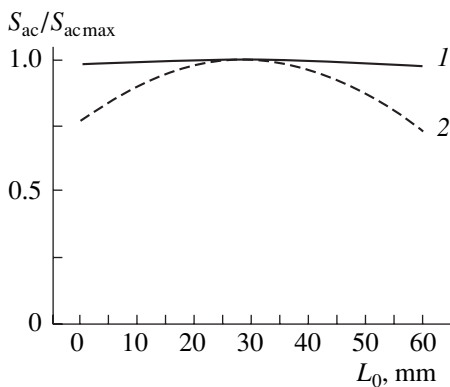
$S_{ac}/S_{ac\,max}$

Fig. 3. Normalized amplitude of the ultrasonic-frequency alternating current, S_{ac}, versus sounding depth L_0 at $\Omega = (1)$ 3 and (2) 6 MHz.
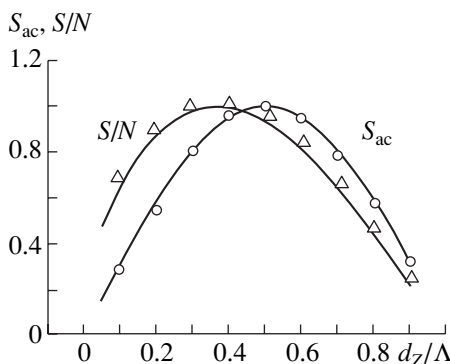
$S_{ac}, S/N$

Fig. 4. Normalized amplitude of the ultrasonic-frequency alternating current, S_{ac}/S_{acmax}, and signal-to-noise ratio, S/N, versus relative aperture d_Z/Λ.

fore, in this expression, the integral over d_Y was set equal to $a_L\sqrt{\pi}$ with a high accuracy and the field amplitudes were assumed to be constant within d_Z.

As follows from expressions (3)–(5), the amplitude and phase of the alternating current at the ultrasonic fre-

quency depend on the position L of the photodetector on the laser beam axis, sounding depth L_0, position of the aperture diaphragm center on the yz plane, relative dimension Kd_Z of the aperture diaphragm, and parameters of the optical and acoustic radiation (E_{L0}, v, k_o, K, θ). At the laser beam axis, expression (3) for the current takes the simpler form

$$i_\Omega \sim E_L^2 J_0(v) J_1(v) [\exp(-H^2/2a_L^2)] a_L^{-1} \sin(Kd_Z/2) \\ \times \sin[K(L_0 + n_0 L)\theta] \cos(\Omega t - Kz_0 + \varphi_a). \tag{6}$$

Figures 2–6 compare the results of calculations by formulas (3)–(6) with experimental values. Expressions (3)–(6) for the alternating current at the ultrasonic frequency were obtained for a nonscattering medium. Since the intensity of collimated (ballistic) light propagating through a scattering medium of length L_1 along the optical radiation axis varies as $\exp(-\mu L_1)$, where μ is the extinction coefficient, the alternating photoelectric current associated with the ballistic light in the scattering medium can be written as

$$i_{\Omega\mu} = \exp(-\mu L_1) i_\Omega(\mu = 0).$$

EXPERIMENTAL

The experimental setup is shown in Fig. 1. The medium under study was an aqueous solution of milk. The liquid was placed in a rectangular organic-glass cell. Its dimension along the laser beam axis was $L_1 = 62$ mm. An ultrasonic transducer with a focusing lens immersed in the solution excited a 3-MHz traveling wave in the z direction. The diameter of the transducer was 23 mm. The acoustic wavelength Λ, focal length, and diameter of the focal spot in water were 0.5, 45, and 1.4 mm, respectively. To suppress reflection from the cell walls, an acoustic absorber was placed in front of the emitter. The radiation of a 25-mW laser operating in the TEM_{00} mode at a wavelength $\lambda = 632.8$ nm crossed the focal region of the acoustic beam perpendicular to its axis (the half-power laser beam diameter at the point of intersection with the acoustic beam was approximately 1.2 mm). The light transmitted through the cell was recorded by a photomultiplier tube (PMT). The surface of the photocathode and the cell walls were perpendicular to the laser beam axis. The aperture diaphragm D was mounted directly on the photocathode. Its dimension d_Y was 1 mm, while the dimension d_Z along the acoustic axis was variable. A 4×4-mm square diaphragm D_1 placed on the outside surface of the cell wall on the photodetector side served to reduce the background illumination falling into the PMT. The ultrasonic transducer and the PMT, together with the aperture diaphragm, could be moved independently along and across the laser beam axis and parallel to the acoustic beam axis. A double-beam oscilloscope was used to adjust the components of the setup and to record phase variations in the alternating current. A frequency-selective voltmeter measured the alternating current

and noise in a 1-kHz-wide band centered at 3 MHz. To set the linear mode of photodetector operation at a small scattering parameter μL_1 of the medium, calibrated neutral color filters (not shown in Fig. 1) were used.

Under the given experimental conditions, the parameter $Q = 0.08$ was much smaller than $\pi/2$ (the width w of the acoustic beam was taken equal to the diameter of the focal spot), the Raman–Nath diffraction parameter $v \cong 0.4$ was less than unity, and the Bragg angle ($\theta = 0.036°$) was very small. In the plane of observation perpendicular to the laser beam, a series of Raman–Nath diffraction spots were present at a distance of more than 2 m; at the same time, at a distance of less than 800 mm, first-order diffraction maxima were within the zero-order diffraction spot.

To see how the amplitude and phase of the alternating current, as well as the value of the direct current, depend on the geometry of the experiment, the photodetector was moved along the laser beam and parallel to the acoustic beam, and the aperture diaphragm was varied at fixed positions of the laser and acoustic emitter. The alternating and direct signal components and the background noise were recorded in each position of the PMT and for all dimensions of the aperture diaphragm. To measure the phase variations in the alternating signal as a function of PMT position, this signal was applied to one of the inputs of the double-beam oscilloscope and compared with the reference signal from the acoustic oscillator applied to the other input. The effect of the scattering parameter on the alternating and direct current components was estimated by measuring the concentration of milk in the cell, which was gradually varied. The current was recorded at each PMT position and milk concentration.

RESULTS AND DISCUSSION

The experimental study of the amplitude S_{ac} and phase F of the alternating current as a function of the geometry of the experiment was conducted in order to corroborate the above model and find the maximum condition for the signal and signal-to-noise ratio.

Figure 2 shows the experimental (dots) normalized amplitude of the alternating current versus the position L of the photodetector on the laser axis (with the acoustic beam axis in the cell) at $L_0 = 37$ mm. In this figure, the normalized amplitudes of the alternating current calculated from formula (6) at ultrasonic frequencies of 3 and 6 MHz for the same L_0 and initial radius of the laser beam $a_0 = 0.5$ mm are also plotted. The theoretical amplitudes of the current at 3 MHz are in good agreement with the experimental data for the same ultrasonic frequency, which validates the model proposed. The functions shown in Fig. 2 are seen to behave as the absolute value of a decaying sinusoidal function. The amplitude of the alternating current vanishes at $L_0 + n_0 L = (m-1)\pi/K\theta$, where $m = 1, 2, \ldots$. The positions of
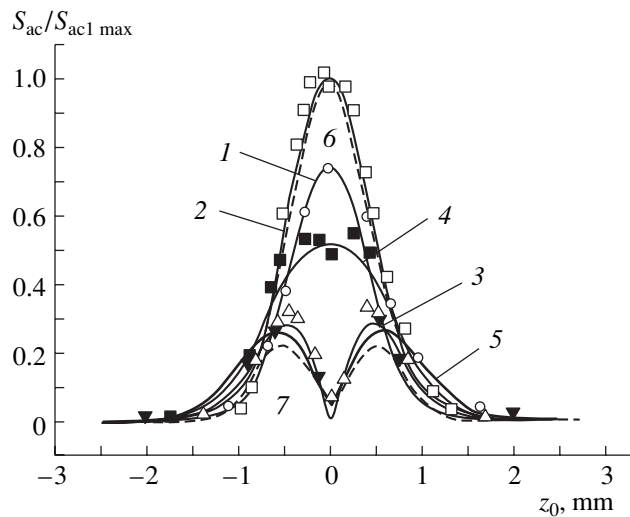

Fig. 5. Normalized amplitude of the ultrasonic-frequency alternating current versus displacement of the photodetector parallel to the acoustic axis at $L_0 = 37$ mm and $L = 1$ (○) 67, 2 (□) 175, 3 (△) 375, 4 (■) 570, and 5 (▼) 750 mm. Solid lines, calculation for $\Omega = 3$ MHz; dashed lines 6 and 7, calculation for $\Omega = 6$ MHz; and symbols, data points.
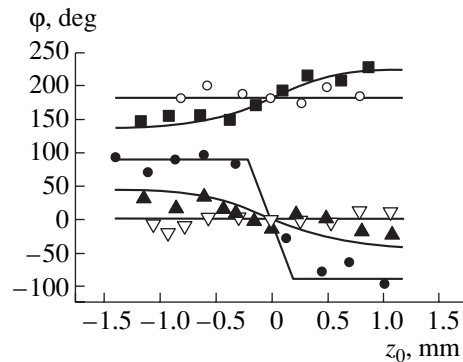

Fig. 6. Theoretical (solid lines) and experimental (symbols) phase φ of the ultrasonic-frequency signal versus displacement of the photodetector parallel to the acoustic axis at $\Omega = 3$ MHz, $L_0 = 37$ mm, and $L = $ (▽) 175, (▲) 280, (●) 375, (■) 475, and (○) 575 mm.

the maxima, $L = L_{m\max}$, can approximately be found from the expression $Ln_0 + L_0 = (2m - 1)\pi/2K\theta$. The period of the sinusoidal functions varies inversely with the square of the ultrasonic frequency. The decrease in the amplitudes of the second and next maxima is attributed to the divergence of the diffracted beams and to an increase in their center distance H with L. At an ultrasonic frequency of 6 MHz, the maxima decrease faster because the Bragg angle is larger.

Figure 3 shows the amplitude S_{ac} of the alternating signal versus L_0 calculated for the same ultrasonic frequencies. Note that the distance between the acoustic beam axis and cell wall, L_0 (Fig. 1), specifies the sounding depth in acoustooptical imaging. This is because the

diffraction waves, which carry information on the optical parameters of the medium, come to the cathode of the photodetector from the acoustic beam region. The value of L_0 can vary from zero to L_1, where L_1 is the dimension of the medium along the laser axis. In our calculations, $L = L_{1\max}$, which corresponds to the first maximum of the current amplitude in expression (6) when the acoustic axis is at the center of the cell ($L_0 = 31$ mm). Note that the distance between the cell and the photodetector, $L_{1\max}$, varied for different ultrasonic frequencies: approximately 180 and 25 mm for 3 and 6 MHz, respectively. As can be seen from Fig. 3, at the lower frequency, the displacement of the ultrasonic emitter along the laser axis from 0 to 60 mm at a fixed position of the photodetector (parameter L) changes the signal insignificantly (by about 2%). At 6 MHz, the same displacement of the region studied significantly (by more than 30%) changes the signal, which should be taken into account in practice. At the same time, the resolution in the x and y directions improves with frequency, because it is proportional to ultrasonic frequency. One should bear in mind that the acoustic attenuation in the medium, which was ignored in our calculations, grows in proportion with the square of the ultrasonic frequency.

The amplitude of the ultrasonic-frequency current in expressions (3) and (6) varies with the relative aperture Kd_Z as a sine function; therefore, it is maximal at $d_Z = (2n-1)\Lambda/2$, where $n = 1, 2, \ldots$. The diaphragm minimal dimension d_Z (at $n = 1$) corresponding to the maximal alternating signal is $d_Z = 0.5\Lambda$. In practice, one has to extract the information-carrying signal from ultrasonic-frequency noise in the receiver bandwidth Δf. It has been shown [5] that the noise current N is proportional to the square root of the product of the photoelectric current direct component S_{dc} by the bandwidth Δf; i.e., $N \sim (S_{dc}\Delta f)^{0.5}$. This means that the noise current is defined mostly by the shot noise of the photocathode due to its illumination by the laser (including scattered) radiation and stray light. The direct component of the photoelectric current, as well as the cathode illumination of the cathode, varies in proportion to the area $d_Z d_Y$ of the aperture diaphragm. Accordingly, at a constant diaphragm dimension in the y axis, d_Y, the shot current N is proportional to $\sqrt{d_Z}$. Thus, the signal-to-noise ratio S/N varies as the function $\sin(0.5Kd_Z)/\sqrt{d_Z}$, which has a maximum at $d_Z \cong 0.375\Lambda$. Figure 4 shows the calculated and experimental normalized amplitude of the alternating signal, S_{ac}, and the signal-to-noise ratio S/N obtained for pure water when the photodetector was placed at $L = L_{1\max}$ on the laser beam axis with $d_Y = 1$ mm. The results are seen to be in good agreement. Since the experiment was performed at an acoustic wavelength of $\Lambda = 0.5$ mm, the dimensions of the aperture diaphragm corresponding to the maximum signal and maximum signal-to-noise ratio were $d_Z = 0.25$ mm

and $d_Z \cong 0.19$ mm, respectively. The experimental results shown in Figs. 2–6 were obtained at $d_Z \cong 0.19$ mm.

The experimental dependence of the alternating current amplitude on the y position of the PMT had a Gaussian shape and coincided with the theoretical one. The curve is omitted because of its clearness.

Figures 5 and 6 show the measured and calculated amplitudes S_{ac} and phase variations F of the alternating signal versus z coordinate of the aperture diaphragm center, z_0, for several PMT positions on the laser axis, L, and the sounding depth $L_0 = 37$ mm. The theoretical and experimental results at 3 MHz are in good qualitative agreement (Fig. 5). The dashed lines in Fig. 5 plot the distributions of the alternating current amplitude for two positions of the photodetector corresponding to the first maximum (curve 6) and first minimum (curve 7) at an ultrasonic frequency of 6 MHz. The shape of the distribution is nearly the same as at 3 MHz because it depends primarily on the laser radiation parameters. The characteristic values of L are the only quantities that change. The amplitude distributions signal parallel to the acoustic axis are associated with the finite cross section of the laser beam.

According to expression (3), when the photodetector position corresponds to the maxima of the alternating current amplitude, $L = L_{n\max}$, the total phase variation of the alternating current, $F = (-Kz_0 - \varphi)$, is a linear function of z because, in this case, $\varphi(z_0, L_{n\max}) = 0$. In the experiments with $L = 175$ and 575 mm, which correspond to the first and second maxima of the alternating current amplitude, the phases φ are other than zero. They can be well approximated by the function $\Delta K z_0$ for $|\Delta K|/K < 0.05$. The nonzero value of $|\Delta K|$ is associated with the shift of the laser spot center along the acoustic axis during the measurements, i.e., with the change in the initial phase φ_a of the acoustic wave. This introduces an additional, positive or negative, phase shift into the signal relative to the reference. Figure 6 plots experimental functions $[\varphi(z_0 L) - \Delta K z_0]$ vs. z for several L. In such a representation, the experimental and theoretical results are in good agreement.

As follows from expression (3) for the alternating current, the signal can be increased by using multislit aperture diaphragms for wide laser beams. The position z_0 of the center of the nth slit should be found from the equation $[Kz_n + \varphi(z_n, L)] = 2\pi n$ [where $n = 0, 1, 2, \ldots$ and $\varphi(z_n, L)$ is defined by expressions (4) and (5)] for each particular position L of the photodetector relative to the exit face of the cell and for each sounding depth L_0. At $L = L_{n\max}$, the phase variation $\varphi = 0$ (these positions correspond to the maximal amplitudes of the current i_Ω), and the center distance between adjacent slits is exactly equal to the acoustic wavelength Λ.

The experimental data reported in this paper were obtained for pure (transparent) water. The dependences of the alternating signal on the PMT coordinates (L and

z) and on the dimension of the aperture diaphragm at milk concentrations *C* from 0 to 0.7%, which corresponds to the scattering parameter μL_1 of 0 to 17.5, were similar to those obtained for pure water and are therefore omitted. The alternating current amplitude at the first maximum varies with the scattering parameter approximately as $\exp(-\mu L_1)$ up to $\mu L_1 = 17.5$. This is apparently an indication that the alternating current is the ballistic component of the laser radiation.

With a further increase in the milk concentration, the spatial characteristics of the alternating current change noticeably. In particular, at $C = 1.1\%$, the current amplitude as a function of *L* becomes proportional to $1/(L + L_0)$, unlike (3). The dependence of the alternating current on the *z* position of the photodetector also changes. These changes in the behavior of the alternating current are associated with the scattered (nonballistic) component of the laser radiation.

To increase the ballistic component of the current, one can use an external heterodyne, for example, tap a portion of the laser radiation and apply it to the photocathode [2]. Let E_h be the field amplitude of the external heterodyne and φ_h as $\varphi_h = \varphi_o + \Delta\varphi$ be its initial phase. Then, the amplitude of the current in expression (3) will increase approximately $E_h/E_0 \exp(-\mu L_1)$ times and the parameter Φ will change to become $\Phi_h = K\theta(Ln + L_0) - \Delta\varphi$. At the laser axis, the amplitude of the alternating current will be proportional to $\sin(K\theta(Ln + L_0) - \Delta\varphi)$. It is easy to check that, by varying the optical path length of the heterodyne wave, one can find $\Delta\varphi$ which maximizes the amplitude of the alternating current irrespective of the position *L* of the photodetector on the laser axis and sounding depth L_0. At a fixed position *L* of the photodetector and a variable sounding depth L_0, the maximum condition for the signal S_{ac} can be provided by applying a control signal from the ultrasonic emitter to a shifter that changes the phase of the heterodyne signal during the scanning of the medium.

Note that the experiments described above determined the amplitude of the alternating current from one measurement (i.e., without averaging) at a signal-to-noise ratio of ≥1. The powers of the optical and acoustic emitters were far from ultimate for biological tissue examination. Also, since the ballistic component is of regular nature, it can be recorded with signal accumulation methods. This could increase the contribution of the ballistic component to the recorded signal S_{ac} and thereby raise μL_1 with the signal-to-noise ratio remaining unchanged.

CONCLUSION

Experimental results for a nonscattering medium and theoretical results obtained under the assumption that the optical and acoustic beams are space-limited and the Bragg angle is small qualitatively agree. This indicates that the model explaining the generation of the ultrasonic-frequency alternating current component as a result of mixing Raman–Nath diffraction waves on the photocathode of the detector is valid. The characteristics of the alternating current were studied theoretically and experimentally only for small diffraction parameters $v < 1$, which allowed us to neglect higher diffraction orders when theorizing. This restriction is well justified because v is usually less than unity for acoustic beams used in diagnostics. For a scattering medium, the experimental amplitude and phase characteristics do not differ from the theoretical ones for scattering parameters up to $\mu L_1 \leq 17.5$. Hence, our model of interaction can be extended to scattering media with the above range of scattering parameter. Both the theory and the experiment show that the amplitude and phase of the ultrasonic-frequency current depend significantly on the position of the photodetector outside and that of the acoustic beam axis inside the medium, the dimension and shape of the aperture diaphragm, and the parameters of the acoustic and optical radiations. Positions of the photodetector that maximize the alternating current were found. The dimensions of the aperture diaphragm that provide the maximal signal and signal-to-noise ratio were found theoretically and corroborated experimentally. The feasibility of dimension optimization to maximize *S/N* is particularly important for measurements in scattering media, where the desired signal may become very weak. The use of different ultrasonic frequencies and an external heterodyne in measurements of the alternating current was considered. The results reported in this paper can be useful in developing a recording system and in choosing optical and acoustic sources for the acoustooptical imaging of absorbing objects in scattering media.

ACKNOWLEDGMENTS

REFERENCES

1. V. V. Tuchin, *Lasers and Fiber Optics in Biomedical Research* (Saratov. Gos. Univ., Saratov, 1998).
2. M. Kempe, M. Larionov, D. Zaslavsky, *et al.*, J. Opt. Soc. Am. **4**, 1151 (1987).
3. L. Wang, Photochem. Photobiol. **67**, 41 (1998).
4. G. S. Kino, *Acoustic Waves: Devices, Imaging, and Analog Signal Processing* (Prentice-Hall, Englewood Cliffs, 1987; Mir, Moscow, 1990).
5. M. I. Perchenko, Yu. P. Sinichkin, A. P. Solov'ev, *et al.*, Proc. SPIE **4002**, 189 (2000).

Translated by A. Khzmalyan