Technical Physics, Vol. 46, No. 7, 2001, pp. 783–788. Translated from Zhurnal Tekhnicheskoĭ Fiziki, Vol. 71, No. 7, 2001, pp. 1–6.

The Fundamental Solution to Equations of Linear Magnetic Hydrodynamics in a Moving Medium

A. A. Aleksandrova* and Yu. N. Aleksandrov**

** Kharkov Military University, Kharkov, 61043 Ukraine*

*** Kharkov Technical University of Radio Electronics, Kharkov, 61726 Ukraine*

e-mail: et@kture.kharkov.ua

Received July 17, 2000

Abstract—The fundamental solution to a system of linear differential equations of magnetic hydrodynamics in a moving medium is obtained. Using the Fourier–Laplace transform, the Green tensor function is calculated as a sum of dyadics. In this way, the integral equations of magnetic hydrodynamics can easily be derived. Particular forms of the fundamental solution that are important in applications are analyzed. © *2001 MAIK "Nauka/Interperiodica".*

INTRODUCTION

Various problems of magnetic hydrodynamics of moving media have attracted much attention recently. The points of interest include the following: the frequency multiplication; the magnetohydrodynamic wave amplification during its reflection by a moving boundary; the diagnostics of moving plasma media by analyzing their interaction with magnetohydrodynamic waves; and the production, reflection, and refraction of such waves in the presence of moving magnetosphere layers, solar wind, etc. Considering a moving medium leads to substantial mathematical difficulties in the equations of magnetic hydrodynamics, which combine the medium velocity **U**, magnetic field strength **B**, density ρ, and entropy.

In this paper, we extend the studies started in [1], where velocity **u** initiated only by perturbed wave motion was considered, and in [2], where an unperturbed motion of an inhomogeneity was additionally taken into account. Those papers are based on the use of the Green function found for a magnetohydrodynamic medium without initial motion. Now, we try to take into consideration the unperturbed motion of an external medium while calculating this Green function. Given this function, we can use the standard method of summing disturbances created by each point of the source to construct the general solution of a boundary-value problem as the convolution of the fundamental solution and the right-hand side of linear magnetohydrodynamic differential equations. Applying the Laplace–Fourier transform, we derive the fundamental solutions to the equations with constant coefficients associated with the boundary-value problem. The problem is stated in a generalized form with boundary and initial conditions included in the instant sources.

STATEMENT AND SOLUTION OF THE PROBLEM

Consider small perturbations in a plasma medium treated as magnetohydrodynamic with nonzero initial velocity $\mathbf{U}_1$, constant density ρ_1, and magnetic field $\mathbf{B}_1$. Then, the classical linear equations of magnetic hydrodynamics are valid [1]. The Green function, or the fundamental solution, of the boundary-value problem satisfies the following system of differential equations with the δ-like right-hand side

$$\begin{aligned}
&\frac{\partial \mathbf{u}}{\partial t} + (\mathbf{U}_1\nabla)\mathbf{u} + \frac{V_{s1}}{\rho_1}\nabla\tilde{\rho} + \frac{1}{4\pi\rho_1}[\mathbf{B}_1\mathrm{rot}\mathbf{b}] \\
&\quad = \mathbf{S}'_u\delta(\mathbf{r}-\mathbf{r}')\delta(t-t'), \\
&\frac{\partial \mathbf{b}}{\partial t} - \mathrm{rot}[\mathbf{U}_1\mathbf{b}] + \mathrm{rot}[\mathbf{B}_1\mathbf{u}] = \mathbf{S}'_b\delta(\mathbf{r}-\mathbf{r}')\delta(t-t'), \\
&\frac{\partial \tilde{\rho}}{\partial t} + \rho_1\mathrm{div}\mathbf{u} + \mathrm{div}(\tilde{\rho}\mathbf{U}_1) = \mathbf{S}'_\rho\delta(\mathbf{r}-\mathbf{r}')\delta(t-t'),
\end{aligned} \tag{1}$$

where **u**, **b**, and $\tilde{\rho}$ are the deviations of velocity, magnetic field, and density from their equilibrium values $\mathbf{U}_1$, $\mathbf{B}_1$, and ρ_1; V_{s1} is the velocity of sound; $\mathbf{S}'_u$ and $\mathbf{S}'_b$ are arbitrary constant vectors; and $\mathbf{S}'_\rho$ is an arbitrary constant. Two vectors in parentheses and brackets mean scalar and vector products.

We denote the velocity, magnetic field, and density obeying system (1) by $\hat{\Gamma}_u$, $\hat{\Gamma}_b$, and $\hat{\Gamma}_\rho$ and write them in the form [3]

$$\begin{aligned}
\hat{\Gamma}_u(\mathbf{r}-\mathbf{r}', t-t') &= \mathbf{G}_{u\rho}(\mathbf{r}-\mathbf{r}', t-t')S'_\rho \\
&+ \hat{\mathbf{G}}_{uu}(\mathbf{r}-\mathbf{r}', t-t')\mathbf{S}'_u + \hat{\mathbf{G}}_{ub}(\mathbf{r}-\mathbf{r}', t-t')\mathrm{S}'_b,
\end{aligned}$$

1063-7842/01/4607-0783$21.00 © 2001 MAIK "Nauka/Interperiodica"

$$\hat{\Gamma}_b(\mathbf{r}-\mathbf{r}', t-t') = \mathbf{G}_{b\rho}(\mathbf{r}-\mathbf{r}', t-t')S'_\rho + \hat{\mathbf{G}}_{ub}(\mathbf{r}-\mathbf{r}', t-t')S'_u + \hat{\mathbf{G}}_{bb}(\mathbf{r}-\mathbf{r}', t-t')S'_b,$$
$$\hat{\Gamma}_\rho(\mathbf{r}-\mathbf{r}', t-t') = G_{\rho\rho}(\mathbf{r}-\mathbf{r}', t-t')S'_\rho + \mathbf{G}_{\rho u}(\mathbf{r}-\mathbf{r}', t-t')S'_u + \mathbf{G}_{\rho b}(\mathbf{r}-\mathbf{r}', t-t')S'_b, \quad (2)$$

where $\hat{\mathbf{G}}_{uu}$, $\hat{\mathbf{G}}_{ub}$, $\hat{\mathbf{G}}_{bu}$, and $\hat{\mathbf{G}}_{bb}$ are second-rank tensors; $\mathbf{G}_{\rho u}$, $\mathbf{G}_{u\rho}$, $\mathbf{G}_{\rho b}$ and $\mathbf{G}_{b\rho}$ are first-rank tensors; and $G_{\rho\rho}$ is a zero-rank tensor. Formula (2) is valid because any tensor can be expanded into a sum of three dyads [4], formally written as scalar products ($\mathbf{a}\cdot\mathbf{b}$) on the right-hand side of (2).

The solutions $\hat{\Gamma}_u$, $\hat{\Gamma}_b$, and $\hat{\Gamma}_\rho$ are sought in the form of the Fourier–Laplace transform with unknown weight functions $\hat{\mathbf{g}}_u(q, \mathbf{p})$, $\hat{\mathbf{g}}_b(q, \mathbf{p})$, and $\hat{g}_\rho(q, \mathbf{p})$ as follows

$$\begin{Bmatrix}\hat{\Gamma}_u(\mathbf{r}-\mathbf{r}', t-t')\\ \hat{\Gamma}_b(\mathbf{r}-\mathbf{r}', t-t')\\ \hat{\Gamma}_\rho(\mathbf{r}-\mathbf{r}', t-t')\end{Bmatrix} = \frac{1}{(2\pi)^4}\int_c dq\int_\infty \exp[-iq(t-t')] + i\mathbf{p}(\mathbf{r}-\mathbf{r}')\begin{Bmatrix}\hat{\mathbf{g}}_u(q, \mathbf{p})\\ \hat{\mathbf{g}}_b(q, \mathbf{p})\\ \hat{g}_\rho(q, \mathbf{p})\end{Bmatrix}d\mathbf{p}. \quad (3)$$

Here, q varies within the infinite limits on a line parallel to the real axis (Req) in the complex plain. Substituting (3) into (1), we get the following equations for the weight functions

$$\{(\mathbf{U}_1\mathbf{p}) - q\}\hat{\mathbf{g}}_u + V_{s1}^2\mathbf{p}\hat{g}_\rho + \frac{1}{4\pi\rho_1}[\mathbf{B}_1[\mathbf{p}\hat{\mathbf{g}}_b]] = -i\mathbf{S}_u,$$
$$\{(\mathbf{U}_1\mathbf{p}) - q\}\hat{\mathbf{g}}_b - \mathbf{U}_1(\mathbf{p}\hat{\mathbf{g}}_b) + [\mathrm{p}[\mathbf{B}_1\hat{\mathbf{g}}_u]] = -i\mathbf{S}_b, \quad (4)$$
$$\{(\mathbf{U}_1\mathbf{p}) - q\}\hat{g}_\rho + \rho_1(\mathbf{p}\hat{\mathbf{g}}_u) = -iS_\rho.$$

After some transformations, the equation for $\hat{\mathbf{g}}_u$ takes the form

$$\hat{\mathbf{g}}_u((\mathbf{U}_1\mathbf{p} - q)^2 - V_{A1}^2(\mathbf{s}_1\mathbf{p})^2) - (V_{s1}^2 + V_{A1}^2)\mathbf{p}(\mathbf{p}\hat{\mathbf{g}}_u) + V_{A1}^2\mathbf{p}(\mathbf{s}_1\hat{\mathbf{g}}_u)(\mathbf{s}_1\mathbf{p}) + V_{A1}^2\mathbf{s}_1(\mathbf{s}_1\mathbf{p})(\mathbf{p}\hat{\mathbf{g}}_u) \quad (5)$$
$$= \hat{\varepsilon}_u(\mathbf{U}_1\mathbf{p} - q) - \frac{V_{s1}^2\mathbf{p}\varepsilon_\rho}{\rho_1} + \{(\mathbf{p}\hat{\varepsilon}_b)[\mathbf{B}_1[\mathbf{p}\mathbf{U}_1]] - [\mathbf{B}_1[\mathbf{p}\hat{\varepsilon}_b]]\},$$

where $V_{A1}^2 = B_1^2/4\pi\rho_1$ is the Alfven speed and $\mathbf{s}_1 = \mathbf{B}_1/B_1$.

Let us introduce the basis $\langle\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\rangle$ related to a chosen direction of the unperturbed magnetic field $\mathbf{s}_1$ such that $\mathbf{e}_2 = \mathbf{s}_1$. This explicitly expresses the fact that the medium is anisotropic with respect to the unperturbed magnetic field. In the new basis, Eq. (5) takes the matrix form

$$\hat{A}\hat{\mathbf{g}}_u = \hat{\Omega}, \quad (6)$$

where

$$\hat{A} = \begin{pmatrix} q'^2 - V_{A1}^2(\mathbf{s}_1\mathbf{p})^2 - (V_{s1}^2 + V_{A1}^2)p_1 & -V_{s1}^2p_1p_2 & -(V_{s1}^2 + V_{A1}^2)p_1p_3\\ -V_{s1}^2p_1p_2 & q'^2 - V_{s1}^2p_2^2 & -V_{s1}^2p_2p_3\\ -(V_{s1}^2 + V_{A1}^2)p_3p_1 & -V_{s1}^2p_2p_3 & q'^2 - V_{A1}^2(\mathbf{s}_1\mathbf{p})^2 - (V_{s1}^2 + V_{A1}^2)p_3^2\end{pmatrix},$$

$$\hat{\Omega} = \begin{pmatrix} \varepsilon_{u1}q' - \frac{V_{s1}^2\varepsilon_\rho p_1}{\rho_1} + \frac{1}{4\pi\rho_1}[\varepsilon_{b1}p_2 - \varepsilon_{b2}p_1 + (\mathbf{p}\hat{\varepsilon}_b)(p_1\mathbf{U}_1\mathbf{B}_1 - U_{11}p_2B_1)]\\ \varepsilon_{u2}q' - \frac{V_{s1}^2\varepsilon_\rho p_2}{\rho_1} + \frac{1}{4\pi\rho_1}[\varepsilon_{b2}p_2 - \varepsilon_{b2}p_1 + (\mathbf{p}\hat{\varepsilon}_b)(p_2\mathbf{U}_1\mathbf{B}_1 - U_{12}p_2B_1)]\\ \varepsilon_{u3}q' - \frac{V_{s1}^2\varepsilon_\rho p_3}{\rho_1} + \frac{1}{4\pi\rho_1}[\varepsilon_{b3}p_2 - \varepsilon_{b2}p_3 + (\mathbf{p}\hat{\varepsilon}_b)(p_3\mathbf{U}_1\mathbf{B}_1 - U_{13}p_2B_1)]\end{pmatrix},$$

$q' = \mathbf{U}_1\mathbf{p} - q$, $\hat{\mathbf{g}}_u = \{g_{u1}, g_{u2}, g_{u3}\}$, $\hat{\mathbf{g}}_b = \{g_{b1}, g_{b2}, g_{b3}\}$, $\mathbf{p} = \{p_1, p_2, p_3\}$, and $\mathbf{U}_1 = \{U_{11}, U_{12}, U_{13}\}$. The solution to (6) can be represented in the form

$$\hat{\mathbf{g}}_u = \frac{1}{\det\hat{A}}\hat{A}^c\hat{\Omega},$$

where matrix $\hat{A}^c = \|A_{ij}\|_{i,j=1,2,3}$ has the following elements:

$$A_{11} = q'^4 - [q'^2(V_{A1}^2 + V_{s1}^2) + V_{A1}^2V_{s1}^2p_2^2](p_2^2 + p_3^2),$$
$$A_{12} = A_{21} = V_{s1}^2p_1p_2(q'^2 - V_{A1}^2(\mathbf{s}_1\mathbf{p})^2),$$
$$A_{13} = A_{31} = [q'^2(V_{A1}^2 + V_{s1}^2) - V_{A1}^2V_{s1}^2(\mathbf{s}_1\mathbf{p})^2]p_1p_3,$$
$$A_{22} = [q'^2 - V_{A1}^2p^2 - V_{s1}^2(p_1^2 + p_3^2)](q'^2 - V_{A1}^2(\mathbf{s}_1\mathbf{p})^2), \quad (7)$$
$$A_{23} = A_{32} = V_{s1}^2p_2p_3(q'^2 - V_{A1}^2(\mathbf{s}_1\mathbf{p})^2),$$

$$A_{33} = q'^4 - [q'^2(V_{A1}^2 + V_{s1}^2) + V_{A1}^2 V_{s1}^2(\mathbf{s}_1\mathbf{p})^2](p_2^2 + p_1^2),$$

and its determinant is

$$\begin{aligned}\det\hat{A} &= [(q - \mathbf{p}\mathbf{U}_1)^2 - V_{A1}^2(\mathbf{s}_1\mathbf{p})^2][(q - \mathbf{p}\mathbf{U}_1)^4 \\ &- (V_{A1}^2 + V_{s1}^2)(q - \mathbf{p}\mathbf{U}_1)^2\mathbf{p}^2 + V_{A1}^2 V_{s1}^2\mathbf{p}^2(\mathbf{s}_1\mathbf{p})^2].\end{aligned} \tag{8}$$

For the solution obtained to take the form as in (2), we expand $\hat{\Omega}$ into the sum of dyads

$$\begin{aligned}\hat{\Omega} &= \hat{\varepsilon}_u(q - \mathbf{p}\mathbf{U}_1) - \frac{V_{s1}^2\hat{\varepsilon}_\rho}{\rho_1}\mathbf{p} + \frac{1}{4\pi\rho_1}(\mathbf{s}_1\mathbf{p})\hat{\varepsilon}_b \\ &+ \frac{1}{4\pi\rho_1}\begin{pmatrix} 0 & -p_1 & 0 \\ 0 & -p_2 & 0 \\ 0 & -p_3 & 0 \end{pmatrix}\hat{\varepsilon}_b \\ &+ \frac{1}{4\pi\rho_1 q}[\mathbf{p}(\mathbf{U}_1\mathbf{B}_1) - \mathbf{U}_1 B_1(\mathbf{s}_1\mathbf{p})]\mathbf{p}\hat{\varepsilon}_b.\end{aligned}$$

Hence, the weight function of the first-rank tensor $\hat{G}_{u\rho}$ takes the form

$$\hat{\mathbf{g}}_{u\rho} = -\frac{V_{s1}^2}{\rho_1}\frac{\hat{A}^c\mathbf{p}}{\det\hat{A}},$$

and the weight functions of the second-rank tensors $\hat{G}_{uu}$ and $\hat{G}_{ub}$ are

$$\hat{\mathbf{g}}_{uu} = (q - \mathbf{p}\mathbf{U}_1)\frac{\hat{A}^c}{\det\hat{A}},$$

$$\begin{aligned}\hat{\mathbf{g}}_{ub} &= \frac{1}{4\pi\rho_1}\frac{\hat{A}^c}{\det\hat{A}}\left\{\mathbf{s}_1\mathbf{p} - \begin{pmatrix} 0 & p_1 & 0 \\ 0 & p_2 & 0 \\ 0 & p_3 & 0 \end{pmatrix}\right. \\ &\left. + \frac{1}{q}[\mathbf{p}(\mathbf{U}_1\mathbf{B}_1) - \mathbf{U}_1 B_1(\mathbf{s}_1\mathbf{p})]\mathbf{p}\right\}.\end{aligned}$$

The weight functions of the other tensors in (2) are found in the same way.

For brevity, we write down in detail only the second-rank tensor $\hat{\mathbf{G}}_{uu}$:

$$\begin{aligned}\hat{\mathbf{G}}_{uu}(\mathbf{r} - \mathbf{r}', t - t') &= \hat{\mathbf{F}}_{uu}\mathbf{I}(\mathbf{r} - \mathbf{r}', t - t'), \\ \hat{\mathbf{F}}_{uu} &= \|F_{ij}\|_{i,j=1,2,3},\end{aligned} \tag{9}$$

Here, the differential operator $\hat{\mathbf{F}}_{uu}$ contains the following components:

$$\begin{aligned}F_{11} &= \left[\left(\frac{\partial}{\partial t} - \mathbf{U}_1\nabla\right)^4\right. \\ &\left. - \left[-(V_{A1}^2 + V_{s1}^2)\left(\frac{\partial}{\partial t} - \mathbf{U}_1\nabla\right)^2 - V_{A1}^2 V_{s1}^2\frac{\partial^2}{\partial x_2^2}\right]\right. \\ &\left.\times\left(\frac{\partial^2}{\partial x_2^2} + \frac{\partial^2}{\partial x_3^2}\right)\right]\left(\frac{\partial}{\partial t} - \mathbf{U}_1\nabla\right),\end{aligned}$$

$$\begin{aligned}F_{22} &= \left[\left(\frac{\partial}{\partial t} - \mathbf{U}_1\nabla\right)^2 - V_{A1}^2\Delta - V_{s1}^2\left(\frac{\partial^2}{\partial x_1^2} + \frac{\partial^2}{\partial x_3^2}\right)\right] \\ &\times\left[\left(\frac{\partial}{\partial t} - \mathbf{U}_1\nabla\right)^2 - V_{A1}^2\frac{\partial^2}{\partial x_2^2}\right]\left(\frac{\partial}{\partial t} - \mathbf{U}_1\nabla\right),\end{aligned}$$

$$\begin{aligned}F_{33} &= \left[\left(\frac{\partial}{\partial t} - \mathbf{U}_1\nabla\right)^4\right. \\ &\left. - \left[(V_{A1}^2 + V_{s1}^2)\left(\frac{\partial}{\partial t} - \mathbf{U}_1\nabla\right)^2 - V_{A1}^2 V_{s1}^2\frac{\partial^2}{\partial x_2^2}\right]\right. \\ &\left.\times\left(\frac{\partial^2}{\partial x_2^2} + \frac{\partial^2}{\partial x_1^2}\right)\right]\left(\frac{\partial}{\partial t} - \mathbf{U}_1\nabla\right),\end{aligned}$$

$$\begin{aligned}F_{12} &= F_{21} = V_{s1}^2\left[\left(\frac{\partial}{\partial t} - \mathbf{U}_1\nabla\right)^2 - V_{A1}^2\frac{\partial^2}{\partial x_2^2}\right] \\ &\times\left(\frac{\partial}{\partial t} - \mathbf{U}_1\nabla\right)\frac{\partial^2}{\partial x_1\partial x_2},\end{aligned}$$

$$\begin{aligned}F_{23} &= F_{32} = V_{s1}^2\left[\left(\frac{\partial}{\partial t} - \mathbf{U}_1\nabla\right)^2 - V_{A1}^2\frac{\partial^2}{\partial x_2^2}\right] \\ &\times\left(\frac{\partial}{\partial t} - \mathbf{U}_1\nabla\right)\frac{\partial^2}{\partial x_3\partial x_2},\end{aligned}$$

$$\begin{aligned}F_{13} &= F_{31} = \left[\left(\frac{\partial}{\partial t} - \mathbf{U}_1\nabla\right)^2(V_{A1}^2 + V_{s1}^2) - V_{s1}^2 V_{A1}^2\frac{\partial^2}{\partial x_2^2}\right] \\ &\times\left(\frac{\partial}{\partial t} - \mathbf{U}_1\nabla\right)\frac{\partial^2}{\partial x_1\partial x_3},\end{aligned}$$

and the Fourier–Laplace transform as a function of the

shift of space and time variables has the form

$$I(\mathbf{r}-\mathbf{r}', t-t') = \frac{1}{(2\pi)^4}\int_{-\infty+i\sigma_0}^{\infty+i\sigma_0} \exp[-iq(t-t')]dq \times \iiint_{\infty} \frac{\exp[i\mathbf{p}(\mathbf{r}-\mathbf{r}')]}{\Delta(\mathbf{p}, q)\delta(\mathbf{p}, q)}d\mathbf{p}. \quad (10)$$

It is seen from (10) that the expression for the Green function involves the integration within infinite limits. To calculate this integral, we have to avoid the poles on the integration path. The poles occur at values of $\mathbf{p}$ and q satisfying the dispersion equations

$$\Delta(\mathbf{p}, q) = (q-\mathbf{U}_1\mathbf{p})^4 - (V_{s1}^2 + V_{A1}^2)(q-\mathbf{U}_1\mathbf{p})^2\mathbf{p}^2 + V_{A1}^2V_{s1}^2\mathbf{p}^2(\mathbf{s}_1\mathbf{p})^2 = 0, \quad (11)$$

$$\delta(\mathbf{p}, q) = (q-\mathbf{U}_1\mathbf{p})^2 - V_{A1}^2(\mathbf{s}_1\mathbf{p})^2 = 0.$$

These poles are associated with free magnetohydrodynamic waves in the moving medium. When the waves are undamped, the poles are on the real axis. Eqs. (1), as the first-order partial differential equations, have several linearly independent solutions. The integral formula (10) gives all of these solutions when various ways of passing around the poles are fixed. The given Green function describes the field created by an instantaneous point source placed at the point $\mathbf{r}'$. We assume that the field is zero until the moment $t = t'$ when the source turns on. This assumption provides a rule of passing around the poles when integrating (10). Indeed, in order for the Green function to be zero for $t < t'$, we have to bypass from above all the poles on the integration path in q. The Mandelstam principle is thereby fulfilled, according to which the energy flux at infinity is directed away from the source [5]. This principle imposes a constraint on the group velocity. The Zommerfeld principle demands that only divergent waves exist and imposes limitations on the wave phase velocity, which does not generally coincide with the group velocity. Therefore, the Zommerfeld principle is not always sufficient, especially when moving media are considered.

The Green function of the magnetohydrodynamic equations for the moving media is thus a tensor function of the observation point $\mathbf{r}$ and the source point $\mathbf{r}'$; its spectral components are given in (9).

The dispersion equations (11) for plane monochromatic magnetohydrodynamic waves were derived by equating the denominator in the Fourier–Laplace expansion for the Green function to zero. The corresponding equations for magnetosonic waves have the form

$$(\omega-\mathbf{U}_1\mathbf{k})^4 - (V_{s1}^2 + V_{A1}^2)(\omega-\mathbf{U}_1\mathbf{k})^2\mathbf{k}^2 + V_{A1}^2V_{s1}^2\mathbf{k}^2(\mathbf{s}_1\mathbf{k})^2 = 0 \quad (12)$$

and, for the Alfven waves, they are

$$(\omega-\mathbf{U}_1\mathbf{k})^2 - V_{A1}^2(\mathbf{s}_1\mathbf{k})^2 = 0. \quad (13)$$

The dispersion relations (12) and (13) contain the scalar product $(\mathbf{U}_1\mathbf{k})$ of wave vector $\mathbf{k}$ and unperturbed velocity $\mathbf{U}_1$ of the medium. It means that the wave propagation in the moving medium depends on the angle between the wave vector and the medium velocity. This expresses the fact that the moving medium is anisotropic with respect to a given direction of the motion velocity. The presence of the product $(\mathbf{s}_1\mathbf{k})$ of wave vector $\mathbf{k}$ and unperturbed magnetic field $\mathbf{s}_1$ reveals the magnetic anisotropy, when the propagation velocity depends on its direction with respect to the magnetic field. Thus, we have two axes of anisotropy, associated with the magnetic field and velocity. This essentially sophisticates the study of magnetohydrodynamic waves in moving media.

The dispersion relations (12) and (13) describe three modes of wave motion induced by different recovering forces. The magnetic tension results in Alfven waves (see (13)). The combined action of the magnetic pressure and pressure of the conducting liquid leads to the formation of two, fast and slow, magnetosonic waves (see (12)). These waves have a different nature as compared to sonic and electromagnetic waves. This makes them interesting for studies, especially for studies of their scattering at various obstacles. While solving the magnetohydrodynamic boundary-value problem in the differential form, the local boundary and initial conditions may be satisfied by either single- or several-mode waves. The integral form avoids this difficult problem due to the physics of the phenomenon. The appearance of scattered waves in the medium leads, under the basic mode action, to formation of induced sources that radiate secondary scattered waves. The interference of the secondary waves gives the necessary modes. This represents the extinction principle of magnetic hydrodynamics [1]. The Green function (9), always implying the representation of the differential equations in the integral form, and the extinction principle lead to a new formalism of solving boundary-value problems by the method of integral equations of linear magnetic hydrodynamics.

As we have discussed, the Green function has an evident interpretation describing the distribution of fields or perturbations created by lumped sources. To calculate the field from some distributed sources, we have to sum the effects from each elementary part of them. In this way, the integral equation can be formulated in the class of generalized function as follows.

Let there be in the magnetohydrodynamic space with parameters $\mathbf{B}_1$, $\mathbf{U}_1$, V_{A1}, V_{s1}, and ρ_1 a volume inhomogeneity $V(t)$ characterized by parameters $\mathbf{B}_2$, $\mathbf{U}_2$,

V_{A2}, V_{s2}, and ρ_2. Using the characteristic function

$$\chi = \chi(\mathbf{r}, t) = \begin{cases} 1, & \mathbf{r} \in V(t) \\ 0, & \mathbf{r} \notin V(t) \end{cases}.$$

Eqs. (1) are extended on the whole space as follows:

$$\frac{\partial \mathbf{u}}{\partial t} + (\mathbf{U}_1\nabla)\mathbf{u} + \frac{V_{s1}}{\rho_1}\nabla\tilde{\rho} + \frac{1}{4\pi\rho_1}[\mathbf{B}_1 \mathrm{rot}\mathbf{b}] = \mathbf{W}_u,$$

$$\frac{\partial \mathbf{b}}{\partial t} - \mathrm{rot}[\mathbf{U}_1\mathbf{b}] + \mathrm{rot}[\mathbf{B}_1\mathbf{b}] = \mathbf{W}_b, \qquad (14)$$

$$\frac{\partial \tilde{\rho}}{\partial t} + \rho_1 \mathrm{div}\mathbf{u} + \mathrm{div}(\tilde{\rho}\mathbf{U}_1) = W_\rho,$$

where

$$\mathrm{W}_u = \chi\left\{(\mathbf{U}_1 - \mathbf{U}_2, \nabla)\mathbf{u} + (V_{s1}^2 - V_{s2}^2)\nabla\tilde{\rho} + \left[\frac{\mathbf{B}_1}{4\pi\rho_1} - \frac{\mathbf{B}_2}{4\pi\rho_2}, \mathrm{rot}\mathbf{b}\right]\right\},$$

$$\mathbf{W}_b = \chi\{\mathrm{rot}[\mathbf{B}_1 - \mathbf{B}_2, \mathbf{u}] - \mathrm{rot}[\mathbf{U}_1 - \mathbf{U}_2, \mathbf{b}]\},$$

$$W_\rho = \chi\{(\rho_1 - \rho_2)\mathrm{div}\mathbf{u} + \mathrm{div}\tilde{\rho}(\mathbf{U}_1 - \mathbf{U}_2)\}.$$

The solution of the inhomogeneous system (14) then has the form

$$\begin{aligned}
\mathbf{u}(\mathbf{r}, t) &= \int dt' d\mathbf{r}' \{\hat{\mathbf{G}}_{u\rho}(\mathbf{r} - \mathbf{r}', t - t')W_\rho(\mathbf{r}', t') \\
&\quad + \hat{\mathbf{G}}_{uu}(\mathbf{r} - \mathbf{r}', t - t')\mathbf{W}_u(\mathbf{r}', t') \\
&\quad + \hat{\mathbf{G}}_{ub}(\mathbf{r} - \mathbf{r}', t - t')\mathbf{W}_b(\mathbf{r}', t')\}, \\
\mathbf{b}(\mathbf{r}, t) &= \int dt' d\mathbf{r}' \{\hat{\mathbf{G}}_{b\rho}(\mathbf{r} - \mathbf{r}', t - t')W_\rho(\mathbf{r}', t') \\
&\quad + \hat{\mathbf{G}}_{bu}(\mathbf{r} - \mathbf{r}', t - t')\mathbf{W}_u(\mathbf{r}', t') \\
&\quad + \hat{\mathbf{G}}_{bb}(\mathbf{r} - \mathbf{r}', t - t')\mathbf{W}_b(\mathbf{r}', t')\}, \\
\tilde{\rho}(\mathbf{r}, t) &= \int dt' d\mathbf{r}' \{\hat{\mathbf{G}}_{\rho\rho}(\mathbf{r} - \mathbf{r}', t - t')W_\rho(\mathbf{r}', t') \\
&\quad + \hat{\mathbf{G}}_{\rho u}(\mathbf{r} - \mathbf{r}', t - t')\mathbf{W}_u(\mathbf{r}', t') \\
&\quad + \hat{\mathbf{G}}_{\rho b}(\mathbf{r} - \mathbf{r}', t - t')\mathbf{W}_b(\mathbf{r}', t')\}.
\end{aligned} \qquad (15)$$

The integration in (15) is performed over the entire space occupied by the field and over the infinite time interval. Formula (15) represents the integral–differential equations of magnetic hydrodynamics in the form of a convolution. Due to the convolution properties, we can take the differential operators out of the integral and arrive at the integral equations of linear magnetic hydrodynamics in the moving medium. Specific cases of these equations were considered in [6].

Let us return to the calculated Green function. It follows from (10) that the denominator of this function involves the product of $\Delta(\mathbf{p}, q)$ and $\delta(\mathbf{p}, q)$ from (11), representing the dispersion relations for magnetosonic and Alfven waves. Therefore, the components of the weight functions ($\hat{\mathbf{g}}_{uu}$) associated with waves of these two types are coupled. It turns out that, in some problems of diffraction of magnetohydrodynamic waves at plane-parallel obstacles or in the case of cylindrical domains with a specific orientation of the unperturbed magnetic field, the Alfven and magnetosonic components of the weight function become uncoupled. This was confirmed by paper [3], where the Green function was calculated in a particular case of a one-dimensional magnetic field. This effect is related to the Green function invariance with respect to a choice of the coordinate basis $\langle \mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3 \rangle$, which we associated with the unperturbed magnetic field. If one of the basis vectors is fixed and constant, e.g., ($\mathbf{e}_2 = \mathbf{s}_1$), the Green function is invariant with respect to a surface transversal to $\mathbf{e}_2$. This is related to the known fact that the Green tensor functions of free space solving Eqs. (1) involving only del operators are invariant over a choice of coordinate basis [7]. The Green function (9) then has the form

$$\hat{\mathbf{G}}(\mathbf{r} - \mathbf{r}', t - t') = \begin{pmatrix} G_A & 0 \\ 0 & G_M \end{pmatrix} = \begin{pmatrix} G_{11} & 0 & 0 \\ 0 & G_{22} & G_{23} \\ 0 & G_{32} & G_{33} \end{pmatrix},$$

demonstrating the separation of the Alfven G_A and magnetosonic G_M components. In [1], the following explicit expression for G_A was derived when there was no unperturbed motion of the medium:

$$G_A = \frac{1}{2V_{A1}}\delta(x_1 - x_1')\delta(x_2 - x_2')\Theta\left(t - t' - \frac{|x_3 - x_3'|}{V_{A1}}\right),$$

where $\Theta(t)$ is the Heaviside function.

When the gas pressure is zero, $V_s \equiv 0$, the Green function (13) becomes simpler and equals

$$\hat{\mathbf{G}}(\mathbf{r} - \mathbf{r}') = \begin{pmatrix} G_A & 0 & 0 \\ 0 & (V_{A1}^2\Delta + \omega^2)F & 0 \\ 0 & 0 & \omega F \end{pmatrix},$$

where

$$G_A = \frac{1}{2V_{A1}}\delta(x_1 - x_1')\delta(x_2 - x_2')\exp\left(\frac{-i\omega|x_3 - x_3'|}{V_{A1}}\right),$$

$$F = \frac{\exp[-i\omega|\mathbf{r} - \mathbf{r}'|/V_{A1}]}{4\pi V_{A1}^2\omega^2|\mathbf{r} - \mathbf{r}'|}.$$

It is seen that parameter V_s does not affect the Alfven component and the magnetosonic components G_{22} and

G_{33} look like the Green function of macroscopic electrodynamics [8].

The acoustic Green function known in the literature can easily be obtained as a particular case of the fundamental solution (9).

When the magnetic field and unperturbed motion of the medium are absent (i.e., $\mathbf{B} \equiv 0$ and $\mathbf{U}_1 \equiv 0$), the function $I(\mathbf{r} - \mathbf{r}')$ takes the form

$$I(\mathbf{r} - \mathbf{r}') = \frac{1}{2\pi}\int_{\infty} \frac{\exp(i\mathbf{p}(\mathbf{r} - \mathbf{r}'))}{\omega^2 - V_{s1}^2 p^2} d\mathbf{p}$$

and is easily calculated with the help of the residue theorem as follows:

$$I(\mathbf{r} - \mathbf{r}') = \frac{1}{4\pi V_{s1}^2} \frac{\exp(-i\omega|\mathbf{r} - \mathbf{r}'|/V_{s1})}{|\mathbf{r} - \mathbf{r}'|}.$$

The components of the differential operator $\tilde{\mathbf{F}}_{uu}$ are simplified to the form

$$1 + \frac{V_{s1}^2}{\omega^2}\left(\frac{\partial^2}{\partial x_2^2} + \frac{\partial^2}{\partial x_3^2}\right), \quad -\frac{V_{s1}^2}{\omega^2}\frac{\partial^2}{\partial x_1 \partial x_2}, \quad -\frac{V_{s1}^2}{\omega^2}\frac{\partial^2}{\partial x_1 \partial x_3},$$

$$-\frac{V_{s1}^2}{\omega^2}\frac{\partial^2}{\partial x_1 \partial x_2}, \quad 1 + \frac{V_{s1}^2}{\omega^2}\left(\frac{\partial^2}{\partial x_1^2} + \frac{\partial^2}{\partial x_3^2}\right), \quad -\frac{V_{s1}^2}{\omega^2}\frac{\partial^2}{\partial x_2 \partial x_3},$$

$$-\frac{V_{s1}^2}{\omega^2}\frac{\partial^2}{\partial x_1 \partial x_3}, \quad -\frac{V_{s1}^2}{\omega^2}\frac{\partial^2}{\partial x_2 \partial x_3}, \quad 1 + \frac{V_{s1}^2}{\omega^2}\left(\frac{\partial^2}{\partial x_1^2} + \frac{\partial^2}{\partial x_2^2}\right).$$

Finally, we can write

$$\hat{\mathbf{G}}(\mathbf{r} - \mathbf{r}') = \left(\hat{\varepsilon} - \frac{V_{s1}^2}{\omega^2}\nabla \times \nabla \times \hat{\varepsilon}\right) I(\mathbf{r} - \mathbf{r}'),$$

where $\hat{\varepsilon}$ is the unit affinor. This function was obtained in [9] while deriving integral equations of acoustics of inhomogeneous liquid. The time dependence is of the form $\exp(i\omega t)$.

REFERENCES

1. A. A. Aleksandrova and N. A. Khizhnyak, *Boundary-Value Problems of Magnetohydrodynamics* (Test-Radio, Kharkov, 1993).
2. A. A. Aleksandrova and Yu. N. Aleksandrov, Zh. Tekh. Fiz. **67** (5), 6 (1997) [Tech. Phys. **42**, 460 (1997)].
3. H. Wetzner, Phys. Fluids **4** (10), 1238 (1961).
4. N. E. Kochin, *Vector Calculus and Principles of Tensor Calculus* (Akad. Nauk SSSR, Moscow, 1951).
5. L. I. Mandel'shtam, *Lectures on Optics, Theory of Relativity, and Quantum Mechanics* (Nauka, Moscow, 1971).
6. A. A. Aleksandrova, Zarubezhn. Radioélektron.: Usp. Sovrem. Radioélektron., No. 3, 25 (1999).
7. L. B. Felsen and N. Marcuvitz, *Radiation and Scattering of Waves* (Prentice-Hall, Englewood Cliffs, 1973; Mir, Moscow, 1978), Vol. 1.
8. N. A. Khizhnyak, *Integral Equations in Macroscopic Electrodynamics* (Naukova Dumka, Kiev, 1986).
9. A. G. Vinogradov and R. Z. Muratov, Dokl. Akad. Nauk SSSR **226** (2), 310 (1976) [Sov. Phys. Dokl. **21**, 29 (1976)].

Translated by V. Gursky

Technical Physics, Vol. 46, No. 7, 2001, pp. 789–795. Translated from Zhurnal Tekhnicheskoĭ Fiziki, Vol. 71, No. 7, 2001, pp. 7–13.

Transformation of a Monochromatic Plane Wane by the Pulse-Periodic Time Modulation of an Infinite Medium

O. N. Rybin and N. I. Slipchenko

Kharkov State Technical University, Kharkov, 310726 Ukraine

Received October 10, 2000

Abstract—An exact solution to the problem of the transformation of a monochromatic plane wave by a finite train of equally spaced rectangular pulses of permittivity and conductivity of an infinite medium is considered. The permittivity pulse train is shifted relative to the conductivity pulse train by an arbitrary time. The problem is studied analytically in terms of the second-order Volterra integral equation describing the electromagnetic wave transformation in a medium with time-dependent parameters. The equation is solved using the resolvent technique. Expressions for the amplitude of the transformed electric field component for any time instant at any spatial point are derived and analyzed. © *2001 MAIK "Nauka/Interperiodica".*

INTRODUCTION

The modulation of the parameters of a propagation medium is an important issue of electromagnetic theory. In this regard, the transformation of an electromagnetic wave by a medium whose parameters are periodically modulated in time is of great interest. Such a transformation was studied in [1–3]. However, approximate analytical approaches to studying this problem are usually based on certain assumptions. For example, the transformation of a plane wave by the harmonic time variation of the conductivity of a semi-infinite medium was studied in [4] in terms of perturbation theory. This technique limits the modulation depth to small values. Furthermore, in this problem, the conductivity was modulated over an infinite time period. Therefore, it is of interest to study the wave transformation in media whose parameters change periodically over a finite time interval.

In this paper, the problem of the transformation of a monochromatic electromagnetic wave in an infinite medium with the permittivity and conductivity modulated in time by a train of equally spaced rectangular pulses is solved exactly. The repetition period, amplitude, and duration of the pulses may be arbitrary (however, the permittivity and conductivity pulses are assumed to have equal durations). Furthermore, the train of the permittivity pulses may be arbitrarily shifted relative to the conductivity pulses.

Exact expressions for the transformed electric field component for any time instant at any spatial point are derived and analyzed.

ELECTRIC FIELD IN A STEPWISE-PERIODIC NONSTEADY MEDIUM

Let the parameters of an infinite medium start to vary arbitrarily at time $t = 0$. According to [5], the electric field in such a medium is described by the Volterra integral equation of the second kind,

$$E(t, x) = F(t, x) + \int_0^t dt' \int_{-\infty}^{\infty} dx' \times K(t, t', x, x')E(t', x'), \quad t > 0, \tag{1}$$

where $E(t, x)$ is the electric field, $K(t, t', x, x')$ is the kernel of the integral equation, and $F(t, x)$ is the free term.

Integral equation (1) may be solved by the resolvent technique [6]:

$$E(t, x) = F(t, x) + \int_0^t dt' \int_{-\infty}^{\infty} dx' \times R(t, t', x, x')F(t', x'), \quad t > 0, \tag{2}$$

where $R(t, t', x, x')$ is the resolvent of Eq. (1).

If the permittivity and conductivity change at $t = 0$ in a stepwise manner (before $t = 0$, the medium is assumed to have zero conductivity; i.e., $\sigma = 0$), the kernel and resolvent of Eq. (1) are described by the difference functions [7]

$$K(t, t', x, x') = -\frac{1}{a^2}\left\{\bar{\sigma}_1 + \frac{1}{2}(1 - a^2)\frac{\partial}{\partial t}\right\}\delta(v_0(t - t') - |x - x'|),$$

$$R(t, t', x, x') = -\frac{1}{2v_1}\int_{\alpha - i\infty}^{\alpha + i\infty} dx' \frac{dp}{2\pi i}(p - \bar{\sigma}_1) \times \frac{(1 - a^2)p + (1 + a^2)\bar{\sigma}_1}{\sqrt{p^2 - \bar{\sigma}_1^2}} e^{p(t - t') - \frac{|x - x'|}{v_1}\sqrt{p^2 - \bar{\sigma}_1^2}}, \tag{3}$$

1063-7842/01/4607-0789$21.00

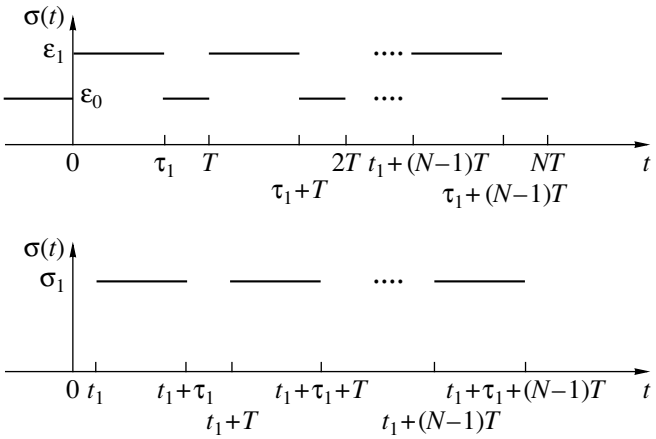


Fig. 1. Time behavior of the parameters of a medium.

where $\alpha > \bar{\sigma}_1$, $\mathrm{Re}(\sqrt{p^2-\bar{\sigma}_1^2}) > 0$, $a = \sqrt{\varepsilon_0/\varepsilon_1}$, $\bar{\sigma}_1 = 2\pi\sigma_1/\varepsilon_1$, $v_1 = c/\sqrt{\varepsilon_1}$, ε_0 is the permittivity of the unperturbed medium (at $t < 0$), ε_1 and σ_1 are the permittivity and conductivity of the perturbed medium, and c is the speed of light in free space.

Let the permittivity and conductivity of the medium change in time as (Fig. 1)

$$\begin{aligned}\varepsilon(t) &= \varepsilon_0 + (\varepsilon_1+\varepsilon_0)\sum_{k=1}^{N}\{\Theta(t-(k-1)T)\\ &\quad -\Theta(t-\tau_1-(k-1)T)\},\\ \sigma(t) &= \sigma_1\sum_{k=1}^{N}\{\Theta(t-t_1-(k-1)T)\\ &\quad -\Theta(t-t_1-\tau_1-(k-1)T)\},\end{aligned} \tag{4}$$

where t_1 is the time shift between the permittivity and conductivity sequences and τ_1 is the pulse duration.

Clearly, the perturbation described by formulas (4) has a period $T = t_1 + \tau_1 + \tau_2$, where τ_2 is the time interval over which the parameters take their quiescent values, and the train contains N pulses.

It can be shown that, on any nth time interval over which the medium parameters remain constant, the electric field is expressed by the formulas

$$\begin{aligned}E_n(t,x) &= F_n(t,x)\\ &\quad + \int_{t_{n-1}}^{t}dt'\int_{-\infty}^{\infty}dx' R_n(t,t',x,x')F_n(t',x'),\\ F_n(t,x) &= E_0(t,x)\\ &\quad + \sum_{i=1}^{n-1}\int_{t_{i-1}}^{t_i}dt'\int_{-\infty}^{\infty}dx' K_i(t,t',x,x')E_i(t',x'),\end{aligned} \tag{5}$$

where $E_0(t, x)$ is the primary field and $R_n(t, t', x, x')$ and $K_n(t, t', x, x')$ are the resolvent and kernel of integral equation (1), their structure being independent of index n of the step in the parameters of the medium.

From the last statement and formulas (3), one can easily obtain expressions for the kernel and resolvent on an arbitrary time interval over which the parameters of the medium remain constant. Specifically, on the intervals over which the permittivity alone takes a new value, expressions for the kernel and resolvent take the form

$$(n-1)T < t < t_1 + (n-1)T, \quad n = 1, \ldots, N:$$

$$\begin{aligned}K_n(t,t',x,x') &= -\frac{1}{2a^2}\frac{\partial}{\partial t}\delta(v_0(t-t')-|x-x'|),\\ v_0 &= c/\sqrt{\varepsilon_0},\\ R_n(t,t',x,x') &= -\frac{1}{2}(1-a^2)\frac{\partial}{\partial t}\delta(v_1(t-t')-|x-x'|).\end{aligned} \tag{6}$$

At the intervals over which both the permittivity and conductivity take new values, expressions for the kernel and resolvent are given by the formulas

$$t_1 + (n-1)T < t < \tau_1 + (n-1)T, \quad n = 1, \ldots, N:$$

$$\begin{aligned}K_n(t,t',x,x') &= -\frac{1}{a^2}\left\{\bar{\sigma}_1 + \frac{1}{2}(1-a^2)\frac{\partial}{\partial t}\right\}\\ &\quad \times\delta(v_0(t-t')-|x-x'|);\\ R(t,t',x,x') &= -\frac{1}{2v_1}\int_{\alpha-i\infty}^{\alpha+i\infty}\frac{dp}{2\pi i}(p-\bar{\sigma}_1)\\ &\quad \times\frac{(1-a^2)p+(1+a^2)\bar{\sigma}_1}{\sqrt{p^2-\bar{\sigma}_1^2}}e^{p(t-t')-\frac{|x-x'|}{v_1}\sqrt{p^2-\bar{\sigma}_1^2}},\\ \alpha &> \bar{\sigma}_1, \quad \mathrm{Re}(\sqrt{p^2-\bar{\sigma}_1^2}) > 0.\end{aligned} \tag{7}$$

At the intervals over which the conductivity alone takes a new value, expressions for the kernel and resolvent are as follows:

$$\tau_1 + (n-1)T < t < t_1 + \tau_1 + (n-1)T, \quad n = 1, \ldots, N:$$

$$\begin{aligned}K_n(t,t',x,x') &= -\bar{\sigma}\delta(v_0(t-t')-|x-x'|),\\ \bar{\sigma} &= 2\pi\sigma_1/\varepsilon_0,\\ R_n(t,t',x,x') &= -\frac{\bar{\sigma}}{v_0}\frac{\partial}{\partial t}\int_{\alpha-i\infty}^{\alpha+i\infty}\frac{dp}{2\pi i}\end{aligned} \tag{8}$$

$$\times \frac{e^{(p-\bar{\sigma})(t-t')-\frac{|x-x'|}{v_0}\sqrt{p^2-\bar{\sigma}^2}}}{\sqrt{p^2-\bar{\sigma}^2}}, \quad \alpha > \bar{\sigma},$$

$$\mathrm{Re}(\sqrt{p^2-\bar{\sigma}^2}) > 0.$$

Finally, at the quiescent intervals over which the parameters of the medium are equal to their original values, we have

$$t_1 + \tau_1 + (n-1)T < t < nT, \quad n = 1, \ldots, N: \qquad K_n(t, t', x, x') = 0 = R_n(t, t', x, x'). \tag{9}$$

TRANSFORMATION OF A MONOCHROMATIC WAVE BY PULSE-PERIODIC MODULATION OF AN INFINITE MEDIUM

Let $E_0(t, x)$ be a unit-amplitude monochromatic plane wave $E_0(t, x) = e^{i(\omega t - kx)}$, where $k = \omega/v_0$. Substituting $E_0(t, x)$ into formulas (5) results in the equations for the transformed electric field at the first period of variations in the medium parameters:

$$E(t, x) = \tilde{A}_1^+ e^{i(\omega_1 t - kx)} + \tilde{A}_1^- e^{-i(\omega_1 t + kx)}, \quad 0 < t < t_1, \tag{10}$$

where

$$\tilde{A}_1^\pm = \frac{a}{2}(1 \pm a), \quad \omega_1 = a\omega;$$

$$E(t, x) = (\tilde{A}_3^+ + \tilde{B}_3^+)e^{-\bar{\sigma}_1 t}e^{i(\Omega t - kx)} + (\tilde{A}_3^- + \tilde{B}_3^-)e^{-\bar{\sigma}_1 t}e^{-i(\Omega t + kx)}, \quad t_1 < t < \tau_1, \tag{11}$$

$$\tilde{A}_3^\pm = \pm\tilde{A}_2 a^2 \frac{(a^2-1)\omega^2 - 2\bar{\sigma}_1(\bar{\sigma}_1 \mp i\Omega)}{2i\Omega(\bar{\sigma}_1 \mp i(\Omega \mp \omega))}e^{(\bar{\sigma}_1 \mp i(\Omega \mp \omega))t_1},$$

$$\tilde{B}_3^\pm = \pm\tilde{B}_2 a^2 \frac{(a^2-1)\omega^2 - 2\bar{\sigma}_1(\bar{\sigma}_1 \mp i\Omega)}{2i\Omega(\bar{\sigma}_1 \mp i(\Omega \mp \omega))}e^{(\bar{\sigma}_1 \mp i(\Omega \mp \omega))t_1},$$

$$\Omega = \sqrt{a^2\omega^2 - \bar{\sigma}_1^2},$$

$$\tilde{A}_2 = \frac{e^{-i\omega t_1}}{2a}(2a\cos\omega_1 t_1 + i(1 + a^2(\sin\omega_1 t_1))), \tag{12}$$

$$\tilde{B}_2 = ie^{i\omega t_1}\frac{1-a^2}{2a}\sin\omega_1 t_1;$$

$$E(t, x) = \tilde{A}_5^+ e^{-\bar{\sigma}t}e^{i(\bar{\Omega}t - kx)} + \tilde{A}_5^- e^{-\bar{\sigma}t}e^{-i(\bar{\Omega}t + kx)}, \quad \tau_1 < t < t_1 + \tau_1,$$

$$\tilde{A}_5^\pm = \mp\left\{\tilde{A}_4\frac{e^{i\omega\tau_1}}{\bar{\sigma} \mp i(\Omega \mp \omega)} + \tilde{B}_4\frac{e^{-i\omega\tau_1}}{\bar{\sigma} \mp i(\Omega \mp \omega)}\right\} \times \frac{\bar{\sigma}(\bar{\sigma} \mp i\Omega)}{i\Omega}e^{(\bar{\sigma} \mp i\Omega)\tau_1},$$

$$\bar{\Omega} = \sqrt{\omega^2 - \bar{\sigma}^2},$$

$$\tilde{A}_4 = -\frac{e^{-(\bar{\sigma}_1 + i\omega)\tau_1}}{2i\Omega}\left\{\bar{\sigma}_1 + (1-a^2)\frac{i\omega}{2}\right\} \times \left\{\tilde{A}_2\left(\frac{\bar{\sigma}_1 - i(\Omega+\omega)}{\bar{\sigma}_1 - i(\Omega-\omega)}e^{i\Omega\tau_1} - \frac{\bar{\sigma}_1 + i(\Omega-\omega)}{\bar{\sigma}_1 + i(\Omega+\omega)}e^{-i\Omega\tau_1}\right) + 2\tilde{B}_2 ie^{-2i\omega t_1}\sin\Omega\tau_1\right\}, \tag{13}$$

$$\tilde{B}_4 = -\frac{e^{-(\bar{\sigma}_1 - i\omega)\tau_1}}{2i\Omega}\left\{\bar{\sigma}_1 - (1-a^2)\frac{i\omega}{2}\right\} \times \left\{\tilde{B}_2\left(\frac{\bar{\sigma}_1 - i(\Omega-\omega)}{\bar{\sigma}_1 - i(\Omega+\omega)}e^{i\Omega\tau_1} - \frac{\bar{\sigma}_1 + i(\Omega+\omega)}{\bar{\sigma}_1 + i(\Omega-\omega)}e^{-i\Omega\tau_1}\right) + 2\tilde{A}_2 ie^{2i\omega t_1}\sin\Omega\tau_1\right\};$$

$$E(t, x) = A_1 e^{i(\omega t - kx)} + B_1 e^{-i(\omega t + kx)}, \quad t_1 + \tau_1 < t < T,$$

$$A_1 = \tilde{A}_4\frac{i\bar{\sigma}}{2\bar{\Omega}}e^{-(\bar{\sigma} + i\omega)t_1} \times \left\{\frac{\bar{\sigma} - i(\bar{\Omega}+\omega)}{\bar{\sigma} - i(\bar{\Omega}-\omega)}e^{i\bar{\Omega}\tau_1} - \frac{\bar{\sigma} + i(\bar{\Omega}-\omega)}{\bar{\sigma} + i(\bar{\Omega}+\omega)}e^{-i\bar{\Omega}\tau_1}\right\} - \frac{\bar{\sigma}}{\bar{\Omega}}\tilde{B}_4 e^{-(\bar{\sigma} + i\omega)t_1 - 2i\omega\tau_1}\sin\bar{\Omega}t_1,$$

$$B_1 = \tilde{B}_4\frac{i\bar{\sigma}}{2\bar{\Omega}}e^{-(\bar{\sigma} - i\omega)t_1} \times \left\{\frac{\bar{\sigma} - i(\bar{\Omega}-\omega)}{\bar{\sigma} - i(\bar{\Omega}+\omega)}e^{i\bar{\Omega}\tau_1} - \frac{\bar{\sigma} + i(\bar{\Omega}+\omega)}{\bar{\sigma} + i(\bar{\Omega}-\omega)}e^{-i\bar{\Omega}\tau_1}\right\} - \frac{\bar{\sigma}}{\bar{\Omega}}\tilde{A}_4 e^{-(\bar{\sigma} - i\omega)t_1 + 2i\omega\tau_1}\sin\bar{\Omega}t_1.$$

Substituting expressions (10)–(13) into formulas (5) and using the mathematical induction method, we obtain recurrent relationships for the electric field at any interval of variations in the permittivity and conductivity, beginning with the second ($n = 2, \ldots, N$). Spe-

cifically, at the time interval $(n-1)T < t < t_1 + (n-1)T$, we have

$$E(t, x) = L_n e^{i(\omega_1 t - kx)} + M_n e^{-i(\omega_1 t + kx)}, \qquad (14)$$

where

$$L_n = \frac{a}{2} e^{-i\omega_1 (n-1)T} \{(1+a) e^{i\omega(n-1)T} A_{n-1} - (1-a) e^{-i\omega n(n-1)T} B_{n-1}\},$$

$$M_n = -\frac{a}{2} e^{i\omega_1 (n-1)T} \{(1-a) e^{i\omega(n-1)T} A_{n-1} - (1+a) e^{-i\omega(n-1)T} B_{n-1}\}.$$

At the time interval $t_1 + (n-1)T < t < \tau_1 + (n-1)T$, we have

$$E(t, x) = (R_n^+ + S_n^+) e^{-\bar{\sigma}_1 t} e^{i(\Omega t - kx)} + (R_n^- + S_n^-) e^{-\bar{\sigma}_1 t} e^{-i(\Omega t + kx)}, \qquad (15)$$

where

$$R_n^\pm = \pm E_n a^2 \frac{(a^2-1)\omega^2 - 2\bar{\sigma}_1(\bar{\sigma}_1 \mp i\Omega)}{2i\Omega(\bar{\sigma}_1 \mp i(\Omega \mp \omega))} \times e^{(\bar{\sigma}_1 \mp i(\Omega \mp \omega))(t_1 + (n-1)T)},$$

$$S_n^\pm = \pm P_n a^2 \frac{(a^2-1)\omega^2 - 2\bar{\sigma}_1(\bar{\sigma}_1 \mp i\Omega)}{2i\Omega(\bar{\sigma}_1 \mp i(\Omega \mp \omega))} \times e^{(\bar{\sigma}_1 \mp i(\Omega \mp \omega))(t_1 + (n-1)T)},$$

$$E_n = \frac{e^{-i\omega t_1}}{2a}(2a\cos\omega_1 t_1 + i(1+a^2)\sin\omega_1 t_1) A_{n-1} - ie^{-i\omega t_1}\frac{1-a^2}{2a} e^{-2i\omega(n-1)T} \sin\omega_1 t_1 B_{n-1},$$

$$P_n = ie^{i\omega t_1}\frac{1-a^2}{2a} e^{2i\omega(n-1)T} \sin\omega_1 t_1 A_{n-1} + \frac{e^{i\omega t_1}}{2a}(2a\cos\omega_1 t_1 - i(1+a^2)\sin\omega_1 t_1) B_{n-1}.$$

At the time interval $\tau_1 + (n-1)T < t < t_1 + \tau_1 + (n-1)T$, we have

$$E(t, x) = C_n^+ e^{-\bar{\sigma} t} e^{i(\bar{\Omega} t - kx)} + C_n^- e^{-\bar{\sigma} r} e^{-i(\bar{\Omega} t + kx)}, \qquad (16)$$

where

$$C_n^\pm = \mp\left\{ W_n \frac{e^{i\omega(\tau_1 + (n-1)T)}}{\bar{\sigma} \mp i(\Omega \mp \omega)} + X_n \frac{e^{-i\omega(\tau_1 + (n-1)T)}}{\bar{\sigma} \mp i(\Omega \pm \omega)} \right\} \times \frac{\bar{\sigma}(\bar{\sigma} \mp i\Omega)}{i\Omega} e^{(\bar{\sigma} \mp i\Omega)(\tau_1 + (n-1)T)},$$

$$W_n = -\frac{e^{-(\bar{\sigma}_1 + i\omega)\tau_1}}{2i\Omega}\left\{ \bar{\sigma}_1 + (1-a^2)\frac{i\omega}{2} \right\} \times \left\{ E_n \left(\frac{\bar{\sigma}_1 - i(\Omega+\omega)}{\bar{\sigma}_1 - i(\Omega-\omega)} e^{i\Omega\tau_1} - \frac{\bar{\sigma}_1 + i(\Omega-\omega)}{\bar{\sigma}_1 + i(\Omega+\omega)} e^{-i\Omega\tau_1} \right) + 2iP_n e^{-2i\omega(t_1 + (n-1)T)} \sin\Omega\tau_1 \right\},$$

$$X_n = -\frac{e^{-(\bar{\sigma}_1 - i\omega)\tau_1}}{2i\Omega}\left\{ \bar{\sigma}_1 - (1-a^2)\frac{i\omega}{2} \right\} \times \left\{ P_n \left(\frac{\bar{\sigma}_1 - i(\Omega-\omega)}{\bar{\sigma}_1 - i(\Omega+\omega)} e^{i\Omega\tau_1} - \frac{\bar{\sigma}_1 + i(\Omega+\omega)}{\bar{\sigma}_1 + i(\Omega-\omega)} e^{-i\Omega\tau_1} \right) + 2iE_n e^{2i\omega(t_1 + (n-1)T)} \sin\Omega\tau_1 \right\}.$$

At the time interval $t_1 + \tau_1 + (n-1)T < t < nT$, we have

$$E(t, x) = A_n e^{i(\omega t - kx)} + B_n e^{-i(\omega t + kx)}, \qquad (17)$$

where

$$A_n = W_n \frac{e^{-(\bar{\sigma} + i\omega)t_1}}{-2i\bar{\Omega}\bar{\sigma}^{-1}}\left\{ \frac{\bar{\sigma} - i(\bar{\Omega}+\omega)}{\bar{\sigma} - i(\bar{\Omega}-\omega)} e^{i\bar{\Omega}t_1} - \frac{\bar{\sigma} + i(\bar{\Omega}-\omega)}{\bar{\sigma} + i(\bar{\Omega}+\omega)} e^{-i\bar{\Omega}t_1} \right\} - \frac{\bar{\sigma} X_n}{\bar{\Omega}} e^{-(\bar{\sigma} + i\omega)t_1 - 2i\omega(\tau_1 + (n-1)T)} \sin\bar{\Omega}t_1,$$

$$B_n = X_n \frac{e^{-(\bar{\sigma} - i\omega)t_1}}{-2i\bar{\Omega}\bar{\sigma}^{-1}}\left\{ \frac{\bar{\sigma} - i(\bar{\Omega}+\omega)}{\bar{\sigma} - i(\bar{\Omega}-\omega)} e^{i\bar{\Omega}t_1} - \frac{\bar{\sigma} + i(\bar{\Omega}-\omega)}{\bar{\sigma} + i(\bar{\Omega}+\omega)} e^{-i\bar{\Omega}t_1} \right\} - \frac{\bar{\sigma} W_n}{\bar{\Omega}} e^{-(\bar{\sigma} - i\omega)t_1 + 2i\omega(\tau_1 + (n-1)T)} \sin\bar{\Omega}t_1.$$

As can be seen from formulas (10)–(17), the modulation of the parameters of a infinite medium splits the primary monochromatic wave into forward and backward monochromatic waves. The modules of the wavenumbers of the new waves and those of the primary wave are the same. The amplitudes and frequencies of these waves are piecewise-constant functions of time.

If the conductivity is high ($\bar{\sigma}_1 > a\omega$ at the time intervals $t_1 + (n-1)T < t < \tau_1 + (n-1)T$ at the time intervals $\tau_1 + (n-1)T < t < \tau_1 + (n-1)T$, where $n = 1, 2, 3, \ldots$), the frequency becomes imaginary and the new waves become aperiodic.

Note that, at $t_1 = 0$ and $a = 1$, expressions (13) coincide with the corresponding expressions in [8], which describe the transformation of a monochromatic plane wave caused by a pulsed modulation of the conductivity of an infinite medium.

It is of interest to study the field at the end of the modulating train of n conductivity and permittivity pulses as a function of n. To this end, we will analyze the coefficients A_n and B_n. We introduce the coefficients that are independent of n:

$$s = \bar{\sigma}/\omega, \quad q = a\sqrt{1-a^2s^2}, \quad \tau = \omega t_1, \quad \tau_d = \omega\tau_1,$$

$$\tau_s = \omega\tau_2, \quad \tau_T = \omega T, \quad g = \sqrt{1-s^2};$$

$$y = \frac{e^{i\tau}}{2a}(2a\cos a\tau - i(1+a^2)\sin a\tau);$$

$$z = -ie^{-i\tau}\frac{1-a^2}{2a}\sin a\tau;$$

$$f = \frac{e^{-(a^2s+i)\tau_d}}{2q}\{2q\cos q\tau_d + i(1+a^2)\sin q\tau_d\},$$

$$h = \frac{e^{-(a^2s-i)\tau_d - 2i\tau}}{2q}\{2a^2s - i(1-a^2)\}\sin q\tau_d;$$

$$u = e^{-(s+i)\tau}\left\{\cos g\tau + i\frac{\sin g\tau}{g}\right\},$$

$$v = -\frac{s}{g}e^{-(s-i)\tau + 2i\tau_d}\sin g\tau.$$

Then, expressions for the coefficients E_n, P_n, W_n, X_n, A_n, and B_n can be written as

$$\begin{aligned} A_{n+1} &= uW_{n+1} + v^*e^{-2int_T}X_{n+1}, \\ B_{n+1} &= ve^{2int_T}W_{n+1} + u^*X_{n+1}; \end{aligned} \tag{18}$$

$$\begin{aligned} W_n &= fE_n - h^*e^{-2int_T}P_n, \\ X_n &= -he^{2int_T}E_n + f^*P_n; \end{aligned} \tag{19}$$

$$\begin{aligned} E_n &= y^*A_{n-1} + ze^{-2i(n-1)t_T}B_{n-1}, \\ P_n &= z^*e^{2i(n-1)t_T}A_{n-1} + yB_{n-1}. \end{aligned} \tag{20}$$

From expressions (18)–(20), after some algebra, we obtain

$$\begin{aligned} A_{n+1}e^{i(n+1)t_T} &= c_{11}e^{int_T}A_n + c_{12}e^{-int_T}B_n, \\ B_{n+1}e^{-i(n+1)t_T} &= c_{21}e^{int_T}A_n + c_{22}e^{-int_T}B_n, \end{aligned} \tag{21}$$

where

$$c_{11} = ufy^* - v^*hy^*e^{2it_T} - uh^*z^*e^{-2it_T} + v^*f^*z^*,$$

$$c_{12} = ufz - v^*hze^{-2it_T} - uh^*ye^{2it_T} + v^*f^*y,$$

$$c_{21} = vfy^* - u^*hy^*e^{-2it_T} - vh^*z^*e^{2it_T} + u^*f^*z^*,$$

$$c_{22} = vfz^* - u^*hze^{2it_T} - vh^*ye^{-2it_T} + u^*f^*y.$$

We introduce new designations for the amplitudes $a_n = A_ne^{int_T}$ and $b_n = B_ne^{-int_T}$ and assume that

$$a_1 = e^{it_T}[(uf - v^*h)y^* - (f^*v^* - uh^*)z^*],$$

$$b_1 = e^{-it_T}[(vf - u^*h)y^* + (f^*u^* - vh^*)z^*].$$

Then, expression (21) can be written in the matrix form:

$$\begin{pmatrix} a_{n+1} \\ b_{n+1} \end{pmatrix} = e^{-s(a^2\tau_d+\tau)}R\begin{pmatrix} a_n \\ b_n \end{pmatrix}, \quad R = \begin{pmatrix} c_{11} & c_{12} \\ c_{21} & c_{22} \end{pmatrix}. \tag{22}$$

By using the mathematical induction method, formula (22) can be written as

$$\begin{pmatrix} a_{n+1} \\ b_{n+1} \end{pmatrix} = e^{-ns(a^2\tau_d+\tau)}R^n\begin{pmatrix} a_1 \\ b_1 \end{pmatrix}. \tag{23}$$

Since $e^{-s(a^2\tau_d+\tau)}\det R = 1$, we obtain [9]

$$R^n = U_{n-1}\left(\frac{c_{11}+c_{22}}{2}\right)R - U_{n-2}\left(\frac{c_{11}+c_{22}}{2}\right)I, \tag{24}$$

where $U_n(\alpha)$ is the Chebyshev nth order polynomial and I is the identity matrix.

By substituting expression (24) into formula (23), we obtain expressions for the forward and backward wave amplitudes on the quiescent portion of an arbitrary $(n+1)$th period of the parameter variations:

$$\begin{aligned} A_{n+1} &= e^{-i(n+1)t_T - n(a^2s\tau_d + i\tau)}\left\{(c_{11}a_1 + c_{12}b_1)\right. \\ &\left.\times U_{n-1}\left(\frac{c_{11}+c_{22}}{2}\right) - U_{n-2}\left(\frac{c_{11}+c_{22}}{2}\right)a_i\right\}, \\ B_{n+1} &= e^{i(n+1)t_T - n(a^2s\tau_d + i\tau)}\left\{(c_{21}a_1 + c_{22}b_1)\right. \\ &\left.\times U_{n-1}\left(\frac{c_{11}+c_{22}}{2}\right) - U_{n-2}\left(\frac{c_{11}+c_{22}}{2}\right)b_1\right\}. \end{aligned} \tag{25}$$

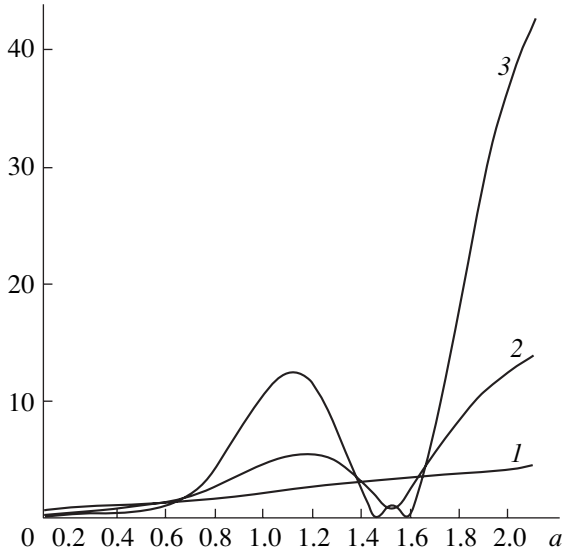


Fig. 2. Amplitude A_n of the forward wave vs. the square root of the relative change in the permittivity, a, at $s = 0.05$, $\tau = 1$, $\tau_d = 2$, and $\tau_T = 5$ for n = (*1*) 3, (*2*) 4, and (*3*) 5.

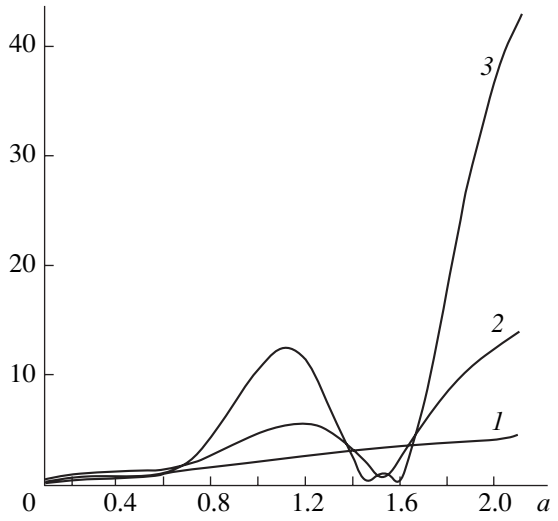


Fig. 3. Amplitude B_n of the backward wave vs. the square root of the relative change in the permittivity, a, at $s = 0.05$, $\tau = 1$, $\tau_d = 2$, and $\tau_T = 5$ for n = (*1*) 2, (*2*) 3, and (*3*) 4.

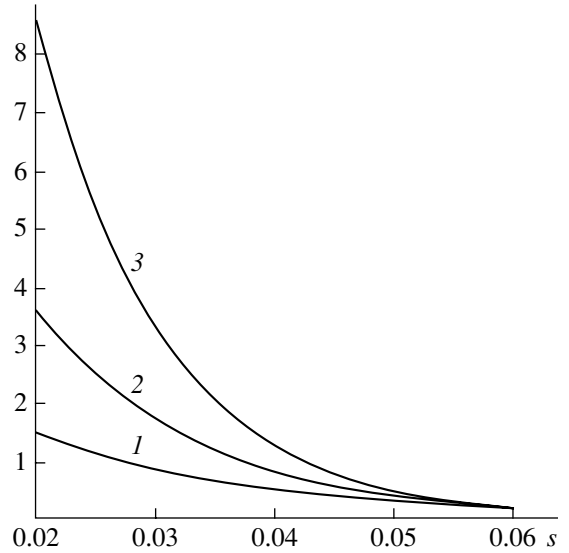


Fig. 4. Amplitude A_n of the forward wave vs. the relative change in the conductivity, s, at $a = 1$, $\tau = 0$, $\tau_d = 2$, and $\tau_T = 5$ for n = (*1*) 3, (*2*) 4, and (*3*) 5.

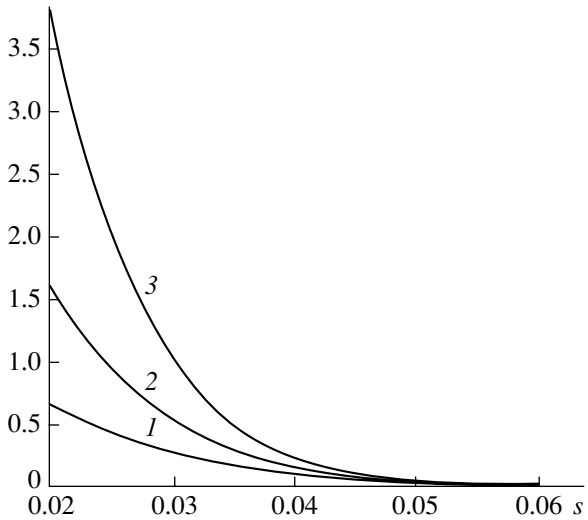


Fig. 5. Amplitude B_n of the backward wave vs. the relative change in the conductivity, s, at $a = 1$, $\tau = 0$, $\tau_d = 2$, and $\tau_T = 5$ for n = (*1*) 6, (*2*) 7, and (*3*) 8.

We numerically estimated the forward and backward wave amplitudes for a and s values typical of the InGaAsP semiconductor ($a \approx 1$, $b \approx 0.05$) [10]. Our calculations showed that the pulse-periodic modulation of the permittivity and conductivity may cause the magnitudes of the forward and backward wave amplitudes to become higher than the amplitude of the primary wave. In passive media ($\sigma \geq 0$), this effect can be achieved by modulating not only the permittivity (Figs. 2, 3), but also by modulating the conductivity of the medium, as is shown in Figs. 4 and 5.

It is seen that the modulation of the conductivity of a medium causes the amplification of the forward and backward waves at low s, i.e., when the conductivity is low or the frequency of the primary signal is high.

Calculations have also shown that the polynomial on the right-hand side of the second equality in formulas (25) vanishes at a certain combination of n, s, τ, τ_d, and τ_T. This means that the backward wave disappears at certain values of n, s, τ, τ_d, and τ_T. Obviously, if at certain values of n, s, τ, τ_d, and τ_T, the forward wave amplitude exceeds the amplitude of the primary wave

and the backward wave amplitude is zero, then the primary wave is amplified.

An important feature in the behavior of the forward and backward wave amplitudes can be revealed by studying the ratios a_{n+1}/a_n and b_{n+1}/b_n as functions of n. Matrix expressions (22) imply

$$\begin{aligned} \frac{a_{n+1}}{a_n} &= c_{11} + \frac{c_{12}}{F_n}, \\ \frac{b_{n+1}}{b_n} &= c_{22} + c_{21}F_n, \end{aligned} \tag{26}$$

where $F_n = a_n/b_n$.

By dividing the upper equality in formulas (26) by the lower one under the assumption that the backward wave is nonzero (i.e., excluding the conditions under which there is no backward wave), we obtain

$$F_{n+1} = \frac{c_{12} + c_{11}F_n}{c_{22} + c_{21}F_n}. \tag{27}$$

It was shown [11] that, when the parameters of the medium vary synchronously ($\tau = 0$), $|F_1| > 1$ at any n, s, τ_d, and τ_T. According to the conformal mapping theory [12], homographic function (27) maps the outer region of a circle of unit radius onto the outer region of a circle of unit radius. Therefore, if $|F_1| > 1$, then $|F_n| > 1$ all the more so. This means that, when the parameters of the medium vary synchronously, the amplitude of the forward wave is higher than the amplitude of the backward wave. Based on physical considerations, one can assume that, when the parameters of the medium are modulated asynchronously ($t_1 \neq 0$), the inequality $|F_1| > 1$ is also valid at any n, s, τ, τ_d, and τ_T. Indeed, our thorough numerical analysis of expression (27) failed to find such parameters n, s, τ, τ_d, and τ_T that comply with the condition $|F_1| \leq 1$. Thus, when the parameters of the medium are modulated asynchronously, the amplitude of the forward wave is, as a rule, higher than that of the backward wave.

CONCLUSION

The transformation of a monochromatic electromagnetic wave by a finite-duration pulse-periodic modulation of the permittivity and conductivity of an infinite medium is considered. The transformed field at any positive time is shown to be a superposition of the forward and backward monochromatic waves. The modules of the wavenumbers of these waves are equal to the wavenumber of the primary wave, and their amplitudes are piecewise-constant functions of time. At quiescent time intervals, the amplitude of the forward wave is higher than that of the backward wave. At certain values of the parameters of the medium and the frequency of the primary wave, the backward wave vanishes. It was also shown numerically that, at the time intervals over which the parameters are quiescent, the amplitudes of the forward and backward waves may exceed the amplitude of the primary wave. In passive media, when the modulating pulses are short, this situation may be achieved by modulating the permittivity and/or the conductivity.

REFERENCES

1. L. B. Felsen and N. Marcuvitz, *Radiation and Scattering of Waves* (Prentice-Hall, Englewood Cliffs, 1973; Mir, Moscow, 1978), Vol. 1.
2. S. I. Averkov and V. P. Boldin, Izv. Vyssh. Uchebn. Zaved., Radiofiz. **23**, 1060 (1980).
3. V. V. Borisov, *Transient Electromagnetic Waves* (Leningr. Gos. Univ., Leningrad, 1987).
4. F. A. Harfoush and A. Taflove, IEEE Trans. Antennas Propag. **39**, 898 (1991).
5. A. G. Nerukh and N. A. Khizhnyak, *Modern Problems of Nonsteady Macroscopic Electrodynamics* (Test-Radio, Kharkov, 1991).
6. V. S. Vladimirov, *Equations of Mathematical Physics* (Nauka, Moscow, 1976; Marcel Dekker, New York, 1971).
7. A. G. Nerukh, O. N. Rybin, and I. V. Shcherbatko, Zh. Tekh. Fiz. **69** (8), 84 (1999) [Tech. Phys. **44**, 945 (1999)].
8. V. V. Borisov, Geomagn. Aéron. **29**, 730 (1989).
9. M. Born and E. Wolf, *Principles of Optics* (Pergamon, Oxford, 1964; Nauka, Moscow 1973).
10. T. D. Visser, H. Block, and D. Lenstra, IEEE J. Quantum Electron. **31**, 1803 (1995).
11. O. N. Rybin, Candidate's Dissertation (Kharkov, 1999).
12. M. A. Lavrent'ev and B. V. Shabat, in *Methods of the Theory of Functions of a Complex Variable* (Nauka, Moscow, 1973), p. 736.

Translated by A. Khzmalyan

Technical Physics, Vol. 46, No. 7, 2001, pp. 796–805. Translated from Zhurnal Tekhnicheskoĭ Fiziki, Vol. 71, No. 7, 2001, pp. 14–20.

Analytical and Numerical Approaches to Calculating the Escape Function for the Emission of Medium-Energy Electrons from Uniform Specimens

L. A. Bakaleĭnikov*, E. Yu. Flegontova*, K. Yu. Pogrebitskiĭ*, Hwack-Joo Lee, Yang-Koo Cho**, Hyun-Min Park**, and Yong-Won Song****

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

***Korea Research Institute of Standards and Science, Taejon, Republic of Korea*

Received November 29, 2000

Abstract—A program for the simulation of electron transport by the Monte Carlo method has been developed. This program implies the model of single scattering and dielectric approach (to calculate the characteristics of an inelastic interaction). The escape functions for aluminum, germanium, and gold in the 0.012–20 keV energy range were calculated. The characteristic lengths determining the electron dynamics (the elastic and inelastic mean free paths, isotropization length, and integral path) were calculated using the differential cross sections for both elastic and inelastic interactions for these elements. The analysis of the relations between the characteristic lengths made it possible to determine the applicability range of the analytical expressions for the emission functions obtained in [1]. The comparison of the results obtained analytically and numerically confirmed the conclusion of [1] about the form of the analytical approximation of the emission function for electrons of various energies and showed the validity of the obtained analytical expressions for the escape lengths of electrons. © *2001 MAIK "Nauka/Interperiodica".*

INTRODUCTION

The development of new materials and technologies is connected with the creation of structures with an ultrasmall active region. The development of diagnostic techniques making it possible to determine the parameters of objects under investigation to a high accuracy is of particular importance. Nowadays, there is a wide range of methods for material diagnostics that use X-rays and electron or ion beams to get various kinds of information on objects under consideration. Among these techniques, methods based on electron emission are of particular importance since they are nondestructive and allow one to obtain information about ultrasmall near-surface layers due to the small effective escape depth (10–100 nm) of emitted electrons.

In spite of the extensive practical use of diagnostic techniques based on electron emission, they are semiempirical in many respects. In this connection, the development of quantitative methods for obtaining information on material properties requires, first of all, the development of adequate theoretical approaches to calculating the basic phenomenon, i.e., electron emission.

A three-step model may be employed to describe electron emission. According to this model, electron emission induced by X-rays can be divided into three stages: the absorption of an X-ray photon by a medium; the generation of primary electrons, i.e., photoelectrons and electrons generated in the course of the relaxation of the photoionized atom; and the motion of primary electrons in the medium (accompanied by the generation of a secondary-electron cascade) and the emission of the produced electrons from the specimen. The last stage is characterized by the escape function $q(E_0,z_0)$. This function is the probability of escape from the specimen for an electron with an energy E_0 produced at a distance z_0 from the surface. In order to find this function, a kinetic equation describing the motion of electrons generated at a depth z_0 with an energy E_0 may be used. In the case of a semi-infinite specimen with a plane surface, this equation has the form

$$\cos\Theta\frac{\partial\Phi}{\partial z} = R_{\text{col}}\Phi + \frac{1}{4\pi}\delta(z-z_0)\delta(E-E_0). \qquad (1)$$

Here, $\Phi(z, \Theta, E)$ is the differential flux density for electrons with energy E moving at a depth z at the angle Θ to the z axis, $R_{\text{col}}\Phi$ is the collision integral describing the interaction of electrons with a medium, and the second term on the right-hand side corresponds to the presence of the electron source with energy E_0 at a depth z_0. Equation (1) must be completed by the boundary conditions $\Phi(0, \Theta, E)|_{\cos\Theta > 0} = 0$ and $\Phi(z, \Theta, E)|_{z\to\infty} \longrightarrow 0$. These conditions mean the absence of both the electron flux from the outside and electrons at large depths. The

escape function is equal to the total electron flux through the surface $z = 0$:

$$q(E_0, z_0) = \int_0^{E_0} \int_0^{2\pi} \int_0^{\pi} \cos\Theta \Phi(0, \Theta, E) \sin\Theta d\Theta d\varphi dE. \quad (2)$$

Equation (1) represents a linear integrodifferential equation. Solving this equation involves severe difficulties. In this connection, two approaches may be used to determine the differential flux density: (1) to take into account some features of electron transport that make it possible to simplify the kinetic equation and find an approximate analytical solution; (2) to numerically solve the kinetic equation.

In [1], assumptions resulting in the simplification of kinetic equation (1) in various ranges of primary-electron energies were analyzed and approximate analytical expressions for the escape functions were found. However, the applicability boundaries of the approximations were obtained just qualitatively. In order to find the applicability ranges and the accuracy of the solutions obtained, it is necessary to compare analytical results with the numerical calculation of the escape functions. A universal technique for numerical simulation of transport phenomena allowing one to take into account the features of the interaction between electrons and a medium is the Monte Carlo method.

THE MODEL DESCRIPTION

In the Monte Carlo method, various models may describe the interaction between an electron and a medium. The most accurate is the model of single scattering. In this model, the result of each interaction of an electron with a medium is defined by the differential cross sections for both elastic and inelastic scattering. In inelastic interaction between an electron and medium, the electron kinetic energy is lost in the course of several competing processes (ionization of inner atomic shells, creation of electron-hole pairs, generation of plasmons, etc.). A uniform description of inelastic interaction allowing for various channels of energy loss is possible by the formalism based on the dielectric permittivity $\tilde{\varepsilon}$. According to this approach, the doubly differential reciprocal mean free path, i.e., the probability of losing energy $\hbar\omega$ and momentum $\hbar q$ by an electron of energy E per unit length, can be represented in the form [2]

$$\frac{d^2 l_{\text{in}}^{-1}}{d(\hbar\omega) d(\hbar q)} = \frac{1}{\pi a_0 q E} \text{Im}\left(-\frac{1}{\tilde{\varepsilon}(q, \omega)}\right). \quad (3)$$

Here, E is the electron energy, which is counted off from the bottom of the conduction band, and $a_0 = \hbar^2/me^2$ is the Bohr radius. The energy-loss function

$$\text{Im}\left(-\frac{1}{\tilde{\varepsilon}(q, \omega)}\right)$$

can be found by the extrapolation of the optical data $\tilde{\varepsilon}(0, \omega)$ in the entire plane (q, ω). A method for such extrapolation is proposed in [3–5].

The distribution in the energy loss in the inelastic interaction is easily calculated from (3) by integrating with respect to the momentum,

$$\frac{dl_{\text{in}}^{-1}}{d(\hbar\omega)} = \int_{\hbar q_-}^{\hbar q_+} \frac{d^2 l_{\text{in}}^{-1}}{d(\hbar\omega) d(\hbar q)} d\hbar q. \quad (4)$$

Here, q_- and q_+ are the minimal and maximal values of the transferred momentum, which are determined from the conservation law for the energy and momentum. The reciprocal mean free path l_{in}^{-1} for the inelastic interaction of an electron with a medium is calculated by integrating (3) over the entire $\hbar\omega$ and $\hbar q$ and may be expressed in the form

$$l_{\text{in}}^{-1} = \int_0^{Q_{max}} \frac{dl_{\text{in}}^{-1}}{d(\hbar\omega)} d\hbar\omega,$$

where $Q_{\max}$ is the maximal possible energy loss.

The calculated reciprocal mean free path for Au and Al is in good agreement with the results reported in [5] in a wide energy range.

In the simulation of an electron trajectory, the energy loss Q is defined by the relation $F(Q) = R$, where R is a random number uniformly distributed in the range [0, 1] and $F(Q)$ is given by the expression

$$F(Q) = \frac{\int_0^Q \frac{dl_{\text{in}}^{-1}}{d(\hbar\omega)} d\hbar\omega}{l_{\text{in}}^{-1}}.$$

We developed an effective way of approximating the function $F^{-1}(R)$. The parameters of this approximation and the differential (with respect to the energy) reciprocal mean free path (4) were calculated for a number of media on the basis of optical data [6]. These data are listed in [7].

The Mott differential cross section was used to calculate the elastic interaction of electrons with atoms of a medium in the Monte Carlo method. This cross section was determined by partial-wave expansion for the elements of the periodic table with atomic numbers Z from 1 to 103 in a wide range of energies. These data are also in [7].

ANALYTICAL APPROXIMATIONS OF THE ESCAPE FUNCTION

In [8], it is shown that the kinetic equation may be simplified in view of different scales of the angular and energy relaxations. The characteristic lengths of the

angular and energy relaxations, i.e., $\lambda(E)$ and $S(E)$, can be determined in the following way:

$$\lambda(E) = 2\left(2\pi N\int_0^\pi \frac{d\sigma_{el}(E,\Theta)}{d\Theta}\Theta^2\sin\Theta d\Theta\right)^{-1},$$

$$S(E) = E/\beta(E).$$

Here, $(d\sigma_{el}(E, \Theta))/d\Theta$ is the differential cross section for the elastic scattering, N is the scatterer density, and $\beta(E)$ is the mean energy loss per unit length expressed by the formula

$$\beta(E) = \int_0^E \hbar\omega\frac{dl_{in}^{-1}}{d\hbar\omega}d\hbar\omega.$$

Other characteristic lengths of the electron propagation in a medium are the mean free path for elastic,

$$l_{el}(E) = (N\sigma_{el})^{-1} = \left(N2\pi\int_0^\pi \frac{d\sigma_{el}(E,\Theta)}{d\Theta}\sin\Theta d\Theta\right)^{-1}$$

Table 1. The elastic and inelastic mean free paths l_{el} and l_{in}, respectively, the isotropization length λ, the energy relaxation length S, and the ratio λ/S for Al

E, keV	$l_{el}\times10^{-9}$ m	$l_{in}\times10^{-9}$ m	$\lambda\times10^{-9}$ m	$S\times10^{-9}$ m	λ/S
0.01	1.15	41.7	1.12	69.6	0.016
0.015	0.890	26.3	0.857	42.8	0.020
0.02	0.693	4.44	0.640	5.25	0.122
0.025	0.576	0.589	0.496	0.726	0.684
0.03	0.508	0.448	0.413	0.635	0.650
0.04	0.447	0.429	0.340	0.771	0.441
0.05	0.429	0.411	0.322	0.895	0.360
0.06	0.422	0.430	0.319	1.10	0.290
0.07	0.423	0.451	0.326	1.32	0.246
0.08	0.453	0.473	0.379	1.56	0.242
0.09	0.462	0.497	0.402	1.83	0.220
0.1	0.474	0.523	0.428	2.11	0.203
0.2	0.631	0.769	0.879	5.40	0.163
0.3	0.780	0.977	1.48	8.53	0.174
0.4	0.919	1.18	2.21	12.0	0.185
0.5	0.106	1.37	3.14	15.9	0.198
0.6	0.117	1.56	3.96	19.9	0.199
0.7	0.129	1.74	4.92	25.0	0.197
0.8	0.141	1.92	6.08	30.1	0.202
0.9	0.152	2.10	7.17	36.0	0.199
1.0	0.163	2.27	8.54	42.1	0.203
2.0	0.266	3.94	25.4	130	0.196
3.0	0.363	5.49	49.2	262	0.188
4.0	0.458	6.98	79.4	425	0.187
5.0	0.553	8.41	116	620	0.186
6.0	0.647	9.82	158	850	0.186
7.0	0.739	11.2	206	1120	0.184
8.0	0.831	12.6	259	1420	0.183
9.0	0.922	13.9	317	1740	0.182
10.0	10.1	15.2	381	2080	0.184
15.0	14.6	21.6	774	4320	0.179
20.0	19.1	27.7	1280	7080	0.181

and inelastic,

$$l_{\text{in}}(E) = (l_{\text{in}}^{-1})^{-1} = \left(\int_0^{Q_{\max}} \frac{dl_{\text{in}}^{-1}}{d\hbar\omega} d\hbar\omega \right)^{-1}$$

collisions. If the energy or angular change is small in a single scattering event, the mean free paths would essentially differ from the relaxation lengths.

In the case in which the angular-relaxation length is less than the length of energy relaxation, i.e., the integral electron path, the introduction of the variable $\sqrt{\lambda(E_0)S(E_0)}$ as a unit length makes it possible to separate the small parameter $\varepsilon = \sqrt{\lambda(E_0)S(E_0)}$ and simplify the kinetic equation by perturbation-theory techniques. Note, first of all, that in the case under consideration, the angular rearrangement in interaction of an electron with a medium is determined basically by the elastic scattering. Therefore, the collision integral $R_{\text{col}}\Phi$ can be approximately expressed as a sum of the elastic

Table 2. The same as in Table 1, for Ge

E, keV	$l_{\text{el}} \times 10^{-9}$ m	$l_{\text{in}} \times 10^{-9}$ m	$\lambda \times 10^{-9}$ m	$S \times 10^{-9}$ m	λ/S
0.01	0.188	19.9	0.235	4.24*E*-06	0.00553
0.015	0.316	8.67	0.379	1.76*E*-06	0.0215
0.02	0.455	3.71	0.579	7.06	0.082
0.025	0.577	1.97	0.807	3.70	0.218
0.03	0.659	1.29	1.00	2.50	0.4
0.04	0.701	0.868	1.12	1.85	0.606
0.05	0.664	0.759	1.02	1.82	0.562
0.06	0.626	0.748	0.925	1.98	0.467
0.07	0.595	0.750	0.845	2.18	0.389
0.08	0.566	0.764	0.757	2.40	0.316
0.09	0.548	0.784	0.714	2.64	0.270
0.1	0.536	0.808	0.683	2.90	0.236
0.2	0.575	1.08	0.713	5.59	0.128
0.3	0.656	1.33	0.908	8.45	0.108
0.4	0.737	1.57	1.17	11.6	0.101
0.5	0.817	1.81	1.48	15.2	0.0977
0.6	0.889	2.04	1.82	19.3	0.0942
0.7	0.958	2.28	2.18	23.6	0.0924
0.8	1.03	2.50	2.60	28.4	0.0917
0.9	1.09	2.73	3.01	33.1	0.0908
1.0	1.15	2.95	3.50	38.0	0.0923
2.0	1.69	4.98	9.44	107	0.0885
3.0	2.14	6.88	17.3	203	0.0855
4.0	2.56	8.70	26.9	320	0.0840
5.0	2.96	10.5	38.0	461	0.0825
6.0	3.34	12.2	50.6	620	0.0815
7.0	3.71	13.8	64.5	803	0.0803
8.0	4.06	15.5	79.7	1010	0.0792
9.0	4.41	17.1	96.1	1230	0.0784
10.0	4.75	18.7	114	1470	0.0774
15.0	6.40	26.5	219	2950	0.0743
20.0	7.99	34.0	352	4830	0.0728

($R_{el}\Phi$) and inelastic ($R_{in}\Phi$) terms

$$R_{el}\Phi = N\int_{4\pi}\frac{d\sigma_{el}(E, \bar{\Omega}'\bar{\Omega})}{d\Omega}\Phi(z, \bar{\Omega}', E)d\bar{\Omega}'$$

$$- N\Phi(z, \bar{\Omega}, E)\int_{4\pi}\frac{d\sigma_{el}(E, \bar{\Omega}\bar{\Omega}')}{d\Omega}d\bar{\Omega}',$$

$$R_{il}\Phi = \int_0^{E_0}\frac{dl_{in}^{-1}(E+\hbar\omega, \hbar\omega)}{d(\hbar\omega)}$$

$$\times\Phi(z, \bar{\Omega}, E+\hbar\omega)d\hbar\omega - \Phi(z, \bar{\Omega}, E)l_{in}^{-1}(E),$$

where N is the scatterer density.

The kinetic equation for the differential flux density is replaced by the equation for the isotropic part of the distribution that depends only on the depth and energy. In terms of dimensionless variables, this equation has the form [8]

$$\frac{1}{3\nu_1(E)}\frac{\partial^2 U}{\partial z^2} - S(E_0)R_{in}U - \delta(z-z_0)\delta(E-1) = 0. \quad (5)$$

Here, $\sqrt{\lambda(E_0)S(E_0)}$ is taken as a unit length; the initial electron energy E_0, a unit energy; and ν_1 is the first eigenvalue of the elastic term in the collision integral. As it is noted in [8], the eigenfunctions of the operator

Table 3. The same as in Tables 1 and 2, for Au

E, keV	$l_{el}\times10^{-9}$ m	$l_{in}\times10^{-9}$ m	$\lambda\times10^{-9}$ m	$S\times10^{-9}$ m	λ/S
0.01	0.510	3.49	0.212	5.75	0.0369
0.015	0.735	2.30	0.226	4.20	0.0538
0.02	0.612	1.76	0.211	3.41	0.0621
0.025	0.430	1.44	0.188	2.88	0.0652
0.03	0.313	1.22	0.169	2.51	0.0672
0.04	0.211	0.999	0.161	2.20	0.0731
0.05	0.181	0.852	0.183	1.91	0.0959
0.06	0.168	0.736	0.218	1.68	0.130
0.07	0.161	0.650	0.262	1.50	0.175
0.08	0.174	0.592	0.322	1.39	0.232
0.09	0.173	0.554	0.394	1.33	0.296
0.1	0.174	0.530	0.483	1.31	0.368
0.2	0.222	0.542	0.869	1.90	0.457
0.3	0.290	0.644	0.873	3.03	0.288
0.4	0.351	0.760	0.891	4.49	0.199
0.5	0.405	0.871	0.982	6.07	0.153
0.6	0.442	0.980	0.982	7.80	0.126
0.7	0.475	1.09	1.04	9.63	0.108
0.8	0.507	1.19	1.11	11.6	0.0965
0.9	0.533	1.29	1.19	13.6	0.0878
1.0	0.560	1.39	1.28	15.6	0.0819
2.0	0.766	2.32	2.46	41.3	0.0596
3.0	0.932	3.18	4.06	76.4	0.0531
4.0	1.08	4.01	5.97	121	0.0491
5.0	1.21	4.80	8.13	174	0.0469
6.0	1.33	5.57	10.5	231	0.0456
7.0	1.45	6.32	13.1	299	0.0439
8.0	1.56	7.06	15.9	371	0.0430
9.0	1.67	7.70	18.9	451	0.0420
10.0	1.77	8.51	22.1	535	0.0413
15.0	2.23	12.0	40.3	1060	0.0381
20.0	2.66	15.4	62.0	1720	0.0361

$R_{el}\Phi$ are the spherical harmonics $Y_{i,k}(\Omega)$, and ν_1 can be found from the equation $R_{el}Y_{1.0} = \nu_1 Y_{1.0}$.

Opportunities for further simplification of electron kinetics are connected to the features of the inelastic interaction of electrons with a medium. The form of the escape function is defined by the relation between the inelastic mean free path l_{in} and the integral electron path S. If the energy losses in the inelastic scattering are comparable to the electron energy and an electron loses all its energy in the course of several inelastic collisions, the inelastic mean free path would be of the same order of magnitude as the integral electron path. If an electron loses all its energy in the first inelastic collision, the number of particles in a phase-volume element does not increase due to collisions. Hence, the inelastic collision integral can be represented in the form

$$R_{in}U = -\frac{U}{l_{in}(E)}. \quad (6)$$

After replacing $R_{in}U$ by (6), Eq. (5) turns into an ordinary differential equation and its solution results in the escape function of the form

$$q(z_0) = C_1 \exp\left(-\frac{z_0}{L_1}\right), \quad (7)$$

where the escape length L_1 is given by

$$L_1(E_0) = \sqrt{\lambda(E_0) l_{in}(E_0)/3|\nu_1(E_0)|}. \quad (8)$$

Table 4. The parameters of the analytical approximation of the escape function and the approximation error for Al

E, keV	$L_1 \times 10^{-9}$ m	C_1	I_1	$L_2 \times 10^{-9}$ m	C_2	I_2
0.015	2.93	0.827	0.0010	5.79	0.694	0.0203
0.02	1.22	0.633	0.0102	2.40	0.534	0.0265
0.025	0.339	0.417	0.0119	0.652	0.348	0.0180
0.03	3.19	0.380	0.0776	6.19	0.321	0.0181
0.04	2.16	0.426	0.0212	4.64	0.314	0.0389
0.05	2.83	0.410	0.0173	5.71	0.331	0.0299
0.1	2.46	0.446	0.0188	4.93	0.363	0.0282
0.2	2.15	0.716	0.0076	4.50	0.559	0.0238
0.3	2.51	0.801	0.0046	5.03	0.654	0.0143
0.4	3.11	0.809	0.0089	6.13	0.669	0.1012
0.5	3.78	0.805	0.0104	7.56	0.663	0.0087
0.6	4.41	0.812	0.0139	8.93	0.667	0.0061
0.7	5.32	0.795	0.0137	10.7	0.659	0.0056
0.8	6.21	0.789	0.0144	12.3	0.659	0.0047
0.9	7.12	0.786	0.0140	14.3	0.649	0.0050
1.0	8.22	0.782	0.0147	16.4	0.652	0.0042
1.5	14.8	0.750	0.0156	29.3	0.629	0.0039
2.0	22.1	0.766	0.0160	43.6	0.642	0.0026
2.5	31.9	0.758	0.0172	62.2	0.641	0.0014
3.0	43.2	0.748	0.0166	83.7	0.637	0.0030
4.0	67.6	0.760	0.0179	131	0.645	0.0025
5.0	98.8	0.754	0.0194	190	0.647	0.0020
6.0	135	0.747	0.0194	259	0.646	0.0023
7.0	176	0.750	0.0203	333	0.649	0.0026
8.0	221	0.751	0.0200	420	0.652	0.0028
9.0	275	0.744	0.0210	517	0.648	0.0035
10.0	327	0.745	0.0211	617	0.650	0.0036
15.0	671	0.740	0.0205	1260	0.649	0.0038
20.0	1100	0.747	0.0223	2080	0.649	0.0049

A different condition takes place for fast electrons. In this case, the inelastic mean free path turns out to be much less than the integral path. This allows one to employ the continuous slowing-down approximation

$$R_{\text{in}}U = \frac{\partial}{\partial E}(\beta(E)U).$$

The kinetic equation is reduced to the diffusion equation [8] in this case. Solving this equation yields

$$q(z_0) = C_2\left[1 - \text{erf}\left(\frac{z_0}{L_2}\right)\right]. \quad (9)$$

Here, the escape length L_2 is related to the lengths of the

Table 5. The same as in Table 4, for Ge

E, keV	$L_1 \times 10^{-9}$ m	C_1	I_1	$L_2 \times 10^{-9}$ m	C_2	I_2
0.012	1.61	0.800	0.0008	3.24	0.651	0.0183
0.014	1.59	0.778	0.0008	3.23	0.628	0.0182
0.015	1.59	0.758	0.0007	3.24	0.608	0.0181
0.016	1.58	0.737	0.0010	3.15	0.599	0.0173
0.017	1.57	0.709	0.0026	3.16	0.579	0.0185
0.018	1.55	0.688	0.0018	3.13	0.558	0.0174
0.02	1.57	0.657	0.0022	3.15	0.534	0.0171
0.025	1.62	0.646	0.0025	3.24	0.528	0.0169
0.03	1.62	0.676	0.0011	3.26	0.552	0.0161
0.04	1.65	0.692	0.0007	3.26	0.570	0.0157
0.05	1.66	0.699	0.0014	3.31	0.577	0.0147
0.1	1.77	0.770	0.0041	3.60	0.628	0.0144
0.2	1.97	0.859	0.0077	3.93	0.703	0.0124
0.3	2.34	0.872	0.0106	4.61	0.721	0.0096
0.4	2.69	0.891	0.0129	5.44	0.729	0.0085
0.5	3.13	0.884	0.0135	6.35	0.726	0.0079
0.6	3.63	0.887	0.0136	7.37	0.727	0.0081
0.7	4.18	0.873	0.0135	8.50	0.715	0.0076
0.8	4.77	0.870	0.0141	9.69	0.712	0.0070
0.9	5.32	0.874	0.0148	10.8	0.715	0.0069
1.0	6.07	0.853	0.0136	12.2	0.708	0.0072
1.1	6.70	0.861	0.0143	13.5	0.710	0.0065
1.2	7.56	0.844	0.0140	15.2	0.698	0.0063
1.3	8.45	0.834	0.0136	16.9	0.694	0.0066
1.5	10.3	0.840	0.0143	20.6	0.699	0.0061
2.0	14.5	0.862	0.0155	29.2	0.709	0.0058
2.5	20.3	0.832	0.0153	40.7	0.688	0.0050
3.0	25.7	0.855	0.0162	52.1	0.699	0.0047
4.0	39.3	0.845	0.0160	79.6	0.693	0.0046
5.0	56.1	0.840	0.0165	112	0.699	0.0040
6.0	74.0	0.847	0.0170	148	0.702	0.0038
7.0	94.3	0.839	0.0178	188	0.698	0.0029
8.0	117	0.843	0.0181	229	0.709	0.0023
9.0	144	0.830	0.0184	277	0.708	0.0021
10.0	168	0.833	0.0176	331	0.700	0.0031
11.0	194	0.841	0.0184	385	0.705	0.0026
12.0	225	0.844	0.0190	442	0.709	0.0015
15.0	324	0.842	0.0187	639	0.707	0.0026
20.0	612	0.774	0.0133	1030	0.711	0.0037

angular and energy relaxations by the relation

$$L_2(E_0) = 2\sqrt{\lambda(E_0)S(E_0)\tau^*}, \quad (10)$$

where the dimensionless coefficient τ^* has the form

$$\tau^* = \int_{E_0}^{0} \frac{dE}{3S(E_0)\lambda(E_0)\beta(E)\nu_1(E)}. \quad (11)$$

The diffusion approximation is valid far from the surface (at depths of the order of the isotropization length and greater). Therefore, solutions (7) and (9) are correct to an accuracy of energy-dependent factors C_1 and C_2, which are defined by the matching condition for the solutions of the kinetic equation in the diffusion region with those in the near-surface layer. In practice, the coefficients C_1 and C_2 may be found by using the results of numerical calculation of the escape function.

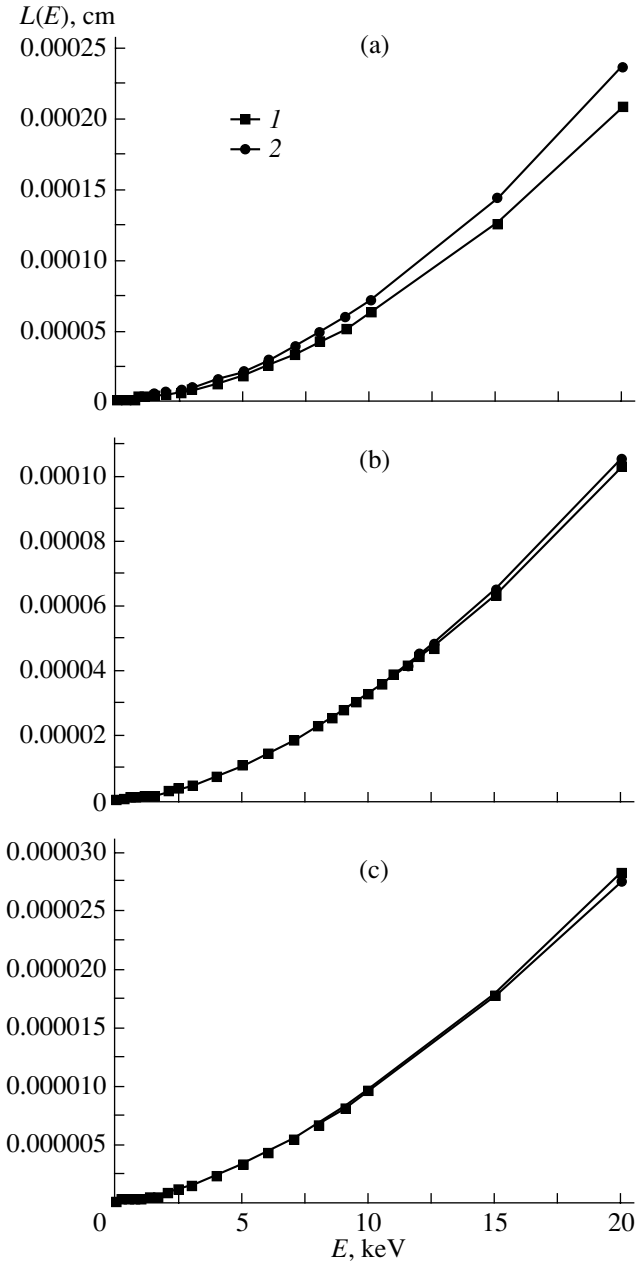
RESULTS AND DISCUSSION

With the program developed, we calculated the escape function for the mediums with rather different atomic numbers Al (Z = 13), Ge (Z = 32), and Au (Z = 79) in a wide range of primary electron energies (from tens of electronvolts to tens of kiloelectronvolts).

The characteristic lengths determining the electron kinetics were calculated in order to determine the

Table 6. The same as in Tables 4 and 5, for Au

E, keV	$L_1 \times 10^{-9}$ m	C_1	I_1	$L_2 \times 10^{-9}$ m	C_2	I_2
0.012	0.626	0.624	0.0050	1.27	0.513	0.0197
0.015	0.631	0.625	0.0047	1.28	0.506	0.0188
0.02	0.626	0.666	0.0043	1.26	0.540	0.0191
0.025	0.616	0.696	0.0025	1.26	0.559	0.0183
0.03	0.616	0.715	0.0027	1.26	0.574	0.0189
0.04	0.640	0.754	0.0019	1.29	0.612	0.0189
0.05	0.654	0.771	0.0013	1.30	0.632	0.0166
0.06	0.668	0.780	0.0019	1.34	0.634	0.0161
0.08	0.691	0.799	0.0038	1.38	0.652	0.0147
0.1	0.721	0.802	0.0056	1.43	0.662	0.0132
0.2	0.888	0.815	0.0114	1.81	0.665	0.0084
0.3	1.08	0.821	0.0138	2.18	0.675	0.0061
0.4	1.26	0.832	0.0150	2.55	0.684	0.0053
0.5	1.45	0.846	0.0151	2.95	0.693	0.0054
0.6	1.65	0.854	0.0154	3.35	0.698	0.0053
0.7	1.84	0.868	0.0160	3.75	0.709	0.0052
0.8	2.08	0.864	0.0149	4.23	0.707	0.0060
0.9	2.30	0.869	0.0149	4.65	0.714	0.0062
1.0	2.56	0.870	0.0141	5.07	0.718	0.0063
1.5	3.84	0.873	0.0132	7.72	0.717	0.0077
2.0	5.25	0.877	0.0137	10.6	0.719	0.0072
2.5	6.98	0.871	0.0133	14.1	0.714	0.0074
3.0	8.69	0.881	0.0142	17.1	0.729	0.0066
4.0	12.6	0.888	0.0145	25.8	0.725	0.0070
5.0	17.5	0.880	0.0145	35.3	0.727	0.0068
6.0	22.3	0.892	0.0153	45.4	0.731	0.0064
7.0	27.9	0.902	0.0168	56.8	0.738	0.0050
8.0	33.9	0.904	0.0169	69.1	0.738	0.0050
9.0	40.8	0.898	0.0174	81.7	0.745	0.0042
10.0	48.4	0.895	0.0161	96.9	0.743	0.0057
15.0	89.1	0.905	0.0175	178	0.751	0.0045
20.0	137	0.921	0.0184	278	0.755	0.0039



Escape lengths for (a) aluminum, (b) germanium, and (c) gold obtained from (*1*) the approximate solution of the kinetic equation and (*2*) the results of the Monte Carlo simulation.

energy ranges in which the assumptions used in the derivation of (7) and (9) are valid. In this study, we used the Mott differential cross section for the elastic scattering, as well as the differential (with respect to the energy) reciprocal mean free path (4). The characteristic lengths for aluminum, germanium, and gold are listed in Tables 1–3. These data show that starting with some energy of the order of hundreds of electronvolts, the ratio λ/S does not exceed 0.2 for all the elements. Note that λ/S decreases with increasing atomic number. At lower energies, except for the region of very small energies ($E < 20$ eV for Al and Ge, and $E < 50$ eV for Au), this ratio is not small, and the processes of angular and energy relaxations cannot be separated. Tables 1–3 show that the inelastic mean free path in an energy range of $E > 1$keV is much less than the integral path. This makes it possible to use the continuous slowing-down approximation. In this connection, the use of expression (9) for the escape function is justified in this energy range. At energies of the order of tens of electronvolts, the mean free path is comparable to the integral path. Hence, expression (7) may be valid to describe the form of the escape function.

In order to solve the problem of the accuracy of the analytical approximations, we compared the dependences of the escape function on the depth $q_{MC}(z)$ obtained by numerical simulation with the analytical expressions (7) and (9). In this study, the quantities L_1, L_2, C_1, and C_2 were treated as free parameters that were determined by numerical minimization of the functional

$$I = \sqrt{\frac{1}{N}\sum_{i=1}^{N}[q_{MC}(z_i) - q(z_i)]^2}. \qquad (12)$$

The parameters obtained, as well as the values of the functional for Al, Ge, and Au, are listed in Tables 4–6. It turns out that in the small-energy range (of the order of tens of electronvolts), expression (7) yields smaller values of I than (9); i.e., expression (7) is a better approximation of the escape function. The reverse situation takes place at energies $E > 1$ keV. In this energy range, the escape function is better approximated by (9). Finally, in the range of the intermediate energies, both analytical approximations have similar accuracy. The value of the functional turns out to be two times as great as that for $E > 1$ keV and $E < 50$ eV. The data obtained are in good agreement with the results of the analysis of characteristic lengths for the electron transport process.

The comparison of escape lengths obtained by minimization of the functional (12) with those calculated by formulas (10) and (11) is of considerable interest. The energy dependences of the escape lengths minimizing the functional (12) and those calculated by the formulas (10) and (11) are shown in the figure. In the case of gold and germanium at electron energies $E_0 > 2$ keV, i.e., in the energy range in which $\lambda \ll S$ and $l_{in} \ll S$, both dependences coincide up to the error of the minimization procedure. In this energy range, the escape lengths calculated by both methods for germanium and gold differ from each other by less than 3%. At the same time, the accuracy in determining the escape length by minimizing the functional (12) is about 10%. For aluminum, the discrepancy in escape lengths turns out to be about 15%. The reason is apparently a relatively large value of $\lambda(E_0)/S(E_0)$.

At an electron energy so low that it is likely lost in the first inelastic collision, one may expect that the escape length L_1 involved in (7) satisfies relation (8). In order to verify this assumption, the escape lengths L_1 obtained by minimization of the functional (12) are compared to those calculated by formula (8). For electrons with energies of 10–25 eV, the discrepancy is about 10%.

CONCLUSION

A program for simulating electron transport by the Monte Carlo method using the single-scattering model was developed. The properties of inelastic interaction were calculated using the dielectric approach. The escape functions for aluminum, germanium, and gold were calculated in an energy range of 0.012–20 keV.

The escape functions calculated were used to determine the applicability ranges for the analytical expressions obtained in [1]. Using the differential cross sections for elastic and inelastic interactions, the characteristic lengths determining the electron kinetics, i.e., the elastic and inelastic mean free paths, the isotropization length, and the integral path, were calculated. The analysis of relations between these quantities made it possible to determine the ranges of applicability for the assumptions used in [1] to derive the analytical expressions for the escape function. The comparison between the analytical and numerical results confirmed the conclusions of [1] about the form of analytical approximation of the escape function for electrons of various energies and showed the validity of the analytical expressions obtained for electron escape lengths.

REFERENCES

1. L. A. Bakaleĭnikov, K. Yu. Pogrebitskiĭ, E. A. Tropp, *et al.*, Nucleus **34**, 1 (1997).
2. D. Pines, *Elementary Excitations in Solids* (Benjamin, New York, 1963; Mir, Moscow, 1965).
3. D. R. Penn, Phys. Rev. B **35**, 482 (1987).
4. Z.-J. Ding and R. Shimizu, Surf. Sci. **222**, 313 (1989).
5. Z.-J. Ding, *Fundamental Studies on the Interactions of kV Electrons with Solids for Applications to Electron Spectroscopies*, PHD Thesis (Osaka Univ., Osaka, 1990).
6. E. D. Palik, *Handbook of Optical Constants of Solids* (Academic, New York, 1985).
7. http://www.ioffe.rssi.ru/ES.
8. L. A. Bakaleĭnikov, S. G. Konnikov, K. Yu. Pogrebitskiĭ, *et al.*, Zh. Tekh. Fiz. **64** (4), 9 (1994) [Tech. Phys. **39**, 354 (1994)].

Translated by M. Fofanov

Technical Physics, Vol. 46, No. 7, 2001, pp. 806–814. Translated from Zhurnal Tekhnicheskoĭ Fiziki, Vol. 71, No. 7, 2001, pp. 21–29.

GASES AND LIQUIDS

Dynamics of the Free Surface of a Conducting Liquid in a Near-Critical Electric Field

N. M. Zubarev and O. V. Zubareva

Institute of Electrophysics, Ural Division, Russian Academy of Sciences, Ekaterinburg, 620016 Russia

e-mail: nick@ami.uran.ru

Received October 18, 2000

Abstract—Near-critical behavior of the free surface of a perfectly conducting liquid in an external electric field is considered. Based on an analysis of three-wave processes using the method of integral estimates, sufficient criteria for hard instability of a planar surface are formulated. It is shown that the higher-order nonlinearities do not saturate the instability, for which reason the growth of disturbances has an explosive character. © *2001 MAIK "Nauka/Interperiodica".*

INTRODUCTION

The electrohydrodynamic instability of the free surface of a conducting liquid in a strong electric field [1, 2] is responsible for many physical processes, such as the initiation and keeping up the emission of charged particles, vacuum breakdown, vacuum discharge, etc. The interaction of an electric field and charges induced by this field on the surface of a liquid (liquid metal) leads to a growth of surface disturbances and the formation of regions with a significant curvature [3–5]. The dispersion law for the waves on a planar surface of a perfectly conducting liquid in an external electric field of strength E has the following form [6]:

$$\omega^2 = g|\mathbf{k}| + \frac{\alpha}{\rho}|\mathbf{k}|^3 - \frac{E^2}{4\pi\rho}|\mathbf{k}|^2, \qquad (1)$$

where ω is the frequency, $\mathbf{k}$ is the wave vector, g is the acceleration of gravity, α is the surface tension coefficient, and ρ is the density of the medium.

It is seen from Eq. (1) that if the condition

$$E^2 < E_c^2 = 8\pi\sqrt{g\alpha\rho}$$

is fulfilled, then $\omega^2 > 0$ at any $|\mathbf{k}|$ and, consequently, the surface disturbances do not grow with time. If the magnitude of the field E, which plays the role of an external governing parameter, exceeds a certain critical value E_c, then there arises a region of wave vectors $|\mathbf{k}|$ for which $\omega^2 < 0$, which corresponds to an aperiodic instability. Thus, the condition $E > E_c$ is a criterion for the surface instability with respect to infinitely small disturbances of the surface shape and of the field of velocities.

It was shown in [7–9], where liquids with various physical properties have been considered, that a nonlinear interaction of three standing waves that form a hexagonal structure can lead to a hard excitation of the instability of a charged surface. In our case, this means that, even at subcritical fields ($E < E_c$), a disturbance of a sufficient magnitude can break the equilibrium of a planar surface. In this connection, there arises a need of constructing criteria for the instability of a charged surface of a conducting liquid with respect to perturbations of a finite magnitude, i.e., criteria that will permit one, proceeding from some initial data such as the shape of the surface and the distribution of velocities, to answer the question of whether or not the initial perturbation will lead to the loss of the stability of a planar boundary and, as a consequence, to an explosive growth of cusp-like structures. This work is devoted to constructing such criteria using the method of integral estimates that was applied previously to obtain the conditions of collapse for the nonlinear Schrödinger equation [10, 11], nonlinear Klein–Gordon equation [12, 13], and various modifications of the Boussinesq equation [14].

In Section 1, we give equations of the vertex-free motion of a perfectly conducting liquid with a free surface in an electric field and give their Hamiltonian formulation. In Section 2, the theory of perturbations in a small parameter, namely, in the characteristic angle of the surface slope, is constructed up to fourth-order terms in the Hamiltonian. The analysis of the surface dynamics is significantly simplified in the case of small "supercriticalities" (i.e., if the field only slightly exceeds the critical value E_c)

$$\varepsilon = (E^2 - E_c^2)/E_c^2, \quad |\varepsilon| \ll 1,$$

when only perturbations with wave numbers close to $k_0 = \sqrt{g\rho/\alpha}$ increase (this value of the wave number corresponds to the so-called dominant harmonic of surface perturbations). This permits us in Section 3 to construct a set of amplitude equations for describing the nonlinear interaction of three standing waves that form a hexagonal structure which is the main interaction at near-critical values of the field E. In Section 4, we

extend the method of integral estimates for several interacting nonlinear waves. By using this method, one can pass from a set of partial differential equations for complex amplitudes A_1, A_2, and A_3 to a second-order differential inequality for the norm

$$X = \int (|A_1|^2 + |A_2|^2 + |A_3|^2) d^2r,$$

by analyzing which we obtain a number of sufficient criteria of the hard excitation of an electrohydrodynamic instability of a charged surface. Note that most of them refer to subcritical values of the external electric field, when the surface is stable in a linear approximation, while the development of the instability is related to three-wave processes. In Section 5, we show, using the example of one-dimensional and square lattices of surface perturbations for which the three-wave interactions degenerate, that the higher-order wave processes do not saturate the instability but, on the contrary, lead to an explosive growth of amplitudes.

INITIAL EQUATIONS

Consider the potential motion of an ideal conducting liquid of infinite depth placed in an external uniform electric field of strength E. Assume that the field-strength vector is directed along the z axis and, correspondingly, in the unperturbed state the boundary of the liquid is a planar horizontal surface $z = 0$. Let the function $\eta(x, y, t)$ specify the deviation of the boundary from the flatness, i.e., the region occupied by the liquid is restricted by a free surface $z = \eta$.

The velocity potential for an incompressible liquid Φ satisfies Laplace's equation

$$\nabla^2 \Phi = 0 \tag{2}$$

with the following conditions at the metal–vacuum boundary and at infinity:

$$\frac{\partial \Phi}{\partial t} + \frac{(\nabla \Phi)^2}{2} = \frac{(\nabla \varphi)^2 - E^2}{8\pi\rho} + \frac{\alpha}{\rho} \nabla_\perp \frac{\nabla_\perp \eta}{\sqrt{1 + (\nabla_\perp \eta)^2}} - g\eta, \tag{3}$$

$$z = \eta, \quad \Phi \longrightarrow 0, \quad z \longrightarrow -\infty,$$

where φ is the electric-field potential.

The first term on the right-hand side of the dynamic boundary condition (nonstationary Bernoulli equation) is responsible for the electrostatic pressure, the second term determines the capillary pressure, and the third term takes into account the effect of the gravitational field. The time evolution of the free surface is determined by the kinematic relation (condition of nonflowing of the liquid through its boundary)

$$\frac{\partial \eta}{\partial t} = \frac{\partial \Phi}{\partial z} - \nabla_\perp \eta \nabla_\perp \Phi, \quad z = \eta.$$

Finally, the electric-field potential φ in the absence of spatial charges satisfies Laplace's equation

$$\nabla^2 \varphi = 0,$$

which should be solved simultaneously with the condition of the equipotentiality of the boundary of the conducting liquid and the condition of the uniformity of the field at an infinite distance from the surface:

$$\varphi = 0, \quad z = \eta,$$

$$\varphi \longrightarrow -Ez, \quad z \longrightarrow \infty.$$

Note that the above-written equations of motion have a Hamiltonian structure and the functions $\eta(x, y, t)$ and $\psi(x, y, t) = \Phi|_{z=\eta}$ are canonically conjugated quantities [15]

$$\frac{\partial \psi}{\partial t} = -\frac{\delta H}{\delta \eta}, \quad \frac{\partial \eta}{\partial t} = \frac{\delta H}{\delta \psi}, \tag{4}$$

where the Hamiltonian H coincides to an accuracy of a constant with the total energy of the system

$$H = \int_{z \le \eta} \frac{(\nabla \Phi)^2}{2} d^3r - \int_{z \ge \eta} \frac{(\nabla \varphi)^2}{8\pi\rho} d^3r + \int \left[\frac{g\eta^2}{2} + \frac{\alpha}{\rho} (\sqrt{1 + (\nabla_\perp \eta)^2} - 1) \right] d^2r.$$

For a further consideration of the problem, it is convenient to represent the Hamiltonian in the form of a surface integral. We introduce a perturbation of the electric-field potential $\tilde{\varphi} = \varphi + Ez$. It can easily be shown that the perturbed potential $\tilde{\varphi}$ satisfies Laplace's equation

$$\nabla^2 \tilde{\varphi} = 0, \tag{5}$$

with conditions

$$\tilde{\varphi} = E\eta, \quad z = \eta, \tag{6}$$

$$\tilde{\varphi} \longrightarrow 0, \quad z \longrightarrow \infty, \tag{7}$$

from which it is seen that the perturbation introduced by the surface $z = \eta$ into the distribution of the electric field decays as $z \longrightarrow \infty$. Taking into account that, in view of the incompressibility of the liquid, a relation $\int \tilde{\varphi}|_{z=\eta} d^2r = 0$ is valid, and neglecting terms whose variation does not contribute to the equations of motion, we obtain, using the first Green's formula,

$$H = \int_s \left[\frac{\psi}{2} \frac{\partial \Phi}{\partial n} + \frac{E\eta}{8\pi\rho} \frac{\partial \tilde{\varphi}}{\partial n} \right] ds + \int \left[\frac{g\eta^2}{2} + \frac{\alpha}{\rho} (\sqrt{1 + (\nabla_\perp \eta)^2} - 1) \right] d^2r,$$

where ds is the surface differential and $\partial/\partial n$ denotes the derivative in the direction of the normal to the surface $z = \eta$.

By eliminating the normal derivatives of the potentials $\tilde{\varphi}$ and Φ, we can reduce the expression for the Hamiltonian to the form

$$H = \int\left[\frac{\psi}{2}(\Phi_z - \nabla\eta\nabla_\perp\Phi)\big|_{z=\eta}\right.$$

$$\left. + \frac{E\eta}{8\pi\rho}(\tilde{\varphi}_z - \nabla\eta\nabla_\perp\tilde{\varphi})\big|_{z=\eta}\right]d^2r \qquad (8)$$

$$+ \int\left[\frac{g\eta^2}{2} + \frac{\alpha}{\rho}(\sqrt{1+(\nabla_\perp\eta)^2} - 1)\right]d^2r,$$

which is more suitable for further transformations.

SMALL-ANGLE APPROXIMATION

Our further problem is to eliminate the spatial variable z from the equations of motion, i.e., to pass from the initial three-dimensional equations to two-dimensional ones. To do this, we should write the integrand in Eq. (8) through the canonical variables η and ψ. Then, there arises a need to solve Eq. (5) with conditions (6) and (7) as well as Eq. (2) with the conditions

$$\Phi = \psi, \quad z = \eta,$$

and condition (3). We use the known solutions to Laplace's equation for the half-spaces $z < 0$ and $z > 0$ for functions that decay at infinity

$$\tilde{\varphi}(x, y, z) = \frac{1}{2\pi}\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}\frac{z\tilde{\varphi}(x', y', 0)}{[(x'-x)^2 + (y'-y)^2 + z^2]^{3/2}}dx'dy', \qquad (9)$$
$$z > 0,$$

$$\Phi(x, y, z) = -\frac{1}{2\pi}\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}\frac{z\Phi(x', y', 0)}{[(x'-x)^2 + (y'-y)^2 + z^2]^{3/2}}dx'dy'. \qquad (10)$$
$$z < 0.$$

Now, we should express the magnitudes of the potentials $\tilde{\varphi}$ and Φ that enter into these relationships at the plane $z = 0$ through their magnitudes at the boundary $z = \eta$, i.e., through the functions $E\eta$ and ψ. Let the characteristic angles of the surface slope be small: $|\nabla_\perp\eta| \ll 1$. In this case, the potentials near the z-0 plane can be expanded into a power series in surface perturbation η:

$$\tilde{\varphi}(x, y, \eta(x, y)) = \sum_{n=0}^{\infty}\frac{\eta^n}{n!}\frac{\partial^n\tilde{\varphi}}{\partial z^n}\bigg|_{z=0},$$
$$\Phi(x, y, \eta(x, y)) = \sum_{n=0}^{\infty}\frac{\eta^n}{n!}\frac{\partial^n\Phi}{\partial z^n}\bigg|_{z=0}. \qquad (11)$$

By differentiating Eqs. (9) and (10) with respect to z, we find that

$$\tilde{\varphi}_z\big|_{z=0} = -\hat{k}\tilde{\varphi}\big|_{z=0}, \quad \Phi_z\big|_{z=0} = \hat{k}\Phi\big|_{z=0},$$

where $\hat{k}$ is the two-dimensional integral operator given by the expression

$$\hat{k}f = -\frac{1}{2\pi}\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}\frac{f(x', y')}{[(x'-x)^2 + (y'-y)^2]^{3/2}}dx'dy'.$$

This relationship can be considered as a consequence of the fact that the Laplacian operator can formally be represented as

$$\nabla^2 = (\partial_z + \hat{k})(\partial_z - \hat{k}),$$

where the left-hand bracket corresponds to solutions that are asymptotically decay as $z \longrightarrow +\infty$ and the right-hand bracket corresponds to solutions that decay as $z \longrightarrow -\infty$.

By eliminating the derivative with respect to z from the expansions (11), we find

$$\tilde{\varphi}(x, y, \eta(x, y)) = \hat{T}_+\tilde{\varphi}(x, y, 0),$$

$$\Phi(x, y, \eta(x, y)) = \hat{T}_-\Phi(x, y, 0),$$

where we introduced nonlinear shear operators

$$\hat{T}_+ = \sum_{n=0}^{\infty}\frac{\eta^n\hat{k}^n}{n!}, \quad \hat{T}_- = \sum_{n=0}^{\infty}\frac{(-\eta)^n\hat{k}^n}{n!}.$$

Let $T_\pm^{-1}$ be the operators that are inverse with respect to the shear operators. Their form can be determined using the method of successive approximations

$$T_\pm^{-1} = 1 \mp \eta\hat{k} - \eta^2\hat{k}^2/2 + \eta\hat{k}\eta\hat{k} + \ldots .$$

Then, we have

$$\tilde{\varphi}(x, y, 0) = E\hat{T}_+^{-1}\eta(x, y),$$

$$\Phi(x, y, 0) = \hat{T}_-^{-1}\psi(x, y).$$

These relationships, along with Eqs. (9) and (10), specify the solutions to Laplace's equations with necessary boundary conditions in the form of infinite series. Using these solutions, we can write the various possible derivatives of the potentials $\tilde{\varphi}$ and Φ that enter into the Hamiltonian through the functions η and ψ. As a result, we find

$$H = \int\frac{\psi}{2}(\hat{T}_+\hat{k}\hat{T}_-^{-1}\psi - \nabla_\perp\eta\hat{T}_+\nabla_\perp\hat{T}_+^{-1}\psi)d^2r$$

$$-\int\frac{E^2\eta}{8\pi\rho}(\hat{T}_-\hat{k}\hat{T}_-^{-1}\eta + \nabla_\perp\eta\hat{T}_-\nabla_\perp\hat{T}_-^{-1}\eta)d^2r$$

$$+\int\left[\frac{g\eta^2}{2}+\frac{\alpha}{\rho}(\sqrt{1+(\nabla_\perp\eta)^2}-1)\right]d^2r.$$

To describe the initial stages of instability development on the surface of a conducting liquid, it is sufficient to restrict ourselves by allowance for a finite number of terms in the expansion of the integrands of the functional H in canonical variables. By omitting terms of higher than the fourth order of smallness for the surface perturbation η and higher than the second order for the potential ψ (this proves to be sufficient at small supercriticalities) and successively integrating by parts, we finally obtain

$$H = H^{(2)} + H^{(3)} + H^{(4)},$$

$$H^{(2)} = \int\left[\frac{\psi\hat{k}\psi}{2}-\frac{E^2\eta\hat{k}\eta}{8\pi\rho}+\frac{g\eta^2}{2}+\frac{\alpha(\nabla\eta)^2}{2\rho}\right]d^2r,$$

$$H^{(3)} = \frac{E^2}{8\pi\rho}\int\eta[(\nabla\eta)^2-(\hat{k}\eta)^2]d^2r, \qquad (12)$$

$$H^{(4)} = -\frac{E^2}{8\pi\rho}\int[\eta\hat{k}\eta\hat{k}\eta\hat{k}\eta+\eta\hat{k}\eta^2\nabla^2\eta]d^2r$$

$$-\int\frac{\alpha(\nabla\eta)^4}{8\rho}d^2r.$$

These expressions, in combination with Eq. (4), represent a two-dimensional reduction of the equations of motion of a conducting liquid in an external electric field that is applicable if the condition of the smallness of the characteristic angles of the surface slope is fulfilled.

AMPLITUDE EQUATIONS

Now, let us consider the nonlinear dynamics of the perturbations of the free surface of a conducting liquid for the case where the magnitude of the external electric field E is close to its threshold value E_c, i.e., $|\varepsilon| \ll 1$. It follows from the dispersion relation (1) that, at small supercriticalities, only surface waves with wave numbers close to k_0 can be excited. The main nonlinear interaction in this case will be the three-wave interaction between the waves whose wave vectors are turned with respect to one another by an angle of $2\pi/3$. This can easily be understood from the conditions

$$\mathbf{k}_1+\mathbf{k}_2+\mathbf{k}_3 = 0, \quad |\mathbf{k}_1| = |\mathbf{k}_2| = |\mathbf{k}_3| = \mathbf{k}_0.$$

Near the threshold, it is natural to pass to envelopes using the following substitutions:

$$\eta(\mathbf{r},t) = \sum_{j=1}^{3}A_j(x_j,y_j,t)e^{i\mathbf{k}_j\mathbf{r}}+\text{c.c.},$$

$$\psi(\mathbf{r},t) = \sum_{j=1}^{3}B_j(x_j,y_j,t)e^{i\mathbf{k}_j\mathbf{r}}+\text{c.c.},$$

where $\mathbf{k}_1 = \{k_0, 0\}$, $\mathbf{k}_2 = \{-k_0/2, \sqrt{3}k_0/2\}$, $\mathbf{k}_3 = \{-k_0/2, -\sqrt{3}k_0/2\}$, and A_j and B_j ($j = 1, 2, 3$) are slow functions of the variables x_j and y_j that form orthogonal coordinate systems with abscissa axes directed along the wave vectors $\mathbf{k}_j$. Such a representation for the functions η and ψ corresponds to a hexagonal structure of the perturbed surface.

Using these relationships for the perturbations η and ψ, we can approximate the integral operator $\hat{k}$ that enters into Hamiltonian (12) by a differential operator. Let us use the following property:

$$\hat{k}e^{i\mathbf{kr}} = |\mathbf{k}|e^{i\mathbf{kr}},$$

which is related to the fact that the Fourier transform of the operator $\hat{k}$ is equal to the modulus of the wave vector. Consider a plane wave of the form

$$e^{i\mathbf{kr}} = e^{i(k_0x+q_xx+q_yy)},$$

whose wave vector is close to k_0 (i.e., $|q_x| \ll k_0$ and $|q_y \ll k_0|$). The quantity $|\mathbf{k}|$ can be expanded in a series in q_x and q_y:

$$|\mathbf{k}| = \sqrt{(k_0+q_x)^2+q_y^2}$$

$$\approx k_0+q_x+(2k_0)^{-1}q_y^2-(2k_0^2)^{-1}q_xq_y^2-(2k_0)^{-3}q_y^4.$$

This means that if we deal with a narrow (in the $\mathbf{k}$ space) wave packet with a carrying wave vector $\mathbf{k} = \{k_0, 0\}$, which can be represented in the form $A(x,y)e^{ik_0x}$, then the operator $\hat{k}$ can be approximated as follows:

$$\hat{k}A(x,y,t)e^{ik_0x} \approx (k_0A-iA_x-(2k_0)^{-1}A_{yy}$$

$$-i(2k_0^2)^{-1}A_{xyy}-(2k_0)^{-3}A_{yyyy})e^{ik_0x}$$

(similar relations are obtained for the amplitudes A_j in the coordinates of x_j and y_j).

Then, inserting the expressions for η and ψ into Hamiltonian (12) and performing necessary averaging, we find (to an accuracy of terms of a higher order of smallness)[1]

$$H = \sum_{j=1}^{3}\int\left(k_0|B_j|^2-2g\varepsilon|A_j|^2\right.$$

$$\left.+\frac{g}{k_0^2}\left|\frac{\partial A_j}{\partial x_j}-\frac{i}{2k_0}\frac{\partial^2A_j}{\partial y_j^2}\right|^2\right)d^2r$$

$$-3gk_0\int(A_1A_2A_3+A_1^*A_2^*A_3^*)d^2r.$$

[1] Note that it is more suitable to perform these calculations in the k representation.

The dynamic equations that describe the time evolution of amplitudes A_j and B_j are found from the relations [16]

$$A_{j_t} = \frac{\delta H}{\delta B_j^*}, \quad B_{j_t} = -\frac{\delta H}{\delta A_j^*},$$

where $j = 1, 2, 3$.

By varying the expression for the averaged Hamiltonian, we obtain the following equations for the amplitudes:

$$A_{j_t} = k_0 B_j,$$

$$B_{j_t} = 2gk_0\varepsilon A_j + \frac{g}{k_0}\left(\frac{\partial}{\partial x} - \frac{i}{2k_0}\frac{\partial^2}{\partial y^2}\right)^2 A_j + 3gk_0^2\frac{A_1^* A_2^* A_3^*}{A_j^*}.$$

By eliminating the amplitudes B_j from these equations and passing to dimensionless quantities using the substitutions

$$\begin{aligned} &\mathbf{r} \longrightarrow \mathbf{r}/(\sqrt{2}k_0), \quad A_j \longrightarrow A_j/k_0, \\ &t \longrightarrow t/\sqrt{2gk_0}, \quad H \longrightarrow 2Hg/k_0^2, \end{aligned} \tag{13}$$

we obtain the following set of equations:

$$A_{1_{tt}} = \varepsilon A_1 + \hat{L}_1^2 A_1 + 3A_2^* A_3^*/2, \tag{14}$$

$$A_{2_{tt}} = \varepsilon A_2 + \hat{L}_2^2 A_2 + 3A_3^* A_1^*/2, \tag{15}$$

$$A_{3_{tt}} = \varepsilon A_3 + \hat{L}_3^2 A_3 + 3A_1^* A_2^*/2, \tag{16}$$

where we introduced operators

$$\hat{L}_j = \frac{\partial}{\partial x_j} - \frac{i}{\sqrt{2}}\frac{\partial^2}{\partial y_j^2}, \quad j = 1, 2, 3.$$

The Hamiltonian corresponding to these amplitude equations is written as follows:

$$H = \int\left[\sum_{j=1}^{3}(|A_j|_t^2 + |\hat{L}_j A_j|^2 - \varepsilon|A_j|^2) - \frac{3}{2}(A_1 A_2 A_3 + A_1^* A_2^* A_3^*)\right] d^2 r. \tag{17}$$

Thus, we obtained equations that describe the initial stages of the development of instability of the surface of a conducting liquid in a near-critical field, when the small-angle approximation is valid and the main nonlinear interaction is the interaction of three standing waves that form a hexagonal lattice. Note that analogous equations describe the instability of the charged surface of liquid helium [8, 17].

CRITERIA FOR EXPLOSIVE INSTABILITY

As is known, hexagonal structures on a charged surface of various liquids are characterized by a hard regime of excitation [7, 8]. For the set of equations (14)–(16), this means the possibility of an unbounded growth of the amplitudes A_j in a finite time. Indeed, in the simplest case, when the amplitudes are independent of the spatial variables, are real, and are equal to one another, i.e., $A_1 = A_2 = A_3 = A(t)$, the time evolution of the quantity A is described by an ordinary differential equation with a quadratic nonlinearity

$$A_{tt} = \varepsilon A + 3A^2/2.$$

Because of its influence, the amplitude grows with the asymptotics $A \longrightarrow 4(t - t_c)^{-2}$ under corresponding initial conditions, i.e., the magnitude of A becomes infinite at the moment t_c. However, what seems to be obvious for spatially uniform (coordinate-independent) solutions requires to be proved in the case of arbitrary amplitudes A_1, A_2, and A_3. In particular, of a significant interest is the situation where the initial perturbation of the surface is localized in a certain region.

Let us show, using the method of differential inequalities, that the nonlinear interaction of amplitudes A_j in terms of model (14)–(16) results in an explosive growth of perturbations of the surface of a conducting liquid and find the sufficient conditions for hard excitation of the instability. To this end, we introduce the norms

$$X_j(t) = \int |A_j|^2 d^2 r; \quad j = 1, 2, 3$$

and consider the time evolution of the following nonnegative quantity:

$$X = \sum_{j=1}^{3} X_j.$$

By analogy with [12], we doubly differentiate X with respect to t

$$X_{tt} = \sum_{j=1}^{3}\int(2|A_j|_t^2 + A_{j_{tt}} A_j^* + A_{j_{tt}}^* A_j) d^2 r$$

$$= \int\left[2\sum_{j=1}^{3}(|A_j|_t^2 + |\hat{L}_j A_j|^2 + \varepsilon|A_j^2|) + \frac{9}{2}(A_1 A_2 A_3 + A_1^* A_2^* A_3^*)\right] d^2 r,$$

after substituting the corresponding right-hand sides of the amplitude equations (14)–(16) for the multiple derivatives $A_{j_{tt}}$ and $A_{j_{tt}}^*$. Then, eliminating the sign-ambiguous cubic nonlinearity from the integrand using

the expression for Hamiltonian (17), we obtain the following relation:

$$X_{tt} + 3H = -\varepsilon X + 5\sum_{j=1}^{3}[|A_j|_t^2 + |\hat{L}_j A_j|^2]d^2r. \quad (18)$$

Now, our problem is to approximate the right-hand side of Eq. (18) using the quantity X and thereby obtain an ordinary differential inequality. From the known Cauchy–Bunyakowsky integral inequality for the functions $|A_j|$ and $|A_j|_t$

$$[\int |A_j|^2 d^2r][\int |A_j|_t^2 d^2r] \geq [\int |A_j||A_j|_t d^2r]^2,$$

it follows that $\int |A_j|_t^2 d^2r \geq X_{j_t}^2/(4X_j)$. Taking also into account the obvious relations $\int |\hat{L}_j A_j|^2 d^2r \geq 0$ for $j = 1$, 2, 3, we obtain from Eq. (18)

$$X_{tt} + 3H \geq -\varepsilon X + \frac{5}{4}\sum_{j=1}^{3}\frac{X_{j_t}^2}{X_j}. \quad (19)$$

Then, note that, as a consequence of the algebraic Cauchy inequality, the following relation is valid:

$$\left[\sum_{j=1}^{3} X_j\right]\left[\sum_{j=1}^{3} X_{j_t}^2/X_j\right] \geq \left[\sum_{j=1}^{3} X_{j_t}\right]^2,$$

and, correspondingly, we have

$$\sum_{j=1}^{3} X_{j_t}^2/X_j \geq X_t^2/X.$$

Substituting the latter inequality into (19), we obtain an ordinary differential inequality

$$X_{tt} + 3X \geq -\varepsilon X + \frac{5}{4}\frac{X_t^2}{X}, \quad (20)$$

which will be the object of our consideration below. Note that analogous inequalities arise when deriving sufficient collapse criteria for various nonlinear partial differential equations [10–14].

The introduction of a new function $Y = X^{-1/4}$ permits us to rewrite inequality (20) in the form of Newton's second law

$$Y_{tt} \leq -\frac{\partial P(Y)}{\partial Y}, \quad P(Y) = -\frac{1}{8}(\varepsilon Y^2 + HY^6), \quad (21)$$

where Y plays the role of the coordinate of a "particle" and P is its potential energy.

Let the velocity of the particle Y_t be negative (in this case $X_t > 0$). Then, multiplying (21) by Y_t, we obtain

$$U_t(t) \geq 0, \quad U(t) = Y_t^2/2 + P(Y),$$

i.e., the particle gains an energy U upon motion. It is understandable that the sufficient criterion for Y to become zero and, correspondingly, for X to become infinity is the condition that the particle encounters no potential barrier even if $U_t = 0$, which corresponds to the equality sign in (20). The explosive growth of amplitudes takes place under the following conditions:

(a) at $\varepsilon < 0$ and $H > 0$ if $Y(t_0) < |\varepsilon|^{\frac{1}{4}}/(3H)^{\frac{1}{4}}$ and $12U(t_0) \leq |\varepsilon|^{\frac{3}{2}}/(3H)^{\frac{1}{2}}$;

(b) at $\varepsilon < 0$ and $H > 0$ if $12U(t_0) > |\varepsilon|^{\frac{3}{2}}/(3H)^{\frac{1}{2}}$;

(c) at $\varepsilon < 0$ and $H \leq 0$; and

(d) at $\varepsilon \geq 0$ if $U(t_0) > 0$.

In (a)–(d), $t = t_0$ corresponds to the starting time moment. In this case, the moment t_c at which the perturbation amplitudes become infinite is estimated as follows:

$$t_c \leq t_0 + \int_0^{Y(t_0)} \frac{dY}{\sqrt{2U(t_0) - 2P(Y)}}.$$

Note that the condition $Y_t(t_0) < 0$ in cases (a) and (c) is by no means necessary: after the reflection from a potential wall, the particle reaches the point $Y = 0$. The above conditions (a)–(d) can be considered as sufficient criteria of the instability of the surface of a conducting liquid with respect to perturbations of a finite amplitude, which distinguishes it from the simplest criterion of linear instability $E > E_c$, which was derived based on the assumption that the perturbations are infinitely small. Note also that conditions (a)–(c) refer to the case of subcritical external fields ($E < E_c$), when the flat surface of the conducting liquid is stable in the linear approximation, i.e., we deal with a hard excitation of an electrohydrodynamic instability.

Thus, if conditions (a)–(d) are fulfilled, Eqs. (14)–(16) describe an infinite growth of amplitudes A_j. In this case, the applicability of model (14)–(16) to the description of the development of an electrohydrodynamic instability is restricted by the condition of the smallness of the amplitudes: in the order of magnitude, the absolute values of the amplitudes $|A_j|$ should not exceed the magnitude of the parameter of supercriticality ε. Otherwise, the model cannot be restricted to the consideration of only three-wave processes. As to the higher-order wave processes, there arises a question of whether they will lead to a stabilization of the instability or will favor an explosive growth of perturbations (the experimental data of [18] and the results of numerical calculations [3, 4] evidence in favor of the latter situation). The complexity of the estimation of their influence is related to the fact that the contribution of higher-order nonlinearities becomes comparable with the contribution of quadratic nonlinearities of the model (14)–

(16) only if the amplitude of surface perturbations is close to the characteristic length of the wave. But in this case conditions of the applicability of our approach to the description of the near-critical dynamics of the charged surface of liquid metal based on the construction of amplitude equations become violated.[2] Nevertheless, it is possible to reveal the influence of the higher-order nonlinearities by considering one-dimensional and square lattices of surface distributions, for which the three-wave interactions degenerate and the dominating interactions are the four-wave ones.

FOUR-WAVE INTERACTIONS

Let us consider perturbations of the boundary of a conducting liquid with such symmetries for which the effect of three-wave processes is negligible. Thereby, we in pure form separate the four-wave interactions that determine the character of the electrohydrodynamic instability at its advanced stages.

First of all, we consider the near-critical behavior of the charged surface of a liquid metal in the assumption of a quasi-one-dimensional character of the arising wave. Let the wave vector be parallel to the abscissa axis. We pass to envelopes by using substitutions

$$\eta(x, y, t) = A(x, y, t)e^{ik_0x} + A_0(x, y, t)e^{2ik_0x} + \text{c.c.},$$

$$\psi(x, y, t) = B(x, y, t)e^{ik_0x} + B_0(x, y, t)e^{2ik_0x} + \text{c.c.},$$

in which the $k_0 \longleftrightarrow 2k_0$ interaction is taken into account. Here, A, B, A_0, and B_0 are slowly varying functions of the spatial variables x and y. Substituting these relations into Hamiltonian (12), we find (to an accuracy of terms of the fourth order of smallness)

$$H = H^{(2)} + H^{(3)} + H^{(4)},$$

$$H^{(2)} = \int\left[k_0|B|^2 - 2g\varepsilon|A|^2 + \frac{g}{k_0^2}\left|\frac{\partial A}{\partial x} - \frac{i}{2k_0}\frac{\partial^2 A}{\partial y^2}\right|^2 + g|A_0|^2 + 2k_0|B_0|^2\right]d^2r,$$

$$H^{(3)} = -2gk_0\int[A^2A_0^* + A^{*2}A_0]d^2r,$$

$$H^{(4)} = \frac{5}{4}gk_0^2\int|A|^4d^2r.$$

The amplitude equations for the perturbations of the free surface are written in the Hamiltonian form as

$$\frac{\partial A}{\partial t} = \frac{\delta H}{\delta B^*}, \quad \frac{\partial B}{\partial t} = -\frac{\delta H}{\delta A^*},$$

$$\frac{\partial A_0}{\partial t} = \frac{\delta H}{\delta B_0^*}, \quad \frac{\partial B_0}{\partial t} = -\frac{\delta H}{\delta A_0^*}.$$

By varying the functional H and then eliminating the quantities B and B_0, we obtain equations of the form

$$A_{tt} = 2gk_0\varepsilon A + \frac{g}{k_0}\left(\frac{\partial}{\partial x} - \frac{i}{2k_0}\frac{\partial^2}{\partial y^2}\right)^2 A + 4gk_0^2A^*A_0 - \frac{5}{2}gk_0^3A^2A^*, \tag{22}$$

$$A_{0tt} = -2gk_0A_0 + 4gk_0^2A^2. \tag{23}$$

Since the characteristic times of changes in the amplitudes at small supercriticalities are small ($\omega^2 \sim \varepsilon$), we neglect the derivatives with respect to time in Eq. (23). Then, the quantity A_0 can be expressed through the amplitude A that plays the role of an order parameter

$$A_0 \approx 2k_0A^2.$$

Using this relation, we eliminate A_0 from Eq. (22) and pass to dimensionless quantities using scalings (13) to obtain for the complex amplitude A

$$A_{tt} = \varepsilon A + \hat{L}_1^2A + sA|A|^2, \quad s = 11/4, \tag{24}$$

to which the following expression for the Hamiltonian corresponds:

$$H = \int[|A_t|^2 + |\hat{L}_1A|^2 - \varepsilon|A|^2 - s|A|^4/2]d^2r. \tag{25}$$

Note that, when neglecting the dependence of the amplitude A on y, Eq. (24) becomes a nonlinear Klein–Gordon equation, i.e., corresponds to the so-called $|\phi|^4$ model. In this form, it can be obtained from the equation for one-dimensional perturbations of the charged surface of liquid helium [8] in the limit of the complete screening of the field under the surface. Note also that if we neglect transverse modulations, then Eq. (24) coincides with that obtained in the Kelvin–Helmholtz theory of instability for the case of a small ratio of the densities of the top and bottom liquids [12]. This is due to the identity of the mathematical description of the planar potential flow of an incompressible liquid and a two-dimensional distribution of an electric field in the absence of spatial electric charges. The allowance for higher-order terms in the expansions in surface perturbations violates this analogy.

Since the term $|A_t|^2$ that is responsible for the kinetic energy and the term $|A|^4$ that is responsible for the four-wave processes enter into the integrand of Hamiltonian (25) with the opposite signs, Eq. (24) admits infinite solutions. This means that the cubic nonlinearity in

[2] The allowance for four-wave interactions along with the three-wave ones proves to be quite possible for an insulating liquid (with a dielectric constant close to unity) placed in a strong electric field, as well as for an insulating liquid with an insignificant free surface charge. This is related to the appearance of small coefficients before the quadratic nonlinearities in corresponding amplitude equations [7, 17].

Eq. (24) does not stabilize the linear instability but, on the contrary, enhances it, leading, under certain conditions, to an explosive growth of the amplitude A of perturbations of the conducting-liquid boundary.

Another possible case when the three-wave processes are degenerate is the interaction of two standing waves whose wave vectors $\mathbf{k}_1$ and $\mathbf{k}_2$ are turned with respect to one another by an angle $\pi/2$ (the vector's coordinates are $\mathbf{k}_1 = \{k_0, 0\}$ and $\mathbf{k}_2 = \{0, k_0\}$). Let us represent the perturbation of the surface η in the form

$$\eta(\mathbf{r}, t) = \sum_{j=1}^{2} [A_j e^{i\mathbf{k}_j\mathbf{r}} + 2k_0 A_j^2 e^{2i\mathbf{k}_j\mathbf{r}}] + (8\sqrt{2} + 10)k_0 \times [A_1 A_2 e^{i(\mathbf{k}_1 + \mathbf{k}_2)\mathbf{r}} + A_1 A_2^* e^{i(\mathbf{k}_1 - \mathbf{k}_2)\mathbf{r}}] + \text{c.c.},$$

and the perturbation of the velocity potential at the liquid boundary ψ in the form

$$\psi(\mathbf{r}, t) = \sum_{j=1}^{2} [k_0^{-1} A_{j_t} e^{i\mathbf{k}_j\mathbf{r}} + 2A_j A_{j_t} e^{2i\mathbf{k}_j\mathbf{r}}] + (8 + 5\sqrt{2}) \times [(A_1 A_2)_t e^{i(\mathbf{k}_1 + \mathbf{k}_2)\mathbf{r}} + (A_1 A_2^*)_t e^{i(\mathbf{k}_1 - \mathbf{k}_2)\mathbf{r}}] + \text{c.c.},$$

where we took into account the nonlinear interactions of the fundamental spatial harmonic k_0 with combination harmonics $2k_0$ and $\sqrt{2}k_0$. This representation for the functions η and ψ corresponds to the symmetry of a square lattice.

Proceeding by analogy to the above-considered quasi-one-dimensional case, we obtain, after passing to dimensionless quantities, the following dynamic equations:

$$A_{1_{tt}} = \varepsilon A_1 + \hat{L}_1^2 A_1 + sA_1|A_1|^2 + \sigma A_1|A_2|^2,$$

$$A_{2_{tt}} = \varepsilon A_2 + \hat{L}_2^2 A_2 + sA_2|A_2|^2 + \sigma A_2|A_1|^2,$$

where $\sigma = 32\sqrt{2} + 65/2$, and the following designations are introduced:

$$\hat{L}_1 = \frac{\partial}{\partial x} - \frac{i}{\sqrt{2}}\frac{\partial^2}{\partial y^2} \quad \text{and} \quad \hat{L}_2 = \frac{\partial}{\partial y} - \frac{i}{\sqrt{2}}\frac{\partial^2}{\partial x^2}.$$

The integral of motion for these equations, corresponding to the conservation of the total energy of a system, is given by the expression

$$H = \sum_{j=1}^{2} \int (|A_j|_t^2 + |\hat{L}_j A_j|^2 - \varepsilon|A_j|^2 - s|A_j|^4/2) d^2r - \sigma \int |A_1|^2 |A_2|^2 d^2r.$$

The first term on the right-hand side of this functional coincides in its structure with Hamiltonian (25) for the quasi-one-dimensional wave. The last term is responsible for the nonlinear interaction of a pair of the waves studied. Note that the coefficient σ before this term exceeds the coefficient s by more than an order of magnitude. This means that the contribution of the interaction $k_1 \longleftrightarrow k_2$ is determining and, consequently, the square structure of the surface perturbations is much more favorable than the one-dimensional one.

In any case, for both the square and one-dimensional lattice (the latter can be considered as a partial case of the square lattice, corresponding to the condition $A_2 = 0$), the four-wave interactions will favor the development of an instability rather than suppress it. Conditions for an explosive growth of the amplitudes A_1 and A_2 can be obtained by considering the evolution of the norm

$$X = \int (|A_1|^2 + |A_2|^2) d^2r.$$

Acting by analogy with Section 4, we obtain the majorizing inequality

$$X_{tt} + 4H \geq -2\varepsilon X + \frac{3}{2}\frac{X_t^2}{X},$$

which coincides with that considered in [12]. The introduction of the variable $Y = X^{-1/2}$ reduces the problem to the analysis of the motion of a particle with a coordinate Y in a potential well $P(Y)$

$$Y_{tt} \leq -\frac{\partial P(Y)}{\partial Y}, \quad P(Y) = -\frac{1}{2}(\varepsilon Y^2 + HY^4).$$

Analyzing this inequality for the case where the velocity of the particle at the initial time moment $t = t_0$ is directed toward the origin (i.e., $Y_t(t_0) < 0$), it can easily be revealed that the quantity Y vanishes, first, at $\varepsilon > 0$ if $U(t_0) > 0$, second, at $\varepsilon < 0$ and $H < 0$, and third, at $\varepsilon < 0$ and $H > 0$ if $U(t_0) > \varepsilon^2/(8H)$ or $Y^2(t_0) < |\varepsilon|/(2H)$. Here, $U(Y)$, just as in Section 4, denotes the total mechanical energy of the particle. Under the above conditions, the norm X becomes infinite in a finite time, which just corresponds to an explosive growth of the amplitudes in the result of four-wave interactions.

All this suggests that the higher-order nonlinearities will not suppress the explosive growth of amplitudes in model (14)–(16). But in this case the above integral criteria (a)–(d) may be considered as sufficient criteria of the explosive growth of perturbations of the surface of a liquid metal in an external electric field.

CONCLUSION

The main result of this work is the construction of sufficient integral criteria of instability for the free surface of a perfectly conducting liquid in a near-critical external electric field. These criteria represent a generalization of the known condition of linear instability ($E > E_c$) to the case where the amplitudes of perturbations of the field of velocities and of the shape of the surface are finite. The criteria found are dynamic in the

sense that they take into account the effect of the velocity distribution in the medium at the initial time moment; the role of the stored kinetic energy can be decisive in the case of the hard mechanism of instability.

An analysis of three-wave and four-wave nonlinear interactions (this corresponds to the allowance for quadratic and cubic nonlinearities in the amplitude equations) showed that the development of the electrohydrodynamic instability has an explosive character, i.e., leads to the appearance of singularities in the solutions in a finite time. This conclusion qualitatively agrees with the results of numerical simulation of the development of the instability of the boundary of liquid metal: it was shown in [4] that the curvature of the surface increases according to a power law characteristic of the explosive instability and causes the formation of singularities of a cusplike type. Note in conclusion that the criteria of hard instability analogous to the criteria (a)–(d) can also be obtained for insulating liquids with induced surface charges [9, 19], for insulating liquids with free surface charges (liquid helium and liquid hydrogen in an electric field refer to this category) [9], and for ferromagnetic liquids in a vertical magnetic field.

ACKNOWLEDGMENTS

This work was supported in part by the Russian Foundation for Basic Research (project no. 00-02-17428) and by the INTAS (project no. 99-1068). We are grateful to N.B. Volkov and A.M. Iskol'dskiĭ for interest in our work; N.M. Zubarev is also grateful to E.A. Kuznetsov for stimulating discussions.

REFERENCES

1. L. Tonks, Phys. Rev. **48**, 562 (1935).
2. Ya. I. Frenkel', Zh. Éksp. Teor. Fiz. **6** (4), 350 (1936).
3. A. L. Pregenzer and B. M. Marder, J. Appl. Phys. **60** (11), 3821 (1986).
4. V. G. Suvorov and E. A. Litvinov, J. Phys. D **33**, 1245 (2000).
5. N. M. Zubarev, Zh. Éksp. Teor. Fiz. **114**, 2043 (1998) [JETP **87**, 1110 (1998)].
6. L. D. Landau and E. M. Lifshitz, *Course of Theoretical Physics*, Vol. 8: *Electrodynamics of Continuous Media* (Nauka, Moscow, 1982; Pergamon, New York, 1984).
7. E. A. Kuznetsov and M. D. Spektor, Zh. Éksp. Teor. Fiz. **71** (1), 262 (1976) [Sov. Phys. JETP **44**, 136 (1976)].
8. L. P. Gor'kov and D. M. Chernikova, Dokl. Akad. Nauk SSSR **228**, 829 (1976) [Sov. Phys. Dokl. **21**, 328 (1976)].
9. N. M. Zubarev, Phys. Lett. A **272**, 119 (2000).
10. E. A. Kuznetsov, J. J. Rasmussen, K. Rypdal, and S. K. Turitsyn, Physica D (Amsterdam) **87**, 273 (1995).
11. P. M. Lushnikov, Pis'ma Zh. Éksp. Teor. Fiz. **62**, 447 (1995) [JETP Lett. **62**, 461 (1995)].
12. E. A. Kuznetsov and P. M. Lushnikov, Zh. Éksp. Teor. Fiz. **108**, 614 (1995) [JETP **81**, 332 (1995)].
13. E. M. Maslov and A. G. Shagalov, Phys. Lett. A **239**, 46 (1998).
14. S. K. Turitsyn, Phys. Rev. E **47**, R796 (1993).
15. V. E. Zakharov, Prikl. Mekh. Tekh. Fiz., No. 2, 86 (1968).
16. V. E. Zakharov and E. A. Kuznetsov, Usp. Fiz. Nauk **167**, 1137 (1997) [Phys. Usp. **40**, 1087 (1997)].
17. D. M. Chernikova, Fiz. Nizk. Temp. **6**, 1513 (1980) [Sov. J. Low Temp. Phys. **6**, 737 (1980)].
18. M. D. Gabovich and V. Ya. Poritskiĭ, Pis'ma Zh. Éksp. Teor. Fiz. **33**, 320 (1981) [JETP Lett. **33**, 304 (1981)].
19. N. M. Zubarev and O. V. Zubareva, Pis'ma Zh. Tekh. Fiz. **26** (9), 65 (2000) [Tech. Phys. Lett. **26**, 389 (2000)].

Translated by S. Gorin

Technical Physics, Vol. 46, No. 7, 2001, pp. 815–824. Translated from Zhurnal Tekhnicheskoĭ Fiziki, Vol. 71, No. 7, 2001, pp. 30–38.

Gas Dynamics and Thermal-Ionization Instability of the Cathode Region of a Glow Discharge. Part I

S. A. Smirnov and G. A Baranov

Efremov Research Institute of Electrophysical Apparatus, St.-Petersburg, 189631 Russia

Received July 10, 2000

Abstract—A model of the cathode sheath of a glow discharge is developed. The model includes the equations for calculating the non-steady-state nonequilibrium physicochemical gas dynamics, cathode temperature, and electric field. The model applies to describing the flow of a viscous, heat-conducting, moderately rarefied gas at Knudsen numbers of about Kn ~ 0.03. The electric field and gas density distributions are determined consistently by renormalizing the values obtained by the Engel–Steenbeck theory. A formula for calculating the time during which a homogeneous volume discharge phase exists is proposed. The formula is based on the relation between the rates of electron production via associative ionization ($A + B \longrightarrow AB^+ + e$) and impact ionization ($A2 + e \longrightarrow A_2^+ + e + e$). Calculations are carried out for nitrogen and air. It is shown that, at high current densities, due to the dissociation and strong heating of the gas, the rate of thermal ionization becomes as high as that of electrical ionization. The calculated ionization time is in reasonable agreement with the measured duration of a uniform anomalous cathode sheath. © *2001 MAIK "Nauka/Interperiodica"*.

INTRODUCTION AND THE FORMULATION OF THE PROBLEM

Gas discharges are widely used in lasers, ozonizers, and plasmochemical reactors. At sufficiently high gas pressures and current densities, a glow discharge transforms into an arc. The discharge contraction terminates the generation of coherent radiation in lasers and deteriorates the efficiency of technological processes in other devices.

The discharge state and the gas parameters are closely related to each other. In a pulsed laser, the discharge gives rise to compression waves [1] that, at high energy deposition, transform into shock waves. Reflecting from the walls, these waves can return into the discharge chamber by the beginning of the next current pulse, thus breaking the gas homogeneity and inducing the transformation of the glow discharge into an arc. The variations in the other gas parameters due to the slow relaxation of large-scale heat perturbations in the so-called "heat mirror" can also affect the discharge stability. Both of these phenomena have been studied rather thoroughly. To decrease their influence, the gas is usually pumped through the discharge chamber and special facilities for suppressing acoustic perturbations are introduced into the gas-flow contour.

The influence of the gas-flow microstructure in the cathode sheath on the parameters and stability of the sheath is a universal but still poorly studied effect. Its importance stems from the key role that the cathode sheath plays in the discharge dynamics. During each current pulse, a fraction of the gas is expelled from the cathode sheath; hence, the sheath parameters change during the pulse, whereas the gas density in the positive column varies only slightly. Thus, for nitrogen at a pressure of $p = 500$ torr and a copper cathode, the cathode sheath thickness is $d_N \approx 1 \times 10^{-5}$ m and the time during which a sound wave passes through the sheath is ~20 ns. This time is much shorter than the time during which a sound wave passes through a 0.03-m-high discharge chamber (about 140 μs).

The evolution of the strongly anomalous cathode sheath of a glow discharge in air and commercial nitrogen was studied, e.g., in [2, 3]. The sequence of processes is as follows. The breakdown of the interelectrode gap is followed by an anomalous discharge with a homogeneous cathode dark space (CDS) and a uniform diffuse glow above it. Then, the discharge breaks into separate channels and a prearc cathode spot arises. In front of this spot, the cathode sheath thickness decreases. This manifests itself as the shift of the diffuse glow toward the cathode and a local increase in the emission intensity. Then, the breakdown of the cathode sheath occurs against the cathode spot. The transition from a homogeneous anomalous sheath to a prearc cathode spot with a decreased cathode potential drop was also observed in experiments in [4]. The reverse transitions and self-oscillations can occur at relatively low discharge currents. At higher currents, the transition to an arc becomes irreversible. In [2, 3], the instability of the cathode sheath was attributed to the explosive autoemission from the cathode micropoints at high electric field strengths; however, such an approach fails

1063-7842/01/4607-0815$21.00

to predict the time during which the CDS remains uniform. In [4], the experimental results were qualitatively explained based on evaluating the volume processes involving metastable particles; the increase in their density increases the impact ionization rate. However, an adequate study of these phenomena requires an accurate quantitative analysis of the relation between the gas parameters and the electric characteristics of the cathode sheath. This is not a simple matter because, in the literature, there are no experimental data on the gas flow structure in the cathode sheath, although the cathode shock waves are well known to propagate beyond the sheath [5, 6]. The complexity of such measurements stems from the fact that the spatial resolution should be fairly high, which is dictated by the small thickness of the cathode sheath at high gas pressures.

Gas dynamics in the cathode region was studied in the following papers. In [7], the energy deposited in the cathode sheath in a XeCl laser at $p = 0.5$–8 bar and current densities of $j = 50$ and 300 A/cm^2 was estimated using a one-dimensional system of Euler equations and the measured jump in the gas density behind the cathode shock wave. In [8], the discharge formation was numerically simulated using the Navier–Stokes equations and equations for the electric parameters of the cathode sheath in the diffusion–drift approximation. At a current density of 100A/cm^2, the gas was heated up to 3000 K during 30 ns. An attempt at analyzing the thermal-ionization instability of the cathode sheath was made in [9], in which the electric field was calculated for a nonuniform gas density assuming the motion of the heat-conducting gas to be one-dimensional. The limiting value of the current was found to agree with the high temperature (e.g., 4300 K) of the gas in the cathode sheath; at such temperatures, the equilibrium electron density is capable of sustaining the discharge current. In [10], cathode sheaths expanding across the electric field were examined using a set of equations for a viscous, heat-conducting, moderately rarefied gas.

The study of gas dynamics in the cathode sheath is necessary for the following reasons.

(i) Although the cathode sheath is characterized by strong gas heating and the high probability of the discharge losing its stability [11], in most of studies, the influence of the temperature on the sheath electric parameters was not analyzed quantitatively. At pressures higher than 5–10 torr, such an analysis becomes even more necessary. Thus, taking into account the nonuniform distribution of the gas temperature in the cathode sheath made it possible to reveal the influence of some gas-dynamic parameters on the limiting current of a transverse glow discharge [12].

(ii) The influence of the emissive properties of the cathode surface on the stability of the cathode sheath cannot be described adequately without allowance for the correspondent changes in the gas properties. Indeed, the change in the secondary emission coefficient changes the electric field, the normal current density, and, consequently, the energy deposition and the gas dynamic structure; the latter affects all the elementary processes in the sheath.

(iii) The gas dynamics of the cathode sheath is more complicated than that in the models presented in [7–10]; hence, a further upgrade of the models is required.

The purpose of this paper is to develop a model of the cathode sheath in order to investigate the mechanism for the onset of the thermal-ionization instability. The model incorporates the system of equations for calculating the non-steady-state nonequilibrium gas dynamics, the electric parameters, and the cathode heat balance.

A criterion for determining the time during which the cathode sheath remains uniform (from the instant of the sheath formation until the onset of a prearc spot) is proposed.

Calculations based on the experimental data from [2, 3] were carried out to verify the model. Since the dimensions of the current area at a copper cathode substantially exceed the CDS thickness, we used the one-dimensional approximation. The current densities and initial pressures were taken as follows: $j = 100$ and 1000 A/cm$_2$ (N_2 at $p_0 = 50$ and 100 torr, respectively) [2] and $j \approx 120$ A/cm^2 (air at $p_0 = 50$ torr) [3].

CATHODE SHEATH MODEL

Scheme of the Processes and Criterion for the Onset of Instability

Fast non-steady-state processes in the strong electric field of the CDS are responsible for the nonequilibrium nature of many of the incorporated processes. In the CDS, the ion current substantially exceeds the current produced by electrons, whose energy is too high for the molecular vibrational states to be efficiently excited. Recall that the gas in the CDS does not glow because the electron energy is higher than the optimum energy for the excitation of molecular electronic (and, all the more so, vibrational) states. The energy deposited in the cathode sheath via the ion current is first transformed into molecular translational energy, and then certain fraction of it relaxes into vibrational energy. The vibrational temperature is lower than the translational; as a result, the rate of molecular dissociation decreases [13]. Under our conditions, this effect is so strong that, for the time during which an anomalous cathode sheath exists, no dissociation would be observed in [2, 3] unless it was accelerated due to other nonequilibrium processes involving fast molecules produced via charge transfer from ions to molecules.

In the strong electric field of the CDS, the average ion velocity is much higher than the thermal velocity of molecules and the efficiency of charge exchange is fairly high [14]. In each charge exchange event, an ion becomes a fast molecule, whereas a molecule becomes a slow ion, which is then accelerated by the electric

field. Collisions of fast molecules (N_2^*) with the bulk slow molecules result in nonequilibrium dissociation ($N_2^* + M \longrightarrow 2N + M + M$, where $M = N_2$, N), which proceeds against the background of non-steady-state gas dynamics and the exponential growth of the rate of associative ionization ($N + N \longrightarrow N_2^+ + e$) due to an increase in the gas temperature. If the gas temperature is high enough, this rate may be as high as the rate of impact ionization ($N_2 + e \longrightarrow N_2^+ + 2e$), which usually enables an increase in the electron current in the cathode sheath and the existence of a discharge.

This point is crucial for the discharge dynamics. Indeed, if the increase in the thermal ionization rate is not balanced by the decrease in the electric field, then the current will increase. This is positive feedback between the gas temperature and the current density, which results in the onset of the cathode sheath instability. Consistent calculations show that, at a constant current density, the above positive feedback is nevertheless balanced by a decrease in both the electric field and cathode potential [15]. Therefore, when the rate of associative ionization is sufficiently high, the current–voltage characteristic of the cathode sheath becomes descending, which is known to result in the instability of the cathode sheath, the growth of the current density, and the decrease in the diameter of the current spot at the cathode [16]. This indicates the correlation between the increase in the rate of associative ionization up to a certain critical value and the onset of a prearc cathode spot.

The inevitable decrease in the electric field in the presence of an extra electron source also follows from condition (2) for the existence of a self-sustained discharge. This condition is derived from the equation for the electron current density

$$\frac{dj_e}{dy} = \alpha(E/N)v_e n_e e + R(T, N_a)e - \beta n_e n_i e, \qquad (1)$$

and allows us to estimate the time during which the cathode sheath remains uniform.

Assuming that $j_e = -\gamma j_+$ at the cathode and $j = j_e$ at the outer boundary of the cathode sheath, from Eq. (1) we obtain

$$\int_0^d (\alpha(E/N)v_e n_e + R(T, N_a) - \beta n_e n_i)dy = \frac{j}{e(1+\gamma)}, \qquad (2)$$

where y is the longitudinal coordinate, α is the Townsend ionization coefficient, β is the recombination coefficient, γ is the secondary emission coefficient, $R(T, N_a)$ is the rate of associative ionization, T is the gas temperature, N_a is the atom density, E is the electric field, N is the molecular density, j_e and j_i are the electron and ion components of the current density, j is the current density, v_e is the electron drift velocity, n_e and n_i are the electron and ion densities, e is the electron charge, and d is the cathode sheath thickness.

As a rule, recombination in the cathode sheath can be neglected because of the low electron density; in addition, $\gamma \ll 1$. Hence, at a constant current density, an increase in $R(T, N_a)$ leads to a decrease in E. Furthermore, taking into account formula (2), the condition for the associative ionization rate to achieve the critical value required for the formation of the descending current–voltage characteristic and, consequently, the cathode sheath transformation can be written as

$$\int_0^d R(T, N_a)dy \sim \int_0^d \alpha(E/N)v_e n_e dy \sim A\frac{j}{e}, \qquad (3)$$

where the coefficient A is on the order of unity.

Calculations show that, to estimate the time during which a homogeneous cathode sheath exists, an accurate determination of A in (3) is not required because $R(T, N_a)$ depends strongly on the temperature. The reason is that as soon as the rate of associative ionization becomes comparable to the rate of electron-impact ionization, the parameters related to the onset of a cathode spot change extremely rapidly. Hence, the condition

$$\frac{e}{j}\int_0^d R(T, \xi)dy = 0.5 \qquad (4)$$

can be regarded as a criterion for the beginning of these changes.

The onset of a prearc cathode spot in an anomalous cathode sheath can be explained as follows. As the temperature increases, condition (4) is first satisfied locally (e.g., in the center of the sheath, where the gas temperature is maximum), whereas the temperature at the sheath periphery is lower due to heat transfer to the current-free region. In the region where condition (4) is satisfied, the cathode potential drop decreases and the boundary of the diffuse glow approaches the cathode surface. The cathode potential drop equalizes over the radius due to the rearrangement of the current density, which decreases at the periphery and increases in the center. Due to the decrease in the current density, the degree of the cathode sheath anomaly and, consequently, the electric field reduce. If the total current is sufficiently high, then its contraction into a prearc spot enhances the rate of thermal ionization up to the critical level, which results in the creation of a self-sustained arc [2–4]. If the current is low, then, after a certain period of time, the prearc spot disappears and the cathode sheath returns into the uniform anomalous regime. This sequence of events may repeat itself again, which corresponds to a self-oscillating mode [4]. The return to an anomalous mode is possibly related to a decrease in the energy deposition and, consequently, a decrease in

the gas temperature due to a substantial reduction of the electric field. Some of these effects are considered in detail in [15].

Nonequilibrium ionization is one of the key processes of our model; it depends on the effective temperature of fast molecules, which is determined based on the following assumptions. The ion velocity distribution function, which is formed in the strong electric field due to the intense charge exchange processes, is close to Maxwellian; the mean ion drift energy is equal to $\pi m v_+^2/4$ [17], where mv_+ is the ion momentum. Due to the high efficiency of charge exchange, the fast molecules have the same distribution; hence, taking into account the gas temperature, the mean energy of the fast molecules after charge exchange is

$$E = \frac{\pi}{4} m v_+^2 + \frac{3}{2} kT. \tag{5}$$

For ions, the analogous formula is given in [14]. Under the conditions adopted in our study, the molecules have equal or nearly equal masses and the efficiency of energy transfer is very high. Hence, the largest amount of energy is transferred in the first (after the charge exchange event) collision between fast and slow molecules. Averaging formula (5) over the colliding molecules (their number is twice the number of the fast molecules), we obtain

$$E_{\text{eff}} = \frac{\pi}{8} m v_+^2 + \frac{3}{2} kT, \tag{6}$$

which corresponds to the effective temperature

$$T_{\text{eff}} = \frac{\pi}{12k} m v_+^2 + T. \tag{7}$$

Here, the correction for the fact that the temperature is nonequilibrium due to the presence of a source of fast molecules is not taken into account because this is a separate problem. Such a simplification is justified by the fact that, in the CDS, the number of fast molecules is much less than that of molecules with thermal velocities.

System of Gas-Dynamic Equations

The Knudsen number computed based on the molecule mean free path and the thickness of an anomalous CDS amounts to 0.05 for air and 0.02 for nitrogen. Hence, the gas flow in the CDS is intermediate between the continuous and thinned modes. This is probably the reason why the preliminary calculations involving the Navier–Stokes equations did not provide satisfactory results. Hereafter, we use the kinetically consistent difference schemes [18], which provide satisfactory results when calculating the flow of a moderately rarefied gas with Kn ≤ 0.1. The equations

$$\frac{\partial \rho}{\partial t} + \frac{\partial \rho v}{\partial y} = \frac{\partial}{\partial y}\left(\tau \frac{\partial}{\partial y}(\rho v^2 + p)\right), \tag{8}$$

$$\frac{\partial \rho v}{\partial t} + \frac{\partial (\rho v^2 + p)}{\partial y} = \frac{\partial}{\partial y}\left(\tau \frac{\partial}{\partial y}(\rho v^3 + 3pv)\right) + \frac{\partial}{\partial y}\left(\mu \frac{\partial v}{\partial y}\right), \tag{9}$$

$$\frac{\partial U}{\partial t} + \frac{\partial (U + p)v}{\partial y} = \frac{\partial}{\partial y}\left(\tau \frac{\partial}{\partial y}(U + 2p)v^2\right) + \frac{\partial}{\partial y}\left(\tau \frac{\partial}{\partial y} p(U + p)/p\right) + \frac{\partial}{\partial y}\lambda \frac{\partial T}{\partial y} + Q(e, j, E), \tag{10}$$

are the differential approximation of these schemes. As compared to the Navier–Stokes equations, they contain additional terms

$$Q(e, j, E) = jE - N\left(\Theta_{O_2} \frac{e_{O_2} - e_{O_2}^0}{\tau_{O_2, R}} + \Theta_{N_2} \frac{e_{N_2} - e_{N_2}^0}{\tau_{N_2, R}}\right) + \nabla\left(RT \sum_{i=1}^{n_K} \frac{D_i^T}{m_i} \nabla(\ln c_i)\right), \tag{11}$$

where t is time; v, ρ, and p are the gas velocity, density, and pressure, respectively; $U = \rho c_p T + 0.5\rho v^2 + N(e_{O_2}\Theta_{O_2} + e_{N_2}\Theta_{N_2})$ is the total energy per unit volume; e_{O_2} and e_{N_2} are the fractions of the vibrationally excited molecules per molecule in the mixture (e^0 stands for their equilibrium values); $\Theta_{O_2} = 2240$ K and $\Theta_{N_2} = 3360$ K are the vibrational energies of O_2 and N_2; c_p is the heat capacity; μ is the viscosity; λ is the thermal conductivity; τ is the mean free path of the molecules; D_i^T is the coefficient of thermodiffusion of the ith component; $\tau_{O_2, R}$ and $\tau_{N_2, R}$ are the times of molecular vibrational relaxation; R is the molar gas constant; c_i and m_i are the mass fraction and molecular weight of the ith component, respectively; and n_K is the number of the mixture components.

Substituting $Q(e, j, E)$ into the energy balance equation (10), we take into account molecular vibrational relaxation, Joule heating of the gas, and diffusion heat transfer. The diffusion coefficient, thermal conductivity, and mixture heat capacity were calculated by the formulas from [19], and the relaxation times were calculated by the formulas from [13].

Note that, from the physical standpoint, differential approximations (8)–(10) are the consequence of finite difference schemes, which are derived based on an analysis of a stepwise–constant one-particle distribution function [18]. Additional terms with τ in equations

appear due to the significant change in the flow parameters over the molecular mean free path. Perhaps these terms are similar to the artificial viscosity, widely used in the calculations of gas flows with high gradients of parameters. In our case, the appearance of these terms is physically even more justified. The larger the Knudsen number, the more important these terms. Under conditions when these terms can be neglected, Eqs. (8)–(10) coincide with the Navier–Stokes equations. For a gas mixture, the equation of state is

$$p = \rho RT\sum_{i=1}^{N_k}\frac{c_i}{m_i}. \tag{12}$$

To determine the temperature at the cathode surface, we took into account the continuity of the heat flux and solved the equation

$$c_C\rho_C\frac{\partial T_C}{\partial t} = \lambda_C\frac{\partial^2 T_C}{\partial y_C^2}, \tag{13}$$

where index C refers to the cathode.

The initial temperature of both the gas and cathode is 300 K. The other boundary conditions at the surface are set in the same manner as in [18].

Equations of the Nonequilibrium Vibrational Chemical Kinetics

The equations

$$\frac{\partial e_{O_2}}{\partial t} + v\frac{\partial e_{O_2}}{\partial y} = \frac{\partial}{\partial y}D_{O_2}\left(\frac{\partial}{\partial y}e_{O_2} + \frac{m_{N_2}}{m}k_{O_2}^T\frac{\nabla T}{T}\right) + \frac{e_{O_2}-e_{O_2}^0}{\tau_{O_2,R}}, \tag{14}$$

$$\frac{\partial e_{N_2}}{\partial t} + v\frac{\partial e_{N_2}}{\partial y} = \frac{\partial}{\partial y}D_{N_2}\left(\frac{\partial}{\partial y}e_{N_2} + \frac{m_{N_2}}{m}k_{N_2}^T\frac{\nabla T}{T}\right) + \frac{e_{N_2}-e_{N_2}^0}{\tau_{N_2,R}}, \tag{15}$$

were used to calculate the degree of excitation of the molecular vibrational degrees of freedom.

To calculate the mass fractions c_i of the mixture components, we used equations from [18]:

$$\frac{\partial c_i}{\partial t} + v\frac{\partial c_i}{\partial y} = \frac{\partial}{\partial y}D_i\left(\frac{\partial}{\partial y}c_i + \frac{m_i}{m}k_i^T\frac{\nabla T}{T}\right) + W_i, \tag{16}$$

$$W_i = m_i\sum_{r=1}^{N_{hr}}(\nu_{sr}-\nu_{sr}^+) \times\left(k_r^+\prod_{k=1}^{N_s}\left(\frac{\rho_k}{m_k}\right)^{\nu_{kr}^+} - k_r^-\prod_{k=1}^{N_s}\left(\frac{\rho_k}{m_k}\right)^{\nu_{kr}^-}\right), \tag{17}$$

where k_i^T is the thermodiffusion coefficient; W_i is the production rate of the ith component; k_r^+, k_r^-, ν_r^+, and ν_r^- are the rate constants and stoichiometric coefficients of the forward and backward reactions; and ρ_k are the partial densities of the components.

The initial mole concentrations of the components were as follows: $\xi_{O_2} : \xi_{N_2} = 0.21 : 0.79$ for air and $\xi_{N_2} = 1$ for nitrogen. The dissociation energy for nitrogen is substantially higher than for oxygen; hence, the dissociation of nitrogen molecules in air starts when the dissociation of oxygen molecules is almost finished [19]. In addition, the ionization potential of N_2 is higher than that of O_2, which permits us to neglect the nitrogen component of the ion current. For this reason, when considering the initial stage of molecular dissociation and gas heating in an air discharge, we simplify the problem by taking into account only the following reactions

$$\begin{gathered}O_2 + O_2^+ \longrightarrow O_2^* + O_2^+,\\ O_2^* + M \longrightarrow 2O + M,\\ O_2 + M \longleftrightarrow 2O + M, \quad O + O \longrightarrow O_2^+ + e,\end{gathered} \tag{18}$$

where O_2^* stands for fast oxygen molecules and M = O_2, N_2,, or O.

For nitrogen, we take into account the reactions

$$\begin{gathered}N_2 + N_2^+ \longrightarrow N_2^* + N_2,\\ N_2^* + M \longrightarrow 2N + M,\\ N_2 + M \longleftrightarrow 2N + M, \quad N + N \longrightarrow N_2^+ + e,\end{gathered} \tag{19}$$

where $M = N_2$, N.

The rates of chemical reactions (18) and (19) were calculated by formulas from [21]. In both cases, the densities of fast molecules were assumed to be equal to the ion density $n_+ = j_+/(ev_+)$, where $v_+ \sim (E/N)^{0.5}$ [20]. To calculate the rate of nonequilibrium dissociation with the participation of fast molecules in reactions (18) and (19), we used the Macherett–Freedman model [21], in which the reaction rate constants are calculated by formulas

$$k^+ = Z(T_{tr}, T_v)k_0^+(T), \tag{20}$$

where k^+ and k_0^+ are the nonequilibrium and equilibrium rate constants of the forward reactions, $Z(T_{tr}, T_v)$ is the coefficient characterizing the degree to which the process is nonequilibrium [21], T_{tr} and T_v are the molecular translational and vibrational temperatures.

Since the bulk of the gas molecules have thermal velocities, we set $T_{tr} = T$, whereas for fast molecules,

we set $T_{tr} = T_{eff}$ [see (7)]. The dissociation of molecules, which also occurs in collisions with electrons, is difficult to calculate when the electron energy distribution in the CDS is nonlocal. In this study, the distribution function was not calculated. In order to estimate the relevant error, we used two approximations: (i) an approximation in which reactions (18) and (19) were only taken into account and (ii) an approximation in which, in addition to the above reactions, electron-impact ionization

$$O_2 + e \longrightarrow O_2^+ + e, \quad N_2 + e \longrightarrow N_2^+ + e, \qquad (21)$$

was taken into account.

Cross sections for reactions (21) were set as $\sigma_1 = 2 \times 10^{-16}$ cm^2 and $\sigma_2 = 0.5 \times 10^{-16}$ cm^2. The electron current in the CDS was assumed to be 10% of the total current. The above value of σ_1 is not lower than the maximum integral cross sections for the dissociation of N_2 and O_2 calculated with allowance for the excitation of different electronic states by the electrons with energies of 7–70 eV [16]. The maximum possible error can be estimated by this value because most of the electrons in the CDS have energies lower than 10 eV [16]. It is likely that the value of σ_2 is more realistic.

Formulas for Calculating the Electric Field for a Nonuniform Gas Density

In order to calculate the electric field distribution in the CDS, the sheath thickness, and the current density at the cathode, the formulas

$$\int_0^d N(y)/[N_0\bar{\delta}(y)]dy = d_N, \qquad (22)$$

$$E(y) = \frac{2U_N}{d_N}\frac{N(y)}{N_0}\bar{E}(y)\left(1 - \int_0^y \frac{N(y)}{N_0\bar{\delta}(y)}dy\right), \qquad (23)$$

$$U = \int_0^d E(y)dy, \qquad (24)$$

which are based on the local renormalization of the sheath parameters with respect to the gas density, were proposed in [9]. Here, N_0 is the molecule density corresponding to the uniform gas temperature T_0; $N(y)$ is the molecule density distribution in the gas flowing through the CDS; U_N and d_N are the normal values of the cathode potential drop and cathode sheath thickness for $N(y) = N_0$; U and d are the cathode potential drop and cathode sheath thickness for the nonuniform distribution $N(y)$; $\bar{E}$ and $\bar{\delta}$ are the normalized coefficients, which depend on the ratio of the current density to the local value of the normal current density, which, in turn, is in proportion to $N^2(y)$.

The experimental data for an air–copper cathode system (U_N = 370 V, pd_N = 0.23 torr cm, and j_N/p^2 = 2.4 A/(m^2 torr^2)) were taken from [16]. For a nitrogen–copper cathode system, we set U_N = 208 V, j_N/p^2 = 4 A/(m^2 torr^2), and pd_N = 0.5 torr cm. The latter value was determined in test calculations [9] and does not contradict the experimental results.

Formulas (22)–(24), inferred from the main scaling laws of the Engel–Steenbeck theory [16], were successfully used to solve a number of applied problems [2, 10, 12]. They describe cathode sheath at a fixed distribution of the gas density. Use of them in our model is justified because the electric parameters change more rapidly than the gas-dynamic ones. Furthermore, these formulas substantially simplify calculations because they do not require solving a set of stiff differential equations. The use of the empirical constants U_N, j_N/p^2, and pd_N is also advantageous. A disadvantage is the insufficient accuracy of calculating the electric field at the high rate of thermal ionization, when, in addition to the Townsend ionization coefficient, the other volume electron sources should also be taken into account. However, as was mentioned above, this circumstance is of minor importance for determining the time at which the current–voltage characteristic becomes descending and the cathode sheath begin to transform.

SIMULATION RESULTS

Figure 1 shows the simulated dependences of the ratio of the integral rate of associative ionization in the CDS to the charged particle flux density (4),

$$R_{int} = e/j\int_0^d R(T, N_a)dy.$$

The dependences are obtained for nitrogen at p_0 = 50 torr and j = 100 A/cm^2. It is seen that condition (4), which determines the critical value of the rate of thermal ionization at which the transformation of the cathode sheath takes place, is satisfied at 35 ns. This value is obtained without allowance for electron-impact dissociation (21). With allowance for electron-impact dissociation, condition (4) is satisfied at 27.1 ns for the cross section σ_1 and at 31.5 ns for the cross section σ_2. These times agree with experiments [2], in which the homogeneity of the cathode sheath was broken within 30–40 ns. In Fig. 1, R_{int} increases very rapidly (from 0.1 to 0.5 in approximately three nanoseconds). The distributions of the gas parameters along the electric field at t = 5, 20, and 35 ns are shown in Figs. 2–5. The thickness of the cathode sheath increases from 0.046 mm at the initial instant to 0.055 mm at t = 35 ns. This thickness is shorter than the normal value d_N = 0.1 mm because the degree of the cathode sheath anomaly is

fairly high. In Fig. 2, there is a local maximum in the gas temperature T_{max}, which occurs due to Joule heating and heat transfer to the cathode. As time elapses, the maximum shifts away from the cathode due to the sheath expansion. At t = 35 ns, we have T_{max} = 7658 K, which is 5.38 times higher than the maximum value of the nitrogen vibrational temperature T_v (1422 K) and 12.6 times less than the maximum value of the effective temperature of fast molecules $T_{eff} = 9.63 \times 10^3$ K. The presence of the maximum of T_{eff} stems from the occurrence of the maximum of E/N across the cathode sheath at a nonuniform gas density distribution [12].

The parameters of the gas and cathode sheath vary with time. In the initial stage (t = 5 ns), the gas in the CDS is slightly rarefied and the pressure distribution is similar to the temperature distribution, with the maximum inside the cathode sheath (Fig. 3). Hence, immediately after the formation of the cathode sheath, the gas begins to be expelled from the region with an elevated pressure toward both the cathode and positive column (Fig. 4). Behind the front of the shock wave propagating from the cathode, there is a jump in the gas density (Fig. 5). As in experiments [7], the jump is peak-shaped. The flow directed toward the cathode compresses the gas near the cathode surface. As a result, a pressure drop is produced there, which pushes the gas toward the positive column.

Both the speed at which gas leaves the cathode sheath and the degree of gas rarefaction increase with time. At t = 35 ns, the maximum velocity is V_{max} = 790 m/s. The gas density behind the shock wave is 2.9 times as high as that in the unperturbed flow. The pressure at the cathode is 14.2 times the initial one. The concentration of nitrogen is maximum ($C_N = 2.76 \times 10^{-4}$) at a certain distance from the cathode (Fig. 6). During 35 ns, the cathode potential drop increases from 632 to 806 V (such a high magnitude is related to the strong anomaly of the cathode sheath) and the cathode surface temperature increases to T_C = 702 K, which is lower than the copper melting point.

In nitrogen at j = 1000 A/cm^2 and p_0 = 100 torr, the cathode spot was experimentally observed to appear within a few nanoseconds [2]. In our calculations carried out under the same conditions without taking into account the electron-impact dissociation of nitrogen, condition (4) is satisfied at t = 2.1 ns. By this time, the maximum values of the parameters are as follows: T_{max} = 9016 K, T_v = 1031 K, $T_{ion} = 115 \times 10^3$ K, V_{max} = 161 m/s, $C_0 = 9.7 \times 10^{-4}$, and T_C = 704 K. Because of the short duration of the process and, consequently, insignificant rarefaction of the gas, the cathode potential drop changes only slightly, being approximately equal to 937 V. With allowance for the electron-impact dissociation of nitrogen (with the cross section σ_1), condition (4) is satisfied at t = 2.07 ns. In spite of a satisfactory agreement between the simulated and measured times, it should be noted that, in this case, there is

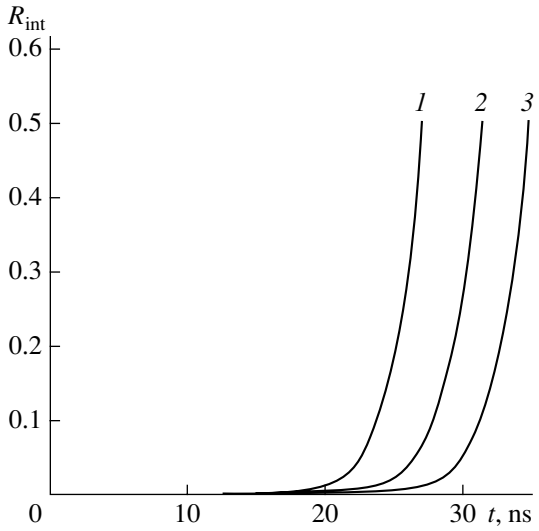


Fig. 1. Time evolution of the relative integral rate of associative ionization with allowance for the electron-impact dissociation of molecules with cross sections (*1*) σ_1 and (*2*) σ_2 and (*3*) with allowance for dissociation in atomic collisions only.

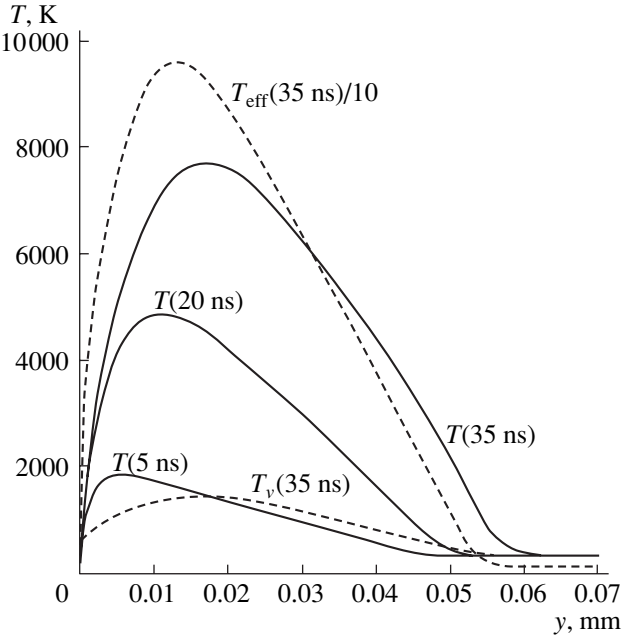


Fig. 2. Profiles of the vibrational, gas, and effective temperatures (T_v, T, and T_{eff}) of fast molecules across the cathode sheath.

a considerable uncertainty in calculating the gas-dynamic characteristics in the early stage of the process. This is related to the fact that, in nitrogen, under the equilibrium conditions corresponding to T = 300 K and p_0 = 100 torr, the mean free time between molecular collisions is relatively long (τ = 0.9 ns). However, this circumstance seems to be of minor importance for

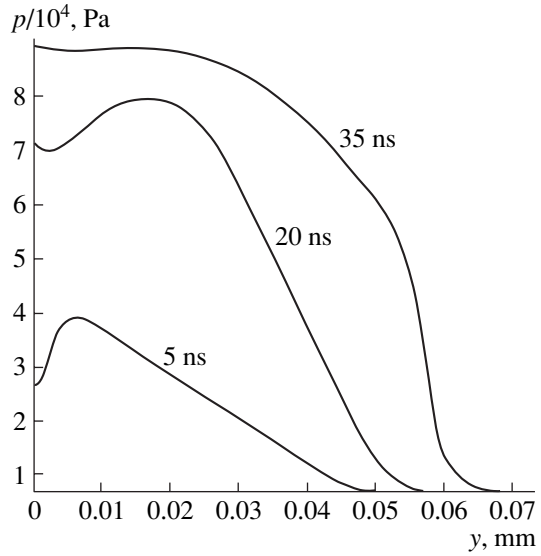


Fig. 3. Profiles of the gas pressure across the cathode sheath.

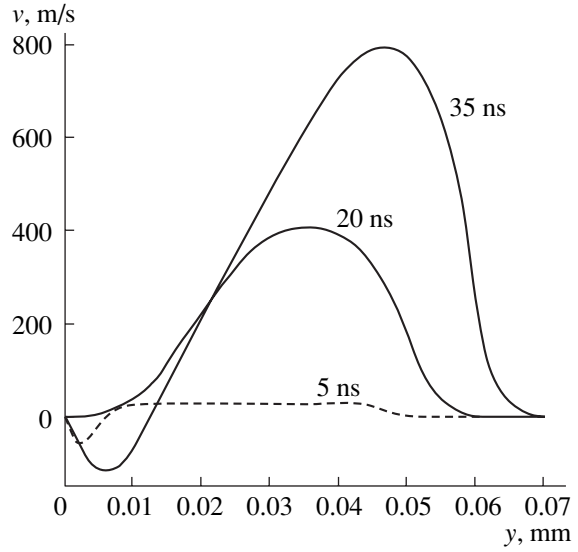


Fig. 4. Profiles of the gas velocity across the cathode sheath.

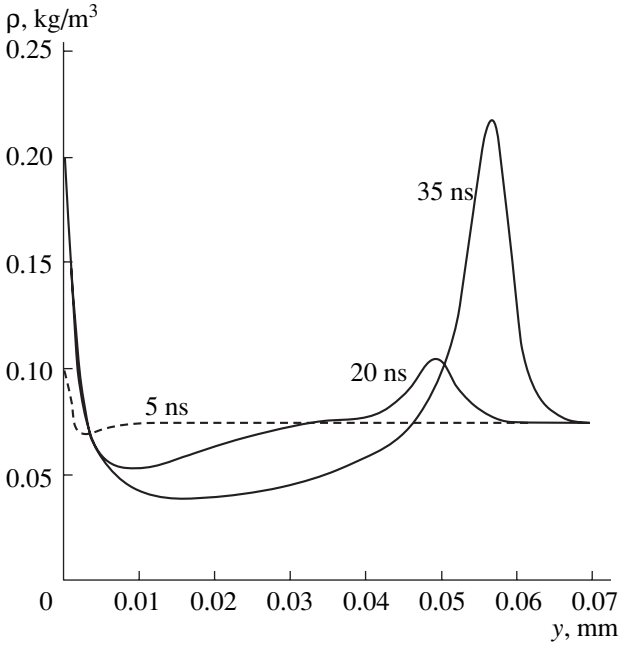


Fig. 5. Profiles of the gas density across the cathode sheath.

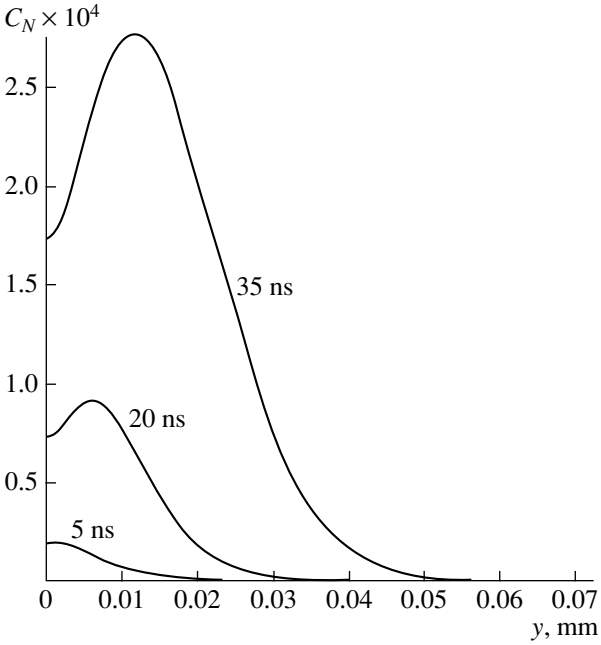


Fig. 6. Profiles of the molar concentration of nitrogen across the cathode sheath.

applying (4) because fast molecules heat the gas in a time much shorter than τ and the accompanying effects of gas motion and rarefaction are of no significance here.

In air at $p_0 = 76$ torr, the diffuse glow is experimentally observed to locally contact the center of the cathode just before the instant $t = 20$ ns [3]. At $t = 10$ ns, the current density is $j \sim 120$ A/cm^2. In simulations, we used the more accurate dependence $j(t)$ derived using an oscillogram from [3]. Condition (4) is satisfied at $t = 14.4$ ns. Within 1.3 ns, the quantity R_{int} increases 10-fold. At $t = 14.4$ ns, the maximum local values of the parameters are as follows: $T_{\max} = 4781$ K, $T_v = 543$ K, $T_{\text{eff}} = 268 \times 10^3$ K, $V_{\max} = 485$ m/s, $T_C = 585$ K, and $C_0 = 8.9 \times 10^{-2}$; the gas density drops by a factor of at most 1.64. The cathode potential drop increases from 729 to 1150 V, and the CDS thickness increases from 0.015 to 0.017 mm. Under the given conditions, the electron-impact dissociation of molecules only slightly affects the process even for the dissociation cross section σ_1.

Under the same conditions, the instability in air develops earlier than in nitrogen, which is related to the

higher value of U_N and lower value of j_N/p^2 in air (the lesser j_N, the higher the degree of the cathode sheath anomaly). Moreover, oxygen molecules decay faster than nitrogen ones, and the rate of their associative ionization increases more rapidly with temperature [13]. For this reason, the electron-impact dissociation of oxygen is of less importance. It follows from the above-said that the main parameters of the cathode sheath, namely U_N, $d_N p$, and j_N/p^2, from which the corresponding anomalous values can be calculated with corrections for the gas density distribution, should be accurately determined.

In the above three experiments, when condition (4) was satisfied, the electric field at the cathode surface was $E_0 = 2.7 \times 10^5$ (N_2, $j = 100$ A/cm^2), 8.31×10^5 (N_2, $j = 1000$ A/cm^2), and 1.48×10^6 V/cm (air). According to the theory of explosive autoemission [3], the above values of E_0 allow us to conclude that the cathode spot in air appears at the same time as in nitrogen at $j = 1000$ A/cm^2, which does not contradict the experimental results.

Note that E_0 increases with time due to gas compression; in nitrogen at $j = 100$ A/cm^2, it increases by 9.6% with respect to the initial value. This relatively small change in E_0 with a more than twofold increase in the gas density at the cathode (see Fig. 5) is related to the opposing action of the two factors. On the one hand, the increase in the gas density must lead to an increase in the electric field. On the other hand, the density growth locally decreases the degree of anomaly of the current density and, consequently, the electric field. As a result, these effects balance each other. In nitrogen at $j = 1000$ A/cm^2, E_0 increases by 2.5%. In air, the increase in E_0 is 59% of the initial value of 9.3×10^5 V/cm. This is related to the change in both the gas density and discharge current.

In conclusion, we note that estimating the electric field at the cathode by the average gas temperature is insufficiently accurate because, during the given computation time, the gas density at the cathode increases, whereas its average density in the CDS decreases. It is seen from Fig. 5 that the thickness of the region with an elevated gas density near the cathode is shorter than 10 μm, which is less than the spatial resolution of the interferometer [7]. At a pressure of 5 bar, this thickness will be 75 times less although the Knudsen number will change slightly because the molecular mean free path and d_N depend on the pressure in the same manner. Such a slight change in the Knudsen number widens the range of applicability of the model.

CONCLUSION

The dynamics of the CDS in a glow discharge has been described with the help of equations of motion of a moderately rarefied gas and the methods of nonequilibrium gas dynamics supplemented with equations for the electric field. An analysis of the conditions for the existence of a self-sustained discharge with allowance for the thermal ionization of the gas yields a criterion for calculating the time during which an anomalous cathode sheath is uniform. The time calculated by this criterion is in reasonable agreement with experimental data. In the cathode sheath, the rate of thermal ionization increases with increasing both the gas temperature and the rate of molecular dissociation. The vibrational temperature increases more slowly than the translational, which substantially decreases the dissociation rate. This effect is partially compensated for by the nonequilibrium distribution over the translational degrees of freedom, which is related to both the ion acceleration in the strong electric field and resonant charge exchange. Under the given conditions, the electron-impact molecular dissociation is not a decisive factor. This problem can be studied in more detail by calculating the nonlocal electron energy distribution function. If the duration of the process is long enough, the discharge stability is affected by the increase in the cathode potential drop, which is caused by gas rarefaction.

ACKNOWLEDGMENTS

This study was supported by the Russian foundation for Basic Research, project no. 98-01-01055.

REFERENCES

1. V. Yu. Baranov, V. M. Borisov, and Yu. Yu. Stepanov, *Inert-Gas Halogenide Electric-Discharge Excimer Lasers* (Énergoatomizdat, Moscow, 1988).
2. R. B. Baksht, Yu. D. Korolev, and G. A. Mesyats, Fiz. Plazmy **3**, 653 (1977) [Sov. J. Plasma Phys. **3**, 369 (1977)].
3. Yu. D. Korolev and G. A. Mesyats, *The Physics of Pulsed Gas Breakdown* (Nauka, Moscow, 1991).
4. Yu. S. Akishev, A. P. Napartovich, S. V. Pashkin, *et al.*, Teplofiz. Vys. Temp. **22**, 201 (1984).
5. P. Delaporte, B. Fontaine, B. Forestier, and M. Sentis, in *Proceedings of the 8th SPIE Gas Flow and Chemical Lasers Symposium, Bellingham, 1990*; Proc. SPIE **1397**, 485 (1990).
6. S. Kosugi, K. Maeno, and H. Honma, Jpn. J. Appl. Phys. **32**, 1480 (1993).
7. G. Schroder, J. Haferkamp, and W. Botticher, J. Appl. Phys. **78**, 4859 (1995).
8. V. M. Fomin, I. V. Shaimova, and V. A. Shveĭgert, in *Proceedings of the II All-Union Conference on Physics of Gas Breakdown, Tartu, 1984*, p. 178.
9. S. A. Smirnov, Preprint No. P-0918, NIIÉFA (TsNIIFTOMINFORM, Moscow, 1993).
10. G. A. Baranov and S. A. Smirnov, in *Proceedings of the XII International Symposium on Gas Flow and Chemical Lasers and High-Power Laser Conference*; Proc. SPIE **3574**, 820 (1998).

11. V. V. Osipov, Usp. Fiz. Nauk **170**, 225 (2000).

12. G. A. Baranov and S. A. Smirnov, Zh. Tekh. Fiz. **69** (11), 49 (1999) [Tech. Phys. **44**, 1305 (1999)].

13. S. A. Losev, V. N. Makarov, and M. Yu. Pogosbekyan, Mekh. Zhidk. Gaza, No. 2, 169 (1995).

14. E. W. McDaniel and E. A. Mason, *The Mobility and Diffusion of Ions in Gases* (Wiley, New York, 1973; Mir, Moscow, 1976).

15. S. A. Smirnov and G. A. Baranov, Zh. Tekh. Fiz. **71** (7), 39 (2001) [Tech. Phys. **46**, 825 (2001)].

16. Yu. P. Raizer, *Gas Discharge Physics* (Nauka, Moscow, 1987; Springer-Verlag, Berlin, 1991).

17. V. E. Golant, A. P. Zhilinskii, and S. A. Sakharov, *Fundamentals of Plasma Physics* (Atomizdat, Moscow, 1977; Wiley, New York, 1980).

18. I. V. Abalkin and B. N. Chetverushin, Mat. model. **4** (11), 19 (1992).

19. Yu. V. Lapin and M. Kh. Strelets, *Internal Flows of Gas Mixtures* (Nauka, Moscow, 1989).

20. H. Raether, *Electron Avalanches and Breakdown in Gases* (Butterworths, London, 1964; Mir, Moscow, 1968).

21. S. A. Losev, A. L. Sergievskaya, V. D. Rusanov, *et al.*, Dokl. Akad. Nauk **346**, 192 (1996).

Translated by N. Ustinovskiĭ

Technical Physics, Vol. 46, No. 7, 2001, pp. 825–830. Translated from Zhurnal Tekhnicheskoĭ Fiziki, Vol. 71, No. 7, 2001, pp. 39–43.

Gas Dynamics and Thermal-Ionization Instability of the Cathode Region of a Glow Discharge. Part II

S. A. Smirnov and G. A. Baranov

Efremov Research Institute of Electrophysical Apparatus, St. Petersburg, 189631 Russia

Received July 10, 2000

Abstract—The influence of the gas flow structure in the cathode sheath of a glow discharge on the discharge stability is studied numerically. The electric parameters are calculated in a diffusion–drift model that consistently takes into account associative dissociation as an additional electron source. The model also includes equations describing both the thermal mode of the cathode and the nonequilibrium physicochemical gas dynamics of a moderately rarefied gas. It is shown that, in a pulsed discharge, the increasing branch of the current–voltage characteristic, which is associated with the gas rarefaction behind the cathode shock wave, can change to a descending branch associated with the intensification of associative ionization. This gives rise to cathode sheath instability. The results of calculations agree well with experiments. © *2001 MAIK "Nauka/Interperiodica"*.

INTRODUCTION

In the first part of this paper [1], we developed a model describing the cathode sheath of a glow discharge. The model includes formulas for calculating the electric field and the set of equations for the dynamics of unsteady flows of a moderately rarefied gas, vibrational molecular kinetics, nonequilibrium dissociation, and cathode temperature. A criterion was also proposed for determining the discharge duration in a homogeneous mode. The transition to a prearc cathode spot was attributed to the fact that the associative ionization rate increased up to a value comparable to the electroionization rate. The testing of the model with the use of experimental data [2, 3] showed that the calculated duration of the homogeneous phase of a strongly anomalous glow discharge agreed satisfactorily with the experimental data.

An advantage of simple formulas used in [1] to calculate the electric field distribution in the cathode dark space (model I) is that they are based on the empirical data on the cathode potential drop, the cathode sheath thickness, and the normal current density. However, these formulas fail to describe the influence of the gas thermal ionization rate on the electric parameters of the cathode sheath. This disadvantage is overcome in this part of the paper because the proposed model of a cathode sheath includes the set of ion and electron balance equations in which the production of charged particles is described by both the Townsend ionization coefficient and thermal ionization rate. The field distribution is determined from Poisson's equation.

DIFFUSION–DRIFT MODEL OF A GLOW DISCHARGE

Here, we present a set of one-dimensional equations for calculating the electric field and the ion and electron densities in the diffusion–drift approximation. These equations are widely used to study discharges (see, e.g., [4, 5]); however, in this paper, they also include the terms related to gas-dynamic and chemical processes:

$$\frac{\partial n_e}{\partial t} + v_e\frac{\partial n_e}{\partial y} = \frac{\partial}{\partial y}D_e\frac{\partial n_e}{\partial y} + \alpha(E/N)v_e n_e - \beta n_e n_i + R(N_a, T), \quad (1)$$

$$\frac{\partial n_i}{\partial t} + v_i\frac{\partial n_i}{\partial y} = \frac{\partial}{\partial y}D_i\frac{\partial n_i}{\partial y} + \alpha(E/N)v_e n_e - \beta n_e n_i + R(N_a, T), \quad (2)$$

$$\frac{\partial^2\varphi}{\partial y^2} = 4\pi(n_i - n_e), \quad (3)$$

$$E = \frac{\partial\varphi}{\partial y}, \quad (4)$$

where t is time, y is the longitudinal coordinate, n_e and n_i are the electron and ion densities, v_e and v_i are the electron and ion velocities, D_e and D_i are the electron and ion diffusion coefficients, φ is the potential, E is the electric field strength, N is the molecular density, $\alpha(E/N)$ is the Townsend ionization coefficient, β is the recombination coefficient, $R(N_a, T)$ is the associative-ionization rate in the processes $A + A \longleftrightarrow A_2 + e$ [1], N_a is the atomic density, and T is the translational gas temperature.

1063-7842/01/4607-0825$21.00

The ion and electron velocities are calculated by the formulas

$$v_e = -\mu_e E, \quad v_i = \mu_i E + v, \tag{5}$$

where μ_e and μ_i are the electron and ion mobilities and v is the gas velocity.

The boundary conditions at the cathode ($y = 0$) are

$$e\mu_i n_i E = \frac{j}{1+\gamma}, \quad n_e = \gamma n_i \frac{\mu_i}{\mu_e}, \quad \varphi = 0, \tag{6}$$

and those at the anode ($y = H$) are

$$\frac{\partial n_e}{\partial y} = 0, \quad n_i = 0, \quad \varphi = \varphi_a, \tag{7}$$

where γ is the secondary emission coefficient; e is the electron charge; j is the current density; φ_a is the anode potential; and the values of μ_e, μ_i, D_e, D_i, $\alpha(E/N)$, and β are taken from [4].

To solve Eqs. (1)–(7) with boundary conditions (6) and (7) numerically, we used a finite-difference scheme of second-order accuracy [4].

The set of equations (1)–(5) with boundary conditions (6) and (7) and equations for calculating the cathode temperature and nonequilibrium physicochemical gas dynamics [1] constitute the model of a cathode sheath (model II). Using this model, we calculated discharges in nitrogen under experimental conditions [2] for current densities of 100 and 1000 A/cm^2 and pressures of $p_0 = 50$ and 100 torr, respectively. An advantage of model II in comparison to model I is that it includes Eqs. (1)–(5) with boundary conditions (6) and (7), which allows us to study in detail the processes associated with the production and transport of charged particles. However, we should note that, when using the diffusion–drift model of the cathode sheath, it is always necessary to fit the results obtained to experimental data. In the first place, this is achieved by properly choosing the secondary emission coefficient γ in order for the calculated values of the cathode potential drop and the sheath thickness to agree more or less satisfactorily with experimental data. It is more difficult to simultaneously obtain good agreement of the three main parameters U_N, $d_N p$, and j_N/p^2 with experimental data. From this point of view, model I has advantages, as was mentioned above. Here, the parameter γ was chosen as 0.25.

RESULTS OF CALCULATIONS

First, we performed test calculations of the electric parameters of the cathode sheath by using Eqs. (1)–(5) with boundary conditions (6) and (7) without including gas dynamics, i.e., for a constant gas density and without allowance for thermal ionization. The nitrogen pressure at a temperature of 300 K was equal to $p_0 = 50$ torr, the discharge gap length was 0.25 mm, and the electrode voltage was $\varphi_a = 450$ V.

Figure 1 shows the longitudinal profiles of the potential, electric field, and electron and ion densities. For the given value of φ_a, the steady-state current density was equal to 28.9 A/cm^2. The sheath thickness and the cathode potential drop were equal to $d = 0.042$ mm and $U_C = 404$ V, respectively. These values agree well with the values $d = 0.052$ mm and $U_C = 406$ V calculated by the Engel–Steenbeck theory for the same current density and the normal values $U_N = 208$ V, $i_N/p^2 4$ A/(m^2 torr^2), and $pd_N = 0.5$ torr cm [6]. It is seen in Fig. 1 that the electric field profile in the cathode sheath is linear. This agrees with the theory and later experiments [7] in which the electric field profile in the cathode sheath was measured in hydrogen at low gas pressures, when gas heating can be ignored.

In most of the cathode sheath, the ion density is higher than the electron density, which increases sharply only near the outer boundary of the sheath. This explains why the diffusion–drift model, which adequately describes the ion current without allowance for local effects, also correctly describes the electric field profile. In the positive column, we obtained $E/p = 40$ V/(cm torr). For a higher current density of 113 A/cm^2, we obtained $E/p = 47$ V/(cm torr). This agrees well with [8], in which both calculations and experiments with a pulsed discharge gave $E/p = 33$–44 V/(cm torr). In experiments with discharges in nitrogen [9], at significantly smaller current densities, the reduced field was $E/p \approx 30$ V/(cm torr). It is seen in Fig. 1 that, near the anode, there is a sheath with a negative space charge and a relatively small potential drop.

Our results do not contradict the well-known data on the cathode sheath and positive column at a constant gas density. However, as was shown in [1, 10], at $p > 5$ torr, gas heating affects the cathode sheath parameters. This effect becomes even more pronounced at higher pressures and as the current density increases above its normal value. For this reason, in order to adequately calculate the cathode sheath, a set of gas-dynamic equations should be incorporated in the model.

Figure 2 shows the longitudinal profiles of the charge density and gas density calculated using the full model II (i.e., using Eqs. (1)–(5) with boundary conditions (6) and (7) and the set of physicochemical gas-dynamic equations [1]). The calculation parameters are the following. The gas pressure in a nitrogen–copper cathode system is $p_0 = 50$ torr, the current density is $j = 100$ A/cm^2, and the interelectrode distance is 0.2 mm. The profiles are given for times $t = 20$ and 45 ns. The time 20 ns corresponds to a discharge mode with a relatively low level of thermal ionization. The time 45 ns corresponds to the mode characterized by intense thermal ionization, i.e., to a prearc spot [1]. It is seen that,

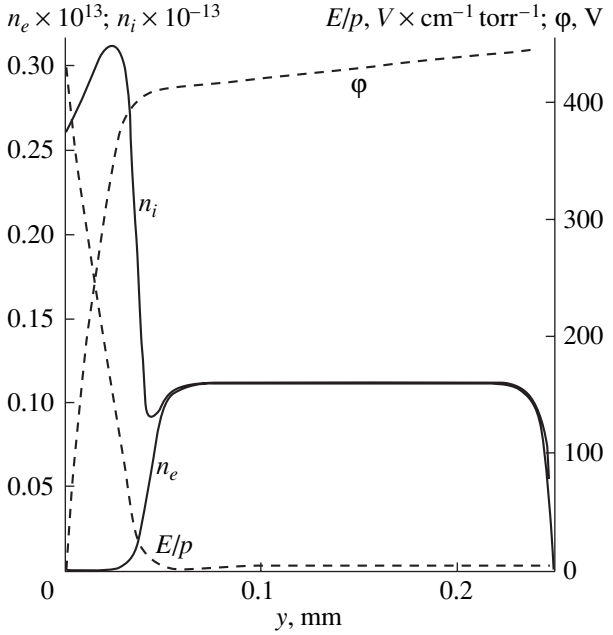


Fig. 1. Longitudinal profiles of the potential, reduced electric field, and ion and electron densities in the discharge gap.

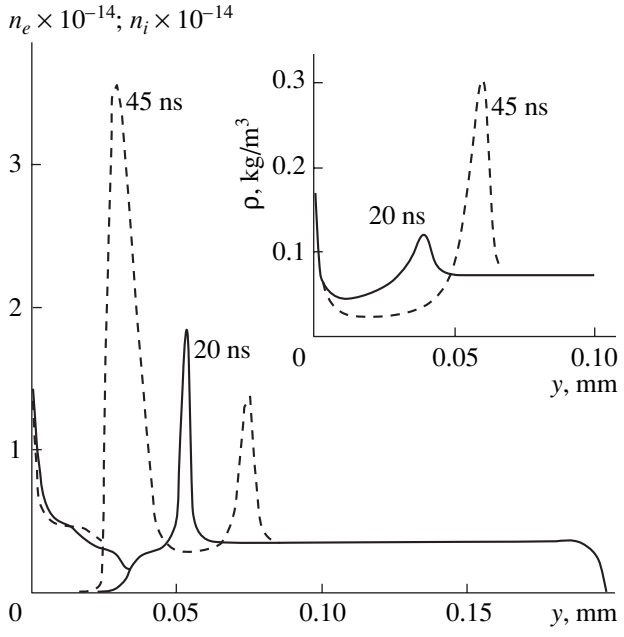


Fig. 2. Longitudinal profiles of the ion and electron densities and the gas density (ρ) in the discharge gap.

with gas dynamics taken into account, the profile of the charge density changes substantially. The first local maximum of the ion density appears on the cathode surface, in the region of local gas compression. The second maximum of both the ion and electron densities occurs at the front of the cathode shock wave. Since, at the times under consideration, the shock wave has already left the sheath, the plasma at the shock front is quasineutral.

At 45 ns, an anomalously high maximum in the ion and electron densities is seen immediately behind the outer boundary of the cathode sheath; the plasma in this region is also quasineutral. By this time, the sheath thickness decreases nearly two times (Fig. 3) and the cathode potential drop decreases from 706 (at 24 ns) to 400 V. This is due to the effect of thermal ionization, which serves as an additional electron source in the region where the strong field has already decayed.

Starting from 36 ns, the maximum gas temperature attains 7685 K in spite of a substantial decrease in the cathode potential drop. From 40 ns, the gas temperature decreases only slightly (Fig. 4). In this case, gas cooling in the outer part of the cathode sheath is balanced by gas heating in the inner part. It is interesting that, under conditions corresponding to the increasing branch of the current–voltage (*I–V*) characteristic (i.e., when associative ionization is still unimportant), models I and II give approximately the same values of the cathode drop and gas temperature in the cathode sheath (Figs. 3, 4).

Figure 4 shows the time dependence of the ratio $R_{\text{int}} = (e/j)\int_0^d R(N_a, T)dy$ of the integral rate of associative ionization in the cathode sheath to the current density (i.e., the double value of the integral of the impact-ionization rate). In [1], it was demonstrated that, in cal-

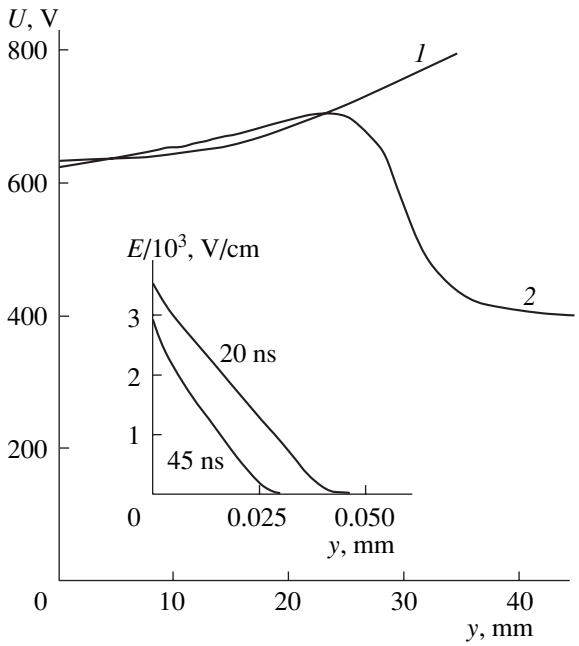


Fig. 3. Longitudinal profiles of the electric field in the cathode sheath and time dependences of the cathode potential drop calculated by (*1*) model I and (*2*) model II.

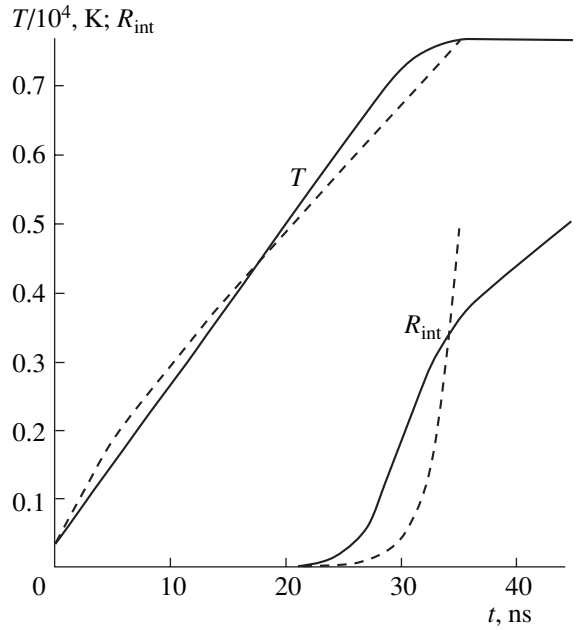


Fig. 4. Time dependences of the maximum gas temperature T in the cathode sheath and the relative integral associative ionization rate calculated by model I (dashed line) and model II (solid line) at a constant current density.

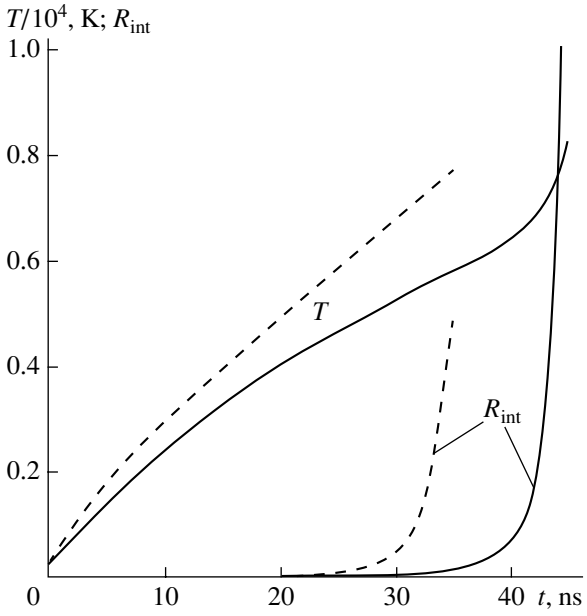


Fig. 5. Same as in Fig. 4, but at a constant cathode potential drop.

culations, the time at which the condition $R_{int} = 0.5$ is satisfied agrees well with the experimental time at which a prearc cathode spot appears. It is seen in Fig. 4 that, 30 ns later, the growth rate of R_{int} calculated by model II decreases due to a decrease in the cathode potential drop, so that R_{int} reaches a value of 0.5 at 45 ns. In model I, this condition is satisfied at 35 ns. In experiments [2], a prearc cathode spot appeared at 30–40 ns. Hence, in spite of the substantial modification of the model, the $R_{int} = 0.5$ criterion introduced in [1] remains valid.

Previous calculations were performed at a fixed value of the current density. However, a decrease in the cathode potential drop under conditions corresponding to the descending branch of the *I–V* characteristic should lead to an increase in the current density. Approximate models explaining the behavior of the cathode sheath on the descending branch of the *I–V* characteristic were proposed, e.g., in [11, 12]. They do not allow us to determine the characteristic growth time of the instability (i.e., the transition time from the subnormal to the normal mode), but show that the diameter of the current region under conditions corresponding to the descending branch of the *I–V* characteristic decreases and the current density increases.

A similar mechanism is realized in the processes under consideration, as was discussed in [1]. This mechanism may first manifest itself not throughout the entire diameter of the cathode sheath, but in the region where the associative ionization rate has a local maximum due to fluctuations in the parameters. The cathode potential drop at the outer boundary of the cathode sheath is smoothed out due to the current flowing into this region. Our one-dimensional model allows us to model this process, assuming that the diameter of the current region is large enough and the current density is redistributed rapidly (i.e., that the diameter of the region where the associative ionization rate attains its critical level is much smaller than the diameter of the current region at the cathode).

To study this effect, we performed calculations under the following conditions. The current density $j = 100$ A/cm^2 was specified only at the initial time. Further, only the interelectrode voltage was kept constant, whereas the current density varied according to the conditions in the interelectrode gap. Since, in our calculations, the potential drop across the cathode sheath is much higher than that through the positive column, these conditions correspond to the modeling of a discharge mode with a slightly varying cathode potential drop. This allowed us not only to determine R_{int}, but also to analyze the behavior of the physical quantities that more definitely indicate the onset of the cathode sheath instability (e.g., the gas temperature and current density).

Figure 5 shows the time dependences of R_{int} and the maximum temperature in the cathode sheath. For comparison, the data obtained by model I at a fixed current density of $j = 100$ A/cm^2 are also shown. It is seen that, until nearly 40 ns, the temperature growth rate calculated by model II is smaller than that calculated by model I. The reason is that the cathode drop in model II increases only slightly and this slight increase is balanced out by a decrease in the current density. The cath-

ode potential drop in model I increases (Fig. 3). For this reason, associative ionization comes into play at later times. However, after the quantity R_{int} has reached a value of 0.1, it begins to grow explosively because of the increased current density; the temperature grows in the same manner. A similar time behavior of R_{int}, U_C, and T_{max} was obtained using model II for nitrogen at $p_0 = 100$ torr and $j = 1000$ A/cm^2. The calculated growth time of the cathode sheath instability in this model agrees satisfactorily with experimental data [2].

Therefore, a model of the cathode sheath has been developed without introducing additional fit parameters. We used an adequate set of equations of gas dynamics with allowance for gas rarefaction. The apparatus of nonequilibrium physicochemical gas dynamics included in the model was developed to describe strong shock waves (references are given in [1]). The diffusion–drift model of the cathode sheath is well known. It adequately describes the quantities necessary for calculating the gas dynamics of the cathode sheath—the electric field and the cathode sheath thickness. For completeness of the picture, we added calculations of the thermal regime of the cathode; this was made not only to determine the boundary condition, but also to check the phase state of the surface. The fact that, in [2], dissociation was observed before the cathode spot was formed and metal vapor appeared points to the adequacy of the model, which, in our case, does not require invoking the poorly predictable mechanism of explosive emission. In [13], the dissociation of hydrogen was also observed experimentally before the formation of a cathode spot.

The data obtained allow us to propose the following scheme of the processes resulting in the thermal-ionization instability of the cathode spot. For a sufficiently high current, when there are resources for increasing the current density and the energy deposition increases monotonically, an irreversible contraction process take place:

$$(N_a, T)\uparrow \longrightarrow R_{int}\uparrow \longrightarrow U_K\downarrow$$
$$\longrightarrow j\uparrow \longrightarrow (jU_K)\uparrow \longrightarrow (N_a, T)\uparrow \ldots.$$

If the total current is insufficiently high for the plasma density to increase significantly as U_C decreases, the temperature can decrease because of heat transfer into the outer region. In this case, we obtain the self-oscillation mode studied experimentally in [14]. The scheme of the processes has the form

$$(N_a, T)\uparrow \longrightarrow R_{int}\uparrow \longrightarrow U_K\downarrow$$
$$\longrightarrow (jU_K)\downarrow \longrightarrow T\downarrow \longrightarrow R_{int}\downarrow$$
$$\longrightarrow U_K\uparrow \longrightarrow (jU_K)\uparrow \longrightarrow (N_a, T)\uparrow \ldots.$$

CONCLUSION

In a pulsed discharge at a constant current density, the increasing branch of the *I–V* characteristic of the cathode sheath is first formed. This branch is associated with an increasing excess of the current density over the normal density because of gas rarefaction. Local maxima of the electron and ion densities are formed at the front of the cathode shock wave. Another maximum of the ion density can occur in a narrow gas-compression region at the cathode.

The intensity of associative ionization increases with time. After it reaches the critical level, the *I–V* characteristics begin descending. Concurrently with the intensification of associative ionization, both the electric field and the cathode sheath thickness decrease. In a narrow region left by the outer boundary of the cathode sheath (i.e., behind the strong field region), charge particles continue to be produced due to thermal ionization. This gives rise to another local maximum characteristic of the unstable state of the cathode sheath.

Under conditions corresponding to the descending branch of the *I–V* characteristic, the discharge becomes unstable. The dynamics of this process depends on the external conditions. At a fixed value of the cathode potential drop, the current density increases as the associative ionization rate increases. After reaching the critical level, the associative ionization rate, the current density, and the gas temperature grow in an explosive manner. The calculated growth time of the instability agree satisfactorily with experimental data. Both models give close values of the cathode potential drop and the maximum gas temperature under conditions corresponding to the increasing branch of the *I–V* characteristic. In both models, the stability criterion defined by the integral associative ionization rate allows one to predict the time during which the prearc cathode spot is formed.

ACKNOWLEDGMENTS

This work was supported in part by the Russian Foundation for Basic Research, project no. 98-01-01055.

REFERENCES

1. S. A. Smirnov and G. A. Baranov, Zh. Tekh. Fiz. **71** (7), 30 (2001) [Tech. Phys. **46**, 815 (2001)].
2. R. B. Baksht, Yu. D. Korolev, and G. A. Mesyats, Fiz. Plazmy **3**, 653 (1977) [Sov. J. Plasma Phys. **3**, 369 (1977)].
3. Yu. D. Korolev and G. A. Mesyats, *The Physics of Pulsed Gas Breakdown* (Nauka, Moscow, 1991).
4. Yu. P. Raĭzer and S. T. Surzhikov, Preprint No. 304, IPM Akad. Nauk SSSR (Moscow, 1987).

5. G. V. Gadiyak, V. A. Shveĭgert, and O. U. Uuémaa, Izv. Sib. Otd. Akad. Nauk SSSR **21** (6), 41 (1988).
6. Yu. P. Raizer, *Gas Discharge Physics* (Nauka, Moscow, 1987; Springer-Verlag, Berlin, 1991).
7. C. Barbeau and J. Jolly, Appl. Phys. Lett. **58**, 237 (1991).
8. V. V. Osipov, V. V. Savin, and V. A. Tel'nov, Izv. Vyssh. Uchebn. Zaved., Fiz., No. 12, 52 (1976).
9. S. V. Pashkin and P. I. Peretyat'ko, Kvantovaya Élektron. (Moscow) **5**, 1159 (1991).
10. S. A. Smirnov, in *Proceedings of the Conference "Physics and Technique of Plasma," Minsk, 1994*, Part 1, p. 168.
11. A. M. Dykhne and A. P. Napartovich, Dokl. Akad. Nauk SSSR **247**, 837 (1979) [Sov. Phys. Dokl. **24**, 632 (1979)].
12. S. Ya. Bronin, V. M. Kolobov, V. N. Sushkin, *et al.*, Teplofiz. Vys. Temp. **18**, 46 (1980).
13. M. M. Keker, M. R. Barrant, and J. D. Craggs, J. Phys. D **5**, 253 (1972).
14. Yu. S. Akishev, A. P. Napartovich, S. V. Pashkin, *et al.*, Teplofiz. Vys. Temp. **22**, 201 (1984).

Translated by N. Larionova

Technical Physics, Vol. 46, No. 7, 2001, pp. 831–839. Translated from Zhurnal Tekhnicheskoĭ Fiziki, Vol. 71, No. 7, 2001, pp. 44–52.

SOLIDS

Interaction and Pinning of Plane Vortices in a Three-Dimensional Josephson Medium and Possible Distances between Two Isolated Vortices

M. A. Zelikman

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

Received July 12, 2000

Abstract—Within a continuous vortex model, exact expressions are obtained for the Josephson and magnetic energies of plane (laminar) vortices, as well as for the energy and force of pinning by cells in a three-dimensional Josephson medium. If the porosity of the medium is taken into account, the Josephson and magnetic energies of the vortex differ from those for the continuum case. The contributions to the pinning energy from the Josephson and magnetic energies have opposite signs. An algorithm for numerically solving a system of difference equations is proposed in order to find the shape and the energy of the vortex in its stable and unstable states. The continuous vortex model is shown to fail in predicting correct values of the Josephson and magnetic energy of the vortex, as well as of the pinning energy components. Expressions for the least possible distances between two isolated vortices are obtained for a small pinning parameter. Analytical results are in close agreement with computer simulation. An algorithm for numerically solving a system of difference equations is proposed in order to find the least possible distances between two isolated vortices when the pinning parameter I is not small. The minimal value of I at which the center-to-center distance N of the vortices equals three cells is 1.428; for $N = 2$, $I_{min} = 1.947$. At $I > 2.907$, the vortices can be centered in adjacent cells. © *2001 MAIK "Nauka/Interperiodica".*

INTRODUCTION

Recent investigations into high-temperature superconductors (HTSCs) have shown that vortical structures are to a great extent responsible for processes arising in the sample. In particular, vortex pinning and interaction specify the processes occurring in HTSCs in an external magnetic field. The movement of vortices results in the non-Joule release of energy, which adversely affects the superconductivity properties. The critical currents and fields can be increased (the fundamental problem in physics and technology) if the vortex behavior, structure, and interaction, as well as pinning mechanisms and forces, are adequately understood.

The structure of vortices, their pinning, and the dependence of pinning on a magnetic field have been the subject of extensive research [1–8]. In [3, 4], the behavior of one-dimensional vortices in an extended Josephson contact is analyzed. Here, it is assumed that the vortices have a continuous phase distribution and that they are pinned by discretely located pinning centers. Actually, however, a Josephson medium has a cellular structure; therefore, pinning depends on the energy necessary to move the vortex center from one cell into another.

In [5], the vortex behavior in a linear chain of SQUIDs was analyzed for the two-dimensional situation, i.e., when the magnetic field of a given loop is taken into account only in the magnetic flux crossing this loop. In the three-dimensional case, however, a vortex consists of coaxial "solenoids" and the magnetic flux through a loop is induced not only by this loop but also by other currents including remote ones. With a decrease in the critical contact current, the vortex size, i.e., the number of loops producing the magnetic flux through the central vortex cell, increases. This compensates for the reduced contribution into the magnetic flux from each of the loops.

In [6, 7], a system of equations for fluxoid quantization in the cells of a three-dimensional ordered Josephson medium was derived to perform a detailed analysis of the structure of possible vortical states. Different configurations of laminar vortices in a three-dimensional ordered Josephson medium, as well as their possible motion, were considered in [8].

In this paper, we calculate the pinning forces for plane (laminar) vortices and their repelling forces in the cells of a three-dimensional Josephson medium. Also, the least possible distances between two isolated vortices are evaluated.

As in [6–8], we will consider a simplified model comprising a cubic lattice with the lattice constant h. The lattice is made up of superconducting wires with each connection containing a Josephson contact and all the contacts having a small size and the same critical current J_c. The vortex axis coincides with the lattice axis. The current distribution is assumed to be planar; that is, the currents are identically distributed in all par-

1063-7842/01/4607-0831$21.00

allel planes that are perpendicular to the vortex axis and spaced at h. Such a model, while being rather simple, enables certain conclusions to be made about the structure of vortices and their dynamics. Results obtained within this model qualitatively agree with those derived from more complicated models.

Based on the condition for fluxoid quantization in a cell, we derived equations [6–8] that describe the distribution of the phase discontinuity values at the contacts for laminar configurations:

$$\varphi_{m+1} - 2\varphi_m + \varphi_{m-1} = I\sin\varphi_m, \quad (1)$$

where $I \equiv 2\pi\mu_0 h J_c/\Phi_0$ is the dimensionless parameter and Φ_0 is a quantum of the magnetic flux.

An equation similar to (1) was first obtained by Frenkel and Kontorova when considering the dynamics of dislocations in a crystal.

It was shown [6–8] that, with an increase in the parameter I, the vortex size decreases and the pinning force increases. Therefore, with increasing I, the least possible distance between the vortices diminishes. Eventually, at high I, the vortex centers may find themselves in the adjacent cells and the concept of discrete vortices loses its meaning. Let us consider the cases of small and nonsmall values of the parameter I separately.

SMALL VALUES OF THE PARAMETER I

When the parameter I is small, the size of vortices is large and so is the distance between their centers. At $I \ll 1$, the distribution becomes quasi-continuous [7] and system (1) rearranges to the sin-Gordon differential equation

$$\frac{d^2\varphi}{dx^2} = I\sin\varphi, \quad (2)$$

where the length x is measured in terms of h.

The solution of Eq. (2) for an individual vortex has the form

$$\varphi(x) = 4\arctan(\exp(-x\sqrt{I})). \quad (3)$$

The plot of function (3) is presented in Fig. 1.

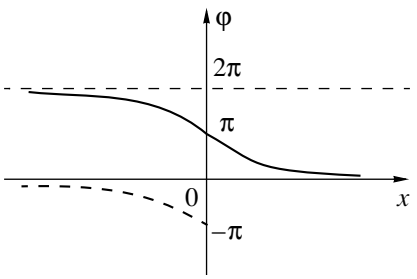


Fig. 1. Solution (3) of Eq. (2) for an isolated vortex.

Since the physically meaningful phase value is defined accurate to the multiple of 2π, 2π can be subtracted from the phase values at $x < 0$ to make the vortex shape symmetric about the coordinate (dashed curve in Fig. 1). This implies the restriction $-\pi < \varphi \le \pi$; in other words, the values of φ_m obtained from Eq. (1) or (2) should be diminished by a multiple of 2π so that φ_m falls into the range specified.

Equation (3) written in the form $d^2\varphi/d\xi^2 = \sin\varphi$, where $\xi = x\sqrt{I}$ is a distance in terms of $h/\sqrt{I}$, has been shown [9] to have the following solutions corresponding to an infinite number of equidistant vortices:

$$x = \frac{1}{2}\int_{\varphi_0}^{\varphi} \frac{d\varphi}{\sqrt{\sin^2\frac{\varphi}{2} + \gamma^2}}, \quad (4)$$

where φ and γ are constants of integration.

The existence of such solutions follows from the fact that Eq. (2) does not include vortex pinning; therefore, the repulsion force acting on a vortex from the right can be balanced only by an equal force acting from the left. That is, its neighbors must be equidistant from it. In the absence of pinning, there is no steady-state solution for two isolated vortices, since mutual repulsion would throw them away to infinity. With pinning, the least possible center-to-center distance is specified by the equality of the repulsion force and the maximal pinning force.

At $I \ll 1$, the repulsion force can be derived from solutions (4) of Eq. (3). To find the pinning force in the cells, the discrete character of the medium should be taken into account.

1. Calculation of the Repulsion Force

For our current configuration of height Z and length Y, the energy is given by [8]

$$E = \varepsilon_0 YZ\sum_{k=-\infty}^{\infty}\left[\frac{1}{2}(\varphi_{k+1} - \varphi_k)^2 + I(1 - \cos\varphi_k)\right], \quad (5)$$

where $\varepsilon_0 = \Phi_0^2/4\pi^2\mu_0 h^3$ is a normalizing constant.

Hereafter, the energies will be expressed in units of ε_0. In the continuous case, expression (5) transforms into the integral

$$E(\varphi(\xi)) = \sqrt{I}YZ\int_{-\infty}^{\infty}\left[\frac{1}{2}(\partial\varphi/\partial\xi)^2 + (1 - \cos\varphi)\right]d\xi. \quad (6)$$

Let us consider a set of equidistant vortices. Substituting (4) into (6) yields the expression for energy per

vortex (per 1 m of height and 1 m of length; $Y = Z = 1$),

$$\frac{E}{\sqrt{I}} = a\left[4(1+\gamma^2)\frac{E(1/\sqrt{1+\gamma^2})}{K(1/\sqrt{1+\gamma^2})} - 2\gamma^2\right] = 8\sqrt{1+\gamma^2}E\left(\frac{1}{\sqrt{1+\gamma^2}}\right) - \frac{4\gamma^2}{\sqrt{1+\gamma^2}}K\left(\frac{1}{\sqrt{1+\gamma^2}}\right), \quad (7)$$

where $K(x)$ and $E(x)$ are the complete elliptic integrals of the first and second kind, respectively, and

$$a = \frac{2}{\sqrt{1+\gamma^2}}K\left(\frac{1}{\sqrt{1+\gamma^2}}\right) \quad (8)$$

is the spatial period (the length per vortex).

The case where the distance between vortices is much greater than the length of a vortex, which is equal to 1 (in dimension units $h/\sqrt{I}$), corresponds to the condition $\gamma \ll 1$. Expanding Eq. (7) in γ and leaving only the first terms, we obtain

$$E/\sqrt{I} = 8 + 2\gamma^2. \quad (9)$$

Using the asymptotic of the function $K(1/\sqrt{1+\gamma^2})$ at small γ, we come to $a = 2\ln(4/\gamma)$; hence, it follows that the energy per unit length and height of an isolated vortex is equal to $E = 8\sqrt{I}$ and the energy per vortex due to vortex interaction is

$$E_{\text{int}} = 2\gamma^2 = 32e^{-a}\sqrt{I}. \quad (10)$$

Since vortices are far apart from each other, each can be assumed to interact only with its left and right nearest neighbors. Thus, expression (10) gives the energy of interaction between two adjacent vortices. Then, the repulsion force of two isolated vortices is given by

$$F = -\frac{\partial E}{\partial r} = -\frac{\sqrt{I}}{h}\frac{\partial E}{\partial a} = 32\frac{I}{h}e^{-a} = 32\frac{I}{h}e^{-N\sqrt{I}}, \quad (11)$$

where N is the number of cells between the centers of the vortices.

Note that expressions (10) and (11) can be derived using a different approach [4].

2. Calculation of Vortex Pinning in Cells

(a) Analytical calculation. Let us obtain the energy of pinning associated with the discrete character of the medium where vortices can be considered as quasi-continuous (i.e., $I \ll 1$). We will use the approach proposed in [10]. In view of the relationship for δ functions

$$\sum_{k=-\infty}^{\infty}\delta(x-k) = \sum_{n=-\infty}^{\infty}\exp(i2\omega nx) = \text{Re}\sum_{n=-\infty}^{\infty}\exp(i2\pi nx), \quad (12)$$

expression (5) at $Y = Z = 1$ takes the form

$$E = \sum_{k=-\infty}^{1} f(\varphi_k) = \int_{-\infty}^{\infty} f(\varphi(x+\alpha))\sum_{k=-\infty}^{\infty}\delta(x-k)dx = \text{Re}\int_{-\infty}^{\infty} f(\varphi(x))\sum_{n=-\infty}^{\infty}\exp(i2\pi n(x-\alpha))dx, \quad (13)$$

where α is the coordinate of the vortex center relative to the cell edge. For instance, if a vortex is centered at the boundary of two cells, $\alpha = 0$; if the centers of a vortex and cell coincide, $\alpha = 0.5$.

If the function $f(x)$ is even, expression (13) takes the form

$$E = 2\sum_{n=-\infty}^{\infty}\cos(2\pi n\alpha) \times \text{Re}\int_0^{\infty} f(\varphi(x))\exp(i2\pi nx)dx. \quad (14)$$

Substituting (3) for a continuous isolated vortex into (14) and leaving the terms with $n = 0$ and ± 1, we obtain

$$E = \frac{2}{\sqrt{I}}\int_0^{\infty} f(4\arctan(\exp(-\xi)))d\xi + \frac{4}{\sqrt{I}}\cos(2\pi\alpha) \times \text{Re}\int_0^{\infty} f(4\arctan(\exp(-\xi)))\exp\left(\frac{i2\pi\xi}{\sqrt{I}}\right)d\xi. \quad (15)$$

(i) Consider the Josephson energy $E = I\sum_{k=-\infty}^{\infty}(1 - \cos\varphi_k)$. In this case,

$$f_1 = I(1 - \cos(4\arctan(e^{-\xi}))) = \frac{2I}{\cosh^2\xi}. \quad (16)$$

Taking into account the evenness of f_1 with respect to ξ, we have [11]

$$E_{J_0} = 4\sqrt{I}\int_0^{\infty}\frac{d\xi}{\cosh^2\xi} = 4\sqrt{I}\lim_{b\to 0}\int_0^{\infty}\frac{\cos(b\xi)d\xi}{\cosh^2\xi} = 4\sqrt{I}\lim_{b\to 0}\frac{b\pi}{2\sinh(b\pi/2)} = 4\sqrt{I}, \quad (17)$$

$$E_{J_1} = 8\sqrt{I}\cos(2\pi\alpha)\int_0^\infty \frac{\cos(2\pi\xi/\sqrt{I})}{\cosh^2\xi} = \frac{8\pi^2}{\sinh(\pi^2/\sqrt{I})} \times \cos(2\pi\alpha) \approx 16\pi^2\exp\left(-\frac{\pi^2}{\sqrt{I}}\right)\cos(2\pi\alpha). \quad (18)$$

The maximal value of E_{J_1} is achieved at $\alpha = 0$, i.e., when a vortex is centered at the boundary of two cells.

(ii) Consider the magnetic energy $E_H = 0.5\sum_{k=-\infty}^{\infty}(\varphi_k - \varphi_{k+1})^2$. In this case,

$$f_2 = 8(\arctan(e^{-(x+\alpha)\sqrt{I}}) - \arctan(e^{-(x+\alpha+1)\sqrt{I}}))^2 = 8\arctan^2\frac{-e^{\alpha\sqrt{I}} - e^{-(\alpha+1)\sqrt{I}}}{e^{-x\sqrt{I}} - e^{-(x+2\alpha+1)\sqrt{I}}}. \quad (19)$$

At $\alpha_1 = -0.5$, the function $f_2(x)$ is even:

$$f_2 = 8\arctan^2\frac{\sinh(\sqrt{I}/2)}{\cosh(x\sqrt{I})}. \quad (20)$$

Expanding arctan into a series, squaring it, and integrating termwise yields

$$E_{H_0} = \frac{8}{\sqrt{I}}\int_{-\infty}^{\infty}\arctan^2\frac{\sinh(\sqrt{I}/2)}{\cosh\xi}d\xi = \frac{16\eta}{\sqrt{I}}\left[1 - \frac{4}{9}\eta + \frac{184}{675}\eta^2 - \frac{704}{3675}\gamma^3 + \ldots\right], \quad (21)$$

where $\eta = \sinh^2(\sqrt{I}/2)$.

Going from the series in η to that in I, we obtain

$$E_{H_0} = 4\sqrt{I} - \frac{1}{9}I^{3/2} + \frac{7}{1350}I^{5/2} + \ldots. \quad (22)$$

To calculate E_{H_1}, we will use methods of the theory of complex variable:

$$E_{H_1} = \frac{32}{\sqrt{I}}\int_{-\infty}^{\infty}\left[\arctan\exp\left(\frac{\sqrt{I}}{2} - \xi\right) - \arctan\exp\left(-\frac{\sqrt{I}}{2} - \xi\right)\right]^2\exp\left(i\frac{2\pi\xi}{\sqrt{I}}\right)d\xi. \quad (23)$$

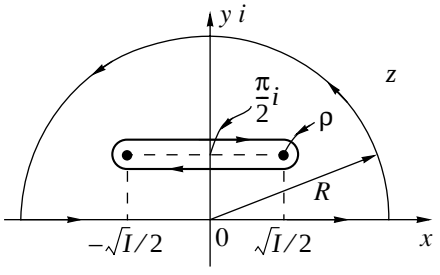


Fig. 2. Contour of integration in the complex plane z.

Writing the function of complex variable arctan$z$ in the form

$$\arctan z = \frac{1}{2i}\mathrm{Ln}\frac{1-iz}{1+iz},$$

we come to

$$E_{H_1} = -\frac{8}{\sqrt{I}}\int_{-\infty}^{\infty}\left[\mathrm{Ln}\frac{1 - i\exp(-z+\sqrt{I}/2)}{1 + i\exp(-z+\sqrt{I}/2)} - \mathrm{Ln}\frac{1 - i\exp(-z-\sqrt{I}/2)}{1 + i\exp(-z-\sqrt{I}/2)}\right]^2\exp\left(i\frac{2\pi z}{\sqrt{I}}\right)dz. \quad (24)$$

The integrand has four singular points:

$$z = \pm\frac{\sqrt{I}}{2} \pm \frac{\pi}{2}i. \quad (25)$$

To obtain E_{H_1}, we take the integral over the contour with a cutout (Fig. 2) and then approach R to infinity and ρ to zero. Formula (24) satisfies the moderate Jordan lemma: the expression in brackets squared steadily approaches zero at $|z| \longrightarrow \infty$ in the sectors $-\varphi_0 \le \arg z \le \varphi_1$ and $\pi - \varphi_2 \le \arg z \le \pi + \varphi_0$. Therefore, the integral over the larger semicircle approaches zero at $R \longrightarrow \infty$. It can easily be shown that the integrals over the smaller circles approach zero at $\rho \longrightarrow 0$. The integrals over the upper and lower sides of the cutout have opposite signs and differ in absolute value, since the function in brackets changes by $2\pi i$ in going around the point $z = (\sqrt{I}/2) + (\pi/2)i$:

$$E_{H_1} = \frac{8}{\sqrt{I}}\exp(-\pi^2/\sqrt{I})\cos(2\pi\alpha) \times \mathrm{Re}\int_{-\sqrt{I}/2}^{\sqrt{I}/2}\left\{\left[\ln\left|\frac{1-\exp(-\xi+\sqrt{I}/2)}{1+\exp(-\xi+\sqrt{I}/2)}\right| - \ln\left|\frac{1-\exp(-\xi-\sqrt{I}/2)}{1+\exp(-\xi-\sqrt{I}/2)}\right|\right]^2 - \left[\left|\frac{1-\exp(-\xi+\sqrt{I}/2)}{1+\exp(-\xi+\sqrt{I}/2)}\right| - \ln\left|\frac{1-\exp(-\xi-\sqrt{I}/2)}{1+\exp(-\xi-\sqrt{I}/2)}\right| - 2\pi i\right]^2\right\}\exp\frac{i2\pi\xi}{\sqrt{I}}d\xi \quad (26)$$

$$= \frac{8}{\sqrt{I}}\exp(-\pi^2/\sqrt{I})\cos(2\pi\alpha)\mathrm{Re}\int_{-\sqrt{I}/2}^{\sqrt{I}/2}\left[\ln\left|\tanh\left(\frac{\xi}{2} - \frac{\sqrt{I}}{4}\right)\right| - \ln\left|\tanh\left(\frac{\xi}{2} + \frac{\sqrt{I}}{4}\right)\right|\right]2\pi i\exp\frac{i2\pi\xi}{\sqrt{I}}d\xi = -\frac{16\pi}{\sqrt{I}} \times \exp(-\pi^2/\sqrt{I})\cos(2\pi\alpha)\int_{-\sqrt{I}/2}^{\sqrt{I}/2}\left[\ln\left|\tanh\left(\frac{\xi}{2} - \frac{\sqrt{I}}{4}\right)\right|\right.$$

$$- \ln\left|\tanh\left(\frac{\xi}{2} + \frac{\sqrt{I}}{4}\right)\right|\Bigg]\sin\frac{2\pi\xi}{\sqrt{I}}d\xi.$$

Expanding the integrand into a series and leaving the first two terms, we arrive at the following expression for the integral [11]:

$$\sqrt{I}\int_{-\sqrt{I}/2}^{\sqrt{I}/2}\left[\ln\left|\frac{\xi - 0.5}{\xi + 0.5}\right|\sin 2\pi\xi + \frac{I}{6}\xi\sin 2\pi\xi\right]d\xi$$

$$= -2\sqrt{I}\int_0^1 \ln\xi\sin 2\pi\xi d\xi + \frac{I\sqrt{I}}{12\pi} = \frac{\sqrt{I}}{\pi}(C + \ln 2\pi \quad (27)$$

$$- \mathrm{Ci}(2\pi) + I/12) = \frac{\sqrt{I}}{\pi}(-S_1(2\pi) + I/12),$$

where C is the Euler constant, Ci(x) is the integral cosine, and

$$S_1(2\pi) = \int_0^{2\pi}\frac{1 - \cos x}{x}dx = 2.43766\ldots.$$

E_{H_1} is maximum at $\alpha = 0$; however, from the initial shift [$\alpha_1 = -0.5$, see expression (20)], it follows that this maximum corresponds to the coincidence of the vortex and cell centers. Taking into account the initial shift for reducing the expression to the single origin for α, we finally obtain

$$E_{H_1} = -16\exp(-\pi^2/\sqrt{I}) \times \cos(2\pi\alpha)(2.43766 - I/12), \quad (28)$$

where $\alpha = 0.5$ corresponds to the position of the vortex center at the center of the cell.

The total vortex energy is

$$E = E_{J_0} + E_{H_0} + (E_{J_1} + E_{H_1})\cos 2\pi\alpha. \quad (29)$$

As in [8], we define the pinning energy E_p as the difference between the maximal and the minimal vortex energies at the different α, i.e., when the vortex center is located at the boundary between two cells ($\alpha = 0$) and at the cell center ($\alpha = 0.5$):

$$E_p = E(0) - E(0.5) = \Delta E_J + \Delta E_H. \quad (30)$$

According to the above analysis, ΔE_J and ΔE_H are given by

$$\Delta E_J = 32\pi^2\exp\left(-\frac{\pi^2}{\sqrt{I}}\right), \quad (31)$$

$$\Delta E_H = -32\exp\left(-\frac{\pi^2}{\sqrt{I}}\right)(2.43766 - I/12). \quad (32)$$

To check the validity of the formulas obtained, we performed a rigorous computer evaluation of the values

$$E_J(\alpha) = I\sum_{k=-\infty}^{\infty}(1 - \cos\varphi_k),$$
$$E_H(\alpha) = \frac{1}{2}\sum_{k=-\infty}^{\infty}(\varphi_{k+1} - \varphi_k)^2. \quad (33)$$

Here, φ_k was calculated with expression (3) for a vortex in a continuous medium:

$$\varphi_k(\alpha) = 4\arctan(\exp(-k - \alpha)\sqrt{I}) \quad (34)$$

for $\alpha = 0$ and $\alpha = 0.5$. The differences between the associated values give ΔE_J and ΔE_H in (30), whereas their half-sums are equal to E_{J0} and E_{H0}, according to (29).

The energies E_{J_0}, E_{H_0}, ΔE_J, and ΔE_H calculated using formulas (17), (22), (31), and (32), respectively, for several values of I are summarized in Tables 1 and 2.

(b) Computer calculation of pinning energy. The foregoing consideration was based on function (34), which corresponds to the solution (3) of differential equation (2), approximating system of difference equations (1) at small I. To find the applicability limit for such an approach, the results obtained should be compared with the exact solution of system (1).

To obtain a numerical solution of system (1), we write it in the form of the recurrent relation

$$\varphi_{m+1} = 2\varphi_m - \varphi_{m-1} + I\sin\varphi_m. \quad (35)$$

Then, knowing the values of the phase discontinuities at two adjacent contacts, one can find the whole configuration. Let us consider an isolated vortex. In the stable equilibrium state [8], the phase discontinuities at the contacts of the central cell are related as $\varphi_{-1} = 2\pi - \varphi_1$.

Table 1. Magnetic and Josephson energies obtained by different approaches

I	Theory (17)	Theory (22)	Numerical model (33), (34)		Exact solution	
	E_{J_0}	E_{H_0}	E_{H_0}	E_{J_0}	E_H	E_J
0.15	1.5492	1.5428	1.5428	1.5492	1.5525	1.5393
0.40	2.5298	2.5022	2.5022	2.5298	2.5452	2.486
0.70	3.347	3.2836	3.2837	3.347	3.3853	3.2393
1.00	4.000	3.894	3.894	4.000	4.073	3.806

Table 2. Pinning energy components obtained by different approaches

I	Theory (31)	Theory (32)	Numerical model (33), (34)		Exact solution			Theory (31), (32)
	ΔE_J	ΔE_H	ΔE_H	ΔE_J	ΔE_H	ΔE_J	E_p	E_p
0.15	2.705*E*-9	–6.6477*E*-10	–6.6479*E*-10	2.705*E*-9	–6.095*E*-8	6.617*E*-8	5.22*E*-9	2.04*E*-9
0.40	5.27*E*-5	–1.2849*E*-5	–1.2852*E*-5	5.27*E*-5	–6.21*E*-4	7.14*E*-4	9.29*E*-5	3.99*E*-5
0.70	2.38*E*-3	–5.734*E*-4	–5.738*E*-4	2.38*E*-3	–1.86*E*-2	2.24*E*-2	3.90*E*-3	1.81*E*-3
1.00	1.63*E*-2	–3.897*E*-3	–3.902*E*-3	1.63*E*-2	–0.095	0.121	0.026	0.0124

Choosing a certain value of φ_1 and substituting $2\pi - \varphi_1$ and φ_1 into (35) as two successive values of φ_k, we start calculating the subsequent values of the discontinuities. From the vortex shape (Fig. 1) it follows that, if any of the next values is less than zero or larger than the preceding one, the starting value of φ_1 is increased or decreased, respectively, and the calculation starts at the beginning. In this way, the whole set of φ_k values for an isolated vortex can be obtained, after which the Josephson and magnetic energies are calculated from formula (5).

A similar calculation can also be carried out for the unstable vortex configuration. In this case, $\varphi_{-1} = \pi$ [8]; therefore, the values of π and φ_1 are taken as two successive values and then the scheme of computation is the same as before.

The computer realization of the algorithms enables the calculation of the Josephson, magnetic, and pinning energies. The results of calculation are presented in Tables 1 and 2.

(c) Analysis of the results. The pinning energies and forces for one-dimensional vortices in a long Josephson contact, which are also described by system (1), were calculated in [4] using data in [10]. The results obtained in this chapter are distinct from those reported in [4, 10].

A comparison of the results based on the analytical expressions with those of the computer simulation leads to the following conclusions.

(1) The exact numerical calculation of the parameters $E_{J_0} = 0.5[E_J(0) + E_J(0.5)]$, $E_{H_0} = 0.5[E_H(0) + E_H(0.5)]$, $\Delta E_J = E_J(0) - E_J(0.5)$, and $\Delta E_H = E_H(0) - E_H(0.5)$ from (33) and (34) (columns 4, 5 in Tables 1, 2) supports (with a high accuracy) the validity of formulas (17), (22), (31), and (32) (columns 2, 3 in Tables 1, 2).

(2) Formulas (17) and (31) for Josephson energies coincide with those in [4, 10] for one-dimensional vortices in a long Josephson contact. This situation is also described by system (1).

(3) Formulas (22) and (32) for magnetic energies differ from the relationships [4, 10]

$$E_H = 4\sqrt{I} \quad \text{and} \quad \Delta E_H = 32\pi^2 \exp\left(-\frac{\pi^2}{\sqrt{I}}\right).$$

While in the formula for E_H, we added small corrections, in the case of ΔE_H, the energies have opposite signs and their absolute values differ approximately by a factor of 4. These differences stem from the fact that, when calculating the magnetic energy E_H, the author of [10] baselessly used the function $f_2 = 0.5(\varphi'^2) = I(1 - \cos\varphi) = f_1$ in (15) instead of function (19). That is why his values for the Josephson energy E_J and the magnetic energy E_H turned out the same.

A comparison of the energy values obtained from the exact numerical solution of the system of the difference equations (columns 6, 7 in Tables 1, 2) with those derived by the approximate approach [solution (3) of differential equation (2)] leads to the following conclusions.

(1) The associated values for the magnetic and the Josephson energies (columns 4, 6 and 5, 7 in Table 1) differ by small yet physically significant amounts.

(2) The exact values of the Josephson, ΔE_J, and magnetic, ΔE_H, components of the pinning energy (columns 5, 7 in Table 2) are several tens and several hundreds of times larger than the values obtained by the approximate approach (columns 4, 6 in Table 2).

(3) Yet, due to the different signs of the Josephson and magnetic components of the pinning energy, the exact values of the pinning energy $E_p = \Delta E_J + \Delta E_H$ (column 8, Table 2) differ insignificantly from the approximate results (column 9, Table 2). The satisfactory agreement between the approximate and exact values of the pinning energy when its Josephson and magnetic components drastically diverge is most probably not accidental and is of fundamental nature. The substantiation of the approximate approach for finding the pinning energy is a separate mathematical problem that is as yet unclear how to solve.

From Table 2, the exact expression for the pinning energy can be written in the form

$$E_p = k64\pi^2 \exp\left(-\frac{\pi^2}{\sqrt{I}}\right), \tag{36}$$

where the coefficient k is in the range from 0.8 to 1 depending on the value of the parameter I.

Note that, in [4], the pinning energy is given by expression (18), i.e., is approximately twice as low as the value from (36).

3. Least Possible Distances between Two Isolated Vortices

To obtain the pinning force of a vortex, we differentiate vortex energy (29) with respect to coordinate:

$$F_p = \frac{\partial E}{\partial r} = \frac{2\pi}{h}(E_{J_1} - E_{H_1})\sin(2\pi\alpha). \quad (37)$$

The least possible center-to-center distance $N_{\min}$ of two isolated vortices is specified by the equality of repulsion force (11) and the maximal pinning force:

$$32\frac{I}{h}e^{-N_{\min}\sqrt{I}} = \frac{2\pi}{h}k32\pi^2\exp\left(-\frac{\pi^2}{\sqrt{I}}\right), \quad (38)$$

hence,

$$N_{\min} = -\frac{1}{\sqrt{I}}\ln\left(\frac{2k\pi^3}{I}\exp\left(-\frac{\pi^2}{\sqrt{I}}\right)\right). \quad (39)$$

In a similar formula for $N_{\min}$ [4], the coefficient $2k$ is lacking. At small I, this may result in a noticeable quantitative discrepancy, as follows from Table 3. The first column of Table 3 lists $N_{\min}$ obtained in [4]; the second column, the values calculated by formula (39); and the third one, the minimal distances $d_{\min}$ between the extreme vortex and its neighbor that were obtained by exactly solving (1) when an external magnetic field [12] is monotonically increased.

When comparing data in Table 3, one should bear in mind that the exact value of $N_{\min}$ can either be equal to $d_{\min}$ or greater than $d_{\min}$ by 1. The reason is the following. At the least possible distance $N_{\min}$ between two isolated vortices, each of them is in the extreme position in a cell. Now suppose that there are other vortices, for instance, on the right of the right vortex, that allow it to move more to the right in the cell. If the left-hand vortex remains in the same cell in this case, $d_{\min} = N_{\min}$; if it passes to the adjacent cell, $d_{\min} = N_{\min} - 1$.

It is evident from Table 3 that, with regard to the aforesaid, the calculations using (39) yield values almost completely coincident with the numerical solution of system (1), unlike the calculations using the formulas in [4]. This agreement confirms the validity of both the theoretical analysis [expression (39)] and the algorithm for the numerical solution of system (1) to obtain $d_{\min}$.

NONSMALL VALUES OF THE PARAMETER *I*

Consider now nonsmall values of I at which a vortex covers several cells. In this case, the transition from the difference equations to the differential equations is impossible. Therefore, the above results are inapplicable in such a case and discrete system (1) should be studied. Let us find possible distances between two isolated plane (laminar) vortices.

Table 3. Least possible distances between vortices obtained by different approaches

I	$N_{\min}$ [4]	$N_{\min}$ from [39]	$N_{\min}$ [12]
0.07	117.9	115.3	115
0.08	102.3	99.8	100
0.10	80.5	78.4	78
0.15	52.0	50.2	50
0.20	38.1	36.5	37
0.30	24.4	23.2	23
0.40	17.8	16.7	17
0.50	13.9	12.9	13
0.60	11.4	10.5	11
0.70	9.6	8.7	9
1.00	6.4	5.7	6

We will analyze system (1) for two vortices remote from the boundary with their centers placed N cells apart in row 1 and row $(N + 1)$ (Fig. 3). To the left of cell 1 and to the right of cell $(N + 1)$, the currents decrease, approaching zero at infinity. At $I > 1$, the values of φ_m outside the central cells of the vortices are small, system (1) is linearized, and its solution takes the form [12]

$$\varphi_m = \varphi_{-1}\gamma^{-m-1} \quad (m \le -1), \quad (40)$$

$$\varphi_m = C_1\gamma^{-m-1} + C_{2\gamma}^{N-m} \quad (1 \le m \le N), \quad (41)$$

$$\varphi_m = \varphi_{N+1}\gamma^{-m-N-1} \quad (m \ge N + 1), \quad (42)$$

where $\gamma = 1 + I/2 - \sqrt{I + I^2/4}$ is the solution of the equation $\gamma^2 - (1 + I)\gamma + 1 = 0$.

The distribution of the currents and the phases will be assumed to be symmetric about the middle of the

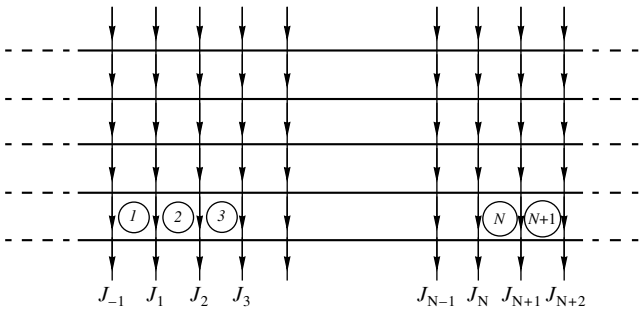


Fig. 3. Current distribution in a plane perpendicular to the axes of two isolated interacting vortices.

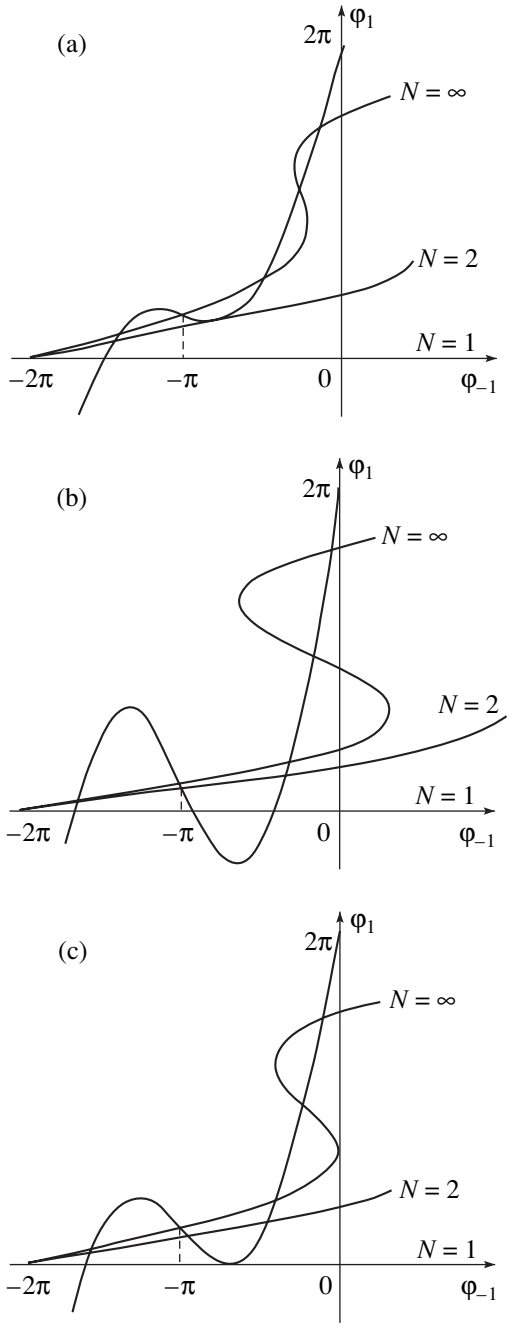


Fig. 4. Graphical solution of system (46)–(47) for different I. Curves emerging from the point $\varphi_1 = 2\pi$ correspond to Eq. (47); those emerging from the point $\varphi_{-1} = -2\pi$ correspond to Eq. (46). I = (a) 2, (b) 4, and (c) 2.9.

distance between the vortices. Then,

$$C_1 = -C_2 = \varphi_1/(1-\gamma^{-N-1}),$$
$$\varphi_2 = \varphi_1 k,$$
$$k = \gamma(1-\gamma^{N-3})/(1-\gamma^{N-1}). \quad (43)$$

Substituting (43) into the boundary conditions for the central cell of the vortex

$$I\sin\varphi_1 = \varphi_{-1} - 2\varphi_1 + \varphi_2 + 2\pi, \quad (44)$$

$$I\sin\varphi_{-1} = \varphi_1 - 2\varphi_{-1} + \varphi_{-2} - 2\pi, \quad (45)$$

we obtain the following system of equations for φ_{-1} and φ_1:

$$\varphi_{-1} = I\sin\varphi_1 + (2-k)\varphi_1 - 2\pi, \quad (46)$$

$$\varphi_1 = I\sin\varphi_1 + (2-\gamma)\varphi_1 + 2\pi. \quad (47)$$

In (46) and (47), only the coefficient k depends on the distance between vortices: $k(\infty) = \gamma$, $k(2) = -1$, $k(3) = 0$, etc. In Fig. 4, dependences (46) and (47) are presented for $I = 2$, 2.9, and 4. It is evident that, at each value of I, system (46)–(47) has solutions in a certain range of N; in other words, the intervortical distance may vary from some minimal value to infinity. This is a result of pinning, since the vortices would move apart to infinity in its absence; that is, there would be no solutions at finite N. Having found N, we can determine the shortest distance at which the pinning forces balance the repulsion of the vortices. It is evident from Fig. 4 that, at $I = 2$, $N_{\min} = 2$; when $I = 4$, the vortices can be centered even in the adjacent cells.

Let us find the minimal I at which the vortices are centered in the adjacent cells. In this case, $N = 1$ and $k = \infty$; then, as follows from (46), φ_1 should be equal to zero (because of the symmetry of the problem). The curve in Fig. 4c with its minimum lying on the φ_{-1} axis corresponds to the lowest I satisfying (46) and (47) at $\varphi_1 = 0$. The exact numerical calculation yields $I_{\min} = 2.907$ at $N = 1$. The minimal value of I when the center-to-center distance N can be equal to 2 is 1.947; for $N = 3$, we obtain $I_{\min} = 1.428$.

CONCLUSION

Within a continuous vortex model, the interaction energies and pinning forces of plane (laminar) vortices in a three-dimensional Josephson medium were calculated.

It is shown that expressions for the vortex magnetic energy and magnetic component of the pinning energy that were derived by other authors in the framework of this model at low values of the pinning parameter are incorrect. The refined expressions for these values were deduced. The values of the Josephson and magnetic energies of the vortices that are found in view of the porosity of the medium diverge unlike the continual case. The contributions to the pinning energy from the Josephson and magnetic energies have opposite signs. The numerical calculation of all the energies within this model showed the validity of the expressions obtained.

An algorithm for numerically solving the system of difference equations was proposed to obtain the shape and the energy of the vortex in its stable and unstable states without using the continuous vortex model.

Computations using this algorithm demonstrated that the continuous vortex model yields improper values of the Josephson and magnetic energy of the vortex, as well as of the pinning energy components. The exact values of these components are tens and hundreds of times larger than those obtained with the continuous vortex model. In spite of this, the exact value of the pinning energy differs from the model results only slightly because of different signs of the Josephson and magnetic components of the pinning energy. This fact most likely is not accidental and is, perhaps, of fundamental nature. The substantiation of using the continuous vortex model for evaluating the pinning energy is a separate mathematical problem. An approach to solve it remains to be elaborated.

From the energies of interaction and pinning calculated, the least possible distances between two isolated vortices at small values of the pinning parameter were evaluated. These results agree well with those of the computer simulation.

An algorithm for numerically solving the system of difference equations in order to find the least possible distances between two isolated vortices at a nonsmall pinning parameter I is proposed. The least value of I for which the center-to-center distance N of the vortices equals three cells is 1.428; for $N = 2$, $I_{min} = 1.947$. At $I > 2.907$, the vortices may be centered in the adjacent cells.

ACKNOWLEDGMENTS

This study was supported by the grant from the Ministry of Science and Technology (“Superconductivity” program).

REFERENCES

1. K.-H. Muller, J. C. MacFarlane, and R. Driver, Physica C (Amsterdam) **158**, 69 (1989).
2. M. S. Rzchowski, S. P. Benz, M. Tinkham, and C. J. Lobb, Phys. Rev. B **42**, 2041 (1990).
3. Y. S. Kivshar and B. A. Malomed, Rev. Mod. Phys. **61**, 763 (1989).
4. V. V. Bryksin, A. V. Gol’tsev, S. N. Dorogovtsev, *et al.*, Zh. Éksp. Teor. Fiz. **100**, 1281 (1991) [Sov. Phys. JETP **73**, 708 (1991)].
5. R. Parodi and R. Vaccarone, Physica C (Amsterdam) **173**, 56 (1991).
6. M. A. Zelikman, Supercond. Sci. Technol. **10**, 469 (1997).
7. M. A. Zelikman, Supercond. Sci. Technol. **10**, 795 (1997).
8. M. A. Zelikman, Supercond. Sci. Technol. **12**, 1 (1999).
9. I. O. Kulik and I. K. Yanson, in *Josephson Effect in Superconducting Tunnel Structures* (Nauka, Moscow, 1970), p. 270.
10. S. Aubry, Lect. Notes Math., No. 925, 240 (1982).
11. I. S. Gradshteĭn and I. M. Ryzhik, *Table of Integrals, Sums, Series, and Products* (Nauka, Moscow, 1963; Academic, New York, 1980).
12. M. A. Zelikman, Zh. Tekh. Fiz. **67** (9), 38 (1997) [Tech. Phys. **42**, 1019 (1997)].

Translated by M. Lebedev

Technical Physics, Vol. 46, No. 7, 2001, pp. 840–843. Translated from Zhurnal Tekhnicheskoĭ Fiziki, Vol. 71, No. 7, 2001, pp. 53–56.

SOLIDS

On One Mechanism of Mechanoelectrical Transformations in Impact-Excited Composites Based on the Cement Binder

T. V. Fursa

Tomsk Polytechnical University, pr. Lenina 30, Tomsk, 634034 Russia

Received August 3, 2000

Abstract—A mechanism of mechanoelectrical transformations produced by impact excitation in cement-binder-based composites is suggested and experimentally verified. The transformations proceed in two stages: material electrification at the point of impact and mechanoelectrical processes at the matrix–binder interface. *© 2001 MAIK "Nauka/Interperiodica".*

Mechanoelectrical transformations in insulating materials subjected to impact excitation are fundamental for nondestructively testing the mechanical strength of structural materials [1]. Studies of sources and mechanisms of these transformations in insulators are aimed at improving the precision and extending the functionality of the nondestructive technique. Experimental and theoretical data for the sources in concretes have been summarized in [2–4]. In this work, we develop a mechanism of the transformations as applied to impact-excited composites consisting of the cement binder and a filler.

We performed vast experimentation using physical simulation. Model materials were cement-based composites with different inclusions.

Experiments were carried out as follows. The sample surface was dynamically struck, and an acoustic and an electromagnetic signal were simultaneously recorded with an Emission-1 two-channel device (Fig. 1). The signal detector, the electromechanical percussion device, and the sample were always arranged in the same way. The exciting mechanical pulse was about 3×10^{-5} s long, and the velocity of the striker was roughly 2 m/s.

Figures 2a and 2b show typical electromagnetic pulses (curves *1*) and acoustic signals (curves *2*) that were recorded in samples of various compositions. Figure 2a refers to the sample composed of cement, sand, and water. Figure 2b refers to the sample of the same composition but having a single metallic inclusion. It is seen that the electromagnetic signals diverge, while the acoustic signals are virtually coincident. For the cement–sand sample, which is a single-component heterogeneous composite, an individual electromagnetic pulse is recorded. For the two-component system (having the metallic inclusion), the electromagnetic response is much more complicated.

The single electromagnetic pulse excited by the impact excitation of the cement stone can be associated with the triboelectric effect, arising when the striker comes into contact with the sample surface (when coming into contact, rubbing, and subsequently separating, materials are known to be electrified by virtue of the triboelectric effect [6]). The impact-induced electrification of materials has been studied in [7]. That this signal is related to the electrification of the sample at the instant the striker touches its surface is corroborated by Fig. 2a. The time delay t between the electromagnetic and acoustic signals results from the different velocities of the electromagnetic and acoustic waves. The process of mechanoelectrical transformations can be thought of as follows. When the striker touches the sample surface, the rubbing materials are electrified and an electrical signal arises. This signal is instantaneously recorded by the electronic detector. After a time, a surface or bulk (depending on the position of the acoustic detector and the sample geometry) elastic wave comes to the acoustic detector and an acoustic signal is recorded. Let us compare the experimental and calculated time delays for the case in Fig. 2a. If the velocity of the surface wave in a cement stone is set equal to ≈2000 m/s and the distance between the point of impact and the acoustic detector is ≈5 cm, the time of arrival of the acoustic wave is ≈25 μs. This value agrees well with the calculated value t = 29 μs. Thus, the electromag-

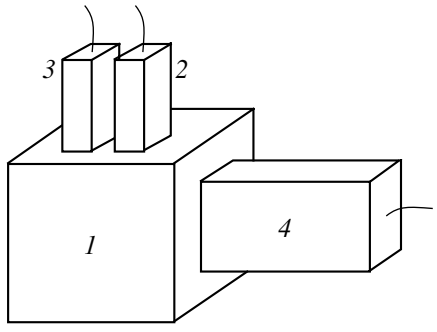


Fig. 1. Experimental scheme: (*1*) sample; (*2*) striker; (*3*) acoustic detector; and (*4*) electronic detector.

1063-7842/01/4607-0840$21.00 © 2001 MAIK "Nauka/Interperiodica"

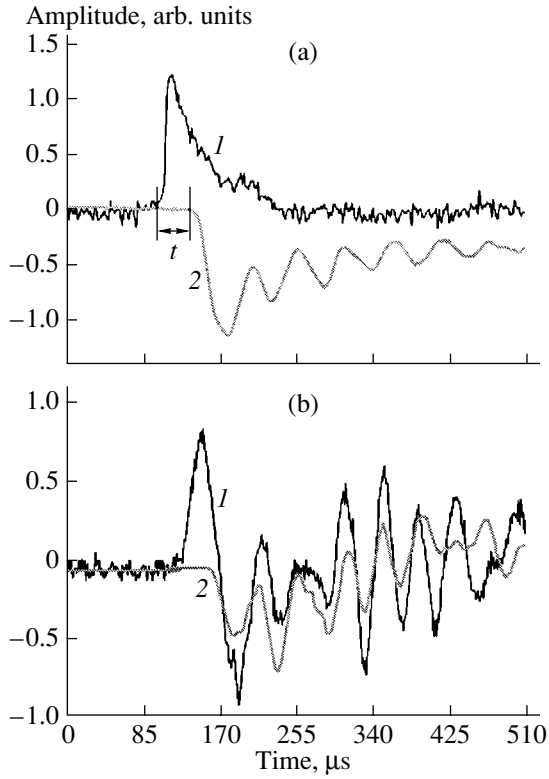


Fig. 2. (*1*) Typical electromagnetic pulses and (*2*) acoustic signals recorded upon the impact excitation of the model samples.

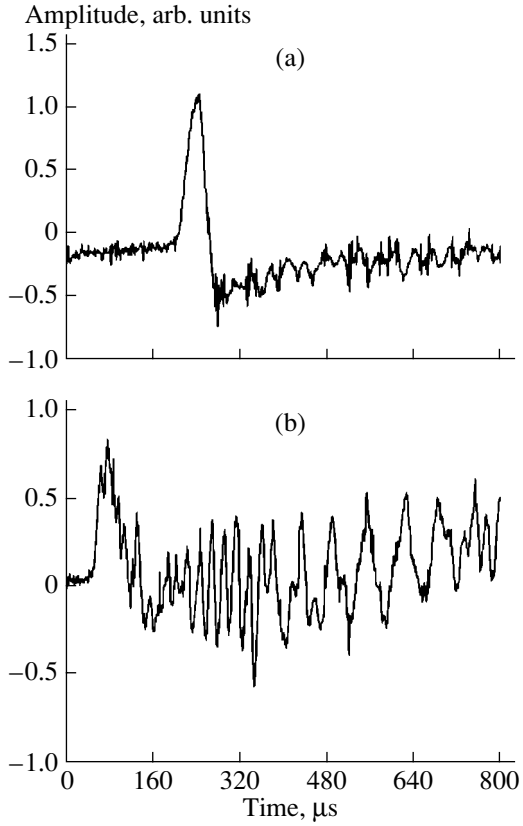


Fig. 3. Typical electromagnetic responses from (a) fine-grain gravel pellets and (b) coarse-grain quartziferous gravel pellets.

netic signal, which is recorded before the acoustic signal, does appear at the instant of impact. The virtual absence of oscillation in the electromagnetic response implies that our impact excitation system suppresses interactions with obstacles whose linear sizes are less than the excitation wavelength [8].

Earlier [1], the impact excitation of concrete, which is a two-component heterogeneous material, was shown to generate an alternating-sign signal. It was speculated that this oscillating electromagnetic response is associated with a change in the dipole moment of the double electrical layer at the boundary of the matrix. The layer is formed by an acoustic wave passing through the matrix. To validate this supposition, we observed the variation of the electromagnetic response shape during the solidification of model concretes. They represented cement–sand mixtures into which a gravel pellet was incorporated. Before the pellets were embedded in the mixture, they had been specially processed and tested. To exclude the effect of the surface condition and scaling factor on the electromagnetic response, the pellets were equally shaped ($80 \times 80 \times 7$ mm) and their surfaces were polished. Of these pellets, we chose those providing impact-induced electromagnetic responses of different types (Fig. 3). The signals of the first type (Fig. 3a) are excited in the fine-grain pellets; the coarse-grain pellets show signals like those in Fig. 3b. Note that the efficiency of the mechanoelectrical transformations grows if gravel is made of a quartziferous rock.

Then, we embedded the pellets in the cement–sand composition, formed samples $100 \times 50 \times 50$ mm in size and tested them during solidification. Typical electromagnetic surfaces recorded 5, 13, 19, and 30 days after sample preparation are depicted in Fig. 4. At the initial time instant (Fig. 4a), an isolated peak is recorded. Subsequently, however, an alternating-sign response appears (Figs. 4b, 4c). Experience suggests that concrete solidifies in about 28 days. The single peak at the initial stage of solidification is possibly associated with the fact that the electrical signal due to mechanical transformations at the matrix–inclusion interface is screened by the conductive wet cement slurry. During solidification, free water combines with the cement to form the cement stone structure and also vaporizes. Therefore, the conductivity decreases. The alternating-sign component of the electromagnetic response even appears on the 13th day of solidification (Fig. 4b). Its amplitude grows up to the 30th day and then remains constant. When refined, these results may be used for

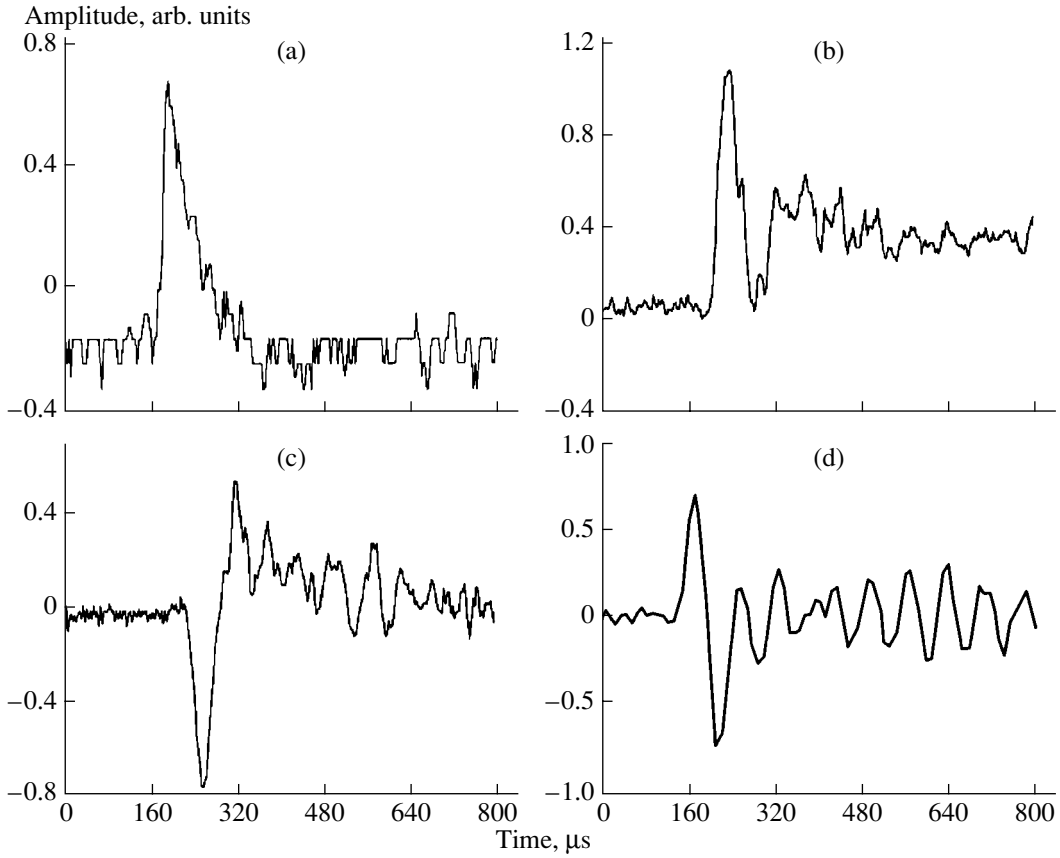


Fig. 4. Typical electromagnetic responses from the model concrete samples at the (a) 5th, (b) 13th, (c) 19th, and (d) 30th day of solidification.

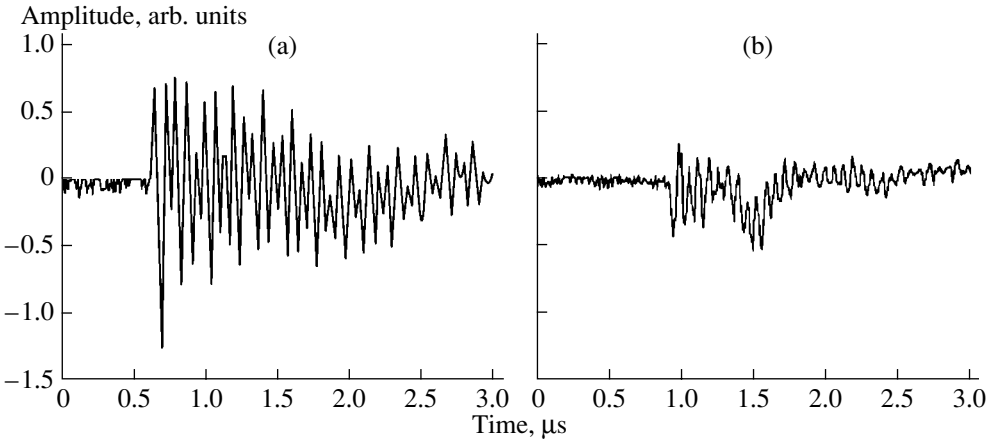


Fig. 5. Electromagnetic responses from the concrete samples with the (a) smooth and (b) rough inclusions.

determining the degree of solidification of concretes. During solidification, the shape of the electromagnetic responses does not depend on whether the inclusion gives a response like that in Fig. 3a or in Fig. 3b. In both cases, the alternating-sign response arises. Therefore, the occurrence of the alternating-sign signal at a certain stage of solidification is related to the acoustic excitation of the double electrical layer at the matrix–inclusion interface when the cement stone solidifies.

From the theory of adhesion bonds in composites consisting of a binder and a filler, the adhesive contact quality depends on the surface condition of the binder [9]. Specifically, if the filler surface is developed, the contact quality improves. If we assume that the adhesive contact quality affects the efficiency of the mechanoelectrical transformations, then the electromagnetic response amplitude must vary according to the surface condition of the inclusion. To check this assumption,

we made model samples composed of the cement matrix and isolated metallic inclusions of equal size. The surfaces of the inclusions were smoothed or roughened. Figure 5 shows typical electromagnetic responses from samples of both types. For the inclusions with the rough surface, the amplitude of the response drops as compared with the smooth inclusions. It seems that, in the former case, the effective surface area of the double electrical layer, which is concurrently excited by the acoustic wave, decreases. This is additional evidence that the alternating-sign electromagnetic response is directly defined by the adhesive contact between the cement matrix and the inclusion.

From the results presented, we suggest the following mechanism of the impact-induced mechanoelectrical transformations in the composites. At the instant of impact, the surface at the point of impact is electrified and an isolated electromagnetic pulse appears. Then, an elastic wave originates in the material. This wave changes the dipole moment of the double electrical layer at the matrix–filler interface, causing an alternating-sign decaying response to arise.

REFERENCES

1. V. L. Chakhlov, Yu. P. Malyshkov, V. F. Gordeev, *et al.*, Izv. Vyssh. Uchebn. Zaved., Stroit., Nos. 5-6, 54 (1995).
2. Yu. P. Malyshkov, T. V. Fursa, V. F. Gordeev, *et al.*, Izv. Vyssh. Uchebn. Zaved., Stroit., No. 12, 31 (1996).
3. T. V. Fursa, V. V. Lasukov, Yu. P. Malyshkov, *et al.*, Izv. Vyssh. Uchebn. Zaved., Stroit., No. 10, 127 (1997).
4. T. V. Fursa, N. N. Khorsov, and E. A. Baturin, Zh. Tekh. Fiz. **69** (10), 51 (1999) [Tech. Phys. **44**, 1175 (1999)].
5. V. F. Gordeev, V. P. Eliseev, Yu. P. Malyshkov, *et al.*, Defektoskopiya, No. 4, 48 (1994).
6. Yu. K. Bivin, V. V. Viktorov, Yu. V. Kulinich, and A. S. Chersin, Izv. Akad. Nauk SSSR, Mekh. Tverd. Tela **17** (1), 183 (1982).
7. Yu. A. Gromov, A. V. Kargapol'tsev, and B. V. Chakhlov, Izv. Vyssh. Uchebn. Zaved., Fiz., No. 4, 120 (1992).
8. T. V. Fursa and V. F. Gordeev, Pis'ma Zh. Tekh. Fiz. **26** (3), 30 (2000) [Tech. Phys. Lett. **26**, 105 (2000)].
9. I. N. Akhverdov, *Foundations of Physics of Concrete* (Stroĭizdat, Moscow, 1981).

Translated by V. Isaakyan

Technical Physics, Vol. 46, No. 7, 2001, pp. 844–846. Translated from Zhurnal Tekhnicheskoĭ Fiziki, Vol. 71, No. 7, 2001, pp. 57–59.

SOLIDS

Creep Duration Analysis in Terms of the Field Theory of Defects

N. V. Chertova and Yu. V. Grinyaev

Institute of Strength Physics and Materials Science, Siberian Division, Russian Academy of Sciences, Akademicheskiĭ pr. 2/1, Tomsk, 634055 Russia

e-mail: chertova@ispms.tsc.ru

Received October 9, 2000

Abstract—From the equations of motion within the field theory of defects, creep curves are derived and a relationship between the applied stress and the time to rupture under different deformation conditions is obtained. The creep duration as a function of the applied stress and the initial strain rate, as well as the ultimate strain, specifying the material rupture, are found. © *2001 MAIK "Nauka/Interperiodica".*

The operation of equipment under severe conditions (high stress and temperature) has resulted in the discovery of the creep effect and the development of creep theory. Engineers have centered on creep analysis [1], i.e., on the evaluation of the time period within which the strain reaches the ultimate value. This problem remains of practical importance nowadays. At its incipient stage, creep theory was developed as an engineering science. Later, it evolved into a branch of continuum mechanics. Simultaneously, physical mechanisms responsible for the creeping effect were studied. Because of their complexity, comprehensive physical description of creep is lacking. Some progress in creep physics has been achieved by invoking the concepts of the dislocation theory [2–4]. A number of dislocation models describing different creep stages and conditions have been constructed. It is argued [3, 4] that, at moderate temperatures, elementary creep events in solids are attributed primarily to the motion of dislocations. Therefore, creep mechanisms will be considered in terms of the field theory, which involves the dynamics of translational defects [5, 6]. Note that the field theory of defects deals with defect ensembles and, according to [7], describes a system on the mesoscopic scale. In contrast, the classical dislocation theory [2] has to do with individual defects and their interactions and thus, implies microscopic methods of description [7].

The dynamic equations in the field theory of defects have the form

$$\begin{aligned} B(\nabla \cdot I) &= -B(\alpha^{\times}.I) - \rho\mathbf{V}, \\ \nabla \cdot \alpha &= 0, \quad \nabla \times I = \frac{\partial \alpha}{\partial t}, \\ S(\nabla \times \alpha) &= -B\frac{\partial I}{\partial t} - S\left(\alpha \cdot \alpha - \frac{\delta}{2}\alpha^2\right) \\ &\quad - B\left(I \cdot I - \frac{\delta}{2}I^2\right) - \eta I - \sigma^{\text{app}}. \end{aligned} \tag{1}$$

From these equations, the equation relating the tensor of the dislocation flux density I to the applied stress σ^{app} was obtained [8]:

$$B\frac{\partial I}{\partial t} + B\left(I \cdot I - \frac{\delta}{2}I^2\right) + \eta I + \sigma^{\text{app}} = 0. \tag{2}$$

Here, α is the dislocation density tensor, $\mathbf{V}$ is the elastic displacement rate, ρ is the material density, η is the viscosity, B and S are the theoretical constants, and δ is the Kronecker delta. The symbols $(\times)$ and $(\cdot)$ stand for the vector and scalar products, respectively, and $(^{\times}.)$ designates the vector product with respect for the first subscripts of the dyad and the scalar product with respect for the second ones. Equation (2), which is helpful in studying the creep process, is written for the uniform distribution of defects, i.e., for space-independent field strengths α and I. It is believed [9] that this assumption is valid near the yield point, where defects are distributed randomly and do not form spatial structures. A great body of experimental data on the creep effect has been obtained from tensile tests of rods. Therefore, in our previous study [8], we considered uniaxial deformation, for which Eq. (2), written in the dimensionless variables $v = -(B/\eta)I_{11}$, $\tau = (\eta/B)t$, and $S = (B/\eta^2)\sigma_{11}$, becomes

$$\frac{dv}{d\tau} = v^2/2 - v + S, \tag{3}$$

where v is the plastic strain rate.

In [8], special attention has been given to the analysis of the functions $v(\tau)$, which specify the creep curves $\varepsilon(t) = \int_0^t v(\tau)d\tau$ under constant stress. In the following, we will carefully investigate the relationship

1063-7842/01/4607-0844$21.00

between the applied stress and the time to rupture of a system.

It has been found [8] that there are two creep modes depending on the applied stress S: stable at $S < 1/2$ and unstable at $S > 1/2$. The corresponding expressions for creep rates are

$$v(\tau) = \frac{p - q[(v_0 - p)/(v_0 - q)]\exp[(p-q)\tau/2]}{1 - [(v_0 - p)/(v_0 - q)]\exp[(p-q)\tau/2]} \quad (4)$$

at $S < 1/2$,

$$v(\tau) = \frac{n + \alpha^2}{n} + \frac{(n^2 + \alpha^2)\cos(\alpha\tau/2)}{n(\cos(\alpha\tau/2) - (n/\alpha)\sin(\alpha\tau/2))} \quad (5)$$

at $S > 1/2$.

Then, the creep curves are given by

$$\varepsilon(\tau) = \varepsilon_0 + p\tau + 2\ln|(p-q)/[(p - v_0) - (q - v_0)\exp[(p-q)\tau/2]]|, \quad (6)$$

$$\varepsilon(\tau) = \varepsilon_0 + \tau - 2\ln|\cos(\alpha\tau/2) - (n/\alpha)\sin(\alpha\tau/2)|, \quad (7)$$

where $2\alpha = p - q$, $n = 1 - v_0$, v_0 is the initial creep rate, and $p, q = 1 \pm \sqrt{1 - 2S}$ are the steady-state rates given by (3) at $dv/d\tau = 0$.

The critical applied stress $S^* = 1/2$ corresponds to the bifurcation point where the creep behavior changes. With this parameter introduced, one can define the stable creep limit $\sigma^* = \eta^2/2B$, which depends on the material parameters.

Expressions (4)–(7) yield the condition

$$\cos(\alpha\tau/2) - (n/\alpha)\sin(\alpha\tau/2) = 0, \quad (8)$$

from which the creep life of a real system at $S > S^*$ is obtained.

Under condition (8), creep rate (5) becomes infinite and the time to rupture is given by

$$t_1^* = (2/\alpha)(\pi - \arctan[\alpha/n]). \quad (9)$$

Curves plotted in Fig. 1 relate the time to rupture of the system to the applied stress at $S > S^*$ and $v_0 =$ (*1*) 0.5 and (*2*) 0.9.

According to [8], when $S < S^*$, solutions of Eq. (3) are strongly dependent on the initial value of v_0. The range of v_0 can be divided into the intervals $0 < v_0 < q$, $q < v_0 < p$, and $v_0 > p$. At $v_0 > p$, the creep rate becomes infinite when the denominator of (4) vanishes; hence, the time to rupture is

$$t_2^* = (2/(p-q))\ln[(v_0 - q)/(v_0 - p)]. \quad (10)$$

Expression (10) is illustrated in Fig. 1 by curves *3* and *4* for $v_0 = 1.6$ and 1.9, respectively. If the initial creep rate lies in the range $q < v_0 < p$, $v(\tau) \longrightarrow q$ at $\tau \longrightarrow \infty$; i.e., the stationary creep mode is observed. In this situation, creep duration analysis turns to the conventional problem (see the first paragraph). If the critical strain ε^* at which the material fails is known, Eq. (6) with $\tau \longrightarrow \infty$ yields the relation of interest between the stress and the time to rupture:

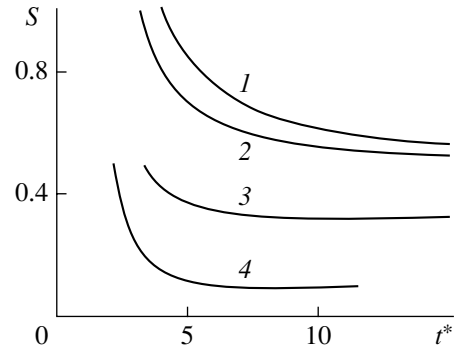


Fig. 1.

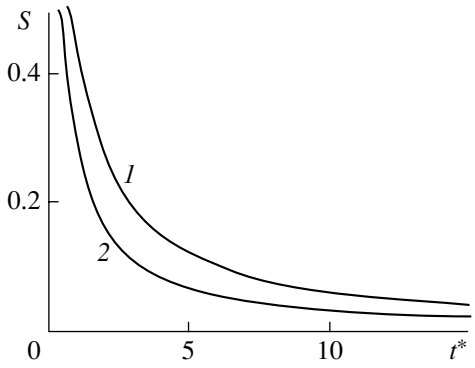


Fig. 2.

$$t_3^* = (\varepsilon_* - \varepsilon_0)/q. \quad (11)$$

This dependence is demonstrated in Fig. 2 for $\varepsilon^* =$ (*1*) 0.65 and (*2*) 0.35. In both cases, the initial strain ε_0 is taken to be equal to 0.01%. This value is in qualitative agreement with Figs. 1 and 3. The latter depicts experimental data [10] for a number of materials tested at various temperatures. The test duration was up to 100 days. Curves *1* and *4* were obtained for carbon steel (0.5% C and 0.24% C) at temperatures of 300 and 432°C, respectively; curve *2*, for nickel steel at 400°C; curves *3* and *5*, for high-alloy nickel–chromium steel at 600 and 700°C, respectively; *7* and *11*, for high-speed steel at 593 and 732°C, respectively; *6*, for cast heat-resistant steel at 800°C; *8*, for lead at room temperature; *9*, for nickel–copper alloy at 600°C; and (*10*), for magnesium alloy at 150°C. Following the conclusions made in [8], we leave without consideration the initial range $0 < v_0 < q$, where attempts to describe the early portion of the creep curve with decreasing rate have failed.

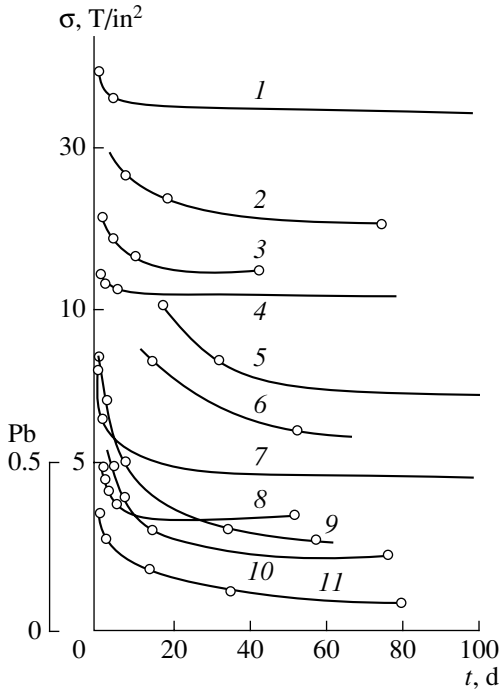


Fig. 3.

In conclusion, it should be noted that a serious effort has been made to find an empirical correlation between the applied stress and the time to rupture during the creep process. Logarithmic, exponential, and hyperbolic expressions have been suggested [10]. However, no universal law that adequately relates these two values under various conditions has been found [10]. We also failed to establish a unified relationship, since creep modes depend on external factors. It is valid to say however, that an increase in the applied stress S at a constant initial creep rate v_0 or a constant critical strain ε^* cuts the time to rupture whatever the creep condition. At a constant stress, the creep life decreases with growing v_0 (Fig. 1) and decreasing ε^* (Fig. 2).

REFERENCES

1. Yu. N. Rabotnov, *Mechanics of Deformable Solid* (Nauka, Moscow, 1979).
2. A. H. Cottrell, *Dislocations and Plastic Flow in Crystals* (Clarendon, Oxford, 1953; Metallurgizdat, Moscow, 1958).
3. J. Friedel, *Dislocations* (Pergamon, Oxford, 1964; Mir, Moscow, 1967).
4. V. M. Rozenberg, *Creep of Metals* (Metallurgiya, Moscow, 1967).
5. Yu. V. Grinyaev and V. E. Panin, Dokl. Akad. Nauk **353**, 37 (1997) [Phys. Dokl. **42**, 108 (1997)].
6. Yu. V. Grinyaev, N. V. Chertova, and V. E. Panin, Zh. Tekh. Fiz. **68** (9), 134 (1998) [Tech. Phys. **43**, 1128 (1998)].
7. H. Haken: *Advanced Synergetics: Instability Hierarchies of Self-Organizing Systems and Devices* (Springer-Verlag, New York, 1983; Mir, Moscow, 1985).
8. Yu. V. Grinyaev and N. V. Chertova, Pis'ma Zh. Tekh. Fiz. **26** (16), 57 (2000) [Tech. Phys. Lett. **26**, 723 (2000)].
9. N. A. Koneva and É. V. Kozlov, Russian Phys. J., No. 2, 165 (1990).
10. A. H. Sully, *Metallic Creep and Heat Resistant Alloys* (Butterworths, London, 1949; Oborongiz, Moscow, 1953).

Translated by A. Sidorova-Biryukova

Technical Physics, Vol. 46, No. 7, 2001, pp. 847–852. Translated from Zhurnal Tekhnicheskoĭ Fiziki, Vol. 71, No. 7, 2001, pp. 60–66.

SOLID-STATE ELECTRONICS

Water After-Etching of *n*-Type Porous Silicon in an Electric Field

B. M. Kostishko and Yu. S. Nagornov

Ul'yanovsk State University, Ul'yanovsk, 432700 Russia

e-mail: Kost@sv.uven.ru

Received October 11, 2000

Abstract—Water after-etching of porous silicon in an external electric field is studied. The application of the electric field, irrespective of its strength and orientation, is shown to decrease the etch rate. When the field vector is directed from the porous layer inward to the sample, the electrode potential in the silicon–electrolyte system sharply changes; for $E > 6$ kV/cm, the changes become periodic. Experimental data suggest the presence of a circulating current in the single-crystal silicon–electrolyte–quantum wire system. © *2001 MAIK "Nauka/Interperiodica".*

INTRODUCTION

Today, porous silicon (*por*-Si) is attracting considerable attention in a number of microelectronics areas. Initially, it was studied as a promising insulator that is compatible with IC technology [1]. Since 1990, when *por*-Si photoluminescence in the visible range was first discovered by Canham [2], the luminescent properties of the material have been the subject of interest. In addition, efforts to elucidate the role of quantum-size effects in the luminescence mechanism have been undertaken.

In the last decade, interest in self-organization phenomena in semiconductors has been aroused because of the development of molecular electronics and nanoelectronics technologies. Self-organization is also observed in *por*-Si. For example, under certain conditions for *por*-Si etching, shunting currents cause ring-shaped and radial structures to form [3]. The water after-etching (WAE) of *por*-Si, that is, the ordered growth of pores when the sample is dipped into an aqueous medium, can also be considered as a self-organization effect [4]. Possibly because the growth of the porous layer during WAE is difficult to control, the WAE mechanism has not been adequately studied to date and the available theoretical models have not been checked experimentally [5, 6].

For a porous layer to form during the anodic etching of silicon, several requirements must be met. The current through the sample should be as high as several to several tens of mA/cm^2, the etchant should contain HF molecules, and external illumination must be used to generate holes in the etching area when *n*-Si is etched. However, during WAE, where the above conditions are not fulfilled, the porous layer does form, the pore growth rate during anodic etching with the current density 5–10 mA/cm^2 and during WAE being the same on the order of magnitude.

The authors of [4] offered theoretical explanations of the WAE processes. They proceed from the assumption that residual HF ions are present only at the mouths of pores and voids on the single-crystal silicon side, thus producing an electric field necessary for etching. Another and, in our opinion, much more intriguing assumption is that there exist circulating currents inside *por*-Si. These assumptions are critically overviewed in this article. It is believed that the external electric field changes the charge currents in the *por*-Si layer and, thereby, the WAE rate.

Thus, this paper is an attempt to give experimental and theoretical justifications of WAE when porous *n*-Si is subjected to external electric fields of various strengths and directions.

EXPERIMENT

In experiments, we used single-crystalline (*c*-Si) phosphorus-doped (100)Si wafers of resistivity 2.4 Ω cm. *por*-Si layers were formed by conventional 40-min electrochemical etching in a Hf : C_2H_5OH = 1 : 1 (volume fractions) electrolyte at a current density of 15 mA/cm^2 under halogen lamp illumination.

The WAE rate in an electric field was measured with a plane capacitor connected to a high-voltage power supply. Immediately after the etching, *por*-Si layers were cut into four fragments. Fragment IV (Fig. 1) (test fragment) was kept in air, fragment II was immersed in distilled water in the absence of the electric field, and fragments I and III were placed between the plates of the capacitor placed into a thin water-filled cuvette. The upper plate was made opaque in order to exclude the optical generation of minority carriers in fragments III and I, and the fragments were arranged so that the electric field in them was oppositely directed (in and out of

1063-7842/01/4607-0847$21.00 © 2001 MAIK "Nauka/Interperiodica"

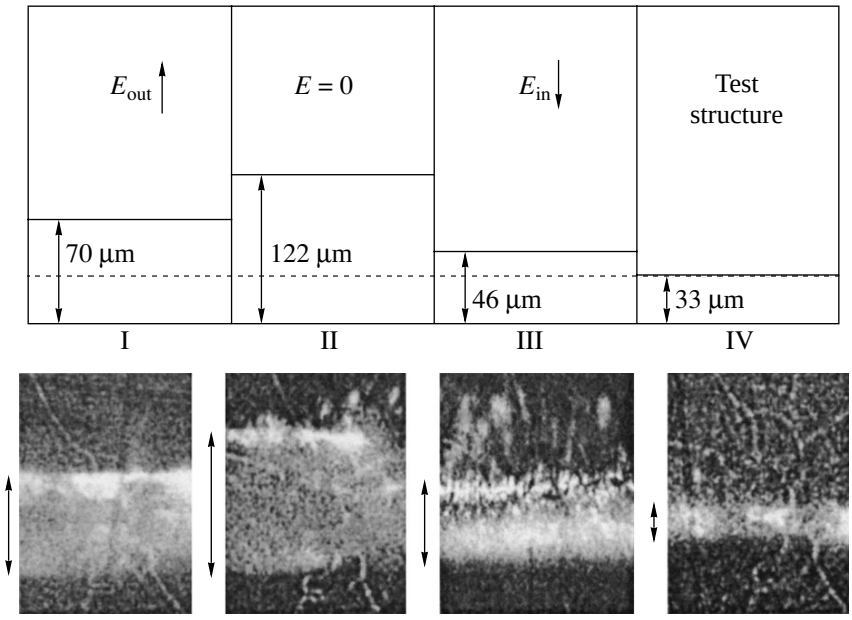


Fig. 1. Water after-etching of porous silicon: sample size and micrographs of the porous layer. I, E_{out}; II, $E = 0$; III, E_{in}; and IV, test structure kept in air. The field strength $E = 5$ kV/cm, the WAE time is 5 h. Arrows in the micrographs indicate the porous layer boundaries.

the layer, respectively; the respective fields are designated E_{in} and E_{out}).

The field strength was 10 kV/cm. WAE lasted from 4 to 8 h. The samples were then dried, and the thicknesses of the porous layers on each of the fragments were measured with an MII-4 interferometer.

The electrode potential U_{el} was determined as follows. An electrode paraffin-insulated from the water was attached to the porous part of the sample with a GaIn paste. The lead reference electrode was immersed immediately in the water. Both were connected to a voltmeter whose readings were applied to an automated reading unit and then to a computer.

Photoluminescence (PL) spectra were taken with a DFS-52 spectrometer at room temperature. The spectra were excited by a DKSSh-150 lamp at a wavelength of 360 nm. The radiation power density was no more than $P = 70$ W/cm^2 to prevent thermal effects on the sample surface.

The chemical composition of the surfaces was determined with a 09IOS-10-005 Auger spectrometer equipped with an electronic analyzer of resolution $\Delta E/E = 0.4\%$ at the residual pressure in the analytic chamber $p = 10^{-7}$ Pa, primary electron energy 3 keV, electron beam current $j = 0.5$ μA, and modulation voltage $U_m = 2$ V.

Ion–plasma etching was performed in a VUP-5 vacuum setup. The pressure of both argon and residual oxygen was maintained at 2.6×10^{-2} Pa, and the plasma discharge voltage and current were 2.5 kV and 25 mA, respectively. During the etching process, the samples were heated by no more than 5°C.

RESULTS AND DISCUSSION

Figure 1 compares the thicknesses of the porous layers obtained after the WAE in the electric field (the thicknesses characterize the WAE rate) with that of the initial *por*-Si. It is seen that the electric field, whatever its direction, decreases the WAE rate, but the thickness (hence, the WAE rate) depends on its direction.

Since WAE is, in essence, the electrochemical reaction of silicon dissolution that continues in water, its rate is proportional to the current of positive charge carriers at the pore mouths that passes through the silicon–electrolyte interface [1]. Therefore, the decrease in the thickness of the after-etched layer for both field directions may be due to a change in the circulating currents, which were introduced in [4]. On the other hand, when the external electric field reaches some critical value, it can modify the type of the chemical reactions in the silicon–electrolyte system [7]. To gain a better insight into mechanisms responsible for the formation of the *por*-Si during WAE, we studied field dependences of the *por*-Si electrode potential.

It turned out that the electrode potential U_{el} remains nearly unchanged during WAE in the absence of the field and when it is directed out of the porous layer ($E_{out} = 1$–10 kV/cm). If, however, the field is reversed (E_{in}) and gradually varies, the potential first jumps (Fig. 2) and then falls (within the time period $\tau \approx 20$ s) down to its initial value. A further increase in the field

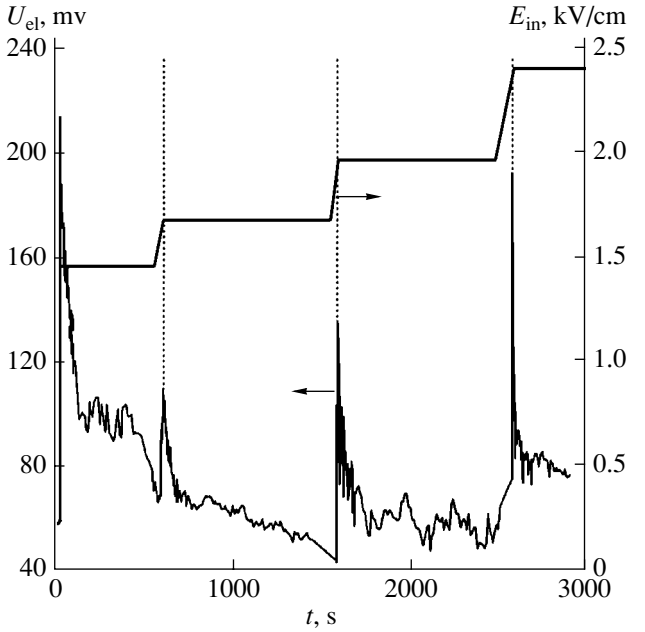


Fig. 2. Variation of the electrode potential of *por*-Si during WAE in the external electric field $E_{in} < 2.5$ kV/cm.

causes another drastic change in U_{el}. When E_{in} exceeds 6 kV/cm, U_{el} starts oscillating. For $E_{in} = 10$ kV/cm, the oscillation period is $T = 16$–17 s (Figs. 3a, 3b).

Note specific features of the periodic behavior of the electrode potential at large field strengths. Both the rise and the fall times of U_{el} do not exceed 1 s (portions *K*–*L* and *L*–*M* in Fig. 3b). After the peaks, U_{el} first falls below its initial level and then ($\tau \approx 4$ s) attains it (*M*–*N* in Fig. 3b). The amplitudes of the jumps are roughly the same (Fig. 3b). The slight discrepancy between the peaks is possibly due to discrete readings of the fast-varying potential U_{el}.

In measuring the WAE rate, the field strength (E = 5 kV/cm) was taken in such a way as to provide the constancy of U_{el}. From [4], two built-in fields may coexist in the porous layer immediately after electrochemical etching. The former (near the pore mouths) is produced by holes on the *c*-Si side and F^- ions of the residual electrolyte. Our previous data suggest that WAE is not observed on samples that were stored in air for a long time or thermally annealed in a vacuum. This finding supports the presence of an electric field at the pore mouths that disappears with charge neutralization.

The latter built-in field arises in quantum wires during *por*-Si formation and is due to sample overetching in the electrolyte. The overetching is accompanied by a gradual decrease in the silicon grain size. This process continues until carrier-depleted regions fully overlap [1]. The closer the grains to the surface, the greater their overetching time and, hence, the smaller their size. Near the *por*-Si surface, the grain size is less than 120 nm. In this system, the bandgap E_g nonuniformly

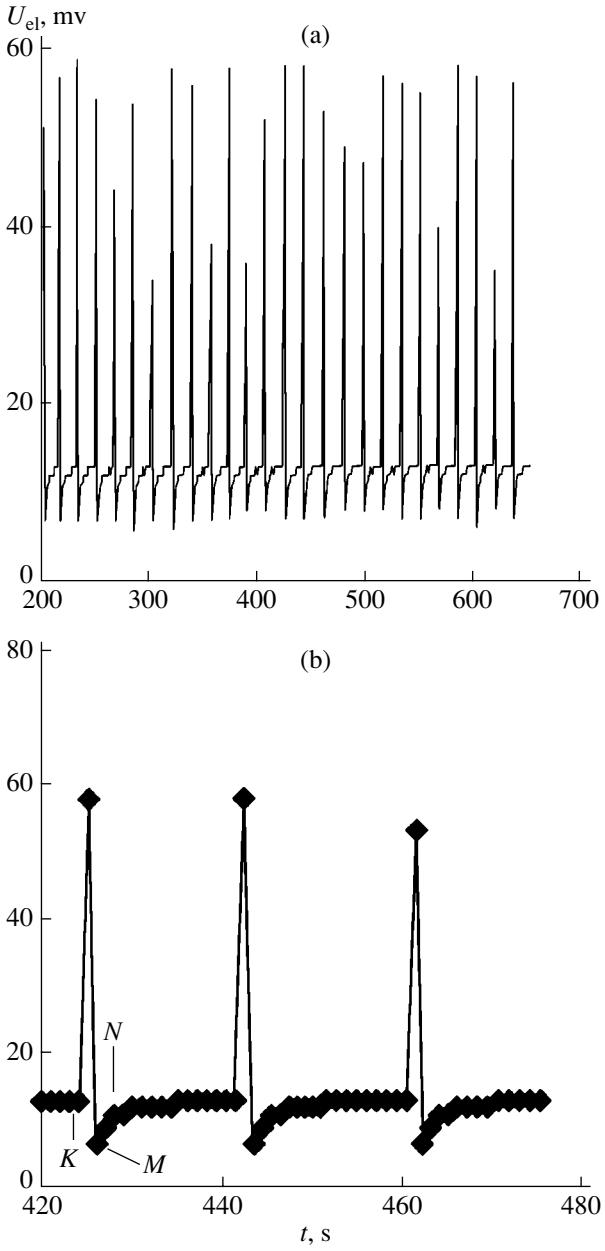


Fig. 3. (a) The same as in Fig. 2 for $E_{in} = 10$ kV/cm and (b) a fragment of panel "a" on an expanded scale.

widens due to the quantum-size effect. The value of E_g smoothly varies with distance from the surface and approaches the value typical of single-crystal silicon (1.1 eV) near the *c*-Si region [4]. In such a variband system, an electric field built in the quantum wires appears. This field accelerates positive charge carriers and makes them move from the wires to the bulk of the semiconductor (Fig. 4, region *C*). Of importance here is the fact that the *por*-Si layers are of the *n*-type; hence, the space charge region (SCR) is formed by mobile minority carriers, i.e., holes.

Figure 5 demonstrates Auger spectra for the initial porous silicon (curve *1*), for the sample after-etched

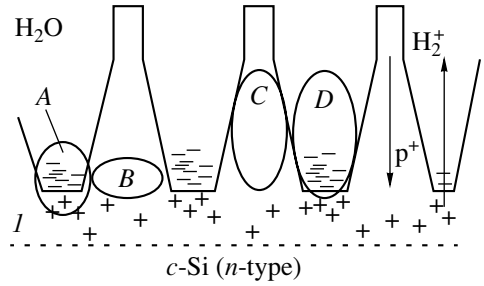


Fig. 4. WAE model. *A*, pore mouth; *B*, quantum wire base; *C*, hole diffusion in quantum wire; *D*, pore interior; and *1*, SCR.

(curve *2*), and for the sample subjected to ion–plasma etching in the argon–oxygen atmosphere for 10 min. The WAE oxidizes the quantum wires, as follows from the shift of the *LVV* peak from $E = 91$ eV to $E = 75$ eV (the latter value is typical of oxidized silicon). Moreover, the peak of the PL spectrum shifts toward shorter wavelengths approximately by 0.3 eV (Fig. 6) after the WAE. This can be explained by a decrease in the quantum wire size because of the formation of the oxide sheath [8]. In this situation, the mean grain size decreases by 1.0 nm [9].

In the Auger spectra, the fluorine peak for the surface of the initial porous layer is absent (Fig. 5, curve *1*); consequently, the surface concentration of fluorine atoms does not exceed 0.05 at. %. The presence of F inside the pores was determined by ion–plasma etching. Investigations show that quantum wires are oxidized and simultaneously fluorinated during ion–plasma etching in argon–oxygen mixtures (Fig. 5, curve *3*). The sensitivity factors for oxidized silicon Si(ox), oxygen, and carbon that were obtained by calibration [10], as well as reference data for the sensitivity factor for fluorine [11], allows the determination of the fluorine–oxidized silicon concentration ratio: $n_F/n_{Si(ox)} = 0.6$. We believe that the fluorination of the *por*-Si surface during ion–plasma processing is the result of the following reactions: the oxidation of HF molecules by oxygen radicals present in the plasma, the formation of dangling bonds on the pore surface due to Ar^+ ion irradiation, and the attachment of fluorine radicals to these bonds. The overall reaction proceeds until the fluorine source in the *por*-Si is exhausted. In our case, the reaction lasted 10–13 min; then, the fluorine concentration on the surface gradually decreased, as evidenced by the Auger data. Thus, the results obtained indicate that the amount of fluorine near the pore mouths is sufficient to maintain the reaction of silicon dissolution during WAE.

The chemical reaction II of silicon dissolution, which is involved in the equation

$$\begin{array}{l} \xrightarrow{\text{(I)}}\ 2H^+ \longrightarrow H_2^{+}\uparrow + h^+ \\ Si + 2HF + 2h^+ \xrightarrow{\text{(II)}} SiF_2 + 2H^+ \\ SiF_2 + 2H_2O \xrightarrow{\text{(III)}} SiO_2 + 2HF + H_2 \\ SiO_2 + 6HF \xrightarrow{\text{(IV)}} H_2SiF_6 + H_2O \end{array} \qquad (1)$$

has been known for a long time [1]. However, during WAE, it can proceed only at the pore mouths in region *A* (Fig. 4) and only in the presence of the electric field produced by F^- ions and holes in the *c*-Si.

According to (1), WAE produced silicon hypofluoride, as a result of which the amount of residual fluorine decreases. In addition, the positive charge of H_2^+ ions is removed (Fig. 4, region *D*), the hole concentration in the SCR of the *c*-Si decreases, and the dissolution reaction eventually ceases. Therefore, a necessary condition for the WAE reaction is a permanent replenishment of the SCR by HF and holes. For this, reaction III and circulating current I in Eq. (1) must exist.

In this context, the effect of the external electric field on the circulating current in the *por*-Si–electrolyte system seems to be the most significant. This current results from several processes. First, H_2^+ molecules produced by reaction II of Eq. (1) rise to the water surface (region *D* in Fig. 4). Simultaneously, the positive charge progressively leaves the H_2^+ molecules for the quantum wires due to tunneling and then is entrained by the built-in field inward to the *c*-Si (region *C* in Fig. 4). Note also that the inside of the quantum wires in *por*-Si of medium porosity is depleted by majority carriers [12]; hence, recombination processes in these regions can be neglected. Thus, the positive charge currents from the pore mouths to the sample surface and in the opposite direction practically balance, and the quasi-equilibrium WAE process is sustained.

In this model, holes thermally generated in the SCR and those due to the reverse saturation current in the porous silicon–electrolyte system are not taken into

account. Indeed, the flux of thermally generated holes can be estimated by the formula

$$j_{gen} = (n_i x_d)/\tau_{gen}, \quad (2)$$

where n_i is the intrinsic carrier concentration in Si, x_d is the SCR width, and τ_{gen} is the characteristic time of carrier generation in the SCR.

For $x_d \approx 1.6\mu m$ [13] and $\tau_{gen} \approx 10^{-5}$ s [14, 15], we have $j_{gen} \approx 1.6 \times 10^{11}$ cm^{-2} s^{-1}. Let us compare this value with the density of the hole current passing through the pore mouths during WAE without the electric field. If we suppose that there are two holes for a dissolved silicon atom [7], the experimentally observed WAE rate 5 nm/s (for the porosity $\eta = 52\%$) is provided by the following hole current from the SCR:

$$j_p = (2\eta\rho_{Si}h)/(tm^a_{Si}), \quad (3)$$

where h is the depth of the layer after-etched, t is the WAE time, ρ_{Si} is the density of *c*-Si, and m^a_{Si} is the atomic mass of Si.

Relationship (3) estimates the hole flux at $j_p \approx 9 \times 10^{15}$ cm^{-2} s^{-1}. Obviously, the current due to thermal generation is negligible in *por*-Si.

It is known [1] that the WAE rate is specified by the current through the Si–electrolyte interface:

$$I = \text{const} n^2_{HF} p_0^2 \exp(\beta V_H/V_T)\exp(V_S/V_T), \quad (4)$$

where n_{HF} is the concentration of HF molecules in the electrolyte, p_0 is the hole concentration in the *c*-Si, β is the symmetry factor, V_H is the Helmholtz potential, $V_T = kT/q$ is the thermal potential, and V_S is the potential at the SCR–electrolyte interface.

The Helmholtz potential is not a governing factor for the current variation. Therefore, the WAE rate depends largely on V_S or the hole concentration in the SCR [1]. With the E_{out} field applied (the external electrode of the high-voltage power supply is under a negative potential), the structure shown in Fig. 4 can be considered as reversed-biased in terms of the Schottky diode model. In this case, the field built in the quantum wires (region *C* in Fig. 4) and, hence, the hole flux in the SCR will diminish. As follows from the above estimates, the thermal generation of holes cannot replenish the SCR by mobile positive-charge carriers, which participate in reaction II of Eq. (1). As a result, the application of the E_{out} field decreases the positive charge concentration in the SCR; according to (4), the WAE rate drops in this case, which is observed experimentally.

For the oppositely directed field (E_{in}), the Schottky diode becomes forward-biased. This also slows down the WAE process because the SCR narrows and the total concentration of holes in region *A* declines (Fig. 4). At the same time, the hole flux inward to the sample must grow in the field thus directed. Eventually,

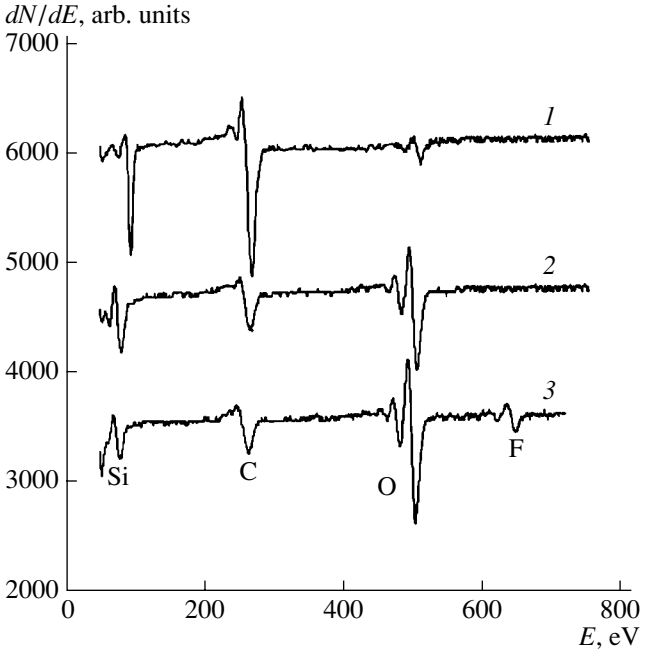


Fig. 5. Auger spectra.

a potential well for holes in region *B* will form (Fig. 4) and the positive charge will be redistributed between the SCR in the *c*-Si and the base of a quantum wire. The accumulation of the positive charge directly in the wire generates a potential across the pore wall and the electrolyte. If this potential (field) exceeds some critical value, the electrode potential U_{el} exhibits a peak (Fig. 3a) because the two-electron reaction (1) of silicon dissolution changes to either of the following four-electron reactions [1, 7]:

$$Si + 4HF + 4h^+ \longrightarrow SiF_4 + 4H^+, \quad (5)$$

$$Si + 4H_2O + 4h^+ \longrightarrow Si(OH)_4 + 4H^+. \quad (6)$$

A four-electron reaction proceeds only on the pore walls, which means that electrolytic polishing conditions set in the Si–electrolyte system and that the interface potential drops to a larger extent. The rate of charge removal doubles; hence, its concentration in the quantum wires and the field strength diminish (region *B* in Fig. 4). Eventually, the four-electron reaction is terminated, and the electrode potential decreases (Fig. 3b, portion *LM*). However, since the hole current into the *c*-Si continues, region *B* will accumulate the critical value of the positive charge again and another peak in the electrode potential curve will appear. In other words, the process will become periodic. Clearly, the period between the U_{el} peaks depends on the field E_{in}. As the field rises, the SCR shrinks, while the sweeping field in the quantum wires grows. Consequently, the charge in region B (Fig. 4) will be accumulated faster and the rate of U_{el} peak appearance will increase.

If the field E_{in} is low (<2.5 kV/cm), the electrode potential exhibits sparse, rather than periodic, peaks (Fig. 2). In this case, the charge accumulated in region *B* (Fig. 4) is likely to exceed the critical value locally and four-electron reaction (5) or (6) proceeds only at these local sites.

The application of the field E_{out} results in neither individual nor periodic peaks of the electrode potential. This is an additional corroboration of our model. In fact, with this external field, the SCR widens and the built-in field of the quantum wires and the hole current into the *c*-Si decrease. Under such conditions, the positive charge cannot be accumulated in the wires and four-electron silicon dissolution is impossible under any circumstances. Thus, with the external field E_{out}, the electrode potential U_{el} in the *por*-Si–electrolyte system will remain unchanged.

CONCLUSIONS

We studied processes taking place during WAE in an external electric field. As follows from the Auger spectroscopy data, the WAE process oxidizes the quantum wires and shifts the PL peaks toward the blue range by approximately 0.3 eV. The application of an electric field markedly affects the WAE rate irrespective of its direction and strength. Moreover, the external field directed from the quantum wires to the *c*-Si can change the type of the silicon dissolution reaction at the pore mouths, as indicated by the electrode potential curves for the *por*-Si. For $E < 2.5$ kV/cm, the curves show individual peaks due to a gradual increase in the field. For $E > 6$ kV/cm, the potential begins to vary periodically, the period being inversely proportional to the field. For the maximum value of the field attained in the experiment, $E = 10$ kV/cm, the period of electrode potential oscillation is 16–17 s. For the oppositely directed field (from the *c*-Si to the quantum wires), the electrode potential remains almost unchanged.

Of most interest, in our opinion, is that the WAE rate decreases regardless of the field direction. This experimental finding directly confirms the existence of the hole current circulating in the porous silicon during after-etching. Without the field, this current equals 8 mA/cm^2 and is comparable to the current passing through the sample during the anodic formation of *por*-Si (15 mA/cm^2 in our case).

Within our model, the current circulating during the WAE process is the result of two built-in fields. The former is produced by fluorine ions at the pore mouths and causes a potential drop there that is necessary to sustain the silicon dissolution reaction and further pore formation. The presence of the fluorine atoms is evidenced by the Auger spectroscopy data for the *por*-Si samples subjected to ion–plasma processing in the Ar + O_2 mixture. The latter built-in field is associated with the quantum-size effect and the nonuniform width of the quantum wires. It is this field that forces holes out of the wires to the *c*-Si bulk.

The experimentally discovered decrease in the WAE rate under an external electric field is explained by changes in the SCR of the Schottky-type *c*-Si–electrolyte system and those in the circular current through the *c*-Si–electrolyte–quantum wire system.

When the field is directed from the quantum wires to the *c*-Si, the Schottky structure becomes forward-biased and the SCR narrows. In this case, the mean hole concentration decreases by a factor of 1.6, according to our estimates. With the oppositely directed field, the mean hole concentration decreases by a factor of 2.5 because of the lesser hole current from the wires. Thus, our model explains why the WAE of porous silicon slows down under an external field.

ACKNOWLEDGMENTS

This work was supported by the Program "Russian Universities: Basic Research" and the Russian Foundation for Basic Research (grants nos. 99-02-17903 and 01-02-06154).

REFERENCES

1. R. L. Smith and S. D. Collins, J. Appl. Phys. **71**, R1 (1992).
2. L. T. Canham, Appl. Phys. Lett. **57**, 1046 (1990).
3. V. S. Kuznetsov and A. V. Prokaznikov, Pis'ma Zh. Tekh. Fiz. **25** (6), 81 (1999) [Tech. Phys. Lett. **25**, 242 (1999)].
4. M. E. Kompan and I. Yu. Shabanov, Fiz. Tekh. Poluprovodn. (St. Petersburg) **29**, 1859 (1995) [Semiconductors **29**, 971 (1995)].
5. M. I. J. Beal, J. D. Benjamin, M. J. Uren, *et al.*, J. Cryst. Growth **73**, 622 (1985).
6. V. Lemann and U. Gusele, Appl. Phys. Lett. **58**, 856 (1991).
7. R. Memming and G. Schwandt, Surf. Sci. **4**, 109 (1966).
8. E. A. Petrova, K. N. Bogoslovskaya, L. A. Balagurov, *et al.*, Mater. Sci. Eng. B **B69-70**, 152 (2000).
9. T. Takagahara and K. Takeda, Phys. Rev. B **46**, 15578 (1992).
10. Sh. R. Atazhanov, B. M. Kostishko, and V. S. Gorelik, Kristallografiya **44**, 551 (1999) [Crystallogr. Rep. **44**, 509 (1999)].
11. L. E. Davis, N. C. MacDonald, P. W. Palmberg, *et al.*, *Handbook of Auger Electron Spectroscopy* (Phys. Electronics Industries Inc., Minnesota, 1976).
12. S. P. Zimin, Fiz. Tekh. Poluprovodn. (St. Petersburg) **34**, 359 (2000) [Semiconductors **34**, 353 (2000)].
13. M. Shur, *Physics of Semiconductor Devices* (Prentice-Hall, Englewood Clifs, 1990; Mir, Moscow, 1992).
14. J. G. Fossum, R. P. Mertens, D. S. Lee, *et al.*, Solid-State Electron. **26**, 569 (1983).
15. M. S. Tyagi and R. Overstraaten, Solid-State Electron. **26**, 577 (1983).

Translated by V. Isaakyan

Technical Physics, Vol. 46, No. 7, 2001, pp. 853–857. Translated from Zhurnal Tekhnicheskoĭ Fiziki, Vol. 71, No. 7, 2001, pp. 67–71.

OPTICS, QUANTUM ELECTRONICS

Optical Waveguide Synthesis in Photopolymers

V. A. Vdovin, A. L. Lonin, and S. N. Mensov

Nizhni Novgorod State University, Nizhni Novgorod, 603600 Russia

Received July 3, 2000

Abstract—The dynamics of the optical synthesis of a waveguide channel in a weakly absorbing photopolymerizable composition, representing a nonlinear focusing medium with memory, was studied by experimental and numerical modeling methods. A qualitative model is proposed that describes the formation of a nondivergent light guide. It is shown that an extended (≥1 cm) light guide can form only in a medium possessing a sufficiently high contrast factor ($\gamma \geq 2$). *© 2001 MAIK "Nauka/Interperiodica".*

Modern photopolymer compositions are widely used for obtaining three-dimensional structures in stereolithography [1] and phase gratings in holography [2]. In these applications, it is sufficient to set the light intensity distribution in a medium that would correspond to the geometry of a structure to be obtained. The polymer density and, hence, the corresponding nonlinear contribution to the refractive index, depend on the polymerization conditions [3]. This circumstance can be employed, in principle, for obtaining structures with refractive index gradients by creating light fields with inhomogeneous intensity distributions. However, the formation of a desired spatial distribution of the light intensity inphotopolymers is complicated by the refraction (both linear and nonlinear) of light beams. Therefore, in order to ensure the formation of an extended nondivergent light guide in a photopolymer, it is necessary to provide for the light beam "self-channeling" in this nonlinear focusing medium with memory.

The self-channeling of a He–Cd laser beam (λ = 0.325 μm) in photopolymer compositions sensitive in the UV spectral range was numerically and experimentally studied by Kewitsh and Yariv [4]. However, strong UV absorption of such compositions allowed the laser radiation self-channeling effect to be obtained only for a distance below one millimeter. Photopolymer compositions [5] exhibiting polymerization under the action of electromagnetic radiation in the visible spectral range exhibit small optical absorption both in the initial stage of polymerization and upon its termination, which makes such media promising structural materials for optical waveguide technology.

In this study, we have numerically and experimentally investigated the possibility of the optical (λ = 0.63 μm) waveguide synthesis in a weakly absorbing photopolymerizable composition using the effect of the light beam self-channeling.

NUMERICAL MODELING OF THE LIGHT BEAM–PHOTOPOLYMER INTERACTION

The polymerization time constant of a photopolymer is greater by several orders of magnitude than the characteristic time of a light wave variation. Therefore, the dynamics of the self-action of a light beam in a nonlinear photopolymerizable medium with memory can be studied within the framework of a quasi-stationary approximation. According to this, the complex amplitude of a cylindrically symmetric light beam $E(r, z)$ possessing a narrow spatial spectrum is described at each time instant by a stationary parabolic equation. With neglect of the optical absorption, this equation can be written as [6]

$$2ik\frac{\partial E(r,z)}{\partial z} = \Delta_\perp E(r,z) + k^2\frac{2\Delta n(r,z,t)}{n_m}E(r,z) \quad (1)$$

with the stationary boundary conditions

$$E(r,0) = E(0,0)\exp\left(-\frac{r^2}{a^2}\right). \quad (2)$$

Here, a is the beam halfwidth, k is the wavevector, and $\Delta n(r, z, t)$ is a positive increment of the refractive index of the photopolymerizing composition. The latter quantity is determined by the process prehistory and has the maximum value $\Delta n_{\max} = n_p - n_m$, where n_p and n_m are the refractive indices of the polymer and monomer, respectively. As is known, variation of the refractive index of a photopolymerizing composition can be described by an exponent [7]

$$\Delta n(r,z,t) = \Delta n_{\max}(1 - \exp[-(H(t,z,t)H_0^{-1})^\gamma]), \quad (3)$$

where $H(r, z, t) = \int_0^t |E(r, z, t')|^2 dt'$ is the exposure, H_0 is the normalization coefficient, and $\gamma > 0$ is a factor determining the contrast of the given photosensitive medium (typically, $1 \leq \gamma \leq 10$).

1063-7842/01/4607-0853$21.00

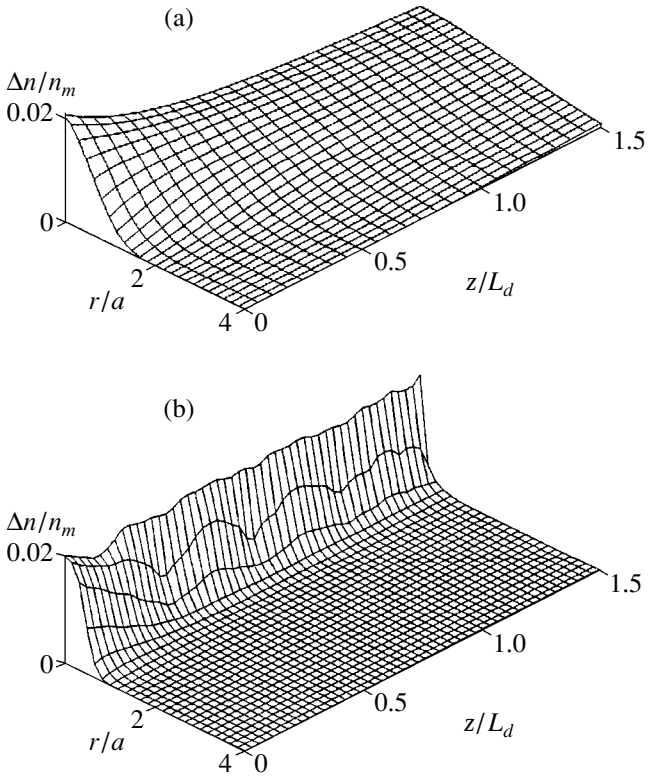


Fig. 1. Refractive index profiles along the polymer track for $\gamma = 1$ (a) and 2 (b).

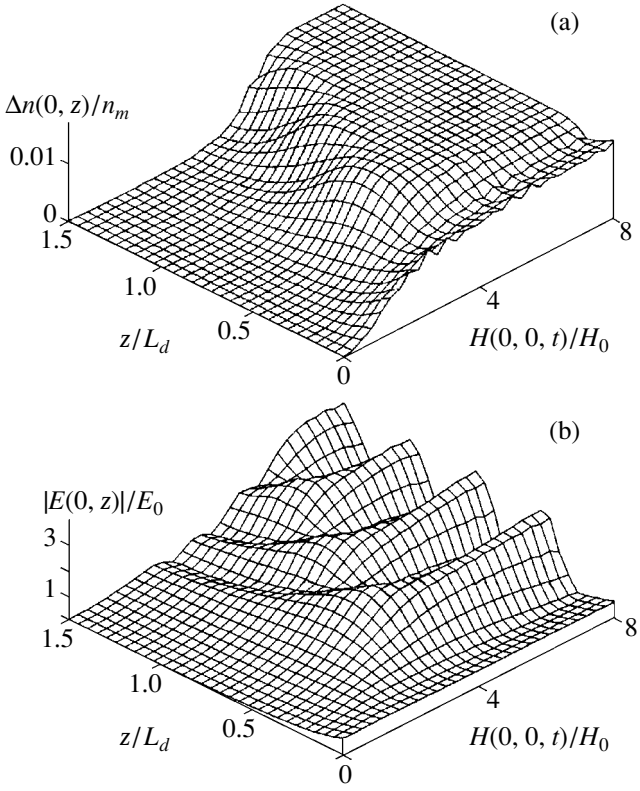


Fig. 2. Axial profiles of the (a) refractive index and (b) dimensionless light field amplitude for various exposures.

The numerical experiment was conducted for a light beam with a width of $a = 20$ μm and $\lambda = 0.63$ μm propagating in a medium with $\Delta n_{\max} = 0.02 n_m$ and various values of the contrast factor γ over a distance on the order of the diffraction length $L_d = ka^2/2$.

It was found that the character of the refractive index variation in a channel formed in the photopolymer along the light beam path depends significantly on the contrast factor γ. For $\gamma \approx 1$, the channel in the photopolymer reproduces the profile of a diverging Gaussian beam (Fig. 1a). For $\gamma \geq 2$, a channel formed in the z direction has a virtually homogeneous profile with the cross section repeating (Fig. 1b) that of the input Gaussian beam, which indicates that the self-channeling effect takes place.

The development of the self-channeling process is clearly illustrated by profiles of the refractive index (Fig. 2a) and the light field amplitude modulus (Fig. 2b) variation along the channel symmetry axis z with time t. As is seen from Fig. 2, the channel length in the z direction is provided by the sequential focusing effect, whereby local intensity maxima determine nonexpanding channel cross section. The channel homogeneity is due to averaging (with allowance for the saturation of the exposure characteristic (3)) of the contribution of these local maxima moving along the channel toward the light source.

THE LENS MODEL OF THE SELF-CHANNELING EFFECT

The process of formation of a nondivergent optical waveguide channel can be explained qualitatively within the framework of the following model. For a Gaussian beam (2) propagating in a homogeneous medium, the spatial distribution of the complex field amplitude is described by the formula

$$E(r, z) = E(0, 0)\left(1 + \frac{z^2}{L_d^2}\right)^{-0.5} \times \exp\left[-\frac{r^2}{a^2}\left(1 + \frac{z^2}{L_d^2}\right)^{-1}\right] \times \exp\left[-\frac{ikr^2}{2}\left(1 + \frac{z^2}{L_d^2}\right)^{-1}\frac{z}{L_d^2}\right]. \quad (4)$$

In the beginning of polymerization, when $(tE^2) \ll H_0$ and the increment of the refractive index is insufficient to produce a significant effect upon the beam, a change in the refractive index follows the field intensity variation $E(r, z)$ according to formula (4) in the incident beam. With an allowance for the exposure characteris-

tic (3), the refractive index increment is described by the expression

$$\Delta n(r, z) = \Delta n_{\max}(1 - \exp(-(t|E(r, z)|^2 H_0^{-1})^{\gamma})). \quad (5)$$

Let us define the longitudinal size l of the inhomogeneity as a distance over which the refractive index increment drops e times. The l value can be estimated using formula (5) for $r = 0$ and $(tE^2) \ll H_0$, which yields

$$\Delta n(0, z) \approx \Delta n_{\max}\left[\frac{t|E(0, 0)|^2}{H_0}\right]^{\gamma}\left(1 + \frac{z^2}{L_d^2}\right)^{-\gamma}, \quad (5a)$$

and, accordingly,

$$l \approx L_d\sqrt{\exp(\gamma^{-1}) - 1}. \quad (6)$$

Relationship (6) shows that the case of $\gamma \approx 1$ corresponds to the formation of an extended polymer track, the transverse size of which increases on moving along z as described by formula (4) for a divergent beam. For large γ, the track length is relatively short and we may consider the inhomogeneity as a thin lens formed near the initial plane $z = 0$ and possessing the transmission coefficient

$$T(r) = \exp(ikd\Delta n(r, z)) \approx \exp\left(ik\Delta n(0, 0)d - \frac{ikr^2}{2}p\right), \quad (7)$$

where p is the optical power of the lens. Taking into account that the refractive index increment at small t in the paraxial region is

$$\Delta n(r, 0) \approx n_{\max}(t|E(0, 0)|^2 H_0^{-1})^{\gamma}\frac{2\gamma}{a^2}r^2, \quad (5b)$$

and using relationships (6) and (7), we obtain an expression for the lens power:

$$p = 2k\gamma\sqrt{\exp(\gamma^{-1}) - 1}\Delta n_{\max}(t|E(0, 0)|^2 H_0^{-1})^{\gamma}. \quad (8)$$

The lens focuses of the initial Gaussian beam $E(r, 0)$ described by formula (4). The corresponding intensity distribution behind the lens in the Fresnel paraxial approximation [8] is

$$|E(r, z)|^2 = \frac{a^2}{L_d^2}\left(\frac{z^2}{L_d^2} + (1 - zp)^2\right)^{-1} \times \exp\left(-\frac{2r^2}{a^2}\left(\frac{z^2}{L_d^2} + (1 - zp)^2\right)^{-1}\right). \quad (9)$$

Using this expression, we may determine the distance z_0 at which the Gaussian beam exhibits a maximum intensity (and minimum width Δr). The plots of z_0 and Δr as functions of the lens power p are presented in Fig. 3. As is seen, the beam width in the cross section $z = z_0$ for $p > 1/L_d$ is smaller than the initial transverse size at $z = 0$. Therefore, the light intensity in this cross section exceeds the initial level and, hence, the polymerization reaction proceeds at a higher rate. This, in turn, leads to the formation of the next local maximum in the refractive index profile in the vicinity of $z = z_0$ in accordance with the results of numerical calculations (see Fig. 2a). As the power (8) of the induced lens increases, the intensity maximum shifts further toward the input plane. This results in the sequential formation of a thin channel with a length of $\sim L_d/2$ in the polymerized composition. Since the light beam at the exit of this channel is similar in many respects to that at the entrance cross section, the process of the polymer channel formation continues further and the next waveguide portion appears as described above. A series of sequential focusing stages results in the formation of an extended channel with a transverse size close to that in the initial cross section.

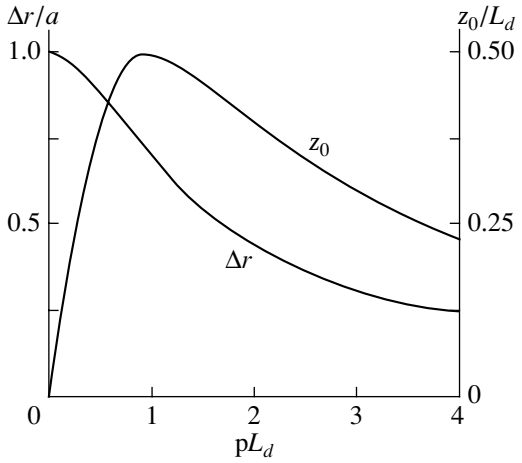


Fig. 3. Plots of the maximum intensity coordinate z_0 and width Δr of a Gaussian beam versus lens power p.

EXPERIMENTAL STUDY OF THE OPTICAL WAVEGUIDE FORMATION

The process of the optical synthesis of a thin waveguide channel in the course of polymerization was experimentally studied using a composition based on an OKM-2 photopolymer with a photoinitiator sensitive in the spectral range of $\lambda < 0.65$ μm. The contrast factor of this medium depends on the radiation intensity (beam power density) I and reaches the values $\gamma_1 \approx 1$ and $\gamma_2 \approx 2$ for $I_1 = 100$ mW/mm^2 and $I_2 = 10$ mW/mm^2, respectively. A light beam with a crossover halfwidth of $a = 15$ μm was obtained by focusing the radiation of a point source. The source was created from a He–Ne laser radiation ($\lambda = 0.6328$ μm) with the aid of a short-focus lens and a filter of low spatial frequencies (a nontransparent foil with a 20-μm hole) placed in the focal plane [9]. The necessary laser beam power was set prior to experiments, by monitoring the radiation intensity

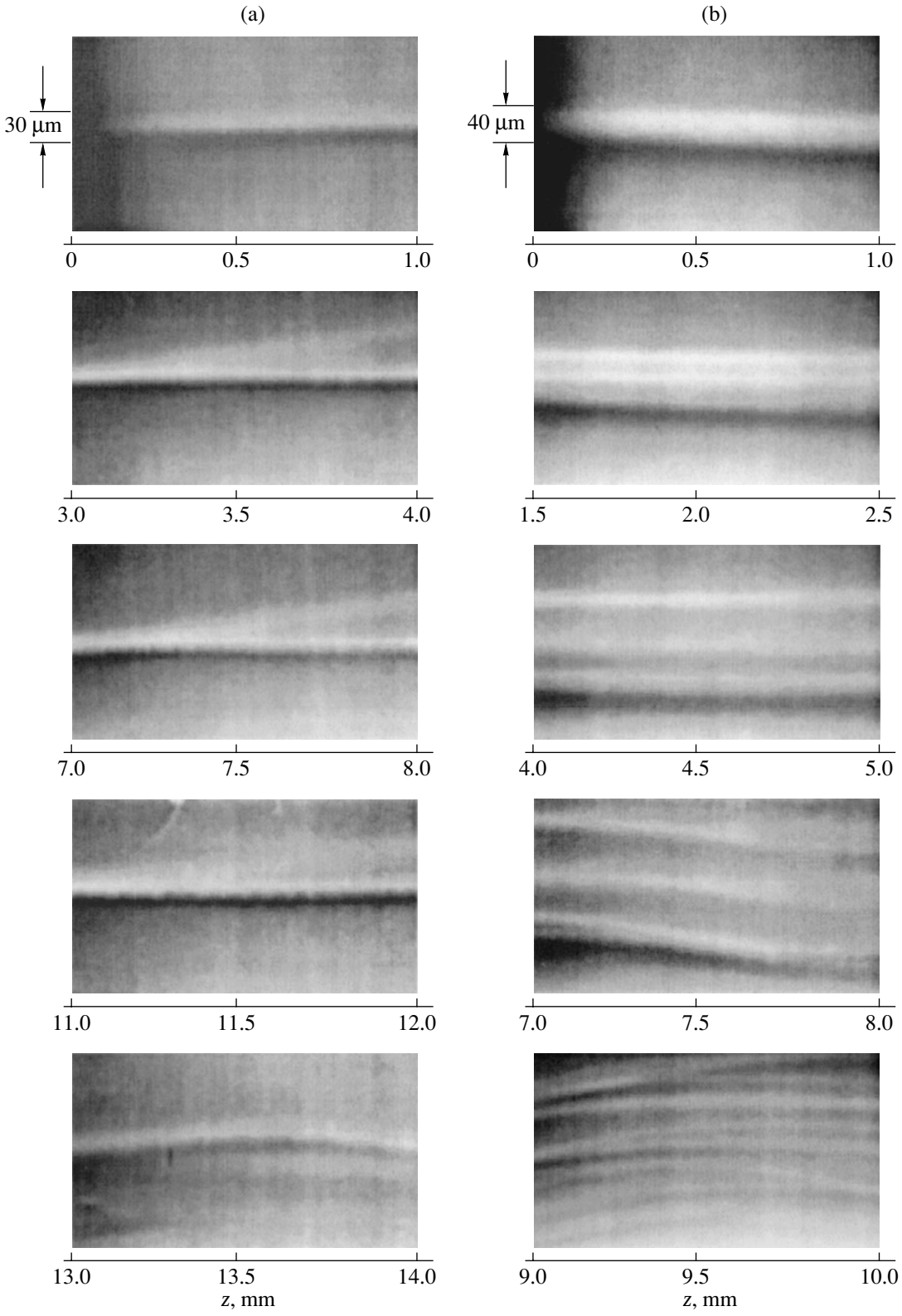


Fig. 4. Photographs showing polymer tracks formed for two different radiation power densities corresponding to (a) $\gamma_1 \approx 1$ and (b) $\gamma_2 \approx 2$.

with a calibrated photodiode and controlling the level with a polarization attenuator.

The polymer track was formed at the edge of a glass cell filled with a photopolymer composition and placed at a beam crossover point. The track was visualized by lateral illumination with an IR ($\lambda = 0.89$ μm) emitting diode. The polymerization process dynamics was monitored with the aid of a video camera (Fig. 4).

For a radiation power density of $I_1 = 10$ mW/mm^2, we obtained a thin channel with a width of 30 μm and a length of 1.5 cm. A photograph of this track is presented in Fig. 4a. The track formed at a greater power density of 100 mW/mm^2 is divergent (Fig. 4b) and becomes inhomogeneous as a result of instability development in the focusing nonlinear medium [7].

Thus, it is possible to synthesize extended nondivergent light guides with a gradient refractive index profile in the cross section using weakly absorbing photopolymerizable compositions with a sufficiently high contrast factor ($\gamma \geq 2$). It should be emphasized that the proposed optical waveguide technology does not require using high-power initiating laser beams and ensures virtually complete use of the photopolymer composition (wasteless technology).

ACKNOWLEDGMENTS

This study was supported by grants from KTsFE (no. 97-0-8.3-82) and GNTS (no. 02.07.01.07).

REFERENCES

1. J. P. Kruth, CIRP Ann. **40** (2), 69 (1991).
2. B. L. Volodin, B. Kippelen, K. Meerholz, *et al.*, Nature **383**, 58 (1996).
3. L. I. Ginsburg, E. I. Krivchenko, and E. I. Egorova, in *Proceedings of the All-Union Conference "Free Radical Polymerization," Gorkii, 1989*, p. 125.
4. A. S. Kewitsh and A. Yariv, Opt. Lett. **21** (1), 24 (1996).
5. G. A. Abakumov, V. A. Vdovin, O. N. Mamysheva, *et al.*, in *Proceedings of the 5th Conference on the Chemistry and Physical Chemistry of Oligomers, Chernogolovka, 1994.*
6. M. B. Vinogradova, O. V. Rudenko, and A. P. Sukhorukov, *The Theory of Waves* (Nauka, Moscow, 1979).
7. G. A. Abakumov, S. N. Mensov, and A. V. Semenov, Opt. Spektrosk. **86**, 1029 (1999) [Opt. Spectrosc. **86**, 927 (1999)].
8. V. A. Zverev, *Image Formation by Wave Fields* (Inst. Prikl. Fiz. Ross. Akad. Nauk, Nizhni Novgorod, 1998).
9. R. J. Collier, C. B. Burckhardt, and L. H. Lin, *Optical Holography* (Academic, New York, 1971; Mir, Moscow, 1973).

Translated by P. Pozdeev

Technical Physics, Vol. 46, No. 7, 2001, pp. 858–863. Translated from Zhurnal Tekhnicheskoĭ Fiziki, Vol. 71, No. 7, 2001, pp. 72–78.

OPTICS, QUANTUM ELECTRONICS

A Dye Laser with a Polyurethane Matrix

V. I. Bezrodnyĭ, N. A. Derevyanko, A. A. Ishchenko, and L. V. Karabanova

Physics Institute, National Academy of Sciences of Ukraine, pr. Nauki 46, Kiev, 252650 Ukraine
e-mail: bezrod@iop.kiev.ua

Received August 2, 2000

Abstract—A dye laser with a polyurethane matrix cured by polycondensation is studied. An increase in the conversion efficiency, service life, optical stability, and tuning range in comparison with lasing media produced by radical polymerization is demonstrated. *© 2001 MAIK "Nauka/Interperiodica".*

Tunable lasers are widely used in basic research and applications. Dye lasers appear to be the most promising and readily available sources of high-power radiation tunable in the visible range. They often exploit liquid solutions of dyes as lasing media. However, the instability of radiation propagation in liquids due to nonlinear processes [1], thermooptical distortions, and degradation of the dye necessitate the use of complex bulky pump-over systems [2]. In addition, the solvents and dyes may pollute the environment. Note also hazards related to the flammability and toxicity of the solvents.

Solid lasing media are an attractive alternative to the dye solutions. However, the list of solid matrices is limited by strict requirements to their service characteristics. There are two main groups of solid matrices: polymer and inorganic microcomposite matrices based on sodium borosilicate porous glass [3, 4] and sol–gel matrices [5]. The dyes can easily be embedded in the polymer matrices. The starting materials for such matrices are readily available, and the matrices are easy to fabricate; therefore, lasers on their basis are cheaper than crystal-based devices.

Poly(methyl methacrylate) (PMMA) and its modifications [6–8], polyurethane acrylate [9, 10], and various epoxy resins [11, 12] are the polymers most widely used as the lasing media of dye lasers. PMMA and polyurethane acrylate matrices are cured by radical polymerization. In the former case, substances (normally, peroxide compounds) initiating the formation of free radicals upon heating are added to the reaction mass. In the latter case, the use of photosensitizers allows one to initiate the radical reaction of polymerization by UV irradiation. The free radicals combine with the dye molecules, which may cause their dissociation. Thus, the PMMA- and polyurethane acrylate-based lasers suffer from dye instability. The excitation in the fundamental absorption band ($S_0 \longrightarrow S_1$ transition) by both weak and intense light leads to the further degradation of the laser dye [10].

Epoxy polymers are cured via the reaction of polycondensation, which is more suitable than the radical reactions used for the systems with organic dyes. However, hardeners employed in the polycondensation reaction exhibit either alkaline or acidic properties and may cause irreversible chemical reactions with many organic dyes. In our previous work [13], we proposed using cross-linked polyurethane as a polymer matrix for passive dye-laser switches. This polymer is also cured by polycondensation. However, in contrast to the polycondensation of epoxy compositions, the reaction takes place at room temperature in a neutral medium. Therefore, dyes of any class can be introduced at the stage of polymerization virtually without degradation. For example, polymethine dyes intensely absorbing in the IR spectral range (>1000 nm) have been found to be stable in polyurethane [14], while decomposing completely and by 80% at the stage of the polymerization of PMMA and polyurethane acrylate, respectively [10].

The purpose of this work was to study polyurethane as a matrix for lasing media with organic dyes and to develop a tunable laser with such an active medium. Note that the polymer matrices of the active media are subjected to higher radiation and thermal loads when interacting with laser emission and pumping radiation than those of passive laser switches. A typical dye concentration in the active medium is larger than that in the passive one by at least one order of magnitude. Hence, the aggregation probability of dye molecules in the former is higher. Dye aggregation is the main reason for a decrease in the optical stability and service time, as well as for worsening the spectral and luminescent properties of dye-containing media [15, 16]. An increase in the dye concentration also stimulates the generation of contact ion pairs, in which anion-to-cation electron phototransfer with the production of neutral radicals is a possibility [17]. The radicals enhance the degradation of the dye. In addition, the pumping of the active lasing media needs power densities larger than those needed for bleaching passive laser switches. The reason is that efficient lasing requires the population inversion as high as possible, whereas *Q* switching is possible even at equal populations of the ground and excitation levels. Still further increase in the load on the

1063-7842/01/4607-0858$21.00

active laser media results from the presence of two, pumping (λ_p) and lasing, wavelengths. Normally, λ_p is much shorter than the lasing wavelength of passively *Q*-switched lasers. The radiation resistance of the polymer decreases with decreasing pumping wavelength [18].

Most laser polymers exhibit two-photon absorption of the pumping radiation. For example, the electron transition at λ_{max} = 266 nm corresponds to the two-photon absorption of the pumping radiation with λ_p = 532 nm. Thus, the degradation of the active laser medium can also be related to the two-photon absorption of the high-power pumping radiation. This is not observed in passive laser switches. Based on the aforesaid and on experience in using polymers in dye lasers, we tried to select a polymer matrix offering a high radiation resistance, a long service time, and a high optical stability.

The radiation resistance of a polymer matrix depends on the viscoelastic properties of the material [6, 7, 19]. A necessary stability can be attained both by the multiple distillation of the initial substances with the subsequent ultrafiltration of the monomers and by the modification of the polymer. Investigations into the laser destruction of polymer matrices [6, 19] show that one must use materials allowing significant elastic strains within a wide range of working temperatures. In particular, low- and high-molecular admixtures in PMMA increase its radiation resistance to the practical level [6, 7]. Lasing media with a cross-linked polyurethane acrylate matrix exhibited the high elasticity within a wide temperature interval [9, 10]. These media withstand high impulsive loads even without adding plasticizers.

In this work, we report the basic parameters of dye lasers based on polyurethane and polyurethane acrylate, which have similar physical properties. Polyurethane is highly transparent in the wide spectral range 0.32–2.20 μm, which almost completely spans the range where dye lasers are used. Curve *1* in Fig. 1 shows the short-wavelength transmission edge of the polyurethane. It is seen that one can use conventional pumping sources, such as nitrogen lasers with λ_p = 337 nm or longer-wave devices. In contrast, the intense short-wavelength absorption of polyurethane acrylate (curve *2* in Fig. 1) makes lasing in the near-UV spectral range impossible.

We determined the radiation resistance of the polyurethane matrix at the wavelength of the second harmonic (λ = 532 nm) of a single-mode single-shot Nd:YAG laser. At the shot duration 15 ns and the spot diameter 200 μm, the threshold of single-shot destruction was no less than 15 J/cm^2. This value is 1.5 times larger than the threshold for polyurethane acrylate under the same conditions [9].

Both lasing media represented triplex structures in which the polymer was sandwiched in two substrates. The polyurethane acrylate structure was produced by the method described in [9, 10], and the polyurethane structure, with the technique in [13, 14]. In experiments, we used xanthene, polymethine, and pyrromethene dyes (rhodamine 6G, astraphloxine, and pyrromethene 597, respectively).

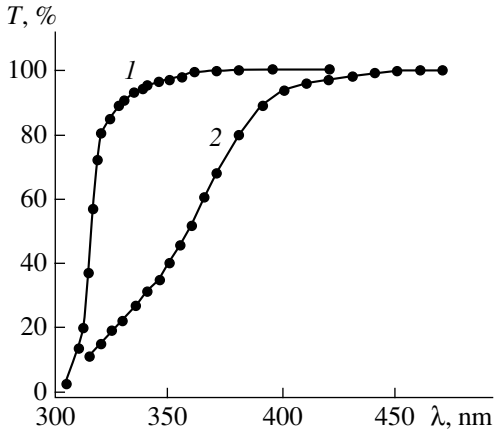


Fig. 1. Short-wavelength transmission edge of (*1*) polyurethane and (*2*) polyurethane acrylate matrices.

The basic parameters of polymer-based dye lasers are service life (the number of shots *N* to failure incident on the same region of the sample), conversion efficiency η, and optical stability. The optical stability was studied on tripleces with quartz substrates. In these experiments, the thickness of the dye-doped polymer film was 300 μm. The optical density *D* in the peak of the basic transition was about unity. The samples were irradiated by a 50-W/cm^2 xenon lamp with the continuous spectrum in the UV and the visible spectral ranges. The optical density of the polymer samples vs. incident energy *E* was measured in the absorption maximum for a given dye using a Shimadzu UV-3100 spectrophotometer. Figure 2 shows optical bleaching curves of the dyes in the polymer matrices. It is seen that the optical stability of the polyurethane samples is substantially higher than that of the polyurethane acrylate ones.

Unexpectedly, the pyrromethene 597 dye with the network structure appeared to be the least stable in both polymers. Conversely, astraphloxine with the unclosed polymethine chain, which is highly sensitive to photochemical reactions, showed the highest optical stability in the polyurethane matrix. As follows from Fig. 3, pyrromethene 597 and rhodamine 6G markedly absorb in the spectral region 320–420 nm, while astraphloxine does not. Therefore, it can be concluded that the UV-induced electron transitions to higher excited states promote the photochemical decomposition of the first two dyes. This should be taken into account when developing optically stable dyes. The higher excited states of dye molecules may nonradiatively be deactivated by vibrational relaxation or by internal conversion via highly excited vibrational states of the polymer to form free radicals of the macromolecules. Since the

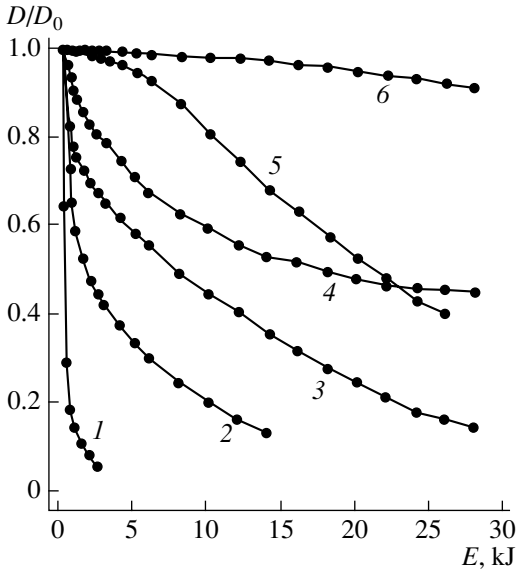


Fig. 2. Photobleaching versus radiation energy curves for (*1*, *4*) pyrromethene 597, (*2*, *6*) astraphloxine, and (*3*, *5*) rhodamine 6G in (*1–3*) polyurethane acrylate and (*4–6*) polyurethane matrices.

overlap of the UV absorption bands of the dyes and the polymer is stronger for polyurethane acrylate, the formation of the free radicals is more probable in this case, which the curves of photodestruction in Fig. 2 indicate.

The service life of the lasing media and the lasing efficiency were studied on triplex samples. The thickness of the dye-doped polymer layer between two glass substrates was 2 mm. The optical density at the pumping wavelength was $D(\lambda_p) = 5$. Two mirrors with the reflection coefficients $R_1 = 99.5\%$ and $R_2 = 30\%$ formed a 16-cm-long nondispersive cavity of the dye laser. We used quasi-longitudinal excitation. The angle between the pumping and laser beams was 15°. The tripleces were placed near the exit mirror at the Brewster angle to the optical axis of the cavity. The second harmonic of a multimode Nd : YAG laser was used for pumping ($\lambda_p = 532$ nm). The pumping pulse energy was 35 mJ, and the pulse duration was 18 ns. A lens with a focal distance of 1 m focused the pumping radiation into a spot of diameter 1.75 mm. The energy density incident on the polymer sample was 1.45 J/cm^2. At a pulse repetition rate of 3 Hz, the mean pumping power density was 4.35 W/cm^2. The service life of lasing media is known to depend on the energy density and power of pumping radiation [8]. We used a high-power pumping radiation for quickly determining the service life of the two polymer-based lasing media. To make the pumping conditions still more severe, thin dye-doped polymer layers were sandwiched in the glass substrates. Note that thick polymer layers made of, for example, poly(methyl methacrylate), have longer service lifes [20] because the side concentration effects (see below) weaken and the thermal load on the polymer layer lowers. The high radiation resistance of both polyurethane and polyurethane acrylate allowed the high-power pumping radiation to be incident on a local area of the thin polymer layer without rotating the disk (rotation is used to increase the service life of the medium).

Figure 4 plots the conversion efficiency η versus the number of pulses N (service life). The efficiency is determined as the ratio of the output energy of the dye laser to the pumping energy. In all samples, the efficiency decreases with increasing N because of the photobleaching of the dye without damaging the polymer matrix. The polyurethane media have longer service lifes than the polyurethane acrylate structures. The table shows that the initial efficiencies η_0 for the polyurethane laser media are larger than those for the polyurethane acrylate media. This effect is especially pronounced for pyrromethene 597 and rhodamine 6G. The reason for the smaller η_0 in the samples made of polyurethane acrylate is the degradation of the dyes during laser fabrication [10]. This was confirmed in [21], where PMMA-based lasing media were made by radical polymerization using benzoyl peroxide as a thermal initiator. The destruction of the dyes is accompanied by the formation of reaction products absorbing in the spectral region of lasing. Such absorption was found to be typical of rhodamine 6G and pyrromethene 597 but is negligible for astraphloxine in polyurethane acrylate; at the same time, it is totally absent in polyurethane, where the dyes do not degrade at the stage of laser fabrication. Hence, η_0 in the polyurethane media is higher.

Operating characteristics of the dye lasers

Lasing medium	Irradiation energy density at which the optical density $D = 1/2D_0$, kJ/cm^2	Number of pulses at which the efficiency $\eta = 1/2\eta_0$, $N \times 10^3$	Initial conversion efficiency η_0 in nondispersive cavity, %	Full width of tuning curve at zero level, nm
Rhodamine 6G in polyurethane	23.0	8.0	34	555–597
Rhodamine 6G in polyurethane acrylate	8.0	0.9	29	558–597
Astraphloxine in polyurethane	160.0	6.5	32	590–620
Astraphloxine in polyurethane acrylate	2.0	0.7	30	590–620
Pyrromethene 597 in polyurethane	18.5	12.2	75	550–610
Pyrromethene 597 in polyurethane acrylate	0.3	2.2	58	553–607

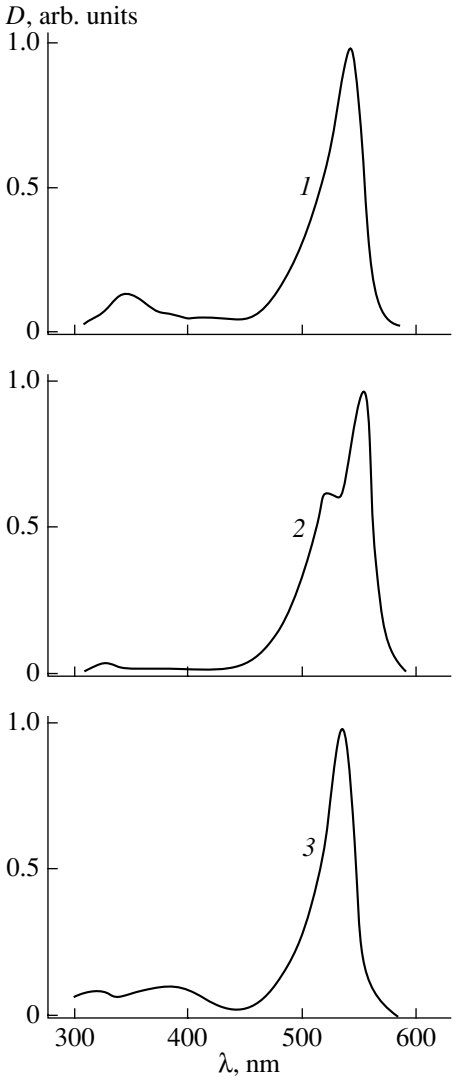


Fig. 3. Absorption spectra of (*1*) rhodamine 6G, (*2*) astraphloxine, and (*3*) pyrromethene 597 in polyurethane.

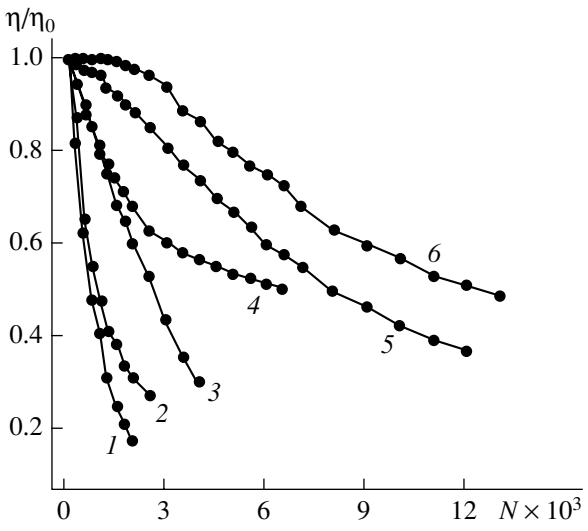


Fig. 4. Normalized conversion efficiency versus number of pulses for (*1*, *4*) astraphloxine, (*2*, *5*) rhodamine 6G, and (*3*, *6*) pyrromethene 597 in (*1–3*) polyurethane acrylate and (*4–6*) polyurethane.

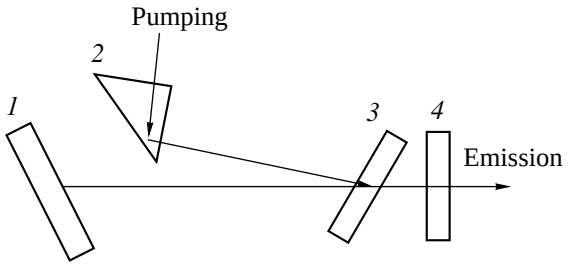


Fig. 5. Optical scheme of the tunable dye laser.

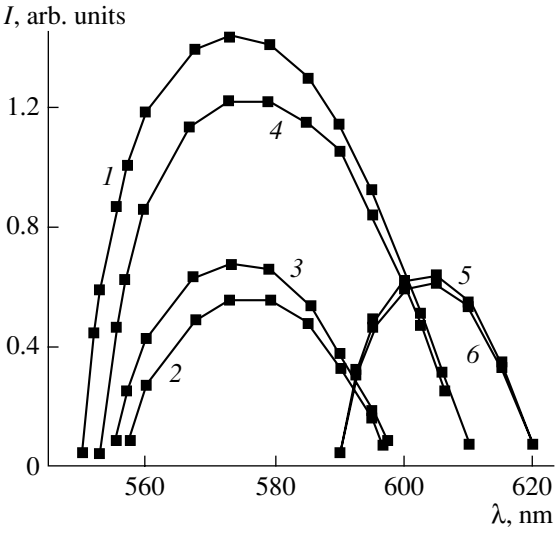


Fig. 6. Tuning curves for (*1*, *4*) pyrromethene 597, (*2*, *3*) rhodamine 6G, (*5*, *6*) astraphloxine in (*1*, *3*, *5*) polyurethane acrylate and (*2*, *4*, *6*) polyurethane.

The insignificant destruction of astraphloxine in polyurethane acrylate accounts for the small difference in η_0 for the lasing structures based on astraphloxine-doped polyurethane and polyurethane acrylate (see table).

To determine the tuning range of the lasers, we used a dispersive Littrow cavity (Fig. 5). Reflection ruled grating *1* (the ruling density 1200 mm^{-1}) serves as a dispersive element. Auxiliary totally reflecting prism *2* minimizes the angle (6°) between the pumping beam and the cavity axis. Lasing medium *3* is placed at the Brewster angle to the optical axis. The reflection coefficient of exit mirror *4* is 30%. The laser was tuned by rotating grating *1*. The spectrum was determined with an SPM-2 specular monochromator. A signal from its exit slit was applied to an FD-24 photodiode and then to an S1-79 oscilloscope. As we did not use expansion of the laser beam, the pumping energy was lowered

down to 10 mJ at a spot size of 2 mm to avoid the radiation-induced damage of the grating. Figure 6 shows tuning curves for the three dyes in the two polymer matrices. The lasers with rhodamine 6G and pyrromethene 597 in polyurethane acrylate (curves *2* and *4*, respectively) have not only smaller efficiencies but also narrower tuning ranges in comparison with the polyurethane lasers (curves *1* and *3*). The latter fact is explained by the red shift of the short-wavelength edges of the corresponding tuning curves. The decrease in the conversion efficiency and the narrowing of the tuning range in the dispersive-cavity polyurethane acrylate-based dye lasers are associated with the decomposition products formed during the radical polymerization of the matrix. The tuning curves for astraphloxine (curves *5* and *6*) are very close to each other. Recall that the decomposition of astraphloxine during the radical polymerization of polyurethane acrylate is the least (5% vs. 14% and 20% for rhodamine 6G and pyrromethene 597, respectively). The optical stability of astraphloxine in polyurethane acrylate is smaller than that of rhodamine 6G (see curves *2* and *3* in Fig. 2). However, the curing time of the oligomer composition with astraphloxine is 6–7 times shorter than that for the sample with rhodamine 6G, because the former dye does not absorb in the spectral region 320–420 nm. This is the reason for the insignificant decomposition of astraphloxine in polyurethane acrylate at the stage of fabrication and for the small difference in η_0 for the matrices studied. The dyes in the polyurethane matrix do not decompose during fabrication; therefore, the polyurethane media have a larger conversion efficiency and a wider tuning range in comparison with the polyurethane acrylate ones.

The dye lasers with pyrromethene exhibit the larger efficiency and the wider tuning range than those with conventional rhodamine 6G. Over the last few years, pyrromethene dyes have been widely used in lasers emitting in the visible range owing to their high conversion efficiency [21, 22]. The photochemical stability of pyrromethene dyes in polyurethane matrices is substantially larger than that in radical-containing matrices made of polyurethane acrylate [9, 10] and PMMA [21]. Thus, the use of the polyurethane matrices makes it possible to improve the stability of the dye lasers.

Summarizing the results obtained, we conclude that the optical stability, service life, conversion efficiency, and tuning range of the lasing media with polyurethane are superior to those of the polyurethane acrylate structures. The reason is the stronger degradation of the dyes in polyurethane acrylate. First, the dyes in this polymer partially decompose even at the stage of radical polymerization. Such a mechanism favors the further activation of the radical reactions when the lasing media interact with light. Second, polyurethane is a highly polarizable [24] polymer. Its dielectric constant exceeds those of polyurethane acrylate, epoxides, and PMMA [24]. That is why the degree of dissociation of salt-like dyes in polyurethane is larger than that in the other polymers. A large number of nucleophilic groups in the polymer chains of polyurethane also contributes to the separation of counterions, since they produce a solvation sheath around the cations of dyes [16]. Thus, the probability of contact ion pairs forming in polyurethane is far smaller than in polyurethane acrylate, epoxides, and PMMA. Therefore, electron phototransfer and the association of the contact ion pairs by electrostatic attraction in the polyurethane matrices are unlikely in contrast to the other polymers [16]. Indeed, if chloride anions in astraphloxine are replaced by tetrafluoroborate, having higher electron affinity and nucleophilicity, the lasing characteristics and the service life of this dye in polyurethane matrix practically do not change. In a polyurethane acrylate matrix, the substitution leads to a slight broadening and a hypsochromic shift of the absorption band and an increase in the optical stability. According to [16], the dependence of these parameters on the counterion properties indicates the formation of contact ion pairs. Interestingly, pyrromethene 597 has the lowest optical stability and the longest service life (see table and curves *3* and *6* in Fig. 4). This is additional evidence of the fact that the higher states of pyrromethene 597 take part in its photochemical decomposition. In studying the service life, we excited only the fundamental absorbing transition $S_0 \longrightarrow S_1$, while the optical stability was studied in the $S_0 \longrightarrow S_2$ transition range as well. Chemically, intraion pyrromethene 597 is a contact ion pair that cannot dissociate at any polarity of the polymer. However, the distance between the charges in this pair is rather large. In addition, the negative charge is markedly localized at the borofluoride bridge because of the high electronegativity of fluorine atoms. Therefore, the intramolecular phototransfer of electrons from boron to nitrogen in pyrromethene 597 is more difficult than in chemically unbounded contact pairs of the cationic dyes (astraphloxine and rhodamine 6G), where the distance between the counter-ions depends on the polarity and the solvating power of the medium. As a result, pyrromethene 597 has the service life longer than astraphloxine and rhodamine 6G in both polymers. As in the case of salt-like cation dyes, the electrostatic attraction of the charges in intraionic pyrromethene 597 facilitates molecular aggregation. For the above reasons, the probability of aggregation in polyurethane is much smaller than that in polyurethane acrylate. The low probability of aggregation, combined with soft polymerization, provide the long service lifes of the dyes in the polyurethane matrix.

Thus, the polyurethane matrix is a promising lasing material for dye lasers. Unlike the other polymers, this medium ensures the stability of dyes of any class. Lasers with the polyurethane matrix may offer long-term stable operation because of their high photochemical stability, conversion efficiency, and lasing characteristics.

REFERENCES

1. V. I. Bezrodnyĭ, V. I. Vashchuk, and E. A. Tikhonov, Zh. Tekh. Fiz. **48**, 151 (1978) [Sov. Phys. Tech. Phys. **23**, 89 (1978)].
2. G. A. Vesnicheva and A. Yu. Ivanov, Zh. Tekh. Fiz. **64**, 201 (1994) [Tech. Phys. **39**, 114 (1994)].
3. V. I. Zemskiĭ, Yu. L. Kolesnikov, and I. K. Meshkovskiĭ, Pis'ma Zh. Tekh. Fiz. **12**, 331 (1986) [Sov. Tech. Phys. Lett. **12**, 136 (1986)].
4. G. B. Al'tshuler, V. A. Bakhanov, and E. G. Dul'neva, Opt. Spektrosk. **62**, 1201 (1987) [Opt. Spectrosc. **62**, 709 (1987)].
5. E. T. Knobbe, B. Dunn, P. D. Fuqua, and F. Nishida, Appl. Opt. **29** (18), 2729 (1990).
6. D. A. Gromov, K. M. Dyumaev, and A. A. Manenkov, Izv. Akad. Nauk SSSR, Ser. Fiz., No. 7, 1364 (1984).
7. D. A. Gromov, K. M. Dyumaev, and A. P. Maslyukov, Izv. Akad. Nauk SSSR, Ser. Fiz., No. 6, 1084 (1985).
8. L. K. Denisov, I. G. Kytina, and V. G. Kytin, Kvantovaya Élektron. (Moscow) **24**, 119 (1997).
9. V. I. Bezrodnyĭ, O. V. Przhonskaya, and E. A. Tikhonov, Kvantovaya Élektron. (Moscow) **9**, 2455 (1982).
10. V. I. Bezrodnyĭ, M. V. Bondar, and G. Yu. Kozak, Zh. Prikl. Spektrosk. **50**, 711 (1989).
11. A. V. Bortkevich, S. A. Geĭdur, and O. O. Karapetyan, Zh. Prikl. Spektrosk. **50**, 210 (1989).
12. Yu. V. Korobkin, O. I. Sidorov, and V. B. Studenov, Zh. Tekh. Fiz. **67** (10), 71 (1997) [Tech. Phys. **42**, 1176 (1997)].
13. V. I. Bezrodnyĭ, L. V. Vovk, and N. A. Derevyanko, Kvantovaya Élektron. (Moscow) **22**, 245 (1995).
14. V. I. Bezrodnyĭ, A. A. Ishchenko, L. V. Karabanova, and Yu. L. Slominskiĭ, Kvantovaya Élektron. (Moscow) **22**, 849 (1995).
15. A. A. Ishchenko, Kvantovaya Élektron. (Moscow) **21**, 513 (1994).
16. A. A. Ishchenko, *Structure and Spectral Luminescent Properties of Polymethine Dyes* (Naukova Dumka, Kiev, 1994).
17. A. A. Ishchenko, N. A. Derevyanko, and A. M. Vinogradov, Zh. Obshch. Khim. **67**, 1191 (1997).
18. M. I. Aldoshin, A. A. Manenkov, and V. S. Nechitaĭlo, Zh. Tekh. Fiz. **49**, 2498 (1979) [Sov. Phys. Tech. Phys. **24**, 1412 (1979)].
19. M. I. Aldoshin, B. G. Gerasimov, and A. A. Manenkov, Zh. Tekh. Fiz. **49**, 2496 (1979) [Sov. Phys. Tech. Phys. **24**, 1411 (1979)].
20. R. S. Anderson, R. E. Hermes, and G. A. Matyushin, Proc. SPIE **3265**, 13 (1998).
21. M. D. Rahn, T. A. King, A. A. Gorman, and I. Hamblett, Appl. Opt. **36**, 5862 (1997).
22. T. G. Pavlopoulos, J. H. Boyer, and K. Thangaraj, Appl. Opt. **31**, 7089 (1992).
23. T. G. Pavlopoulos, J. H. Boyer, and G. Sathyamoorthi, Appl. Opt. **37**, 7797 (1998).
24. *Encyclopedia of Polymers* (Sovetskaya Éntsiklopediya, Moscow, 1972), Vol. 2.

Translated by A. Chikishev

Technical Physics, Vol. 46, No. 7, 2001, pp. 864–867. Translated from Zhurnal Tekhnicheskoĭ Fiziki, Vol. 71, No. 7, 2001, pp. 79–82.

RADIOPHYSICS

RF Oscillation Generated by an Electromagnetic Shock Wave in Coupled Transmission Lines with Anomalous and Normal Dispersion

A. M. Belyantsev and A. B. Kozyrev

Institute for Physics of Microstructures, Russian Academy of Sciences, Nizhni Novgorod, 603600 Russia

e-mail: kozyrev@ipm.sci-nnov.ru

Received June 1, 2000; in final form, November 10, 2000

Abstract—The generation of rf oscillation when an electromagnetic shock runs in synchronism with a backward wave in coupled transmission lines with normal and anomalous dispersions and ferrite-related nonlinearity is considered. The oscillation damping can be significantly reduced if a large portion of the rf energy flux is directed to the anomalous-dispersion line, where rf losses are relatively low. *© 2001 MAIK "Nauka/Interperiodica".*

Earlier [1], we have shown that rf oscillation may arise behind the electromagnetic shock (EMS) front in a nonlinear transmission line (TL) with spatial dispersion. The reason is that the EMS front is unstable against an rf wave running in synchronism with it (that is, direct video-to-radio pulse conversion is possible during the interaction of the waves). It has also been found that the duration of the radio pulse and its filling frequency depend on rf losses in the TL. Moreover, it has been noted [2] that rf losses have a minor effect on the damping rate of the shock-generated rf oscillation when the rf energy is rapidly removed from the EMS front at the synchronous excitation of backward waves (backward spatial harmonics). Consequently, wider radio pulses may be generated under such conditions.

If the rf energy in a TL is lost mostly in a saturated nonlinear medium, such as ferrite, the effect of rf losses on the oscillation damping can apparently be suppressed if the portion of the rf energy in the lossy region is decreased. In this work, we show that this possibility can be realized, e.g., in an electrodynamic system like coupled TLs with anomalous and normal dispersions (hereafter, anomalous and normal lines, respectively) where the EMS front synchronously initiates backward waves.

Such a system, namely, a quasi-coaxial TL electrodynamically coupled with a ladder-type slow-wave system through the electric field in the gap is shown in Fig. 1a. The dispersion characteristics of this system that were obtained with a lumped-parameter equivalent circuit (Fig. 1b) [3] are depicted in Fig. 2a. The frequency vs. wavenumber dependence has two branches separated by the stopband.

For the equivalent circuit of the transmission lines in Fig. 1b, the Kirchhoff differential–difference equations for currents and voltages were numerically integrated by the Runge–Kutta method. As follows from the results of simulation, the leading edge of a video pulse applied to the input of the normal line is transformed into the EMS front as the pulse travels along the line [1]. Behind the EMS front, oscillations with a frequency equal to the frequency of the synchronous wave $[\nu_s = \nu_p(\omega)]$ are generated. The number of oscillations grows in proportion to the travel. The oscillations run from the shock front to the TL input and are released on the matched load $R_{in}^{(2)}$ (Fig. 1b). The nonlinearity behind the shock front saturates, and the oscillations propagate as in a linear medium. Typical voltage waveforms in the 100th section of the normal and anomalous TLs $V_{100}^{(1)}/V_s$ and $V_{100}^{(2)}/V_s$, respectively, normalized to the EMS amplitude V_s) are depicted in Fig. 3. Both waveforms contain oscillations at the frequency of synchronous wave and also at a higher frequency because of the synchronism with the forward wave of the higher frequency branch of the dispersion curve.

The field structure of an rf wave traveling in coupled TLs (and, eventually, in the region where the EMS field and the field of the shock-induced rf wave overlap) is conveniently characterized by the ratio $V_\omega^{(2)}/V_\omega^{(1)}$ (where $V_\omega^{(1)}$ and $V_\omega^{(2)}$ are the rf oscillation amplitudes at the nodes of the normal and anomalous lines). In

1063-7842/01/4607-0864$21.00

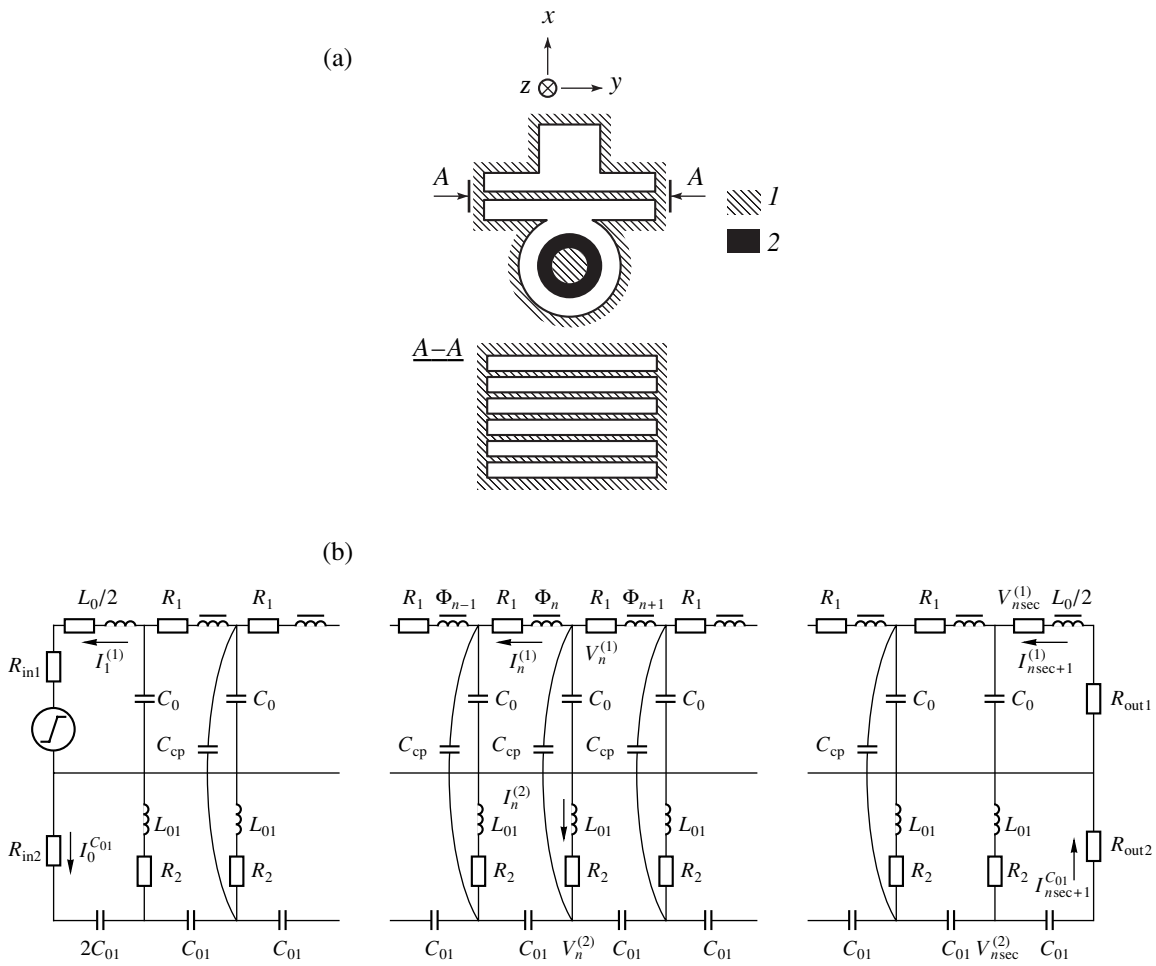


Fig. 1. (a) Coupled quasi-coaxial TL and ladder-type slow-wave system (T-shaped waveguide with grating) and (b) its equivalent circuit. (*1*) Metal; (*2*) ferrite.

Fig. 2b, the ratio $V_\omega^{(2)}/V_\omega^{(1)}$ is plotted against wavenumber φ for several values of the coupling factor C_{cp}/C_0. The basic feature of the voltage waveforms is that the radiation source (shock front) runs in the normal (ferrite-filled coaxial) line, while the major portion of the energy of the synchronous shock-generated wave moves in the anomalous one. In coupled TLs one of which is filled by ferrite, wave damping in the 500–1000 MHz range is largely due to rf losses in the magnetized ferrite, whereas losses due to the skin effect are negligible. Therefore, the damping of the wave generated turns out to be one order of magnitude lower than when the shock wave is in synchronism with a backward spatial harmonic in nonlinear TLs with capacitive cross links (provided that the losses in the ferrite are the same) [2]. Figure 2c plots the calculated parameter N_e against the real part of the wavenumber for the EMS structure formed in the coupled normal and anomalous TLs. Here, N_e is the number of rf oscillations (in the radio pulse generated behind the shock wave) whose amplitude differs from that of the first oscillation (i.e., immediately behind the shock front) by a factor of less than *e*. (Note that the rf oscillation amplitude exponentially decays with distance from the shock front.) The growth of N_e with decreasing coupling coefficient C_{cp}/C_0 fully correlates with the decrease in $V_\omega^{(2)}/V_\omega^{(1)}$.

Note that the possibility of spatially separating the powers of an exciting EMS and an EMS-induced backward wave greatly simplifies the output problem. In this situation, one can isolate the "source" of an input video pulse (at the input of the normal line) and the "receiver" of the oscillations (resistive load at the input of the anomalous one).

Of great importance is the efficiency of rf generation in coupled TLs. This parameter is defined as the ratio of the energy spent to generate oscillations to the total energy delivered to the EMS front. Estimates show that the generation efficiency depends not only on the magnetization reversal rate of the ferrite (as in the case of TLs with capacitive cross links [1, 2]) but also on the

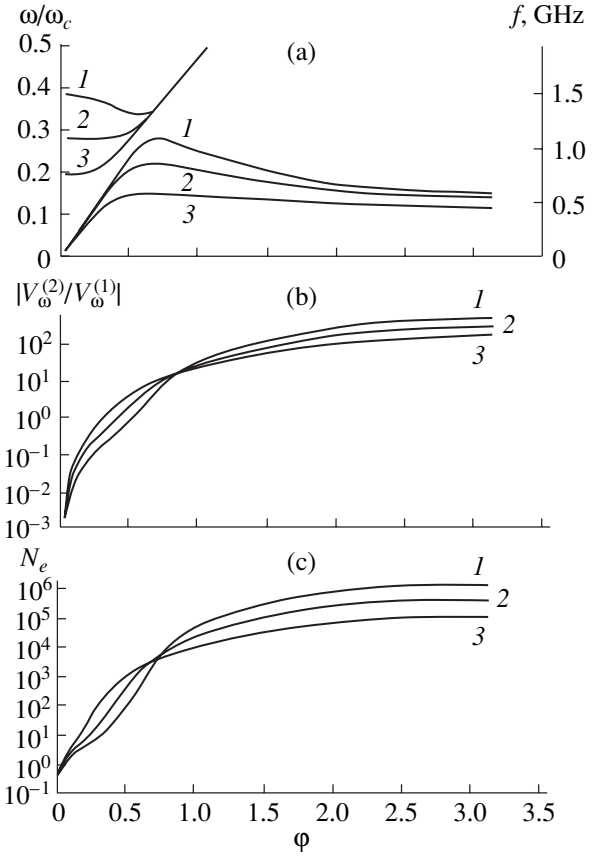


Fig. 2. (a) Relative frequency ω/ω_c, (b) ratio $V_\omega^{(2)}/V_\omega^{(1)}$ of the rf amplitudes in the normal and anomalous lines, and (c) N_e vs. wavenumber φ for the coupling factor C_{cp}/C_0 = (*1*) 0.08, (*2*) 0.16, and (*3*) 0.4. For the definition of N_e, see the text. The parameters of the equivalent circuit are $\omega_c = 2/(L_0C_0)^{1/2} = 2.44 \times 10^{10}$ Hz, $L_{01}/L_0 = 22.5$, and $C_{01}/C_0 = 0.1$ (L_0 is the inductance of an individual cell for saturated nonlinearity).

field structures of the shock and the shock-generated waves (the structures define the interaction efficiency of the waves: the farther the synchronous wave field penetrates into the normal line, the greater the interaction efficiency). If $V_\omega^{(2)}/V_\omega^{(1)} = 1$–10 near the backward wave, the energy efficiency of generation is given by

$$\eta_{\text{eff}} \approx \left(1 - \frac{\tau_f}{T}\right)\frac{1}{1 + V_\omega^{(2)}/V_\omega^{(1)}}.$$

Here, τ_f is the EMS front duration in a dispersionless medium [4]. As follows from the numerical simulation, the energy efficiency of backward wave generation in the case of synchronism corresponding to the minimal the group velocity is about 10%, that is, much smaller

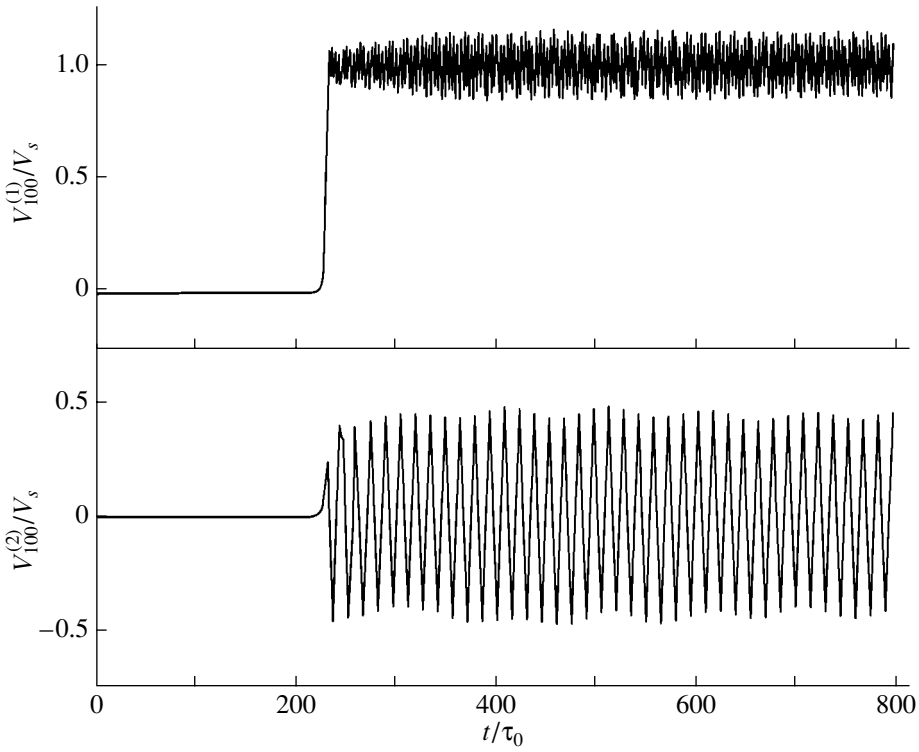


Fig. 3. Voltage waveforms for the 100th cell of the normal ($V_{100}^{(1)}/V_s$) and anomalous ($V_{100}^{(2)}/V_s$) lines (normalized to the shock amplitude). The parameters of the equivalent circuit are $q_0 = \alpha\gamma\tau_0/(1 + \alpha^2) = 0.256$, $\eta = 0.8$, $m_0 = -0.5$, $V_s p/Z_0 M = 8.63905$, $L_{01}/L_0 = 22.5$, $C_{01}/C_0 = 0.1$, and $C_{cp}/C_0 = 0.16$.

than in the case of synchronism with the backward spatial harmonic in TLs with capacitive cross links [2].

To summarize, for the EMS front locked in synchronism with the backward wave in coupled TLs (a ladder-type system and a quasi-coax), where a major part of the energy propagates outside the ferrite, the damping of rf oscillations can be much less than when the EMS is synchronous with the forward wave [1] or a spatial backward harmonic [2]. In view of this circumstance, our work may lay a foundation for designing a device that generates high-power (50–100 MW) rf pulses as long as 10^3 periods with the filling frequency 500–1000 MHz. As in the case when the EMS front excites forward synchronous waves in ferrite-filled TLs [1], the filling frequency can electronically be tuned within a wide range.

ACKNOWLEDGMENTS

This work was supported by the Russian Foundation for Basic Research (project no. 99-02-18046).

REFERENCES

1. A. M. Belyantsev, A. I. Dubnev, S. L. Klimin, *et al.*, Zh. Tekh. Fiz. **65** (8), 132 (1995) [Tech. Phys. **40**, 820 (1995)].
2. A. M. Belyantsev and A. B. Kozyrev, Zh. Tekh. Fiz. **70** (6), 78 (2000) [Tech. Phys. **45**, 747 (2000)].
3. A. M. Belyantsev and A. V. Gaponov, Radiotekh. Élektron. (Moscow) **9** (7), 1188 (1964).
4. A. M. Belyantsev, A. V. Gaponov, and G. I. Freĭdman, Zh. Tekh. Fiz. **35**, 678 (1965) [Sov. Phys. Tech. Phys. **10**, 531 (1965)].

Translated by V. Isaakyan

Technical Physics, Vol. 46, No. 7, 2001, pp. 868–872. Translated from Zhurnal Tekhnicheskoĭ Fiziki, Vol. 71, No. 7, 2001, pp. 83–87.

On the Effect of a Magnetic Field on Fast Charged Particles Passing through a Substance

N. D. Naumov

Received June 14, 2000; in final form, December 13, 2000

Abstract—A transport equation in the small-angle approximation is obtained for a curvilinear beam of fast charged particles passing through a substance in a nonuniform magnetic field. Green functions for this equation are found for an annular beam in a weak-focusing field and for a helical beam in the nonuniform magnetic field.

INTRODUCTION

Theoretical models considering fast charged particles passing through a substance in a magnetic field are of interest in high-energy physics, astrophysics, and physics of the Earth. The use of analytical methods here is validated by the smallness of the single-scattering angle. This greatly simplifies the elastic part of the collision integral in a transport equation. With the collision integral represented in the differential form, one can obtain a number of solutions of a transport equation in the small-angle approximation [1–3].

However, results available have been obtained for the case when charged particles move along the magnetic field, i.e., for a straight beam. For a curvilinear beam, a transport equation in the small-angle approximation has not been derived. For example, the injection of a charged particle beam at an angle to the direction of a uniform magnetic field was studied with a kinetic equation [4–6]. In this work, a transport equation in the small-angle approximation was derived for a charged particle beam passing through a substance in a nonuniform magnetic field. Also, Green functions for this equation were found for annular and helical beams.

CURVILINEAR COORDINATE SYSTEM

The propagation of a particle beam is conveniently considered in a system of curvilinear coordinates (s, η, ζ):

$$\mathbf{x} = \mathbf{Y}(s) + \eta\mathbf{n} + \zeta\mathbf{b},$$

where $\mathbf{Y}(s)$ is the particle trajectory along the beam axis; s is the trajectory length reckoned from the point of injection; and $\mathbf{t}$, $\mathbf{n}$, and $\mathbf{b}$ are the vectors of the Frenet trihedral related to the curve $\mathbf{Y}(s)$.

Although the Frenet trihedral vectors are mutually perpendicular, the coordinate system introduced is not orthogonal, because the scalar products of the basis vectors

$$\mathbf{e}_1 = \frac{\partial\mathbf{x}}{\partial s} = \sigma\mathbf{t} + \kappa(\xi\mathbf{b} - \eta\mathbf{n}),$$

$$\mathbf{e}_2 = \frac{\partial\mathbf{x}}{\partial\eta} = \mathbf{n}, \quad \mathbf{e}_3 = \frac{\partial\mathbf{x}}{\partial\xi} = \mathbf{b}$$

are not all equal to zero. Here, $\sigma = 1 - k\eta$ and k and κ are the curvature and the torsion of the curve $\mathbf{Y}(s)$. The equation of motion for an axial particle in an external magnetic field $B_0 = \mathbf{B}(\mathbf{Y}(s))$ in view of energy losses can be expressed as

$$mv\mathbf{u}' = \frac{e}{c}[\mathbf{v}\mathbf{B}_0] - \varepsilon_0\mathbf{t}. \tag{1}$$

Here, $\mathbf{u} = \gamma v\mathbf{t}$, $\gamma = 1/\sqrt{1 - u^2/c^2}$, $\varepsilon_0 = \varepsilon(W)$, and $W = mc\sqrt{u^2 + c^2} - mc^2$ is the kinetic energy of the axial particle. The function $\varepsilon(W)$ is the stopping power of the substance. For fast charged particles, it is found from the Bethe–Bloch formula [7]. Differentiation with respect to s is designated by prime, and […] in (1) means the vector product.

If the decomposition of the magnetic field vector at the beam axis is used,

$$\mathbf{B}_0 = B_{01}\mathbf{t} + B_{02}\mathbf{n} + B_{03}\mathbf{b},$$

Eq. (1) can be written in the form

$$u'\mathbf{t} + ku\mathbf{n} = \frac{e}{mc}(B_{02}\mathbf{b} - B_{03}\mathbf{n}) - \frac{\varepsilon_0}{mv}\mathbf{t}. \tag{2}$$

Hence, for the trajectory curvature, we have $k = -\rho B_{03}$. Hereafter, we use the designation $\rho = e/mcu$. Next, $\mathbf{n}\mathbf{B}_0 = 0$; i.e., when the charged particle moves, the magnetic field vector lies in the so-called rectifying plane related to the particle trajectory.

Using the inherent time of motion of the axial particle,

$$\tau = \int_0^s \frac{dx}{u(x)},$$

as a longitudinal variable instead of s; we pass to the parametric representation of the particle trajectory. Then, Eq. (1) can be written as

$$m\dot{\mathbf{u}} = \frac{e}{c}[\mathbf{u}\mathbf{B}_0] - \gamma\varepsilon_0\mathbf{t}, \tag{3}$$

where differentiation with respect to τ is designated by dot.

From the definition of the curvature and the torsion of a spatial curve, $k^2 = [\mathbf{u}\dot{\mathbf{u}}]^2/u^6$ and $\kappa = \ddot{\mathbf{u}}[\mathbf{u}\dot{\mathbf{u}}]/k^2u^6$ [8], we can relate these trajectory parameters with the external field and the particle velocity using Eq. (3):

$$k = \frac{\rho}{u}\sqrt{u^2B_0^2 - (\mathbf{u}\mathbf{B}_0)^2},$$

$$\kappa = -\frac{\rho}{u}\left(\mathbf{u}\mathbf{B}_0 + \frac{\rho}{uk^2}\mathbf{B}_0[\mathbf{u}\dot{\mathbf{B}}_0]\right).$$

In particular, $\kappa = -\rho\mathbf{u}\mathbf{B}_0/u$ if the magnetic equation is homogeneous.

TRANSPORT EQUATION

In the presence of an external magnetic field **B**, the transport equation in the continuous slowing-down approximation is given by

$$\boldsymbol{\Omega}\frac{\partial N}{\partial \mathbf{x}} + \frac{e}{c}[\boldsymbol{\Omega}\mathbf{B}]\frac{\partial N}{\partial \mathbf{p}} = \frac{\partial \varepsilon N}{\partial T} + I_{el},$$
$$I_{el} = n_a\int d\boldsymbol{\Omega}'\frac{d\Sigma}{d\boldsymbol{\Omega}'}(T|\boldsymbol{\Omega}', \boldsymbol{\Omega})[N(\mathbf{x}, \boldsymbol{\Omega}', T) - N(\mathbf{x}, \boldsymbol{\Omega}, T)]. \tag{4}$$

Here, $N(\mathbf{x}, \boldsymbol{\Omega}, T)$ is the flux density; e, m, and T are the particle charge, mass, and kinetic energy, respectively; $\mathbf{p} = \boldsymbol{\Omega}\sqrt{T(T + 2mc^2)}/c$; $\varepsilon = \varepsilon(T)$; n_a is the number of atoms per unit volume; and I_{el} is the elastic collision integral, where $d\Sigma(T|\boldsymbol{\Omega}', \boldsymbol{\Omega})/d\boldsymbol{\Omega}'$ is the cross section of elastic scattering from the state $\boldsymbol{\Omega}'$ to the state $\boldsymbol{\Omega}$.

To write the transport equation in the coordinate system considered, one should take advantage of the general expression for a gradient of function in nonorthogonal curvilinear coordinates [8]:

$$\frac{\partial N}{\partial \mathbf{x}} = \frac{\partial N}{\partial s}\mathbf{e}^1 + \frac{\partial N}{\partial \eta}\mathbf{e}^2 + \frac{\partial N}{\partial \zeta}\mathbf{e}^3,$$

where $\mathbf{e}^i$ are the vectors of the reciprocal basis that are calculated from the vectors $\mathbf{e}_i$ and, in our case, have the form

$$\mathbf{e}^1 = \frac{1}{\sigma}\mathbf{t}, \quad \mathbf{e}^2 = \mathbf{n} + \zeta\frac{\kappa}{\sigma}\mathbf{t}, \quad \mathbf{e}^3 = \mathbf{b} - \eta\frac{\kappa}{\sigma}\mathbf{t}.$$

Using the decomposition of the vectors **p** and **B** in the Frenet trihedral vectors,

$$\mathbf{p} = p_1\mathbf{t} + p_2\mathbf{n} + p_3\mathbf{b}, \quad \mathbf{B} = B_1\mathbf{t} + B_2\mathbf{n} + B_3\mathbf{b},$$

and Eq. (4) for flux density, we obtain the equation in the curvilinear coordinates:

$$\frac{1}{\sigma}\frac{\partial N}{\partial s} + \left[p_2\frac{k}{\sigma} + \frac{e}{cp}(p_2B_3 - p_3B_2)\right]\frac{\partial N}{\partial p_1} + \left[\frac{1}{\sigma}(\kappa p_3 - kp_1) + \frac{e}{cp}(p_3B_1 - p_1B_3)\right]\frac{\partial N}{\partial p_2} + \left[\frac{e}{cp}(p_1B_2 - p_2B_1) - p_2\frac{\kappa}{\sigma}\right]\frac{\partial N}{\partial p_3} + \left(\frac{p_2}{p} + \zeta\frac{\kappa}{\sigma}\right)\frac{\partial N}{\partial \eta} + \left(\frac{p_3}{p} - \eta\frac{\kappa}{\sigma}\right)\frac{\partial N}{\partial \zeta} = \frac{\partial \varepsilon N}{\partial T} + I_{el}. \tag{5}$$

SMALL-ANGLE APPROXIMATION

In the small-angle approximation, $\boldsymbol{\Omega} = \mathbf{p}/p \approx \mathbf{t} + \alpha\mathbf{n} + \beta\mathbf{b}$. Therefore, the elastic collision integral can be simplified to the form (if the effect of the external field on collisions is neglected)

$$I_{el} = L(T)N, \quad L(T) = \frac{1}{4}\chi^2(T)(L_1 + L_2),$$

$$L_1 = \frac{\partial^2}{\partial \alpha^2}, \quad L_2 = \frac{\partial^2}{\partial \beta^2}.$$

Here, $\chi^2(T)$ is the mean square of the scattering angle per unit length.

If the ratios of the beam cross size to the radius of curvature and to the radius of torsion are small, i.e., a beam is narrow, the values of $k\eta$, $\kappa\eta$, $\kappa\zeta$ are also small. Moreover, the energy inhomogeneity of the beam will also be small $\tau = T/W - 1$. Therefore, we can put $\chi^2(T) \approx \chi^2(W)$ due to the smallness of the scattering angle. Then, the components of the momentum vector correct to the second order of smallness are given by

$$p_1 = mu(1 + \Gamma\tau), \quad p_2 = mu\alpha, \quad p_3 = mu\beta,$$

where $\Gamma = (W + mc^2)/(W + 2mc^2)$.

For further rearrangements, we should substitute the variables α and β for p_2 and p_3, as well as use the decomposition of the external field in the vicinity of the beam axis:

$$B_1 = B_{01} + H_1, \quad B_2 = H_2, \quad B_3 = B_{03} + H_3,$$

where H_i are first-order quantities.

As a result, Eq. (5) up to first-order terms can be expressed in the form

$$\frac{\partial N}{\partial s} + (\alpha + \kappa\zeta)\frac{\partial N}{\partial \eta} + \left[\lambda\beta - k^2\eta - \rho H_3 - \left(k\tau\Gamma + \alpha\frac{u'}{u}\right)\right]\frac{\partial N}{\partial \alpha} + (\beta - \kappa\eta)\frac{\partial N}{\partial \zeta} + \left(\rho H_2 - \lambda\alpha - \beta\frac{u'}{u}\right)\frac{\partial N}{\partial \beta} = \sigma\frac{\partial \varepsilon N}{\partial T} + \Lambda N, \quad (6)$$

where $\lambda = \kappa + \rho H_1$ and $\Lambda = L(W)$.

At the final stage of transformation of transport equation (6), τ should be substituted for T. Here, we must take into account that $u' = u\Gamma W'/W$ because $u = \sqrt{W(W + 2mc^2)}/mc$. On the other hand, from Eq. (2) it follows that $u' = -\varepsilon_0/mv$. Using the expression for the particle velocity $v = W/mu\Gamma$, we arrive at the natural result: the kinetic energy of axial particles passing through a substance depends on its stopping power $W' = -\varepsilon_0$.

Eventually, for the function $F = \varepsilon N$, we obtain the following equation correct to the second-order terms:

$$\frac{\partial F}{\partial s} + \left[\frac{\varepsilon_0}{W}(\tau + k\eta) - \tau\frac{\partial \varepsilon_0}{\partial W}\right]\frac{\partial F}{\partial \tau} + (\alpha + \kappa\zeta)\frac{\partial F}{\partial \eta} + \left[\lambda\beta - k^2\eta - \rho H_3 - \Gamma\left(k\tau - \alpha\frac{\varepsilon_0}{W}\right)\right]\frac{\partial F}{\partial \alpha} + (\beta - \kappa\eta)\frac{\partial F}{\partial \zeta} + \left(\rho H_2 - \lambda\alpha + \beta\Gamma\frac{\varepsilon_0}{W}\right)\frac{\partial F}{\partial \beta} = \Lambda F. \quad (7)$$

Equation (7) has been derived for a narrow beam of fast charged particles passing through a substance in a nonuniform magnetic field with regard for energy losses and multiple elastic scattering. Obtaining analytical solutions of this equation is a rather complicated task. A simpler case is the transport equation in the small-angle approximation with allowance for multiple elastic scattering only:

$$\frac{\partial N}{\partial s} + (\alpha + \kappa\zeta)\frac{\partial N}{\partial \eta} + (\lambda\beta - k^2\eta - \rho H_3 - k\tau\Gamma)\frac{\partial N}{\partial \alpha} + (\beta - \kappa\eta)\frac{\partial N}{\partial \zeta} + (\rho H_2 - \lambda\alpha)\frac{\partial N}{\partial \beta} = \Lambda N. \quad (8)$$

ANNULAR BEAM

Equation (8) is simplified at $\kappa = 0$, i.e., when the beam axis is a plane curve. As a specific example, let us consider the propagation of an annular beam in a slightly focusing field with the field index q:

$$\mathbf{B} = B_0[qk\zeta\mathbf{n} + (1 + qk\eta)\mathbf{b}].$$

In this case, Eq. (8) can be written as

$$MN = 0, \quad M = \frac{\partial}{\partial s} + M_1 + M_2, \quad (9)$$

where $k_1 = k\sqrt{1 - q}$, $k_2 = k\sqrt{q}$,

$$M_1 = \alpha\frac{\partial}{\partial \eta} - (k_1^2\eta + k\tau\Gamma)\frac{\partial}{\partial \alpha} - \Lambda_1,$$

$$M_2 = \beta\frac{\partial}{\partial \zeta} - k_2^2\zeta\frac{\partial}{\partial \beta} - \Lambda_2.$$

A solution of Eq. (9) can be obtained by the Green function method:

$$\vartheta(s)N(s, X, T) = \int G(X, X_0, s)N(0, X_0, T)dX_0,$$

$$MG(X, X_0, s) = \delta(s)\delta(X - X_0).$$

Here, $\vartheta(x)$ is the Heaviside step function $X = \{x_1, x_2\}$, $x_2 = \{\zeta, \beta\}$, and $x_1 = \{\eta, \alpha\}$. It is easy to see that the Green function has the form

$$G(X, X_0, s) = \vartheta(s)F_1(x_1, x_{10}, s)F_2(x_2, x_{20}, s),$$

where the functions F_i satisfy the equation

$$\left(\frac{\partial}{\partial s} + M_i\right)F_i = 0 \quad (10)$$

with the initial condition $F_i(x_i, x_{i0}, 0) = \delta(x_i - x_{i0})$.

To determine the functions F_i, we should first replace x_i by new variables ξ_i and γ_i:

$$\xi_1 = f_1 - k_1\eta_0, \quad \xi_2 = f_2 - k_2\zeta_0,$$

$$\gamma_1 = g_1 - \alpha_0, \quad \gamma_2 = g_2 - \beta_0,$$

where the functions f_i and g_i are expressed as ($\psi_i = k_i s$)

$$f_1\left(k_1\eta + \frac{k}{k_1}\tau\Gamma\right)\cos\psi_1 - \alpha\sin\psi_1 - \frac{k}{k_1}\tau\Gamma,$$

$$f_2 = k_2\zeta\cos\psi_2 - \beta\sin\psi_2,$$

$$g_1 = \alpha\cos\psi_1 + \left(k_1\eta + \frac{k}{k_1}\tau\Gamma\right)\sin\psi_1,$$

$$g_2 = \beta\cos\psi_2 + k_2\zeta\sin\psi_1.$$

These functions are the integrals of the system of ordinary differential equations

$$\eta' = \alpha, \quad \alpha' = -k_1^2\eta - k\tau\Gamma,$$

$$\zeta' = \beta, \quad \beta' = -k_2^2\zeta.$$

The above change of variables eliminates the terms with first-order derivatives from Eq. (10) (subscript i is

here omitted):

$$\frac{\partial F}{\partial s} - \frac{\chi^2}{4}\left(\sin^2\psi\frac{\partial^2 F}{\partial \xi^2} - \sin 2\psi\frac{\partial^2 F}{\partial \xi \partial \gamma} + \cos^2\psi\frac{\partial^2 F}{\partial \gamma^2}\right) = 0. \quad (11)$$

To solve Eq. (10), one can apply double Fourier transformation with respect to the variables ξ and γ. In doing so, we come to an ordinary differential equation for the Fourier transform. This equation is easily integrable; eventually,

$$F_i = \frac{k_i}{\pi D_i}\exp\left[-\frac{1}{D_i^2}(A_i\xi_i^2 - 2B_i\xi_i\gamma_i + C_i\gamma_i^2)\right],$$

$$A_i = \frac{\chi^2}{2k_i}\left(\psi_i + \frac{1}{2}\sin 2\psi_i\right), \quad B_i = -\frac{\chi^2}{2k_i}\sin^2\psi_i,$$

$$C_i = \frac{\chi^2}{2k_i}\left(\psi_i - \frac{1}{2}\sin 2\psi_i\right), \quad D_i^2 = C_iA_i - B_i^2.$$

In the case of a delta-like source, the expression for the particle concentration is

$$n = \frac{k_1k_2n_0}{\pi\sqrt{C_1C_2}}\exp\left(-\frac{k_1^2\eta^2}{C_1} - \frac{k_2^2\zeta^2}{C_2}\right).$$

Thus, the coefficients C_1 and C_2 characterize a beam widening due to multiple elastic scattering.

HELICAL BEAM

In a similar way, one can find the Green function for a beam in a uniform magnetic field when the beam axis is a spiral line. In this case, the curvature and the torsion of the beam axis do not depend on s: $k = |\nu|\sin\theta$ and $\kappa = \nu\cos\theta$, where $\nu = -\rho B_0$ and θ is the angle between the magnetic field and the direction of beam injection. For a helical beam, Eq. (8) can be expressed in the form

$$\frac{\partial N}{\partial s} + (\alpha + \kappa\zeta)\frac{\partial N}{\partial \eta} + (\beta - \kappa\eta)\frac{\partial N}{\partial \zeta} - k(k\eta + \tau\Gamma)\frac{\partial N}{\partial \alpha} = \Lambda N. \quad (12)$$

In this case, to eliminate the terms with first-order derivatives, one should make the change of variables

$$\xi_1 = (\nu\eta - Q)\cos\psi - (\alpha + \kappa\zeta)\sin\psi + Q - \nu\eta_0.$$

$$\xi_2 = \frac{\zeta}{\nu}(k^2 + \kappa^2\cos\psi) + \kappa\left(\eta - \frac{Q}{\nu}\right)\sin\psi + \alpha\frac{\kappa}{\nu}(\cos\psi - 1) - \nu sR - \nu\zeta_0,$$

$$\xi_3 = \mu\alpha + \frac{k^2}{\nu^2}[(\nu\eta - Q)\sin\psi + \kappa\zeta(\cos\psi - 1)] + \kappa sR - \alpha_0,$$

where

$$\psi = \nu s, \quad \mu = (\kappa^2 + k^2\cos\psi)/\nu^2,$$

$$Q = (\kappa\beta - k\tau\Gamma)/\nu, \quad R = k(k\beta + \kappa\tau\Gamma)/\nu^2.$$

As a result, Eq. (12) becomes similar to Eq. (11):

$$\frac{\partial N}{\partial s} - \frac{1}{4}\sum_{i,j=1}^{3} a_{ij}\frac{\partial^2 N}{\partial \xi_i \partial \xi_j} = 0.$$

Here, for short, we introduced the symmetric matrix of the coefficients $a_{ij}(s)$:

$$a_{11} = \chi^2\left[\sin^2\psi + \frac{\kappa^2}{\nu^2}(\cos\psi - 1)^2\right],$$

$$a_{22} = \frac{\chi^2}{\nu^4}[\kappa^2\nu^2(\sin\psi - 1) + (\kappa^2\sin\psi + \nu sk^2)^2],$$

$$a_{33} = \chi^2\left[\mu^2 + \frac{\kappa^2}{\nu^6}k^4(\sin\psi - \nu s)^2\right],$$

$$a_{12} = \kappa\chi^2\frac{k^2}{\nu^3}(\cos\psi - 1)(\nu s - \sin\psi),$$

$$a_{13} = -\frac{\kappa}{\nu}a_{12} - \mu\chi^2\sin\psi,$$

$$a_{23} = \kappa\frac{\chi^2}{\nu^5}[\mu\nu^4(\cos\psi - 1) + k^2(\sin\psi - \nu s)(\kappa^2\sin\psi + \nu sk^2)].$$

The use of Fourier transformation allows us to obtain the Green function

$$G = \frac{\vartheta(s)\nu^2}{2\pi\sqrt{\pi D}}\exp\left(\frac{1}{D_2}[(A_{22}A_{33} - A_{23}^2)\xi_1^2 + (A_{11}A_{33} - A_{13}^2)\xi_2^2 + (A_{11}A_{22} - A_{12}^2)\xi_3^2 + 2(A_{13}A_{23} + A_{33}A_{12})\xi_1\xi_2 + 2(A_{12}A_{23} + A_{22}A_{13})\xi_1\xi_3 + 2(A_{12}A_{13} + A_{11}A_{23})\xi_2\xi_3]\right),$$

where

$$A_{ij} = \int_0^s a_{ij}(x)dx,$$

$$D^2 = A_{11}A_{22}A_{33} + 2A_{12}A_{13}A_{23} - A_{11}A_{23}^2 - A_{22}A_{13}^2 - A_{33}A_{12}^2.$$

The practical value of the results obtained is that one can estimate the parameters of a beam of fast charged particles passing through a substance in an external magnetic field. These estimates apply until the beam appreciably widens due to the spread of the particle velocity and multiple elastic scattering.

REFERENCES

1. L. R. Kimel' and O. N. Salimov, Zh. Tekh. Fiz. **42**, 1154 (1972) [Sov. Phys. Tech. Phys. **17**, 918 (1972)].
2. N. D. Naumov, Zh. Tekh. Fiz. **62** (2), 178 (1992) [Sov. Phys. Tech. Phys. **37**, 212 (1992)].
3. S. D. Kovalev, A. I. Kuzovlev, and D. B. Rogozkin, Zh. Tekh. Fiz. **63** (6), 37 (1993) [Tech. Phys. **38**, 447 (1993)].
4. A. S. Artamonov and V. A. Gorbunov, Zh. Tekh. Fiz. **50**, 1605 (1980) [Sov. Phys. Tech. Phys. **25**, 936 (1980)].
5. V. S. Remizovich and S. N. Taraskin, Zh. Tekh. Fiz. **51**, 1356 (1981) [Sov. Phys. Tech. Phys. **26**, 778 (1981)].
6. S. P. Andreev and A. V. Koshelkin, Dokl. Akad. Nauk SSSR **289**, 593 (1986) [Sov. Phys. Dokl. **31**, 570 (1986)].
7. V. S. Remizovich, D. B. Rogozkin, and M. I. Ryazanov, *Fluctuations of Charged Particle Ranges* (Énergoatomizdat, Moscow, 1988).
8. V. A. Il'in and É. G. Pozdnyak, in *Foundations of Mathematical Analysis* (Nauka, Moscow, 1973), Part 2.

Translated by M. Astrov

Technical Physics, Vol. 46, No. 7, 2001, pp. 873–876. Translated from Zhurnal Tekhnicheskoĭ Fiziki, Vol. 71, No. 7, 2001, pp. 88–91.

Generation of High-Power Electron Beams in Magnetron Guns with Secondary-Emission Cathodes

Yu. Ya. Volkolupov, A. N. Dovbnya, V. V. Zakutin, M. A. Krasnogolovets, N. G. Reshetnyak, and V. P. Romas'ko

Uskoritel' Research Center

Kharkov Institute of Physics and Technology, National Scientific Center of Ukraine, Akademicheskaya ul. 1, Kharkov, 61108 Ukraine

Received October 10, 2000

Abstract—High-power electrons beams generated in a single injection magnetron gun with secondary-emission cathodes and in a set of such guns are studied. Hollow electron beams of current 50–100 A, electron energy 30–100 kV, and peak power 1–5 MW are obtained. The beams can be used as electron sources in accelerators and ordinary and multibeam high-power microwave devices. © *2001 MAIK "Nauka/Interperiodica".*

INTRODUCTION

In recent years, injection magnetron guns with ordinary [1–6] and inverse [6, 7] secondary-emission cathodes (SECs) have attracted particular interest. These sources offer a number of advantages (long service life, high current density, simple design, hollow beams, etc.), which allow their use as reliable high-power microwave sources in the technology of accelerators [3] and high-speed high-voltage devices [8]. In these guns, the cathode is subjected to backward bombardment by primary electrons (the secondary emission coefficient of the cathode material is above unity), which gain energy when moving in a decaying electric field. The primary electrons can be produced by cold emission, emission from insulating inclusions on the cathode surface, or emission from an additional thermionic cathode [6]. Under these conditions, the electrons are multiplied by secondary emission, their density builds up in an avalanche-like manner, an electron layer is produced near the cathode, and an electron beam is formed and extracted from the gun. At the early stage of electron layer formation, secondary-emission multiplication takes place because the electrons, when following a cycloidal path, gain energy in the decaying electric field. Once the electrons have been accumulated (steady-state stage), the process is governed by electric fields due to the oscillation of the space charge density. Of interest is the stable generation of electron beams with a high peak power in SEC magnetron injection guns. In this work, we study the generation of high-power electron beams in a single gun and in a set of the guns. Also, we trace a correlation between the beam current density and the electric and magnetic fields and evaluate the cross size of the beams.

EXPERIMENT

Experiments were performed with the setup shown in Fig. 1. The magnetron gun is fed by modulator *1*, which forms a voltage pulse of amplitude 4–200 kV, duration 4 μm, and repetition rate 10–50 Hz. A negative-polarity flat-top pulse with an overshoot is applied to copper cathode *5*. Anode *6* (stainless steel or copper) is grounded through the resistor R_3. The secondary emission process is triggered in the decaying electric field produced by the falling edge (duration 0.6 μs, steepness 50–100 kV/μs) of an overshoot that is specially produced on the top of the cathode voltage pulse [4]. The magnetic field is generated by solenoid *4*. The beam current is measured with Faraday cup *7* and resistor R_4; the cathode voltage, with the divider R_1R_2; the anode current, with the resistor R_3 and monitor *2*; and the beam size, with prints on an X-ray film and on a molybdenum foil that are placed on the Faraday cup. The magnetron gun is positioned in stainless-steel vacuum chamber *3*, where a pressure of ~10^{-6} torr is maintained.

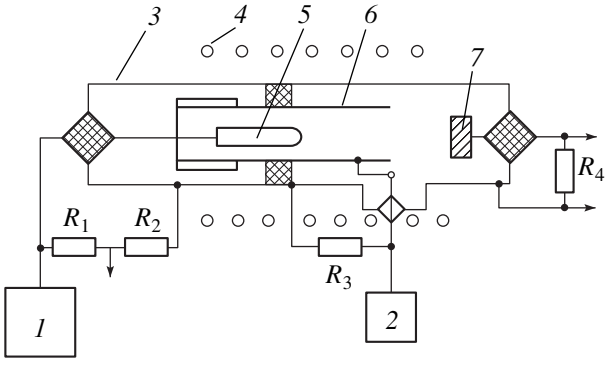


Fig. 1. Experimental setup.

1063-7842/01/4607-0873$21.00 © 2001 MAIK "Nauka/Interperiodica"

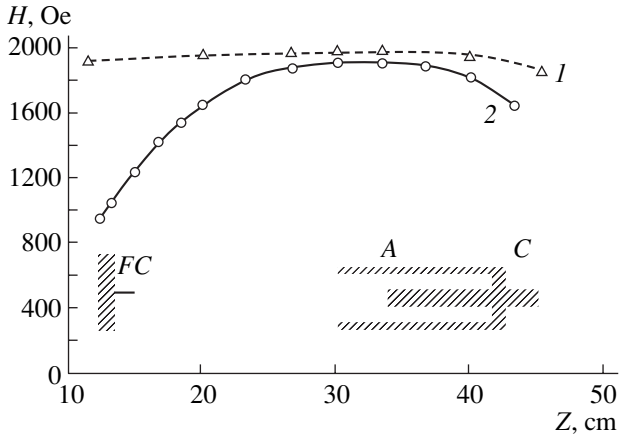


Fig. 2. Longitudinal distribution of the magnetic field. *A*, anode; *C*, cathode; and *FC*, Faraday cup.

In the experiments, two approaches to produce magnetic fields necessary for generating and transporting the beam were employed: a pulsed discharge through a solenoid [9] and dc feeding of a solenoid [10]. Figure 2 shows the distribution of the magnetic field along the solenoid axis for both cases, as well as the arrangement of the magnetron gun and the Faraday cup.

In the former case (Fig. 2, curve *1*), the axial field of the solenoid can be very uniform (±5%) and high (5000–6000 Oe). One, however, should take into account the decay of the pulsed magnetic field when it diffuses through the walls of the cavity and the vacuum chamber. As a result, its longitudinal distribution may change (see, e.g., [9, 11]). The repetition rate of the magnetic field pulses depends on the value of the reservoir capacitor and the switch selected. If the pulse rate is low, the cathode surface is contaminated within the time between the pulses. Due to electron bombardment, the contaminants may fall into the anode–cathode gap, causing its vacuum breakdown [12].

In the latter case, a high density of the feed power and water cooling are necessary to produce a high permanent magnetic field. Here, the field strength is limited by the heat being released in the solenoid and the field distribution is less uniform (Fig. 2, curve *2*). Also, due to current ripple in the solenoid, the repetition rate of pulses driving the modulator must be related to the mains frequency.

RESULTS AND DISCUSSION

1. Generation of High-Power Beams in a Single Magnetron Gun

In this case, high-power beams can be produced if the cathode diameter is large. The beam current (hence, power) can be increased by raising the cathode–anode voltage. However, the voltage cannot exceed some critical value, since the interelectrode gap may break down. The larger the cathode diameter, the smaller the field at the cathode and, hence, the lower the breakdown probability. Therefore, higher voltages can be applied. In our experiments, the cathode diameters were varied from 40 to 80 mm and the anode diameters, from 50 to 140 mm. The voltage amplitudes were between 20 and 120 kV.

The stable current generation mode was achieved in the ~100-mm-long magnetron gun with a cathode diameter of 40 mm and an anode diameter of 78 mm. The voltage pulse amplitude (hereafter, the flat-part amplitude is meant) was 100 kV. The beam current was found to be about 50 A, which corresponds to the microperveance ~1.6 and the peak power ~5 MW. The magnetic field was ~1800 Oe in this case. Figure 3 plots the beam current at the Faraday cup against the cathode voltage pulse amplitude. It is seen that the beam current obeys the Child's law. During measurements, to each voltage value, there corresponded the optimum value of the magnetic field at which the beam current amplitude was the highest. It turned out that the gun readily withstands the voltage 120 kV; hence, the peak power as high as 8 MW can be attained. In the experiments, we also varied the electron current to the anode. The anode current was within 1–10% of the beam current and depended on experimental conditions.

Next, we studied the beam parameters as functions of the electric and magnetic fields. When the cathode voltage is higher or lower than the optimum value (with the magnetic field fixed), the conditions for secondary-emission multiplication are violated and the process ceases. If the top of a voltage pulse is harmonically modulated, electron bunches arise at the exit of the gun. In time, they appear when the sinusoid falls (forced modulation [4]). The beam current vs. magnetic field dependence shows that, as the magnetic field grows (with the cathode amplitude fixed), the beam current at the Faraday cup first sharply grows, exhibits a plateau, and then sharply falls. Such behavior reflects changes in the electron trajectories and in the conditions under which the electrons gain energy in the anode–cathode gap with increasing magnetic field. When the magnetic field amplitude varies, so does the shape of a beam current pulse. At a cathode voltage of 55 kV and a magnetic field of 700 Oe, the beam current pulse breaks down into spikes of amplitude ~30 A and duration 10–30 ns. When the magnetic field grows to ~1200 Oe, the spikes consolidate and the beam current pulse becomes flat.

In the ~100-mm-long gun having a copper cathode of diameter 40 mm and a stainless steel anode of diameter 50 mm, we obtained a beam current of 50 A at a voltage of 30 kV and a magnetic field of 2200 Oe. This corresponds to the microperveance ~10 and the peak power ~1.5 MW.

In [6], single magnetron guns of both ordinary (cathode diameter 50 mm, anode diameter 60 mm) and inverse (cathode diameter 54 mm, anode diameter 43 mm) type generated the beam current ~100 A in a

magnetic field of ~1800 Oe at a voltage of 40 kV (the cathodes were made of the BeCu alloy or stainless steel). This corresponds to the microperveance ~12 and the peak power ~4 MW, which is consistent with our results.

2. Production of Multiple Electron Beams in a Set of Magnetron Guns

Another way of increasing the beam power is the use of a system of parallel-connected magnetron injection guns. The current, power, microperveance, and size of each of the beams are small in this case, but the aggregate beam current and power in the system are significant.

In producing stable electron beams with a system of parallel-connected SEC magnetron beams, difficulties like those with parallel-connected switches may arise. When a partial beam is generated or when one of the guns is broken down, the voltage amplitude at the cathodes of the other guns decreases. This may quench generation or cause beam instability. In actual systems, the voltage decrease at the cathodes of the other guns that is related to the presence of parasitic inductances and capacitances takes place over a period of several nanoseconds. It was shown [3] that the formation time of the electron beam in SEC guns (hence, the rise time of the beam current pulse) may be decreased to 2 ns (if secondary-emission multiplication is triggered by the nanosecond falling edge of the voltage pulse with a steepness of more than 300 kV/μs). Thus, in the system of the guns, the electron layer is produced and the beam is generated even if one of the guns is broken down. If the feeding modulator has a low output resistance, the amplitude of the generation-driving pulse decreases insignificantly. In our experiments, the modulator resistance was 2000 and the decrease was less than 20%. It was demonstrated [14] that beam generation continues when the cathode voltage changes by 30% (with the magnetic field fixed), which exceeds the above value. Thus, the difficulties associated with beam generation in the system of the guns are quite surmountable.

The electrical strength of a system of magnetron guns may be somewhat reduced because of an increase in the total surface area of the electrodes (the breakdown voltage is inversely proportional to $S^{0.1}$, where S if the surface area of the electrodes). If the number of the guns is eight or more, the decrease in the breakdown voltage is 15–20%. In a set of eight guns with the cathode and anode diameters 5 and 22 mm, respectively, the breakdown voltage was shown to diminish roughly by 30% (from 70 to 50 kV) as compared with a single gun with the same electrode geometry. Note that this decrease is not only due to the above effect but also results from the asymmetric arrangement of the guns, poor vacuum conditions, and other adverse factors.

In our experiments, high-power beams were produced in a system of six or eight guns. The cathodes and anodes were made of copper, and the gun length was ~100 mm.

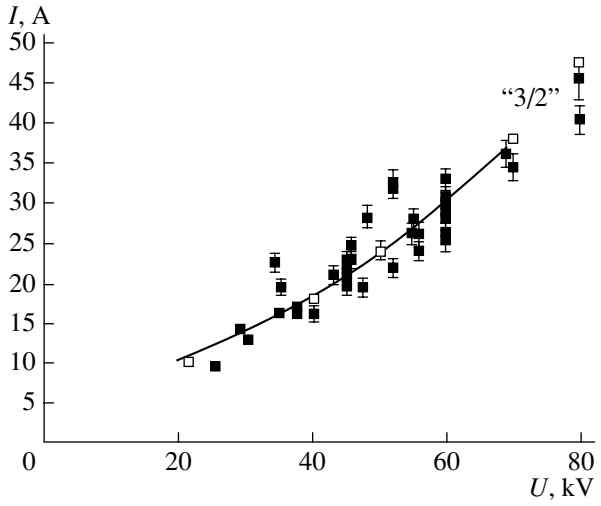


Fig. 3. Beam current vs. cathode voltage pulse amplitude.

In the first case, the system generates six beams. All the guns were arranged in a circle of diameter 60 mm. The anode and cathode diameters were 5 and 26 mm, respectively. For a cathode voltage of 40 kV and a magnetic field strength of ~2000 Oe, each gun generates a beam current of 18 A with a microperveance of ~2.5. The total beam current was 100 A with a peak power of 4 MW.

In the second case, the eight guns were arranged in a circle of diameter ~70 mm. The cathode and anode diameters were 5 and 22 mm, respectively. For a cathode voltage of ~30 kV and a magnetic field strength of ~2000 Oe, the total beam current was ~60 A with a peak power of ~2 MW. The microperveance of each of the beams was ~2.

3. Beam Size

In the SEC guns, the beam traveled a distance of 50–100 mm from the anode surface. It was found that they have a ring-shaped cross section with the uniform azimuth distribution of the intensity. The inner diameter of the ring is roughly equal to the cathode diameter. The "wall" thickness was found to be 1–2 mm.

The outer and inner diameters of the beam produced in the single magnetron gun (with cathode and anode diameters of 40 and 70 mm, respectively) were 45 and 41 mm, respectively. For the magnetron diode (cathode and anode diameters 80 and 140 mm, respectively), it was found that, when the electric field is nonuniform within 5% in the transverse direction, the distinct beam trace is surrounded by the wide diffuse region. For a more nonuniform field, the beam disrupts in the azimuth direction (over a length of about 20 mm). This indicates that the electrode coaxiality must meet stringent requirements.

In the multibeam magnetron system, the beams were uniformly arranged in a circle of the given diameter (60 or 70 mm). The sizes of each of the beams equaled those of the beam in the individual gun (the outer and inner diameters ~9 and ~5 mm, respectively). The intensity of the beams was uniformly distributed in the azimuth direction.

CONCLUSION

Our experiments show that magnetron guns with secondary-emission cathodes can provide high-power electron beams. We established a correlation between the beam current and the electric and magnetic fields. Beam generation in a multiple gun system and its electrical strength were considered.

REFERENCES

1. J. F. Skowron, Proc. IEEE **6** (3), 69 (1973).
2. S. A. Cherenshchikov, Élektron. Tekh., Ser. 1, No. 6, 20 (1973).
3. A. N. Dovbnya, V. V. Zakutin, V. F. Zhiglo, *et al.*, in *Proceedings of the 5th European Particle Accelerator Conference*, Ed. by S. Myers, A. Pacheco, R. Rascual, *et al.* (Inst. of Physics Publ., Bristol, 1996), Vol. 2, p. 1508.
4. A. N. Dovbnya, V. V. Mitrochenko, *et al.*, in *Proceedings of the 1997 Particle Accelerator Conference, Vancouver, 1997*, Ed. by M. Comyn, M. K. Craddok, M. Reiser, and J. Thomson (IEEE, Piscataway, 1997), Vol. 3, p. 2820.
5. A. V. Agafonov, V. P. Tarakanov, and V. M. Fedorov, Vopr. At. Nauki Tekh., Ser. Yad.-Fiz. Issled., Nos. 2, 3 (29, 30), 134 (1997).
6. Y. M. Saveliev, W. Sibbett, and D. M. Parkes, Phys. Plasmas **4**, 2319 (1997).
7. G. I. Churyumov, Izv. Vyssh. Uchebn. Zaved., Radioélektron. **40** (7), 77 (1997).
8. A. I. Vishnevskiĭ, A. I. Soldatenko, and A. I. Shendakov, Izv. Vyssh. Uchebn. Zaved., Radioélektron. **11**, 555 (1968).
9. H. Knoepfel, *Pulsed High Magnetic Fields* (North-Holland, Amsterdam, 1970; Mir, Moscow, 1972).
10. I. I. Kifer, *Ferromagnetic Materials Testing* (Gosénergoizdat, Moscow, 1962).
11. V. V. Zakutin and A. M. Shenderovich, Prib. Tekh. Éksp., No. 5, 238 (1976).
12. A. N. Dovbnya, V. V. Zakutin, N. G. Reshetnyak, *et al.*, Vopr. At. Nauki Tekh., Ser. Yad.-Fiz. Issled., No. 1 (28), 53 (1997).
13. N. I. Aĭzatskiĭ, A. N. Dovbnya, V. V. Mitrochenko, *et al.*, Vopr. At. Nauki Tekh., Ser. Yad.-Fiz. Issled., No. 3 (34), 38 (1999).
14. V. V. Zakutin, A. N. Dovbnya, N. G. Reshetnyak, *et al.*, Zh. Tekh. Fiz. **71** (3), 78 (2001) [Tech. Phys. **46**, 348 (2001)].

Translated by V. Isaakyan

Technical Physics, Vol. 46, No. 7, 2001, pp. 877–882. Translated from Zhurnal Tekhnicheskoĭ Fiziki, Vol. 71, No. 7, 2001, pp. 92–97.

Focusing of Atomic and Molecular Beams in Electric Fields of Various Configuration

V. V. Smirnov

Research Institute of Physics, St. Petersburg State University, St. Petersburg, 198904 Russia

e-mail: smir@paloma.spbu.ru

Received October 4, 2000

Abstract—The properties of an axisymmetric electrostatic lens used for focusing atomic or molecular beams were studied. The lens was formed by a dip in an electric field in the axial direction. Two types of interaction between the particles and the electric field were studied: quadratic and linear in field. An analytical approximation of the dependence of the focal distance of the lens on the beam energy and the parameters of the electric field was obtained. Chromatic and spherical aberrations of the lens were determined. © *2001 MAIK "Nauka/Interperiodica".*

INTRODUCTION

Such systems as a Fresnel zone plate [1] or a macroscopic electromagnetic lens are often used in atomic optics for focusing atomic beams. In the latter case, either focused laser radiation [2] or a magnetic lens [3] (as in electron microscopy) is used. The magnetic lens focusing ensures maximum resolution. The limiting resolution is determined by the aberrations of the focusing lens (mainly, by a chromatic aberration). It was suggested that, if an atomic beam were cooled, a resolution of about 50 nm could be attained, provided that, under otherwise identical conditions, the diffraction limit of the system was 6.5 nm [3]. It was also suggested in the preceding works [4, 5] that an object of atomic size could be used as a lens for focusing particle beams.

Electric fields of axisymmetric configuration focusing atomic or molecular beams are also of interest. In this work, the properties of such lenses are studied theoretically.

INTERACTION QUADRATIC IN FIELD

The interaction of an atom or molecule with an electric field **E** significantly weaker than the atomic field is described by the formula [6]

$$V = -\frac{1}{2}\alpha E^2, \tag{1}$$

where α is the polarizability of the atom or molecule.

This interaction is responsible for the quadratic Stark effect in the atom.

It is well known that the field potential U near the axis of a cylindrical lens has the form

$$U(r, z) = U_0(z) - \frac{r^2}{4}U_0''(z), \tag{2}$$

where U_0 is the field potential on the axis and (r, z) are cylindrical coordinates.

The field strength is

$$E_z(r, z) = E_0(z) - \frac{r^2}{4}E_0''(z),\ E_r(r, z) = -\frac{r}{2}E_0'(z),$$

where $E_0 = -U_0'$.

The potential determined by Eq. (2) should satisfy the following equation:

$$\Delta U(r, z) = \frac{r^2}{4}E_0'''(z).$$

Therefore, this potential determines the approximate solution of the Laplace equation subject to the condition

$$r \ll 2\sqrt{\left|\frac{E_0'}{E_0'''}\right|}. \tag{3}$$

Let the field distribution along the axis be expressed in terms of a dimensionless profile $f(z)$ of the distribution: $E_0(z) = E_e f(z)$. Thus, Eq. (1) for the interaction potential can be written as

$$V(r, z) = -\frac{1}{2}\alpha E_e^2 g(r, z),$$
$$g = \left(f - \frac{r^2}{4}f''\right)^2 + \frac{r^2}{4}f'^2. \tag{4}$$

The values of αE_e^2 and g specify the scale of interaction and spatial distribution of the interaction potential, respectively. On the axis, $g(0, z) = f(z)^2$. The movement of the atom is determined by the force applied to it, i.e., by the derivative of the potential. Therefore, the zero potential level can be selected arbitrarily. Let the

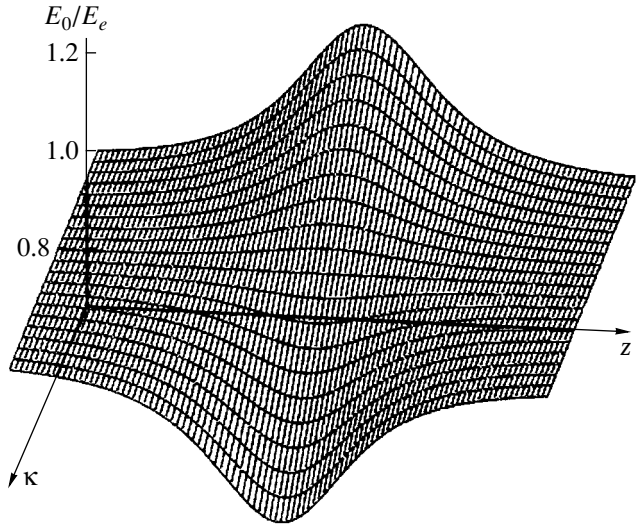


Fig. 1. Family of distributions of the relative strength $E_0(z)/E_e$ of the electric field along the lens axis for various values of the parameter κ (–0.3 < κ < 0.3).

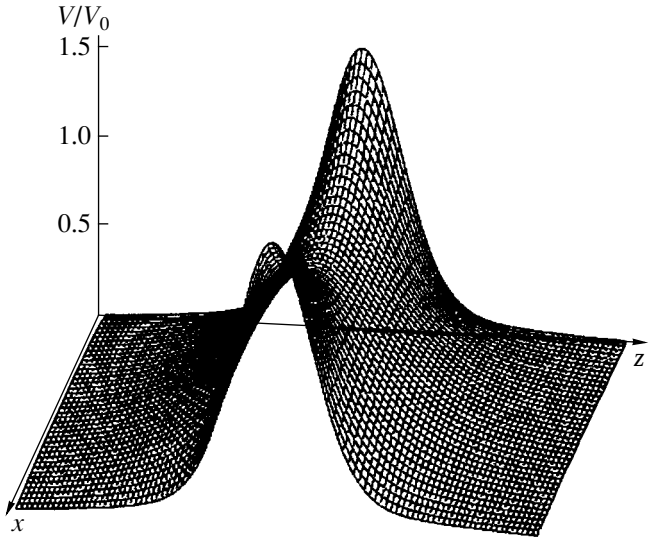


Fig. 2. Spatial distribution of the relative value $V(x, y, z)/V_0$ of the quadratic-interaction potential (1) in the plane $y = 0$ for $\kappa = 0.2$ ($r = |x|$).

potential be equal to zero for atoms incident on the lens along its axis. In this case, $g \longrightarrow g - f(-\infty)^2$.

Let us consider a family of field distributions along the axis with a dip near the origin of coordinates. These distributions can be described using the approximation

$$f = \left(1 - \kappa S^{-\frac{3}{2}}\right), \quad S = 1 + \left(\frac{z}{R}\right)^2, \tag{5}$$

where κ is the parameter specifying the relative value of the field dip near the origin of coordinates and R is the length of the region of field variation.

For $0 < \kappa < 1$, the field distribution has a minimum (dip) near the center of the lens; for other values of κ, there is a maximum (hump) near the center of the lens. The family of distributions of the relative strength of the electric field on the lens axis for various values of the parameter κ (–0.3 < κ < 0.3) is shown in Fig. 1. Electric fields of such a configuration can be generated using appropriate electrodes. It should be noted that the form of approximation selected above is less important than its qualitative behavior. Within the framework of approximation (5), the condition (3) is reduced to

$$\frac{r}{R} \ll \frac{2S}{\sqrt{5|4S-7|}},$$

i.e., virtually to $r \ll R$. Therefore, the axial length R of the field dip specifies the upper limit of the lens radius.

The interaction potential (1) at the center of a lens formed by an electric field of configuration (5) is described by the equation

$$V_0 = \frac{1}{2}\alpha E_e^2 \kappa(2-\kappa). \tag{6}$$

The spatial distribution of the relative value V/V_0 of the shifted interaction potential (1) for $\kappa = 0.2$ is shown in Fig. 2. For $0 < \kappa < 2$, the distribution of the interaction potential along the axis of the lens has the shape of a barrier. On the axis, the barrier height is V_0. For $z = 0$, the radial distribution of the barrier height is

$$V(r, 0) = \frac{1}{2}\alpha E_e^2 \kappa'(2-\kappa'), \quad \kappa' = \kappa\left(1 + \frac{3}{4}\left(\frac{r}{R}\right)^2\right),$$

so that the ratio between the barrier heights at the

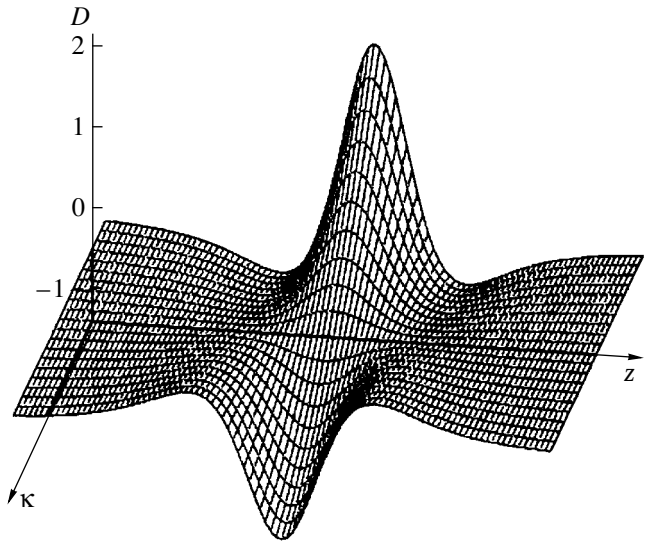


Fig. 3. Family of distributions of the coefficient $D(z)$ along the lens axis for various values of the parameter κ (–0.3 < κ < 0.3).

boundary and at the center is

$$\frac{V(R,0)}{V_0} = \frac{7}{4}\frac{(2-7\kappa/4)}{(2-\kappa)}.$$

Thus, if $0 < \kappa < 2$, the electrostatic lens filters the atomic beam by energy (atoms with energies lower than the barrier height are stopped). The force exerted on each atom is

$$\mathbf{F} = -\frac{\partial V}{\partial \mathbf{r}} = \frac{1}{2}\alpha E_e^2 \frac{\partial g}{\partial \mathbf{r}},$$

or, in the componentwise form,

$$\begin{bmatrix} F_r \\ F_z \end{bmatrix} = \frac{1}{2}\alpha E_e^2 \begin{bmatrix} \frac{r}{2}(f'^2 - 2f\cdot f'') + \frac{r^3}{4}f''^2 \\ 2f\cdot f' - \frac{r^2}{2}f\cdot f''' + \frac{r^4}{8}f''\cdot f''' \end{bmatrix}.$$

It can be seen that the atoms are attracted to the axis of the lens if the coefficient $D(z) = f'^2 - 2f\cdot f''$ is negative.

In the approximation selected, the coefficient $D(0)$ at the center of the lens is equal to $2\kappa(\kappa - 1)$. Therefore, atoms are attracted to the axis of the lens near its center if $0 < \kappa < 1$, i.e., if the axial distribution of the field strength has its minimum at the center of the lens. The family of the distribution curves of the coefficient D along the lens axis for various values of the parameter κ ($-0.3 < \kappa < 0.3$) is shown in Fig. 3. The coefficient D specifies the radial component of the force applied to each atom. It is seen from Fig. 3 that regions of attraction alternate with regions of repulsion. This makes the radial motion of atoms through the lens nonmonotonic, which can be seen from the atomic trajectories shown in Fig. 4.

If $0 < \kappa < 0.1$, the atomic beam is focused (in the paraxial approximation) at all energies E higher than V_0. If $\kappa > 0.1$, the atomic beam is focused only at energies lower than a certain critical value; at energies higher than this value, the atomic beam is defocused.

The relative focal distances and values of the spherical and chromatic aberrations of the electrostatic lens providing focusing of atomic beams were calculated from the atomic trajectories (Tables 1–3).

For $0 < \kappa < 0.3$, the dependence of the focal distance F on the atomic energy E in the focal region can be approximated by the expression

$$\frac{F}{R} \cong a\left(\frac{E}{V_0}\right)\left(\frac{E}{V_0} - 1\right)^b b^{1+b}, \qquad (7)$$

where $a = 3.11$, $b = 1 + 4.02\kappa$, and V_0 is determined by Eq. (6).

Equation (7) was derived analytically from the theoretical expression for the focal distance. The approximation parameters were determined by interpolation of the data obtained. Equation (7) describes the chromatic aberration of the lens.

Table 1. Dependence of the focal distance F on the radial position r of the atom trajectory near the lens (impact parameter)

r/R	F/R		
	$\kappa = 0.03$	$\kappa = 0.1$	$\kappa = 0.3$
0.05	9.95	12.93	25.03
0.10	9.87	12.87	24.78
0.15	9.71	12.60	24.38
0.20	9.47	12.32	23.79
0.25	9.18	11.95	23.05
0.30	8.85	11.48	22.21
0.35	8.46	11.01	21.27
0.40	8.03	10.46	20.18
0.45	7.58	9.87	19.09
0.50	7.10	9.25	17.90
0.55	6.59	8.58	16.70
0.60	6.08	7.95	15.46
0.65	5.55	7.27	14.21
0.70	5.00	6.59	12.90
0.75	4.43	5.89	11.68
0.80	3.85	5.18	10.41
0.85	3.24	4.44	9.17
0.90	2.60	3.67	7.91
0.95	1.84	2.88	6.68

Note: Relative energy of atoms E/V_0 is 2. The results obtained for $\kappa = 0.1$ correspond to atomic trajectories shown in Fig. 3.

INTERACTION LINEAR IN FIELD

The polarizability of polar molecules is due mainly to electron and orientational polarization [7]. In weak

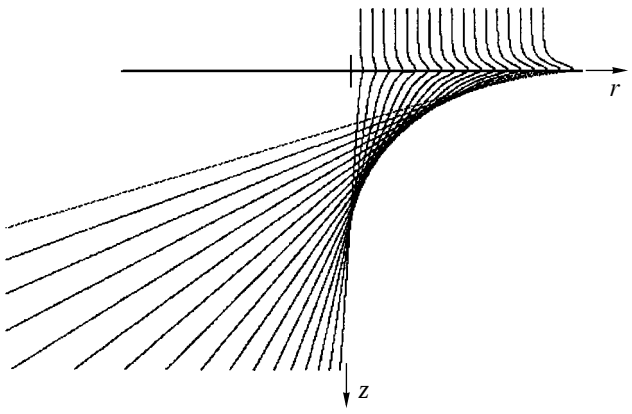


Fig. 4. Trajectories of atoms with a relative energy $E/V_0 = 2$ in the quadratic potential field described by Eq. (1) for $\kappa = 0.2$. Vertical and horizontal scale bars at the center are equal to $2R$ each.

Table 2. Dependence of the focal distance F on the relative energy of atoms for various values of the parameter κ

E/V_0	F/R			
	$\kappa = 0.03$	$\kappa = 0.1$	$\kappa = 0.2$	$\kappa = 0.3$
1.4				6.1
1.5	3.8	4.4	5.9	9.8
1.6				16.0
1.7			10.8	27.3
1.8			14.4	53.9
2.0	10.0	12.9	25.0	
2.5	19.2	28.6		
3.0	31.9	55.8		

fields, both polarization mechanisms cause the interaction described by Eq. (1). In stronger fields (though weaker than atomic fields), the orientational polarization reaches the saturation value. In many cases, the electron polarizability of molecules with a large dipole moment is small and can be neglected.

The interaction of a molecular dipole $\boldsymbol{\mu}$ with an electrostatic field $\mathbf{E}$ is described by the following equation:

$$V = -\boldsymbol{\mu} \cdot \mathbf{E}. \tag{8}$$

Movement of a molecule through the lens can be described in various approximations. Movement of a molecular dipole can be regarded as a combination of rotational and translational motion with further averaging of the rotational component according to the initial conditions. For simplicity's sake, the movement of the molecule is considered in this work as a translational motion in the rotation-averaged potential field described by Eq. (8).

The interaction potential (8) is averaged over the angular coordinate using the Boltzmann distribution [7]:

$$\bar{V} = \frac{-\int \mu E \cos\Theta \exp\left(\frac{\mu E \cos\Theta}{kT}\right) d\cos\Theta}{\int \exp\left(\frac{\mu E \cos\Theta}{kT}\right) d\cos\Theta} = -\bar{\mu} E. \tag{9}$$

Thus, the average value of the dipole moment projection on the field direction is

$$\bar{\mu} = \mu\left(\frac{\cosh\eta}{\sinh\eta} - \frac{1}{\eta}\right), \quad \eta = \frac{\mu E}{kT}.$$

For small values of η ($\eta \ll 1$), we obtain that $\bar{\mu} = \mu^2 E/2kT$; i.e., an averaged interaction potential of the form (1) quadratic in field with orientational polarizability $\alpha = \mu^2/kT$ is obtained. It should be noted that the exponential function under the integral in Eq. (9) is often replaced by its expansion for small η ($\eta \gg 1$) [7]. In this case, a somewhat different averaged value of the dipole moment is obtained: $\bar{\mu} = \mu^2 E/3kT$. This estimation seems to be less consistent.

For large values of η ($\eta \gg 1$), we obtain that

$$\bar{\mu} = \mu\left(1 - \frac{1}{\eta}\right),$$

i.e., the averaged moment tends to the maximum value of μ. The averaged interaction is linear in field:

$$\bar{V} = -\mu E + kT. \tag{10}$$

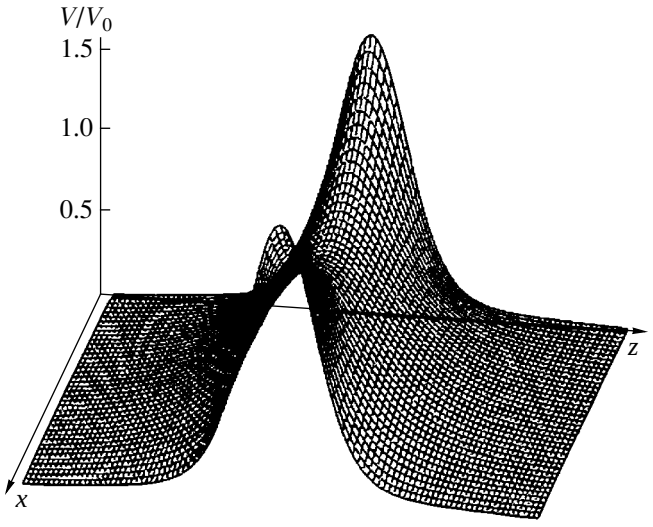


Fig. 5. Spatial distribution of the relative value $V(x, y, z)/V_0$ of the linear-interaction potential in the plane $y = 0$ for $\kappa = 0.2$ ($r = |x|$).

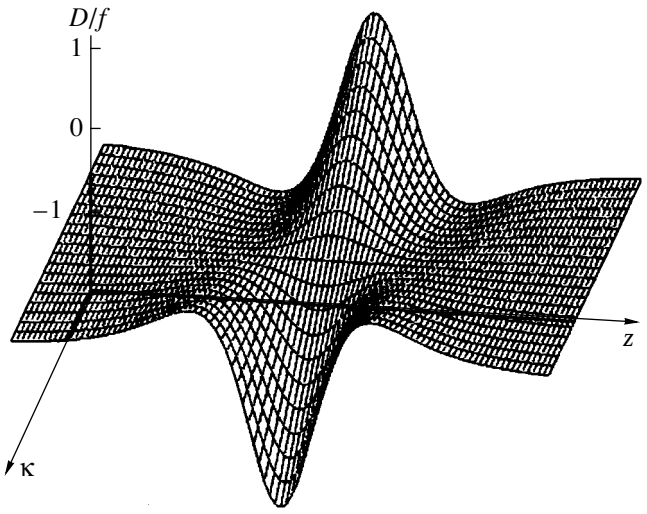


Fig. 6. Family of distributions of the coefficient $D(z)/f$ along the lens axis. The coefficient $D(z)/f$ specifies the radial force component in the case of a linear interaction.

Table 3. Dependence of the coefficients of the third-order and fifth-order spherical aberrations (S_3 and S_5, respectively) on the relative energy of atoms for various values of the parameter κ

E/V_0	S_3R^2				S_5R^4			
	κ = 0.03	κ = 0.1	κ = 0.2	κ = 0.3	κ = 0.03	κ = 0.1	κ = 0.2	κ = 0.3
1.4				1.09				19.5
1.5	0.94	1.17	1.33	1.40	13.6	11.6	9.56	8.09
1.6				1.16				5.74
1.7			0.70	0.93			7.05	5.17
1.8			0.91	1.07			4.25	2.97
2.0	0.89	0.93	1.02		2.74	2.40	1.97	
2.5	0.85	0.87			1.10	1.05		
3.0	0.80	0.78			0.76	0.76		

Table 4. Dependence of the focal distance F on the radial position r of the molecule trajectory

r/R	F/R		
	κ = 0.1	κ = 0.2	κ = 0.3
0.05	10.03	11.34	13.43
0.10	9.93	11.22	13.25
0.15	9.74	10.98	12.93
0.20	9.49	10.66	12.48
0.25	9.18	10.27	11.95
0.30	8.80	9.81	11.33
0.35	8.38	9.29	10.67
0.40	7.93	8.75	9.95
0.45	7.45	8.16	9.23
0.50	6.94	7.56	8.48
0.55	6.42	6.94	7.72
0.60	5.88	6.33	6.98
0.65	5.33	5.69	6.24
0.70	4.76	5.06	5.51
0.75	4.20	4.44	4.80
0.80	3.60	3.80	4.06
0.85	2.99	3.12	3.32
0.90	2.33	2.41	2.52
0.95	1.55	1.57	1.63

Note: Relative energy of molecules E/V_0 is 2.

Table 5. Dependence of the focal distance F on the relative energy of molecules for various values of the parameter κ

E/V_0	F/R		
	κ = 0.1	κ = 0.2	κ = 0.3
1.5	3.76	4.01	4.36
2.0	10.03	11.34	13.43
2.5		23.50	
3.0	32.59	42.10	64.03
3.5	49.67		

Table 6. Dependence of the coefficients of the third-order and fifth-order spherical aberrations (S_3 and S_5, respectively) on the relative energy of molecules for various values of the parameter κ

E/V_0	S_4R^2			S_5R^4		
	κ = 0.1	κ = 0.2	κ = 0.3	κ = 0.1	κ = 0.2	κ = 0.3
1.5	1.14	1.17	1.08	14.40	15.50	17.48
2.0	0.83	0.92	1.10	3.98	4.52	5.19
2.5		1.23			1.42	
3.0	0.97	1.25	1.91	0.84	1.06	1.46
3.5	0.94			0.72		

If the electron polarizability α_e of the molecule is taken into account, the averaged interaction takes the form

$$\bar{V} = -\frac{1}{2}\alpha_e E^2 - \mu E + kT.$$

The last term kT has no effect on the molecule movement and can be neglected.

If the field configuration is the same as for the quadratic interaction (4), the linear interaction (10) takes the form

$$\bar{V} = -\mu E_e \sqrt{g}$$

(the interaction energy at infinity is taken to be zero, i.e., $\sqrt{g} \longrightarrow \sqrt{g} - f(-\infty)$). The height of the barrier at the center of the lens is

$$V_0 = \mu E_e \kappa. \quad (11)$$

The spatial distribution of the relative value V/V_0 of the shifted linear-interaction potential (10) for $\kappa = 0.2$ is shown in Fig. 5.

Thus, the force applied to the molecule is

$$\bar{\mathbf{F}} = -\frac{\mu}{2E}\frac{\partial E^2}{\partial \mathbf{r}} = \frac{\mu E_e}{2\sqrt{g}}\frac{\partial g}{\partial \mathbf{r}}$$

(the value of g in the denominator is not shifted).

Therefore, focusing of particle beams in the case of interaction linear in field should be performed under almost the same conditions as in the case of interaction quadratic in field. In the paraxial approximation, focusing is specified by the coefficient D/f. The family of distributions of the coefficient D/f along the lens axis for various values of the parameter κ ($-0.3 < \kappa < 0.3$) is shown in Fig. 6.

Focusing at energies $E > V_0$ takes place within a wider range of values of the parameter κ ($0 < \kappa < 0.25$) than in the case of quadratic interaction (1).

Comparison of Figs. 3 and 6 shows that the focusing power of the electrostatic lens is higher in the case of interaction linear in field.

Relative focal distances and values of aberrations of the electrostatic lens focusing a molecular beam are given in Tables 4–6.

The dependence of the focal distance F on the atomic energy in the focal region for various values of the parameter κ ($0 < \kappa < 0.3$) can be approximated by Eq. (7) with $a = 3.80$, $b = 1 + 1.24\kappa$, and V_0 determined by Eq. (11).

CONCLUSION

The properties of an axisymmetric electrostatic lens used for focusing atomic or molecular beams were analyzed in this work. The lens was formed by an axial dip of an electric field. Two types of interaction between the particles and the electric field were considered: quadratic and linear in field. An analytical approximation (7) of the dependence of the focal distance on the beam energy and the parameters of the electric field was obtained. As follows from this dependence, the chromatic aberration of the lens is smaller in the case of interaction linear in field (parameter b is smaller). The third-order relative spherical aberration $S_3R^2 \sim 1$ depends only slightly on the type of interaction and parameters of the electric field and the particle beam. The fifth-order relative spherical aberration S_5R^4 depends only slightly on the type of interaction and parameters of the electric field, but increases with a decrease in the beam energy.

ACKNOWLEDGMENTS

This work was supported by the Russian Foundation for Basic Research, project no. 00-03-32920.

REFERENCES

1. O. Carmal, M. Sigel, T. Sleator, *et al.*, Phys. Rev. Lett. **67**, 3231 (1991).
2. J. E. Bjorkholm, R. R. Freeman, A. Ashkin, and D. B. Pearson, Phys. Rev. Lett. **41**, 1361 (1978).
3. W. G. Kaenders, F. Lison, A. Richter, *et al.*, Nature **375**, 214 (1995).
4. O. A. Vorob'ev and V. V. Smirnov, in *Proceedings of the International Conference of Young Scientists and Specialists "Optics-99," St. Petersburg, 1999*, p. 9.
5. V. V. Smirnov, J. Phys. D **31** (13), 1548 (1998).
6. A. A. Radtsig and B. M. Smirnov, *Reference Data on Atoms, Molecules, and Ions* (Atomizdat, Moscow, 1980; Springer-Verlag, Berlin, 1985).
7. J. D. Jackson, *Classical Electrodynamics* (Wiley, New York, 1962; Mir, Moscow, 1963).

Translated by K. Chamorovskiĭ

Technical Physics, Vol. 46, No. 7, 2001, pp. 883–891. Translated from Zhurnal Tekhnicheskoĭ Fiziki, Vol. 71, No. 7, 2001, pp. 98–105.

Application of the Matrizant Method to Calculate the Third-Order Aberration Coefficients for a Sector Magnetic Field Including Fringing-Field Effects

S. N. Mordik and A. G. Ponomarev

Institute of Applied Physics, National Academy of Sciences of Ukraine, ul. Petropavlovskaya 58, Sumy, 40030 Ukraine

e-mail: iapuas@gluk.apc.org

Received October 11, 2000

Abstract—The matrizant method is used to study the nonlinear dynamics of charged particles in magnetic sector analyzers. The calculations of matrizants (transfer matrices) allow for both fringing-field effects due to the stray field and higher harmonics of the sector magnetic field (up to the third order). For the rectangular distribution of the field components along the optical axis, analytical expressions for the aberration coefficients (including dispersion ones) are derived up to the third order. In the simulation of real fields with a nonzero stray-field width, the smooth distribution of the field components is employed. The aberration coefficients for this distribution were calculated by means of a conservative numerical method. *© 2001 MAIK "Nauka/Interperiodica".*

Among the approaches to solving the problems of nonlinear beam dynamics, rigorously conservative (providing the conservation of the phase-space volume of the beam at each step of calculation) matrix methods, including the matrizant method [1–6], for the calculation of ion–optical systems are worth noting. Conservatism is of particular importance for studying the nonlinear dynamics of the phase set in extended systems involving several ion–optical elements. The mathematical rigor of the matrizant method allows us to avoid controversial assumptions when calculating ion–optical systems. Another advantage of the method is the simplicity of algorithmization. This makes it possible to employ modern software for analytically finding solutions. Recently, numerical codes based on the matrizant method have been developed for the optimization of probe-forming systems. These methods allow one to solve the problem of nonlinear beam dynamics for dispersion-free rectangular-axis media in the Cartesian coordinate system for the upper triangular matrix of the coefficients $P^{(3)}$ [2, 3]. In this paper, the vector of the dispersion phase moments is introduced to study the dispersion properties of sector magnetic systems. In the case of a real sector magnetic field, the matrix of the $P^{(3)}$ coefficients that is obtained by the method of embedding in the phase-moment space has an upper triangular form in the conventional curvilinear orthogonal coordinate system [7] for both the coordinate phase moments and the dispersion phase moments. In order to take into account the influence of the fringing-field effects on the dynamics of charged particle beams, two models of the longitudinal distribution of the magnetic field (rectangular and smooth) are considered. The rectangular model implies that the stray-field region is absent. In contrast to the model of a short cutoff fringing field (SCFF) [8], the magnetic field strength in the rectangular model is described by a step function (its first-order derivative is equal to the delta function). This model allows one to take into account the influence of the fringing-field effects on the beam dynamics (in particular, in mass analyzers with magnetic shields), as well as to study the possibility of solutions for the smooth model converging to those for the rectangular one when the stray-field width tends to zero. The smooth field model is introduced to approximate the field distribution with a sufficient accuracy in order to take into account the effect of the fringing stray fields on the beam dynamics in an ion–optical system. Unlike the well-known third-order transfer matrices of magnetic sector analyzers [9, 10], the matrizants obtained in this paper make it possible to include the fringing effects for both the rectangular and the smooth field models. This provides the conservatism at each step of calculations.

We introduce a natural coordinate system (x, y, s) that is related to a plane curve uniquely defined by a constant radius of curvature ρ. This coordinate system totally coincides with that employed by Brown [7]. A Cartesian coordinate system ($\tilde{x}$, $\tilde{y}$, $\tilde{z}$) with the origin at the point where an axial particle starts moving and our coordinate system with the origin at the center of the radius of curvature for a reference particle are related as

$$\tilde{x} = (x + \rho)\cos(s/\rho) - \rho,$$

$$\tilde{y} = y, \quad \tilde{z} = (x + \rho)\sin(s/\rho).$$

1063-7842/01/4607-0883$21.00

Consider the nonrelativistic motion of a particle in the coordinate system chosen. In view of the fact that the Lamé coefficients in this system are $h_1 = 1$, $h_2 = 1$, and $h_3 = 1 + x/\rho$, the trajectory equations can be represented in the form [11]

$$\begin{aligned} x'' + \frac{JT'}{\vartheta}x' - \frac{h_3}{\rho} &= \frac{qT'}{m\vartheta}(y'B_s - h_3B_y),\\ y'' &= \frac{JT'}{\vartheta}y' = \frac{qT'}{m\vartheta}(h_3B_x - x'B_s),\\ J &= \frac{d}{ds}\left(\frac{\vartheta}{T'}\right) = \frac{q}{mh_3}(x'B_y - y'B_x) - \frac{2\vartheta x'}{T'h_3\rho},\\ T' &= \sqrt{h_3^2 + x'^2 + y'^2}, \end{aligned} \tag{1}$$

where prime denotes differentiation with respect to s; T is the magnitude of a trajectory length element for a simultaneous increment in all three coordinates; and ϑ, m, and q are the constant velocity, mass, and charge of the particle, respectively.

In view of the symmetry condition $V(x, y, s) = -V(x, -y, s)$, the expansions of the scalar magnetic potential $V(x, y, s)$ and the magnetic field induction $\mathbf{B}(x, y, s)$ correct to the third order in the vicinity of the axial trajectory are given by

$$\begin{aligned} -V(x, y, s) &= V_{01}(s)y + V_{11}(s)xy + \frac{1}{2}V_{21}(s)x^2y\\ &+ \frac{1}{6}V_{03}(s)y^3 + \frac{1}{6}V_{31}(s)x^3y + \frac{1}{6}V_{13}(s)xy^3,\\ B_x(s) &= V_{11}(s)y + V_{21}(s)xy + \frac{1}{2}V_{31}(s)x^2y + \frac{1}{6}V_{13}(s)y^3,\\ B_y(s) &= V_{01}(s) + V_{11}(s)x + \frac{1}{2}V_{21}(s)x^2\\ &+ \frac{1}{2}V_{03}(s)y^2 + \frac{1}{6}V_{31}(s)x^3 + \frac{1}{2}V_{13}(s)xy^2,\\ B_s(s) &= \frac{1}{h_3}\left(V'_{01}(s)y + V'_{11}xy + \frac{1}{2}V'_{21}(s)x^2y + \frac{1}{6}V'_{03}(s)y^3\right). \end{aligned} \tag{2}$$

It is clear that

$$V_{j1}(s) = \left(\frac{\partial^j B_y(s)}{\partial x^j}\right)\Bigg|_{\substack{x=0\\y=0}}.$$

Let

$$V_{01}(s) = (B_y(s))|_{\substack{x=0\\y=0}} = B_0(s),$$

$$V_{11}(s) = \left(\frac{\partial B_y(s)}{\partial x}\right)\Bigg|_{\substack{x=0\\y=0}} = G(s),$$

$$V_{21}(s) = \left(\frac{\partial^2 B_y(s)}{\partial x^2}\right)\Bigg|_{\substack{x=0\\y=0}} = W(s),$$

$$V_{31}(s) = \left(\frac{\partial^3 B_y(s)}{\partial x^3}\right)\Bigg|_{\substack{x=0\\y=0}} = O(s).$$

In practice, the index of magnetic field attenuation (field index)

$$n = \rho\left(\frac{1}{B_y}\frac{\partial B_y}{\partial x}\right)\Bigg|_{\substack{x=0\\y=0}}$$

is often used instead of $G(s)$.

In view of the fact that the scalar magnetic potential V must satisfy the Laplace equation $\Delta V = 0$ (where Δ is the Laplacian operator), we obtain the expansions of the potential components correct to the third order at the axial trajectory in the form

$$\begin{aligned} V_{03}(s) &= -(B_0''(s) + W(s) + hG(s)),\\ V_{13}(s) &= 2hB_0''(s) + h^2G(s) - G''(s) - O(s) - hW(s), \end{aligned} \tag{3}$$

where $h = 1/\rho$.

The magnetic induction, as well as the coefficients $G(s)$, $W(s)$, and $O(s)$, can be represented as

$$\begin{aligned} B_0(\tau) &= \check{B}_0\Theta(\tau), \quad G(\tau) = \check{G}\Theta(\tau),\\ W(\tau) &= \check{W}\Theta(\tau), \quad O(\tau) = \check{O}\Theta(\tau). \end{aligned} \tag{4}$$

In the rectangular field model,

$$\Theta(\tau) = u_+(\tau - s_0) - u_+(s - \tau), \tag{5}$$

where $u_+(t)$ is the step function [12] that satisfies the conditions

$$\frac{d}{dt}u_+(t) = \delta_+(t),$$

$$\int_{a+0}^{b}\varphi(\varepsilon)\delta_+(\varepsilon - t)d\varepsilon = \begin{cases} 0 & t < a, \quad t \geq b,\\ \varphi(t+0) & a \leq t < b, \end{cases} \tag{6}$$

$$\int_{a+0}^{b}\varphi(\varepsilon)\delta_+^{(r)}(\varepsilon - t)d\varepsilon = \begin{cases} 0 & t < a, \quad t \geq b,\\ (-1)^r\varphi^{(r)}(t+0), & a \leq t < b. \end{cases}$$

In the smooth field model,

$$\Theta(\tau) = \begin{cases} 1 & s_1 \leq \tau \leq s_2,\\ 0 & \tau < s_0,\ \tau > s\\ \dfrac{1}{1 + eC_0 + C_1\eta(\tau) + C_2\eta^2(\tau) + C_3\eta^3(\tau)} & s_0 \leq \tau \leq s_1,\\ \dfrac{1}{1 + eC_4 + C_5\nu(\tau) + C_6\nu^2(\tau) + C_7\nu^3(\tau)} & s_2 < \tau \leq s, \end{cases} \tag{7}$$

where s_0 and s are the entrance and exit points, respectively, for the charged particle in the sector magnetic field s_1 and s_2 define the boundaries of the stray fields, and $\nu = (\Delta p)/p$ is the momentum spread of the charged particles.

In order to study the dispersion properties of sector magnetic systems, we introduce the vector of the dispersion phase moments $\{\mu, \mu'\}^T$. Since $\mu' = 0$ in sector magnetic analyzers, we will describe the nonlinear dynamics of charged particles in sector magnetic fields with the vector $\bar{\hat{Q}}^{(3)}_{x, x', y, y', \mu} = \{x, x', y, y', \mu, x^2, xx', x'^2, y^2, yy', y'^2, xy, x'y, xy', x'y', x\mu, x'\mu, y\mu, y'\mu, \mu^2, x^3, x^2x', xx'^2, x'^3, xy^2, xyy', xy'^2, x'y^2, x'y^2, x'yy', x'y'^2, y^3, y^2y', yy'^2, y'^3, yx^2, yxx', yx'^2, y'x^2, y'xx', y'x'^2, x^2\mu, xx'\mu, x'^2\mu, y^2\mu, yy'\mu, y'^2\mu, xy\mu, x'y\mu, xy'\mu, x'y'\mu, x\mu^2, x'\mu^2, y\mu^2, y'\mu^2, \mu^3\}^T$ involving fifty-five phase moments of the first, second, and third orders.

When embedded in the space of the phase moments $\bar{\hat{Q}}^{(3)}_{x, x', y, y', \mu}$, nonlinear equations of motion (1) for a charged particle beam in a sector magnetic field are replaced by an extended set of linear differential equations. In the matrix form, this set has the form

$$\frac{d}{ds}(\bar{\hat{Q}}^{(3)}_{x, x', y, y', \mu}) = P^{(3)}\bar{\hat{Q}}^{(3)}_{x, x', y, y', \mu}, \qquad (8)$$

where

$$P^{(3)}(s) = \begin{Vmatrix} P^{1,1} & P^{1,2} & P^{1,3} \\ 0 & P^{2,2} & P^{2,3} \\ 0 & 0 & P^{3,3} \end{Vmatrix}$$

is the upper triangular matrix of coefficients.

The solution of the thus-obtained linear set of equations is an approximation of the solution of the initial nonlinear system with a given order of approximation with respect to the phase variables.

Let us employ the method of embedding in phase-moment space [1] to determine the block elements of the coefficient matrix $P^{(3)}$. After rearranging, we arrive at

$$P^{1,1} = \begin{Vmatrix} H^{1,1} & 0 & H^{1,3} \\ 0 & H^{2,2} & 0 \\ 0 & 0 & H^{3,3} \end{Vmatrix},$$

$$P^{1,2} = \begin{Vmatrix} H^{1,4} & H^{1,5} & 0 & H^{1,7} & 0 & H^{1,9} \\ 0 & 0 & H^{2,6} & 0 & H^{2,8} & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{Vmatrix},$$

$$P^{1,3} = \begin{Vmatrix} H^{1,10} & H^{1,11} & 0 & H^{1,13} & 0 & H^{1,15} & H^{1,16} & H^{1,17} & 0 & H^{1,19} \\ 0 & 0 & H^{2,12} & 0 & H^{2,14} & 0 & 0 & 0 & H^{2,18} & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{Vmatrix},$$

$$P^{2,2} = \begin{Vmatrix} H^{4,4} & 0 & 0 & H^{1,7} & 0 & 0 \\ 0 & H^{5,5} & 0 & 0 & 0 & 0 \\ 0 & 0 & H^{6,6} & 0 & H^{6,8} & 0 \\ 0 & 0 & 0 & H^{7,7} & 0 & H^{7,9} \\ 0 & 0 & 0 & 0 & H^{8,8} & 0 \\ 0 & 0 & 0 & 0 & 0 & H^{9,9} \end{Vmatrix},$$

$$P^{1,3} = \begin{Vmatrix} H^{4,10} & H^{4,11} & 0 & 0 & H^{4,14} & 0 & 0 & H^{4,16} & 0 & 0 \\ 0 & H^{5,11} & 0 & 0 & 0 & H^{5,15} & 0 & 0 & 0 & 0 \\ 0 & 0 & H^{6,12} & H^{6,13} & 0 & 0 & H^{6,16} & 0 & H^{6,18} & 0 \\ 0 & 0 & 0 & 0 & H^{7,14} & H^{7,15} & 0 & H^{7,17} & 0 & H^{7,19} \\ 0 & 0 & 0 & 0 & 0 & 0 & H^{8,16} & 0 & H^{8,18} & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{Vmatrix},$$

$$P^{3,3} = \begin{Bmatrix} H^{10,10} & 0 & 0 & 0 & H^{10,14} & 0 & 0 & 0 & 0 & 0 \\ 0 & H^{11,11} & 0 & 0 & 0 & H^{11,15} & 0 & 0 & 0 & 0 \\ 0 & 0 & H^{12,12} & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & H^{13,13} & 0 & 0 & H^{13,16} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & H^{14,14} & 0 & 0 & H^{14,17} & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & H^{15,15} & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & H^{16,16} & 0 & H^{16,18} & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & H^{17,17} & 0 & H^{17,19} \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & H^{18,18} & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & H^{19,19} \end{Bmatrix},$$

$$H^{1,1} = H^{7,7} = H^{17,17} = \begin{Bmatrix} 0 & 1 \\ -k & 0 \end{Bmatrix},$$

$$H^{1,3} = H^{7,9} = H^{17,19} = \begin{Bmatrix} 0 \\ h \end{Bmatrix},$$

$$H^{1,4} = H^{7,14} = \begin{Bmatrix} 0 & 0 & 0 \\ -h^3 + 2hg - \frac{1}{2}w & 0 & \frac{1}{2}h \end{Bmatrix},$$

$$H^{1,5} = H^{7,15} = \begin{Bmatrix} 0 & 0 & 0 \\ \frac{1}{2}b_2 - \frac{1}{2}hg + \frac{1}{2}w & b_2 & -\frac{1}{2}h \end{Bmatrix},$$

$$H^{1,7} = H^{7,17} = \begin{Bmatrix} 0 & 0 \\ 2h^2 - g & 0 \end{Bmatrix},$$

$$H^{1,10} = \begin{Bmatrix} 0 & 0 & 0 & 0 \\ \frac{1}{2}h^4 + h^2 g - hw + \frac{1}{6}o & 0 & -2h^2 + \frac{3}{2}g & 0 \end{Bmatrix},$$

$$H^{1,9} = H^{7,19} = \begin{Bmatrix} 0 \\ -h \end{Bmatrix},$$

$$H^{1,11} = \begin{Bmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ -\frac{1}{2}h^2 g - \frac{1}{2}g_2 + \frac{3}{2}hw - \frac{1}{2}o & -g_1 & \frac{1}{2}g & 0 & -g & 0 \end{Bmatrix},$$

$$H^{1,17} = \begin{Bmatrix} 0 & 0 \\ -2h^2 + g & 0 \end{Bmatrix},$$

$$H^{1,14} = \begin{Bmatrix} 0 & 0 & 0 \\ h^3 - 2hg + \frac{1}{2}w & 0 & \frac{3}{2}h \end{Bmatrix},$$

$$H^{3,3} = H^{9,9} = H^{19,19} = \begin{Bmatrix} 0 \\ 0 \end{Bmatrix},$$

$$H^{1,15} = \begin{Bmatrix} 0 & 0 & 0 \\ -\frac{1}{2}b_2 + \frac{1}{2}hg - \frac{1}{2}w & 0 & \frac{1}{2}h \end{Bmatrix}, \quad H^{1,19} = \begin{Bmatrix} 0 \\ h \end{Bmatrix},$$

$$H^{2,2} = H^{8,8} = H^{18,18} = \begin{Bmatrix} 0 & 1 \\ -g & 0 \end{Bmatrix},$$

$$H^{2,6} = H^{8,16} = \begin{Bmatrix} 0 & 0 & 0 & 0 \\ -2hg & 0 & -b_1 & 0 \end{Bmatrix},$$

$$H^{2,8} = H^{8,18} = \begin{Bmatrix} 0 & 0 \\ g & 0 \end{Bmatrix},$$

$$H^{2,12} = \begin{Bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & -\frac{3}{2}g & -\frac{1}{2}h \end{Bmatrix}, \qquad (9)$$

$$H^{2,13} = \begin{Bmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ -h^2 g & g_1 & -\frac{1}{2}g & 0 & -2h^2 + g & 0 \end{Bmatrix},$$

$$H^{2,16} = \begin{Bmatrix} 0 & 0 & 0 & 0 \\ 2hg - w & 0 & b_1 & n \end{Bmatrix}, \quad H^{2,18} = \begin{Bmatrix} 0 & 0 \\ -g & 0 \end{Bmatrix},$$

$$H^{4,4} = H^{14,14} = \left\{ \begin{array}{ccc} 0 & 2 & 0 \\ -k & 0 & 1 \\ 0 & -2k & 0 \end{array} \right\}, \qquad H^{4,7} = H^{17,15} = \left\{ \begin{array}{cc} 0 & 0 \\ h & 0 \\ 0 & 2h \end{array} \right\},$$

$$H^{4,11} = \left\{ \begin{array}{cccccc} 0 & 0 & 0 & 0 & 0 & 0 \\ \frac{1}{2}b_2 - \frac{1}{2}hg + \frac{1}{2}w & b_1 & -\frac{1}{2}h & 0 & 0 & 0 \\ 0 & 0 & 0 & b_2 - hg + w & 2b_1 & -h \end{array} \right\},$$

$$H^{4,10} = \left\{ \begin{array}{ccccc} 0 & 0 & 0 & 0 \\ -h^3 + 2hg - \frac{1}{2}w & 0 & \frac{1}{2}h & 0 \\ 0 & -2h^3 + 4hg - w & 0 & h \end{array} \right\}, \qquad H^{4,14} = \left\{ \begin{array}{ccc} 0 & 0 & 0 \\ 2h^2 - g & 0 & 0 \\ 0 & 4h^2 - 2g & 0 \end{array} \right\},$$

$$H^{4,17} = \left\{ \begin{array}{cc} 0 & 0 \\ -h & 0 \\ 0 & -2h \end{array} \right\}, \quad H^{5,5} = H^{15,15} = \left\{ \begin{array}{ccc} 0 & 2 & 0 \\ -g & 0 & 1 \\ 0 & -2g & 0 \end{array} \right\},$$

$$H^{5,11} = \left\{ \begin{array}{cccccc} 0 & 0 & 0 & 0 & 0 & 0 \\ -2hg & 0 & 0 & -b_1 & h & 0 \\ 0 & -4hg & 0 & 0 & -2b_1 & 2h \end{array} \right\}, \qquad H^{5,15} = \left\{ \begin{array}{ccc} 0 & 0 & 0 \\ g & 0 & 0 \\ 0 & 2g & 0 \end{array} \right\},$$

$$H^{6,6} = H^{16,16} = \left\{ \begin{array}{cccc} 0 & 1 & 1 & 0 \\ -g & 0 & 0 & 1 \\ -k & 0 & 0 & 1 \\ 0 & -k & -g & 0 \end{array} \right\}, \qquad H^{6,8} = \left\{ \begin{array}{cc} 0 & 0 \\ 0 & 0 \\ h & 0 \\ 0 & h \end{array} \right\}, \quad H^{6,16} = \left\{ \begin{array}{ccccc} 0 & 0 & 0 & 0 \\ g & 0 & 0 & 0 \\ 2h^2 - g & 0 & 0 & 0 \\ 0 & 2h^2 - g & g & 0 \end{array} \right\},$$

$$H^{6,12} = \left\{ \begin{array}{cccc} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ \frac{1}{2}b_2 - \frac{1}{2}hg + \frac{1}{2}w & b_1 & -\frac{1}{2}h & 0 \\ 0 & \frac{1}{2}b_2 - \frac{1}{2}hg + \frac{1}{2}w & b_1 & -\frac{1}{2}h \end{array} \right\},$$

$$H^{6,13} = \left\{ \begin{array}{ccccc} 0 & 0 & 0 & 0 & 0 \\ -2hg + w & -b_1 & 0 & 0 & h \\ -h^3 + 2hg - \frac{1}{2}w & 0 & \frac{1}{2}h & 0 & 0 \\ 0 & -2hg + w & -b_1 - h^3 + 2hg - \frac{1}{2}w & 0 & \frac{3}{2}h \end{array} \right\},$$

$$H^{6,18} = \begin{Bmatrix} 0 & 0 \\ 0 & 0 \\ -h & 0 \\ 0 & -h \end{Bmatrix},$$

$$H^{10,10} = \begin{Bmatrix} 0 & 3 & 0 & 0 \\ -k & 0 & 2 & 0 \\ 0 & -2k & 0 & 1 \\ 0 & 0 & -3k & 0 \end{Bmatrix}, \quad H^{10,14} = \begin{Bmatrix} 0 & 0 & 0 \\ h & 0 & 0 \\ 0 & 2h & 0 \\ 0 & 0 & 3h \end{Bmatrix},$$

$$H^{12,12} = \begin{Bmatrix} 0 & 3 & 0 & 0 \\ -g & 0 & 2 & 0 \\ 0 & -2g & 0 & 1 \\ 0 & 0 & -3g & 0 \end{Bmatrix},$$

$$H^{11,11} = \begin{Bmatrix} 0 & 2 & 0 & 1 & 0 & 0 \\ -g & 0 & 1 & 0 & 1 & 0 \\ 0 & -2g & 0 & 0 & 0 & 1 \\ -k & 0 & 0 & 0 & 2 & 0 \\ 0 & -k & 0 & -g & 0 & 1 \\ 0 & 0 & -k & 0 & -2g & 0 \end{Bmatrix},$$

$$H^{11,15} = \begin{Bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ h & 0 & 0 \\ 0 & h & 0 \\ 0 & 0 & h \end{Bmatrix},$$

$$H^{13,13} = \begin{Bmatrix} 0 & 2 & 0 & 1 & 0 & 0 \\ -k & 0 & 1 & 0 & 1 & 0 \\ 0 & -2k & 0 & 0 & 0 & 1 \\ -g & 0 & 0 & 0 & 2 & 0 \\ 0 & -g & 0 & -k & 0 & 1 \\ 0 & 0 & -g & 0 & -2k & 0 \end{Bmatrix},$$

$$H^{13,16} = \begin{Bmatrix} 0 & 0 & 0 & 0 \\ h & 0 & 0 & 0 \\ 0 & 0 & 2h & 0 \\ 0 & 0 & 0 & 0 \\ 0 & h & 0 & 0 \\ 0 & 0 & 0 & 2h \end{Bmatrix},$$

where

$$\chi_B = \frac{m\vartheta}{q}, \quad h = \frac{1}{\rho} = \frac{\breve{B}_0}{\chi_B}, \quad b_1 = \frac{B_0'(s)}{\chi_B},$$

$$b_2 = \frac{B_0''(s)}{\chi_B}, \quad k = h^2 - g, \quad g = -\frac{G(s)}{\chi_B},$$

$$g_1 = -\frac{G'(s)}{\chi_B}, \quad w = \frac{W(s)}{\chi_B}, \quad o = -\frac{O(s)}{\chi_B}.$$

The relation between the coefficient g and the field index n can be represented in the form

$$g = nh^2. \tag{10}$$

The solution of Eq. (8) is expressed in terms of the matrizant as follows:

$$\bar{\hat{Q}}^{(3)}_{x,x',y,y',\mu} = X(P^{(3)}, s/s_0)\bar{\hat{Q}}^{(3)}_{x_0,x_0',y_0,y_0',\mu}, \tag{11}$$

where $\bar{\hat{Q}}^{(3)}_{x_0,x_0',y_0,y_0',\mu}$ is the initial particle coordinates, $X(P^{(3)}, s/s_0)$ is the matrizant (transfer matrix) of the third order with respect to the phase variables $\hat{Q}^{(3)}_{x,x',y,y',\mu}$.

The matrizant $X(P^{(3)}, s/s_0)$ has the same upper triangular block structure,

$$X(P^{(3)}, s/s_0) = \begin{Bmatrix} X^{1,1} & X^{1,2} & X^{1,3} \\ 0 & X^{2,2} & X^{2,3} \\ 0 & 0 & X^{3,3} \end{Bmatrix}$$

as the coefficient matrix $P^{(3)}$, and satisfies the differential equation

$$\begin{gathered} X'(P^{(3)}, s/s_0) = P^{(3)}X(P^{(3)}, s/s_0), \\ X(P^{(3)}, s/s_0) = I, \end{gathered} \tag{12}$$

where I is the unit matrix.

In the rectangular field model, the integrals in (12) can be taken by quadrature. Hence, the elements of the matrizant $X(P^{(3)}, s/s_0)$ will have an analytical form. The solutions of the linearized equations

$$\begin{gathered} \frac{dX^{1,1}(s/s_0)}{ds} = P^{1,1}(s)X^{1,1}(s/s_0), \\ X^{1,1}(s/s_0) = I \end{gathered}$$

can be represented in the form

$$X^{1,1} = \begin{Bmatrix} r_{11} & r_{12} & 0 & 0 & d_{11} \\ r_{21} & r_{22} & 0 & 0 & d_{21} \\ 0 & 0 & q_{11} & q_{12} & 0 \\ 0 & 0 & q_{21} & q_{22} & 0 \\ 0 & 0 & 0 & 0 & 1 \end{Bmatrix}, \tag{13}$$

where

$$r_{11} = \cos(\sqrt{k}(s-s_0)),\ r_{12} = \frac{1}{\sqrt{k}}\sin(\sqrt{k}(s-s_0)),$$

$$r_{21} = -\sqrt{k}\sin(\sqrt{k}(s-s_0)),\ r_{22} = \cos(\sqrt{k}(s-s_0)),$$

$$q_{11} = \cos(\sqrt{g}(s-s_0)),\ q_{12} = \frac{1}{\sqrt{g}}\sin(\sqrt{g}(s-s_0)),$$

$$q_{21} = -\sqrt{g}\sin(\sqrt{g}(s-s_0)),\ q_{22} = \cos(\sqrt{g}(s-s_0)),$$

$$d_{11} = \frac{h}{k}(1-\cos(\sqrt{k}(s-s_0))),\ d_{21} = \frac{h}{\sqrt{k}}\sin(\sqrt{k}(s-s_0)).$$

For a uniform field ($n = 0$),

$$r_{11} = \cos(h(s-s_0)),\quad r_{12} = \frac{1}{h}\sin(h(s-s_0)),$$

$$r_{21} = -h\sin(h(s-s_0)),\quad r_{22} = \cos(h(s-s_0)),$$

$$q_{11} = 1,\quad q_{12} = s-s_0,\quad q_{21} = 0,\quad q_{22} = 1,$$

$$d_{11} = \frac{1}{h}(1-\cos(h(s-s_0))),$$

$$d_{21} = \sin(h(s-s_0)).$$

Let b_{ij} be the elements of the block matrix $X^{1,1}$, where $i = 1, \ldots, 5$ and $j = 1, \ldots, 5$. Then, in the framework of the rectangular field model, analytical solutions of Eq. (15) for the diagonal matrix blocks $X^{k,k}$ ($k = 2, 3$) can be found by a little algebra. For example, for the $y'xx'$ line of the matrix block $X^{3,3}$, from $y'xx' = (b_{41}x_0 + b_{42}x_0' + b_{43}y_0 + b_{44}y_0' + b_{45}\mu)(b_{11}x_0 + b_{12}x_0' + b_{13}y_0 + b_{14}y_0' + b_{15}\mu)(b_{21}x_0 + b_{22}x_0' + b_{23}y_0 + b_{24}y_0' + b_{25}\mu)$, we get

$$X^{3,3}_{16,15} = r_{11}r_{12}q_{21},\quad X^{3,3}_{16,16} = (r_{11}r_{22} + r_{21}r_{12})q_{21},$$

$$X^{3,3}_{16,17} = r_{12}r_{22}q_2,\quad X^{3,3}_{16,18} = r_{11}r_{21}q_{22},$$

$$X^{3,3}_{16,19} = (r_{11}r_{22} + r_{21}r_{12})q_{22},\quad X^{3,3}_{16,20} = r_{12}r_{22}q_{22},$$

$$X^{3,3}_{16,27} = (r_{11}d_{21} + r_{21}d_{11})q_{21},$$

$$X^{3,3}_{16,28} = (r_{12}d_{21} + r_{22}d_{11})q_2,$$

$$X^{3,3}_{16,29} = (r_{11}d_{21} + r_{21}d_{11})q_{22},$$

$$X^{3,3}_{16,30} = (r_{12}d_{21} + r_{22}d_{11})q_{22},$$

$$X^{3,3}_{16,33} = d_{11}d_{21}q_2,\quad X^{3,3}_{16,34} = d_{11}d_{21}q_{22},$$

taking into account that $b_{11} = r_{11}$, $b_{12} = r_{12}$, $b_{21} = r_{21}$, $b_{22} = r_{22}$, $b_{15} = d_1$, $b_{25} = d_{21}$, $b_{43} = q_{21}$, $b_{44} = q_{22}$, and $b_{13} = b_{14} = b_{23} = b_{24} = b_{41} = b_{42} = b_{45} = 0$; the other matrix elements of this line are equal to zero.

By applying the same procedure to all the components of the phase moments of the second and third orders, we find analytical solutions for all the elements of the diagonal blocks $X^{2,2}$ and $X^{3,3}$. For the off-diagonal blocks $X^{i,k}$ ($k > i$), the general formula applies [1]:

$$X^{i,k}(s/s_0) = \sum_{j=1+i}^{k}\int_{s_0}^{s} X^{i,i}(s/\tau)P^{i,j}(\tau)X^{i,k}(\tau/s_0)d\tau. \quad (14)$$

Thus, we determine the second-order aberration coefficients in the curvilinear coordinate system chosen by the formula

$$X^{1,2}(s/s_0) = \int_{s_0}^{s} X^{1,1}(s/\tau)P^{1,2}(\tau)X^{2,2}(\tau/s_0)d\tau. \quad (15)$$

For instance, the dispersion aberration coefficient $\langle x|\mu^2\rangle$ is given by

$$X^{1,2}_{1,15}(s/s_0) = \langle x|\mu^2\rangle$$

$$= \int_{s_0}^{s} X^{1,1}_{1,2}(s/\tau)[H^{1,4}_{2,1}(\tau)(X^{1,1}_{1,5}(\tau/s_0))^2$$

$$+ H^{1,4}_{2,3}(\tau)(X^{1,1}_{2,5}(\tau/s_0))^2 + H^{1,7}_{2,1}(\tau)X^{1,1}_{1,5}(\tau/s_0)$$

$$+ H^{1,9}_{2,1}(\tau)]d\tau = \frac{1}{k^3}\Big((C_x-1)\Big(\frac{4}{3}h^5 - \frac{7}{3}h^3k$$

$$+ hk^2 + khg - \frac{8}{3}h^3g + \frac{2}{3}h^2w\Big)$$

$$- S_x^2\Big(\frac{1}{3}h^5 + \frac{1}{6}h^3k - \frac{2}{3}h^3g + \frac{1}{6}h^2w\Big)$$

$$+ S_x\sqrt{k}(s-s_0)\Big(h^5 - h^3k - 2h^3g + \frac{1}{2}hkg + \frac{1}{2}h^2w\Big)\Big),$$

where

$$H^{1,4}_{2,1}(\tau) = -h^3 + 2hg - \frac{1}{2}w,\quad H^{1,4}_{2,3}(\tau) = \frac{1}{2}h,$$

$$H^{1,7}_{2,1}(\tau) = 2h^2 - g,\quad H^{1,9}_{2,1}(\tau) = -h,$$

$$X^{1,1}_{1,2}(s/\tau) = \frac{1}{\sqrt{k}}\sin(\sqrt{k}(s-\tau)),$$

$$X^{1,1}_{1,5}(\tau/s_0) = \frac{h}{k}(1-\cos(\sqrt{k}(\tau-s_0))), \quad (16)$$

$$X^{1,1}_{2,5}(\tau/s_0) = \frac{h}{\sqrt{k}}\sin(\sqrt{k}(\tau-s_0)),$$

$$C_x = \cos(\sqrt{k}(s-s_0)),\quad S_x = \sin(\sqrt{k}(s-s_0)).$$

In the case of the uniform field ($g = 0$, $w = 0$, $k = h^2$), we arrive at the well-known expression [7] for this aberration coefficient:

$$\langle x|\mu^2\rangle = -\frac{1}{2h}\sin^2(h(s-s_0)).$$

The third-order aberration coefficients in our coordinate system are calculated by the formula

$$X^{1,3}(s/s_0) = \int_{s_0}^{s}(X^{1,1}(s/\tau)P^{1,2}(\tau)X^{2,3}(\tau/s_0) + X^{1,1}(s/\tau)P^{1,3}(\tau)X^{3,3}(\tau/s_0))d\tau. \quad (17)$$

Here, the elements of the off-diagonal block $X^{2,3}$ are

$$X^{2,3}(s/s_0) = \int_{s_0}^{s}X^{2,2}(s/\tau)P^{2,3}(\tau)X^{3,3}(\tau/s_0)d\tau.$$

The third-order transfer matrix of the phase moments $\hat{Q}_{\tilde{x},a,\tilde{y},b,\mu}$ in the Cartesian coordinate system has the from

$$R^{(3)}(s/s_0) = A^{(3)}_{(x,y,s)\to(\tilde{x},\tilde{y},\tilde{z})}X(P^{(3)}, s/s_0)A^{(3)}_{(\tilde{x},\tilde{y},\tilde{z})\to(x,y,s_0)}, \quad (18)$$

where $A^{(3)}_{(\tilde{x},\tilde{y},\tilde{z})\to(x,y,s_0)}$ and $A^{(3)}_{(x,y,s)\to(\tilde{x},\tilde{y},\tilde{z})}$ are the matrices of phase-moment transformations from the Cartesian to the natural coordinate system at the entrance and from the natural to the Cartesian system at the exit of a sector magnetic analyzer, respectively. These matrices are easily found if we take into consideration that, in going from the curvilinear to the Cartesian coordinate system, the coordinates do not change and the angles obey the formulas

$$a = \frac{d\tilde{x}}{d\tilde{z}} = \frac{x'}{1+hx}, \quad b = \frac{d\tilde{y}}{d\tilde{z}} = \frac{y'}{1+hx}. \quad (19)$$

The third-order aberration coefficients have the form

$$\langle\tilde{x}|\hat{Q}^{(3)}_{\tilde{x},a,\tilde{y},b,\mu}\rangle = R^{(3)}_{1,k}, \quad \langle a|\hat{Q}^{(3)}_{\tilde{x},a,\tilde{y},b,\mu}\rangle = R^{(3)}_{2,k},$$
$$\langle\tilde{y}|\hat{Q}^{(3)}_{\tilde{x},a,\tilde{y},b,\mu}\rangle = R^{(3)}_{3,k}, \quad \langle b|\hat{Q}^{(3)}_{\tilde{x},a,\tilde{y},b,\mu}\rangle = R^{(3)}_{4,k}, \quad (20)$$

where k is the serial number of the phase variable.

The total number of the third-order aberration coefficients is $4 \times 55 = 220$. We will list only those for which the contribution of the fringing-field effects with the zero width of the stray field is significant.

In a sector electrostatic field, the contribution of the fringing-field effects with the zero width of the stray field shows up as corrections to the second-order aberration coefficients with respect to x and a ($\langle\tilde{x}|\tilde{x}\,a\rangle$, $\langle a|\tilde{x}^2\rangle$, $\langle a|\tilde{x}\,a\rangle$, $\langle a|a^2\rangle$, $\langle\tilde{x}|y^2\rangle$, $\langle\tilde{x}|yb\rangle$, $\langle\tilde{x}|b^2\rangle$, $\langle a|y^2\rangle$, $\langle a|yb\rangle$, and $\langle a|b^2\rangle$) [13]. In a sector magnetic field, the contribution is essential for the aberration coefficients $\langle\tilde{x}|y^2\rangle$, $\langle\tilde{x}|yb\rangle$, $\langle a|y^2\rangle$, $\langle a|yb\rangle$, and $\langle a|b^2\rangle$:

$$\langle\tilde{x}|y^2\rangle = R^{(3)}_{1,9} = \frac{1}{4k(k-4g)}((1-C_x)(8g^2h-2kgh-4gw) - kw(C_y^2-S_y^2-2C_c)) + \zeta\left(-\frac{h^2}{2\sqrt{k}}S_x\right),$$

$$\langle\tilde{x}|yb\rangle = R^{(3)}_{1,10} = \frac{1}{\sqrt{k}\sqrt{g}(k-4g)}(\sqrt{k}wC_y - \sqrt{g}wS_x) + \zeta\left(\frac{h}{\sqrt{k}}S_x\right),$$

$$\langle\tilde{x}|b^2\rangle = R^{(3)}_{1,11} = \frac{1}{2kg(k-4g)} \times ((1-C_x)(4g^2h-kgh-2gw) - kwS_y^2),$$

$$\langle a|y^2\rangle = R^{(3)}_{2,9} = \frac{1}{2\sqrt{k}(k-4g)} \times (S_x(4g^2h-kgh+kw-2gw) - 2\sqrt{k}\sqrt{g}wC_yS_y) + \zeta\left(-h^2C_x+\frac{h^2}{2}C_y^2+h\sqrt{g}S_y\right),$$

$$\langle a|yb\rangle = R^{(3)}_{2,10} = \frac{w}{(k-4g)}(-S_x-S_y^2+C_y^2) + \zeta\left(hC_x-\frac{h^2}{\sqrt{g}}S_yC_y-h(C_y^2-S_y^2)\right),$$

$$\langle a|b^2\rangle = R^{(3)}_{2,11} = \frac{1}{2\sqrt{k}\sqrt{g}(k-4g)} \times (S_x(4hg\sqrt{g}-hk\sqrt{g}-2\sqrt{g}w)+\sqrt{k}wS_yC_y) + \zeta\left(\frac{h^2}{2g}S_y^2-\frac{h}{\sqrt{g}}S_yC_y\right),$$

where $C_x = \cos(\sqrt{k}\,(s-s_0))$, $S_x = \sin(\sqrt{k}\,(s-s_0))$, $C_y = \cos(\sqrt{g}\,(s-s_0))$, and $S_y = \sin(\sqrt{g}\,(s-s_0))$. The parameter $\zeta = 1$ if the function is described by Eq. (6). If we put $\zeta = 0$ [$B'(s) = 0$ and $B''(s) = 0$], we would obtain the equations [7] that are widely used in designing ion–optical systems where the stray fields are taken into account by replacing the real field by the idealized field equivalent by rotation.

Thus, for a rectangular longitudinal distribution of a sector magnetic field, the analytical expressions for all matrizant elements, and hence, for all aberration coefficients are obtained. In the case of a smooth longitudinal field distribution, the matrizant was calculated with the conservative numerical method of shuttle sums [6].

The convergence of the solution of the beam dynamics problem within the smooth model to that for the rectangular model was studied by the software developed. It is found that the "rectangular" and "smooth" aberration coefficients approach each other when the stray field width tends to zero. Hence, both the analytical and the numerical solutions are reliable.

REFERENCES

1. A. D. Dymnikov and R. Hellborg, Nucl. Instrum. Methods Phys. Res. A **330**, 323 (1993).
2. V. Brazhnik, V. Khomenko, S. Lebed, and A. Ponomarev, Nucl. Instrum. Methods Phys. Res. B **104**, 69 (1995).
3. V. Brazhnik, S. Lebed, W. Kwiatek, *et al.*, Nucl. Instrum. Methods Phys. Res. B **130**, 104 (1997).
4. A. D. Dymnikov *et al.*, Nucl. Instrum. Methods Phys. Res. A **403**, 195 (1998).
5. A. H. Azbaid, A. D. Dymnikov, and G. Martinez, Nucl. Instrum. Methods Phys. Res. B **158**, 61 (1999).
6. A. D. Dymnikov, Nucl. Instrum. Methods Phys. Res. A **363**, 435 (1995).
7. K. L. Brown *et al.*, Rev. Sci. Instrum. **35**, 481 (1964).
8. A. S. Kuzema, O. R. Savin, and I. Ya. Chertkov, *Analyzers for Magnetic Mass Spectrometers* (Naukova Dumka, Kiev, 1987).
9. H. Matsuda and H. Wollnik, Nucl. Instrum. Methods **77**, 283 (1970).
10. Y. Fujita, H. Matsuda, and T. Matsuo, Nucl. Instrum. Methods **144**, 279 (1977).
11. M. Szilagyi, *Electron and Ion Optics* (Plenum, New York, 1988; Mir, Moscow, 1990).
12. G. A. Korn and T. M. Korn, *Mathematical Handbook for Scientists and Engineers* (McGraw-Hill, New York, 1961; Nauka, Moscow, 1968).
13. S. N. Mordik and A. G. Ponomarev, Zh. Tekh. Fiz. **71** (4), 105 (2001) [Tech. Phys. **46**, 466 (2001)].

Translated by M. Fofanov

Technical Physics, Vol. 46, No. 7, 2001, pp. 892–896. Translated from Zhurnal Tekhnicheskoĭ Fiziki, Vol. 71, No. 7, 2001, pp. 106–110.

The Formation of Developed Morphology on the Indium Phosphide Surface by Ion Argon Beam Sputtering

I. P. Soshnikov*, A. V. Lunev*, M. É. Gaevskiĭ*, S. I. Nesterov*, M. M. Kulagina*, L. G. Rotkina*, V. T. Barchenko, I. P. Kalmykova**, A. A. Efimov**, and O. M. Gorbenko*****

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

***St. Petersburg State University of Electrical Engineering, St. Petersburg, 197376 Russia*

****Institute of Analytical Instrument Making, Russian Academy of Sciences, Ruzhskiĭ pr. 26, St. Petersburg, 198103 Russia*

Received September 12, 2000

Abstract—Self-organizing structures on the InP surface that are formed by ion-beam sputtering in the energy range 0.1–15 keV are investigated. It is shown that the processing of the InP surface by monochromatic argon beams can give rise to the formation of two, "grass" and "cone-in-pit," morphologies. The formation of the relief is treated in terms of a qualitative model including the processes of sputtering, cascade mixing, and surface transport. The model adequately predicts the fluence dependence of the density and size of morphological features. In addition, it enables one to clarify conditions under which the morphologies form, as well as to explain the effect of target temperature on the demarcation line between the morphologies. It is demonstrated that the morphology may become anisotropic in the case of mask etching. In particular, the application of regularly spaced strips as masks makes it possible to produce a texture-like surface structure. *© 2001 MAIK "Nauka/Interperiodica".*

INTRODUCTION

Structures with a developed surface morphology have received much attention due to their potential application in micro- and optoelectronics [1, 2]. A basic challenge in this field is obtaining a regular morphology [1]. At the same time, it is common knowledge that the sputtering of some semiconductors (such as InP) may cause a developed quasi-regular surface morphology of the so-called grass type to form [3–5]. Such a morphology can be used as an element of device structures. However, data for the formation of the relief by ion sputtering are insufficient for applications [3–5]. These investigations are also of fundamental interest, since they provide a better insight into the mechanism behind relief formation during ion bombardment.

In this study, we investigate the structure of the InP surface exposed to argon-ion beams with an energy of 0.1–15 keV.

EXPERIMENT

We experimented with standard FIÉT-4 InP(001) wafers with a dopant concentration of ~10^{17} cm^{-3}. The wafers were cut into 250- to 450-μm-thick samples about 1 cm^2 in area. The damaged surface layer was removed by dynamical chemical polishing [6].

The samples were irradiated with two special setups based on VUP-5 vacuum stations (NPO Elektron, Sumy, Ukraine). The ion gun of one of the setups generates a monochromatic neutralized ion beam of energy from 0.1 to 1.2 keV, diameter about 50 mm, and particle flux density $j \sim 10^{15}$ particle/cm^2 s.

The fluence and the particle flux density were determined by sputtering satellite GaAs samples, since this process has been well studied [3, 7–11]. In our experiments, the satellites were partially masked and then, together with the sample under study, exposed to the ion beam. Next, the height of a step between the masked and uncovered regions of the satellite was determined. The fluence Φ and the particle flux density j are related to the etch depth h as

$$\Phi = \frac{2N_A\rho}{M}, \quad j = \frac{2N_A\rho}{M}\frac{h}{\tau}, \tag{1}$$

where $N_A = 6 \times 10^{23}$ is the Avogadro number, ρ = 5.35 g/cm^2 is the GaAs density [12], M = 143.79 g is the GaAs molar weight [12], and τ is the exposure time.

The gun of the second setup produces a monochromatic beam of Ar ions with an energy from 0.1 to 15 keV. The diameter of the beam was 3–15 mm (depending on focusing conditions), and the particle flux density j reached ~10^{15} particle/cm^2 s (≈250 μA/cm^2). In this case, the current density was determined using a calibrated-aperture Faraday cup. The sources of accelerated particles are described more closely in [13]. The temperature of the samples did not exceed 80°C in our experiments.

The surface relief was investigated by scanning electron microscopy (SEM). The surface morphology

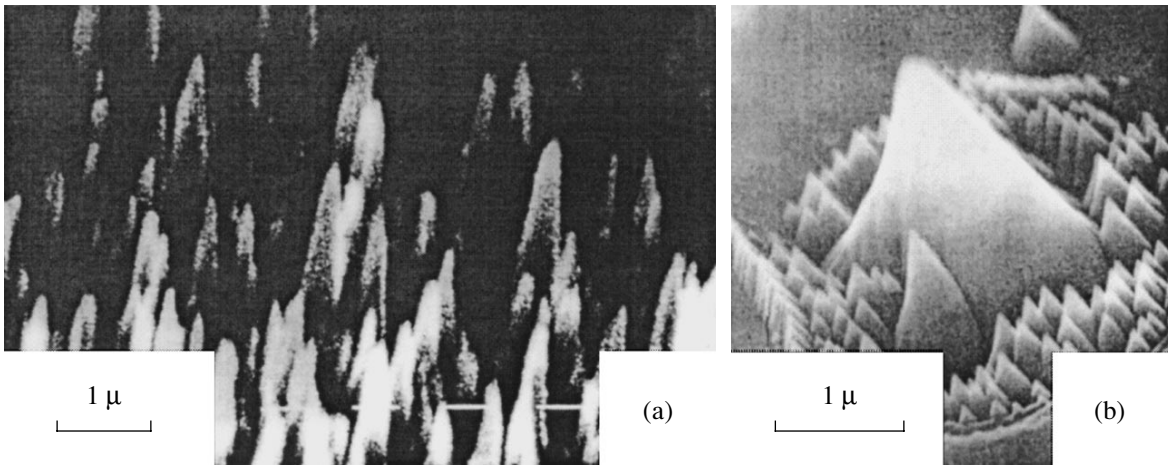


Fig. 1. SEM image of the InP surface exposed to the (a) neutralized Ar ion beam with $E = 0.6$ keV and $j = 10^{15}$ particle/cm^2 s and (b) Ar ion beam with $E = 5$ keV and $j \sim 10^{15}$ particle/cm^2 s.

was analyzed with a special program allowing envelope subtraction, noise filtering, Fourier transformation of images, etc. [14].

RESULTS AND DISCUSSION

Sputtering by Ar ions, for example, with an energy $E \approx 1$ keV (particle flux density $j \approx 10^{15}$ particle/cm^2 s, or ≈100 μA/cm^2) gives rise to a developed ("grassy") surface (Fig. 1a) consisting of closely spaced vertical conical features. The apex angle of the cones was found to be $\Theta = 12 \pm 5°$. Note that such a surface poorly reflects visible light.

As the energy and/or ion flux density increases, isolated cones and/or clusters of cones appear on the smooth surface (Fig. 1b). The apex angle of these cones depends on the ion energy and is roughly twice as large as the angle at which the sputtering yield is maximum, ranging from 60° to 80°. The reflection of such a surface remains high in the visible range.

The demarcation line between the two morphologies is graphically represented in Fig. 2. The line is adequately described by the relationship $Ej = \text{const} \approx 0.10 \pm 0.05$ W/cm^2. At power densities below or equal to this value, the grassy surface develops; otherwise, the cone(s)-in-pit morphology forms. Note that the constant may depend on the sample temperature [4].

To characterize the surface morphology, we processed the SEM image using a Wien noise filter, as well as envelope-subtraction and Fourier-transform techniques. The resulting Fourier transforms of the top-view image of the grassy surface (Fig. 3a) are shown in Figs. 3b and 3c. They were obtained with and without envelope subtraction, respectively.

At $Ej > 0.10$ W/cm^2, the concentration of the cones depends on the radiant exposure only slightly and their heights h are randomly distributed in the interval $0 < h < h_0$, where h_0 is the value close to the etch depth (Fig. 4).

At the same time, the size distribution of the features that form at $Ej > 0.10$ W/cm^2 has a well-pronounced peak whose position depends on the ion fluence and energy (Fig. 4). Note that the variance of the cone sizes is estimated at $\delta D/D \le 10\%$. However, if the surface is pretreated inappropriately, the variance may be as high as $\delta D/D \approx 35\%$. For this morphology, the concentration of the features depends on the ion flux density and/or energy. The fluence dependences of the density and the characteristic (mean) height are given in Fig. 5. It is seen that the feature size and density vary as $h \sim \Phi^{1/4}$ and $\sigma \sim \Phi^{-1/2}$, respectively.

The Fourier transforms of the images from the grassy surface are typical of isotropic objects; i.e., the cones of close sizes are arranged isotropically.

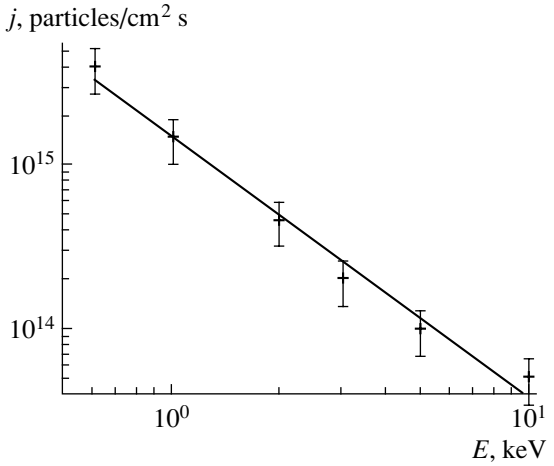


Fig. 2. Conditions for the formation of the grass-type and cone-in-pit morphologies. The target temperature is not greater than 80°C.

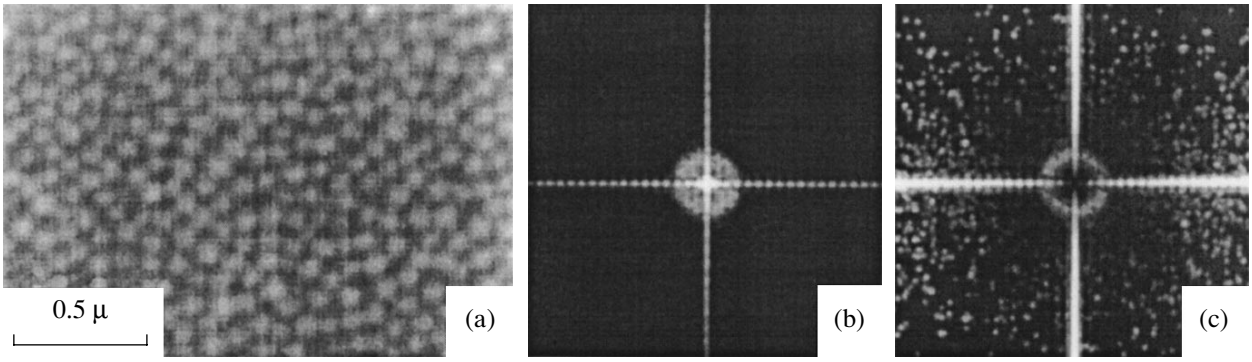


Fig. 3. (a) SEM image of the unmasked InP surface exposed to the neutralized Ar ion beam with $E = 0.6$ keV, $j \sim 10^{15}$ particle/cm^2 s, and $\Phi \sim 10^{18}$ particle/cm^2; (b, c) Fourier transforms of the image shown in Fig. 3a with and without envelope subtraction, respectively.

To trace the effect of boundaries on the morphology type, a series of samples were masked by strips made by the interference photolithography method. The strip widths were 0.4 and 0.1 μm, and the spacings between them were 1.6 and 0.45 μm, respectively. The surface morphology on the masked samples and its Fourier transforms are shown in Fig. 5. Here, the Fourier transforms are typical of textured objects. The sizes of the cones formed on the unmasked and masked parts of the surface diverge, and so do the variances of the sizes: $\delta D/D_{\text{nonmask}} \approx 10\%$ and $\delta D/D_{\text{mask}} \approx 15\%$.

The type of surface morphology can be explained in the framework of the model of spontaneous coalescent relief formation [3, 5]. In this model, features of the grassy surface result from sputtering, cascade collisions, and surface transport. When these processes combine, the nuclei of conical whiskerlike features appear on the amorphized surface. The nuclei grow due to the redistribution of the target material. With such a relief formation mechanism, the concentration of the cones as a function of the fluence varies because of coalescence; i.e., some cones take up others. The basic equation of the model (the equation for surface evolu-

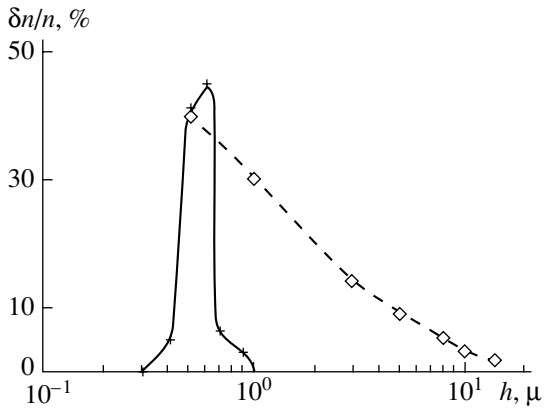


Fig. 4. Size distribution of the surface features for the (□) cone(s)-in-pit and (+) grass-type morphologies. (□) E = 5 keV, $j \sim 10^{15}$ particle/cm^2 s, and $\Phi \approx 5 \times 10^{18}$ particle/cm^2 and (+) E = 0.6 keV, $j \sim 10^{15}$ particle/cm^2 s, and $\Phi \approx 10^{18}$ particle/cm^2.

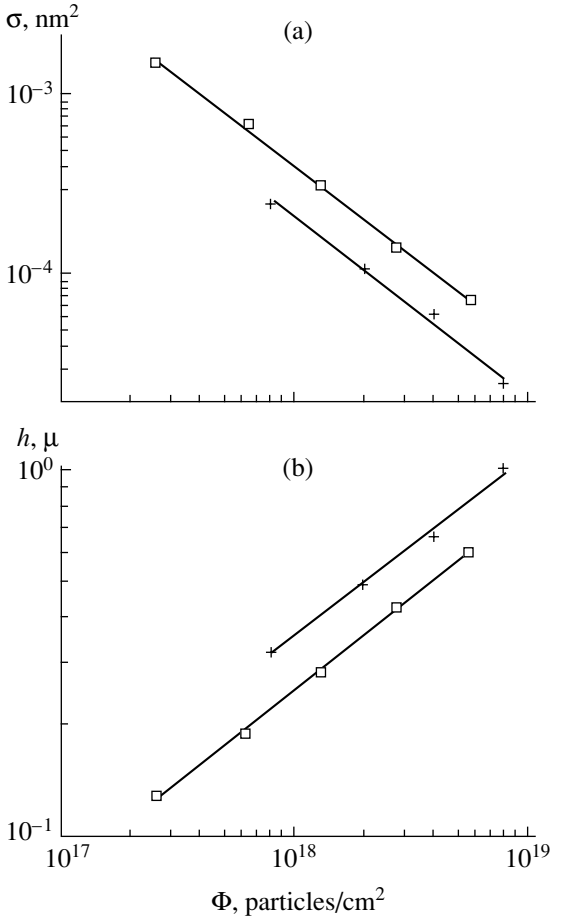


Fig. 5. (a) Concentration and (b) height h of grassy surface features as a function of fluence for E = (+) 0.6 and (□) 5 keV.

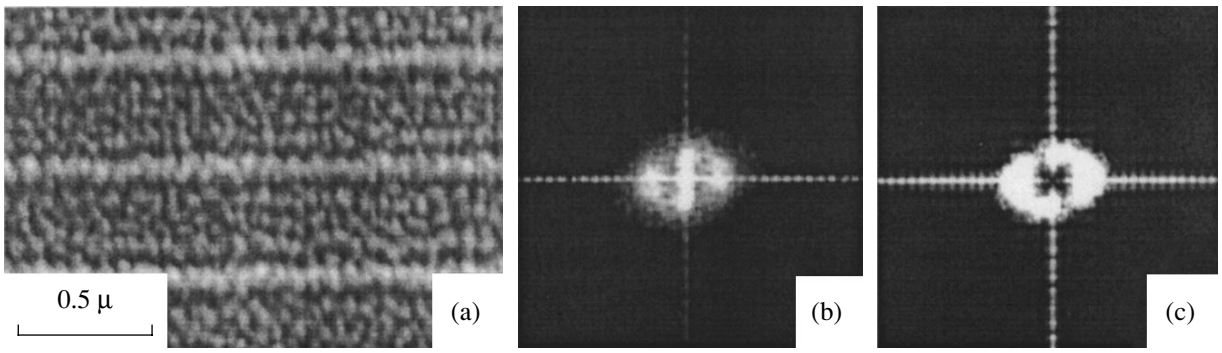


Fig. 6. (a) SEM image of the InP surface mask-etched by the neutralized Ar ion beam for $E = 0.6$ keV, $j \sim 10^{14}$ particle/cm^2 s, and $\Phi \sim 10^{18}$ particle/cm^2; (b, c) Fourier transforms of the image shown in Fig. 6a with and without envelope subtraction, respectively.

tion) is as follows [15, 16]:

$$\left(\frac{\partial}{\partial t} - v\frac{\partial}{\partial r}\right)S = n^{-1}(J_{\text{sput}} + J_{\text{trans}} + J_{\text{mix}}), \qquad (2)$$

where t is the exposure time; J_{sput}, J_{trans}, and J_{mix} are the atomic fluxes due to sputtering, diffusion, and cascade mixing, respectively, on the target surface; S is the surface area; and v is the local sputtering rate, which depends on the type of ions, target material, as well as on the energy, angle of incidence, and flux density of ions.

Note that, at $J_{\text{trans}} > J_{\text{sput}} + J_{\text{mix}}$, it is possible to derive a locally synergistic solution corresponding to the grassy surface. Such equations are solved in the coalescence theory [17]. Detailed analysis of Eq. (2) will be presented in following articles.

The coalescent mechanism of surface morphology allows one to explain the fluence dependences of the cone concentration and size ($\sigma \sim \Phi^{-1/2}$ and $h \sim \Phi^{-1/4}$, respectively). In our model, an increase in the flux density and/or ion energy may destroy the nuclei and/or disturb transport fluxes responsible for the growth of the surface features. In other words, the model predicts the presence of a threshold for the surface morphology evolution, which was observed in the experiments. At the same time, simulation experiments where the rate of transport processes was varied showed that the threshold for grass formation may shift toward greater energies and flux densities. Since transport processes may depend on the target temperature, an increase in the temperature may change the type of surface morphology. Such an effect was observed when InP samples were bombarded by Ar^+ ions at elevated target temperatures [4].

CONCLUSION

Thus, we have investigated self-organizing structures on the InP surface exposed to monochromatic argon ion beams with an energy between 0.1 and 15 keV. It is shown that the ion-beam treatment of the InP surface may lead to the development of grass-type and cone-in-pit morphologies. The effect is explained with the qualitative model including the processes of sputtering, cascade mixing, and surface transport. The model adequately predicts fluence dependences of the concentration and size of morphological features, conditions under which the morphologies form, and the effect of target temperature.

It is shown that the presence of etch-region boundaries may induce the partial anisotropy of the morphology. In particular, the application of regularly spaced masking strips results in the textured structure.

ACKNOWLEDGMENTS

This work was supported in part by the Russian Foundation for Basic Research (grant no. 00-02-17006) and by the Federal Scientific and Technical Program (project no. 141).

REFERENCES

1. J. Joannopoulos, R. Meade, and J. Winn, *Photonic Crystals* (Princeton Univ. Press, Princeton, 1995).
2. E. Yablonovitch, J. Mod. Opt. **41**, 173 (1994).
3. J. B. Malherbe, Crit. Rev. Solid State Mater. Sci. **19** (2/3), 55 (1994).
4. O. Wada, J. Phys. D **17**, 2429 (1984).
5. I. P. Soshnikov, A. V. Lunev, M. E. Gaevski, *et al.*, Proc. SPIE **3048**, 404 (1997).
6. K. Sangwal, *Etching of Crystals: Theory, Experiment, and Application* (Elsevier, Amsterdam, 1987; Mir, Moscow, 1990).
7. I. P. Soshnikov, A. V. Lunev, Yu. A. Kudryavtsev, and N. A. Bert, Nucl. Instrum. Methods Phys. Res. B **59/60**, 115 (1996).
8. I. P. Soshnikov and N. A. Bert, Fiz. Tverd. Tela (St. Petersburg) **35**, 2250 (1993) [Phys. Solid State **35**, 1118 (1993)].
9. I. P. Soshnikov and N. A. Bert, Zh. Tekh. Fiz. **66** (6), 84 (1996) [Tech. Phys. **41**, 567 (1996)].
10. J. Comas and C. B. Cooper, J. Appl. Phys. **38**, 2956 (1967).

11. S. R. Bhattacharya, D. Ghose, and D. Basu, Nucl. Instrum. Methods Phys. Res. B **47**, 253 (1990).

12. V. T. Barchenko and A. Sokolovskiĭ, Izv. Leningr. Élektrotekh. Inst. im. V. I. Ul'yanova **303**, 42 (1982).

13. *Landolt-Börnstein, Numerical Data and Functional Relationships in Science and Technology* (Springer-Verlag, Berlin, 1982), Vol. 17e.

14. O. M. Gorbenko, D. V. Kurochkin, and A. O. Golubok, in *Proceedings of the 1st International Conference on Digital Signal Processing and Its Application, Moscow, 1998*, Vol. 3, p. 197.

15. G. Carter, in *Erosion and Growth of Solids Simulated by Atom and Ion Beams*, Ed. by G. Kirikidis, G. Garter, and J. L. Whitton (Martinus Nijhoff Publ., Dordrecht, 1986).

16. M. G. Stepanova, in *Proceedings of the 3rd International Conference on Computer Simulation of Radiation Effects in Solids, COSIRES'96, Univ. of Surrey, Guildford, 1996*, Vol. 1, p. 85.

17. S. A. Kukushkin and V. V. Slezov, *Disperse Systems on Solid Surface* (Nauka, St. Petersburg, 1996).

Translated by Yu. Vishnyakov

Technical Physics, Vol. 46, No. 7, 2001, pp. 897–904. Translated from Zhurnal Tekhnicheskoĭ Fiziki, Vol. 71, No. 7, 2001, pp. 111–118.

EXPERIMENTAL INSTRUMENTS AND TECHNIQUES

Structure and Properties of a Hard Alloy Deposited on a Copper Substrate by Means of a Pulsed Plasma Spray Technology

A. D. Pogrebnyak*, M. V. Il'yushenko*, O. P. Kul'ment'eva*, Yu. N. Tyurin, A. P. Kobzev***, Yu. F. Ivanov****, V. S. Ivaniĭ*****, and V. S. Kshnyakin*******

** Sumy Institute for Surface Modification, pr. Shevchenko 17, Sumy, 40030 Ukraine*

*** Paton Electric Welding Institute, National Academy of Sciences of Ukraine, 11 Bozhenko St., Kiev, 03680 Ukraine*

**** Frank Laboratory of Neutron Physics, Joint Institute of Nuclear Research, Dubna, Moscow oblast, 141980 Russia*

***** Tomsk State Academy of Architecture and Civil Engineering, Solyanaya pl. 2, Tomsk, 634050 Russia*

****** Sumy State Pedagogical University, Romenskaya St., 87, Sumy, 40030 Ukraine*

Received October 20, 2000

Abstract—A pulsed plasma spray generator based on a new principle is developed for depositing ceramic-metal, ceramic, and metal coatings on solid substrates. Calculations of the plasma-generator parameters are presented. A hard alloy (W–Co) was deposited on a copper substrate to demonstrate the use of the plasma generator. A Rutherford backscattering (RBS) technique, x-ray diffraction microanalysis of phase composition, diffraction transmission electron microscopy (TEM), and hardness and adhesion measurements were used to examine the hard-alloy coating. It is shown that the coating consists of W–Co crystals with hexagonal and cubic lattices and contains ~25-nm cobalt α- and β-phase crystallites, with W_3Co_3C particles revealed at the crystallite edges. © *2001 MAIK "Nauka/Interperiodica".*

INTRODUCTION

Current progress in high-quality coating deposition technology relies on the development of techniques in which high velocities of deposited materials are attained by means of detonation devices [1, 2], various rocket combustion chambers [3, 4], or electromagnetic rail guns [5–7]. Investigations have shown that a dense coating that reliably adheres to a substrate surface can be created (even without heating) by using metal alloy particles with velocities ranging from 600 to 1000 m/s [8].

In the devices where high-velocity jets of combustion products issue from rocket combustion chambers, at least 10 m^3 of gas are required to deposit 1 kg of a tungsten-carbide coating [4]. The highest attainable velocity of particles 45 ± 10 μm in diameter is 600–650 m/s. To obtain a high-velocity flow of combustion products, 30–150 m^3/h of a combustible premixed gas are burned in combustion chambers. However, the deposition rate and coating quality do not increase proportionately with the rate of heat release in the general case [3]. Normally, large amounts of premixed-gas components are necessary only to attain a high velocity of the jet of combustion products, whereas uniform distribution of the powder over the jet cross section has not been obtained to this day. As a consequence, the thermal efficiency of the utilization of combustion products has yet to be improved.

Electromagnetic techniques for powder acceleration and heating are also being developed [5–7]. In an electromagnetic rail gun, both working gas and powder are accelerated and heated by the ponderomotive forces due to the current traversing the rail electrodes and the plasma layer. The energy transferred to the plasma is proportional to the current, and a current of 150 kA is required to attain a velocity of 2–4 km/s. The corresponding acceleration time is 120 ms, and a plasma temperature as high as 20 000°C can be attained [6].

A more efficient type of rail gun used in coating deposition was discussed in [7]. In the axial rail gun proposed in that study, a plasma "piston" speed of 10 km/s can be attained by using a current of 20 kA. The velocity of the plasma jet issuing from the rail gun can be as high as 4 km/s. As a result, the powder is accelerated to a speed sufficient to create high-quality tungsten-carbide and aluminum-oxides hard-alloy coatings.

The applicability of electromagnetic systems is limited because of arc localization in the plasma piston and erosion of rail electrodes. Moreover, these systems have complicated devices for switching currents of amplitude 20–150 kA and frequency 2–10 Hz, which restrict their effectiveness and reliability.

In the gaseous-detonation systems employed in coating deposition facilities [1–3], coating powders are accelerated to 1000 m/s and heated to their melting points. These systems are characterized by deposition rates comparable to those attained with the use of rocket combustion chambers, whereas the corresponding rates of heat release are several times lower. In a detonation system, a high-velocity gas jet is obtained as a result of combustion in a detonation regime. This process is independent of the burned amount of a combus-

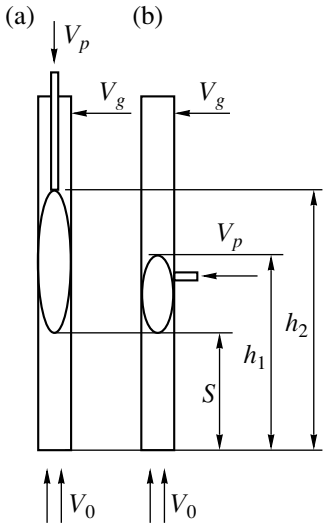


Fig. 1. Schematic of gas and powder supply to the combustion chamber of a detonation gun: (a) axial injection; (b) radial injection.

tible mixture. Despite these advantages, detonation deposition techniques have not been applied because they involve very complicated powder feed rate control units. Moreover, since a single pulsed jet of combustion products produced by a detonation system has a limited intensity, this coating deposition technique is characterized by a relatively low reliability.

In most detonation techniques for coating deposition [1–2, 9–12], the coating powder and gases are fed into a combustion chamber. The combustion chamber of a detonation facility (see Fig. 1) is filled with gases (V_g) and powder (V_p) from three directions. The gases are introduced at the closed end of the combustion chamber through a feed rate control unit. Air (V_0) flows into the combustion chamber through its open end, reaches the location of the gas–powder mixture, and impedes the mixture's movement. When the powder is introduced along the combustion-chamber axis (Fig. 1a), it mixes with the combustible gas and fills a large volume $V = f(h_2 - S)$. Uniform conditions for powder heating and acceleration cannot be achieved in this setting. When powder is introduced radially through a side wall (Fig. 1b), a more compact and dense powder cloud of volume $V = f(h_1 - S)$ is obtained. As a result, uniform conditions for energy transfer from combustion products to the entire powder mass are created.

Coating deposition devices with radial powder injection and mass flow-rate control systems have been developed [12]. The gases and powder are introduced periodically when the combustion-chamber pressure decreases after a high-velocity jet of combustion products is ejected.

HIGH-ENERGY PLASMA JET TECHNIQUE

Rail and coaxial guns should be singled out as pulsed plasma spray accelerators that provide the highest efficiency of surface modification and coating deposition. Experimental and theoretical investigations of the processes taking place in rail and coaxial guns were summarized in [7].

Currently, a new class of modification and coating-deposition techniques is being developed. These techniques are based on electromagnetic acceleration of the products of premixed gas combustion [12–20]. It was proposed in [12, 13] to use a special chamber (*1* in Fig. 2) for gaseous mixture preparation and detonation. The chamber is separated from the plasma gun. The

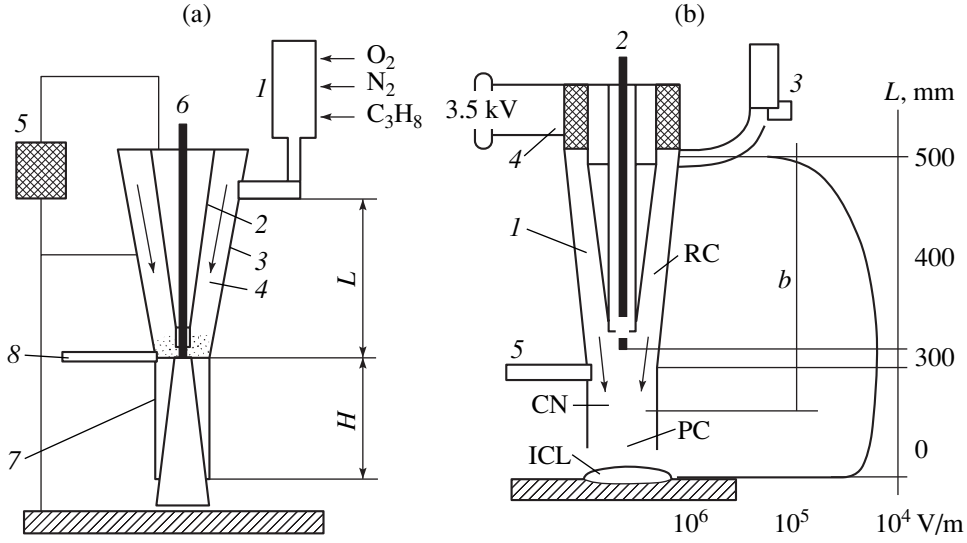


Fig. 2. Schematic of the a pulsed-jet plasma spray facility based on an electromagnetic principle of additional energy input: (a) design of the pulsed plasma spray gun; (b) profile of electric field strength.

plasma gun consists of an inner conical electrode (*2*) and an outer cathode (*3*). In the annular gap of length *L* (*4*) between the coaxial electrodes, an electric field of intensity *E* is generated by means of a high-voltage source (*5*). The central electrode holds an expendable metal rod (*6*). A typical rod is made of a refractory metal or alloy. The plasma gun has a barrel (*7*), in which the powder is heated and accelerated. The barrel length *H* depends on the elemental composition and size distribution of the powder. When a hard alloy is deposited, $H = 300$ mm. The powder is injected into the barrel through a tube (*8*).

The gaseous mixture components are fed into the detonation chamber. Their mixing is followed by the initiation of a detonation. Then, the burned gas flows out of the detonation chamber into the electrode gap and closes the circuit containing the voltage generator. The conductive layer of combustion products is accelerated by gasdynamic and electrodynamic forces. The powder injected into the barrel is heated and accelerated by the plasma jet. The heated expendable metal rod evaporates and supplies an alloying element to the plasma jet. When ejected from the plasma gun, a plasma jet closes the electric circuit between the electrode and coated surface as anode and cathode, respectively. The current carried by the jet generates a magnetic field pulse, while the plasma and powder are further heated through the Joule heat release.

The energy parameters of a pulsed plasma jet can be determined by solving the well-known problem of detonation propagation in electric field between two coaxial bodies of revolution (electrodes). The problem was simplified: the average temperature, velocity, pressure, and density of combustion products were determined on the working-chamber axis without taking into account the variation of its cross section [14, 15].

The geometric and energy parameters of the two-dimensional time-dependent problem of detonation propagation averaged over the working-chamber electrode gap are calculated as

$$\langle X\rangle = \frac{l}{h}\int_0^h X\,dh, \tag{1}$$

where *X* may be *B*, *E*, *J*, *w*, …, and *h* is the average annular-gap width (*B* is the magnetic induction, *E* is the electric field strength, *J* is the electric current, and *w* is the plasma velocity; see Fig. 3).

The integral in (1) was calculated in the normal direction to the electrode surface. The parameters were calculated as depending only on time and the distance *l* along the working-chamber element up to the location where the detonation was initiated. The working-chamber length was set equal to *L*. In the approximation adopted here, the location of a detonation front is $l = l_D(t)$, and the detonation velocity is $D = dl_D/dt$. The detonation front is represented as an array of line segments perpendicular to the working-chamber electrode surfaces. The detonation intensity is characterized by the Mach number

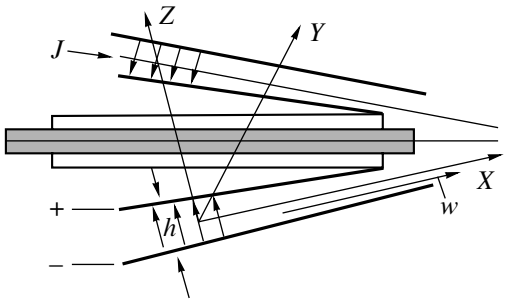


Fig. 3. Schematic of the working chamber used to calculate the electromagnetic enhancement of detonation.

$$M_D = \frac{D}{a_0} = \frac{dl_D/dt}{a_0}, \tag{2}$$

where a_0 is the additional velocity induced by electromagnetic acceleration.

It was assumed that the vector of electric current $\langle j\rangle$ is perpendicular to the intensity $\langle E\rangle$ and the plasma-jet velocity $\langle w\rangle$ is parallel to the electrode surface element. These assumptions are based on the fact that the electrode gap is relatively narrow ($h = 6$–8 mm).

The gas flow behind the detonation wave is governed by a system of partial differential equations parameterized by the working-chamber length *L*, the cross-sectional area *A* of the annular gap, and the angle β between the surface elements of the electrodes. The detonation-wave characteristics were calculated by Whitham's method as functions of the distance *l* traveled by a detonation wave [21]. The Mach number $M_D = M_D(l)$ was determined by solving the differential equation, with the acoustic characteristic c^+ as an integration constant. The flow parameters were expressed in terms of the Mach number of a steady detonation wave. Combining the equations of fluid dynamics, we obtain

$$\frac{dp}{dl} + \rho a\frac{dw}{dl} = \frac{j_h}{w+a} \times\left(\frac{\gamma - l}{\sigma}j_h + \frac{a}{c}B\right) - \frac{\gamma P w}{w+a}\frac{l}{A}\frac{dA}{dl}, \tag{3}$$

$$(w+a)\frac{d}{dl} = \frac{\partial}{\partial t} + (w+a)\frac{\partial}{\partial l}, \tag{4}$$

where γ is the ratio of specific heats of the burned gas, *p* is pressure, *P* is density, *w* is the average gas velocity behind the detonation wave, *a* is the additional velocity induced by electromagnetic acceleration, j_h is electric current, σ is the plasma conductivity, *c* is a dimensionless constant related to the total concentration of positive ions, *t* is time, and *l* is a current location in the working chamber.

The following values were substituted into the equations: $B = 0$, $j_h = \sigma E_0$. Here, E_0 is the averaged vector of electric field strength ahead of the detonation wave. It is calculated as the potential difference across the electrode gap divided by its width. We change from M_D and l to the following dimensionless quantities:

$$Z = \sqrt{l - M_j^2/M_D^2}, \quad x = l/L. \tag{5}$$

Using a known function $Z = Z(x)$, we derive formulas for flow characteristics behind the detonation wave:

$$D = \frac{a_0 M_j}{\sqrt{(l - Z^2)}}, \tag{6}$$

$$P = \frac{P_0 M_j^2 \gamma}{(\gamma + l)(l - Z)}, \tag{7}$$

$$\rho = \frac{\rho_0(\gamma + l)}{(\gamma - Z)}, \tag{8}$$

$$U = \frac{M_j a_0}{\gamma + l}\sqrt{(l + Z)(l - Z)}, \tag{9}$$

$$T = \frac{T_0 \gamma M_j^2 (\gamma - Z)}{(\gamma + l)^2 (l - Z)}, \tag{10}$$

where M_D is the detonation Mach number; p is pressure; ρ is the plasma density; U is the total velocity; and T and T_0 denote the plasma temperature after and before its electromagnetic acceleration, respectively.

The form of $Z(x)$ is determined by solving the following ordinary differential equation in terms of elementary functions:

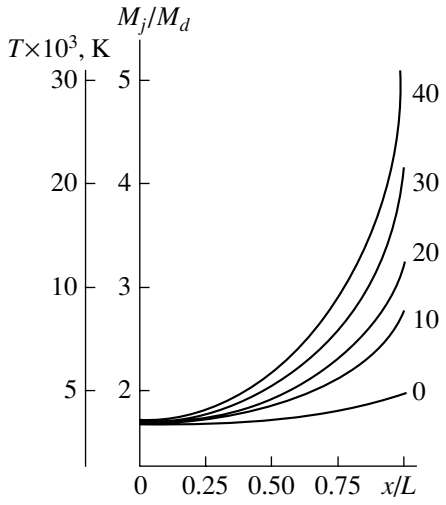


Fig. 4. Increase in detonation intensity and temperature of combustion products in the working chamber at various electric field strengths. Number at curves are the values of E (V/m).

$$\left(l = \sqrt{\frac{\gamma}{(\gamma - Z)(l + Z)}}\right)(1 + Z + \sqrt{\gamma(\gamma - Z)(l + Z)}) \times \frac{dZ}{dx} = K(l - Z)^2\sqrt{l - Z^2} - \gamma(l - Z^2)\frac{l}{A}\frac{dA}{dx}. \tag{11}$$

To calculate the geometric parameters of a plasma gun, we set $\gamma = 1.2$, $M_j = 5.0$ (Mach number of a steady detonation wave), $\rho_0 = 1$ kg/m^3 (initial density of combustion products), and $E_0 = 35$ V/m (electric field strength) and use the total cross-sectional area $A = A(l)$ of the annular gap perpendicular to the working-chamber axis. The ordinary differential equation was solved numerically under the initial condition $Z(x) = 0$. The solution was obtained in implicit form.

The solution to Eq. (11) was represented as a formula for calculating a relative detonation intensity. When $dA/dx = 0$, K is the scaling parameter expressed as

$$K = \frac{\delta E_0^2 (\gamma - l)(\gamma + l)^2 L}{\gamma a_0 \rho_0 M_j^3}, \tag{12}$$

where

$$\delta = c\frac{(kT)^{3/2}}{e^2 \ln\Delta\sqrt{8\pi m_e}}, \tag{13}$$

k is the Boltzmann constant, T is temperature, m_e is the electron mass, e is the electron charge, and $\ln\Delta$ is the Coulomb logarithm (a function of the working-chamber plasma temperature and degree of ionization).

In the numerical integration, we treated $\ln\Delta$ as a constant parameter, setting $\ln\Delta = 10$. We also assumed that the cross-sectional area of the annular gap is constant, $dA/dx = 0$, and $E_0 = 35$ V/m. As a result, we found the scaling parameter as

$$K = K_0\frac{(\gamma - Z)^{3/2}}{(l - Z)^{3/2}}. \tag{14}$$

Figure 4 shows the graphs of T and M_j/Md for the gas behind the detonation front as increasing functions of the current location x/L in the electrode gap for several values of E. The plasma-jet energy exhibits nonlinear growth, and its peak values are reached in the outlet cross section of the plasma gun.

An analysis of Eqs. (6)–(14) shows that the plasma-jet temperature, pressure, velocity, and density can be varied over wide intervals. These characteristics depend on the working-chamber length, cone angle, electrode-gap width, and electric field intensity (Fig. 5).

The results of a numerical analysis of the ordinary differential equation show that the Joule heating begins to contribute to the acceleration of a detonation wave as E_0 is increased. We suppose that the conductivity behind the detonation wave is mainly due to electrons.

This assumption is valid when the ion concentration is comparable to the electron concentration in the gas. As the wave accelerates, the plasma-jet temperature increases, and its velocity and density increase accordingly. The electric field strength in the working chamber decreases as the plasma piston moves further, because the electric field intensity in the gap decreases with increasing gap width, whereas the plasma conductivity increases. These trends stabilize the energy conversion and result in a gradual energy supply to the detonation wave. When the electric field strength is higher than 100 V/m, the plasma conductivity rapidly increases, and a 100-mm working-chamber is sufficiently long for an electric breakdown to occur.

The working-chamber length is the most important determinant of the energy characteristics of a pulsed plasma-jet flow. By varying the working-chamber length, the following technical characteristics of a pulsed plasma jet can be obtained: a jet energy flux of 10^4–10^7 W/cm^2, a temperature of 5×10^3–3×10^4 K, and a velocity of 2–8 km/s. The highest energy of plasma jets produced by the plasma gun are obtained when the working-chamber length exceeds 0.5 m (see Fig. 5).

Our studies showed that the plasma-jet velocity varies insignificantly as the working-chamber length is varied from 0.3 to 0.5 m. This observation was taken into account in optimizing the dimensions of a plasma generator to be used in coating deposition. The plasma gun described in [12, 13, 22] was designed for use in surface modification and high-temperature deposition of high-quality metal, hard-alloy, and metal-oxide coatings. The premixed-gas components and powder are continuously fed into the plasma generator, which improves the cost efficiency of coating deposition technologies and reduces the costs of technological facilities.

ANALYSIS OF COATINGS

The pulsed plasma spray technology was used to deposit an aluminum-oxide coating in [14] and a hard-alloy coating in the present study. The alloy consisted of tungsten carbide (88%) and cobalt (12%). The coating was deposited on a 4 mm thick copper substrate. The deposition rate was about 1 m^2 per hour for a coating of thickness 0.6–0.8 mm.

The 88%WC+12%Co coating was examined by using an ion Rutherford backscattering (RBS) technique, an elastic-resonance technique, diffraction transmission electron microscopy (TEM), and measurements of coating hardness and adhesion to the copper substrate.

The adhesion of the coating to the substrate was determined for specimens made of M-00 grade copper. Ten measurements were performed by indenting the tested surface with a diamond pyramid. The adhesion strength was calculated as $H_v = 4P/b^2$, where P is load

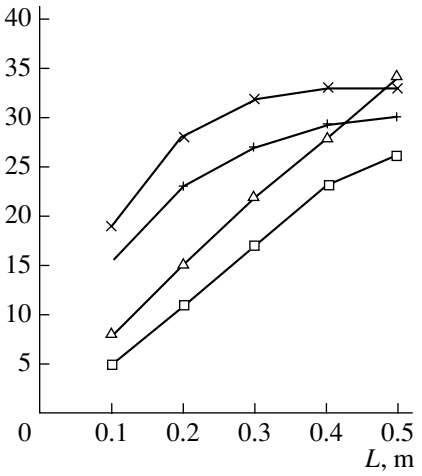


Fig. 5. Characterization of the plasma flow issuing from the working chamber as depending on its length: △—$T \times 1000$ K, +—$P \times 3$ MPa, ×—$V \times 0.3$ km/s, and □—$\rho \times 0.3$ kg/m^3.

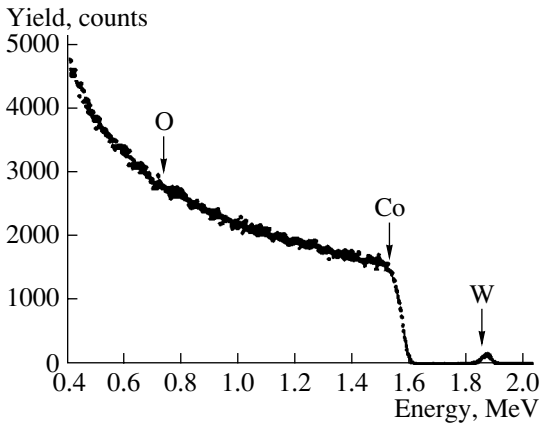


Fig. 6. Energy spectrum of the Rutherford backscattering of He^+ ions; arrows indicate the kinematical limits of elements.

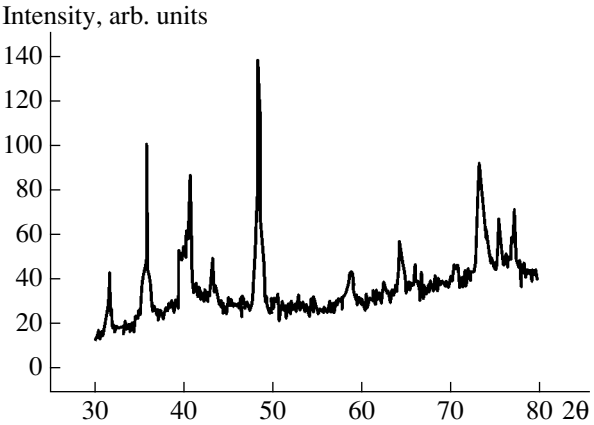


Fig. 7. Fragments of x-ray diffraction patterns obtained for specimens with WC–Co coatings.

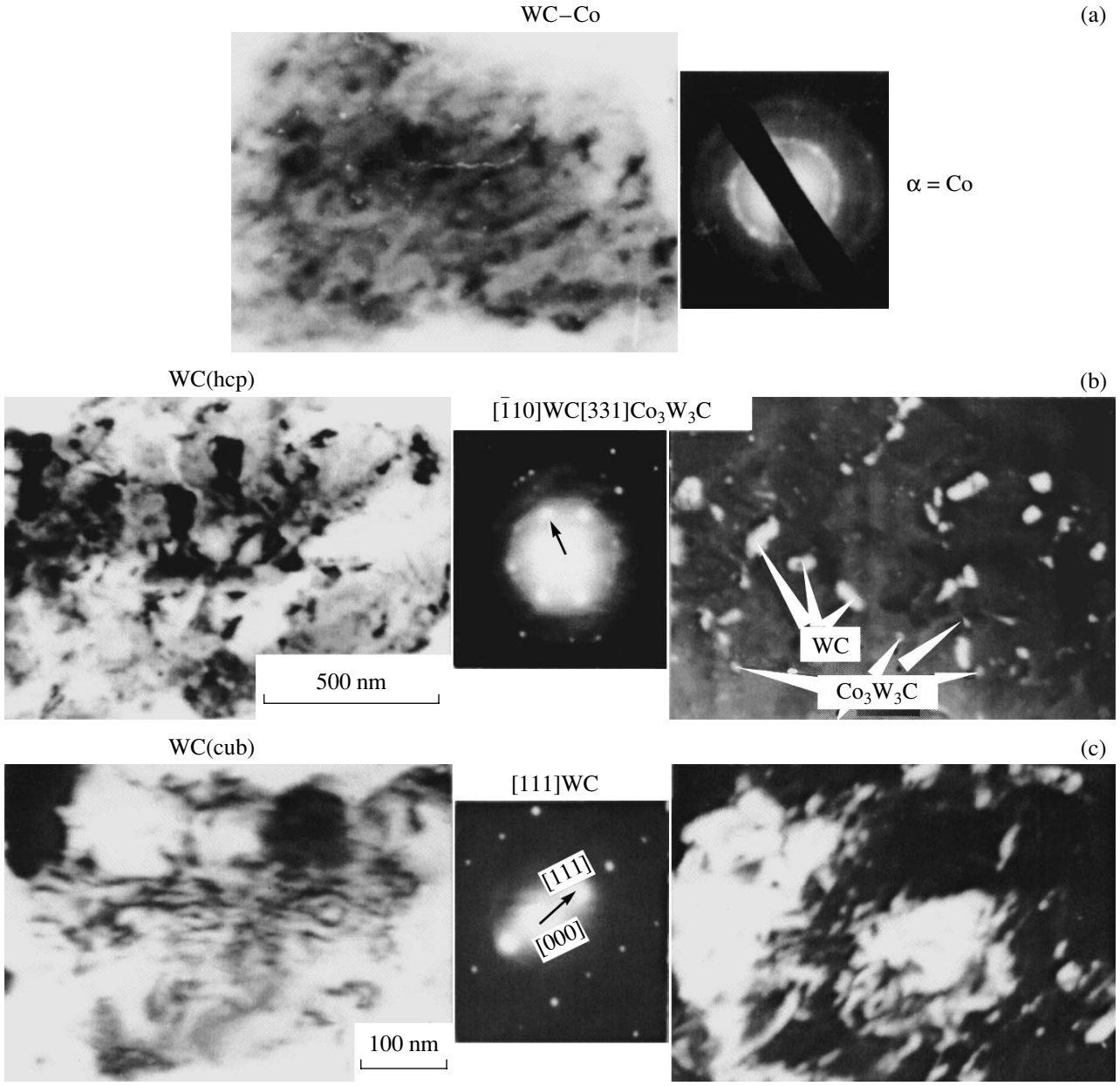


Fig. 8. Diffraction TEM photographic images of a WC–Co coating: (a) cobalt polycrystals with average crystallite size 25 nm in a diffraction pattern of local areas containing Co polycrystals; (b) tungsten-carbide areas with face-centered close-packed lattice, crystallites of average size 0.15 μm, and W_3Co_3C nanoparticles of average size ≈15 nm at grain boundaries (in diffraction patterns of local WC–W_3Co_3C areas); (c) tungsten-carbide areas with cubic lattice and dislocation substructure inside crystallites (dark and light areas) in diffraction patterns of coating areas containing cubic tungsten carbide.

and *b* is the indentation width. Our measurements showed that the average adhesion strength is 250 MPa, while the lowest and highest strengths are 210 and 280 MPa, respectively. The hardness of coatings varied from 8000 to 1.28×10^4 N/mm.

The ion RBS technique and proton elastic resonance provide information about the chemical composition of the outer coating layer. Figure 6 shows the RBS energy spectrum. The spectrum exhibits distinct tungsten and oxygen peaks. The Co kinematical limit is indicated by an arrow. The coating analysis by means of RBS and proton elastic resonance techniques showed that the deposited layer contains Co, W, C, and O in the form of WC_{89}, Co_8, C_2, and O_2. It is worth noting that the tungsten concentration in the layer is low (about 1 at. %), whereas the carbon concentration is high (about 30 at. %).

The results of a chemical analysis of a coating reported in [23] showed that HVOF (High Velocity Oxygen–Fuel) spray deposition increased the tungsten percentage to 84.38%, and deposition by means of a

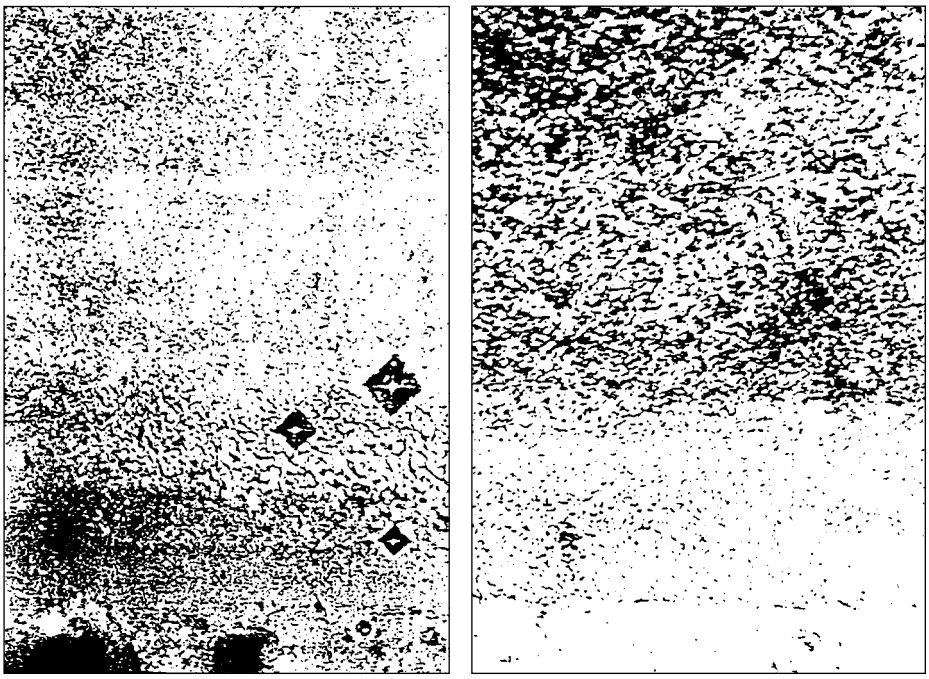


Fig. 9. Transverse thin section of WC–Co coating with microhardness indentations.

high-velocity plasma (HEP process) raised it to 87.38%. Partial coating decrystallization was detected in [24]. The cobalt percentage changed to 12.98 and 9.22%, respectively. The tungsten and cobalt percentages in the powder were 82.9% and 11.61%, respectively. The carbon percentage changed from 4.09% in the powder to 2.54% after an HVOF deposition and to 2.52% after an HEP deposition.

The elemental analysis performed with the use of RBS and proton resonance techniques, as well as x-ray diffraction analysis, demonstrated good agreement of our results with those obtained in [23, 24], where coatings were deposited by means of HVOF and HEP processes. The inner layers of the coating deposited by means of a pulsed plasma spray consist of phases that are virtually identical with those present in the starting powder material.

The analysis of coating phase composition was based on Cu_k emission. The X-ray patterns obtained show that the predominant coating phase is the face-centered close-packed WC lattice. The presence of other phases was inferred from reflections at angles ranging from 37° to 47° (see Fig. 7). Several lines overlap in this angular interval, which complicates the analysis. The interplanar spacings calculated from the reflections that we managed to single out suggest that the following coating phases are present: W_2C, Co_7W_6, Co_3W, W, and hexagonal Co. The composite phases detected in the intercrystallite space are in amorphous states, as observed in [23, 24]. This is explained by the effect of high temperature on the formation of a coating [24].

An analysis of diffraction TEM images of the WC–Co cermet coating showed that the coating has a polycrystalline structure involving WC crystallites with hexagonal lattices, cobalt α- and β-phases, and WC crystallites with cubic lattices (see Fig. 8). The average size of WC crystallites with face-centered close-packed lattices is 0.15 μm, and the average cobalt crystallite size is about 25 nm. W_3Co_3C phase particles of size 15 nm were observed at crystallite edges. A dislocation substructure was observed inside the WC crystallites with cubic lattices.

Figure 9 shows a transverse thin section of a WC–Co coating with indentations made by the diamond pyramid of a hardness tester (the scale is 1 cm ≈ 200 μm). It demonstrates that the coating hardness varies within 800–12 800 N/mm^2.

CONCLUSIONS

The magnetogasdynamic acceleration of detonation products is described. Based on mathematical modeling and calculations, a method is proposed for creating high-energy pulsed plasma jets. Design parameters are calculated for a plasma gun to be used in coating deposition. The plasma-jet energy flux can be varied from 10^4 to 10^7 W/cm^2; temperature, from 5×10^3 to 3×10^4 K; velocity, from 2000 to 8000 m/s.

The plasma gun was used to deposit a 88%WC + 12%Co powder coating on a copper substrate. An analysis of phase composition showed that the coating consists of WC crystallites with hexagonal and cubic lattices of average size 0.15 μm, cobalt α- and β-phases of

size about 25 nm, and the composite carbide W_3Co_3C at crystallite edges. An x-ray diffraction analysis revealed the presence of W_2C, Co_7W_6, Co_3W, W, and hexagonal Co in the coating. The coating hardness was found to reach 1.28×10^4 N/mm^2; the coating–substrate adhesion, 250 MPa.

ACKNOWLEDGMENTS

This work was supported, in part, by the Ukrainian Scientific and Technical Center, project no. 1472. We appreciate the support provided by the National Research Institute for Metals, which made possible the studies conducted in Tsukuba, Japan.

REFERENCES

1. T. P. Gavrilenko, Y. A. Nikolaev, and V. Y. Ulianitsky, in *Proceedings of the 15th International Thermal Spray Conference, Nice, 1998*, pp. 1485–1488.
2. M. L. Thorpe and H. J. Richter, in *Proceedings of the 13th International Thermal Spray Conference "Thermal Spray: International Advances in Coatings Technology," Orlando, 1992* (ASM International, Materials Park, 1992), pp. 137–148.
3. H. Kreye, R. Schweitzker, and S. Zimmerman, in *Proceedings of the 9th National Thermal Spray Conference "Thermal Spray: Practical Solutions for Engineering Problems, 1996* (ASM International, Materials Park, 1996), pp. 450–456.
4. G. R. Heath and R. J. Dumola, in *Proceedings of the 15th International Thermal Spray Conference, Nice, 1998*, pp. 1495–1500.
5. M. Miyamoto, T. Sakurai, and M. Tago, in *Proceedings of the 15th International Thermal Spray Conference, Nice, 1998*, pp. 1501–1506.
6. J. R. Uglum, J. L. Bacon, D. G. Davis, *et al.*, in *Proceedings of the 1st United Thermal Spray Conference, 1997* (ASM International, Materials Park, 1998), pp. 373–391.
7. A. D. Lebedev and B. A. Uryukov, *High Pressure Pulsed Plasma Accelerators* [in Russian] (Novosibirsk, 1990).
8. R. B. Bhagat, M. F. Amatean, A. Papyrin, *et al.*, in *Proceedings of the 1st United Thermal Spray Conference, 1997* (ASM International, Materials Park, 1998), pp. 361–376.
9. Yu. A. Kharlamov, *Gas-Detonation Facilities for Deposition of Coatings* [in Russian] (Voroshilovgrad, 1998).
10. A. I. Zverev, S. Yu. Sharivker, and E. A. Astakhov, *Detonation Deposition of Coatings* (Sudostroenie, Leningrad, 1979).
11. A. P. Semenov, Yu. P. Fed'ko, and A. I. Grigorov, in *Survey of Mechanical Engineering Institute* (1977), p. 167.
12. Yu. N. Tyurin and A. P. Arzubov, USSR Inventor's Certificate No. 1045491 (1983).
13. K. A. Yuschenko, Y. S. Borisov, and Y. N. Tuyrin, European Patent No. 0531527. A1 (1991).
14. Y. N. Tyurin and A. D. Pogrebnjak, Surf. Coat. Technol. **111**, 269 (1999).
15. V. A. Levin, Doctoral Dissertation in Engineering, Moscow (1975).
16. D. H. Edwards and T. R. Lawrence, Proc. R. Soc. London, Ser. A **286**, 415 (1965).
17. I. B. Helliwell, J. Fluid Mech. **16**, 243 (1963).
18. G. D. Salamandra, Fiz. Goreniya Vzryva **12**, 229 (1976).
19. Yu. A. Burenin and G. A. Shvetsov, Fiz. Goreniya Vzryva **13**, 130 (1977).
20. R. I. Soloukhin, Yu. A. Yakobi, and V. I. Yakovlev, Fiz. Goreniya Vzryva **13**, 481 (1977).
21. G. B. Whitham, *Linear and Nonlinear Waves* (Wiley, New York, 1974; Mir, Moscow, 1977).
22. Yu. N. Tyurin, USSR Inventor's Certificate No. 879 862 (1981).
23. J. Nerz, B. Kushner, and A. Rotolico, J. Therm. Spray Technol. **1** (2), 147 (1992).
24. C. J. Li, A. Ohmori, and Y. Harada, J. Therm. Spray Technol. **5** (1), 69 (1996).

Translated by A. Betev

Technical Physics, Vol. 46, No. 7, 2001, pp. 905–909. Translated from Zhurnal Tekhnicheskoĭ Fiziki, Vol. 71, No. 7, 2001, pp. 119–123.

EXPERIMENTAL INSTRUMENTS AND TECHNIQUES

Pulsed Heating of Thin Powder Layers by Intense Microwaves

G. M. Batanov, N. K. Berezhetskaya, I. A. Kossyĭ, A. N. Magunov, and V. P. Silakov

General Physics Institute, Russian Academy of Sciences, ul. Vavilova 38, Moscow, 117942 Russia

e-mail: kossyi@fpl.gpi.ru

Received November 13, 2000

Abstract—Thin (0.05 cm) layers of mechanical mixtures of conductive and insulating powders are heated by microwave pulses with an intensity of about 10 kW/cm^2. For mW-power microwaves, the absorption thickness in the mixtures is found to be on the order of 1 cm. For intensities of 10 kW/cm^2 and above, pulse durations of 5–8 ms, and the number of pulses in a train of 5 or less, the powder layer melts and the characteristic spatial scale of absorption decreases to ≈0.05 cm. The reflection factor drops during a microwave pulse. Ways of improving the absorption efficiency by initiating surface breakdowns and plasma formation in pores between coarse grains in the bulk of the powders are considered. © *2001 MAIK "Nauka/Interperiodica".*

(1) In recent years, the microwave heating of various materials has provoked much interest (see, e.g., [1]). For ceramic (oxide) powders, its efficiency is not high because of low dielectric losses. It, however, can be improved by passing to the mm-wave range [2]. A particular challenge is the efficient heating of thin powder layers.

Therefore, materials that have a large absorption factor in a wide wavelength range are of great interest. Examples are those obtained by sintering powders of metals and their oxides at high temperatures [3]. Unfortunately, the characteristic absorption thickness and the type of conduction for such materials remain a mystery; it is known, however, that they offer direct-current conduction as well.

It has been found [4] that surface discharges may appear at high (≈10^4 W/cm^2 or higher) radiation intensities if conductive surface inclusions are embedded in an insulating matrix. In this case, the energy is released near the surface, the release being associated with breakdowns at the metal–insulator interface and the formation of a thin plasma in the gas near the surface.

Two questions arise in this respect. What is the absorption factor in a heterogeneous mechanical mixture of powders of metals and their oxides? If the absorption factor is high, is nonlinear absorption similar to that taking place at surface breakdowns on heterogeneous surfaces possible?

(2) In experiments, we used powders of silicon, titanium, aluminum, tin oxide, lead oxide, copper oxide, alumina, as well as powders of their mixtures. The grain sizes were 10–40 μm on average. The powders were placed between 2-mm-thick glass plates, and the plates were ground together until a 0.2- to 0.5-mm-thick powder layer forms. The samples were fixed in a special support and placed into the caustic surface of a gyrotron beam normally to the axis of a microwave beam (Fig. 1). The wavelength and the power of the gyrotron radiation were 4 mm and 120–180 kW (power density at the beam axis 10–15 kW/cm^2), respectively. A Gaussian beam was used. The gyrotron power and the signal reflected from the sample were controlled with a quasi-optical coupler. The microwave pulse duration was varied between 1 and 8 ms. The time spacing between pulses in a train was 2 min.

With a powder layer placed between radiotransparent plates, surface exposure to the beam is eliminated; in this way, the occurrence of a discharge due to microwave breakdowns and its propagation toward the beam inside the powder are avoided.

The properties of the powder layers were estimated at a low (milliwatt) power level on a specially designed quasi-optical bench. Measurements made on reference quartz plates of different thickness showed that the sum

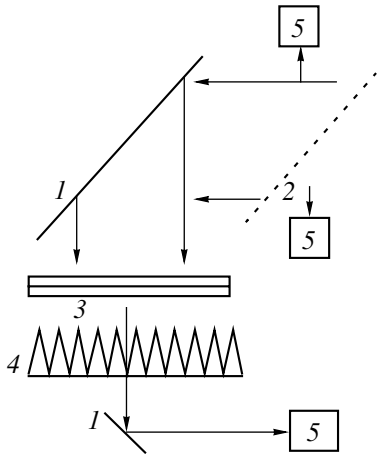


Fig. 1. Heating of thin conductive + insulating powder mixtures: (*1*) metal mirrors; (*2*) mica plate of quasi-optical coupler; (*3*) sample; (*4*) absorbing loads; and (*5*) microwave detectors.

1063-7842/01/4607-0905$21.00 © 2001 MAIK "Nauka/Interperiodica"

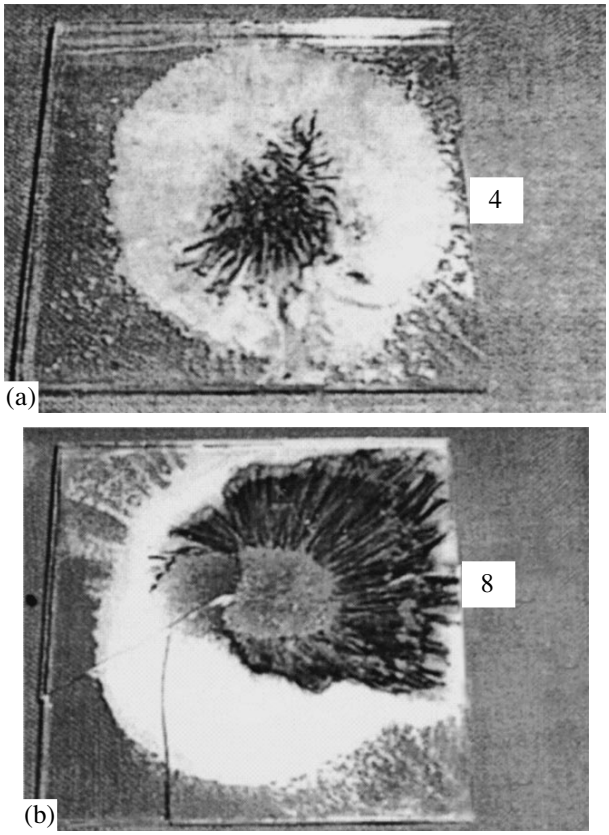


Fig. 2. Samples irradiated: (a) 25% of Si powder + 75% of PbO powder, intensity ~10 kW/cm^2, five 7-ms pulses; (b) 50% Ti + 50% PbO, intensity ~10 kW/cm^2, one 8-ms pulse.

of the reflection and transmission factors is no more than ≈0.05 in error.

For the glass plate used in the experiments, the reflection factor was found to be R = 0.25–0.32 and the transmission factor, T = 0.46–0.53; hence, the absorption factor A lies between 0.22 and 0.23. For the two glass plates, R = 0.40–0.43, $T \approx 0.34$, and A = 0.23–0.26. Such high values of the reflection and transmission factors may introduce high errors in the properties of the layers estimated. Therefore, for the powder characterization at the low power level of the beam, we employed a half-wave (1 mm) plate made of fused quartz. The table lists results for the different powders.

Table

Composition	R	T	$R + T$	A
PbO	0.143	0.907	1.043	−0.043
CuO	0.112	0.928	1.04	−0.04
Al	0.912	0.138	1.05	−0.05
Si	0.313	0.295	0.608	+0.392
SnO	0.38	0.16	0.54	+0.46
0.5Ti + 0.5CuO	0.646	0.422	1.068	−0.068
0.5Si + 0.5Al_2O_3	0.525	0.422	0.947	+0.05
0.5SnO + 0.5Al_2O_3	0.617	0.389	1.006	−0.006
0.5Al + 0.5PbO	0.759	0.257	1.016	−0.016

As follows from the table, the thin insulating powders of lead and copper oxides do not absorb, and the sum $R + T$ for them deviates from unity within the measurement error. These powders also feature a low reflection factor.

The aluminum powder has the high reflection factor. Presumably, this is due to contacts between the grains. This powder behaves as a continuous medium. However, the radiation power transmitted through the dispersed material is also significant.

In the silicon and tin oxide powders, which conduct like semiconductors, the absorption is high. This parameter is close to the absorption factor of a thin (0.45 mm) *n*-type silicon wafer (phosphorus-doped to a resistivity of 4.5 Ω cm). As for the aluminum powder, such behavior is apparently associated with intergranular contacts.

The mixtures of the conductive and insulating powders absorb to a minor extent, but their reflections are rather high (≈0.5) (see the table).

(3) When the mixture of the silicon and lead oxide powders (1 : 1) is subjected to a single microwave pulse of intensity higher than 10 kW/cm^2 and duration of 5–7 ms, one or two small molten spots appear on the surface. After 2–4 pulses, the small local spots expand, and after five pulses, they merge into a large continuous spot. A similar situation is observed for the mixture of tin oxide and lead oxide powders (1 : 1) (Fig. 2). Note that the melting process is nonuniform across the depth: on the side of beam incidence, the material solidified is strongly adhered to the glass substrate, while near the second (lower) glass, the mixture remains intact in some areas.

When the duration pulse decreases to 2 ms, melting spots are not observed after four or five pulses. Similarly, no such spots appear when the pulses last 7 or 8 ms but their intensity drops to 4–5 kW/cm^2.

A somewhat different situation arises in the case of the titanium + PbO (1 : 1) mixture. A continuous melting spot appears even after the application of two 4-ms 10-kW/cm^2 pulses.

The titanium and copper oxide in the mixture (1 : 1) vigorously react to form metallic copper. This reaction is initiated when the power density is reduced to 5.0–5.5 kW/cm^2 and the pulse narrows to 1 ms. As the power density is reduced further, longer pulses initiate the reaction.

Finally, a radically distinct situation is observed for the titanium + alumina (1 : 1) powder. At the power density 8–9 kW/cm^2 and pulse duration 7 ms, only individual melting spots appear. Each pulse radially splashes the powder between the plates, which is apparently associated with the heating of the gas in the pores.

Radial powder splashing from the beam center is also observed in the case of the silicon powder. Here, again, gas heating in the pores between the grains is a possible reason for the splashing.

The variation of the signal reflected from the targets during a pulse is demonstrated in Fig. 3. The signal varies with the number of pulses in a train. As follows from the oscillograms (Fig. 3a), it smoothly decreases even during the first pulse. In the next pulse, the signal drops twofold for 2–3 ms. In subsequent pulses, the signal drops within 1 ms and its final level continuously decreases.

To study the variation of the electrical parameters of the powder mixtures under the action of microwave pulses, we performed special measurements with the above-mentioned microwave bench. The radiation parameters were the following: wavelength 2.5 cm, power density about 10 kW/cm^2, and pulse duration 5 μs. Coaxial electrodes terminated by a 50 Ω-resistor were placed into the sample between the plates, and a radiation-induced voltage pulse was measured. A unipolar (negative) voltage pulse arose on the central electrode only if the silicon powder or the silicon + copper oxide mixture was placed between the plates (Fig. 4). The average levels of the signal for both the silicon powder and its mixture with copper oxide (1 : 3 in volume fractions) were the same, about 0.2 V. For the mixtures in proportions 1 : 1 and 3 : 1, the mean level of the signal rose to 1.0–1.5 V. For the 1 : 1 and 1 : 3 mixtures, bright flashes on the surface were observed. For the silicon powder and its mixture with copper oxide in proportion 3 : 1, the flashes were absent.

(4) According to [5], the characteristic absorption thickness in a planar layer of a mixture of conductive and nonconductive balls is given by

$$h_0 \approx \lambda|\varepsilon|^2/18\pi\alpha\varepsilon_i \quad \text{at} \quad \delta = c/(4\pi\sigma\omega)^{12} > a_0,$$

$$h_0 \approx (4a_0/9\alpha)(\lambda\sigma/c)^{1/2} \quad \text{at} \quad a_0 > \delta \quad \text{and} \quad |\varepsilon| \gg 1.$$

Here, $a_0 \ll \lambda$ is the radius of the balls, $n_0 = 3\alpha/4\pi a_0^3$ is the concentration of the conductive balls, ε is the dielectric constant of the conductive balls, ε_i is its imaginary part, σ is the conductivity of the conductive balls, ω is the circular frequency of the radiation, α is the volume fraction occupied by the conductive balls, and δ is the skin depth. For the silicon + alumina mixture, $\delta > a_0$ and $h_0 \approx 0.44$ cm if $\alpha = 0.5$. In the case of the titanium + copper oxide powder, $\delta < a_0$ and $h_0 \approx 0.4$ cm for the fraction of diameter 4×10^{-3} cm and $\alpha = 0.5$. Hence, at the characteristic thickness of the powder layer $h = 0.05$ cm, the absorption factor $A = (1 - R)[1 - \exp(-h/h_0)] = 0.05$–$0.06$, which agrees with experimental data obtained at the low power level.

That the powder melts when irradiated by a train of intense pulses (about 10 kW/cm^2) suggests that the absorption factor grows and becomes much greater than that observed at the weak fields. The melting non-

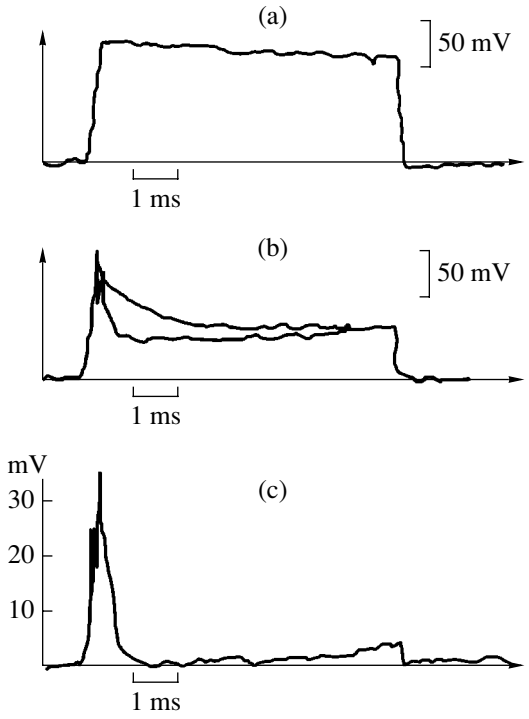


Fig. 3. Variation of the reflection pulse with the number of microwave pulses applied for 25% Si + 75% PbO: (a) first pulse, (b) second and third pulses, and (c) fourth pulse.

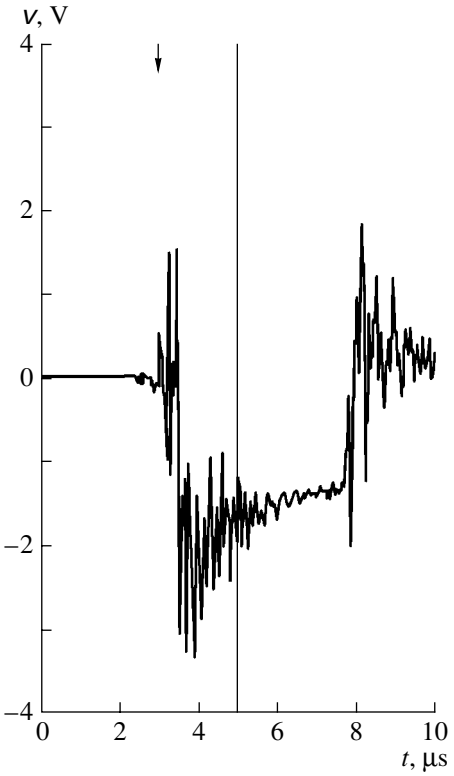


Fig. 4. Voltage pulse across the coaxial electrodes for 50% Si + 50% CuO. The termination is 50 Ω.

uniformity across the sample means that the characteristic absorption thickness is less than the layer thickness (0.05 cm); that is, it dropped by one order of magnitude compared with the characteristic absorption scale at the low intensities.

When estimating the reflection coefficient, one should take into account that the waves reflected by individual grains are coherent, since the intergranular spacing is much smaller than the radiation wavelength. The value of the reflection coefficient must be proportional to the reflection cross section of an individual grain. For a conducting ball, this section is given by $S_r = 8\pi(2\pi)^4 a_0^6/3\lambda^4$ [5]. For an absorbing nonconducting ball, this section should be smaller by a factor of $(\varepsilon - 1)^2/(\varepsilon + 2)^2$. For quartz (as in our case), the coefficient equals 0.25. At the low intensities, this seems to be a possible reason for the discrepancy between the reflection coefficient of the lead oxide + copper oxide powder mixture and that for the powder mixtures half-composed of conducting grains. However, such a discrepancy is also observed when the nonconducting and conducting grains reflect the intense microwave radiation. This allows us to draw a conclusion concerning processes taking place in the powders irradiated by high-amplitude waves. We see (Fig. 3) that the signal reflected drastically drops during an intense pulse. The question now arises as to why the scattering cross section of a conducting ball decreases. It seems likely that we are dealing with the phenomenon of surface microwave breakdown akin to that observed in [4]. In fact, the electric field of the wave (≈3 kV/cm) is much less than the breakdown field for air (≈30 kV/cm). In pores between the grains, the latter is still higher because, for the grain size 10–40 μm, the rate of electron loss by diffusion, $\nu_d \approx D_e/a_0^2 \approx 10^9$ s^{-1}, far exceeds the rate of three-body electron attachment in air, $\nu_a \approx 10^8$ s^{-1}, under normal pressure. The three-fold enhancement of the field at the conductive balls is insufficient for the breakdown threshold to be overcome. Therefore, a plasma can possibly be produced only at the surface breakdown of the insulator [4]. Pores filled by the plasma represent a so-called "cellular plasma," since this medium is, in essence, plasma-filled cells in the powder. As the concentration of the charged particles in the pores grows, so does the absorption factor until the skin depth becomes less than the pore diameter. Subsequently, the absorption decreases and the concentration ceases to rise. It can be assumed that the absorption is the highest at $\delta \approx a_0$, since the pore size is on the order of a_0. Then, from this condition, one can easily derive the minimum effective absorption thickness: $h_{\min} \approx 2^{1/2}\lambda/9\pi\alpha = 0.02$ cm, which roughly coincides with the experiment. In this case, we have a three-component system similar to that considered in [6].

Evidence in support of the above assumptions is unipolar voltage pulses generated in the powders with the coaxial electrodes (Fig. 4). It is just these signals that were recorded in a pulsed microwave discharge in gases under high pressures [7].

It follows from the aforesaid that the high absorption factors are associated with plasma-assisted surface breakdowns. Then, it remains to be seen whether the energy absorbed is sufficient to melt the powder components. If the absorption factor of the plasma-containing medium is set equal to unity ($A_{pl} \approx 1$), it is easy to evaluate the energy density released in the powder volume per pulse:

$$q_0 = s\tau/h \approx 10^4 \times 7 \times 10^{-3}/0.05 \approx 1.4 \text{ kJ/cm}^{-3}.$$

The energy necessary for melting is estimated from the relationship

$$q \approx \alpha(\Delta H_1 + Cp_1\Delta T_i)\rho_1/M_1 + (1-\alpha)Cp_2\Delta T\rho_2/M_2.$$

Here, ΔH_1 is the energy necessary for melting the material with the lowest melting point; and Cp, ρ, and M are the specific heat at constant pressure, density, and molecular weight of the powder components, respectively. For the silicon + lead oxide mixture, $q \approx 1.7$ kJ/cm^3. Thus, the pulse durations and intensity used in this work are appropriate in terms of the energy consumption. The formation of the molten spots after initial pulses suggests that the absorption cross section in some regions exceeds their geometric areas. The increase in the molten area after the application of the second and following pulses may indicate the formation of a semiconducting phase.

That the mixtures offer high local absorption factors is most impressively supported by experiments on the initiation of the solid-phase redox reaction in the 0.5Ti + 0.5CuO mixture. In this mixture, a high temperature (~10^3 K) is required to initiate the reaction. However, even for $A \approx 1$, the mean temperature of the mixture cannot exceed 200 K.

Another evidence in favor of the high locality of energy evolution is the formation of molten spots in the $0.5Al + 0.5Al_2O_3$ mixture. Here, to melt aluminum, the energy density of about 3 kJ/cm^3 is necessary, while the mean energy density in the related experiment could not be higher than 1.6 kJ/cm^3.

(5) Let us summarize the basic results of this article. We measured the reflection and absorption factors of thin (0.05 cm) layers of insulating powders, as well of their mixtures with semiconductor and metallic powders. The measurements were performed with the quasi-optical method at the low level of the microwave power. The absorption factors were found to be low (no more than 0.05), which agrees with the calculations for the mechanical mixtures considered. The reflection factors of the mixtures with conductive grains turned our to exceed those of the insulating powders several-fold.

It was established that individual melting spots appear on the surface of the conductive + insulating powder mixtures once a pulse of intensity about 10 kW/cm^2 has been applied. With a train of 4 to

7 pulses of duration 5–8 ms, the spots merge into a continuous molten area. The intense local heating of the material was observed. Melting is nonuniform across the powders, which means that the characteristic absorption scale is less than 0.05 cm.

The reflection factor of the powder mixtures drops during the microwave pulse. A voltage pulse is generated across the electrodes placed into the powder. The results obtained at high intensities suggest the initiation of surface breakdowns where insulating and conductive grains come into contact, and also the formation of a plasma between pores. In our opinion, it is plasma formation that is responsible for the effective absorption of microwaves.

ACKNOWLEDGMENTS

The authors thank L. Kolik and V. Plotnikov, who provided the operation of the gyrotron, and N. Malykh for the assistance in low-power quasi-optical measurements.

This work was supported by the International Science and Technology Center (contract no. 908)

REFERENCES

1. D. J. Agrawal, J. Mater. Educ. **19**, 49 (1999).
2. Yu. V. Bykov and A. G. Eremeev, in *High-Frequency Discharge in Wave Fields* (Inst. Prikl. Fiz., Gorki, 1988), pp. 265–289.
3. Yu. N. Kovneristyĭ, I. N. Lazareva, and A. A. Ravaev, *Microwave Radiation Absorbents* (Nauka, Moscow, 1982).
4. G. M. Batanov, E. F. Bol'shakov, A. A. Dorofeyuk, *et al.*, J. Phys. D **29**, 1641 (1996).
5. L. D. Landau and E. M. Lifshitz, in *Course of Theoretical Physics*, Vol. 8: *Electrodynamics of Continuous Media* (Nauka, Moscow, 1982; Pergamon, New York, 1984), Parag. 93, 92.
6. E. A. Galstyan and A. V. Ravaev, Izv. Vyssh. Uchebn. Zaved., Radiofiz. **30**, 1243 (1987).
7. S. I. Gritsinan, A. A. Dorofeyuk, I. A. Kossyĭ, and A. N. Magunov, Teplofiz. Vys. Temp. **25**, 1068 (1987).

Translated by V. Isaakyan

Technical Physics, Vol. 46, No. 7, 2001, pp. 910–913. Translated from Zhurnal Tekhnicheskoĭ Fiziki, Vol. 71, No. 7, 2001, pp. 124–127.

BRIEF COMMUNICATIONS

Oscillatory Regimes of Vapor Condensation

V. A. Schweigert and A. L. Alexandrov

Institute of Theoretical and Applied Mechanics, Siberian Division, Russian Academy of Sciences, Novosibirsk, 630090 Russia
e-mail: schweig@site.itam.nsc.ru

Received July 17, 2000

Abstract—The problem of the condensation of supersaturated vapor in an open system at a constant rate of production of a monomer and a continuous flow of a carrier gas that removes the products of condensation from the system is considered. It is shown both analytically and by numerical experiment that with a decreasing rate of the carrier gas to below a critical magnitude, the condensation regime becomes oscillatory; namely, time oscillations of the cluster-size distribution in the vapor to be condensed are set in. The cause of the phenomenon is in the suppression of the rate of nucleation and the presence of large clusters. © *2001 MAIK "Nauka/Interperiodica".*

INTRODUCTION

The nucleation and growth of clusters in a supersaturated vapor are one of the most common phenomena in nature and engineering. In the simplest case, the kinetics of condensation is described by the attachment and evaporation of monomer species, i.e., $A_m + A \longleftrightarrow A_{m+1}$.

The energy of attachment of a monomer $\varepsilon_m = E_m - E_{m-1}$, where E_m is the total binding energy of a cluster, increases with increasing cluster size to a limiting value ε_0 for a macroscopic sample of the substance. There exists a critical nucleus size m_* such that for clusters of size $m < m_*$ the evaporation of monomer species is predominant, while the clusters with $m > m_*$ grow steadily. According to Zel'dovich's classical theory of condensation applicable to the case $m_* \gg 1$, the rate of condensation is determined by the probability of the appearance of clusters of critical size.

For a fixed amount of a substance to be condensed, the appearance of a sufficient number of clusters exceeding the critical size is a decisive factor in the process of condensation. Because of the decrease in the concentration of the monomer, the critical size increases and the appearance of new condensation centers ceases. A redistribution of the substance from clusters $m < m_*$ to clusters $m > m_*$ occurs [1]. Such is the picture of condensation in a closed system with initial conditions in the form of a supersaturated monomer existing in the system.

We consider an open system, where the monomers appear continuously over the whole volume due to some external source and the condensation products are removed from the system by a flow of a carrier gas, also uniformly over the whole volume. It can be shown analytically that the equilibrium cluster-size distribution in such a system becomes unstable as the flow rate decreases. The results of numerical experiments show that this instability leads to the appearance of oscillations in the cluster-size distribution function.

THEORY OF OSCILLATORY REGIMES OF CONDENSATION

The kinetics of the "monomeric" growth in an open reactor with a spatially uniform concentration of clusters n_m ($m = 2,\ldots, \infty$) can be described by the following equations:

$$\frac{\partial n_m}{\partial t} = j_{m-1,m} - j_{m,m+1} - \frac{n_m}{\tau}, \qquad (1)$$

where $j_{m-1,m} = \nu_{m-1} n_1 n_{m-1} - \eta_m n_m$ is the rate of cluster growth from the size $m-1$ to the size m; τ is the time of the gas flow through the reactor; $\nu_m = \nu_0 \nu'_m$ and $\eta_m = \nu_{m-1} n_s(-\beta(\varepsilon_m - \varepsilon_0))$ are the rate constants of the reactions of the monomer condensation and evaporation, respectively; $\beta = 1/kT$; T is the gas temperature; and n_s is the density of the saturated vapor.

For simplicity, the clusters are considered to be spherical structureless particles, so that $\nu'_m = (m^{1/3} + 1)^{2/3} \times (1 + 1/m)^{-1/2}$, and ν_0 is a constant geometrical factor that describes the cross sections of the reactions of condensation. Since the rate of removal is the same for all the clusters, the concentration of the monomer n_1 is given by the condition of the conservation of the total amount of the substance

$$\sum_{m=1}^{\infty} m n_m = \text{const}. \qquad (2)$$

For large clusters ($m \gg 1$), the following approximation can be used for the binding energy: $E_m = m\varepsilon_0 - Cm^{2/3}$, where the second term describes the effect of the

1063-7842/01/4607-0910$21.00

surface (the constant C characterizes the surface tension of the substance to be condensed). Then, $\varepsilon_m \approx dE_m/dm = \varepsilon_0 - 2C/3m^{1/3}$, and the critical size of the nucleus $m_* = (2\beta C/3\ln(n_1/n_s))^3$ can be found from the condition of equilibrium between condensation and evaporation $j_{m_*, m_*-1} = 0$. We consider the case $m_* \gg 1$; however, the critical size is assumed to be sufficiently small, so that the characteristic time of nucleation is small in comparison with both the time of removal of the condensation products τ and the period of oscillations.

In this case, the steady-state cluster-size distribution for $m < m_*$ is found from the condition of quasi-equilibrium ($\eta_m n_m \gg j_{m, m-1} \approx 0$) as

$$n_{m \le m_*} \approx n_1 \exp(m\ln(n_1/n_s) + \beta(E_m - m\varepsilon_0)). \quad (3)$$

For clusters of sizes $m > m_*$, the usual approximation of the theory of condensation is applicable, i.e., the neglect of the process of evaporation and the description of the cluster size by a continuous variable rather than by a discrete one

$$\frac{\partial n_m}{\partial t} + v_0 n_1 \frac{\partial}{\partial m} m^{2/3} n_m = -\frac{n_m}{\tau} \quad (4)$$

with a boundary condition $n_{m=m_*} = \psi(n_1) \equiv n_1 \exp(-4\beta^3 C^3/27\ln(n_1/n_s)^2)$, which follows from Eq. (3) and the above expression for m_*. Consider the stability of the quasi-equilibrium solution

$$n_1, n_m = \psi(n_1)(m_*/m)^{2/3}\exp(-\gamma),$$

$$\gamma = 3(m^{1/3} - m_*^{1/3})/v_0\tau$$

with respect to small perturbations δn_1, $\delta n_n \propto \exp(\lambda t)$. The solution of the linearized equation (4) has the following form:

$$\delta n_m = \delta n_1 \left(\frac{m_*}{m}\right)^{2/3} \left(\frac{\partial \psi}{\partial n_1} e^{-\gamma(1+\lambda\tau)} + \frac{\psi}{\lambda\tau n_1} e^{-\gamma}\right).$$

By substituting the last expression into the normalization condition (2) and neglecting the contribution from clusters of size $1 < m \le m_*$, we obtain the following dispersion relation: $1 - 2m_*^{2/3}\psi(v_0 n_1 \tau)^4 \Omega/9n_1 = 0$, which is valid at $3m_*^{1/3} \ll v_0 n_1 \tau$, when the size distribution for small clusters $m \le m_*$ is steady-state. Here, $\Omega = -(1/\lambda\tau + m_*/(1+\lambda\tau)^4)$ and in the derivative $\partial\psi/\partial n_1 \approx m_*\psi/n_1$ we allow for only exponentially fast changes in ψ. At the boundary of the stability region, the real part of the eigenvalue λ is zero and $\lambda = i\omega$, which corresponds to undamped oscillations of the solution in time. Then, the frequency of oscillations ω is found from the condition $\mathrm{Im}(\Omega(i\omega)) = 0$ in the form

$$m_* = \frac{(1+\omega^2\tau^2)^4}{\omega^2\tau^2(\omega^2\tau^2 - 1)}.$$

The last equation has two positive roots $\omega\tau > k = \sqrt{5+\sqrt{17}/2} \approx 1.510$ and $\omega\tau < k$; we chose the smaller one. Thus, the dimensionless oscillation frequencies lie in a narrow range $1 < \omega\tau \le k$, and the oscillation period is determined by τ, i.e., the time during which the gas passes through the reactor. The minimum critical size $m_* \approx 39.66$ that allows the development of instability corresponds to $\omega\tau = k$. The equilibrium cluster distribution becomes unstable for τ greater than the critical value

$$\tau_* = \frac{m^{-5/12}}{v_0 n_s}\left(\frac{9}{2\mathrm{Re}(\Omega)}\right)^{1/4}\left(\frac{n_1}{n_s}\right)^{m_*/8-1},$$

where $\mathrm{Re}(\Omega) = m_*(6\omega^2\tau^2 - 1 - \omega^4\tau^4)/(1+\omega^2\tau^2)^4$.

Numerical solutions of Eqs. (1) and (2) show that the development of instability causes self-sustaining oscillations of the cluster-size distribution function at $\tau > \tau_*$. To experimentally observe these oscillations, one should ensure a monomer density $n_1 = 2$ to $10n_s$ at moderate gas temperatures $\beta C = 3$ to 8 (Fig. 1).

NUMERICAL EXPERIMENT

The above consideration of condensation was performed in the approximation of monomeric growth. Therefore, to check the validity of the results obtained,

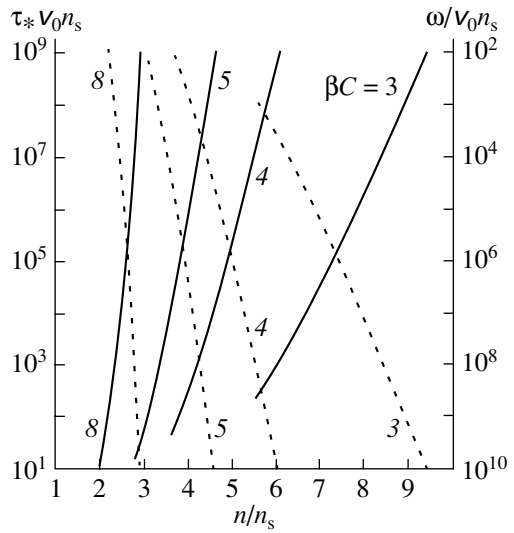


Fig. 1. Dimensionless critical value τ_* of the characteristic time during which the carrier-gas passes through the reactor (solid lines) and the corresponding frequency of oscillations of the solution (dashed lines) for various values of the parameter $\beta C = 3$ to 8.

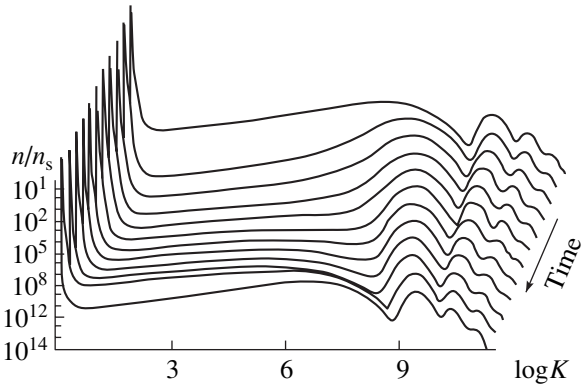


Fig. 2. One period of oscillations of the cluster concentration depending on the cluster size plotted as a function of the logarithm of the size and normalized to the density of the saturated vapor n_s. The first and the last curves are identical.

we additionally performed a more accurate numerical simulation of the process of condensation in an open system with a constant generation of monomers and allowance for all of the reactions of coalescence of growing clusters ($A_m + A_k \longrightarrow A_{m+k}$), including the attachment of dimers, trimers, and all greater clusters. Under conditions of high supersaturations and cluster concentrations that are much greater than those in the equilibrium vapor, the back reactions of cluster decomposition ($A_{m+k} \longrightarrow A_m + A_k$) are relatively slow and can be neglected. Although this approximation has not been substantiated strictly, it is admissible for estimating the role of coalescence reactions in the development of an oscillatory regime, since it only enhances their effect on the process of condensation. Thus, to numerically describe condensation in an open system, we solved the following equations:

$$\frac{\partial n_m}{\partial t} = \sum_{1 \le k \le \frac{m}{2}} \nu_{m-k,k} n_{m-k} n_k - n_m \sum_{k \ne m} \nu_{m,k} n_k - 2\nu_{mm} n_m^2 - \eta_m n_m + \eta_{m+1} n_{m+1} - \frac{n_m}{\tau}, \quad (5)$$

$$\frac{\partial n_1}{\partial t} = N_0 - n_1 \sum_{k>1} \nu_{1,k} n_k - 2\nu_{11} n_1^2 + \sum_{k>1} \eta_k n_k - \frac{n_1}{\tau}.$$

Here, $\nu_{i,k} = \nu_0 (i^{1/3} + k^{1/3})^{2/3} (1/i + 1/k)^{-1/2}$ is the rate constant of the reaction of coalescence of two clusters of sizes i and k, and N_0 is the rate of monomer generation. The values of the concentrations n_m are normalized to n_s. The numerical experiment was performed in wide ranges of the parameters N_0 and τ. In the experiment, we studied the time evolution of the cluster-size distribution or, to be exact, the dependence of the concentration of clusters on the size n_m. The computing technique applied permitted us to trace the values of m up to 10^{11}, which was sufficient for estimating the total amount of the substance in the reactor; for $m > 100$, the size was considered as a real variable with a step of about $0.05m$. The initial value of n_m does not affect the result, since after a certain time there is always established a final regime of condensation, stationary or oscillatory, depending on the magnitude of τ. Each observation was repeated under conditions in which the regime of coalescence was switched-off (by setting the coalescence constants in Eq. (5) to be zero), which corresponded to only monomeric growth; Eqs. (5) became equivalent to Eq. (1). For those values of τ for which oscillations of n_m were observed, the results of the simulation of monomeric and coolescence growth were very close. The period of oscillations of the distribution function remained virtually unaltered and corresponded to τ; the observed values of n_m varied within a few percent.

Figure 2 displays the results of the simulation of the behavior of the cluster-size distribution function for the oscillatory regime of condensation. For clarity, in this example we chose a sufficiently large value of the parameter $\beta C = 20$ (that characterizes the temperature of the system and the surface tension of the substance) and the rate of generation of monomers N_0 that yields a supersaturation n/n_s varying between 40 and 60 during an oscillation period. The distribution function exhibits wavelike perturbations that propagate from small sizes to greater sizes and damp at large sizes. The nature of the oscillations can qualitatively be explained by the fact that the accumulation of large clusters suppresses the process of nucleation; the monomer is consumed at the surface of these clusters, and the formation of new particles ceases. However, as the condensation products are progressively removed from the reactor (with a characteristic time τ), the larger clusters also are removed from the system, and the nucleation process is reactivated. At a large rate of removal, the condensation is equilibrium; as τ exceeds a certain critical value, the accumulation of large clusters in the system becomes periodic. In Fig. 2, the concentration of clusters in the range of sizes from 10^2 to 10^{10} changes by about four orders of magnitude during an oscillation period. These results show that if we were able to obtain such supersaturations in a real experiment, the oscillations of the cluster density could easily be observed in a condensing vapor, e.g., with the help of laser beam scattering.

At present, we do not know experimental works that permit a comparison with the numerical results obtained. In our opinion, the experimental observation of oscillatory regimes of condensation, which requires that a strongly supersaturated vapor be obtained ($n/n_s \simeq 50$), is possible, for example, upon the laser evaporation of a liquid in a reactor with a flowing cold gas. High supersaturations can also be reached by generating monomers via a chemical reaction. In any case, good stirring of the monomer in the volume of the reactor with a flowing gas should be provided in order to obtain spatial uniformity and a constant rate of the removal of the condensation products.

Low-frequency oscillations of cluster density were observed in experiments with plasmas containing microparticles [2, 3]. High vapor supersaturations were obtained upon plasma decomposition of silane. However, the kinetics of growth and transfer of clusters in a plasma reactor is much more complex than in the case of vapor condensation. Nevertheless, we can suppose that the observed oscillations have a similar physical origin, namely, a periodic suppression of the process of nucleation, when the accumulation of large clusters in an open system initially occurs more rapidly than the process of their removal.

CONCLUSION

We considered the process of vapor condensation in an open system with an external source of a monomer and a constant rate of removal of condensation products. A theoretical analysis and numerical experiments show the existence of oscillatory regimes of condensation or, in other words, the appearance of self-sustaining oscillations of the size distribution of growing clusters as the rate of the removal of condensation products becomes smaller than a certain critical value. The cause of the oscillations is the appearance and the further growth of clusters with a size exceeding a critical value suppressing the process of nucleation until the growing clusters are removed from the system, after which a new period of oscillations begins. Ways for the experimental observation of the phenomenon predicted in this work are suggested.

ACKNOWLEDGMENTS

This work was supported in part by the INTAS (project no. 96-0235) and by the INTAS–Russian Foundation for Basic Research (project no. IR-97-775).

REFERENCES

1. E. M. Lifshitz and L. P. Pitaevskiĭ, *Physical Kinetics* (Nauka, Moscow, 1979; Pergamon, Oxford, 1981).
2. C. K. Yeon and K. Whang, J. Vac. Sci. Technol. A **13**, 2044 (1995).
3. G. Praburam and J. Goree, Phys. Plasmas **3**, 1212 (1996).

Translated by S. Gorin

Technical Physics, Vol. 46, No. 7, 2001, pp. 914–918. Translated from Zhurnal Tekhnicheskoĭ Fiziki, Vol. 71, No. 7, 2001, pp. 128–132.

BRIEF COMMUNICATIONS

Temporal Representations of the Coefficients of Aerodynamic Forces and Moments Experienced by a Rectangular Wing under Conditions of Static Hysteresis

I. V. Kolin, T. I. Trifonov, and D. V. Shukhovtsov

Received July 31, 2000

Abstract—Results are presented of experimental investigations into the static hysteresis of the aerodynamic characteristics of a rectangular wing with an aspect ratio of $\lambda = 5$, obtained in static tests of a wing model in a wind tunnel. Variation with time of the aerodynamic forces and moments are analyzed in the range of angles of attack where a transition occurs from one branch of the hysteresis loop to the other. © *2001 MAIK "Nauka/Interperiodica".*

INTRODUCTION

Experimental investigations in wind tunnels of models of rectangular wings with large aspect ratios, as well as airplanes with straight wings of large aspect ratios, have demonstrated that alongside the areas of unique dependence of the aerodynamic forces and moments on the experimental conditions (angle of attack α, angle of slide β, and the Reynolds number Re), there exist areas where uniqueness is lacking, that is, regions of aerodynamic hysteresis [1–4]. There are a number of studies devoted to the hysteresis observed in the aerodynamic characteristics of wings with large aspect ratios [1–6]. The authors of [1] considered static hysteresis of the lift coefficient $c_y = c_y(\alpha)$ observed for a wing of the NASA-23012 profile with a large aspect ratio ($\lambda = 5$) in a range of Reynolds numbers $1 \times 10^6 < \mathrm{Re} < 4 \times 10^6$ at subsonic velocities. Data on the hysteresis of aerodynamic forces and moments in a range of $0.2 \times 10^6 < \mathrm{Re} < 0.8 \times 10^6$ for wings of large aspect ratio and various thickness ratios ($\bar{c} \geq 0.12$) were obtained in [2–6].

The present paper describes the results of experimental investigations of aerodynamic characteristics of a model of a rectangular wing with the aspect ratio $\lambda = 5$ under conditions of hysteresis. Static aerodynamic dependences of the lift $c_y(\alpha)$ and moments $m_z(\alpha)$ and $m_x(\alpha)$ are presented alongside schematic patterns of the airflow over the wing, visualized with the use of silk threads. Tests employing an aerodynamic balance and visualization were performed simultaneously. With a view to increasing the accuracy of the positions of the hysteresis region boundaries, we studied variations of $c_y(t)$, $m_z(t)$, and $m_x(t)$, as well as their spectra, $A_{c_y}(f)$, $A_{m_z}(f)$, and $A_{m_x}(f)$, obtained for constant ($\alpha = \mathrm{const}$) or slowly varying angles of attack.

EXPERIMENTAL TECHNIQUE

Experimental investigations of the wing were carried out in a low-speed subsonic wind tunnel with an open-ended test section, at $\mathrm{Re} = 0.33 \times 10^6$. Figure 1 schematically shows the wing model secured to the tail sting of a dynamic device. The device was rigidly attached to the turntable of the wind tunnel. The wing of NASA-0018 profile had the following geometrical parameters: profile thickness ratio $\bar{c} = 18\%$, wing area

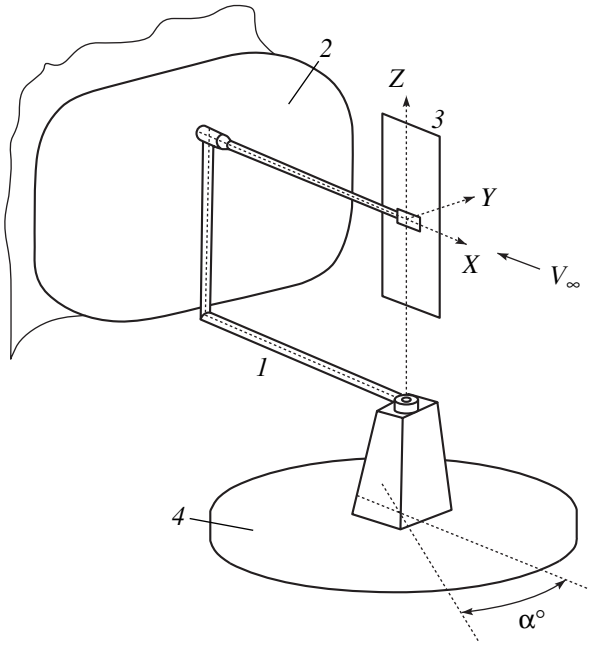


Fig. 1. Scheme of arrangement of the dynamic device with the wing model inside the test section of the wind tunnel. (*1*) OVP-102B unit, (*2*) diffuser of the wind tunnel, (*3*) wing model, and (*4*) turntable of the wind tunnel test section.

1063-7842/01/4607-0914$21.00 © 2001 MAIK "Nauka/Interperiodica"

S_1 = 0.288 m^2, mean airfoil chord b_2 = 0.24 m, and wingspan l_1 = 1.2 m. The aerodynamic forces and moments were measured with a built-in strain-gauge balance in a set of coordinates 0*XYZ* associated with the wing. The model's center of mass was at $\bar{x}_T$ = 0.5 (in units of b_2). The angle of attack α was set by the turntable and varied in a range of –3 to 27°. The angle of attack was changed continuously by turning the turntable with an angular speed of $\dot{\alpha}$ = 0.5 degrees/s. Photographs of the wing with the silk threads indicating the air flow were used to roughly estimate the location of the boundary between areas of stalling and nonstalling flows.

INVESTIGATION RESULTS

Figure 2 displays the static aerodynamic characteristics $c_y(\alpha)$, $m_z(\alpha)$, and $m_x(\alpha)$ of the rectangular wing model for an increasing ($\dot{\alpha} > 0$) and decreasing ($\dot{\alpha} < 0$) angle of attack in a range of –3° to 27° ((open diamonds). It is seen that in the range of $12.5° \le \alpha \le 19°$, a static hysteresis takes place in the curves of $c_y(\alpha)$, $m_z(\alpha)$. The $m_x(\alpha)$ curve is unique, and in the range of the angles of attack investigated, the values of coefficient m_x are small. Figure 2 also schematically shows the patterns of airflow over the wing. The blank areas correspond to the nonstalling flow (NSF) and the shaded area to the stalling flow (SF). The visualization results confirm the data of the balance measurements of the coefficients of aerodynamic forces and moments by demonstrating the difference between patterns of airflow over the wing corresponding to the upper and lower branches of the hysteresis loop. The upper branch of the hysteresis loops in the dependences $c_y(\alpha)$ and $m_z(\alpha)$ at small angles of attack corresponds to the occurrence of flow separation near the trailing edge of the wing. The separation area expands with increasing angle of attack. With transition to modes of the lower branch of the hysteresis loop, the separation line shifts toward the leading edge of the wing. If the angle of attack is reduced in conditions corresponding to the lower branch of the hysteresis loop, the flow separation area shrinks; however, the separation line is always found near the leading edge of the wing.

Figure 2 also shows static dependences $c_y(\alpha)$ and $m_z(\alpha)$ obtained in two series of repetitive tests of the model. In the first series of tests (solid circles), the angle of attack increased from an initial value of α_i = 13° to the final value of α_f = 19°. In this case, transition from the upper to the lower branch of the hysteresis loop occurred in dependences $c_y(\alpha)$ and $m_z(\alpha)$. In the second series of tests (solid diamonds), the angle of attack decreased, with the initial and final angles of attack chosen at α_i = 15° and α_f = 12°, respectively. In these experiments, it was found that the dependences $c_y(\alpha)$ and $m_z(\alpha)$ followed upper or lower branches of the hysteresis loop depending on whether the initial angle of attack α_i of the model was set before or after putting the wind tunnel into operation.

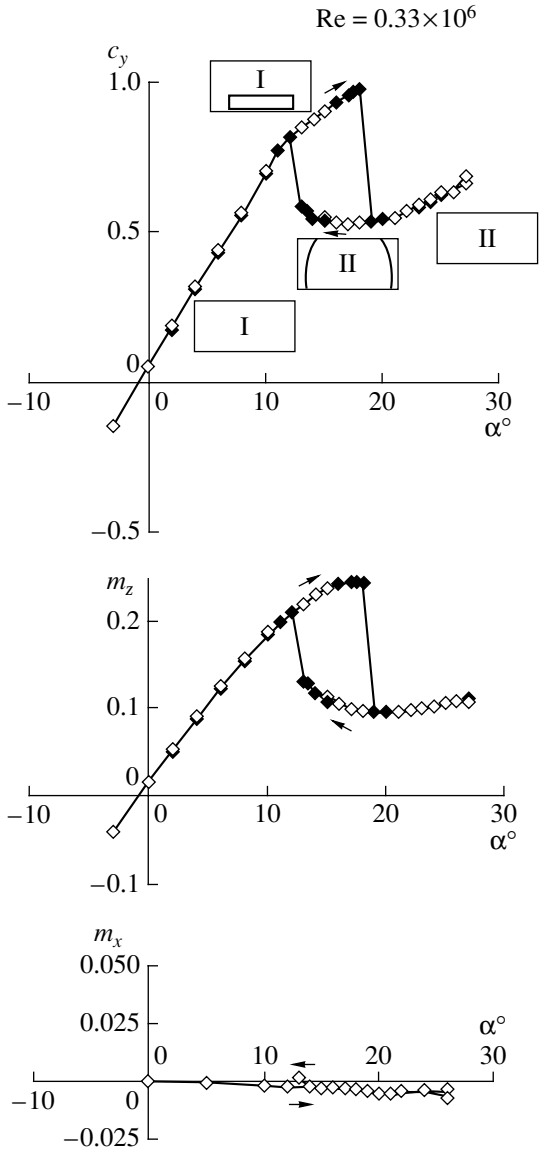


Fig. 2. Variations of the coefficients of the static aerodynamic forces c_y and moments m_z and m_x with the angle of attack. (I) NSF, (II) SF.

When the initial angle of attack α_i was set before starting the wind tunnel, the observed magnitudes of the force and moment always corresponded to the lower branch of the hysteresis loop. If the wind tunnel was first started and then the angle of attack decreased in increments down to α_f = 12°, the forces and the moments changed to magnitudes corresponding to the upper branch of the hysteresis loop. It should be noted that the data obtained in both series of repetitive tests of the model were in good agreement with the initial dependences $c_y(\alpha)$ and $m_z(\alpha)$ (open diamonds in Fig. 2).

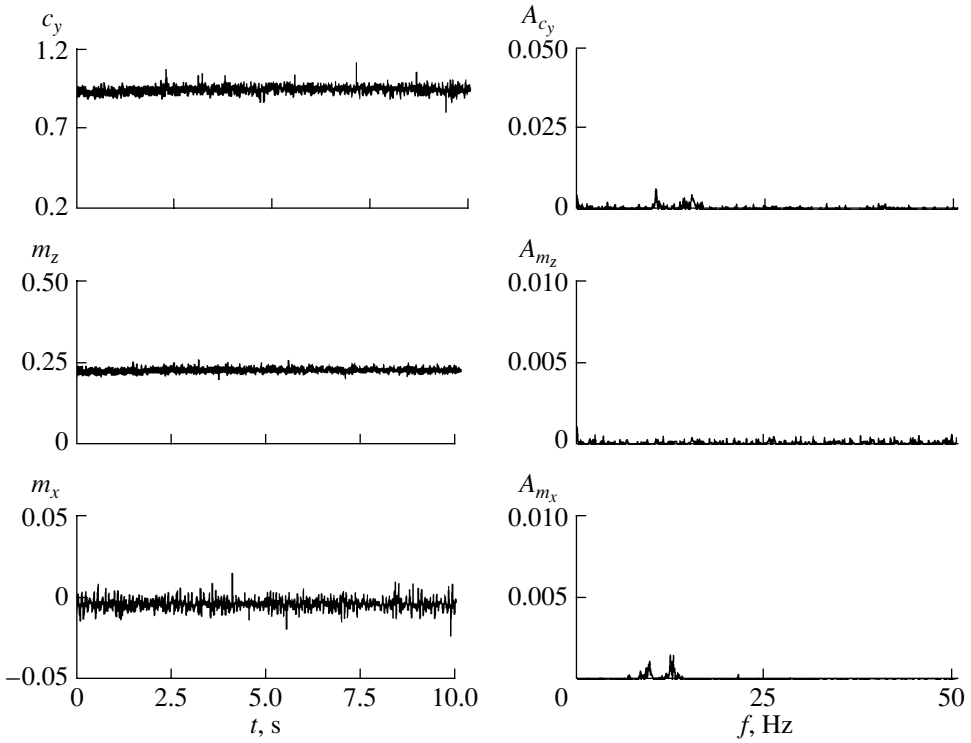


Fig. 3. Temporal representations and frequency spectra of the coefficients of the aerodynamic forces and moments corresponding to the upper branch of the static hysteresis loop. Angle of attack $\alpha = 13°$, set starting from lower values.

The above-described static aerodynamic dependences $c_y(\alpha)$, $m_z(\alpha)$, and $m_x(\alpha)$ are averaged temporal representations of $c_y(\alpha, t_k)$, $m_z(\alpha, t_k)$, and $m_x(\alpha, t_k)$, measured for each angle of attack α at the time instants $t_k = k\Delta t$, where $k = 1, 2, \ldots, n$ are the measurement points and Δt is the time increment.

As an example, Figs. 3, 4 show the characteristic temporal representations of $c_y(t)$, $m_z(t)$, and $m_x(t)$, as well as their frequency spectra, $A_{c_y}(f)$, $A_{m_z}(f)$, and $A_{m_x}(f)$, measured in modes corresponding to the upper and lower branches of the static hysteresis loop at $\alpha = 13°$. The temporal representations corresponding to the upper branch of the static hysteresis loop are regular, and the respective frequency spectra are plane (Fig. 3). Such time processes have a corresponding point-type attractor. The temporal representations of $c_y(t)$, $m_z(t)$, and $m_x(t)$ corresponding to the lower branch of the hysteresis loop are irregular (Fig. 4), and in their frequency spectra, a number of harmonics become evident. Besides, in the dependences $A_{c_y}(f)$, $A_{m_z}(f)$, and $A_{m_x}(f)$, a continuous low-frequency component is present. A mathematical representation of such processes is the concept of a strange attractor [7]. Thus, comparing the dependences $c_y(t)$, $m_z(t)$, and $m_x(t)$, as well as the frequency spectra $A_{c_y}(f)$, $A_{m_z}(f)$, and $A_{m_x}(f)$, corresponding to the upper and lower branches of the hysteresis loop at the same angles of attack, it can be noted that the differences detected in such measurements are both quantitative and qualitative.

Let us consider the results of the investigations of variations with time of the forces and moments in the course of the transition from the upper to the lower branch of the hysteresis loop. To this end, the experiments included tests with continuous slow variation of the angle of attack at a constant rate of $\dot{\alpha} = 0.5$ deg/s. Analysis of the dependences $c_y(t)$ and $m_z(t)$ obtained by continuously changing the angle of attack in a range of α from 17° to 21° has revealed a drastic change in the coefficients c_y and m_z. Figure 5 shows the representations $c_y(t)$ and $m_z(t)$ in the time interval where c_y and m_z pass from the upper to the lower branches of the hysteresis loop. As seen from the graphs, this time interval lies in the range $t \approx 1.55$–1.7 s and is equal to $\Delta t \approx 0.15$ s. Analysis of the dependences $c_y(t)$ and $m_z(t)$ has shown that in this time interval, there is not a single steady-state mode between the initial and final points of the transition. If the considered variation of the dependences $c_y(t)$ and $m_z(t)$ accompanying the transition from the upper to the lower branch of the hysteresis loop is approximated by a linear function, then, using known transition time Δt and the rate of change of the angle of attack $\dot{\alpha}$, it is possible to evaluate the range of the angles of attack $\Delta\alpha$ in which the transition occurs. From the experiment results, it follows that $\Delta\alpha = \Delta t\dot{\alpha} \approx 0.075$ degrees.

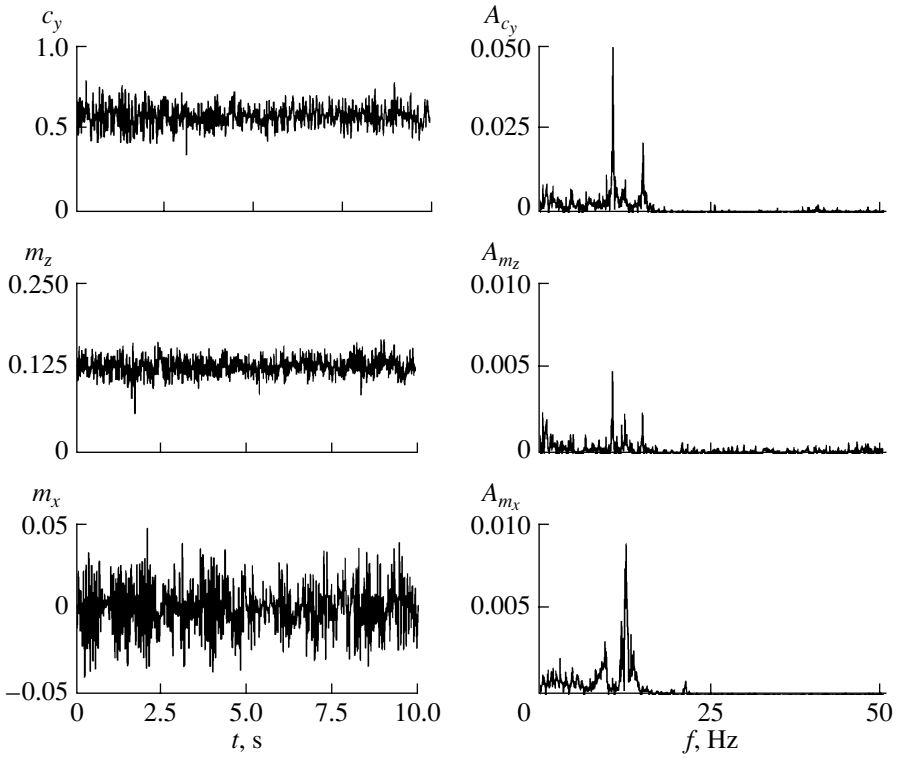


Fig. 4. Temporal representations and frequency spectra of the coefficients of the aerodynamic forces and moments corresponding to the lower branch of the static hysteresis loop. Angle of attack $\alpha = 13°$, set starting from higher values.

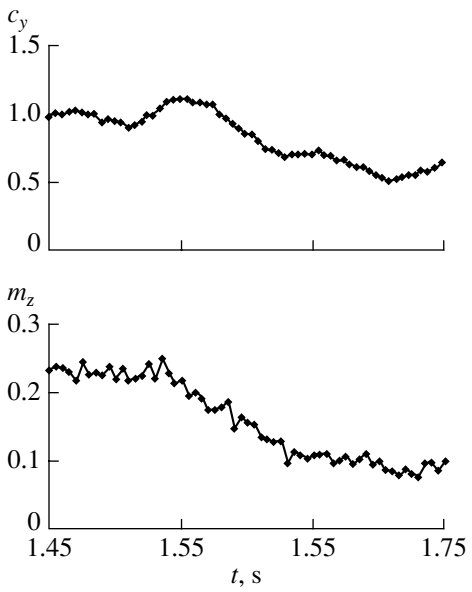


Fig. 5. Variations of $c_y(t)$ and $m_z(t)$ measured under continuous change of the angle of attack in the range $17° < \alpha < 21°$.

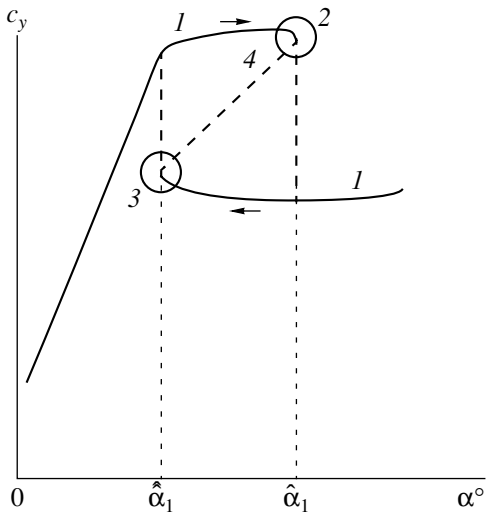


Fig. 6. Schematization of the hysteresis curves. (*1*) stable branch, (*2*) fold I, (*3*) fold II, (*4*) unstable branch.

The results of the study of the upper and lower branches of the static hysteresis, as well as the transitions from one branch to the other, can be useful in refining mathematical models roughly describing the hysteresis phenomenon in the dependences of the aerodynamic forces and moments, such as the mathematical model of hysteresis accepted in catastrophe theory. From the point of view of this theory, a hysteresis is a pair of elementary catastrophes called folds. In the fold catastrophe, the control space is one-dimensional. The bifurcation set represents a point [8–10]. The static dependence $c_y(\alpha)$ is divided into three segments; two

segments corresponding to the stable steady states and one, located within the hysteresis loop, corresponding to the unstable branch of the solutions (Fig. 6). At the points corresponding to the bifurcation angles of attack, a sudden transition occurs from one hysteresis branch to the other.

CONCLUSIONS

From the analysis of the experimental variations with time of aerodynamic forces and moments measured at constant angles of attack, as well as under continuous slow variation of the angle of attack in the airflow of a wind tunnel, the following principal conclusions can be made.

Time dependences $c_y(t)$, $m_z(t)$, and $m_x(t)$ of the aerodynamic forces and moments corresponding to the upper branch of the hysteresis loop are characterized by a point-type attractor and those at the lower branch by a strange attractor.

Transitions of the dependences $c_y(\alpha)$ and $m_z(\alpha)$ from one hysteresis branch to the other occur as a result of small changes of the preset angles of attack.

In the course of transition from the upper to the lower branch of the hysteresis loop, no intermediate steady-state modes are realized, which confirms the instability of the collapsing pattern of airflow over the wing when the bifurcation angles of attack are reached.

REFERENCES

1. A. I. Kur'yanov, G. I. Stolyarov, and R. I. Shteĭnberg, Uch. Zap. TsAGI **10** (3), 12 (1979).
2. É. A. Karavaev, Yu. A. Prudnikov, and E. A. Chasovnikov, Uch. Zap. TsAGI **17** (6), 95 (1986).
3. G. I. Stolyarov and V. G. Tabachnikov, in *Proceedings of the Central Institute of Aerohydrodynamics* (Tsentr. Aérogidrodin. Inst., Moscow, 1985), Vol. 2290, p. 71.
4. I. V. Kolin, T. I. Trifonova, and D. I. Shukhovtsov, Preprint No. 87, TsAGI (Zhukovsky Central Institute of Aerohydrodynamics, Moscow, 1996).
5. Yu. A. Kolmakov, Yu. A. Ryzhov, G. I. Stolyarov, and V. G. Tabachnikov, in *Proceedings of the Central Institute of Aerohydrodynamics* (Tsentr. Aérogidrodin. Inst., Moscow, 1985), Vol. 2290, p. 84.
6. I. V. Kolin, T. I. Trifonova, K. F. Latsoev, *et al.*, Preprint No. 85, TsAGI (Zhukovsky Central Institute of Aerohydrodynamics, Moscow, 1996).
7. P. Berge, Y. Pomeau, and C. Vidal, *Order within Chaos: Towards a Deterministic Approach to Turbulence* (Wiley, New York, 1986; Mir, Moscow, 1991).
8. V. I. Arnold, *Catastrophe Theory* (Nauka, Moscow, 1990; Springer-Verlag, Berlin, 1986).
9. J. M. T. Thompson, *Instabilities and Catastrophes in Science and Engineering* (Wiley, Chichester, England, 1982; Mir, Moscow, 1985).
10. I. Stewart, *Les Chroniques de Rose Polymath* (E. Belin, Paris, 1982; Mir, Moscow, 1987).

Translated by N. Mende

Technical Physics, Vol. 46, No. 7, 2001, pp. 919–922. Translated from Zhurnal Tekhnicheskoĭ Fiziki, Vol. 71, No. 7, 2001, pp. 133–136.

BRIEF COMMUNICATIONS

Bunching of Particles with Nonzero Longitudinal Emittance

A. S. Chikhachev

Received November 12, 1999; in final form, August 2, 2000

Abstract—The dynamics of the bunching of particles with nonzero longitudinal emittance is studied in plane geometry. A self-consistent solution to the model time-dependent kinetic equation is obtained for a collisionless system of particles. The solution is shown to be nonunique. Two acceleration modes resulting in the creation of a virtual cathode and, therefore, in the bunching of particles are basically possible. © *2001 MAIK "Nauka/Interperiodica".*

(1) The initial stage of particle acceleration in a planar diode is usually described under the assumption that the longitudinal phase volume is zero. The phase volume also remains zero in the presence of a velocity spread at the cathode [1] if the particles are emitted by a planar surface. However, under real conditions, the region from which the accelerated particles are extracted is extended in the longitudinal direction. If, in addition, the longitudinal velocity spread is nonzero, the particles have a finite phase volume; i.e., they are characterized by nonzero emittance.

In this paper, we consider the dynamics of a bunch with finite longitudinal emittance. We assume that the particles possess nonzero transverse velocities and that large transverse dimensions of the bunch are fairly large.

Let us consider a system consisting of an emitter and a control grid placed at $z = 0$ and L, respectively. The emitter is at a zero potential, and the control potential $\Phi L(t)$ is applied to the grid (Fig. 1). The potential $\Phi(z)$ satisfies the equation

$$\Delta\Phi = -4\pi en. \tag{1}$$

Here, e is the particle charge and n is the density. If the longitudinal size R of the region occupied by the particles meets the inequality $R > L$, the following relationship is valid there:

$$\Phi = -2\pi enz^2 + z\left(\frac{\Phi_L}{L} + 2\pi enL\right). \tag{2}$$

The equation of particle motion has the form

$$\ddot{z} = \omega^2(z)(z + z_0(t)), \tag{3}$$

where $\omega^2(t) = 4\pi e^2 n(t)/m$, m is the mass of an accelerated particle, and $z_0 = -(\Phi_L/4\pi enL) > 0$.

In order to provide a kinetic description of the particles characterized by equation of motion (3), one should devise an invariant of this equation, which must be valid at any function $\omega(t)$. This invariant must replace the Hamiltonian used in the steady-state problem of particle motion in a diode [1, 2]. For a homogeneous time-dependent equation of motion, we can use the Kapchinsky–Vladimirsky invariant [3]. After being transformed, this invariant also applies to an ensemble described by inhomogeneous equation (3). Let $z_1 = z - \xi$, where $\xi(t)$ is defined by the equation $\ddot{\xi} = \omega^2(t)(\xi + z_0)$. Then, $z_1(t)$ satisfies the homogeneous equation

$$\ddot{z}_1 = \omega^2(t)z_1. \tag{4}$$

The invariant for the longitudinal motion has the form

$$I = \frac{(R_1\dot{z}_1 - \dot{R}_1 z_1)^2}{\varepsilon_0^2} + \frac{z_1^2}{R_1^2}. \tag{5}$$

The condition $dI/dt \equiv 0$ implies that the function $R_1(t)$ satisfies the equation

$$\ddot{R}_1 - \omega^2(t)R_1 = \frac{\varepsilon_0^2}{R_1^2}. \tag{6}$$

Assuming that the distribution function has the form

$$f = \kappa\frac{\sigma(1-I)}{\sqrt{1-I}},\ \sigma(x) = 0,\ x < 0,\ \sigma = 1,\ x \geq 0,$$

we obtain the formulas for the density,

$$n = \frac{\pi\kappa\varepsilon_0}{R_1}\sigma(r_1 - z_1), \tag{7}$$

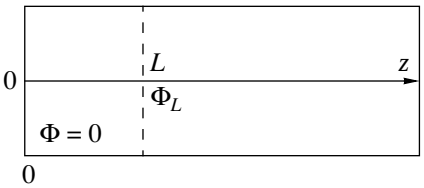


Fig. 1. Electrode geometry. The control potential $\Phi_L(t)$ is applied to the grid placed at $z = L$.

1063-7842/01/4607-0919$21.00

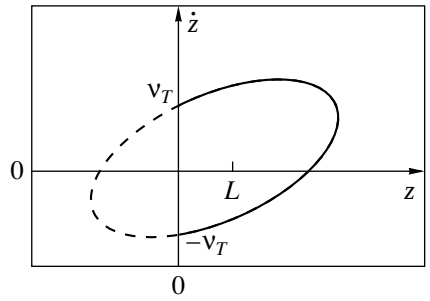


Fig. 2. Domain on the phase plane occupied by the accelerated particles: a portion of the ellipse at $z > 0$.

current density j_z,

$$j_z = enz\frac{\dot{R}_1}{R_1} + enR_1\left(\frac{\xi}{R_1}\right), \tag{8}$$

and rms longitudinal velocity spread:

$$\overline{(\Delta\dot{z})^2} = \frac{\varepsilon_0^2}{2R_1^2}\left(1 - \frac{z_1^2}{R_1^2}\right). \tag{9}$$

Formula (9) yields that, at the cathode (at $z = 0$),

$$\overline{(\Delta\dot{z})^2}\Big|_{z=0} = v_T^2 = \frac{\varepsilon_0^2}{2R_1^2}\left(1 - \frac{\xi^2}{R_1^2}\right), \tag{10}$$

where v_T is the thermal velocity spread at the cathode.

Expressions (7)–(10) suggest an appropriate choice of the model distribution function that gives a qualitatively correct description of the initial conditions for the particles. Since the thermal velocity at the cathode remains constant, expression (10) provides an additional relationship between functions $\xi(t)$ and $R_1(t)$:

$$\xi(t) = \pm R_1\sqrt{1 - R_1^2/R_*^2}, \tag{11}$$

where $R_*^2 = \varepsilon_0^2/2v_T^2$.

The particles that leave the emitter and enter the accelerating gap occupy the phase space region bounded by an ellipse (Fig. 2). Near the cathode, the particles can move in both the positive and negative directions along the z axis.

Note that the above expressions for the distribution function f and potential $\Phi(z, t)$ are the exact solution to a model Vlasov self-consistent system—the kinetic equation for the particles and the equation for the electrostatic potential.

(2) In accordance with relationship (11), two acceleration modes are possible. If we assume that $\xi = +\sqrt{1 - R_1^2/R_*^2}$, then this relationship, together with Eqs. (6) and (4), yields an expression for $z_0(t)$:

$$z_0 = -\frac{R_1^3}{R_*^2\sqrt{1 - R_1^2/R_*^2}} + \frac{\varepsilon_0^2(1 - 2R_1^2/R_*^2)}{\omega^2 R_1\sqrt{1 - R_1^2/R_*^2}} - \frac{\dot{R}_1^2 R_1(3 - 4R_1^2/R_*^2)}{\omega^2 R_*^2(1 - R_1^2/R_*^2)^{3/2}}. \tag{12}$$

By virtue of expression (7), we obtain

$$\omega^2(t) = 4\pi n_0 e^2/m = \frac{4\pi^2\kappa\varepsilon_0}{mR_1(t)} = \omega_0^2\frac{R_{10}}{R_1(t)}, \tag{13}$$

where $\omega_0^2 = 4\pi n_0 e^2/m$, n_0 is the density at the initial instant and $R_{10} = R_1(t)|_{t=0}$.

If $\dot{R}_1(t)|_{t=0} = 0$, Eq. (6) yields

$$\dot{R}_1^2 = 2\omega_0^2 R_{10}(R_1 - R_{10}) - \varepsilon_0^2\left(\frac{1}{R_1^2} - \frac{1}{R_{10}^2}\right). \tag{14}$$

Therefore, when the emittance ε_0 is nonzero, the condition $R_{10} > 0$ must be satisfied. This means that, at $t = 0$, the longitudinal size of the cloud of charged particles near the emitter is $R_0 = R_{10} + R_{10}\sqrt{1 - R_{10}^2/R_*^2}$. Below, we use the following notation: $R_1/R_* = \eta$, $R_{10}/R_* = \eta_0$, and

$$\frac{\varepsilon_0^2}{\omega_0^2 R_{10}R_*^3} = \frac{\varepsilon_0^2}{\omega_0^2\eta_0 R_*^4} = \frac{4v_T^4}{\omega_0^2\eta_0\varepsilon_0^2} = \mu.$$

Then, expression (12) can be rewritten as

$$\frac{z_0(1 - \eta^2)}{R_*} = -7\eta^3 + 9\eta^5 + 2\eta_0\eta^2(3 - 4\eta^2) + \mu\left[\frac{1}{\eta^2} - 2\eta^2 - \frac{3\eta^2}{\eta_0^2} + 4\frac{\eta^4}{\eta_0^2}\right]. \tag{15}$$

According to Eqs. (2), (3), and (5) and the definition of z_0, the function $U = -\frac{e\Phi}{m}$ can be written as

$$U = \frac{2\pi e^2 n}{m}(z + z_0)^2 + -\frac{2\pi e^2 n}{m}z^2 = \frac{\omega_0^2 R_{10}}{R_1}(z + z_0)^2 - \frac{\omega_0^2}{R_1^2}R_{10}^2 z_0^2. \tag{16}$$

If $z_0(t)|_{t=0} > 0$ and z_0 decreases at $t > 0$ and then becomes negative, the first term in formula (16) vanishes at $z = -z_0$ and the second term describes a poten-

tial barrier that blocks the particles if the following condition is valid:

$$\frac{\omega_0^2 R_{10}}{R_1} z_0^2 > \frac{v_T^2}{2}. \quad (17)$$

In this situation, the particles become separated from the cathode and a bunch is formed (see also [4]).

It should, however, be noted that the above equations can be used only when there is no potential barrier in the accelerating gap; i.e., this description is valid at $z_0 \geq 0$. Presumably, to create a barrier that blocks the particles, the function $U_L(t)$ must contain a region in which it sharply drops by $\sim v_T^2/2$ after reaching the point $z_0 = 0$. In accordance with this, we further assume that such a drop occurs and that a bunch is formed at time t when $z_0(t) = 0$. To make the process of the bunch formation possible, the control potential applied to the grid must satisfy the relationship

$$U_L = -\frac{e\Phi_L(t)}{m} = \frac{\omega_0^2 R_{10}}{R_1(t)}(z_0(t) + L/2) \quad (18)$$

at $z_0 > 0$.

For $\eta = \eta_0$ and $z_0 > 0$, it follows from Eq. (15) that, at the initial instant,

$$\frac{z_0\sqrt{1-\eta^2}}{R_*} = \eta\frac{1-2\eta_0^2}{\eta_0} - \eta_0^3, \quad (19)$$

if $\eta_0 < \frac{1}{\sqrt{2}}$ at $\mu > 1$ or $\eta_0 < \mu^{1/5}$ at $\mu \ll 1$. If the condition for η_0 to be small is satisfied, z_0 becomes zero when $\eta < \sqrt{3}/2$. In fact, at $\eta^2 = 3/4$, we obtain

$$\frac{z_0}{R_*} = -\left(\frac{3\sqrt{3}}{4} + \frac{4\mu}{3}\right) < 0.$$

(3) Under the above conditions, the bunch size is

$$R(t) = R_1(1 + \sqrt{1 - R_1^2/R_*^2}) = R_*\eta(1 + \sqrt{1-\eta^2}).$$

The function $R(\eta)$ is shown in Fig. 3 (curve *I*). At $\eta = \sqrt{3}/2$, the bunch size reaches its maximum $R_{\max} = R_*(3\sqrt{3}/4)$. Note that the value $R = R_*$ can be reached not only at $\eta = 1$, but also at $\eta < 1$, because, in addition to $\eta = 1$, the equation $\eta(1 + \sqrt{1-\eta^2}) = 1$ has a solution $\eta \simeq 0.54$. At $\eta = 0.54$, the equation $z_0 = 0$, which is the condition for the bunch to leave the cathode, has the following form:

$$-0.67 + 1.06\eta_0 + \mu\left(2.87 - \frac{0.53}{\eta_0^2}\right) = 0.$$

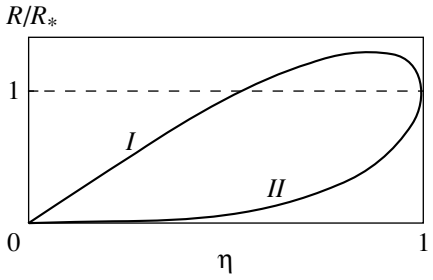


Fig. 3. The size of the charged particle bunch vs. parameter η in two different acceleration modes.

From this equation, we obtain

$$\mu = \frac{0.67 - 1.06\eta_0}{2.87 - 0.53/\eta_0^2}.$$

Therefore, $\mu > 0$ when

$$0.42 < \eta_0 < 0.63. \quad (20)$$

Thus, a bunch of size $\sim R_*$ leaves the cathode if the initial size of the region occupied by the particles is $R_0 \simeq 0.8R_*$. The size $\sim R_*$ of the bunch at the instant it leaves the cathode seems to be optimal, because, in this case, as is seen in Fig. 2, the phase volume of the bunch separates from the cathode. The parameter μ determines the density at the given emittance and velocity v_T.

Let us consider the second acceleration mode, for which $\xi = -\sqrt{1 - R_1^2/R_*^2}$. In this case, the right-hand side of Eq. (15) changes its sign [as does the right-hand side of Eq. (19)]. The particles are extracted ($z_0 > 0$) at $\mu > 1$ when $\eta_0^2 > 1/2$ and at $\mu \ll 1$ when $\eta_0 > \mu^{1/5}$. At $\eta^2 = 3/4$, we have $z_0 > 0$, whereas, as $\eta \longrightarrow 1$, we obtain

$$\frac{z_0(1-\eta^2)}{R_*} \longrightarrow -(1-\eta_0^2)\left(2 + \frac{\mu}{\eta_0^2}(1+\eta_0)\right).$$

This means that the particles leave the cathode when $\frac{\sqrt{3}}{2} < \eta < 1$. The function $R(\eta) = R_*\eta(1 - \sqrt{1-\eta^2})$ is monotonic (see Fig. 3, curve *II*). In this acceleration mode, the width of the charged particle cloud near the emitter at small η is much smaller than in the first mode.

(4) If, at the end of the pulse, η differs slightly from η_0, the pulse duration can easily be calculated. Formula (14) yields

$$\dot{\eta}^2 = 2\omega_0^2\eta\left[\eta - \eta_0 - \frac{\mu}{2}\left(\frac{1}{\eta^2} - \frac{1}{\eta_0^2}\right)\right]. \quad (21)$$

Therefore,

$$t = \frac{1}{\omega_0}\int_{\eta_0}^{\eta} d\eta \Big/ \left(\sqrt{2\eta_0\left(1+\frac{\mu}{2}\frac{\eta+\eta_0}{\eta^2\eta_0^2}\right)}\sqrt{\eta-\eta_0}\right).$$

At the initial stage, the main contribution to the integral is due to the points at which $\eta \sim \eta_0$. Then, we have

$$t \simeq \frac{1}{\omega_0}\left(\frac{2(\eta-\eta_0)}{\eta_0+\mu/\eta_0^3}\right)^{1/2}.$$

To determine the pulse shape, we should calculate $z_0(t)$ from Eq. (18).

Note that the additional term in the current at $z = 0$ has different signs in the two acceleration modes considered above. In the first mode, the second term in relationship (8) is

$$\sim -\frac{R_1\dot{R}_1 2v_T^2}{\varepsilon_0^2\sqrt{1-R_1^2/R_*^2}}.$$

This term is proportional to v_T^2; at $t = 0$, it is zero because $\dot{R}_1\big|_{t=0} = 0$. In the second mode, it has a different sign, so that this term increases the current. The additional current becomes very high as $R_1 \longrightarrow R_*$ ($\eta \longrightarrow 1$). However, this point is inaccessible in both modes.

Thus, in this paper, we have constructed a nonsteady model for accelerating a bunch with a nonzero longitudinal emittance. It shows that two acceleration modes are possible. The duration of the accelerating pulse and the effect of the longitudinal velocity spread are qualitatively determined.

REFERENCES

1. V. L. Granovskiĭ, *Electric Current in a Gas* (Nauka, Moscow, 1971).
2. V. M. Smirnov, Zh. Éksp. Teor. Fiz. **50**, 1005 (1966) [Sov. Phys. JETP **23**, 668 (1966)].
3. I. M. Kapchinskiĭ, *Theory of Linear Resonant Accelerators: Particle Dynamics* (Énergoatomizdat, Moscow, 1988).
4. A. S. Chikhachev, in *Proceedings of the XVI Conference on Accelerators, Protvino, 1998.*

Translated by A. Khzmalyan

Technical Physics, Vol. 46, No. 7, 2001, pp. 923–925. Translated from Zhurnal Tekhnicheskoĭ Fiziki, Vol. 71, No. 7, 2001, pp. 137–139.

BRIEF COMMUNICATIONS

Effect of Gas Flows on Fullerene Formation Process

D. V. Afanas'ev, G. A. Dyuzhev, and A. A. Kruglikov

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

Received November 20, 2000.

Abstract—The effect of helium flow on the fullerene formation process is studied. It has been shown that an annular flow of helium directed into the fullerene formation zone along the discharge axis reduces the content of fullerenes in soot. If the velocity of the additional flow of helium pumped through a hole in the cathode becomes close to that of the gas-plasma jet, practically no fullerenes are formed. © *2001 MAIK "Nauka/Interperiodica".*

Further experimental studies are needed for achieving insight into the process of fullerene formation in an arc discharge. Useful data can be obtained from studies on the effect that injecting various gases has on the efficiency of fullerene formation.

(1) An arc was burning between two vertical graphite electrodes 6 mm in diameter in a cylindrical vacuum chamber 180 mm in diameter and 300 mm in length. After evacuation, the chamber was filled with helium at a certain pressure and a flow of helium was produced using a closed helium contour incorporating a standard GR-A5-5 air blower.

The experimental installation is schematized in Fig. 1. Helium from blower *5* was fed into the hollow toroidal chamber of gas-distribution manifold *4*. The end cover of the chamber facing the arc had a annular slot of a width of 1 mm. Values of the annular slot diameter, 2*R*, used in the experiments were 4, 9.2, and 14 cm, and the distances *h* from the manifold cover to middle of the interelectrode gap were 1.5, 2, and 2.5 cm. Secured to the lower flange of the vacuum chamber was filter *6*, with the filtering element of one of two layers of common sheeting fabric. Having passed the filter, helium returned into the blower. In all the experiments, the flow rate of helium was constant at 5×10^4 cm^3/s.

In a different series of experiments, helium was introduced directly into the discharge area along its axis. For this purpose, an opening was provided in cathode *2*, 2.6 mm in diameter, through which helium flowed into the arc. In these experiments, the flow of helium was controlled with choke valve *7* and the flow rate was measured with flow meter *8*.

(2) The idea of an experiment with annular flow is as follows. It is known that the transformation of carbon vapor into fullerenes occurs in a fanlike jet exiting the interelectrode gap [1, 2]. It is of interest to know the reason why the fullerene formation process is affected by the fact that, at a certain distance from the discharge axis, the carbon transformation products are entrained by a jet of cold helium.

The results are shown in Fig. 2, which is the dependence of the fullerene content α in the soot deposited on the chamber walls (curve *1*) and on the filter (curve *2*) on the radius R of the annular slot (α_0 being the fullerene content in the soot when no helium is pumped through). The value of α was determined by the standard method of absorption of toluene solution [3].

It is seen that at R = 2 and 4.6 cm, the ratio α/α_0 is well below unity. At first sight, these results appear to disagree quantitatively with the experimental data of [4] and calculations in [2], where it has been shown that the area of fullerene formation is in the region $1 < R <$ 3.5 cm. However, if a jet divergence angle of 90° is assumed, then at R = 2, 4.6, and 7 cm, the jet "boundary" will be located away from the discharge axis at distances of 0.5, 2.7, and 4.5 cm, respectively. These values correlated with the results of [2, 4] quite well.

Thus, if carbon clusters, being "midway" towards transforming into fullerenes, enter a jet of cold gas, the efficiency of fullerene formation will be drastically

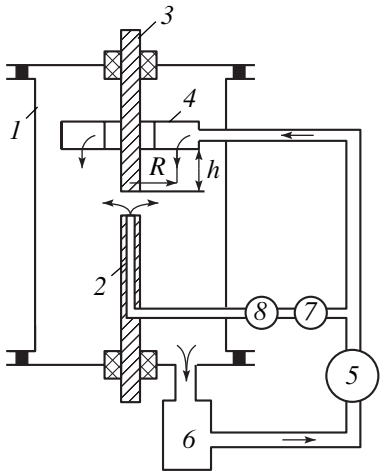


Fig. 1. Experimental setup: (*1*) vacuum chamber; (*2*) cathode; (*3*) anode; (*4*) gas-distribution manifold; (*5*) blower; (*6*) filter; (*7*) choke valve; (*8*) flow meter.

1063-7842/01/4607-0923$21.00 © 2001 MAIK "Nauka/Interperiodica"

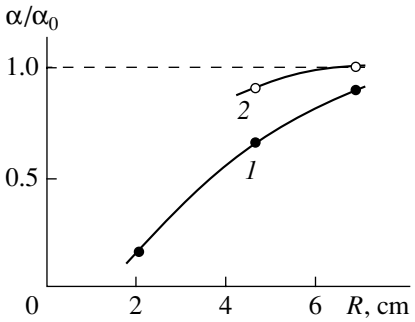


Fig. 2. Dependence on the radius of annular gas flow of the fullerene content in the soot depositing on the chamber walls (*1*) and the filter (*2*). P_{He} = 100 torr, I = 70 A, d = 3 mm.

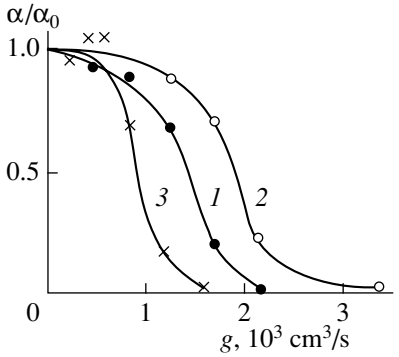


Fig. 3. Dependence on the flow rate of helium pumped through the cathode of the fullerene content in the soot depositing on the chamber walls. P_{He} = 100 (*1*, *2*), 600 torr (*3*); I = 100 (*1*, *3*), 200 A (*2*); d = 4 mm.

reduced. Note that if intermediate transformation products stay in the cold jet long enough, then the efficiency of fullerene formation rises (in Fig. 2, curve *2* is higher than curve *1*).

The experiments have shown that an annular gas jet effectively entrains the carbon soot. The quantity of soot deposited on the chamber walls is much less than on the filter.

(3) Experiments with pumping helium through a cathode are based on the following concept. The velocity of an arc gas-plasma jet is determined by thermal and magnetic processes and depends on the arc current I, helium pressure P, radius of the annular slot r_0, and the gas temperature T [5]:

$$v_0 = 1.5 \times 10^3 \frac{I[\mathrm{A}]}{r_0[\mathrm{cm}]}\left(\frac{T[\mathrm{eV}]}{P[\mathrm{torr}]}\right)^{1/2} [\mathrm{cm/s}]. \qquad (1)$$

With the given operating parameters of the arc, this velocity can be increased by pumping a gas along the arc.

Figure 3 shows the dependence of the ratio α/α_0 on the rate of helium flow through the cathode g. It is seen that in different arc modes, the ratio α/α_0 varies but weakly, as long as the flow rate is low, until at a certain value of the flow rate the ratio α/α_0 drops rather drastically, and at high flow rates there are practically no fullerenes in the soot. The helium pressure in the chamber and the arc current were maintained constant at different flow rates, whereas the arc voltage increased with the rising flow rate.

In the table, some characteristics of the arc modes used are given: helium pressure P_{He}, arc current I, arc voltage with zero helium flow rate U_0, the flow rate value for $\alpha/\alpha_0 = 0.5$, and the corresponding increase of the arc voltage ΔU. If it is assumed that the entire additional power supplied to the arc is spent in heating the cold helium introduced through the cathode,

$$I\Delta U = \rho v C_P \Delta T, \qquad (2)$$

then the increase in the gas temperature ΔT and its velocity v can be calculated. In the last column of the table, v_0 values calculated by formula (1) are given.

It is seen from the table that v and v_0 have well correlated functional dependences on I and P_{He}; however, their numerical values differ considerably. Probably, this difference is caused by the fact that formula (1) has been derived for an axial plasma jet. It is not known by how much the transformation of the axial cathode and anode jets into a radial fanlike jet will reduce the jet velocity. Therefore, we consider the quantity v to be more trustworthy than v_0.

The physical explanation of the dependence of the quantity α/α_0 on g presented in Fig. 3 is as follows. As soon as the velocity of the gas pumped through the cathode becomes close to that of the gas-plasma jet itself, the time spent by carbon clusters (precursors of fullerenes) in the zone optimal for the formation of fullerenes becomes shorter and fullerene formation is disrupted. If the width of the formation zone is ~ 2.5 cm and the jet velocity ~ 1×10^4 cm/s, then the time spent in the optimal zone is ~ 2.5×10^{-4} s, which is close to the characteristic time of transformation into fullerenes

Table

P_{He}, Torr	I, A	U_0, V	α_0, %	g, 10^3 cm^3/s	ΔU, V	ΔT, K	v, 10^4 cm/s	v_0, 10^4 cm/s
100	100	27	7.5	1.4	3	1780	1.1	2.7
100	200	38	6.8	1.9	2	1750	1.5	5.4
600	100	34	11.1	0.9	9	1380	0.55	1.1

of two-ring clusters calculated in [6] for the number of atoms in an initial cluster $N > 44$.

ACKNOWLEDGMENTS

This study was carried out in the framework of the program "Fullerenes and Atomic Clusters" (complex project no. 3) and supported in part by the Russian Foundation for Basic Research (project no. 00-02-16928).

REFERENCES

1. O. A. Nerushev and G. I. Sukhinin, Zh. Tekh. Fiz. **67** (2), 41 (1997) [Tech. Phys. **42**, 160 (1997)].
2. N. I. Alekseev, F. Chibante, and G. A. Dyuzhev, Zh. Tekh. Fiz. **71** (6), 122 (2001) [Tech. Phys. **46**, 761 (2001)].
3. D. V. Afanas'ev, I. O. Blinov, A. A. Bogdanov, *et al.*, Zh. Tekh. Fiz. **64** (10), 76 (1994) [Tech. Phys. **39**, 1017 (1994)].
4. G. A. Dyuzhev and V. I. Karataev, Fiz. Tverd. Tela (St. Petersburg) **36**, 2795 (1994) [Phys. Solid State **36**, 1528 (1994)].
5. S. Ramakrishnan, A. Stokes, and J. Lowke, J. Phys. D **11**, 2267 (1978).
6. N. I. Alekseev and G. A. Dyuzhev, Zh. Tekh. Fiz. **71** (5), 67 (2001) [Tech. Phys. **46**, 573 (2001)].

Translated by B. Kalinin

Technical Physics, Vol. 46, No. 7, 2001, pp. 926–928. Translated from Zhurnal Tekhnicheskoĭ Fiziki, Vol. 71, No. 7, 2001, pp. 140–142.

BRIEF COMMUNICATIONS

On Luminescence Enhancement in the Narrow-Gap Phase of PbS–CdS Heterogeneous Semiconductor

A. G. Rokakh and N. B. Trofimova

Chernyshevsky State University, ul. Astrakhanskaya 83, Saratov, 410026 Russia

e-mail: semiconductor@sgu.ssu.runnet.ru

Received December 18, 2000

Abstract—A great enhancement of photoluminescence from CdS-doped PbS [1] can be attributed to the recombination flux from the wide- into narrow-gap phase due to the field of the variband transition region. A semiconductor heterostructure model accounting for the enhancement of luminescence in the PbS–CdS system is presented. The concentration profile of nonequilibrium charge carriers and the integral intensity of luminescence at the interface between the narrow- and wide-gap phase are numerically studied. © *2001 MAIK "Nauka/Interperiodica".*

As is known, PbS–CdS semiconductor compounds feature a limited homogeneity range and form ternary solid solutions $Pb_{1-x}Cd_xS$ at small x [1–4]. The material is inhomogeneous in composition: CdS precipitates and PbS lines are observed in X-ray diffraction patterns [1, 5]. Earlier, it was demonstrated [1, 6, 7] that $Pb_{1-x}Cd_xS$ films contain a heterogeneous CdS–PbS material consisting of the wide-gap (CdS) matrix and embedded islands of the narrow-gap solid solution $Pb_{1-x}Cd_xS$ with x varying from 0.02–0.2 to 0.24–0.3.

The lasing effect in $Pb_{1-x}Cd_xS$ crystals at $T = 10$ – 30 K [5], as well as spontaneous and stimulated emission from the $Pb_{1-x}Cd_xS$ films at $T = 300$ K [1], indicates a high luminescence intensity in the spectral range corresponding to the narrow-gap phase (2.3–3 μm).

Our study of luminescence from the $Pb_{1-x}Cd_xS$ films shows that the position of the luminescence peak is nearly unaffected by the charge composition for $0 < x < 0.7$ and corresponds to the solid solution with $x = 0.06$. In contrast, the intensity of luminescence I_l is sensitive to the initial charge composition; specifically, as the CdS fraction grows, I_l sharply grows and peaks at $0.4 < x < 0.6$ (see table).

Auger spectroscopy, plasma resonance, and X-ray diffraction data support the presence of the wide-gap CdS phase and a number of solid solutions with x = 0.02–0.3, 0.06–0.8, and 0.24–0.3 in this material. Therefore, it would be reasonable to elaborate a model of luminescence enhancement in heterostructures with narrow-gap solid solution inclusions.

In this paper, we numerically study the spatial evolution of the one-dimensional nonequilibrium charge carrier distribution near the narrow-gap phase.

For the linearly varying bandgap and the z-dependent field of the variband transition region (Figs. 1a, 1b), we seek a solution of the diffusion–drift equation in the form of the exponentially varying change in the nonequilibrium carrier concentration [8]:

$$\Delta n(z) = C\exp(-z/l) + B\exp(az), \qquad (1)$$

where

$$a = \frac{eE}{2kT}, \quad E = \frac{E_{g0} - E_g(z)}{ez}.$$

Here, l is the effective diffusion–drift length, E is the electric field strength, and E_{g0} and $E_g(z)$ are the band-

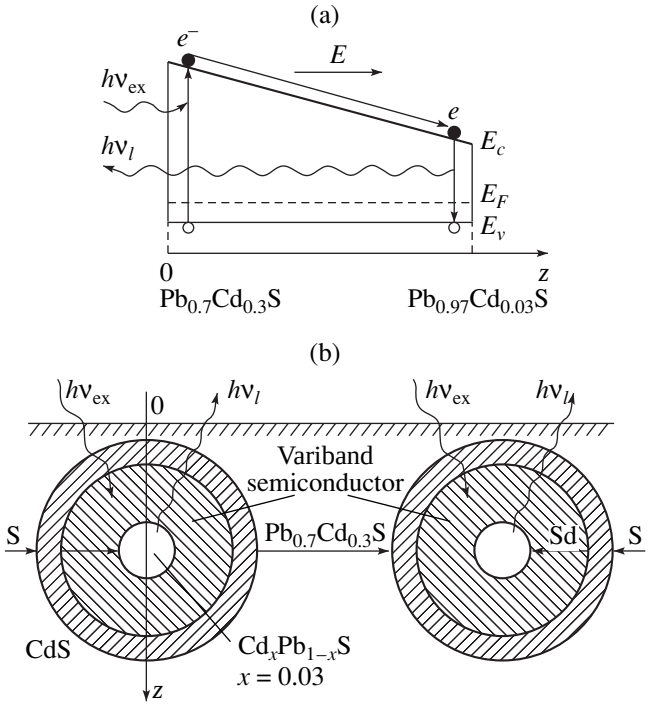


Fig. 1. (a) Energy diagram and (b) geometry of the variband structure.

1063-7842/01/4607-0926$21.00 © 2001 MAIK "Nauka/Interperiodica"

gaps of the wide- and the narrow-gap solutions, respectively.

At $z \longrightarrow \infty$ $\Delta n(z) \longrightarrow 0$ and the second term can be omitted. For a variband transition region of finite thickness d, when $d \lesssim l$, it is appropriate to introduce, along with the recombination rate S at the outer surface of the variband transition region, the recombination rate at its inner surface, S_d. Then, with the boundary conditions

$$\left(\frac{d}{dz}\Delta n - 2a\Delta n\right)_{z=0} = \frac{S}{D_n}\Delta n(0),$$
$$\left(\frac{d}{dz}\Delta n - 2a\Delta n\right)_{z=d} = \frac{S_d}{D_n}\Delta n(d), \qquad (2)$$

and $\alpha d \gg 1$ and $\alpha l_+ \gg 1$, which is true for strongly absorbable excitation of energy $h\nu \gg E_{g0}$ and intensity I_0, the solution takes the form

$$\Delta n(z) = c_1 \exp(-z/l_+) + c_2 \exp(z/l_-), \qquad (3)$$

where

$$l_\pm = [(a^2 + L_n^{-2})^{1/2} \pm a]^{-1}, \qquad (4)$$

$$c_1 = K(l_+^{-1} + S_d/D_n)\exp(d/l_-), \qquad (5)$$

$$c_2 = K(l_+^{-1} + 2a - S_d/D_n)\exp(-d/l_+), \qquad (6)$$

$$K = \frac{I_0 \alpha D_n^{-1}}{(\alpha^2 + 2a\alpha - L_n^{-2})\left[\left(l_+^{-1} + 2a + \frac{S}{D_n}\right)\left(l_+^{-1} + \frac{S_d}{D_n}\right)\exp d/l_- - \left(l_+^{-1} - \frac{S}{D_n}\right)\left(l_+^{-1} + 2a - \frac{S_d}{D_n}\right)\exp d/l_+\right]}, \qquad (7)$$

L_n is the diffusion length, D is the diffusion coefficient, and α is the absorption factor.

From (3)–(6) follows $|c_1| > |c_2|$ and $c_1 > 0$. If $(l_+^{-1} + 2a) > S_d/D_n$, then $c_2 > 0$. In this case, the nonequilibrium carriers are accumulated at the inner surface of the variband transition region near the narrow-gap phase. It is seen from Fig. 2 that the built-in electric field strength is limited from below, $E \geq 100$ V/cm, when S_d/D_n lies in the interval 10^3 cm^{-1}–5×10^5 cm^{-1}.

The top-most curve Δn (z) corresponds to $S_d/D_n \sim 10^3$ cm^{-1} at the built-in field strength $E = 800$ V/cm.

The luminescence intensity in the case of linear interband recombination can be expressed as

$$I_l = \int_0^d \frac{\Delta n(z)}{\tau_n} dz. \qquad (8)$$

Substituting (3) into (8), one obtains, in view of (4)–(7),

$$I_l = \frac{1}{\tau_n}\{c_2 l_-[\exp(d/l_-) - 1] + c_1 l_+ \exp(d/l_+)\},$$

where τ_n is the carrier lifetime. Figure 3 depicts d dependences of the quantity $I_l x \tau_n$ for $I_0 = 1.3 \times 10^{23}$ photon/(cm^2 s), $D_n = 0.3$ cm^2/s, $S_d/D_n = 2.4 \times 10^4$ cm^{-1}, and $\alpha = 10^5$ cm^{-1}. It is seen that, as the variband region becomes thicker, I_l increases and tends to saturation. Therefore, for the maximum luminescence intensity, the films of choice are those with a variband region thickness d of about 4–5 μm. An I_l vs. E curve for the built-in field E ranging from 10^2 to 10^3 V/cm was calculated in a similar way (Fig. 4). This curve peaks at $E = (2–3) \times 10^2$ V/cm. To this value, there correspond

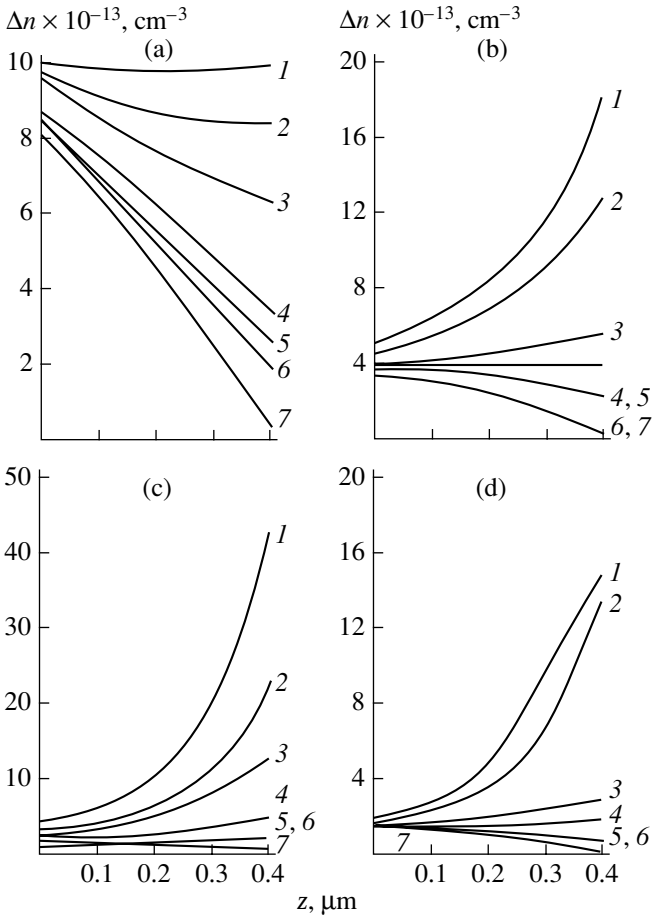


Fig. 2. Nonequilibrium carrier concentration across the variband transition region for S_d/D_n = (*1*) 5×10^3, (*2*) 10^4, (*3*) 2×10^4, (*4*) 5×10^4, (*5*) 8×10^4, (*6*) 10^5, and (*7*) 5×10^5 cm^{-1}. E = (a) 100, (b) 500, (c) 800, and (d) 1000 V/cm.

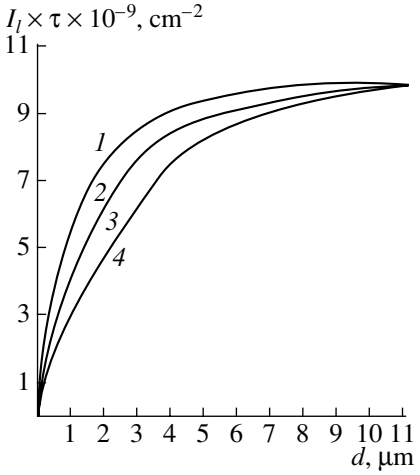


Fig. 3. Integral photoluminescence intensity vs. surface recombination rate near the narrow-gap phase in $Pb_{1-x}Cd_xS$ layers at d = 0.4 μm, L_n = 0.73 μm, and α = 10^5 cm^{-1}. E = (*1*) 100, (*2*) 500, (*3*) 800, and (*4*) 1000 V/cm.

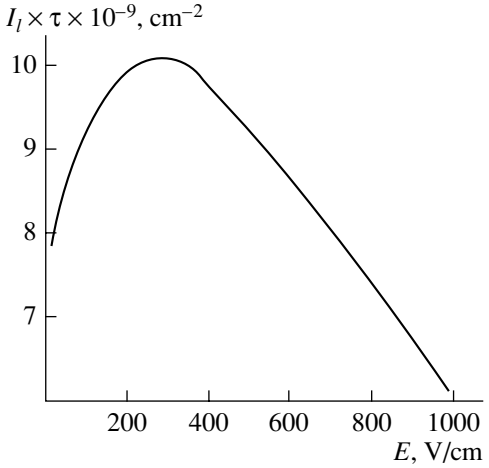


Fig. 4. Integral photoluminescence intensity vs. built-in field strength for $S/D_n = 2.4 \times 10^4$ cm^{-1}, α = 10^5 cm^{-1}, d = 0.4 μm, and L_n = 0.73 μm.

the wide-gap and narrow-gap compositions x = 0.08 and 0.03, respectively, which correlates with the X-ray diffraction data. Thus, intense luminescence from $Pb_{1-x}Cd_xS$ films is provided if the variband transition region is sufficiently thick (~4 μm); S, S_d, and α are small; L_n is large; and the built-in field of the transition region is optimal, i.e., (2–3) × 10^2 V/cm.

Another conclusion is the following. Calculating the intensity I_l as a function of S_d/D_n for E from 100 to 1000 V/cm, one can see that S_d markedly affects I_l only at $d \sim L_n$. At $d < L_n$ or $d > L_n$ $I_l \neq I_l(S_d)$.

Since the nonequilibrium carrier lifetime in the wide-gap (II–VI) phase is several orders of magnitude greater than that in the narrow-gap phase, the luminescence enhancement may also reach several orders of magnitude.

Note that the enhancement of long-wave luminescence from Pb-doped ZnS was observed in [9]; this fact, however, was not treated numerically in that work.

Table

Charge composition, x	Intensity I_l, arb. units
0	6
0.1	45
0.2	47
0.3	96
0.4	132
0.5	117
0.6	156
0.7	35

REFERENCES

1. S. I. Zolotov, N. B. Trofimova, and A. É. Yunovich, Fiz. Tekh. Poluprovodn. (Leningrad) **18**, 631 (1984) [Sov. Phys. Semicond. **18**, 393 (1984)].
2. Kepi Wu and J. N. Zemel, J. Vac. Sci. Technol. **6**, 497 (1969).
3. G. S. Oleĭnik and G. A. Mizetskiĭ, Izv. Akad. Nauk SSSR, Neorg. Mater. **19**, 1799 (1983).
4. O. D. Mukhamed'yarov and G. A. Kitaev, Pis'ma Zh. Tekh. Fiz. **6**, 1330 (1980) [Sov. Tech. Phys. Lett. **6**, 571 (1980)].
5. A. R. Calawa, J. A. Mroczkovscky, and T. S. Harman, J. Electron. Mater. **1**, 191 (1972).
6. A. G. Rokakh, Pis'ma Zh. Tekh. Fiz. **10**, 820 (1984) [Sov. Tech. Phys. Lett. **10**, 344 (1984)].
7. A. G. Rokakh, S. V. Stetsyura, N. B. Trofimova, and N. V. Elagina, Neorg. Mater. **35**, 552 (1999).
8. V. R. Kovalenko, G. P. Peka, and L. G. Shepel', Fiz. Tekh. Poluprovodn. (Leningrad) **12**, 1421 (1978) [Sov. Phys. Semicond. **12**, 839 (1978)].
9. I. J. Mita, J. Phys. Soc. Jpn. **20**, 1822 (1965).

Translated by A. Sidorova-Biryukova