Technical Physics, Vol. 45, No. 2, 2000, pp. 141–146. Translated from Zhurnal Tekhnicheskoĭ Fiziki, Vol. 70, No. 2, 2000, pp. 1–5.

Calculation of the Electrostatic Capacitance of a System of Wires of Circular, Elliptic, and Rectangular Cross Section in the Presence of Conducting Plane

Z. M. Narkun

Grodno State University, Grodno, 230023 Belarus

Received March 27, 1998; in final form, January 30, 1999

Abstract—A method is proposed for calculation of the capacitances of a system of infinitely long wires of circular, elliptic, and rectangular cross section in the presence of conducting plane. The method is based on exact construction of the electrostatic field potential. An algorithm for obtaining approximate calculation formulas is provided. Several particular cases are considered. © *2000 MAIK "Nauka/Interperiodica".*

Handbook [1] contains approximate formulas for calculation of the capacitance per unit length of single-, two-, and three-wire transmission lines comprising infinitely long wires of circular cross section in the presence of conducting plane. In this work, this problem is solved for any finite number of wires of circular, elliptic, and rectangular cross section.

The electrostatic system under consideration is plane-parallel; therefore, in what follows, a plane perpendicular to the wire axes is assumed to be the *xOy* coordinate plane. In this plane, circles, ellipses, line segments, and the straight line, along which the *xOy*-plane intersects the wire system, are considered instead of wires and the conducting plane.

All quantities with linear dimensions are replaced by dimensionless ones, that is, related to some chosen linear unit. Let the half-plane $x < 0$ contain N_1 circles Γ_j of radius R_j centered at points O_j ($j = \overline{1, N_1}$) and N_2 ($N_1 + N_2 = N$) ellipses Γ_j centered at points O_j ($J = \overline{N_1 + 1, N}$); a part of (or all) ellipses can degenerate into line segments. Circles and ellipses are situated externally relative to each other and have common points neither with each other nor with the ordinate axis $x = 0$.

Mathematically, the problem of determination of the potential φ of an electrostatic field consists in finding a solution to the Laplace equation

$$\Delta\varphi = 0 \tag{1}$$

in the half-plane $x < 0$ and outside the circles Γ_j ($j = \overline{1, N}$) (domain G). The solution must be finite at infinity and satisfy boundary conditions

$$\varphi|_{x=0} = 0, \tag{2}$$

$$\varphi|_{\Gamma_j} = f_j, \quad j = \overline{1, N}, \tag{3}$$

where f_j equal zero or unity depending on the capacitance calculated.

In order to solve the problem, let us designate mirror images of curves Γ_j relative to straight line $x = 0$ by $\bar{\Gamma}_j$ and images of centers O_j by $\bar{O}_j$. Relate each circle Γ_j and $\bar{\Gamma}_j$ ($j = \overline{1, N_1}$) to local Cartesian coordinates $x_jO_jy_j$ and $\bar{x}_j\bar{O}_j\bar{y}_j$ and polar coordinates (ρ_j, Θ_j) and $(\bar{\rho}_j, \bar{\Theta}_j)$ and relate each ellipse Γ_j and $\bar{\Gamma}_j$ ($j = \overline{N_1 + 1, N}$) to local Cartesian coordinates $x_jO_jy_j$ and $\bar{x}_j\bar{O}_j\bar{y}_j$ and elliptic coordinates (μ_j, ϑ_j) and $(\bar{\mu}_j, \bar{\vartheta}_j)$ [2], so that equation of the circle Γ_j in the related polar coordinates will be written as

$$\rho_j = R_j, \quad j = \overline{1, N_1},$$

and equation of the ellipse Γ_j in the related elliptic coordinates, as

$$\mu_j = \mu_j^0, \quad \mu_j^0 = \text{const}, \quad j = \overline{N_1 + 1, N}.$$

In particular, if some $\mu_k^0 = 0$, the ellipse transforms into a line segment of length h_k. In what follows, we will take into account that polar axes are parallel to the *Ox*-axis, axes $\bar{O}_j\bar{y}_j$ are mirror images of axes O_jy_j, and axes $\bar{O}_j\bar{x}_j$ are chosen so that coordinate systems $\bar{x}_j\bar{O}_j\bar{y}_j$ are right-handed.

We will seek the field potential φ in the form

$$\varphi(M) = A + \sum_{j=1}^{N}\left(A_j \ln\frac{1}{\rho_j} + \bar{A}_j \ln\frac{1}{\bar{\rho}_j}\right)$$

1063-7842/00/4502-0141$20.00

$$+\sum_{j=1}^{N_1}\sum_{k=1}^{\infty}[(a_k^j\cos k\Theta_j + b_k^j\sin k\Theta_j)\rho_j^{-k}$$

$$+(\bar{a}_k^j\cos k\bar{\Theta}_j + \bar{b}_k^j\sin k\bar{\Theta}_j)\bar{\rho}_j^{-k}] \quad (4)$$

$$+\sum_{j=N_1+1}^{N}\sum_{k=1}^{\infty}[(a_k^j\cos k\vartheta_j + b_k^j\sin k\vartheta_j)\exp(-k\mu_j)$$

$$+(\bar{a}_k^j\cos k\bar{\vartheta}_j + \bar{b}_k^j\sin k\bar{\vartheta}_j)\exp(-k\bar{\mu}_j)],$$

where $M \in G$ is an arbitrary point; (ρ_j, Θ_j) and $(\bar{\rho}_j, \bar{\Theta}_j)$ are coordinates of this point in the polar systems with origins O_j and $\bar{O}_j$ $(j = \overline{1, N_1})$, respectively; (μ_j, ϑ_j) and $(\bar{\mu}_j, \bar{\vartheta}_j)$ are coordinates in the elliptic systems with origins O_j and $\bar{O}_j$ $(j = \overline{N_1+1, N})$, respectively; and coefficients are to be determined from the boundary conditions and the condition at infinity.

Note that, should in formula (4) (and in the following equations) the upper index in the sum over j be smaller than the lower index [which is so provided that neither circles ($N_1 = 0$) nor ellipses ($N_2 = 0$ and $N = N_1$) are present], the corresponding sum is assumed to be zero. All indexed quantities with the initial index less than the final index are also assumed to be zero, for example, $A_s = 0$ when $s = \overline{1, 0}$.

In order to ensure finiteness of function (4) at infinity and satisfy condition (2), it is sufficient to set

$$A = 0, \quad \bar{A}_j = -A_j, \quad \bar{a}_k^j = -(1)^{k+1}a_k^j,$$

$$\bar{b}_k^j = (-1)^k b_k^j, \quad j = \overline{1, N}, \quad k = 1, 2, \ldots.$$

In order to determine the remaining coefficients A_j, a_k^j, and b_k^j, let us express all variables in formula (4) in terms of the local coordinates related to Γ_s $(s = \overline{1, N})$ using the summation theorems for separated solutions to the Laplace equation written in polar [3] and elliptic [4] systems and take into account that axes of Cartesian coordinates are not necessarily parallel to each other and new designations are introduced for both polar and elliptic coordinates [2]. Apply conditions (3) and equate the coefficients of functions cos and sin of the corresponding arguments in the obtained equations. As a result, we obtain an infinite linear system of equations.

In order to illustrate this scheme, we perform detailed calculations for the case of one ellipse. For the arbitrarily located Cartesian coordinates with different origins, formulas (14) and (15) from [4] can be written in the following form:

$$\exp(-k\mu_1)\cos k\vartheta_1$$

$$= \sum_{\nu=0}^{\infty}\varepsilon_\nu[Ac_\nu^{(k)}(\mu_{21}, \vartheta_{21}; h_2/h_1, \alpha_{21})\cosh\nu\mu_2\cos\nu\vartheta_2$$

$$+As_\nu^{(k)}(\mu_{21}, \vartheta_{21}; h_2/h_1, \alpha_{21})\sinh\nu\mu_2\sin\nu\vartheta_2],$$

$$\mu_2 < \mu_{21},$$

$$\exp(-k\mu_1)\sin k\vartheta_1 \quad (5)$$

$$= \sum_{\nu=0}^{\infty}\varepsilon_\nu[As_\nu^{(k)}(\mu_{21}, \vartheta_{21}; h_2/h_1, \alpha_{21})\cosh\nu\mu_2\cos\nu\vartheta_2$$

$$-Ac_\nu^{(k)}(\mu_{21}, \vartheta_{21}; h_2/h_1, \alpha_{21})\sinh\nu\mu_2\sin\nu\vartheta_2],$$

$$\mu_2 < \mu_{21},$$

where

$$Ac_\nu^{(k)}(\mu, \vartheta; h, \alpha)$$

$$= (-1)^k\sum_{n=0}^{\infty}\sum_{r=0}^{\infty}\frac{k(n+2r+k-1)!(n+\nu+2r+k-1)!}{h^{2r+k}n!(n+\nu)!r!(k+r)!(2r+k-1)!}$$

$$\times\exp[-(2n+\nu+2r+k)\mu]$$

$$\times\cos[(2n+\nu+2r+k)\vartheta + (2r+k)\alpha].$$

The coefficients $As_\nu^{(k)}(\ldots)$ can be obtained from $Ac_\nu^{(k)}(\ldots)$ by replacing cos with sin, and (μ_i, ϑ_i) are the coordinates of an arbitrary point in ith elliptic coordinate system related to the corresponding Cartesian system through formulas

$$x_i = \frac{1}{2}h_i\cosh\mu_i\cos\vartheta_i, \quad y_i = \frac{1}{2}h_i\sinh\mu_i\sin\vartheta_i;$$

$$i = 1, 2,$$

where μ_{21} and ϑ_{21} are the coordinates of the old origin O_1 expressed in the new system, α_{21} is the angle between axes O_1x_1 and O_2x_2 measured from axis O_1x_1 counterclockwise, $\varepsilon_0 = 1$, and $\varepsilon_\nu = 2$ for $\nu > 0$.

Potential (4) has the form

$$\varphi(M) = A_1\left(\ln\frac{1}{\rho_1} - \ln\frac{1}{\rho_1}\right)$$

$$+\sum_{k=1}^{\infty}[(a_k^1\cos k\vartheta_1 + b_k^1\sin k\vartheta_1)\exp(-k\mu_1) \quad (6)$$

$$+(-1)^{k+1}(a_k^1\cos k\bar{\vartheta}_1 - b_k^1\sin k\bar{\vartheta}_1)\exp(-k\bar{\mu}_1)].$$

In order to satisfy the boundary condition on ellipse Γ_1, let us express all variables in formula (6) through μ_1

and ϑ_1 with the help of formula (6) from [4],

$$\ln\frac{1}{\rho_1} = \ln\frac{4}{h_2} - \mu_{21} + 2\sum_{\nu=1}^{\infty}\frac{1}{\nu}(\cosh\nu\mu_2\cos\nu\vartheta_{21}\cos\nu\vartheta_2 + \sinh\nu\mu_2\sin\nu\vartheta_2)\exp(-\nu\mu_{21}),$$

formulas (5), and the source function written in elliptic coordinates [5],

$$\ln\frac{1}{\rho_1} = \ln\frac{4}{h_1} - \mu_1 + 2\sum_{\nu=1}^{\infty}\frac{1}{\nu}\exp(-\nu\mu_1)\cos\nu\frac{\pi}{2}\cos\nu\vartheta_1.$$

As a result, we obtain (using designations explained below) the following equation:

$$\varphi(M) = A_1\left[-\mu_1 + 2\sum_{\nu=1}^{\infty}\frac{1}{\nu}\exp(-\nu\mu_1)\cos\frac{\nu\pi}{2}\cos\nu\vartheta_1 + \bar{\mu}_{11} - 2\sum_{\nu=1}^{\infty}\frac{1}{\nu}(\cosh\nu\mu_1\cos\nu\bar{\vartheta}_{11}\cos\nu\vartheta_1 + \sinh\nu\mu_1\sin\nu\bar{\vartheta}_{11}\sin\nu\vartheta_1)\exp(-\nu\bar{\mu}_{11})\right] + \sum_{\nu=1}^{\infty}(a_\nu^1\cos\nu\vartheta_1 + b_\nu^1\sin\nu\vartheta_1)\exp(-\nu\mu_1)$$
$$-\sum_{k=1}^{\infty}(-1)^k a_k^1\sum_{\nu=0}^{\infty}\varepsilon_\nu[Ac_\nu^{(k)}(\bar{\mu}_{11},\bar{\vartheta}_{11};1,\bar{\alpha}_{11})\cosh\nu\mu_1\times\cos\nu\vartheta_1 + As_\nu^{(k)}(\bar{\mu}_{11},\bar{\vartheta}_{11};1,\bar{\alpha}_{11})\sinh\nu\mu_1\sin\nu\vartheta_1]$$
$$+\sum_{k=1}^{\infty}(-1)^k b_k^1\sum_{\nu=0}^{\infty}\varepsilon_\nu[As_\nu^{(k)}(\bar{\mu}_{11},\bar{\vartheta}_{11};1,\bar{\alpha}_{11})\cosh\nu\mu_1\times\cos\nu\vartheta_1 - Ac_\nu^{(k)}(\bar{\mu}_{11},\bar{\vartheta}_{11};1,\bar{\alpha}_{11})\sinh\nu\mu_1\sin\nu\vartheta_1].$$

Satisfying boundary condition $\varphi(M)|_{\Gamma_1} = \varphi(M)|_{\mu_1=\mu_1^0} = 1$, by virtue of the uniqueness of expansion into the Fourier series, we obtain the following infinite system of linear algebraic equations:

$$A_1(\bar{\mu}_{11} - \mu_1^0) - \frac{1}{2}\sum_{k=1}^{\infty}(\xi_{0k}^1 A_k^1 - \tilde{\omega}_{0k}B_k^1) = 1,$$

$$\varepsilon_\nu^1 A_1 + A_\nu^1 - \sum_{k=1}^{\infty}(\xi_{\nu k}^1 A_k^1 - \tilde{\omega}_{\nu k}B_k^1) = 0, \qquad (7)$$

$$\tilde{\varepsilon}_\nu^1 A_1 + B_\nu^1 - \sum_{k=1}^{\infty}(\tilde{\xi}_{\nu k}^1 A_k^1 + \omega_{\nu k}^1 B_k^1) = 0;\ \nu = 1, 2, \ldots,$$

which is obtained if we set $N_1 = 0$ and $N_2 = 1$ in the system presented below for arbitrary N_1 and N_2; new variables and coefficients are also defined in the description of arbitrary system.

Performing similar calculations in the general case, we obtain the following system:

$$-A_s\ln q_s + \sum_{j=1,\,j\neq s}^{N} A_j\ln\frac{\bar{\rho}_{sj}}{\rho_{sj}} + \sum_{k=1}^{\infty}\left[-A_k^s q_s^k + \sum_{j=1,\,j\neq s}^{N}(\sigma_{0k}^{sj}A_k^j + \tilde{\tau}_{0k}^{sj}B_k^j) + \sum_{j=N_1+1}^{N}(\alpha_{0k}^{sj}A_k^j + \tilde{\beta}_{0k}^{sj}B_k^j)\right] = f_s.$$

$$A_q(\bar{\mu}_{qq} - \mu_q^0) + \sum_{j=1,\,j\neq q}^{N} A_j(\bar{\mu}_{qj} - \mu_{qj}) + \frac{1}{2}\sum_{k=1}^{\infty}\left[\sum_{j=1}^{N_1}(\gamma_{0k}^{qj}A_k^j + \tilde{\delta}_{0k}^{qj}B_k^j) - \xi_{0k}^q A_k^q + \tilde{\omega}_{0k}^q B_k^q + \sum_{j=N_1+1,\,j\neq q}^{N}(\eta_{0k}^{qj}A_k^j + \tilde{\zeta}_{0k}^{qj}B_k^j)\right] = f_q,$$

$$-\frac{1}{\nu}q_s^\nu A_s + \sum_{j=1,\,j\neq s}^{N}\sigma_\nu^{sj}A_j + A_\nu^s + \sum_{k=1}^{\infty}\left[-L_{\nu k}q_s^{\nu+k}A_k^s + \sum_{j=1,\,j\neq s}^{N_1}(\sigma_{\nu k}^{sj}A_k^j + \tilde{\tau}_{\nu k}^{sj}B_k^j) + \sum_{j=N_1+1}^{N}(\alpha_{\nu k}^{sj}A_k^j + \tilde{\beta}_{\nu k}^{sj}B_k^j)\right] = 0,$$

$$\sum_{j=1,\,j\neq s}^{N}\tilde{\sigma}_\nu^{sj}A_j + B_\nu^s + \sum_{k=1}^{\infty}\left[-L_{\nu k}q_s^{\nu+k}B_k^s + \sum_{j=1,\,j\neq s}^{N_1}(\tilde{\sigma}_{\nu k}^{sj}A_k^j - \tau_{\nu k}^{sj}B_k^j) + \sum_{j=N_1+1}^{N}(\tilde{\alpha}_{\nu k}^{sj}A_k^j - \beta_{\nu k}^{sj}B_k^j)\right] = 0, \qquad (8)$$

$$\varepsilon_\nu^q A_q + \sum_{j=1,\,j\neq q}^{N}\tau_\nu^{qj}A_j + A_\nu^q + \sum_{k=1}^{\infty}\left[\sum_{j=1}^{N_1}(\gamma_{\nu k}^{qj}A_k^j + \tilde{\delta}_{\nu k}^{qj}B_k^j) - \xi_{\nu k}^q A_k^q + \tilde{\omega}_{\nu k}^q B_k^q\right.$$

$$+\sum_{j=N_1+1,\,j\neq q}^{N}(\eta_{\nu k}^{qj}A_k^j+\tilde{\zeta}_{\nu k}^{qj}B_k^j)\Bigg]=0,$$

$$\tilde{\varepsilon}_\nu^q A_q+\sum_{j=1,\,j\neq q}^{N}\tilde{\tau}_\nu^{qj}A_j+B_\nu^q$$

$$+\sum_{k=1}^{\infty}\Bigg[\sum_{j=1}^{N_1}(\tilde{\gamma}_{\nu k}^{qj}A_k^j-\delta_{\nu k}^{qj}B_k^j)-\tilde{\xi}_{\nu k}^{q}A_k^q-\omega_{\nu k}^{q}B_k^q$$

$$+\sum_{j=N_1+1,\,j\neq q}^{N}(\tilde{\eta}_{\nu k}^{qj}A_k^j-\zeta_{\nu k}^{qj}B_k^j)\Bigg]=0;$$

$$s=\overline{1,N_1};\quad q=\overline{N_1+1,N};\quad \nu=1,2,\ldots,$$

where

$$A_\nu^s=a_\nu^s R_\nu^s R_s^{-\nu},\quad B_\nu^s=b_\nu^s R_s^{-\nu};\quad s=\overline{1,N_1},$$

$$A_\nu^q=a_\nu^q\exp(-\nu\mu_q^0),\quad B_\nu^q=b_\nu^q\exp(-\nu\mu_q^0);$$

$$q=\overline{N_1+1,N};\quad q_s=\frac{R_s}{2l_s},\quad L_{\nu k}=\frac{(\nu+k-1)!}{\nu!(k-1)!},$$

l_s is the distance from point O_s to line $x=0$;

$$\varepsilon_\nu^q=\frac{2}{\nu}\Bigg[\exp(-\nu\mu_q^0)\cos\nu\frac{\pi}{2}$$

$$-\exp(-\nu\bar{\mu}_{qq})\cosh\nu\mu_q^0\cos\nu\bar{\vartheta}_{qq}\Bigg],$$

$$\tilde{\varepsilon}_\nu^q=-\frac{2}{\nu}\exp(-\nu\bar{\mu}_{qq})\sinh\nu\mu_q^0\sin\nu\bar{\vartheta}_{qq},$$

$$\sigma_\nu^{sj}=\frac{1}{\nu}\Bigg[\left(\frac{R_s}{\rho_{sj}}\right)^\nu\cos\nu\Theta_{sj}-\left(\frac{R_s}{\bar{\rho}_{sj}}\right)^\nu\cos\nu\bar{\Theta}_{sj}\Bigg],$$

$$\sigma_\nu^{sj}=\frac{1}{\nu k}\Bigg[(-1)^k\left(\frac{R_s}{\rho_{sj}}\right)^\nu\left(\frac{R_j}{\rho_{sj}}\right)^k\cos(\nu+k)\Theta_{sj}$$

$$-\left(\frac{R_s}{\bar{\rho}_{sj}}\right)^\nu\left(\frac{R_j}{\bar{\rho}_{sj}}\right)^k\cos(\nu+k)\bar{\Theta}_{sj}\Bigg],$$

$\tau_{\nu k}^{sj}$ is obtained from $\sigma_{\nu k}^{sj}$ by changing the sign between terms from minus to plus,

$$\tau_\nu^{qj}=\frac{2}{\nu}[\exp(-\nu\mu_{qj})\cos\nu\vartheta_{qj}$$

$$-\exp(-\nu\bar{\mu}_{qj})\cos\nu\bar{\vartheta}_{qj}]\cosh\nu\mu_q^0;$$

$\tilde{\sigma}_\nu^{sj}$, $\tilde{\sigma}_{\nu k}^{sj}$, $\tilde{\tau}_{\nu k}^{sj}$, and $\tilde{\tau}_\nu^{qj}$ are obtained from corresponding quantities without tilde by replacing cos with sin and cosh with sinh;

$$\alpha_{\nu k}^{sj}=R_s^\nu[Lc_\nu^{(k)}(\rho_{sj},\Theta_{sj};h_j,\alpha_{sj})$$

$$-(-1)^k Lc_\nu^{(k)}(\bar{\rho}_{sj},\bar{\Theta}_{sj};h_j,\bar{\alpha}_{sj})]\exp k\mu_j^0;$$

$\beta_{\nu k}^{sj}$ is obtained from $\alpha_{\nu k}^{sj}$ by changing sign minus between terms with sign plus; $\tilde{\alpha}_{\nu k}^{sj}$ and $\tilde{\beta}_{\nu k}^{sj}$ are obtained from corresponding quantities without tilde by replacing $Lc_\nu^{(k)}(\ldots)$ with $Ls_\nu^{(k)}(\ldots)$; functions $Lc_\nu^{(k)}(\ldots)$, $Ls_\nu^{(k)}(\ldots)$, $Dc_\nu^{(k)}(\ldots)$, and $Ds_\nu^{(k)}(\ldots)$ are defined in [2];

$$\gamma_{\nu k}^{qj}=2R_j^k[Dc_\nu^{(k)}(\mu_{sj},\vartheta_{qj};h_q,\alpha_{qj})$$

$$-(-1)^k Dc_\nu^{(k)}(\bar{\mu}_{qj},\bar{\vartheta}_{qj};h_q,\bar{\alpha}_{qj})]\cosh\nu\mu_q^0,$$

$$\delta_{\nu k}^{qj}=2R_j^k[Dc_\nu^{(k)}(\mu_{qj},\vartheta_{qj};h_q,\alpha_{qj})$$

$$+(-1)^k Dc_\nu^{(k)}(\bar{\mu}_{qj},\bar{\vartheta}_{qj};h_q,\bar{\alpha}_{qj})]\sinh\nu\mu_q^0,$$

$$\eta_{\nu k}^{qj}=2[Ac_\nu^{(k)}(\mu_{qj},\vartheta_{qj};h_q/h_j,\alpha_{qj})$$

$$+(-1)^k Ac_\nu^{(k)}(\bar{\mu}_{qj},\bar{\vartheta}_{qj};h_q/h_j,\bar{\alpha}_{qj})]\exp k\mu_j^0\cosh\nu\mu_q^0,$$

$$\zeta_{\nu k}^{qj}=2[Ac_\nu^{(k)}(\mu_{qj},\vartheta_{qj};h_q/h_j,\alpha_{qj})$$

$$+(-1)^k Ac_\nu^{(k)}(\bar{\mu}_{qj},\bar{\vartheta}_{qj};h_q/h_j,\bar{\alpha}_{qj})]\exp k\mu_j^0\sinh\nu\mu_q^0,$$

$$\xi_{\nu k}^{q}=2(-1)^k\exp k\mu_q^0\cosh\nu\mu_q^0 Ac_\nu^{(k)}(\bar{\mu}_{qq},\bar{\vartheta}_{qq};1,\bar{\alpha}_{qq}),$$

$$\omega_{\nu k}^{q}=2(-1)^k\exp k\mu_q^0\sinh\nu\mu_q^0 Ac_\nu^{(k)}(\bar{\mu}_{qq},\bar{\vartheta}_{qq};1,\bar{\alpha}_{qq}),$$

$\tilde{\gamma}_{\nu k}^{qj}$, $\tilde{\delta}_{\nu k}^{qj}$, $\tilde{\eta}_{\nu k}^{qj}$, $\tilde{\zeta}_{\nu k}^{qj}$, $\tilde{\xi}_{\nu k}^{q}$, and $\tilde{\omega}_{\nu k}^{q}$ can be obtained from corresponding quantities without tilde by replacing $Xc_\nu^{(k)}(\ldots)$ with $Xs_\nu^{(k)}(\ldots)$ and interchanging cosh and sinh; $(\rho_{sj},\Theta)sj)$ and $(\bar{\rho}_{sj},\bar{\Theta}_{sj})$ are the coordinates of points O_j and $\bar{O}_j$ written in polar coordinates with origin O_s; $(\mu_{qj},\vartheta_{qj})$ and $(\bar{\mu}_{qj},\bar{\vartheta}_{qj})$ are the coordinates of the same points written in elliptic coordinates with origin O_q; α_{pr} and $\bar{\alpha}_{pr}$ are the angles between axis $O_p x_p$ and axes $O_r x_r$ and $\bar{O}_r\bar{x}_r$, respectively, measured from the latter axes counterclockwise; and $p,r=\overline{1,N}$.

Since circles and ellipses have common points neither with each other nor with the line $x=0$, system (8) has a completely continuous form [6] and, by virtue of the uniqueness of solution to the considered problem, has a unique solution belonging to domain l^2, which can be found by the method of reduction. It follows from

formula (4) that

$$\int_{\Gamma_j} \frac{\partial \varphi}{\partial n} dl = 2\pi A_j; \quad j = \overline{1, N}, \tag{9}$$

which implies that, in order to calculate any capacitance, it is sufficient to know only coefficients A_j, $j = \overline{1, N}$. The coefficients and free terms of the reduced system (of minimum dimensionality),

$$-A_s \ln q_s + \sum_{j=1,\, j\neq s}^{N} A_j \ln \frac{\bar{\rho}_{sj}}{\rho_{sj}} = f_s,$$

$$A_q(\bar{\mu}_{qq} - \mu_q^0) + \sum_{j=1,\, j\neq q}^{N} A_j(\bar{\mu}_{qj} - \mu_{qj}) = f_q, \tag{10}$$

$$s = \overline{1, N_1}, \quad q = \overline{N_1 + 1, N},$$

completely and uniquely determine both the dimensions and mutual positions of wires and the boundary conditions. System (10) allows one to determine all coefficients A_j; therefore, it is natural to use it for the approximate calculation of electric capacitances.

Note that, for $s = \overline{1, 0}$ or $a = \overline{N_1 + 1, N_1}$, equation (10) does not contain the first or the second equation, respectively.

Let us consider several particular cases.

(a) $N_1 = 1$ and $N_2 = 0$. In this case, system (10) takes the form

$$-A_1 \ln q_1 = 1.$$

Then

$$C_l \approx 2\pi\varepsilon/(2l_1/R_1),$$

which is consistent with [1, formulas (4)–(1)] if we take into account the properties of capacitance [1, formula (B-18)].

If $N_1 = 0$ and $N_2 = 1$, system (10) takes the form

$$A_1(\bar{\mu}_{11} - \mu_1^0) = 1.$$

Then

$$C_l \approx 2\pi\varepsilon/(\bar{\mu}_{11} - \mu_1^0),$$

which is consistent with [7, formula (13)] with an allowance for the same properties of capacitance. In order to calculate $\bar{\mu}_{11}$, one can use a method described in [7].

(b) $N_1 = 1$ and $N_2 = 1$. System (10) takes the form

$$\begin{aligned} -A_1 \ln q_1 + A_2 \ln(\bar{\rho}_{12}/\rho_{12}) &= f_1, \\ A_1(\bar{\mu}_{21} - \mu_{21}) + A_2(\bar{\mu}_{22} - \mu_2^0) &= f_2. \end{aligned} \tag{11}$$

For $f_1 = 1$ and $f_2 = 0$ (or vice versa), we obtain

$$C_{12} = C_{21} \approx \pi\varepsilon[\mu_{21} - \bar{\mu}_{21} + \ln(\rho_{12}/\bar{\rho}_{12})]/\Delta \tag{12}$$

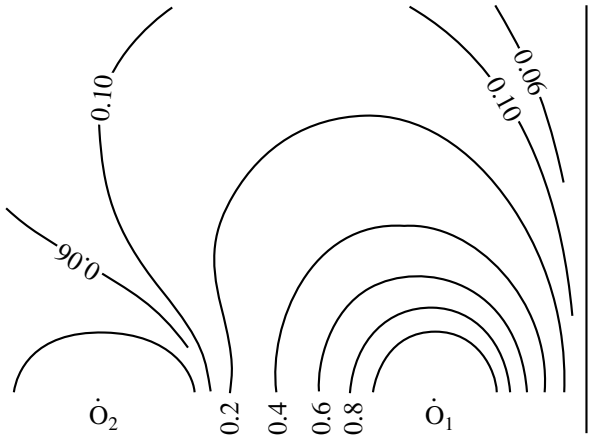


Figure.

[system (11) yields $C_{12} \neq C_{21}$, and we present here an arithmetical mean of the two quantities].

For $f_1 = 1$ and $f_2 = 1$, we obtain

$$C_{10} \approx 2\pi\varepsilon[\ln(\bar{\rho}_{12}/\rho_{12}) - (\bar{\mu}_{22} - \mu_2^0)]/\Delta, \tag{13}$$

$$C_{20} \approx 2\pi\varepsilon[\ln q_1 + (\bar{\mu}_{21} - \mu_{21})]/\Delta, \tag{14}$$

where

$$\Delta = (\bar{\mu}_{21} - \mu_{21})\ln(\bar{\rho}_{12}/\rho_{12}) + (\bar{\mu}_{22} - \mu_2^0)\ln q_1.$$

In particular, if the circle is centered at point $O_1(-l_1, 0)$ in system xOy, the ellipse with semiaxes a and b is centered at point $O_1(-l_2, 0)$, $l_2 > l_1$, and the ellipse major axis is perpendicular to line $x = 0$, then ($c = \sqrt{a^2 - b^2}$)

$$\mu_{21} = \operatorname{Arccosh}\frac{l_2 - l_1}{c}, \quad \bar{\mu}_{21} = \operatorname{Arccosh}\frac{l_1 + l_2}{c},$$

$$\bar{\mu}_{22} = \operatorname{Arccosh}\frac{2l_2}{c}, \quad \mu_2^0 = \operatorname{Arccosh}\frac{a}{c},$$

Table

l_1	l_2	Value of $C_{12}/(2\pi\varepsilon)$		Relative error, %
		"exact"	approx.	
15.5	36.5	0.207196	0.203787	1.6452
15.5	46.5	0.129713	0.132149	–1.8774
25.5	46.5	0.222773	0.214002	3.9374
25.5	56.5	0.143918	0.143126	0.5500
25.5	66.5	0.107671	0.108142	–0.4381
35.5	56.5	0.232711	0.222177	4.5269
35.5	66.5	0.153371	0.151506	1.2160
35.5	76.5	0.116508	0.116210	0.2554
35.5	86.5	0.094395	0.094538	–0.1518

$$\rho_{12} = l_2 - l_1, \quad \bar{\rho}_{12} = l_1 + l_2.$$

In order to calculate the errors of formulas (12)–(14), a numerical experiment was performed which consisted in solving system (8) using the method of reduction and comparing the "exact" value of capacitance with the approximate one. In this case, system (8) separates into two systems: a homogeneous system for B_ν^1 and B_ν^2 with a trivial solution and an inhomogeneous system for A_1, A_2, A_ν^1, and A_ν^2 ($\nu = 1, 2, \ldots$). The table presents several numerical results in this particular case for an ellipse with semiaxes $a = 5$ and $b = 3$ and a circle with radius $R_1 = 5$ versus l_1 and l_2.

Note that the dimensionality of reduced system used to obtain the "exact" solution depends strongly on the mutual arrangement of wires and the conducting plane. In order to obtain values presented in the table with six or three correct decimal places, it is sufficient to take ten or five equations, respectively. If, for example, $l_1 = 6$ and $l_2 = 17$, then one should take 38 equations in order to obtain six correct decimal places.

The figure shows equipotential field lines. Due to the symmetry, the lines are shown only for $y \geq 0$. Calculations were performed for $a = 3$, $b = 2$, $R_1 = 2$, $l_1 = 5$, and $l_2 = 16$.

For arbitrary N_1 and N_2, it is easy to write, using determinants, an approximate formula for the calculation of any capacitance similarly to formulas [1, formulas (6)–(21)].

The proposed method for the exact and approximate calculation of capacitances can be used by engineers and scientists of various specialities in their research and practical work.

REFERENCES

1. Yu. Ya. Iossel', É. S. Kochanov, and M. G. Strunskiĭ, *Calculation of Electric Capacitance* (Énergoizdat, Leningrad, 1981).
2. Z. M. Narkun, Differ. Uravn. **19**, 654 (1983).
3. Z. M. Narkun, in *Studies in Mathematics and Physics: A Collection of Articles* (Grodn. Gos. Univ, Grodno, 1978), pp. 144–147.
4. Z. M. Narkun, Differ. Uravn. **15**, 357 (1979).
5. P. M. Morse and H. Feshbach, *Methods of Theoretical Physics* (McGraw–Hill, New York, 1953; Inostrannaya Literatura, Moscow, 1960), Vol. 2.
6. L. V. Kantorovich and B. I. Krylov, *Approximate Methods of Higher Analysis* (Fizmatgiz, Moscow, 1962).
7. V. K. Boĭko and Z. M. Narkun, Élektrichestvo, No. 2, 76 (1991).

Translated by A. Kondrat'ev

Technical Physics, Vol. 45, No. 2, 2000, pp. 147–152. Translated from Zhurnal Tekhnicheskoĭ Fiziki, Vol. 70, No. 2, 2000, pp. 6–11.

The Hall Effect and Oscillating Decay of a Magnetic Field

D. A. Shalybkov and V. A. Urpin

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

Received September 28, 1998; in final form, August 2, 1999

Abstract—The decay of a strong magnetic field in a conducting matter is considered. It is shown that nondissipative Hall currents can considerably change the behavior of the field when it decays. The nonlinear character of the Hall effect leads to the generation of fields of high multipolarity even for most simple initial magnetic configurations. In particular, the evolution of an initially dipole configuration may give rise not only to quadrupole or higher poloidal harmonics but also to a toroidal field that is other than zero only inside the conductor. The nonlinear Hall currents relate different harmonics to each other and, in a sufficiently strong field, may provide efficient energy exchange between them. Due to this redistribution of the magnetic energy, the evolution of different harmonics has an oscillating character. The oscillation period is determined by the characteristic time of Hall drift and may be fairly short in strong magnetic fields. *© 2000 MAIK "Nauka/Interperiodica".*

INTRODUCTION

The conductivity of a matter subjected to sufficiently strong magnetic field is anisotropic and is described by tensors. As a result, the decay of strong magnetic fields is an essentially nonlinear process and may noticeably differ from a simple exponential law, which, in particular, describes the decay of the fundamental harmonics in a uniform conducting sphere (see, for example, [1]). In this study, we investigate the influence of the Hall effect and its associated conductivity anisotropy on the magnetic field evolution. The Hall current arises owing to electron drift perpendicularly to both the electric and the magnetic fields and is nondissipative, since it does not increase the entropy density of the matter. In magnetized media where $\omega_B\tau \gg 1$ ($\omega_B = eB/mc$ is the gyrofrequency of electrons, and τ is their relaxation time), the Hall drift may change the geometry of both the electric currents and the magnetic field and thus indirectly influence the decay of the magnetic fields. The particular case of a purely toroidal magnetic configuration was considered in [2]. It was shown that Hall currents may considerably accelerate magnetic field decay. This is due to the fact that Hall drift forms regions with a high electric current density and a highly nonuniform magnetic field, where dissipation is accelerated. In addition, it was found that the field symmetry in the process of evolution also can change. For example, a toroidal field that is initially mirror-symmetric with respect to the equatorial plane may become mirror-asymmetric due to the Hall effect and then, once the field has decreased, restore the symmetry.

In this work, we consider the influence of Hall currents on the evolution of poloidal magnetic configurations. As is known (see, for example, [3]), in these configurations, $\mathbf{B} \neq 0$ neither inside nor outside a conductor (we recall that, in toroidal fields, $\mathbf{B} = 0$ outside the conductor). The evolution of poloidal fields may differ from that of purely toroidal fields. First, the Hall currents may excite toroidal harmonics even from an initially purely poloidal field, whereas the reverse is not true. Hall coupling between different field components makes possible energy transfer between the components. As a result, ohmic damping of the currents may be accompanied by amplitude oscillations of different components (on the background of a monotonic decrease in the total magnetic energy). These oscillations are similar to well-known helical oscillations in a plasma (see, for example, [4]), which are also associated with electron drift.

A number of attempts were made to investigate the influence of Hall currents on magnetic field dissipation [5, 6]. However, in order to simplify the numerical calculations, the authors of [5, 6] used the angle-averaged magnetic field. In their formulas, the terms responsible for the generation of the toroidal field with only one harmonic of the poloidal field present were erroneously omitted. In addition, in [5], the Hall parameter was considered to be small, $\omega_B\tau \ll 1$. In this case, the Hall drift of electrons is small as compared with the ohmic dissipation rate and, correspondingly, so are deviations from the damping obtained in the linear approximation.

Notice that nonlinear effects associated with the conductivity anisotropy and, in particular, with the Hall effect are of particular significance for astrophysical applications (see, for example, [7]), where the magnetic fields in a number of objects (white dwarfs and neutron stars) may be as high as $\sim 10^8$–10^{13} G and strongly magnetize the plasma.

1063-7842/00/4502-0147$20.00

In Section 1, fundamental equations describing the evolution of a magnetic field in a matter with anisotropic conductivity are given. In Section 3, results of the numerical calculations are discussed.

(1) In the magnetohydrodynamic approximation, the evolution of magnetic field **B** obeys the Faraday law of induction

$$\frac{\partial \mathbf{B}}{\partial t} = -c\nabla \times \mathbf{E} \tag{1}$$

and the Ampere law

$$\nabla \times \mathbf{B} = \frac{4\pi}{c}\mathbf{j}, \tag{2}$$

where **E** is the electric field and **j** is the current density.

When macroscopic motions in the matter are absent, Ohm's law takes the form (see, for example, [1])

$$\mathbf{E} = R_\parallel \mathbf{j}_\parallel + R_\perp \mathbf{j}_\perp + R_H \mathbf{j} \times b, \tag{3}$$

where $R_\parallel$ and $R_\perp$ are the components of the electrical resistance tensor along and transverse to the magnetic field, respectively; R_H is the Hall component; and $\mathbf{b} = \mathbf{B}/|\mathbf{B}|$ is the unity vector along the magnetic field.

The Hall current in matter is due to charge drift perpendicularly to the electric and magnetic fields. This current is nondissipative, since it does not contribute to the entropy density Q of the matter:

$$Q = \mathbf{E}\mathbf{j} = R_\parallel j_\parallel^2 + R_\perp j_\perp^2. \tag{4}$$

In the relaxation time approximation (see, for example, [8]), expressions for the components of the electrical resistance tensor take a simple form

$$R_\parallel = R_\perp \equiv R_0 \frac{m}{en\tau}, \quad R_H = \frac{B}{nce}, \tag{5}$$

where $e = |e|$ and m are the charge and the effective mass of an electron, respectively.

In this approximation, we have

$$\frac{R_H}{R_0} \equiv \alpha = \omega_B \tau. \tag{6}$$

At $\omega_B\tau \ll 1$, $R_H \ll R_0$ and charge transfer is scarcely affected by the magnetic field. If $\omega_B\tau \gg 1$, $R_H \gg R_0$ and the influence of the magnetic field is significant. Notice that the Hall resistance R_H does not depend on the relaxation time because of the nondissipative character of Hall currents.

Equations (1)–(3) with allowance for (5) and (6) set up a closed system for solving the given problem. For simplicity, let us consider the field evolution in a uniform isothermic conducting sphere. In this case, the values of R_0 and α/B do not depend on the coordinates. Then, equation of induction (1) takes the form

$$\frac{\partial \mathbf{B}}{\partial t} = -\frac{cR_0}{4\pi}\left\{\nabla \times (\nabla \times \mathbf{B}) + \frac{\alpha}{B}\nabla \times ((\nabla \times \mathbf{B}) \times \mathbf{B})\right\}. \tag{7}$$

An arbitrary magnetic field may be represented as the sum of the poloidal, $\mathbf{B}_p$, and toroidal, $\mathbf{B}_t$, components. For axially symmetric configurations in a spherical coordinate system, we have (see, for example, [3])

$$\mathbf{B} = \mathbf{B}_p + \mathbf{B}_t = \nabla \times A(r, \theta)\mathbf{e}_\varphi + B_\varphi(r, \theta)\mathbf{e}_\varphi, \tag{8}$$

where $\mathbf{e}_\varphi$ is the unit azimuth vector and r, θ, and φ are spherical coordinates.

Substituting (8) into (7) yields a system of two equations that describe the evolution of the poloidal and toroidal magnetic fields:

$$\frac{\partial \mathbf{B}_p}{\partial t} = -\frac{c^2 R_0}{4\pi}\{\nabla \times (\nabla \times \mathbf{B}_p) + \beta\nabla \times ((\nabla \times \mathbf{B}_t) \times \mathbf{B}_p)\}, \tag{9}$$

$$\frac{\partial \mathbf{B}_t}{\partial} = -\frac{c^2 R_0}{4\pi}\{\nabla \times (\nabla \times \mathbf{B}_t) + \beta\nabla \times ((\nabla \times \mathbf{B}_p) \times \mathbf{B}_p + (\nabla \times \mathbf{B}_t) \times \mathbf{B}_t)\}, \tag{10}$$

where $\beta = \alpha/B$.

Notice that $\nabla \times \mathbf{B}_p$ is a toroidal vector; consequently, $(\nabla \times \mathbf{B}_p) \times \mathbf{B}_t = 0$. For $\alpha = 0$ (i.e., for negligible Hall currents), the toroidal and poloidal components evolve independently. In this case, the eigenmodes of both components decay exponentially with the characteristic decay time increasing with decreasing mode number. The characteristic decay time of the fundamental poloidal mode (the dipole field) is

$$\tau_0 = \frac{4a^2}{\pi c^2 R_0}, \tag{11}$$

where a is the radius of the sphere.

At $a \gg 1$, the poloidal and toroidal components prove to be strongly coupled. It is seen from equation (10) that, in the general case, $(\nabla \times \mathbf{B}_p) \times \mathbf{B}_p \neq 0$ even for the purely dipole initial configuration, which results in the generation of the toroidal field. The toroidal field, in its turn, generates higher harmonics of the poloidal field according to the nonlinear term in the right part of equation (9). Notice that the Hall current does not generate the poloidal field when the initial configuration is purely toroidal. However, the toroidal field may generally come about from the poloidal one for the purely dipole configuration. As a result, to study the influence of the Hall current on the decay of the poloidal field, there is no need to consider a complex initial superposition of the toroidal and poloidal fields, as was argued

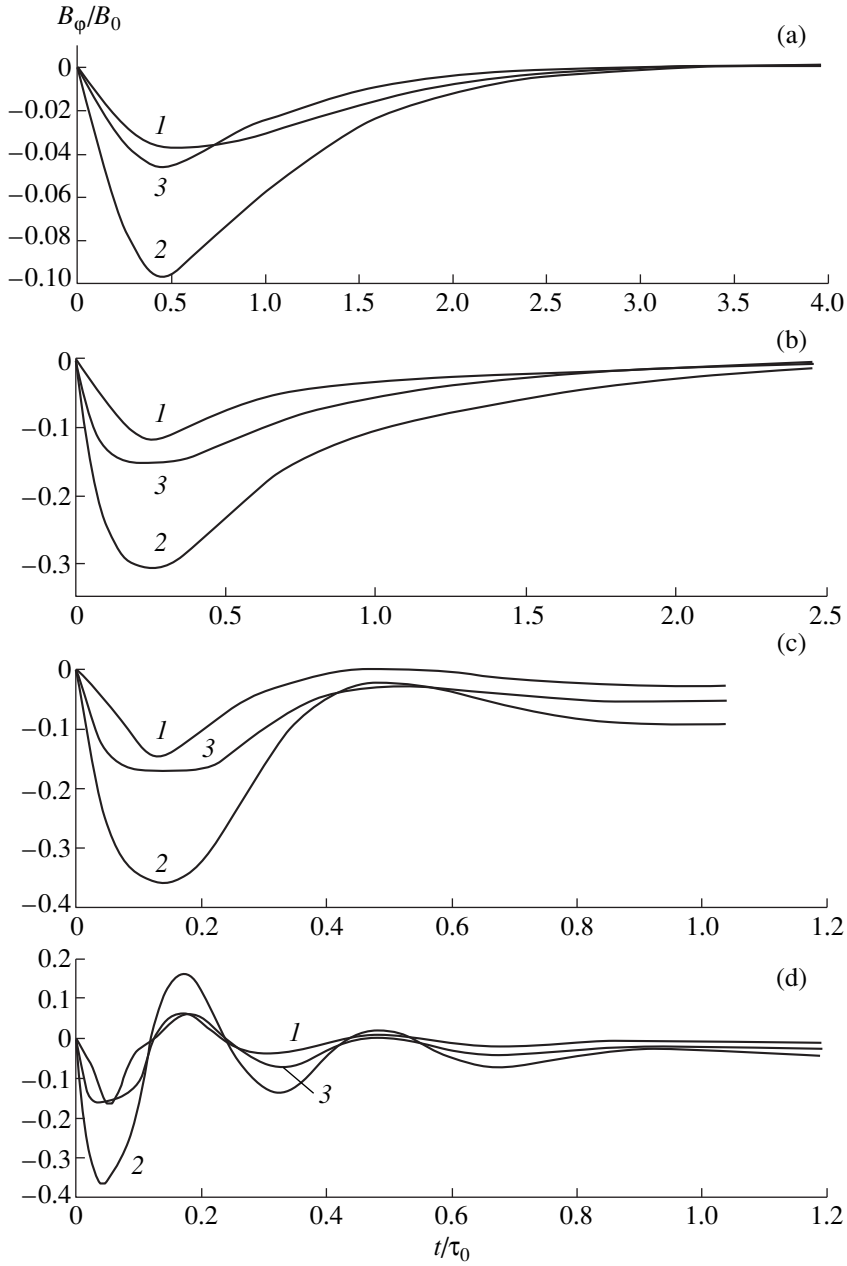


Fig. 1. Time dependence of the toroidal field for $\theta = 140°$ and α = (a) 0.2, (b) 1, (c) 2, and (d) 5. The curves are obtained for r = (*1*) 0.3*a*, (*2*) 0.7*a*, and (*3*) 0.9*a*. The field is normalized to the initial value of the equatorial poloidal magnetic field. The time is normalized to τ_0 (11).

in [5, 6]. For an initial distribution, we chose

$$A(r, \theta) = \pi B_0 \sin(\theta)\left[\frac{\sin(\pi r/a)}{(\pi r/a)^2} - \frac{\cos(\pi r/a)}{\pi r/a}\right], \quad B_\varphi = 0 \tag{12}$$

as the simplest.

The normalizing constant in (12) was taken such that the surface value of the equatorial field was equal to B_0. Vector potential (12) corresponds to the fundamental dipole mode for a uniform sphere with the characteristic decay time (11) at $\alpha = 0$.

For equations (9) and (10), the boundary conditions were the following. First, the conducting sphere was supposed to be in a vacuum. Hence, $\mathbf{B}_t$ vanishes on the surface, and bfB_p is joined together with the vacuum magnetic field when $r = a$. In addition, both the toroidal field $\mathbf{B}_\varphi$ and the potential A must go to zero when $r \longrightarrow 0$.

(2) Equations (9) and (10) were solved numerically using conventional explicit difference methods (see, for

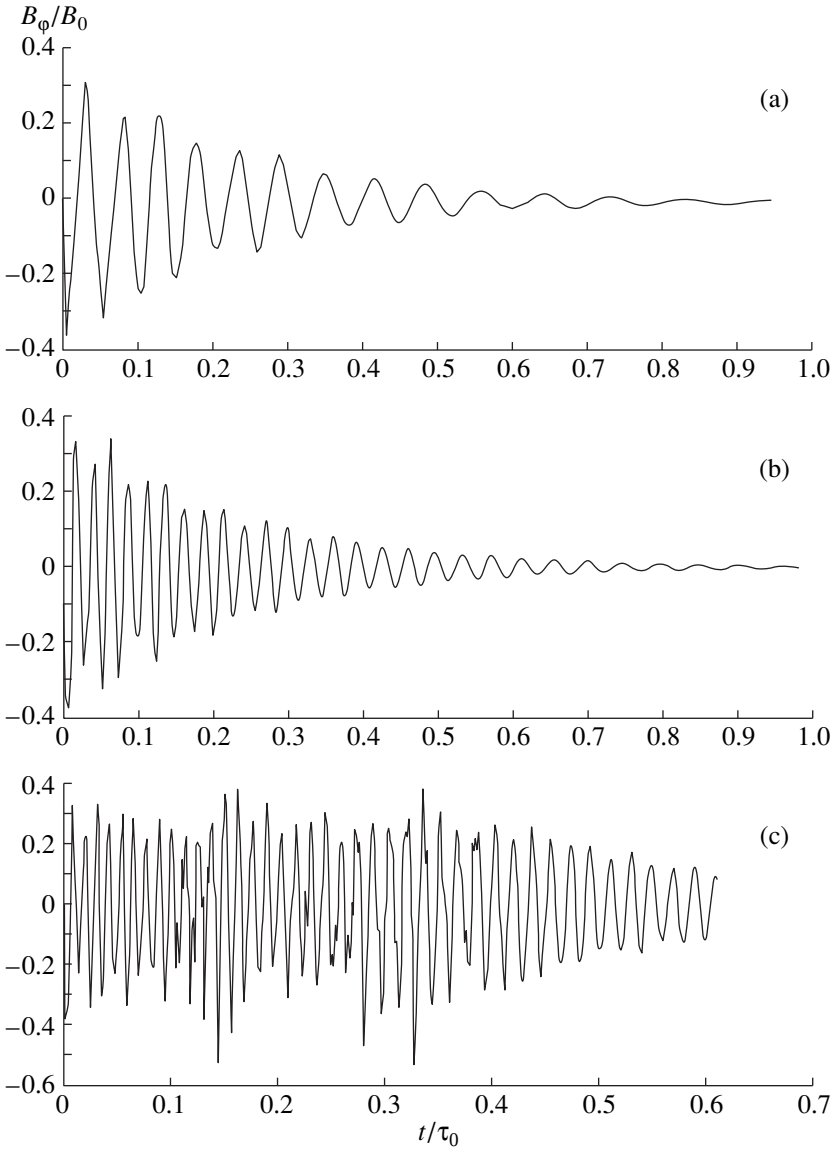


Fig. 2. The same as in Fig. 1 for α = (a) 25, (b) 50, and (c) 100 at $r = 0.7a$.

example, [2]). If we introduce the dimensionless time $\tau = (c^2R_0/4\pi a^2)t$, then (9) and (10) contain the only parameter $\alpha_0 = \beta B_0$, which is equal to the Hall coefficient at the initial time instant. We evaluated the field evolution for $\alpha_0 = 0.2$, 1, 2, 5, 25, 50, and 100. In all the calculations, the external boundary condition was approximated by the first nine multipoles.

Hall currents lead to the generation of higher modes from the initial dipole mode. Both poloidal and toroidal modes are generated. The amplitudes of the poloidal modes rapidly drop with the mode number. The toroidal field, generated at the initial stage of the evolution, is antisymmetric with respect to the equatorial plane ($B_\varphi = 0$ at $\theta = \pi/2$). Such a behavior of the toroidal field results from the symmetry of the initial configuration and the absence of the flux through the azimuth plane at the initial time instant.

The nonlinearity of the problem makes the individual modes, including the initial dipole one, oscillate. The oscillations are caused by nondissipative energy exchange between the individual modes and are similar to well-known helical oscillations in a magnetized plasma (see, for example, [9]). At $\alpha_0 \leq 1$, the dissipative processes go more rapidly than the Hall ones and the oscillations have no time to set up; conversely, at $\alpha_0 \gg 1$, the oscillations are well pronounced (Figs. 1, 2). They are most defined for the case of the toroidal field.

In Figs. 1 and 2, time variations of the toroidal field are presented for various values of the radius and for $\theta = 0.8\pi$. Initially, as long as the poloidal field is not noticeably attenuated, the oscillation period is τ_{H0} =

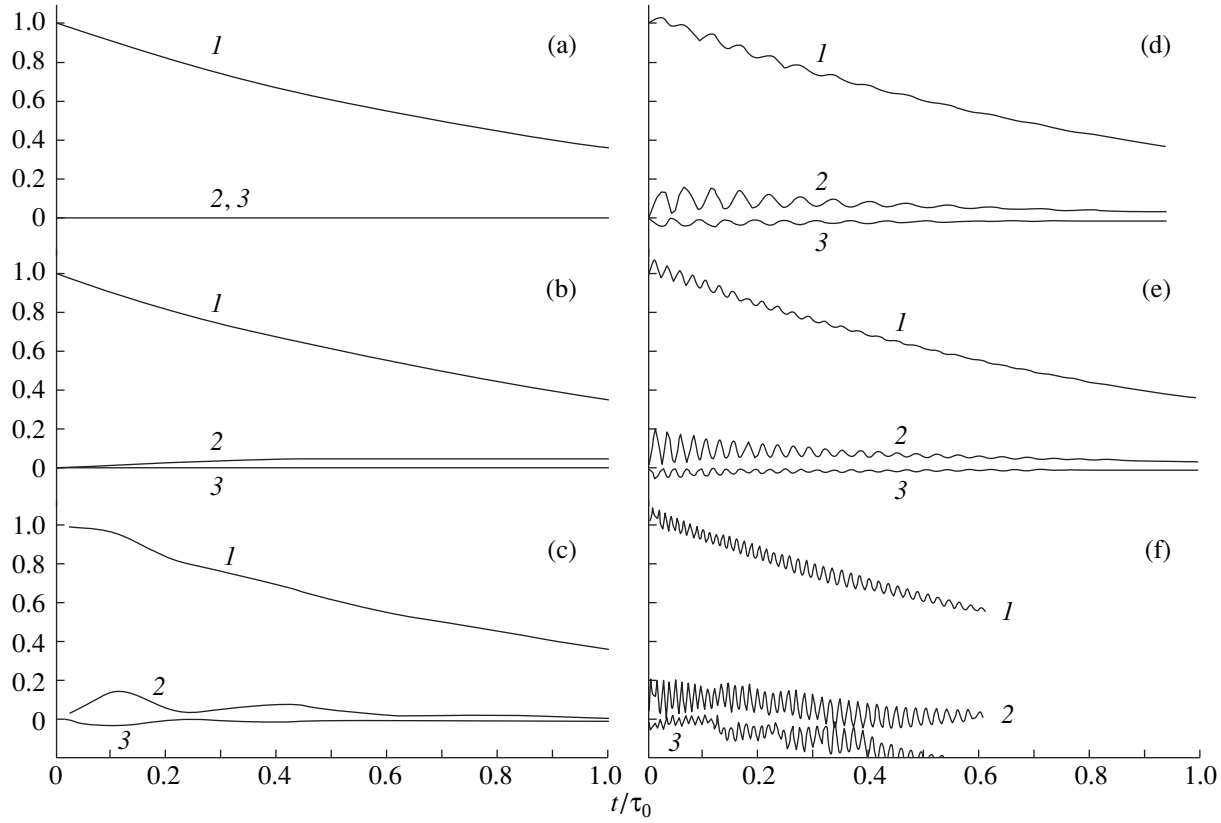


Fig. 3. Evolution of the poloidal multipoles on the spherical surface for l = (*1*) 1, (*2*) 3, and (*3*) 5 and α = (a) 0.2, (b) 1, (c) 5, (d) 25, (e) 50, and (f) 100. The multipoles are normalized to the initial value of the equatorial dipole field.

τ_0/α_0 [τ_0 is given by (11)]. Then, it increases as the poloidal field decays. The variation of the period with time is approximately described by the formula

$$\tau_H \approx \tau_{H0}\exp(-t/\tau_0). \qquad (13)$$

The oscillation amplitude depends on r and θ. The value of $\mathbf{B}_\varphi$ reaches its maximum when $\theta \simeq 40°$ and 140° (we recall that $\mathbf{B}_\varphi$ is antisymmetric).

It is interesting to note that the oscillation amplitude of the toroidal field peaks (~0.3) even at $\alpha_0 \sim 1$. With a further increase in α, it remains unchanged.

As to the total magnetic-field energy, the influence of the Hall currents on its damping is insignificant for all α_0's. During the evolution, the total magnetic-field energy practically does not differ from the value obtained without regard for the Hall currents. This is due to fact that the energy of higher multipoles, excited by the Hall currents, does not exceed 10% of the dipole field energy. The last statement is well illustrated by the behavior of poloidal multipoles on the surface (Fig. 3). We recall that the external boundary condition was approximated by multipoles with numbers $l \le 9$. However, even for $l \ge 5$, their contribution to the total magnetic-field energy is negligible. Therefore, in Fig. 3, only the multipoles with l = 1, 3, 5 are shown (those with even numbers do not contribute because of the unsuitable parity). Like the toroidal magnetic field, the poloidal modes oscillate, the oscillations being inherent in not only newly generated modes with $l > 1$ but also in the initial dipole component. For the dipole component, however, the oscillation amplitude is small (~10% of the nonoscillating part). For l = 3 and l = 5, the oscillations are stronger: for l = 3, the oscillating part is close to the average value, and for l = 5, the evolution occurs even with the sign changed. The period of the poloidal oscillations is obviously the same as for toroidal, since both are due to intermode energy exchange.

(3) We directly evaluated the influence of Hall currents on the evolution of a magnetic field in a conducting sphere. A very simple model allowing us to characterize the field decay with regard for the Hall currents was used. It turned out that, even for the simplest initial configuration, the dipole field outside the sphere, the field evolution is fairly complicated. The Hall currents lead to energy exchange between different modes and hence to the generation of higher multipoles of both the poloidal and the toroidal fields. The most interesting feature of the evolution seems to be that all modes of the magnetic field oscillate. The nature of these oscillations is easy to understand by a simple example of linear waves in a magnetized plasma.

Consider small disturbances of both toroidal, $\delta\mathbf{B}_t$, and poloidal, $\delta\mathbf{B}_p$, fields in the presence of an undisturbed poloidal field $\mathbf{B}_{p0}$. For simplicity, let the wavelength of the disturbances be much smaller than both the radius of the sphere and the characteristic scale of $\mathbf{B}_{p0}$ variation. Then, the behavior of $\delta\mathbf{B}_p$ and $\delta\mathbf{B}_t$ is described by linearized equations (9) and (10):

$$\frac{\partial \delta\mathbf{B}_p}{\partial t} = -\frac{c^2 R_0}{4\pi}[\nabla\times(\nabla\times\delta\mathbf{B}_p) + \beta\nabla\times((\nabla\times\delta\mathbf{B}_t)\times\mathbf{B}_{p0})], \quad (14)$$

$$\frac{\partial \delta\mathbf{B}_t}{\partial t} = -\frac{c^2 R_0}{4\pi}[\nabla\times(\nabla\times\delta\mathbf{B}_t) + \beta\nabla\times((\nabla\times\delta\mathbf{B}_p)\times\mathbf{B}_{p0})]. \quad (15)$$

As in the case of a large-scale field, the inclusion of the Hall currents causes terms proportional to $(\sim\beta)$ to appear that relate the toroidal and the poloidal magnetic fields. In our approximation, one can consider that $\delta\mathbf{B}_p$ and $\delta\mathbf{B}_t$ are proportional to $\exp(i\omega t - i\mathbf{kr})$, where ω and $\mathbf{k}$ are the frequency and the wave vector, respectively. For a plane wave, equations (14) and (15) are transformed to the form

$$(i\omega + \omega_\sigma)\delta\mathbf{B}_p = -\frac{c^2 R_0\beta}{4\pi}(\mathbf{kB}_{p0})\mathbf{k}\times\delta\mathbf{B}_t, \quad (16)$$

$$(i\omega + \omega_\sigma)\delta\mathbf{B}_t = -\frac{c^2 R_0\beta}{4\pi}(\mathbf{kB}_{p0})\mathbf{k}\times\delta\mathbf{B}_p, \quad (17)$$

where $\omega_\sigma = c^2 R_0 k^2/4\pi$ is the reciprocal of the ohmic dissipation time.

Expressions (16) and (17) provide a better insight into the relation between the poloidal and toroidal fields. The Hall current proportional to $\delta\mathbf{B}_t$ changes the value and, consequently, the energy of the poloidal field. The toroidal field, in its turn, is changed under the action of the Hall current that is proportional to $\delta\mathbf{B}_p$. Since the Hall currents are nondissipative, they influences $\mathrm{Re}\,\omega$ only, whereas $\mathrm{Im}\,\omega$ is determined by ohmic dissipation. From (16) and (17), we have the dispersion relationship

$$\omega = \pm\omega_\sigma\beta(\mathbf{kB}_{p0})/k + i\omega_\sigma. \quad (18)$$

In strong magnetic fields ($\beta\mathbf{B}_{p0} \gg 1$), equation (18) is a wave with a period $P = 2\pi k/\omega_\sigma\beta(\mathbf{kB}_{p0})$ and an amplitude slowly decreasing with a characteristic time $\sim\omega_\sigma^{-1}$. Such waves (sometimes called helicons) are due to the Hall effect and can exist only in a magnetized plasma.

Similarly to the small-scale disturbances, the large-scale field also oscillate with a period $\tau_H = \tau_0/\alpha$. With time, α decreases; conversely, the period increases, approaching τ_0. The oscillation amplitude can reach ~0.1 that of the initial dipole mode. For the toroidal component, the amplitude rapidly reaches saturation with increasing α_0 and becomes α_0-independent as $\alpha_0 \geq 1$. The maximal amplitude is ~0.3 the initial value of the poloidal field on the equator. For $\alpha \leq 1$, the influence of the Hall currents becomes negligible.

To conclude, notice that no more than 10% of the energy of the initial dipole component evolves into higher-order multipoles. The influence of the Hall currents on the damping rate of the magnetic energy is therefore insignificant, which basically differs from the case of a purely toroidal field.

ACKNOWLEDGMENTS

This work was financially supported by the Russian Foundation for Basic Research (grant no. 97-02-18096) and the INTAS (grant no. 96-0154).

REFERENCES

1. L. D. Landau and E. M. Lifshitz, *Electrodynamics of Continuous Media* (Nauka, Moscow, 1982; Pergamon, Oxford, 1960).
2. V. A. Urpin and D. A. Shalybkov, Zh. Éksp. Teor. Fiz. **100**, 1272 (1991) [Sov. Phys. JETP **73**, 703 (1991)].
3. S. Chandrasekhar and K. Prendergrast, Proc. Nat. Acad. Sci. **42**, 5 (1956).
4. E. M. Lifshitz and L. P. Pitaevskiĭ, *Physical Kinetics* (Nauka, Moscow, 1979; Pergamon, London, 1981).
5. A. Muslimov, H. Van Horn, and M. Wood, Astrophys. J. **442**, 758 (1995).
6. T. Naito and Y. Kojima, Mon. Not. R. Astr. Soc. **266**, 598 (1994).
7. V. Urpin and D. Shalybkov, Astron. Astroph. **294**, 117 (1995).
8. A. Kingsep, K. Chukbar, and V. Yan'kov, Rev. Plasma Phys. **16**, 243 (1990).
9. B. M. Askerov, *Kinetic Effects in Semiconductors* (Nauka, Moscow, 1970).

Translated by Yu. Vishnyakov

Technical Physics, Vol. 45, No. 2, 2000, pp. 153–156. Translated from Zhurnal Tekhnicheskoĭ Fiziki, Vol. 70, No. 2, 2000, pp. 12–15.

Loss of Electrons by Fast H^- Ions in Gases

V. I. Radchenko, D. A. Kozhukhov, and V. N. Kudryavtsev

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

Received August 26, 1998

Abstract—The characteristic angles and cross sections for a $(\bar{1}0) + (\bar{1}1)$ reaction in which H^- atoms with an energy of 0.1–10 000 MeV lose electrons when scattered on C, N, or O atoms were calculated for the relativistic case of the Bethe scattering theory. The cross sections for N_2, O_2, and CO_2 molecular targets were obtained using the additivity rule. The results presented are compared with known experimental and theoretical data.

INTRODUCTION

In practice, particle beam optimization usually means a decrease in the cross size of the beam in some "reference" zone of its path. Hence, the need for lowering its emittance arises. As a rule, the emittance value is determined by equipment-related factors at the stage of beam generation and acceleration, etc. However, if the energy is transported at large distances by high-energy hydrogen atoms, a beam of H^0 atoms is formed by passing initial H^- ions through a gas or a thin-film target. Because of this, the emittance of the H^0 beam includes the principally unavoidable "physical" component, which arises because of the scattering of the particles in the target. Therefore, the minimum emittance and other angular characteristics of the beam that are optimized will be governed by a $(\bar{1}0)$ process of H^- ion neutralization in the target. Also, in beam technologies where the particle charge changes in a target, this process must be optimized in terms of energy, depending on the requirements for the beam in the reference zone. The cross section $\sigma_{\bar{1}0}$ of electron loss by the H^- ions and the cross section σ_{01} of electron loss by the H^0 atoms allow us to evaluate the ion-to-atom conversion efficiency [1], that is, to find the maximum part of the H^- beam energy that can be transformed to the atomic beam energy. From the above, it follows that the study of H^- neutralization is essential for solving the problem of hydrogen atom beam optimization.

The goal of this work is to evaluate the characteristic angles and the cross sections σ_Σ of the process $(\bar{1}0) + (\bar{1}1)$ when one or two electrons are lost by H^- atoms with an energy $E = 10^{-1}$–10^4 MeV that collide with C, N, or O atoms. The calculations were made within the dipole moment approximation (DMA) for a target atom [1].

THEORY

The DMA concept sounds as follows. Assume that any atom of the target is a structureless particle in the first approximation. These particles are the sources of an electric field that can be described by the Thomas–Fermi potential at small distances from the nucleus and by the dipole potential at large distance. We introduce the notion of effective dipole moment of an atom. The mean value of the squared effective dipole moment is represented as the sum of the mean values of the squared radius vectors $\langle r_{nl}^2 \rangle$ of individual electrons occupying the shells with quantum numbers n and l (all values are taken in atomic units). Here, the relation between v_{nl} and v, where v_{nl} is the mean orbital velocity of an electron and v is the impact velocity, as well as the polarization of the electron shell of a target atom in the field of an incident particle, are heuristically included. The differential cross sections of scattering of the particles on the dipole potential and on the Thomas–Fermi potential join together in such a way that small-angle scattering is dipole scattering.

The formulas for the differential cross sections of scattering on these potentials were derived in [1] in the framework of the Bethe theory of scattering. In the DMA, the expression for the differential cross section of dipole scattering,

$$\frac{d\sigma^d_{\alpha_f\alpha_i}(\theta)}{d\Omega} = 4\left(\frac{M_A}{m}\right)^2\left(\frac{d}{ea_0}\right)^2 \times \frac{k_f}{k_i}\left|F^A_{\alpha_f\alpha_i}(\mathbf{q})\right|^2\frac{\cos^2\gamma}{q^2}, \quad (1)$$

is the basic one. Here, θ is the scattering angle in the laboratory frame of reference; m and M_A, the masses of an electron and incident particle A, respectively; e, the unit charge; a_0, the Bohr radius; $\mathbf{k}_i$ and $\mathbf{k}_f$, the wave vectors of particle A before and after impact, respectively; $\mathbf{q} = \mathbf{k}_f - \mathbf{k}_i$; γ, the angle between the vectors $\mathbf{q}$

1063-7842/00/4502-0153$20.00

and **d**; **d**, the vector of the effective dipole moment of a target atom; and $F^A_{\alpha_f\alpha_i}$, the form factor of the incident particle A, the internal state of which changes from α_i to α_f due to the impact.

Characteristic angles and the cross sections of the process of electron loss by H^- ions in the atomic (C, N, and O) targets (angles are given in mrad, and cross sections, in units of 10^{-18} cm^2)

E, MeV	$\theta_{1/2}$	C	N	O
0.1	103	1679	1502	1242
0.12	94.1	1517	1372	1169
0.15	84.1	1320	1207	1059
0.2	72.7	1083	1000	906
0.3	59.3	798	743	696
0.4	51.3	634	593	566
0.5	45.9	527	494	477
0.7	38.7	395	371	364
1	32.4	288	272	269
1.15	30.2	254	240	239
1.67	25.1	180	171	172
2	22.9	152	145	146
3	18.7	104	99.2	100.6
5	14.5	64.0	61.1	62.3
6.9	12.3	47	44.8	45.8
10	10.2	32.8	31.4	32.1
10.4	10.0	31.6	30.2	31
12	9.31	27.5	26.3	27
14.9	8.35	22.4	21.4	21.9
20	7.20	16.9	16.1	16.5
25	6.42	13.6	13	13.4
30	5.86	11.4	11	11.3
40	5.06	8.74	8.37	8.61
50	4.52	7.11	6.81	7.01
70	3.79	5.24	5.02	5.17
100	3.14	3.83	3.68	3.78
200	2.15	2.2	2.11	2.17
300	1.70	1.65	1.59	1.63
500	1.25	1.23	1.18	1.21
700	1.00	1.05	1.01	1.03
800	0.912	0.994	0.953	0.982
1000	0.777	0.920	0.882	0.909
1500	0.567	0.826	0.793	0.817
2000	0.448	0.784	0.752	0.775
3000	0.310	0.746	0.716	0.738
5000	0.185	0.722	0.692	0.714
10000	0.084	0.708	0.680	0.700

The instantaneous dipole moment **d** of an atom in (1) is [1]

$$\frac{\mathbf{d}}{ea_0} = -\sum_{i=1}^{Z_B}\mathbf{r}_i, \tag{2}$$

where Z_B is the atomic number of a target atom and $\mathbf{r}_i$ is the radius vector of electrons in the units of a_0.

Averaging (1) over target atoms necessitates averaging sum (2) squared:

$$\langle d^2\rangle = \int\left(\sum_{i=1}^{Z_B}\mathbf{r}_i\right)^2\left|\Psi_B(r_1, \ldots, r_{Z_B})\right|^2 d\tau_B, \tag{3}$$

where Ψ_B is the wave function of atom B and $d\tau_B$ is an element of the configuration space of Z_B electrons of atom B.

Squaring (2) and taking integral (3), we find

$$\langle d^2\rangle = \sum_{i=1}^{Z_B}\langle r_i^2\rangle = \sum_{n,l}N_{nl}\langle r_{nl}^2\rangle, \tag{4}$$

where N_{nl} is the number of electrons on the subshells with quantum numbers n and l; for the filled subshells, $N_{nl} = (2l + 1)$.

According to expression (4), the correction for the averaging of the dipole moment of an atom over the impact time will be made separately for every orbital electron. We will assume that every electron contributes to the instantaneous dipole moment $\mathbf{d}_{nl} = -e\mathbf{u}_{nl}$, where $\mathbf{u}_{nl} = \langle r_{nl}^2\rangle^{1/2}\mathbf{r}_{nl}/r_{nl}$. Notice that the differential scattering cross section (1) equals zero if the vectors **d** and **q** are perpendicular to each other. It is therefore necessary first to find the impact-time-average projection of the vector $\mathbf{d}_{nl}$ onto the wave vector **q** and then to average the found projection squared over **q** directions in the space. As the magnitude of the vector $\mathbf{d}_{nl}$ remains unchanged, one averages the function $\cos\gamma_{nl}$, where γ_{nl} is the angle between the vectors $\mathbf{d}_{nl}$ and **q**. The impact-time-average value of $\cos\gamma_{nl}$ should be taken with some weighting coefficient, since the instantaneous dipole moment $\mathbf{d}_{nl}$ at the initial, $t \approx 0$, and final, $t \approx \tau$, instants of impact is smaller than at $t \approx \tau/2$, where τ is the impact time. As a weighting coefficient, we used the function

$$p(t) = \frac{2}{\tau}\sin^2\frac{\pi t}{\tau},$$

normalized to unity in the interval from 0 to τ.

Omitting intermediate calculations, we write for (3) the improved version of expression (4):

$$\langle d^2\rangle = \frac{2}{3}\sum_{n,l}N_{nl}\langle r_{nl}^2\rangle\frac{1-\cos\alpha_{nl}}{\alpha_{nl}^2[1-(\alpha_{nl}/2\pi)^2]^2}, \tag{5}$$

where the angle is

$$\alpha_{nl} = 5.262\frac{\langle r_{nl}^2\rangle}{\langle r_{nl}\rangle}\frac{v_{nl}}{v}. \quad (6)$$

When $\alpha_{nl} \longrightarrow 0$, the value of $\langle d^2\rangle$ represents the squared instantaneous dipole moment of an atom averaged over **d** orientations in the space (relative to **q**) and projected onto the vector **q**.

In the field of an H^- ion, the electron shell of a target atom polarizes, shifting (relative to its center of mass) away from the instantaneous position of the ion by a distance ρ. The distance ρ is the averaged sum of the projections of the radius vectors of atomic electrons onto the axis connecting the atom and the ion. Within the DMA, ρ can be interpreted as an increase in the effective distance p over which the colliding particles interact. Let the cross section σ_Σ^p corrected for the polarization be related to the initial cross section σ_Σ as

$$\sigma_\Sigma^p = \sigma_\Sigma\left(1+\frac{\rho}{p}\right)^{-4}, \quad (7)$$

since the scattering cross section is proportional to the matrix element **M** squared (expression (34) from [1]) within the instantaneous DMA. This element, in its turn, is in inverse proportion to the squared distance between the particles (expression (32) from [1]). In [2], ρ/p was estimated as

$$\frac{\rho}{p} = \frac{0.265Z_B}{Idv}\mu(0.0255Id^{1/2}v^{-1}), \quad (8)$$

where

$$\mu(x) = e^{-x}\bar{E}_i(x) - e^{x}E_i(-x), \quad (9)$$

$\bar{E}_i$ and E_i are modified and ordinary exponential integrals, I is the ionization potential of an atom, and $d = \sqrt{3\langle d^2\rangle}$; the numeric coefficients in expression (8) take into account that I is measured in electronvolts and v, in units of 10^9 cm/s.

RESULTS OF CALCULATIONS

With the corrections for electron motion and for the polarization of the electron shell of a target atom in the field of an incident H^- ion, we determined the characteristic angles and the total cross sections for the hydrogen particles scattered on C, N, or O atoms in the process $(\bar{1}0) + (\bar{1}1)$. In this work, by a characteristic angle we mean the value of θ for which the product $\sin\theta d\sigma(\theta)/d\Omega$ is maximal. The mean values of $\langle r_{nl}\rangle$, $\langle r_{nl}^2\rangle$, and v_{nl}, entering into formulas (5) and (6), were obtained using tabulated Rutaan–Hartree–Fock uncorrelated one-electron wave functions [4]. The atomic ionization potentials necessary for the calculation of the corrections to the cross sections for the polarization of a target atom were taken from [5].

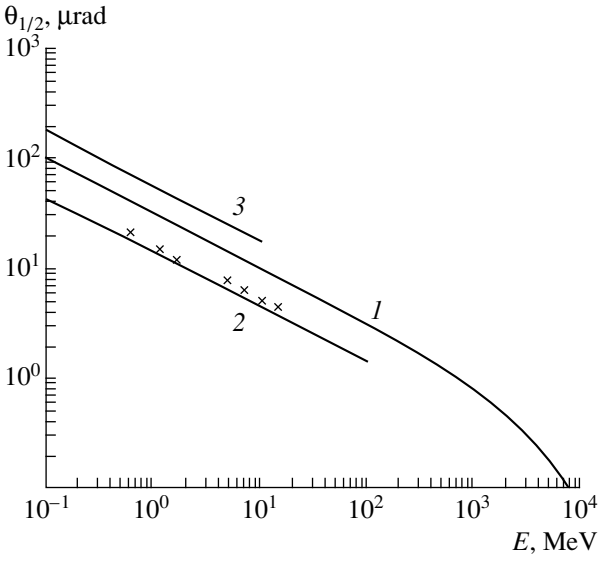


Fig. 1. Energy dependence of the characteristic scattering angles $\theta_{1/2}$ for electron loss by the H^- ions. Theoretical results: (*1*) this work, (*2*) [7], and (*3*) [8]. Data points (×) were taken from [2, 6] for the CO_2 target.

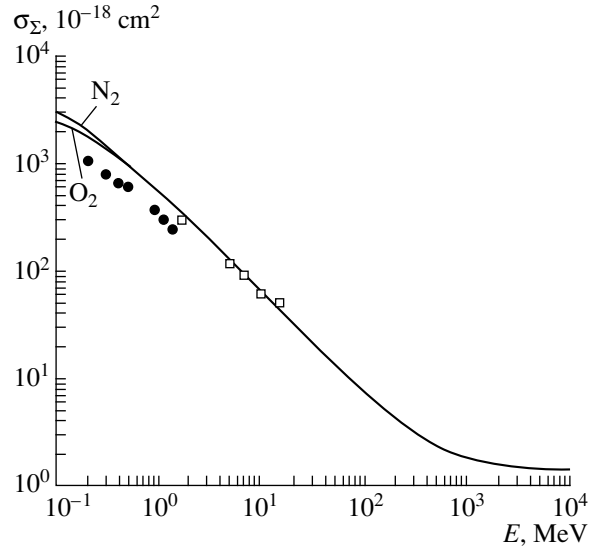


Fig. 2. Energy dependence of the cross sections σ_Σ for electron loss by the H^- ions in the N_2 and O_2 targets. Solid curves represent the results of this work. The cross sections σ_Σ obtained in [9] for nitrogen and oxygen atomic targets (for ion energies E = 1–10 000 MeV) and doubled according to the additivity rule coincide with our results. Data points: (□) O_2 target, energy range E = 1.67–14.9 MeV [10]; (●) N_2 target, E = 0.9–1.3 MeV [11] and 0.2–0.5 MeV [12].

The results of the calculations for the H^- energies in the interval E = 0.1–10 000 MeV are presented in the table. The energy dependences of the specific angles $\theta_{1/2}$ and the cross sections σ_Σ, as well as the comparison

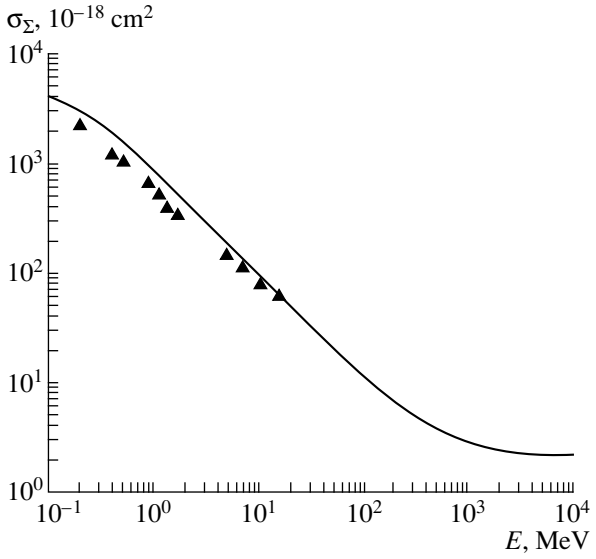


Fig. 3. Energy dependence of the cross section σ_Σ for electron loss by the H^- ions in the CO_2 target. The solid curve, this work; data points (▲) for E = 1.67–14.9 [10], 0.9–1.3 [11], and 0.2–0.5 MeV [12].

of our results with other theoretical and experimental results, are shown in Figs. 1–3. The cross sections for the molecular targets N_2, O_2, and CO_2 were obtained using the additivity rule. The cross sections σ_Σ for the atomic nitrogen and oxygen targets agree with the results of [9].

Thus, we can draw the following conclusions:

(1) The specific angles $\theta_{1/2}$ do not depend on the type of the target and vary as $E^{-1/2}$ in the energy range $E < 100$ MeV. With an increase in E, the energy dependence becomes stronger. For $E \approx 10000$ MeV, the specific angles $\theta_{1/2} \propto E^{-1}$.

(2) The angles $\theta_{1/2}$ calculated in this work are approximately twice as large as those measured in [6]. It seems that the main source of this discrepancy is in the fact that the DMA is a two-particle approximation. The calculation [7] carried out with Born's three-particle approximation gives the values of $\theta_{1/2}$ approximately 2.3 times less than those obtained in the framework of the DMA.

(3) The energy dependences of the cross sections σ_Σ are similar for all of the targets. Their plots are decreasing curves, which asymptotically tend to some limit as $v \longrightarrow c$. For $E > 1$ MeV, the cross sections σ_Σ for all of the atomic targets are practically the same (the discrepancy is below 7%).

(4) The theoretical curves for $\sigma_\Sigma(E)$ fall more rapidly than the experimental ones, crossing the latter in the energy range E = 10–20 MeV [2]. In this range, all of the corrections to the cross section σ_Σ can be neglected with an error below 3%. Therefore, when E = 10 MeV, for which the related measurements are easy to perform, it is appropriate to use the DMA for the systematic construction of the $\sigma_\Sigma(Z_B)$ curves. The DMA is a very convenient approximation, so such a problem will not involve fundamental difficulties.

REFERENCES

1. V. I. Radchenko, Zh. Éksp. Teor. Fiz. **105**, 834 (1994).
2. V. I. Radchenko, Doctoral Dissertation in Mathematical Physics (Yekaterinburg, 1996).
3. M. Inokuti, Rev. Mod. Phys. **43**, 297 (1971).
4. E. Clementi and C. Roetti, At. Data Nucl. Data Tables **14**, 177 (1974).
5. A. A. Radtsig and B. M. Smirnov, *Parameters of Atoms and Atomic Ions: A Reference Book* (Énergoatomisdat, Moscow, 1986).
6. G. D. Ved'manov, Yu. G. Lazarev, and V. I. Radchenko, Zh. Tekh. Fiz. **70**, 221 (2000).
7. J. A. Johnstone, Nucl. Instrum. Methods Phys. Res., Sect. B **52**, 1 (1990).
8. Y. T. Lee and J. C. Y. Chen, Phys. Rev. A **19**, 526 (1979).
9. K. Riesselmann, L. W. Anderson, L. Durand, *et al.*, Phys. Rev. A **43**, 5934 (1991).
10. V. I. Radchenko and G. D. Ved'manov, Zh. Éksp. Teor. Fiz. **107**, 3 (1995) [JETP **80**, 1 (1995)].
11. G. I. Dimov and V. G. Dudnikov, Zh. Tekh. Fiz. **36**, 1239 (1966) [Sov. Phys.–Tech. Phys. **11**, 919 (1966)].
12. I. Kovacs, Nucl. Instrum. Methods Phys. Res. **51**, 224 (1967).

Translated by S. Bulanov

Technical Physics, Vol. 45, No. 2, 2000, pp. 157–160. Translated from Zhurnal Tekhnicheskoĭ Fiziki, Vol. 70, No. 2, 2000, pp. 16–19.

Critical Heat Load upon Pool Liquid Boiling

G. N. Kruzhilin* and E. V. Lykov**

** Institute of Problems of Harmless Development of Atomic Power Engineering, Russian Academy of Sciences, Moscow, 113191 Russia*

*** Rostov State Academy of Agricultural Machine Building, Rostov-on-Don, 344101 Russia*

Received March 30, 1998; in final form, December 1, 1998

Abstract—The theoretical derivation of the formula for calculating heat load, given in the treatise titled "The Theory of Heat and Mass Transfer" (edited by A.I. Leont'ev, Bauman State Technical University, Moscow, 1997) is noted to be incorrect. A corrected deduction of the formula, performed using the dimensional method, is considered. A brief review of the whole problem is given. It is noted that the "second crisis of boiling" is due to the effect of the "spheroidal" state in liquids. © *2000 MAIK "Nauka/Interperiodica".*

Numerous works concerning the important scientific and engineering problem of the critical heat load during boiling of a liquid were published, but, in our opinion, no clear understanding of the problem has been reached up to now. For example, in the treatise [1] on the heat transfer published in 1997 (edited by A.I. Leont'ev), a significant inexactness crept into Section 2.6, Chapt. VIII (written by B.M. Mironov), when deriving Kutateladze's formula concerning the critical heat load upon boiling of a liquid on a flat surface in the case of free convection, i.e., in a pool. The author writes: "Suppose that the critical velocity of vapor that characterizes the boiling crisis depends on the relation between the buoyancy forces, surface tension forces, and the inertia of the vapor flow. Then

$$U_{cr} = f[g, (\rho' - \rho''), \rho'', \sigma]\ldots. \quad (1)$$

In accordance with the π theorem, the process is determined in this case by a single similarity parameter

$$K_1 = \frac{U_{cr}\sqrt{\rho''}}{\sqrt[4]{\sigma g(\rho' - \rho'')}}\ldots. \quad (2)$$

The critical velocity of the vapor phase is

$$U_{cr} = \frac{q_{cr}}{r\rho''g},$$

and, consequently,

$$K_1 = \frac{q_{cr}}{rg\sqrt{\rho''}\sqrt[4]{\sigma g(\rho' - \rho'')}}. \quad \text{(VIII.120)}$$

Experiments on the boiling of nonmetallic heat-transfer agents performed under various conditions confirmed the main conclusions of the hydrodynamic theory of crises. The magnitudes of the parameter K_1, according to these experimental data, lie in the interval of 0.13–0.16."

In the above text, the following designations were assumed: U (m/s), the average normalized velocity of the vapor; ρ', ρ'' (kg s^2/m^4), the densities of the liquid and the vapor, respectively; σ (kg/m), the surface tension of the liquid; g (m/s^2), the gravitational acceleration; q_{cr} (kcal/m^2 s), the critical specific heat load; and r (kcal/m^2 s), the heat of vaporization. Equation (1) relates five physical quantities and contains three dimensions: kg, m, and s. Therefore, according to the π theorem, two dimensionless numbers should correspond to this equation. One of these can contain the quantities U_{K_p}, g, ρ'' and σ; the second one includes g, ρ'', $(\rho' - \rho'')$, and σ. As a result, from the first table of the quantities we obtain, using the theory of dimensions, a dimensionless group

$$K_2 = U_{cr}\left(\frac{\rho''}{g\sigma}\right)^{1/4} \quad (3)$$

and from the second table, a dimensionless group

$$K_3 = \frac{\rho''}{\rho' - \rho''}. \quad (4)$$

Thus, according to the theory of dimensions, equation (1) is associated with two dimensionless groups (3) and (4), rather than a single group B that is given in the textbook [1]. It is expedient to note that this inexactness first appeared long ago, in the first works of Kutateladze concerning this problem. As is well-known, formula (VIII.120) was for the first time suggested by Kutateladze as long ago as in 1951, and since that time it received rather wide acceptance. The parameter K_1 itself was obtained from the set of Navier–Stokes equations for the liquid phase and also for the vapor for the case of equal velocities of the vapor and the liquid at the interface, as well as of pressures with allowance for the surface tension [2]. The position of the boundary between the liquid and the vapor upon boiling cannot be specified mathematically rigorously because of the

1063-7842/00/4502-0157$20.00 © 2000 MAIK "Nauka/Interperiodica"

chaotic motion of vapor bubbles and, correspondingly, local volumes of the liquid. Therefore, this set of equations cannot be closed, and the set of dimensionless numbers that is obtained from this set of equations cannot be full. It seems that it is precisely for this reason that neither Kutateladze himself nor anyone else has subsequently used this method for the derivation of formula (VIII.120). Instead, in his monograph [3], Kutateladze develops this formula from elementary dimensional considerations. We cite:

> Let us consider the conditions for the stability of a two-phase incompressible boundary layer over an unbounded horizontal porous surface, assuming that the liquid has a vanishing viscosity. In this case, the factors that can interact are only the kinetic energy of the injected gas and the gravitational and surface forces in the two-phase boundary layer. On the order of magnitude, the dynamic pressure of the gas is $\rho'' U^2$, and the work done to displace the liquid from the created gas 'bell jar' (gaseous or vapor film) is $g\delta(\rho' - \rho'')$. Here, δ is the average thickness of the created gas layer. On the order of magnitude, the value of δ is determined by the Laplace constant $\sqrt{\sigma/g(\rho' - \rho'')}$, which directly follows from the dimensional considerations. Since, in the model at hand, the development of crisis (loss of stability of the nucleate structure of the two-phase boundary layer) is equally probable at any site of the extended porous surface, then the ratio of the quantities considered must be a number, i.e.,
>
> $$\frac{\rho'' U_{cr}^2}{\sqrt{\sigma g(\rho' - \rho'')}} = \text{const}.$$
>
> Here, U_{cr} has the sense of the reduced critical rate of bubbling, i.e., critical volume gas flow through the unit surface of the plate. Taking (for convenience) the square root of this expression, we obtain the following stability criterion:
>
> $$K_1 = \frac{U_{cr}\sqrt{\rho''}}{\sqrt[4]{\sigma g(\rho' - \rho'')}}.$$
>
> For the above-assumed conditions ($\mu \longrightarrow 0$, horizontal unbounded plate, unbounded volume of quiescent incompressible liquid), we have $K = \text{const} = C$.
>
> Upon boiling, $U_{cr} = q_{cr}/(r\rho'' g)$; therefore, $q_{cr} = Cr g \sqrt{\rho''}\sqrt[4]{\sigma g(\rho'' - \rho')}$. According to experimental data, $C = 0.14$." The derivation given in [3] is not rigorous. It is quite proofless to assume that in the case at hand the main factor is the ratio of $\rho'' U^2$ to the work "spent on the displacement of the liquid from a layer of thickness δ close to the diameter of vapor bubbles that are formed at the surface.

It is undoubtedly that in reality the major part of the work done by the gas acting on the liquid is converted into the kinetic energy of the liquid and its division into fine droplets, but the magnitude of this part cannot be estimated based on the dimension theory. On the other hand, this work is being done not at the expense of the kinetic energy of the gas phase (since, e.g., upon pure barbotage, the kinetic energy of the gas remains unaltered during its motion), but rather at the expense of its potential energy (with the pressure decreasing by ΔP). One also can only be fogged by the argument on the "stability of the nucleate structure of the two-phase boundary layer," since the motion of a two-phase vapor–water flow is chaotic in its nature, as was noted above. For the same reason, we cannot even speak of the "boundary layer" in this case.

Nevertheless, the concept of stability was applied to boiling in [4]. By that time (1959), it was known that the previously existed concept of the occurrence of the boiling crisis due to the coalescence of vapor bubbles at the heating surface has not been confirmed by experimental data on the boiling of water and organic liquids. Instead, a concept was developed that the boiling crisis occurs because of the pressing liquid away from the surface by the flow of the gas formed. For the first time, this concept was substantiated in [5]. Its validity is clearly illustrated by the fact that the critical heat load upon water boiling at a horizontal tube was shown experimentally [6] to decrease from 0.960106 to 0.308106 kcal/(m^2 h) (i.e., by a factor of 3.1) with pressure decreasing from 1.0 to 0.09 atm. However, it is known that with decreasing pressure the limiting radius of the operating centers of vapor formation increases and, correspondingly, the number of the centers decreases. For this reason, the coalescence of nucleating gas bubbles and the formation of a continuous vapor film at the heating surface with decreasing pressure had to occur at a correspondingly greater critical heat load. In reality, the opposite takes place for the reason that a decrease in pressure leads to a sharp decrease in the density of vapor, and consequently, at a specified heat load, the volume of the vapor formed and the rate of its pressure increase, leading to the pressing of the liquid away from the surface, thereby decreasing the flow of the liquid toward the surface and, consequently, deteriorating liquid cooling and preventing the occurrence of the heat transfer crisis with increasing wall temperature.

In the monograph of Zuber [4], the interaction between streams (layers) of vapor moving upward from the heating surface and liquid moving downward toward this surface are considered. Based on mathematical calculations, a conclusion was made that the stability of such stratified motion of liquid is disturbed under the condition

$$K_1 = \sqrt{\frac{\rho'}{\rho' + \rho''}} = \text{const}. \qquad (5)$$

Zuber [4] refers to Rayleigh, who in his time mathematically studied the stability of combined horizontal stratified flow of two liquids—a heavier one below and a lighter one above—and received a similar result.

As is well-known, the problem of the transition from a laminar, i.e., stratified, motion in a tube into a turbulent motion is also studied as a problem of stability. However, it is, naturally, impossible to speak of the stability of the turbulent motion of a liquid. Still less we may speak of a two-phase gas–liquid boundary layer, since in this case the boundaries of the phases continuously change in shape and in space.

To generalize, it is clearly impossible to teach students the K_1 criterion after Kutateladze. Therefore, the idea of V.M. Mironov to obtain formula VIII.120 using the π theorem of the method of dimensions may only be welcomed. To fully realize this idea, we again turn to equation (1). The equation expresses the dependence between the parameters of the process of motion of a gas–water mixture. The equation is correct. Nevertheless, as we told above, it does nor permit one to obtain the K_1 criterion because of the existence of five physical quantities in it. However, this difficulty can be overcome, since in this case we can refine the formulation of the problem so as to reduce the number of determining physical quantities to four and thus obtain the sought-for criterion. We can do this because the quantities U and ρ'' affect the process of interaction between the vapor and the liquid only in a combination, as a product ρU^2, i.e., in the form of the flux of momentum of the vapor, where $\rho'' U^2$ [kg/(m^2 s)] is the flux of mass, since the quantity $g''U^2$ enters into the equation of motion of the liquid [7]. Thus, instead of the table of five physical quantities, we should consider the table of only four quantities:

$$\rho'' U^2, \quad g, \quad (\rho' - \rho''), \quad \text{and} \quad \sigma.$$

With this table, the theory of dimensions yields only one number K_1 in the form (2). According to the theory of similarity, this means that the numerical value of this number determines all hydrodynamic characteristics of pool boiling, such as the velocities of the liquid and vapor in the various points of the volume; the average vapor content in the various planes and directions in the volume; the inflow of the liquid to the heating surface G [kg/(m^2 h)], which undoubtedly decreases as the vapor velocity U increases because of pressing away of the liquid from the heating surface by the vapor formed at the surface. It is this process of pressing away of the liquid by the vapor that causes the beginning of the boiling crisis. For this reason, the critical heat load corresponds to a quite certain value

$$G_{cr} = \frac{q_{cr}}{r},$$

and the K_1 number also has a quite certain value $(K_1)_{cr} = C$. The vapor velocity U_{cr} in the criterion K_1 transforms into the ratio $q_{cr}/(r\rho''g)$ and, as a result, this criterion takes on the form of formula (VIII.120).

Here, it is expedient to note that formula (2) is more correct from the theoretical viewpoint than the analogous formula (5), since the latter was obtained on the assumption of the interaction of the stratified flows of the vapor and liquid upon pool boiling, which is very far from truth, whereas formula (2) is derived by the method of dimensions strictly in the framework of mathematical physics without any speculative arguments. Nevertheless, in our opinion, the above-mentioned work [4] remains in its content the most important one from the standpoint of both the extensive sweep of the available (at that time) publications on the problem and the fullness of the treatment. For the history of the problem, of interest is the fact (noted by Zuber [4]) that the criterion K_1 in the form (2) was for the first time published in 1941 [8], i.e., ten years before this was done by Kutateladze.

It is well known from experiments that when a liquid is underheated, i.e., when its temperature t_l is below the saturation temperature t_s, the critical heat load upon boiling is higher and increases with increasing the underheating $\Delta t_n = t_s - t_l$. The reason is first in that, upon underheating, the heat consumption spent on the vaporization of 1 kg of the liquid increases, and the effective heat of vaporization becomes, instead of r,

$$r_{eff} = r + C\Delta t_n,$$

where c is the heat capacity of the liquid, or

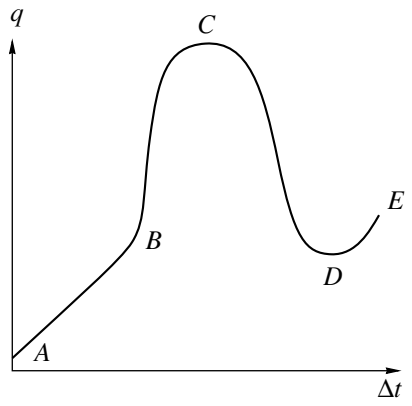
$$r_{eff} = i'' - i',$$

where i is the heat content.

For this reason, with the replacement of r by r_{eff}, a greater quantity q_{cr} corresponds to formula VIII.120. But in reality, the experimental data exhibit even greater increase in q_{cr}, since the "subcooled liquid" takes correspondingly larger amount of heat from the heating surface and then transfers it into the volume of the boiling liquid. As a result, the vapor distribution over the volume changes, vapor formation near the heating surface becomes relatively smaller, and, correspondingly, the effect of pressing liquid by the vapor away from the heating surface becomes decreased. However, it is difficult to estimate this second effect by calculation.

The forced circulation of the liquid increases the critical heat load. This refers to the pool boiling and the more so to the motion of a boiling liquid in a tube. In this case, a forced inflow of the liquid to the heating surface is ensured, for which reason the critical "pressing away of the liquid" from the surface is reached at correspondingly more intense vapor formation, i.e., higher critical heat load. There are known experiments when at a small horizontal heating surface cooled by a water jet, a critical load was reached that many times exceeded that obtained upon pool boiling.

Special heat transfer conditions are realized upon boiling of a liquid with a large vapor content in a tube in the regime of disperse annular flow, when a film of liquid moves at the wall of the tube and the bulk stream



Variation of the specific heat load q upon boiling of a liquid as a function of the temperature difference Δt ($t = t_s - t_{sat}$).

("core of the flow") contains vapor with fine liquid droplets. Such conditions are realized, in particular, in the cooling channels of boiling-water nuclear reactors. Under these conditions, a reliable cooling of the wall is only reached until a liquid film is retained at the wall, i.e., until a turbulent transport of liquid droplets from the bulk stream to the wall compensates the evaporation of the liquid and breaking liquid droplets away from the wall. This very complex process, especially in the presence of a flow of vapor generated at the wall and moving from the wall, cannot be quantitatively calculated theoretically and therefore is determined only on the basis of a vast body of available experimental data.

Note also that in both [1] and [3], the so-called second crisis of heat transfer is discussed. Actually, this second crisis is merely a myth. It is clear even from the fact that this crisis is absent upon barbotage, when the hydrodynamics of the process is similar to that upon pool boiling of a liquid. To be concrete, we turn to a vessel inside which a horizontal plate with orifices 5–7 mm in diameter is mounted; let a layer of liquid exist above it, through which a gas is injected, passing upward through the orifices in the plate. If the gas rate is small, streams of the liquid will flow through the same orifices downward. Such a scheme is used in the apparatus for the purification of gases from impurities, with the difference that, instead of plates with orifices, layers of wire networks, balls, or even small hollow cylinders are usually employed. It is natural that with increasing gas flow rate, the downward flow of liquid decreases and finally ceases entirely; i.e., a crisis occurs, which the technologists call flooding. Correspondingly, the gas rate G (kg/s) and its reduced velocity U (m/s) are called critical in this case. Naturally, when the gas rate decreases and U becomes equal to U_{cr}, the downward liquid flow is restored; i.e., the crisis ceases. In this case, thus, no second crisis takes place and cannot take place.

Now, let us turn to the figure, where a curve of water boiling is shown at the atmospheric pressure in the form of the dependence of the specific heat load on the temperature difference. The figure shows that up to point B the heat transfer from the heating surface to the liquid occurs by natural (free) convection, i.e., without boiling. Then, the region of nucleate boiling begins, until a critical heat load is reached at point C [of about 106 kcal/(m h)] at a critical temperature difference $t = t_s - t_{sat} = 25°C$ and the heat-transfer coefficient $\alpha = 40.102$ kcal/(m^2 h K). After the crisis begins, a transient process occurs until the point D is reached, where stable film boiling starts with a specific heat load of only about 33.103 kcal/(m^2 h) [4] and a very large temperature difference $t = t_s - t_{sat}$ reaching several hundred Kelvins, which may be accompanied by even an overburning (melting) of the heating surface. It is for this reason, i.e., because of the very high temperature of the wall after reaching point D, that the restoration into the precritical boiling regime is impossible in a boiling crisis even in spite of the above-mentioned very low specific heat load. The point is that under these conditions, upon the breakthrough of the liquid through the film to the hot wall, an instantaneous partial vaporization of liquid occurs, as well as its repulsion from the wall, similar to a burst. However, in this case, no wetting of the heating surface by the liquid occurs, which might ensure its intense cooling, because of the so-called spheroidal state (Leidenfrost effect), which for water at the atmospheric pressure manifests itself beginning from about 150°C. This process, thus, is mainly of thermodynamic origin. By the way, technologists call this process "precooling," because they deal with it upon the filling of the system with the cooling agent under the conditions of the initial, rather large temperature difference $\Delta t = t_s - t_{sat}$. Although this is a separate problem, we note here that the precooling occurs the faster, the greater the flow velocity and the smaller the initial temperature difference $\Delta t = t_s - t_{sat}$.

REFERENCES

1. *The Theory of Heat and Mass Transfer*, Ed. by A. I. Leont'ev (Izd. MGTU im. N. É. Baumana, Moscow, 1997), pp. 550–551.
2. S. S. Kutateladze, Izv. Akad. Nauk SSSR **4**, 529 (1951).
3. S. S. Kutateladze, *Introduction to the Theory of Heat Transfer* (Atomizdat, Moscow, 1979), pp. 305–306.
4. N. Zuber, *Hydrodynamic Aspects of Boiling Heat Transfer* (Univ. of California, Los Angeles, 1959).
5. G. N. Kruzhilin, Izv. Akad. Nauk SSSR **5**, 701 (1949).
6. S. M. Braunlich, Thesis in Chem Eng. (Mass. Inst. of Technology, 1941), pp. 305–319.
7. L. G. Loĭtsyanskiĭ, *Mechanics of Liquids and Gases* (Gosizdat, Moscow, 1950), p. 141.
8. C. Bonilla and C. W. Perry, Trans. Am. Inst. Chem. Eng. **37**, 685 (1941).

Translated by S.N. Gorin

Technical Physics, Vol. 45, No. 2, 2000, pp. 161–167. Translated from Zhurnal Tekhnicheskoĭ Fiziki, Vol. 70, No. 2, 2000, pp. 20–27.

GASES AND FLUIDS

Transport Damping by Fluctuations and Cluster-Type Expansions in Kinetics

S. N. Gordienko and S. S. Moiseev

Institute of Space Research, Russian Academy of Sciences, ul. Profsoyuznaya 84/32, GSP-7, Moscow, 117810 Russia

Received June 30, 1998

Abstract—A mechanism of transport damping by fluctuations is proposed and analyzed in detail for the thermal diffusivity of a one-dimensional gas in the case when the externally induced temperature gradient is weaker than fluctuating gradients. Thermal diffusivity is calculated as a function of parameters of a one-dimensional gas characterized by a homogeneous potential of interaction between particles: for the potential $V(x) = Q/|x|^k$, thermal diffusivity is $\chi \sim (Qn^k/T)^{1/(k-1)}\nu_T/n$. The nonanalytic form of thermal diffusivity as a function of concentration is elucidated. The nonanalytic behavior of thermal diffusivity as a function of concentration is explained by a nontrivial symmetry of the problem under analysis at long times, after the initial fluctuations have dissipated. It is shown that spontaneous generation of macroscopic structures developing through selective growth of fluctuations of a certain type in a medium with an externally induced temperature gradient controls the transport properties of the medium when $k \approx 1$. © *2000 MAIK "Nauka/Interperiodica".*

INTRODUCTION

In recent years, self-organization processes in turbulence have been investigated with increasing interest. Helical turbulence is of special importance in this perspective. Helicity leads to flow quasi-laminarization, because eddies tend to broaden, and the so-called inverse cascade develops. However, these eddies are characterized by lower viscosity. Therefore, their lifetime is longer, and their contribution to the turbulent ensemble plays a dominant role. As a consequence, the average turbulent viscosity is reduced together with the Reynolds stresses. These trends have been observed experimentally in turbulent flows (e.g., see [1, 2]). The decrease in turbulent viscosity due to helicity was also obtained theoretically [3, 4]. In other words, the role of nonlinear effects tends to decrease in helical turbulence. One may expect that the domain of flow regimes characterized by low effective Reynolds numbers (Stokes creeping flows) will tend to expand. It remains unclear whether helicity is the only characteristic responsible for quasi-laminarization. In this paper, we show that the answer is negative. When fluctuations of parameters of a nonequilibrium state of the fluid are characterized by alternating signs, the flow tends to partially laminarize. Physically, this occurs when the fluctuation intensity in a system of this type is sufficiently high, and, therefore, relaxation effectively corresponds to a lower value of the nonequilibrium parameter. In particular, when temperature gradient fluctuates in a medium characterized by unstable convection, the average anomalous thermal diffusivity decreases because the direction of heat flux is not uniquely defined. The corresponding instantaneous dynamics are substantially more complicated. For example, when symmetric fluctuations take place in a one-dimensional flow, the following processes develop. Fluctuations that amplify the initial gradient result in a faster heat transfer. In other words, the local gradients that stimulate heat transfer are rapidly smoothed out while the remaining gradients correspond to a slower heat transfer. The influence of fluctuations on the thermal diffusivity χ is determined by the competition of the two processes. In the analysis below, we elucidate the conditions under which χ decreases.

In this paper, we consider a simple model of transport damping by fluctuations and substantiate the reasoning given above by a rigorous analysis, which provides guidelines for constructing a consistent theory of phenomena of the class discussed here. Moreover, the problem analyzed below is of interest in itself.

THE ROLE OF FLUCTUATIONS IN TRANSPORT PHENOMENA

It is well known that the Boltzmann-type kinetic equations used to describe transport processes in gases and plasmas do not allow for fluctuations, as shown in [5, 6]. Therefore, they can adequately describe transport phenomena only when the externally induced gradients of flow variables are relatively steep. This can be illustrated by considering thermal diffusivity as an example. Fluctuations taking place in a gas are characterized by a temperature gradient $|\nabla T|_{\text{fl}}$ whose magnitude is determined by the law of interaction between particles and depends on the particular interaction model considered. Accordingly, there can exist two essentially different regimes of heat diffusion depending on the average temperature gradient $|\nabla T|$ across the

1063-7842/00/4502-0161$20.00

gas, one corresponding to $\nabla T_{\text{fl}} \ll |\nabla T|$ and the other taking place when $|\nabla T|_{\text{fl}} \geq |\nabla T|$. In the former case, temperature gradients associated with fluctuations are negligible as compared to the average temperature gradient; i.e., the "direction" of the system's evolution toward thermodynamic equilibrium is well defined at any point. It is the regime described by Boltzmann-type equations and characterized by monotonic growth of entropy (by the *H*-theorem) at every point, i.e., by a uniquely defined by "direction" of relaxation. In the latter regime ($|\nabla T|_{\text{fl}} \geq |\nabla T|$), the "direction" of relaxation toward equilibrium is difficult to define locally for the medium; i.e., the local "direction" of heat flux may vary, which essentially complicates the physics of relaxation.

The hypothesis that transport processes can be damped by fluctuations is based on a quite general qualitative analysis. This effect appears to be particularly important for studies of relaxation in a turbulent fluid (in this case, one should obviously consider relaxation toward a steady turbulent state rather than a thermodynamically equilibrium state). Indeed, the characteristic difference v_λ in velocity between points in a fluid separated by a distance λ is $v_\lambda \sim \lambda^q$, where $0 < q < 1$ (some examples of flows with various q were given in [3]). Accordingly, the "characteristic" gradient associated with turbulent fluctuations with length scales on the order of λ is estimated as $dv_\lambda/d\lambda \sim \lambda^{q-1}$, infinitely increasing toward smaller λ. Thus, there must always exist a range of length scales in which one should expect a substantial influence of fluctuations on momentum transfer. Moreover, an increase in fluctuation intensity associated with helicity is known to lower turbulent viscosity [3, 4].

IRREVERSIBLE EVOLUTION EQUATION ALLOWING FOR FLUCTUATIONS

The analysis of kinetic phenomena in rarefied gases is based on the Boltzmann equation. Unfortunately, this approach cannot be applied to transport phenomena in systems of classical particles that execute a one-dimensional motion, i.e., are described by the Hamiltonian

$$H = \sum_{i=1}^{i=N} \frac{p_i^2}{2m} + \sum_{1 \leq i < j \leq N} V(x_i - x_j), \quad (1)$$

where $V(x_i - x_j)$ is the potential of interaction between the ith and jth particles with coordinates x_i and x_j, respectively.

Indeed, the momentum and energy conservation laws uniquely determine the particle velocities after a binary collision, which obviously eliminates the contribution of binary collisions to the collision integral. Therefore, to obtain a kinetic equation for 1D classical particles, it is not sufficient to allow only for binary collisions. One must also take into account collisions involving a greater number of particles. On the other hand, it is well known that allowance for three-body or higher order collisions generally leads to certain difficulties associated with divergences similar to those considered in [7]. However, since now there exists a method for eliminating this divergences (at least, in the lowest-order approximations), we do not expand on the details of the method here. To estimate thermal diffusivity $|\nabla T|_{\text{fl}} \ll |\nabla T|$, we invoke an approach based on the relation between a transport coefficient and the mean free path. In addition, we derive a kinetic equation that describes relaxation in the fluctuation-dominated case $|\nabla T|_{\text{fl}} \geq |\nabla T|$ by applying the method developed in [8, 9].

Differentiating the equations of motion entailed by (1) with respect to time, we obtain

$$\dot{x}_i = v_i, \quad \dot{v}_i = a_i, \quad \dot{a}_i = \sum_{j=1,\, j \neq i}^{j=N} U_{ij}, \quad i = \overline{1, N}, \quad (2)$$

where the dot denotes a time derivative and

$$U_{ij} = -\frac{1}{m} \frac{\partial^2 V(x_i - x_j)}{\partial x_i^2} (v_i - v_j), \quad i \neq j.$$

We define the following "microcanonical" distribution functions:

$$f_1(t, 1) = \sum_{i=1}^{i=N} \delta_i(1), \quad (3)$$

$$f_2(t, 1, 2) = \sum_{1 \leq i \leq j \leq N} (\delta_i(1)\delta_j(2) + \delta_i(2)\delta_j(1)), \quad (4)$$

where

$$\delta_i(1) = \delta(v_1 - v_i(t))\delta(x_1 - x_i(t))\delta(a_1 - a_i(t)),$$

$$\delta_i(2) = \delta(v_2 - v_i(t))\delta(x_2 - x_i(t))\delta(a_2 - a_i(t)).$$

The "microcanonical" distribution functions thus defined obey the equations (see explanations below)

$$\left(\frac{\partial}{\partial t} + \hat{L}_1\right) f_1(t, 1) = d(t, 1), \quad (5)$$

$$\left(\frac{\partial}{\partial t} + \hat{L}_1 + \hat{L}_2\right) f_2(t, 1, 2) = d(t, 1) f_1(t, 2) + d(t, 2) f_1(t, 1), \quad (6)$$

where

$$\hat{L}_m = v_m \frac{\partial}{\partial x_m} + a_m \frac{\partial}{\partial v_m}, \quad m = \overline{1, 2};$$

$$d(t, 1) = \int \hat{q}(1, 3) f_2(t, 1, 3) d3,$$

$$\hat{q}(1,3) = \frac{1}{m}\frac{\partial^2 V(x_1 - x_3)}{\partial x_1^2}(v_1 - v_3)\frac{\partial}{\partial a_1}.$$

For the "microcanonical" distribution functions, equations (5) and (6) are the exact equations of evolution (note also that the function f_2 should be treated with certain accuracy at $x_1 = x_2$ [9]). In conventional analyses, ensemble-averaged functions, rather than "microcanonical" distribution functions (3) and (4), are considered from the start. This immediately leads to the Born–Bogolyubov–Green–Kirkwood–Yvon (BBGKY) hierarchy. The approach employed here admits a similar treatment (the additional argument introduced into the distribution function does not preclude this treatment), but a different method is better suited for the objectives of the present analysis. We particularly point to the fact that the representation of the right-hand side in equation (6) is not the only possible one in dealing with "microcanonical" distribution functions. The factorization employed here is characterized by the absence of a term associated with binary interaction between particles located at points x_1 and x_2 on the left-hand side of equation (6).

Note that the formulation of the problem of kinetic theory is drastically changed by introducing distribution functions with an additional argument (acceleration). Indeed, the conventional approach is focused on description of interactions between particles in terms of a generalized distribution function. Introducing an additional argument, we obtain an inverse problem: assuming that an arbitrary spatial distribution of forces acting on the particles of a medium is prescribed, one calculates the particle distribution function that corresponds to the distribution of forces (accelerations). Since the admissible distribution of accelerations is not subject to any constraints, except for those directly entailed by the equations of motion, this approach can be applied to analyze fluctuations on an arbitrary scale.

Let us now derive a closed equation describing the evolution of the function f_2 in the long-time limit ($t \longrightarrow +\infty$).

By (5), we have

$$f_1(t,1) = e^{-t\hat{L}_1} f_1(t=0,1) + \int_0^t e^{-\tau\hat{L}_1} d(t-\tau,1)\,d\tau. \quad (7)$$

Note that $\lim_{v_1 \to +\infty} f_1 = 0$, which implies that

$$\lim_{\tau\to\infty} e^{-\tau\hat{L}_1} f_1(t=0,1) = 0 \quad \text{at} \quad a_1 \neq 0. \quad (8)$$

Relation (8) is extremely important. It is not analogous to any equation for a distribution function with the normal number of arguments (velocity and coordinate) and follows from the fact that the operator $e^{-\tau\hat{L}_1}$ indefinitely increases the velocity v_1 for $a_1 \neq 0$ as $\tau \longrightarrow \infty$. The limit in (8) means that the Poincaré cycles are eliminated, which is quite obvious in the case of a single particle driven by a constant force in a box with absolutely elastic walls. Formally, the additional argument was introduced into the distribution function intentionally to ensure that relation (8) should hold. Combining (7) with (8), we obtain

$$f_1(t,1) = \int_0^{+\infty} e^{-\tau\hat{L}_1} d(t-\tau,1)\,d\tau \quad \text{for} \quad t \gg \tau^* = \frac{v_{ch}}{a_{ch}}, \quad (9)$$

where v_{ch} and a_{ch} are the characteristic velocity and acceleration of particles, respectively.

Note that the differential equations in (2), which were obtained by raising the order of the equations of motion, admit solutions that do not correspond to any physical motion of particles obeying the second Newton law. Indeed, the solutions to the equations of motion depend on $2N$ arbitrary constants in the one-dimensional case while solutions to (2) depend on $3N$ arbitrary constants. In the course of further analysis, we will discuss the unphysical solutions to (2). Here, we only note that unphysical solutions to (2) involve a v_{ch} that increases with time elapsed; that is, relations (8) and (9) are satisfied only within the manifold of physical solutions and in its neighborhood where v_{ch} is a slowly increasing function of time (say, a function that grows slower than $a_{ch}t$).

The passage to the limit as $t \longrightarrow +\infty$, which resulted in (9), is nonuniform and makes the theory irreversible in time. This is manifested by the fact that the single-particle distribution function is expressed in terms of a local operator applied to a two-particle distribution function. The physical meaning of the resulting irreversibility is clarified by the inequality in (9): the system becomes irreversible in time after the particles have interacted. In other words, the irreversibility reflects the difference between the "precollision" and "postcollision" correlations in a dynamical system. Thus, expression (8) is merely a formal representation of the physical fact that the system "forgets" initial correlations and "dynamic" correlations develop as a result of an interaction between particles.

Substituting (9) into (6), we obtain a closed time-irreversible equation that describes the evolution of the "microcanonical" two-particle distribution function f_2 at long times t:

$$\left(\frac{\partial}{\partial t} + \hat{L}_1 + \hat{L}_2\right) f_2(t,1,2) = d(t,1)\int_0^{+\infty} e^{-\tau\hat{L}_2} d(t-\tau,2)\,d\tau + d(t,2)\int_0^{+\infty} e^{-\tau\hat{L}_1} d(t-\tau,1)\,d\tau. \quad (10)$$

The set of differential equations in (2), which were obtained by differentiating the equations of motion, admits solutions that do not correspond to any physical

motion of particles. Indeed, whereas the solutions to the equations of motion depend on $2N$ arbitrary constants, the solutions to (2) depend on $3N$ arbitrary constants. Therefore, the function f_2 must satisfy some additional conditions eliminating the "microcanonical" distribution functions that do not correspond to any physical motion. We formulate these conditions here without derivation (see a detailed discussion in [8, 9]), because they are not required for the analysis that follows:

$$a_1 \int_0^{+\infty} e^{-\tau \hat{L}_1} d(t-\tau, 1)\, d\tau = -\frac{1}{m}\int \frac{\partial V(x_1 - x_2)}{\partial x_1} f_2(t, 1, 2)\, d2. \quad (11)$$

Relation (11), being valid at the initial moment of time, remains valid as time elapses; in essence, it is a constraint imposed on the choice of initial conditions.

Note the following detail characteristic of the formalism developed here: when we raise the order of the equations of motion, i.e., analyze the set of differential equations (2), we admit the existence of unphysical motions of particles. Using (8), (9), and (11), we return to the set of physical solutions, but this is done after the particles have "essentially" interacted.

KINETIC EQUATION FOR SMOOTHED DISTRIBUTION FUNCTIONS

To obtain results concerning thermal diffusivity, we need to know only the symmetries of equation (10). Therefore, the results presented in this section are not strictly necessary for the subsequent analysis, but they are important for understanding the physics that underlies the formalism developed here.

Let us now change from "microcanonical" distribution functions to smoothed ones by averaging (10) and (11) over an ensemble. Linear in f_2 terms are averaged in a straightforward manner. The only difficulty is associated with the right-hand side of equation (10), which contains products of two-time two-particle distribution functions, $f_2(t', 1, 2) f_2(t, 3, 4)$. When the gas temperature T is high as compared to the energy V_{ch} of interaction between particles separated by a distance characteristic of a one-dimensional gas, determined by the number $n = N/L$ of particles per unit length, where N is the total number of particles and L is the total length of the system, the following approximation is warranted:

$$\overline{f_2(t', 1, 2) f_2(t, 3, 4)} \longrightarrow \overline{f_2(t', 1, 2)}\,\overline{f_2(t, 3, 4)}, \quad (12)$$

where a bar denotes an ensemble average.

However, a replacement of the type of (12) can be used on the right-hand side of (10) for arbitrary T and V_{ch}. Indeed, we may treat it as a definition of an ensemble of systems described by a kinetic equation. Note that this definition is consistent with the system's dynamics. Note also another essential distinction between the resulting equation and the standard BBGKY hierarchy: the distribution functions governed by the hierarchy cannot be factored not because of the integral terms, which are identical to the integral terms in (10). The reason lies in the term on the left-hand side that describes direct interaction between particles located at points x_1 and x_2. In the proposed approach, this term does not appear because (8) allows one to average products of two-time "microcanonical" distribution functions (cf., for example, [10]). Let us show how the formalism developed here eliminates the terms that preclude factorization. According to [10], we can write

$$\left\langle \sum_i \delta(x - x_i(t))\delta(x' - x_i(t)) \right\rangle = \delta(x - x') \left\langle \sum_i \delta(x - x_i(t)) \right\rangle,$$

and the factorization becomes impossible when this expression defines a singular distribution. In the present formalism, terms that preclude factorization could be contained only in a singular part of the expression

$$\frac{\partial}{\partial a} \lim_{\varepsilon \to +0} \int_0^{\varepsilon} \left\langle \sum_i \delta(a - a_i(t) + \dot{a}_i(t)\tau)\delta(a' - a_i(t)) \right\rangle d\tau = \lim_{\varepsilon \to +0} \frac{\partial}{\partial a} \int_0^{\varepsilon} \left\langle \sum_i \delta(a - a' + \dot{a}_i\tau)\delta(a' - a_i(t)) \right\rangle d\tau,$$

which vanishes as $\varepsilon \longrightarrow +0$. Therefore, terms analogous to those describing the direct interaction between particles located at x_1 and x_2 in the conventional BBGKY hierarchy are not obtained by averaging the product of singular functions on the right-hand side of (10) at times satisfying the inequality in (9).

We have already noted that a kinetic theory involving conventional distribution functions does not admit any local operator that expresses the single-particle distribution function in terms of the two-particle one and much less any possibility of representing a higher order distribution function as a functional of the two-particle distribution function. On the other hand, in his lectures on the Boltzmann equation [11], Uhlenbeck hypothesized that all distribution functions can be expressed as functionals of the two-particle distribution function. Uhlenbeck's hypothesis is true as applied to the distribution functions with additional arguments considered in this paper. Expression (9) provides an example of such a functional for the single-particle distribution function. Here, we do not consider similar functionals for higher order distribution functions.

Note also that the irreversibility of the type considered here is merely a property of a dynamical system

that manifests itself as $t \longrightarrow +\infty$. Thus, the introduction of irreversibility and the analysis of a system in terms of smoothed distribution functions can be performed independently.

ONE-DIMENSIONAL GAS WITH A POWER-LAW INTERACTION POTENTIAL

Let us now consider a one-dimensional gas characterized by a power law of interaction between particles:

$$V(x) = \frac{Q}{|x|^k} \text{ where } k > 0 \text{ and } Q > 0. \quad (13)$$

The constraints for k and Q ensure the existence of the kinetic limit as $N \longrightarrow +\infty$ and $L \longrightarrow +\infty$ while $n = N/L$ = const. It should also be noted that the limit is approached in an essentially different manner when $0 < k \le 1$ and $1 < k$. When $1 < k$, the total potential energy of the system increases as the kinetic limit is approached as a linear function of N and the forces acting on the particles are also well defined. When $0 < k \le 1$, the total potential energy increases as N, i.e., more rapidly than any linear fiction of N^{2-k}, but the forces acting on the particles remain well defined in the kinetic limit because the following improper integral, which determines the contribution of remote particles to the force acting on a particle, is convergent at the upper limit:

$$\int_{1/n}^{+\infty} \frac{dx}{x^{1+k}}.$$

Thus, the Hamiltonian of the system is ill defined for $0 < k \le 1$, but the equations of its motion are still well defined in the kinetic limit.

According to a dimensional analysis, thermal diffusivity can be expressed as

$$\chi = \phi\left(\frac{Qn^k}{T}\right)\frac{v_T}{n}, \quad (14)$$

where $v_T = \sqrt{2T/m}$ and ϕ is an unknown function that cannot be determined by dimensional analysis only.

On the other hand, when the potential is given by (13), kinetic equation (10) for smoothed distribution functions is invariant under a two-parameter transformation group G_1: each element of the group transforms a solution to equation (10) into another solution to the same equation. The group is defined by the relation

$$g(\lambda_1, \lambda_2) f_2 = \lambda_1^{-2k}\lambda_2^{-2} f_2(\lambda_1\lambda_2 t, \lambda_1 x_1, \lambda_2^{-1} v_1, \lambda_1^{-1}\lambda_2^{-2} a_1, \lambda_1 x_2, \lambda_2^{-1} v_2, \lambda_1^{-1}\lambda_2^{-2} a_2). \quad (15)$$

Moreover, the following law of multiplication is valid under the group:

$$g(\lambda_1, \lambda_2) g(\lambda_1^*, \lambda_2^*) = g(\lambda_1\lambda_1^*, \lambda_2\lambda_2^*).$$

Note that the nontrivial symmetry associated with G_1 is characteristic of kinetic equation (10), where collisions of arbitrary high order are allowed for. It can easily be shown that the symmetry does not hold for kinetic equations that allow for collisions only up to a finite order. In other words, the symmetry associated with G_1 manifests itself only when the "postcollision" correlations have fully developed.

The formal interpretation of the group G_1 is very simple: the problem under analysis is characterized by both physical dimensionalities (meters, grams, and seconds) and formal ones, the latter being described by G_1. Taking the formal dimensionalities into account, one can write the unknown function ϕ in (14) up to a constant as

$$\chi = \text{const}\left(\frac{Qn^k}{T}\right)^{1/(k-1)}\frac{v_T}{n}, \quad \phi(y) = \text{const}\, y^{1/(k-1)}. \quad (16)$$

SHORT-RANGE POWER-LAW INTERACTION POTENTIALS ($k > 1$)

Let us discuss the meaning of the results obtained. First of all, we note that the mean free path and the characteristic time of three-body collisions are estimated, respectively, as follows (recall that binary collisions do not contribute to the collision term):

$$l_{st} \sim \frac{1}{n^2}\left(\frac{T}{Q}\right)^{1/k} \text{ and } \tau_{st} \sim \frac{1}{v_T}\frac{1}{n^2}\left(\frac{T}{Q}\right)^{1/k}. \quad (17)$$

Thus, the thermal diffusivity χ_B evaluated within the framework of Boltzmann's approach is

$$\chi_B \sim v_T l_{st} = \frac{v_T}{n^2}\left(\frac{T}{Q}\right)^{1/k}. \quad (18)$$

The thermal diffusivities given by (16) and (18) can be compared by comparing the times $\tau^* \sim mv_T/(Qn^{(k+1)})$ and τ_{st} [see (9) and (17)]:

$$\frac{\tau^*}{\tau_{st}} \sim \left(\frac{T}{Qn^k}\right)^{(k-1)/k}. \quad (19)$$

Boltzmann's approach corresponds to the limit $n \longrightarrow 0$, when the time τ^* at which kinetic equation (10) becomes valid is infinitely longer that τ_{st}. Thus, a Boltzmann-type equation allowing for three-body collisions is merely inapplicable at times for which kinetic equation (10) is valid. Therefore, transport coefficients evaluated on the basis of Boltzmann's approach correspond to intermediate asymptotics that remain valid up to times when the true "postcollision" correlations have developed.

The applicability of Boltzmann's approach is limited to times at which one may neglect the fluctuations (associated with "postcollision" correlations) that cannot be described by a Boltzmann-type equation [5]. Moreover, expression (18) was derived without taking

into account the "mean field" that acts on particles in a nonuniform medium and affects transport phenomena.

Also recall that the *H*-theorem was proved by allowing for binary collisions only. When the collision term allows for three-body (and higher order) collisions, the *H*-theorem cannot be proved.

In addition, note that thermal conduction is described by linearized equation (10) when temperature substantially varies only over length scales $L \gg l^* = v_T\tau^*$ while a Boltzmann-type equation can be linearized when temperature substantially varies over $L \gg l_{st}$, where $l^* \gg l_{st}$. On the other hand, when the intensity of external noise is such that the system behaves as a closed one at times on the order of τ_{st} and cannot be treated as closed at times on the order of τ^*, equation (10) cannot describe its dynamics at any time. This suggests that the applicability of the proposed kinetic equations can be analyzed from a different perspective.

The thermal diffusivity given expressed by (16) is not an analytic function of *n*. This fact is of key importance here. Indeed, Bogolyubov relied on a direct analogy with cluster expansions for equilibrium problems in assuming that there exist analytic expansions of transport coefficients in powers of density (see [12, 13]). However, this issue has permanently evoked a substantial interest since it was shown in [7] that such an expansion actually does not exist.

LONG-RANGE POWER-LAW INTERACTION POTENTIALS ($0 < k < 1$)

When $0 < k < 1$, the thermal diffusivity allowing for fluctuations is again expressed as

$$\chi \sim \left(\frac{Qn^k}{T}\right)^{1/(k-1)} \frac{v_T}{n}, \tag{20}$$

while the thermal diffusivity calculated on the basis of Boltzmann's approach relying on the mean free path is

$$\chi_B \sim v_T l_{st} = \frac{v_T}{n^2}\left(\frac{T}{Q}\right)^{1/k}. \tag{21}$$

However, the physical interpretation of (20) and (21) is essentially different from the interpretation of their counterparts corresponding to short-range power-law potentials of interaction between particles. Indeed, thermal diffusivity is increased by fluctuations at $T \gg Qn^k$ in the case of long-range interaction, whereas the converse is true in the short-range case. Moreover, recall that Boltzmann-type equations are valid in the long-range case at times much shorter that τ_{st}. Without expanding on this observation, we illustrate it by a well-studied example: a quasi-ideal plasma [$N_D = (T/e^2n^{1/3})^{3/2}$] is described by the Lennard–Balescu kinetic equation at $t > 1/\omega_{pe}$, where $\omega_{pe}^2 = 4\pi ne^2/m$ is the plasma frequency, while the mean free time is $\tau_{st} \sim N_D/\omega_{pe}$ [13]. Clearly, these results are explained by the role played by collective processes in a system dominated by long-range forces.

THERMAL DIFFUSIVITY FOR $k = 1$

The intermediate case of $k = 1$ is of special interest. First of all, note that, when $k = 1$, *n* and *T* have the same dimensionality under the group G_1. Therefore, the theory must involve a parameter γ that can be treated as a dimensionless group both with respect to physical units of measure and under G_1:

$$\gamma = \frac{Qn}{T}.$$

On the other hand, the parameters *n* and *T* cannot be combined to obtain a quantity having the dimensionality of thermal diffusivity (i.e., transformed as the product of length squared with time) under the group G_1. Thus, we cannot construct a linear transport theory allowing for fluctuations with a small but constant ∇T when $k = 1$.

Let us now consider the case of an arbitrary *k* when there exists a temperature gradient slowly varying in space:

$$\frac{dT}{dx} = \left(\frac{dT}{dx}\right)_0 \cos(qx), \tag{22}$$

where $1/q$ is substantially greater than the mean free path.

The thermal diffusivity χ_q that characterizes a linear response to a temperature perturbation described by (22) is partially determined by the requirement that it should have correct dimensionalities both in physical units and under the group G_1. Accordingly, we write

$$\chi_q = v_T l_{fl} F_1(k, ql_{fl}) \quad \text{for} \quad k \neq 1, \tag{23}$$

where

$$l_{fl} = \frac{1}{n}\left(\frac{Qn^k}{T}\right)^{1/(k-1)},$$

and

$$\chi_q = \frac{v_T}{q} F_2(\gamma), \tag{24}$$

for $k = 1$, where $\gamma = Qn/T$ and F_1 and F_2 are unknown functions of the their arguments.

The relation between F_1 and F_2 can be specified by using the fact that χ_q for $k \neq 1$ becomes χ_q for $k = 1$ as $k \longrightarrow 1$ while *n*, *q*, and *T* are held constant:

$$\lim_{k \to 1} l_{fl} F_1(k, ql_{fl}) = \frac{F_2(\gamma)}{q}. \tag{25}$$

PHYSICAL MEANING OF THE RESULTS

First, note that any attempt to interpret (24) in terms of a mean free path Λ of a particle leads to a physically meaningless result: $\Lambda \sim 1/q$. However, expression (24) for thermal diffusivity can be understood by observing that a temperature gradient selectively affects fluctuations of a certain type (see Introduction) and this leads to spontaneous generation of spatial macroscopic structures (self-organization). In a real space, structures of this type manifest themselves as spontaneously developing fluid flows and particle beams. In view of (24) and (25), this suggests that, when k is close to unity, heat transfer is dominated by the contribution of macroscopic structures spontaneously generated in the phase space as a result of the effect of a temperature gradient on the structure of fluctuations.

When k is close to unity, the small-scale motion of particles weakly depends on the sign of $\Delta = 1 - k$, whereas the behavior of thermal diffusivity drastically changes with the sign of Δ at high T. Physically, this behavior is readily explained within the framework of the theory developed here by considering the structure of large-scale fluctuations as depending on the sign of Δ. Let us consider a segment of length R on the axis in the one-dimensional problem analyzed above. Treating the medium as a perfect gas (which should be a good approximation at high T), we can estimate the intensity of a particle number-density fluctuation on the segment, δN_R, as

$$\langle (\delta N_R)^2 \rangle \sim nR. \tag{26}$$

The contribution of such a fluctuation to the potential energy of the system is then estimated as

$$\delta U_R \sim QnR^{1-k}. \tag{27}$$

However, the Gibbs distribution rules out fluctuations contributing more than T to the energy of a system in thermal equilibrium. Therefore, the system cannot exhibit large-scale fluctuations when $k < 1$. This is an important distinction between a system characterized by a long-range potential of interaction between particles and a system for which $k > 1$, which must manifest itself in the behavior of transport coefficients.

CONCLUSION

In this paper, we show that fluctuations may affect transport properties of a medium in a nontrivial manner. Considering thermal diffusivity as an example, we have found conditions under which fluctuations impede relaxation. Moreover, under certain conditions ($k \approx 1$), fluctuations result in spontaneous generation of macroscopic structures that dominate the kinetic properties of a medium. Note that the argumentation and technique used here do not essentially rely on the fact that the problem is one-dimensional and can be extended to the multidimensional case.

It should be particularly noted that the passage from (7) to the limit in (9) is incorrect for a rarefied gas characterized by a potential of interaction with a radius much shorter than the mean distance between particles, because most of the particles are not accelerated at any time [i.e., the generalized distribution function reduces to $f(t, r, v)\delta(a)$] and irreversibility in time cannot be introduced by following the scheme employed here (i.e., by formally setting $\tau^* = +\infty$). Such problems call for a separate analysis; they have been analyzed in most detail for billiards of various kinds.

ACKNOWLEDGMENTS

We thank S.I. Anisimov for his interest in this study and discussions of the method and results described here.

This work was supported by the Russian Foundation for Basic Research, project nos. 98-02-17229 and 98-02-17441, and under the Support Program for Leading Scientific Schools, project no. 96-15-96448.

REFERENCES

1. *Handbook of Turbulence*, Ed. by W. Frost and T. H. Moulden (Plenum, New York, 1977).
2. G. G. Branover and A. B. Tsinober, *Magnetohydrodynamics of Compressible Media* (Nauka, Moscow, 1970).
3. A. V. Belyan, S. S. Moiseev, and O. G. Chkhetiani, *Turbulent Viscosity in Helical Turbulence* (IKI RAN, Moscow, 1992).
4. A. V. Belyan, S. S. Moiseev, and O. G. Chkhetiani, Dokl. Akad. Nauk **344**, 34 (1994).
5. M. A. Leontovich, Zh. Éksp. Teor. Fiz. **5**, 211 (1935).
6. S. Moiseev and O. Onishenko, Physica B **228**, 83 (1996).
7. J. R. Dorfman and E. G. D. Cohen, Phys. Lett. **16**, 124 (1965).
8. S. N. Gordienko, Zh. Éksp. Teor. Fiz. **106**, 436 (1994).
9. S. N. Gordienko, Fiz. Plazmy **23**, 754 (1997).
10. B. B. Kadomtsev, Zh. Éksp. Teor. Fiz. **33**, 151 (1957).
11. M. Kac, *Probability and Related Topics in Physical Sciences* (Am. Math. Soc., Providence, RI, 1957; Nauka, Moscow, 1965).
12. N. N. Bogolyubov, *Selected Papers* (Naukova Dumka, Kiev, 1970).
13. A. V. Shelest, *Bogolyubov's Method in Dynamical Theory of Kinetic Equations* (Nauka, Moscow, 1990).

Translated by A.S. Betev

Technical Physics, Vol. 45, No. 2, 2000, pp. 168–173. Translated from Zhurnal Tekhnicheskoĭ Fiziki, Vol. 70, No. 2, 2000, pp. 28–33.

GASES AND FLUIDS

Supersonic Air-Scoop Flows of a Weakly Ionized Gas in External Electromagnetic Field

Yu. P. Golovachev and S. Yu. Sushchikh

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

Received October 20, 1998

Abstract—Steady supersonic flows of a weakly ionized gas in external electric and magnetic fields computed for the problem of flow control in the air scoop of a supersonic vehicle. The numerical simulation was based on a second-order accurate marching scheme for the Euler equations. Effects of plasma and electromagnetic-field parameters on flow structure and distributions of flow variables are analyzed. © *2000 MAIK "Nauka/Interperiodica".*

INTRODUCTION

Effects due to magnetogasdynamic interactions in internal flows have been studied by numerous authors as related to problems of thermal-into-electrical energy conversion (e.g., see [1–3]). In this paper, we address a different problem, which must be solved to develop a new method for scoop flow control in a supersonic vehicle. The method makes use of gas preionization upstream of the scoop and flow structure control by means of external electric and magnetic fields [4]. Without going into the details of the preliminary ionization methods, we should note that they may rely on injection of an easily ionizable species or ionization by electron beams or electric discharges. A typical supersonic air scoop is a duct of a rectangular cross section with angular walls. The plasma parameters and external electromagnetic fields required for effective flow control were evaluated here by computing flows in such ducts. In particular, we analyzed the three-dimensional effects due both to the duct geometry and to the interaction between the flow and electromagnetic field. The computations were performed in the magnetogasdynamic approximation for an inviscid gas model. Numerical solutions were computed with the use of an explicit second-order accurate shock-capturing marching scheme.

STATEMENT OF THE PROBLEM

We consider steady supersonic flows of a preionized gas in electric and magnetic fields. Simple estimates show that, under conditions characteristic of hypersonic vehicles, these flows are described by the system of equations of magnetogasdynamics in which the effects due to external electromagnetic fields are represented by the ponderomotive force $\mathbf{F} = \mathbf{j} \times \mathbf{B}$ and ohmic heating rate $Q = \mathbf{j} \cdot \mathbf{E}$, where $\mathbf{j}$ is the conductive current density, $\mathbf{B}$ is magnetic induction, and $\mathbf{E}$ is the electric-field strength. Under the conditions of this study, the magnetic Reynolds number $\mathrm{Re}_m \ll 1$, which allows one to treat magnetic induction as a known quantity equal to the external magnetic induction. Then, if $\mathbf{E}$ is prescribed, the ponderomotive force and ohmic heating rate can be calculated by invoking generalized Ohm's law

$$\mathbf{j} = \mu_e(\mathbf{j} \times \mathbf{B}) = \sigma(\mathbf{E} + \mathbf{V} \times \mathbf{B}), \tag{1}$$

where μ_e is the electron mobility, σ is electrical conductivity, and $\mathbf{V}$ is the gas velocity.

Further assumptions that can be introduced under the conditions of this study include $\mathrm{Re} \gg 1$ for the flow Reynolds number and $\alpha \ll 1$ for the degree of ionization. They allow one to use an inviscid gas model described by the Euler equations and neglect the effect of ionization on thermodynamic properties.

Under these assumptions, the original system of equations for dimensionless variables is written in Cartesian coordinates as

$$\frac{\partial \mathbf{E}}{\partial x} + \frac{\partial \mathbf{F}}{\partial y} + \frac{\partial \mathbf{G}}{\partial z} = \mathbf{H}, \tag{2}$$

where the flux vectors for $\mathbf{E}$, $\mathbf{F}$, and $\mathbf{G}$ and the vector of source terms $\mathbf{H}$ are defined by the formulas

$$\mathbf{E} = \begin{pmatrix} \rho u \\ p + \rho u^2 \\ \rho uv \\ \rho uw \\ Hu \end{pmatrix}, \quad \mathbf{F} = \begin{pmatrix} \rho u \\ \rho uv \\ p + \rho v^2 \\ \rho vw \\ Hv \end{pmatrix},$$

1063-7842/00/4502-0168$20.00

$$\mathbf{G} = \begin{pmatrix} \rho w \\ \rho u w \\ \rho v w \\ p + \rho w^2 \\ H w \end{pmatrix}, \quad \mathbf{H} = \begin{pmatrix} 0 \\ F_x \\ F_y \\ F_z \\ Q \end{pmatrix}.$$

Here, ρ, p and H are the specific density, pressure, and total heat of the gas, respectively, and u, v, and w are the velocity projections onto the x, y, and z axes in a coordinate system whose axis is aligned with the channel centerline. When the external magnetic and electric fields are parallel to the y and z axes, the projections of ponderomotive force are written in dimensionless variables as

$$F_x = -\frac{S}{1+\beta_e^2}\left[\beta_e w + u\left(1 + \frac{R_F}{u}\right)\right], \tag{3}$$

$$F_y = 0, \tag{4}$$

$$F_z = \frac{S}{1+\beta_e^2}\left[-w + \beta_e u\left(1 + \frac{R_E}{u}\right)\right], \tag{5}$$

and the source term of the energy equation is

$$Q = \frac{S}{1+\beta_e^2} R_E(\beta_e w + u + R_E). \tag{6}$$

Here, the electric-field strength is expressed in terms of the parameter $R_E = E/(V_0 B)$, where V_0 is the reference gas velocity. When the duct operated as an MHD generator, the electric-field strength is sought in terms of the external load ratio $k = E/(V_0 B)$ while the source terms in equation (2) are calculated by formulas (3)–(6) in which R_E/u is replaced by k.

Interaction between the flow and electromagnetic field is characterized by the parameter R_E, the Stuart number $S = \sigma B^2 L/(\rho_0 V_0)$, and the Hall parameter $\beta_e = \mu_e B$, where L is the reference length and ρ_0 is the reference density. Formula (3) shows that the flow acceleration regime corresponds to negative values of R_E satisfying the condition

$$R_E V_0 < -(\beta_e w + u). \tag{7}$$

Otherwise, the flow is decelerated by electric and magnetic fields.

In the present study, it was assumed that the projection of gas velocity onto the x axis exceeds the sonic velocity in the entire computational domain. Under this condition, the system of equations (2) is hyperbolic with respect to x and an initial–boundary value problem is posed for this system, with free-stream parameters at $x = 0$ used as initial conditions. Impermeability conditions are set on the duct walls.

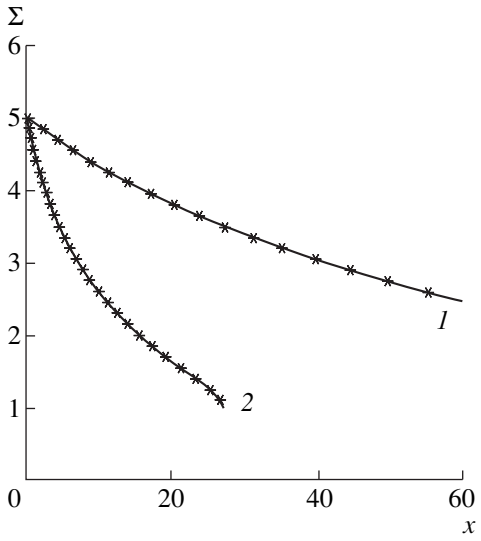


Fig. 1. Comparison of analytical (curves) and numerical (symbols) solutions for S = 0.02: k = (*1*) 0.25 and (*2*) 0.75.

NUMERICAL METHOD

The system of equations (2) was integrated by means of the explicit marching scheme described in [5], which is second-order accurate as applied to smooth solutions, conservative, and characterized by nonincreasing total variation. Each step along the marching coordinate involves predictor and corrector steps. The Riemann problem is solved only at the corrector stage. Increments are calculated for the nonconservative variables and are limited by means of the operator minmod. The source terms, which take into account the interaction between the flow and the external electromagnetic field, are calculated at the predictor and corrector steps in terms of the current values of nonconservative variables.

The numerical method was tested by analyzing problems that have exact analytical solutions. As a problem of this type, we considered the one-dimensional steady plasma flow in a duct of constant cross section in a constant external magnetic field operating in the regime of an MHD generator. Figure 1 shows graphs of the Mach number distributions along the duct obtained by analytical solution [1] (solid curve) and computation (symbols). The numerical solution of this one-dimensional problem was computed by the three-dimensional code to test the algorithm.

For the numerical method to be applicable, the problem must be hyperbolic with respect to x. This property may be violated even in supersonic flows if the direction of velocity vector substantially deviates from the x direction. This disadvantage is eliminated by changing to a new coordinate system, with axes λ and ξ parallel

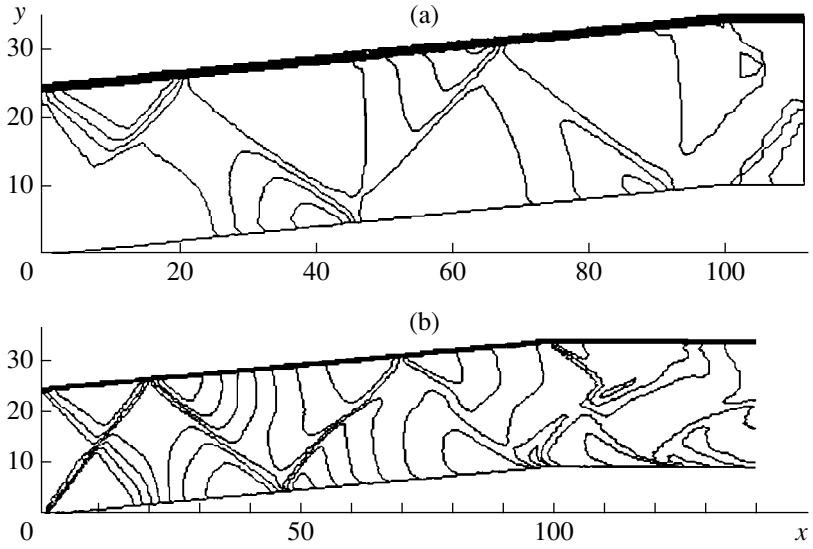


Fig. 2. Lines of constant density in a planar duct with angular walls for $M_0 = 1.4$: (a) computed in Cartesian coordinates; (b) computed in coordinates tied to streamlines.

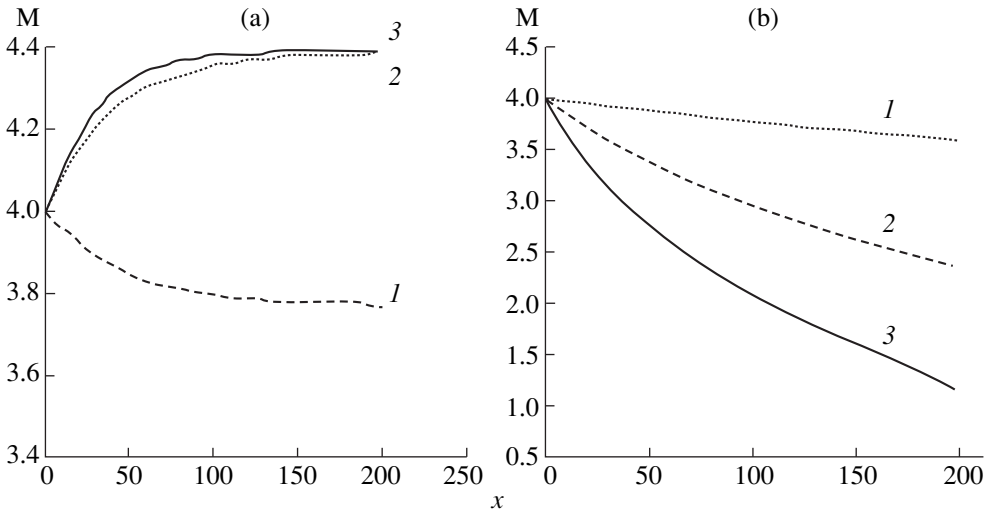


Fig. 3. Effects of (a) the Stuart number and (b) parameter R_E on the variation of the Mach number in nonparallel electric and magnetic fields for $\beta_e = 0$: (a) $R_E = -1.1$; S = (*1*) 0.05, (*2*) 0.10, and (*3*) 0.15; (b) S = 0.001; R_E = (*1*) –0.75, (*2*) –0.25, and (*3*) 0.25.

and normal to the streamlines, respectively. In the two-dimensional case, new independent variables are defined by the formulas [6]

$$dx = \frac{u}{q}d\lambda + Ud\xi, \quad dy = \frac{v}{q}d\lambda + Vd\xi,$$

where $q = \sqrt{u^2 + v^2}$ is flow velocity and U and V are geometric variables having the dimensionality of inverse velocity.

In the new system, the vector quantities again are defined by their Cartesian coordinates, and the governing equations are identical in form to those written in Cartesian coordinates,

$$\frac{\partial \mathbf{F}}{\partial \lambda} + \frac{\partial \mathbf{F}}{\partial \xi} = \mathbf{H},$$

where

$$\mathbf{E} = \begin{pmatrix} K \\ Ku + pV \\ Kv - pU \\ KH \\ U \\ V \end{pmatrix}, \quad \mathbf{F} = \begin{pmatrix} 0 \\ -p\sin\theta \\ -p\cos\theta \\ 0 \\ -\cos\theta \\ -\sin\theta \end{pmatrix}, \quad \mathbf{H} = \begin{pmatrix} 0 \\ F_x \\ F_\xi \\ Q \\ 0 \\ 0 \end{pmatrix},$$

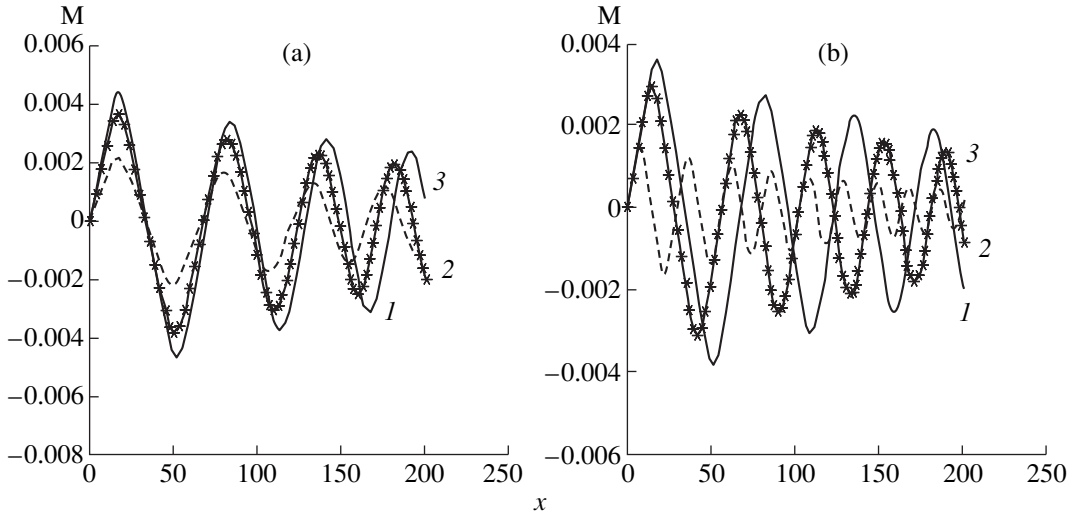


Fig. 4. Variation of the transverse gas velocity along the centerline of a duct with parallel walls for $M_0 = 4$, $R_E = 0.25$, and $S = 0.001$: (a) $l/h = 25$; $\beta_e = 0.25$ (*1*), 0.50 (*2*), and 0.75 (*3*); (b) $\beta_e = 0.5$; $l/h = 20$ (*1*), 25 (*2*), and 50 (*3*).

$K = \rho(uV - vU)$, $\cos\theta = u/q$, and $\sin\theta = v/q$.

Here, the first four equations are the mass, momentum, and energy conservation laws. The last two equations are geometric relations entailed by the transformation of coordinates. In the problem under consideration, λ is the marching coordinate. The numerical algorithm employed is similar to that used in Cartesian coordinates.

The wider possibilities associated with the use of coordinates tied to streamlines are illustrated by Fig. 2. Here, we present numerical results obtained for a planar flow in a duct with angular walls in the absence of electromagnetic field and lines of constant density plotted with the increment $\Delta\rho/\rho_0 = 0.03$. Figure 2a shows results calculated in Cartesian coordinates; Fig. 2b, those calculated in the coordinates λ and ξ. The computation in the Cartesian coordinates were stopped at $x \approx 110$ because the hyperbolicity of the system of equations with respect to x was violated. The marching algorithm of computation in the coordinates tied to streamlines remains effective in a greater part of the flow region, ensuring a finer resolution of weak shocks.

DISCUSSION OF RESULTS

In accordance with the assumptions made in the computations, we used the simplest physical model of homogeneous perfect gas with the ratio of specific heats $\gamma = 1.4$. The results are presented in dimensionless form. We used the velocity and density at the inflow boundary of the computational domain as reference values of these quantities. Pressure was normalized to γM_0^2, where M_0 is the Mach number at the inflow boundary.

First, we discuss the results computed for one-dimensional flows in a duct with straight parallel walls for the zero value of the Hall parameter. Figure 3b demonstrates the effects of magnetogasdynamic interaction parameters on the Mach-number distribution along the duct. At $R_E = -1.1$ (Fig. 3a), the duct flow accelerates. The lower Mach number obtained for $S = 0.05$ (curve *1*) is explained by the faster increase in sonic velocity as compared to gas velocity. Each value of R_E is associated with a particular peak value of the Mach number, reached at $S = 0.1$ in the case in question. When R_E is positive or negative (but small in absolute value), interaction with electric field decelerates the flow (see Fig. 3b). When $R_E = 0.25$ (curves *3*), the gas velocity decreases and becomes equal to the sonic velocity.

In the presence of Hall current and electromagnetic field of the same geometry, the gas velocity in a duct with straight parallel walls has a component parallel to the *z*-axis. Figures 4a and 4b show, respectively, the distributions of transverse velocity and duct length-to-width ratio l/h along the duct axis for various values of the Hall parameter. When R_E has the value indicated above, the flow decelerates. The oscillatory behavior of the transverse velocity is explained by reflection of the flow from the duct walls. Variations in the Hall parameter and duct width produce a significant effect on the oscillation amplitude and period, and the distance x at which the effect of walls was negligible decreased as the ratio l/h was increased.

The next two figures show the lines of constant gas density in a channel with an angular bottom wall. Here, the duct height-to-width ratio at the scoop cross section is 2, and the curves are plotted with the increment $\Delta\rho/\rho_0 = 0.15$. Figures 5a and 5b show planar flow patterns that take place in the absence of external fields and in an electromagnetic field without Hall current,

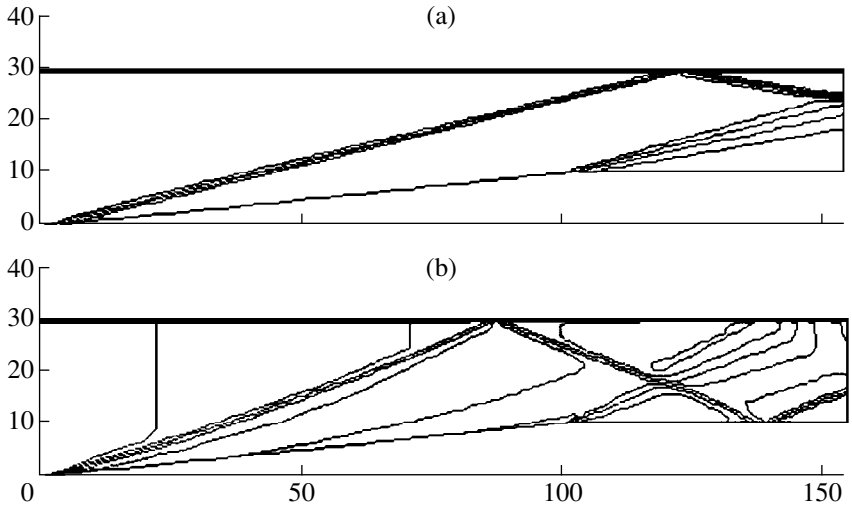


Fig. 5. Lines of constant density in a planar gas flow through a duct with angular walls for $M_0 = 6$: (a) $E = B = 0$; (b) S = 0.005, $R_E = -0.5$, $\beta_e = 0$.

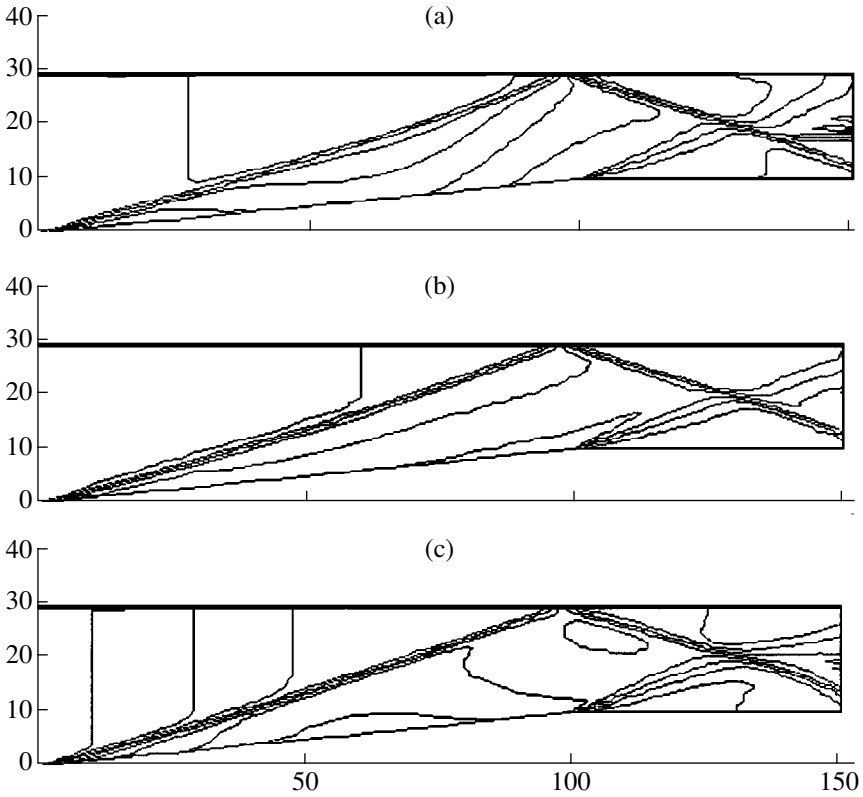


Fig. 6. Lines of constant density in a three-dimensional flow in a duct with angular walls for $M_0 = 6$, S = 0.005, $R_E = -0.5$, and $\beta_e = 1.0$: $z = h/2$ (a), 0 (b), and $-h/2$ (c).

respectively. A comparison of the figures demonstrates a considerable effect of the electromagnetic field on the flow structure, which manifests itself by changes in the locations of shock waves and the broadening of shock waves and rarefaction-wave fans. Figures 6a–6c show results computed for the three-dimensional flow developing in the presence of the Hall current and the lines of constant density in the central cross section of the

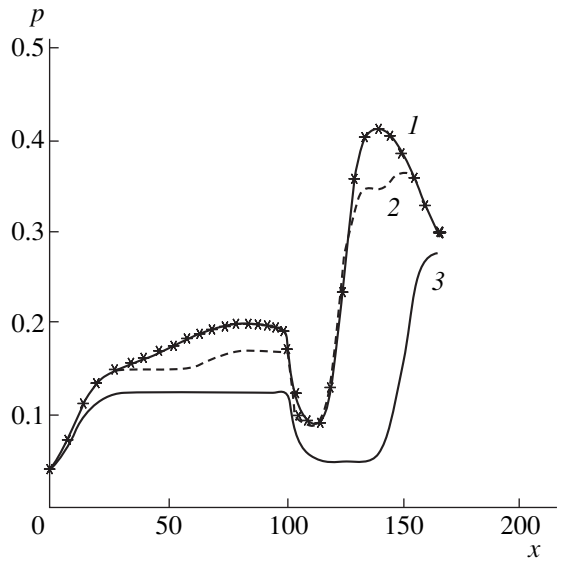


Fig. 7. Pressure distributions over the bottom duct wall for various orientations of the magnetic field for $M_0 = 4$, $z = 0$, $S = 0.005$, $\beta_e = 0.5$, and $k = 0.5$: (*1*) $B_y = 0$, $B_z = B$; (*2*) $B_y = B$, $B_z = 0$; (*3*) $B_y = B_z = 0$.

duct (Fig. 6a) and on the duct walls (Figs. 6b and 6c). A comparison with Fig. 5b makes it possible to estimate the effect of the Hall current.

Figure 7 demonstrates the effect of orientation of magnetic induction in the case when the duct operates in the regime of an MHD generator with the external load ratio $k = 0.5$. It is obvious the effect of magnetic field is maximized when the vector of induction is aligned with the *z*-axis.

Thus, the results presented here demonstrate the feasibility of control of the parameters and shock-wave structure of supersonic scoop flow by means of external electric and magnetic fields. For the length scale $L = 1$ cm, the magnetic induction $B = 2T$, the Mach number $M_0 = 6$, and the density ρ_0 corresponding to the height $H = 20$ km, the Stuart numbers required for effective control are attained when the plasma conductivity is several Ω^{-1} m^{-1}.

ACKNOWLEDGMENTS

This work was supported by the Russian Foundation for Basic Research, project no. 98-01-01121.

REFERENCES

1. G. Sutton and A. Sherman, *Engineering Magnetohydrodynamics* (Toronto, 1965; Mir, Moscow, 1968).
2. A. B. Vatazhin, G. A. Lyubimov, and S. A. Regirer, *Magnetohydrodynamic Flows in Ducts* (Nauka, Moscow, 1970).
3. V. V. Breev, A. V. Gubarev, and V. P. Panchenko, *Supersonic MHD Generators* (Énergoatomizdat, Moscow, 1988).
4. Yu. P. Golovachev, S. A. Il'in, and S. Yu. Sushchikh, Pis'ma Zh. Tekh. Fiz. **23** (16), 1 (1997) [Tech. Phys. Lett. **23**, 615 (1997)].
5. A. V. Rodionov, Zh. Vychisl. Mat. Mat. Fiz. **27**, 1853 (1989).
6. W. H. Hui and D. L. Chu, Comput. Fluid Dyn. J. **4**, 403 (1996).

Translated by A.S. Betev

Technical Physics, Vol. 45, No. 2, 2000, pp. 174–178. Translated from Zhurnal Tekhnicheskoĭ Fiziki, Vol. 70, No. 2, 2000, pp. 34–37.

An Injector of Xenon Macroscopic Pellets for Quenching the Fusion Reaction in a Tokamak

P. V. Reznichenko, I. V. Vinyar, and B. V. Kuteev

St. Petersburg State Technical University, St. Petersburg, 195251 Russia

Received July 7, 1998; in final form, February 18, 1999

Abstract—An injector for the formation of solid xenon pellets and their injection into the tokamak plasma is designed on the basis of a light-gas gun. Experimental results and calculations on production of xenon pellets with a diameter of 4 mm and a length of up to 25 mm directly inside of the barrel are presented. © *2000 MAIK "Nauka/Interperiodica".*

INTRODUCTION

When a discharge in a tokamak is quenched, a high-temperature plasma falls on the diverter plates and may destroy them. In order to prevent a local plasma-energy release in the diverter, it was proposed to rapidly quench the discharge and to convert the plasma energy into the radiative form by injecting macroscopic pellets of frozen noble gases with large nucleus charges [1]. Injection of pellets of solidified gases into nuclear fusion installations is widely used for both fuel introduction and plasma diagnostics [2]. Injectors based on a light-gas gun are most frequently used for fuel introduction. Gaseous hydrogen (or deuterium) is frozen in the form of pellets in the breech of the gun. The small mass of hydrogen facilitates the formation of pellets from the condensed phase, because surface-tension forces prevent liquid drops from spreading over the barrel stacks. In the case of comparatively heavy inert gases, it was necessary to develop an injector with another method for pellet formation, which is described below.

INJECTOR DESIGN AND OPERATING PRINCIPLE

Figure 1a shows a schematic diagram of the injector. A 0.2-m-long copper rod *2* with a rectangular cross section of 8×24 mm is soldered to a thin-walled section of the barrel *1*. The other end of the rod is immersed into a vessel with liquid nitrogen *3*. The inner diameter of the barrel is 4 mm, and its length is 0.8 m. A semiconductor temperature sensor *4* with a measurement range of 5–273 K is attached to the rod at a distance of 8 mm from the barrel axis. The barrel has an electromagnetic impulse valve *5* connected to a cylinder *6* with a pressure regulator filled with compressed helium or hydrogen. The internal volume of the valve chamber is 70 cm^3, and the valve opens in 1 ms at a helium (hydrogen) pressure of up to 12 MPa. The barrel end enters the diagnostic chamber *7*, which is equipped with an optical sensor *8* with two photodiodes located along the barrel axis at a distance of 10 mm from each other. Microphone *9* is attached to the chamber flange. The photodiodes are illuminated with a laser beam *10* expanded to a diameter of 15 mm.

The xenon-filled cylinder *11* with a pressure regulator is connected to the diagnostic chamber, impulse valve, and vacuum pump *13* with a system of pipelines with valves *12* and *14–16*.

In order to visualize the pellet formation process, a unit *17* with an optical window could be attached to the initial part of the barrel instead of the valve (see Fig. 1b). In this case, the impulse valve was coupled to the side surface of unit *17*, and a light beam was directed to the barrel through the windows of the diagnostic chamber.

For qualitative estimation of the strength of produced pellets, a piezoelectric pressure gage *18* could be set at the barrel end instead of at the diagnostic chamber (Fig. 1b). The magnitude and shape of the pressure impulse, which is produced in the barrel in front of the accelerated pellet *19*, allowed us to evaluate the mechanical strength of the frozen pellet.

The operating principle of the injector is demonstrated by Fig. 1. A segment of the barrel is cooled to a temperature of 135–140 K, at which xenon supplied to the barrel through valve *12* and the diagnostic chamber is frozen. After the formation of pellets according to the technique described below, the cylinder with xenon is shut off, and pump *13* evacuates the barrel from both sides of the particle through the valves *5*, *16*, *14*, and the diagnostic chamber. The valve's design allows it to be maintained in the open state for several tens of seconds without overheating the windings. After the barrel is evacuated, valve *5* is closed, and compressed helium

1063-7842/00/4502-0174$20.00

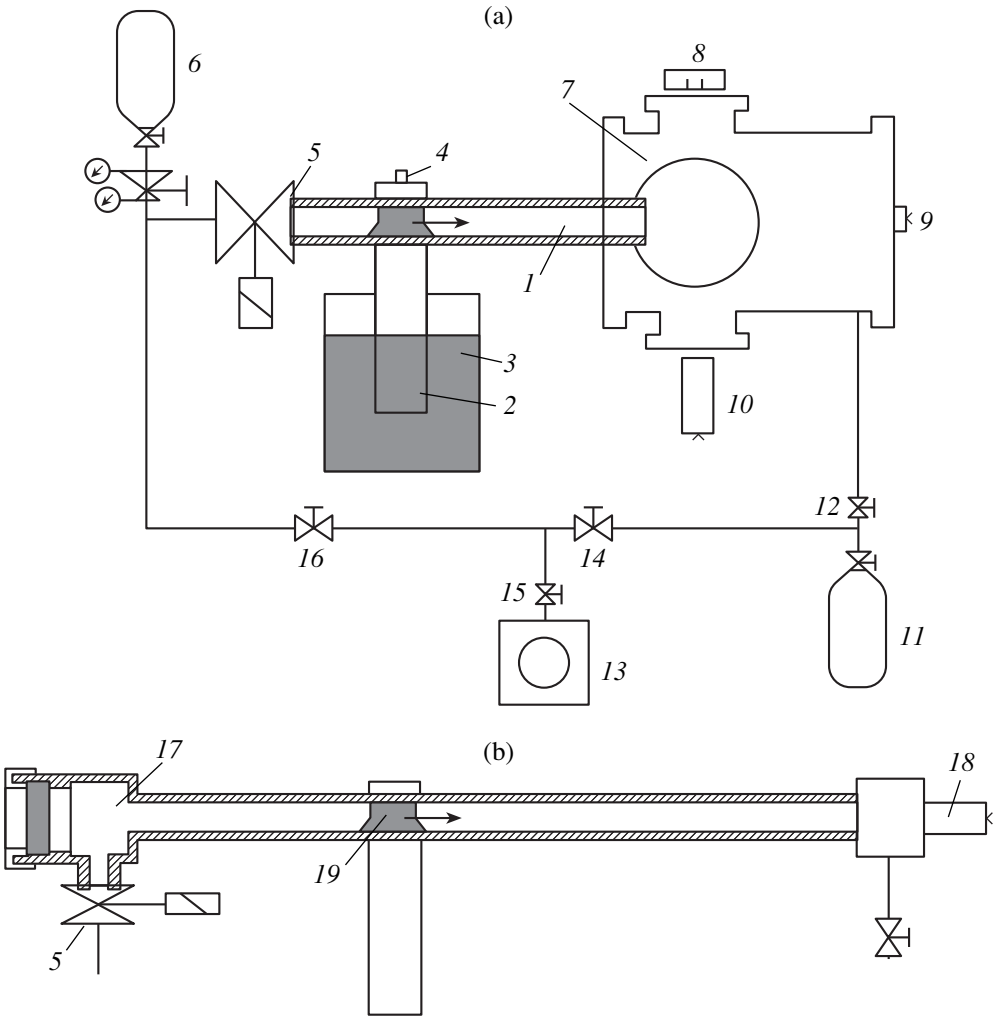


Fig. 1. (a) Schematic diagram of the pellet injector and (b) modification of its barrel.

is supplied to its chamber from cylinder *6*. When the valve is open pulsewise, the gas at a pressure of 1–10 MPa enters the barrel and accelerates the pellet.

EXPERIMENTAL TECHNIQUE AND RESULTS

Among noble gases with large nucleus charges, xenon was selected for experiments, because the temperature of its triple point is 161.4 K. This allowed us to easily obtain the solid phase by using liquid nitrogen for cooling.

First experiments have shown that, at a pressure below that at the triple point for xenon (0.08 MPa), pellets were shaped in the form of a loose snowlike cylinder, which was easily destroyed when being accelerated. For pressures and temperatures above the triple point, it was revealed that, in contrast to liquid hydrogen [3], liquid xenon spilled along the barrel without forming a compact pellet. The ratio of the surface tension coefficient to the density for liquid xenon near the triple point is almost seven times lower than for liquid hydrogen; therefore, the surface tension forces are insufficiently strong to hold a xenon drop in a tube with an inner diameter of 4 mm. We developed the following technique for shaping solid pellets.

After the barrel is evacuated to 2–3 Pa and the copper rod is cooled to 135–140 K, gaseous xenon was supplied to the barrel from the side of the diagnostic chamber at a pressure of 0.1 MPa (see Fig. 2a). A thin layer of solid xenon appeared on the barrel's walls 5–10 s later. Valve *14* was closed, valve *5* was fully uncovered, and valves *15* and *16* were slightly opened, thus providing for the slow evacuation of xenon from the side of the impulse valve (Fig. 1a) and creating a weak xenon flow through the cooled section of the barrel, as shown in Fig. 2b. The xenon flow and the rate of the liquid-xenon layer build-up were so small that the liquid had enough time to solidify almost without spilling along the barrel. One could observe through the barrel

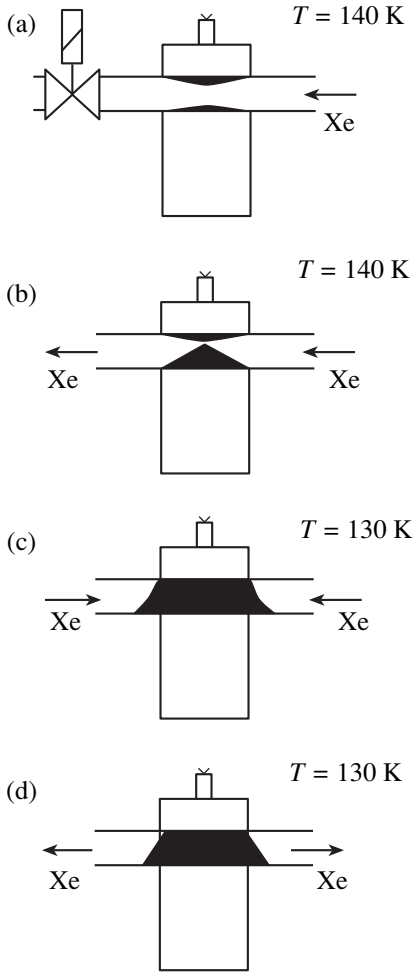


Fig. 2. Stages of formation of a xenon pellet inside the barrel: (a) gas admission and formation of a solid xenon layer on the barrel walls; (b) xenon blow-through and formation of a solid xenon pyramid; (c) xenon condensation from the opposite sides and formation of a truncated pyramid; and (d) xenon sublimation in a vacuum.

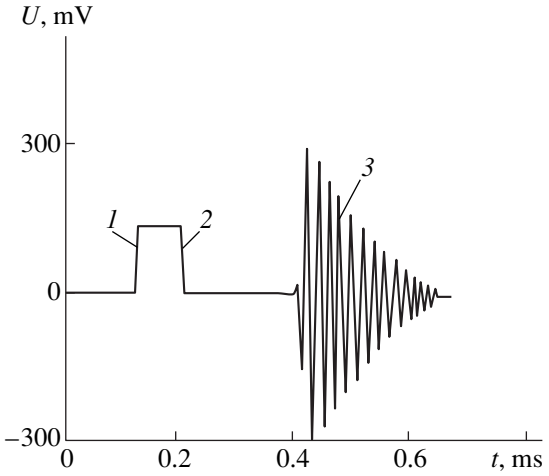


Fig. 3. Oscillogram of the (*1*, *2*) photodiode and (*3*) microphone signals from the flying pellet.

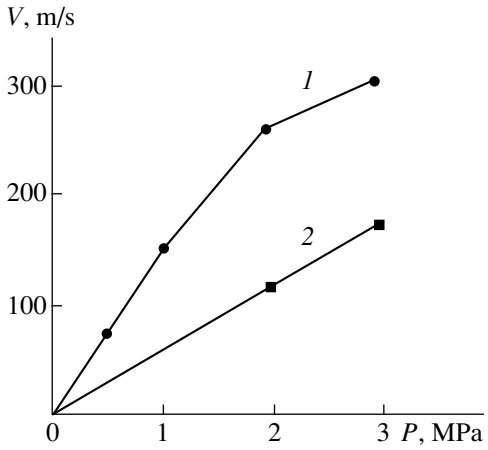


Fig. 4. Velocities of a xenon pellet as a function of gas pressure: (*1*) helium and (*2*) nitrogen.

how the barrel cross section was being filled with solid xenon from the bottom upwards, forming a sort of pyramid (Fig. 2b). Approximately 100 s later, the pyramid vertex reached the upper point of the internal surface of the barrel, and the xenon flow through the barrel stopped. Valve *14* was then slightly opened, and the xenon pressure decreased to 0.02 MPa due to evacuation. Valve *15* was then closed, and the gas continued condensing from two opposite sides on the solid xenon strap formed in the barrel (Fig. 2c). In this case, the temperature of the copper rod was reduced to 125–130 K by extending its part immersed in the vessel with liquid nitrogen. The pellet acquired the final form of a truncated pyramid 120–180 s later (Fig. 2c). Before the shot, the pellet was kept for 60–90 s in a vacuum (see Fig. 2d) in order to weaken its freezing to the barrel as a result of sublimation.

This technique allowed us to stably produce pellets with a diameter of 4 mm and a length of up to 25 mm. The formation time was ~360 s. In order to test the strength of the formed pellet, the impulse valve was sometimes opened, and helium at a pressure of up to 0.1 MPa was admitted into the barrel. A correctly manufactured particle withstood such a pressure without letting helium pass into the diagnostic chamber, thus showing that it was formed from xenon ice and not from snow.

The shot was executed by helium compressed to a pressure of 1–10 MPa entering into the barrel from the impulse valve. The pellet flew into the diagnostic chamber and successively shielded both photodiodes (PDs) from laser light incident on them. At this instant, the electronic system shaped a pulse with a typical form, which is shown in Fig. 3. Without a delay, this pulse triggered a photoflash, which illuminated the pel-

let in the flight for 40 ns. A photo or video camera recorded a shadow photograph of the particle. The latter hit the chamber flange with a membrane, to the exterior side of which a microphone is fixed. The microphone signal and the signals from the PDs were recorded by a storage oscilloscope. The velocity of pellets was calculated from the distance between the PDs and the microphone and the time delays between the signals on the oscilloscope. Figure 3 is a typical oscillogram showing the leading edge *1* of the pulse shaped by the first PD at the moment when it is shaded by the front part of the pellet; the duration of the pulse itself characterizes the velocity and size of the particle during shielding of both PDs; its trailing edge *2* is shaped at the moment when the rear part of the particle bypasses the second PD; and the microphone signal *3* is shaped at the moment the particle hits it. The ratio of the distance between the first PD and the microphone to the time interval between signals *1* and *3* determined the pellet's velocity. The time interval between signals *1* and *2* and the distance between the PDs allowed us to estimate the pellet's length using the velocity obtained.

Figure 4 shows the experimental results obtained for the velocities of pellets as a function of the pressure of the propelling gas (helium or nitrogen). The maximum velocity (300 m/s) was reached for 25-mm-long particles with a diameter of 4 mm. Such a velocity is sufficient for injecting a particle behind the diverter layer and quenching the discharge in the surface plasma layer with energy re-emission to the entire interior surface of the tokamak chamber. However, in order to quench the discharge completely, the velocity of pellets must be increased by using, for example, a double-stage light-gas gun. Since pistons in the second stage of such guns often have to be replaced because of wear [4], we proposed to simultaneously shape a one-shot piston from a gas together with a pellet [5]. Gas pressure pulses with a peak power above 100 MPa in front of pellets were obtained in experiments that consisted of shooting into a barrel preliminarily filled with helium at a pressure of 0.01–0.4 MPa and closed by a pressure piezosensor (see Fig. 1b). Figure 5 shows typical pulse shapes obtained when nitrogen at a pressure of 5 MPa was used for pellet acceleration and the initial nitrogen pressure in front of a pellet was 0.2 and 0.3 MPa. We see that the strength of xenon pellets and the duration of pressure impulses are sufficient for using the developed technique for shaping one-shot pistons in double-stage guns, thus solving the problem of their durability. Moreover, by using the model for calculating double-stage guns developed in [6], one can evaluate the velocity of solid hydrogen pellets for the case when the gas in the second stage is compressed by a piston similar to that manufactured by us in experiments. If the initial helium pressure in the second stage before compression by a xenon piston was 0.4 MPa and after the compression it reached

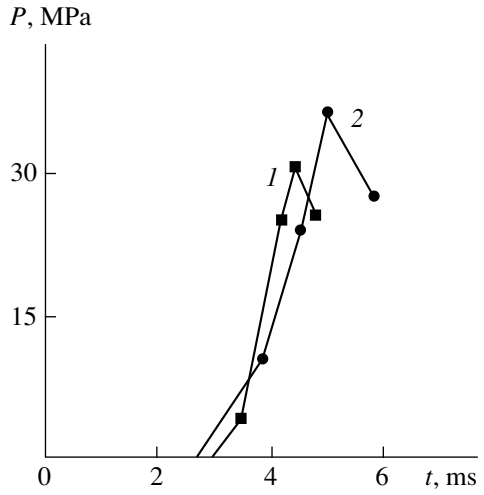


Fig. 5. Pressure impulses P produced in the gas in front of the accelerated pellet. The initial nitrogen pressure is (*1*) 0.2 and (*2*) 0.3 MPa.

100 MPa, then, assuming the compression to be adiabatic, we obtain a temperature of over 3000 K for the compressed gas. The pellet's velocity is approximately equal to the sound velocity in the gas multiplied by certain correcting coefficients that take into account friction, heat transfer from the gas to the barrel, actual gas properties, the ratio of the accelerated gas mass to the pellet's mass, and construction parameters of the double-stage gun. If the dimensions of the barrel, pellet, and the both stages of the gun are selected correctly, all these coefficients increase the pellet's velocity by a factor of 1.8–2.5 with respect to the sound velocity [6]. Taking into account that the experimental sound velocity in the compressed gas was 3.2 km/s, we may conclude that the use of a piston made of solid xenon would allow us to accelerate a solid hydrogen pellet with a size of 1–2 mm to a velocity exceeding 3 km/h in the second stage of a double-stage gun.

CONCLUSION

This paper describes a new technique for shaping solid xenon particles that can be used in injectors to introduce pellets in tokamak plasmas in order to prevent the discharge quench and plasma energy ejection to the diverter. The technique was implemented by an injector design in which pellets with a diameter of 4 mm and a length of up to 25 mm were formed and accelerated to velocities of up to 300 m/s. The duration of the pellet formation and acceleration cycle was about 360 s. It was shown experimentally that, when being accelerated, pellets can produce a pressure impulse with an amplitude above 100 MPa in the gas in front of them. Therefore, they can be utilized as one-shot pistons in double-stage light-gas guns to accelerate pel-

lets, including those of solid hydrogen isotopes, to velocities exceeding 3 km/s.

REFERENCES

1. B. V. Kuteev, V. Yu. Sergeev, and S. Sudo, Nucl. Fusion **35**, 1167 (1995).
2. S. Milora, W. Houlberg, L. Lenguel, *et al.*, Nucl. Fusion **35**, 657 (1995).
3. S. Combs, Rev. Sci. Instrum. **64**, 1679 (1993).
4. S. Sudo, T. Baba, M. Kanno, *et al.*, Fusion Technol. **20**, 387 (1991).
5. I. V. Vinyar, RF Inventor's Certificate No. 1699298, Byull. Izobret. **15**, 223 (1997).
6. *Ballistic Setups and Their Use in Experimental Studies*, Ed. by N. A. Zlatin and G. I. Mishin (Nauka, Moscow, 1974).

Translated by A. Seferov

Technical Physics, Vol. 45, No. 2, 2000, pp. 179–183. Translated from Zhurnal Tekhnicheskoĭ Fiziki, Vol. 70, No. 2, 2000, pp. 38–42.

SOLIDS

Polarization Model of Strengthening of Thermoplastic Materials Containing Ultradisperse Inorganic Fillers

L. S. Pinchuk,[1] S. V. Zotov,[1] V. A. Gol'dade,[1] A. V. Vinogradov,[2] A. A. Okhlopkov,[2] and S. A. Sleptsova[2]

[1] *Institute of Mechanics of Metallopolymeric Systems, Academy of Sciences of Belarus, Gomel, 246652 Belarus*

[2] *Yakutsk State University, Yakutsk, 677891 Russia*

Received September 16, 1998

Abstract—The polarization mechanism of strengthening was studied in thermoplastic polymers filled with ultradisperse powders of refractory SiMeON compounds obtained by plasmachemical synthesis. Thermostimulated depolarization current measurements showed the presence of spontaneous polarization charge in the filler and in filled thermoplastic compositions. Effects of the electric polarization field of the filler particles on the strength and structure of a boundary surface layer in a thermoplastic polymer binder contacting with the filler were studied on model samples. Material in the boundary layer exhibited an increase in the strength and a change in the degree of crystallinity and in the melting temperature. © *2000 MAIK "Nauka/Interperiodica".*

INTRODUCTION

Silicon nitride based solid solutions, in which nitrogen atoms are partly replaced by oxygen, and silicon atoms are partly replaced by aluminum, boron, or yttrium, can be synthesized in the form of ultradisperse powders with a particle size in the 0.01–0.10 μm range [1]. Polymers filled with these substances possess unusual mechanical properties, sometimes being significantly superior to those of conventional plastics [2, 3]. This effect is usually attributed to a high density of defects in the crystalline structure of these polymers and a large specific surface of dispersed particles [1]. The latter circumstance implies that a considerable fraction of the polymeric binder in a plastic composition occurs in the state of an interphase boundary layer (interface) possessing special properties. In addition, there are grounds to believe that some features of the high-frequency (or radio-frequency, RF) plasma technologies used for the production of ultradisperse fillers lead to polarization of the final particles.

The purpose of this work was to study the charged state of ultradisperse particles obtained by the methods of plasmachemical synthesis from inorganic refractory compounds and determine the mechanisms of influence of the particle polarization on the mechanical properties of thermoplastic compositions filled with these particles.

EXPERIMENTAL

Ultradisperse powders of inorganic compounds of the sialon type were originally obtained by methods of plasmachemical synthesis [1]. According to this technology, powdered silicon and oxides of aluminum, yttrium, and boron are introduced into a flow of RF nitrogen plasma. The temperature of flow in the region of powder introduction into the plasma is typically 5600–6200 K. In order to cool the plasma to $T = 2300$ K (a temperature at which stable nitrogen-containing silicon compounds can exist) and increase the concentration of active nitrogen in the plasma, the reaction zone was additionally supplied with ammonia. Since the initial solid materials were completely evaporated, the reactions of synthesis occur in the gas phase and the final products are condensed in the form of an ultradisperse silicide (UDS) powder:

$$\text{Si + Me oxide + N} \xrightarrow{\text{RF nitrogen plazma}} \text{SiMeON.}$$

The sample compositions were prepared using thermoplastic polymer binders characterized by significant polarizability and large lifetime of the polarized state, including high-density poly(ethylene) (HDPE), poly(vinyl butyral) (PVB), and poly(tetrafluoroethylene) (PTFE). The electric polarization of the samples prepared in the form of polymeric films on an aluminum foil substrate was produced using two methods. Corona electrets (CE) were obtained by treating samples in the field of corona discharge (at a discharge voltage $U = -25$ kV). Thermoelectrets (TE) were prepared in the electric field between two electrodes, one of these being the grounded substrate foil of a sample, and the other contacting with the sample layer (through an insulator film). A negative potential applied to the second electrode ensured a field strength of $E = 20$ kV/cm in the polymeric layer. The charging time ($t = 10$ min) and the temperature ($T = 333$ K) were the same for all samples.

1063-7842/00/4502-0179$20.00 © 2000 MAIK "Nauka/Interperiodica"

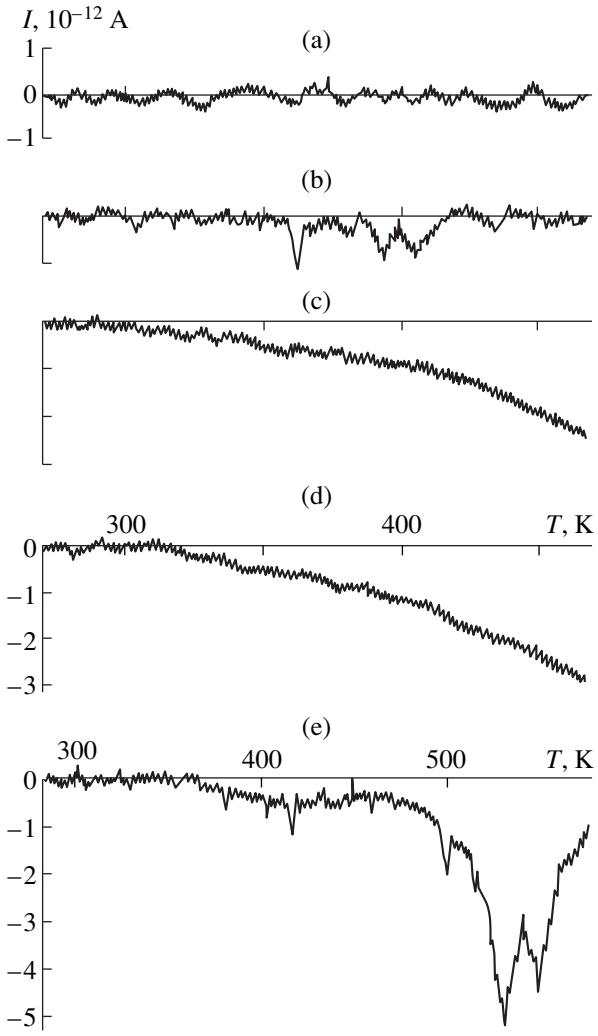


Fig. 1. The plots of TSD current versus temperature for (a) initial (unfilled) HDPE, the same polymer filled with (b) 0.5, (c) 2, and (d) 20% UDS, and (e) pure UDS.

Processes occurring in the thermoplastic polymers during their filling with UDS were modeled in TE samples prepared using disks cut from a PTFE rod. The disks were polarized, according to the scheme described above, by heat treatment at 613 K for 1 h.

The thermostimulated depolarization (TSD) current spectra were obtained by the method of thermal electret analysis, and the surface polarization charge density was measured by a static induction technique according to the State Standard GOST 25209-82.

The mechanical properties of TE samples were studied using methods stipulated by the State Standard GOST 11262-80. Data presented below represent the values averaged over not less than 20 independent measurements. The temperatures corresponding to an endothermal peak in the DTA curves were determined on a Q-1500 derivatograph. The ratio of crystalline and non-crystalline fractions in the samples was determined with the aid of a DRON-3.0 diffractometer.

SPONTANEOUS POLARIZATION CHARGE OF UDS PARTICLES

Measurements of the relative dielectric permittivity (ε_r) of the samples showed that UDS is "inferior" to the thermoplastic polymer binders with respect to insulating properties. Indeed, the permittivity of HDPE increases upon filling from the initial value $\varepsilon_r = 1.56$ to $\varepsilon = 2.25$ in the composition containing 20% UDS.

Polymeric samples formed by the conventional hot pressing method (without exposure to electric fields from external sources) were studied in the TSD regime. As seen in Fig. 1 showing typical TSD spectra, the initial HDPE (Fig. 1a) is electrically neutral. Even slight filling (0.5% UDS) results in the appearance of peaks in the current versus temperature plot (Fig. 1b). The HDPE samples containing 2 and 20% UDS (Figs. 1c and 1d, respectively) exhibit a steady-state charge flow resembling the current of conduction. The stable direction and significant magnitude of this charge allow this phenomenon to be considered as the conduction current. The current is most probably due to the motion of free charge carriers under the action of local electric fields generated by UDS particles. This is confirmed by the TSD spectrum of a tablet pressed (400 MPa) from the UDS powder (Fig. 1e). Here, the stable peaks observed at $T = 500$–570 K are evidence of the presence of the high-temperature traps of charge carriers. Apparently, there is a much greater number of these traps in the entire temperature range below 2000 K, and the state of carriers captured in these traps is more stable as compared to that in the low-temperature traps inherent in polymers.

If this hypothesis is valid, the spontaneous polarization charge of UDS particles would affect the polarizability of filled polymers, the adhesion of UDS particles to binder, and the crystallization of polymer in the vicinity of these particles, thus determining the structure and mechanical properties of filled plastic compositions.

POLARIZABILITY OF FILLED THERMOPLASTIC COMPOSITIONS

Figure 2 shows the TSD spectra of CE and TE samples based on the thermoplastic polymer binders studied. There are several features to be noted in these curves.

(1) The spectra of all CE samples (Figs. 2a, 2c, and 2e) exhibit the appearance of stable peaks due to the negative charge relaxation (in PVB, a double peak). These peaks correspond to the homocharge, that is, have the same sign as that induced by the polarizing corona discharge. As the temperature increases, the TSD currents cross the zero level (change sign) and

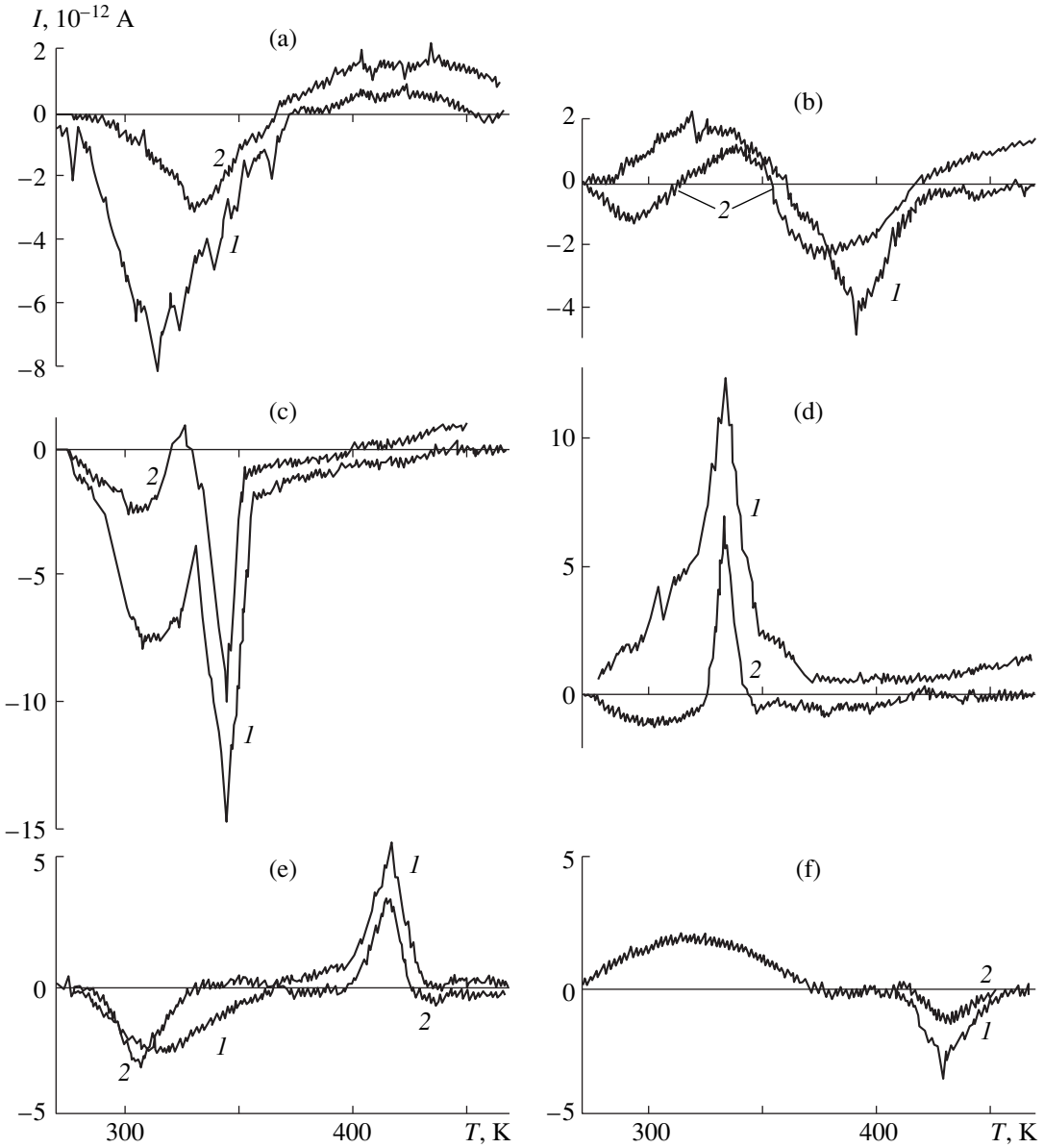


Fig. 2. The plots of TSD current versus temperature for (a, c, e) the crown electrets and (b, d, f) thermoelectrets based on (a, b) HDPE, (c, d) PVB, and (e, f) PTFE: (*1*) initial thermoplastic polymer; (*2*) the same polymer containing 0.5% UDS.

exhibit a smooth halo (in PTFE, a sharp peak) corresponding to relaxation of a positive space charge formed in the field of the homocharge.

(2) The TSD spectra of TE samples prepared from the initial thermoplastic polymers (Figs. 2b, 2d, and 2f, curves *1*) exhibit peaks or halos (T = 370–420 K) corresponding to relaxation of a positive heterocharge. This charge is generated by a mechanism analogous to that leading to the electric double layer formation. According to this, a negative potential is applied to the polarizing electrode and the positive charge is induced in the sample surface layer (predominantly by the relaxation polarization mechanism [4]). As the temperature increases, the TSD current crosses the zero level (changes sign) to reflect relaxation of the negative space charge induced in the bulk of the sample.

(3) The TE spectra of filled HDPE (Fig. 2b, curve *2*) and PVB (Fig. 2d, curve *2*) exhibit characteristic low-temperature (~300 K) peaks of a negative current corresponding to the Maxwell–Wagner polarization. This effect is related to displacement of the free charge carriers toward the thermoplastic binder–UDS particle interphase boundaries in the sample [4].

(4) The TSD currents observed in all filled compositions are lower as compared to the values in the corresponding unfilled binders. Apparently, a spontaneous charge trapped on the high-temperature traps in the course of the UDS formation is only weakly released

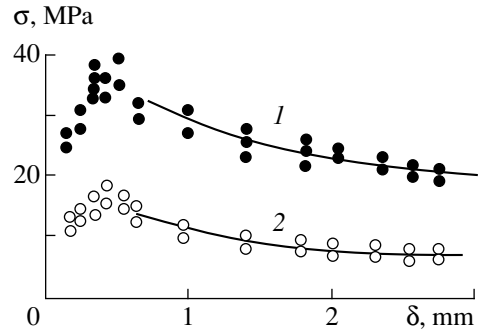


Fig. 3. The plots of (*1*) ultimate strength and (*2*) yield stress versus sample thickness for PTFE upon electric polarization.

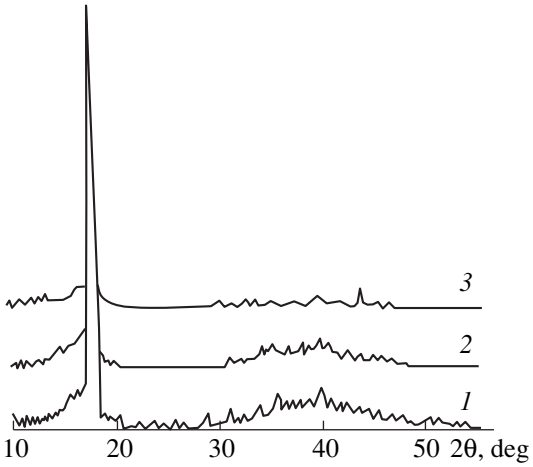


Fig. 4. X-ray diffractograms of PTFE samples: (*1*) initial; (*2*) heat treated; (*3*) electrically polarized during heat treatment.

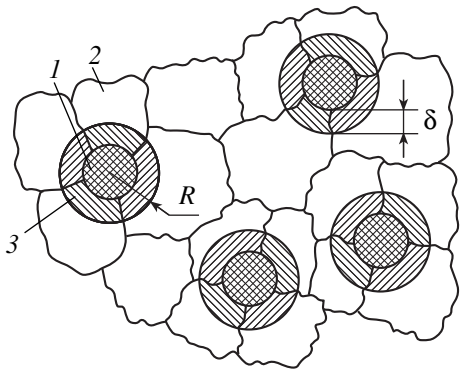


Fig. 5. Schematic diagram illustrating the structure of UDS-filled PTFE: (*1*) UDS particles; (*2*) PTFE particles; (*3*) boundary layer at the PTFE–UDS interface; *R* is the effective radius of action of the field of spontaneous polarization charge of UDS particles; δ is the boundary layer thickness.

during the TSD procedure in the temperature interval studied. This charge release contributes to the TSD current to a lower extent than does the leak of a polarization charge (induced by the external field) via the conductivity chains formed by UDS particles in the binder matrix.

Thus, no increase in the charge of electrets based on thermoplastic polymer binders was observed upon their filling with UDS particles. We may suggest that the mechanism of influence of the polarization charge of UDS particles on the properties of filled plastic compositions consists in the electric stimulation of physicochemical processes occurring at the polymer–UDS particle interphase boundaries during the composition formation. In electrets, these processes, determining the structure of transition layers and the mechanical properties of plastics, proceed in the field of the spontaneous charge of UDS particles.

MODEL OF STRENGTHENING OF UDS-FILLED PLASTICS

We have compared the structural parameters and mechanical properties of thermoplastic polymers cured in the presence of an electric field and without field. The results of comparative mechanical tests are presented in the table. As seen, the recrystallization of PTFE in the electric field leads to an increase in the ultimate tensile strength σ_u and the yield stress σ_y as compared to those of the control samples. The heat treatment alone somewhat decreases the strength compared to the initial level, probably because the annealing is accompanied by breakage of the oriented supermolecular structures present in the initial PTFE rods. The relative elongation at break ε was virtually the same in all samples.

The ultimate strength depends on the thickness δ of polarized samples (Fig. 3). Apparently, the polarization process results in the formation of a surface transition layer at the sample–electrode boundary, the strength of which is higher as compared to that of PTFE in the bulk. The maximum at δ = 400–500 μm indicates that the thickness of this strengthened layer is about δ/2 = 200–250 μm. Evidently, such a thin layer cannot insignificantly contribute to the elongation of "thick" samples at break.

Crystallization in the electric field leads to a change in the supermolecular structure of PTFE. The temperatures of characteristic points of the endothermal peak of melting in the DTA curves of samples listed in the table show that the melting temperature depends on the charged state of samples. A decrease in this temperature is evidence of the growing mobility of macromolecules in the crystalline phase of PTFE, which is equivalent to an increase in the flexibility of these macromolecules. This is most probably caused by weakening of the intermolecular bonds as a result of the electric polarization of macromolecules.

The X-ray diffractograms presented in Fig. 4 are typical of PTFE, exhibiting a "crystalline" peak (2θ ~ 18°) and a halo due to a noncrystalline phase (30°–60°).

Effect of heat treatment and electric polarization on the mechanical properties and the melting peak parameters of PTFE

Treatment	Mechanical properties			Melting peak, K		
	σ_u, MPa	σ_y, MPa	ε, %	onset	maximum	termination
Heat treatment + electric polarization	39.49	14.25	214	578	583	590
Heat treatment without polarization	29.10	11.73	203	587	593	601
Control (without treatment)	30.33	13.93	216	584	591	598

Note that the halo is most pronounced in the initial samples, decreases upon the thermal treatment, and virtually vanishes in the polarized samples. This is indicative of the transformation of the pseudocrystalline structures (typical of PTFE) into standard crystal lattice structures and of the removal of technological impurities from polarized samples.

CONCLUSION

The above data allow us to describe the mechanism of strengthening in the UDS-filled PTFE as follows. The plastic composite structure is formed in the course of pressing, whereby PTFE and UDS particles are brought into contact (Fig. 5). A very large difference in the melting temperature between UDS and PTFE (a few thousand Kelvin), virtually no relaxation of the spontaneous charge of UDS takes place during the formation of the filled plastic composition. In the vicinity of UDS particles, the PTFE melt occurs under the action of an electric field of the polarization charge of these particles. Provided that the effective radius of action of the field is R and the average distance between particles is less than $2R$, virtually all the polymer binder volume would be exposed to the electric field.

The field of UDS particles produces the main effect on the boundary layer of binder (with the thickness δ) in contact with the particles. The electric polarization of this layer may proceed by the Maxwell–Wagner mechanism, by the injection of charge carries from UDS, and by the relaxation polarization with a heterocharge formation [4]. As a result, the crystal structure in this boundary layer is modified and the strength of this layer, determining mechanical properties of the filled composition, is increased.

REFERENCES

1. T. N. Miller, Neorg. Mater. **15**, 557 (1979).
2. O. A. Andrianova, Ya. S. Semenov, Yu. V. Demidova, *et al.*, Mekh. Kompozit. Mater., No. 4, 599 (1991).
3. A. A. Okhlopkova, A. V. Vinogradov, A. P. Krasnov, *et al.*, Trenie Iznos **18** (1), 114 (1997).
4. *Electrets*, Ed. by G. Sessler (Springer, Berlin, 1987).

Translated by P.P. Pozdeev

Technical Physics, Vol. 45, No. 2, 2000, pp. 184–188. Translated from Zhurnal Tekhnicheskoĭ Fiziki, Vol. 70, No. 2, 2000, pp. 43–47.

SOLIDS

The Effect of Domain Wall Motion on the Nondiagonal Impedance Component in Amorphous Wires with Circular Magnetic Anisotropy

N. A. Buznikov, A. S. Antonov, and A. L. Rakhmanov

Scientific Center for Applied Problems in Electrodynamics, Russian Academy of Sciences, Moscow, 127412 Russia

Received October 26, 1998

Abstract—The effect of the domain wall motion on the nondiagonal impedance component in amorphous ferromagnetic wires with circular anisotropy was theoretically studied. The frequency spectrum of the electromotive force (emf) induced in the pick-up coil wound around the wire is analyzed. For sufficiently small amplitudes of the ac current passing in the wire, the emf frequency equals doubled frequency of the current. For the ac current amplitudes exceeding certain threshold value, all even harmonics appear in the frequency spectrum of the signal. The emf strongly depends on the longitudinal component of the applied magnetic field at any value of the ac current amplitude. These results can be important for developing frequency transducers controlled by magnetic field. © *2000 MAIK "Nauka/Interperiodica".*

INTRODUCTION

It is well known that the impedance of ferromagnetic wires usually depends only slightly on the applied magnetic field. However, recent studies of soft magnetic materials such as amorphous wires with circular anisotropy revealed the possibility of significant variations in the microwave impedance at low magnetic fields [1–3]. This phenomenon is referred to as giant magnetoimpedance. The large sensitivity of impedance to low magnetic fields is promising for the development of magnetic field sensors [4].

Another effect is observed when alternating current is passed along an amorphous wire to which a constant magnetic field is applied. In this case, the current flowing in the wire generates emf in the pick-up coil wound around the wire [5]. In a wire possessing circular anisotropy, the magnetic moment induced by the applied magnetic field has both longitudinal and transverse components with respect to the wire axis. This gives rise to an axial component of the ac (rf) magnetic moment with an amplitude strongly dependent on the applied magnetic field.

When the rf current magnitude in the wire is not too high (i.e., the amplitude of the magnetic field generated by the current at the wire surface is small compared to that of the anisotropic field), the emf induction in the coil is not accompanied by complete remagnetization of the sample. In this case, the emf value is proportional to a nondiagonal component of the wire impedance [5, 6]. Dependence of the nondiagonal impedance component on the wire parameters and applied magnetic field strength for a single-domain sample was studied in detail [6, 7]. A single-domain approximation is applicable to wires with the length below a certain critical value (of the order of several centimeters) [4]. However, long amorphous wires have a complicated domain structure [3, 4, 8], which should be taken into account in the calculations of impedance.

In this work, we have studied the effect of the domain wall motion on the nondiagonal impedance component in amorphous magnetically soft ferromagnetic wires with circular anisotropy. In the case of negligible skin effect, we derived explicit relationships for the emf induced in the coil wound around the wire. For sufficiently small amplitudes of the ac current flowing in the wire, the frequency of the coil output signal equals the doubled frequency of the current. At greater ac current amplitudes, all even harmonics appear in the frequency spectrum of the signal. It is also shown that, at any value of the ac current amplitude, the emf strongly depends on the applied magnetic field strength, provided this value does not exceed the anisotropic field characteristic of the wire.

NONDIAGONAL IMPEDANCE OF A MULTIDOMAIN WIRE

The distribution of easy anisotropic axes in ferromagnetic wires is determined mainly by the effect of magnetostriction and, hence, by the distribution of quench-induced strains arising during the manufacturing of these wires. Based on this picture, the magnetic properties of samples with small negative magnetostriction are usually described within the framework of a model assuming the existence of two regions in the wire: a central region (core) with the axial anisotropy and an outer region (shell) with the circular anisotropy [2–4]. The shell of a long wire has a domain structure

1063-7842/00/4502-0184$20.00

of "bamboo" [3, 8]. The ground state of such a multidomain system is characterized by opposite direction of the magnetization vectors in the neighboring domains along the φ axis, whereas in an external field H_e these vectors acquire equal components along the z axis (Fig. 1).

Let us assume for simplicity that the shell of the wire contains only one domain wall. The obtained results are easily extended to the case of a multidomain sample. We shall also neglect the effect of the core, since its dimensions are usually small, in wires with the diameter below 50 μm [4]. Let us consider a wire with length l, radius a, and possessing a circular anisotropy. In the applied magnetic field H_e directed along the wire axis, the magnetization vectors of domains rotate through an angle θ with respect to the anisotropy axis (Fig. 1). The value of this angle can be found from the condition of minimum free energy [9]:

$$\sin\theta = H_e/H_A, \quad H_e < H_A, \qquad \sin\theta = 1, \quad H_e \geq H_A, \tag{1}$$

where H_A is the anisotropy field of the wire.

Alternating current $I(t) = I_0\sin(\omega t)$ flowing in the wire induces emf in the pick-up coil wound coaxially to the wire. This emf is proportional to the nondiagonal component of the sample impedance [5, 6]. The effect is related to the fact that the magnetization vector has both longitudinal and transverse components with respect to the wire axis. Using a general expression written in the form of series [7], we obtain a formula for the nondiagonal impedance component $Z^0_{\varphi z}$ for a single-domain sample in the absence of skin effect:

$$Z^0_{\varphi z} = -i4\pi\omega a\tilde{\mu}\cos\theta\sin\theta/3c^2. \tag{2}$$

Here,

$$\tilde{\mu} = \frac{4\pi\gamma M_s}{\gamma H_Z\cos^2\theta - i\alpha\omega - \omega^2/4\pi\gamma M_s} \tag{3}$$

is the effective magnetic permeability of the wire [6], M_s is the saturation magnetization, γ is the gyromagnetic ratio, and α is the dissipation parameter in the Landau–Lifshitz equation characterizing [9, 10]. It follows from (1) and (2) that the nondiagonal impedance component is nonzero if the applied constant field H_e magnetizing the sample is lower than the anisotropic field.

Let us consider a two-domain wire with the domain wall located in the middle of the sample. The symmetry of the problem suggests that φ components of the electric field have opposite directions in different domains. As a result, the total emf V_φ induced in the coil vanishes and becomes nonzero only when the domain wall is displaced from its position in the middle of the sample. Then, an ac current flowing in the sample generates the emf V_φ in the coil, which is related not only to rotation

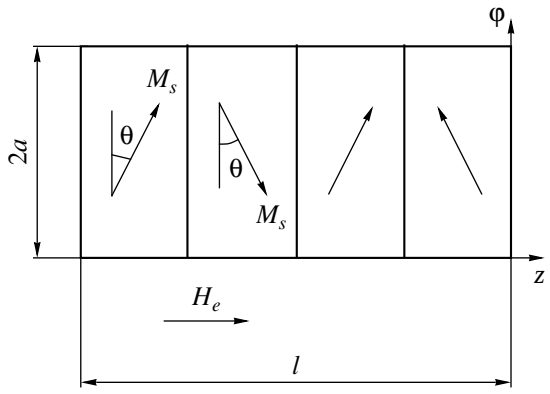


Fig. 1. Schematic diagram of the magnetization distribution in a multidomain wire with circular magnetic anisotropy.

of the magnetization vector but to the domain wall motion as well. Upon neglect of the magnetization vector rotation induced by the current variations, we obtain the following expression for the nondiagonal impedance component $Z_{\varphi z}(t) = Z^0_{\varphi z}2z(t)/l$, where $z(t)$ is the time-dependent position of the domain wall relative to the middle of the sample. Thus, the time dependence of the emf V_φ induced in the coil is given by the following relationship:

$$V_\varphi(t) = \mathrm{Re}\{Z_{\varphi z}(t)I(t)\} = (2/l)\mathrm{Re}\{Z^0_{\varphi z}z(t)I(t)\}. \tag{4}$$

DOMAIN WALL MOTION UNDER THE ACTION OF ALTERNATING MAGNETIC FIELD

The domain wall motion under the action of a var magnetic field generated by the alternating current is described by the following equation [9, 10]:

$$m\frac{d^2z}{dt^2} + \beta\frac{dz}{dt} + f(z) = 2M_sH_\varphi(\rho, t)\cos\theta. \tag{5}$$

Here, $m = (K_1/A)^{1/2}(m_ec/e)^2/2\pi\cos^2\theta$ is the domain wall mass per unit area (K_1 is the anisotropy constant and A is the exchange constant), β is the coefficient of friction proportional to the ac losses related to eddy currents [10], $f(z)$ is the restoring force, $2M_sH_\varphi(\rho, t)\cos\theta$ corresponds to the pressure acting on the domain wall, $H_\varphi(\rho, t)$ is the magnetic field induced by the ac current, which depends on the radial coordinate ρ. For the values of parameters typical of the Fe- and Co-based magnetically soft wires ($K_1 \cong 250$ erg cm^{-3}, $A \cong 10^{-8}$ erg cm^{-1}, $M_s \cong 500$ G) and $\cos\theta \cong 1$, the mass of the domain wall is very small, $m \approx 10^{-10}$ g cm^{-2}. This circumstance allows us to neglect the first term in (5) at frequencies up to a value of the order of several tens of gigahertz.

In the absence of skin effects, $H_\varphi(\rho, t)$ varies linearly across the sample cross section from zero at $\rho = 0$ up to

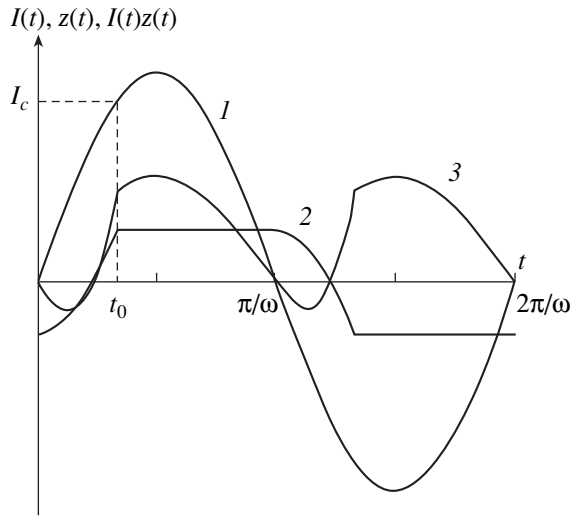


Fig. 2. Time variation of (*1*) $I(t)$, (*2*) $z(t)$, and (*3*) $I(t)z(t)$.

$2I(t)/ac$ at $\rho = a$. For the quantitative description of the domain wall motion, let us replace $H_\varphi(\rho, t)$ in (5) by the value $\langle H_\varphi(t)\rangle$ averaged over the cross section:

$$H_\varphi(\rho, t) \approx \langle H_\varphi(t)\rangle = \frac{2}{a^2}\int_0^a \frac{2I(t)\rho^2}{ca^2}d\rho = \frac{4I(t)}{3ca}. \quad (6)$$

The restoring force in equation (5) describing the domain wall motion is related either to variation of the total magnetization energy in the wire due to changes of the domain size or to the interaction of the domain wall with defects [11]. Let us assume the wire to be homogeneous and consider first the domain wall motion in the case when it is possible to neglect the restoring force action. This approximation is valid at sufficiently high frequencies and small values of the current amplitude, when the domain wall displacement from the equilibrium position is small and does not cause any significant changes in the magnetostatic energy of the sample. Neglecting the restoring force in equation (5) and taking into account that $\beta/m \gg \omega$, we obtain the following relationship from (5) and (6):

$$z(t) = -\frac{8M_sI_0\cos\theta}{3\beta\omega ca}\cos(\omega t) + z_0. \quad (7)$$

In the approximation under study, the domain wall motion is symmetric with respect to the middle of the wire. Therefore, the constant z_0 in (7) must be equal to zero. Substituting (7) with $z_0 = 0$ into (4), we arrive at the following expression for the emf V_φ induced in the coil:

$$V_\varphi(t) = \frac{32\pi M_s\cos^2\theta\sin\theta}{9c^3l\beta}\tilde{\mu}I_0^2\sin(2\omega t + \delta - \pi), \quad (8)$$

where

$$\delta = \arctan\{\gamma H_A\cos^2\theta/\alpha\omega - \omega/4\pi\gamma\alpha M_s\}. \quad (9)$$

Thus, at small values of the ac current amplitude, the emf in the pick-up coil is proportional to a square of the current amplitude and the emf frequency is equal to doubled frequency of the current. Since the value of $4\pi\gamma M_s$ is sufficiently large for systems under study (it is usually assumed that $4\pi\gamma M_s \cong 10^{-11}$ s^{-1} and $\gamma H_A \cong 10^{-8}$ s^{-1}), we obtain from (1) and (9) that

$$\delta = \arctan\{\gamma(H_A^2 - H_c^2)/\alpha\omega H_A\}. \quad (10)$$

It follows from (8) and (10) that the phase of V_φ varies with the field strength H_e and frequency ω from $-\pi$ (high fields and frequencies) to $-\pi/2$ (low fields and frequencies).

The magnetostatic energy of the sample grows with the amplitude of domain wall oscillations. This effect causes the growth of restoring force $f(z)$ in (5) and leads to the corresponding restrictions on the freedom of domain wall displacements. Let us consider now the qualitative effects produced by the growth of restoring force $f(z)$ on the domain wall motion and on the emf generated in the pick-up coil. We shall describe the effect of restoring force within the framework of the following simplest model. Let the restoring force be negligibly small when the current grows, unless the domain wall displacement attains certain critical value z_c. At $z = z_c$, the domain wall is affected by a large restoring force hindering its further displacement. Let the wall remain immobile unless the current changes its sign. After the change in the sign of current, the restoring force vanishes, and the wall starts its motion in the opposite direction according to (7) (with $z_0 \neq 0$) until attaining the position $-z_c$. Thus, the time dependence of the domain wall position within the framework of this model is as follows:

$$z(t) = \begin{cases} \pm z_c, \ |I| > I_c, \\ -\dfrac{8M_sI_0\cos\theta}{3\beta\omega ca}[1 + \cos(\omega t)] + z_c, \ |I| < I_c, \ \dfrac{dI}{dt} < 0, \\ \dfrac{8M_sI_0\cos\theta}{3\beta\omega ca}[1 - \cos(\omega r)] - z_c, \ |I| < I_c, \ \dfrac{dI}{dt} > 0, \end{cases} \quad (11)$$

where I_c is the value of current at which the domain wall attains the position z_c.

A relationship between z_c and I_c is determined by the following obvious formulas:

$$z_c = \frac{4M_sI_0\cos\theta}{3\beta\omega ca}[1 - \cos\tau_0], \qquad \tau_0 = \arcsin(I_c/I_0). \quad (12)$$

The $I(t)$ and $z(t)$ curves are shown in Fig. 2. This figure also illustrates the behavior of their product $I(t)z(t)$,

which determines the frequency spectrum of the emf V_φ according to equation (4). Function $I(t)z(t)$ is neither even nor odd and has a period half as small as that of the field H_φ induced by the ac current. Hence, the Fourier expansion of function $I(t)z(t)$ should contain a constant component, the main harmonic with the frequency 2ω, and harmonics with the frequencies multiple of 2ω. Using (11) and (12), we obtain the following relationship for the function $I(t)z(t)$:

$$I(t)z(t) = \frac{8M_s\cos\theta}{3\pi\beta\omega ca}I_0^2 \times\left[\frac{a_0}{2} + a_2\cos(2\omega t) + b_2\sin(2\omega t) + \sum_{k=2}^{\infty} a_{2k}\cos(2k\omega t) + b_{2k}\sin(2k\omega t)\right], \quad (13)$$

where the coefficients in the Fourier expansion have the form

$$a_0 = (I_c/I_0)^2, \quad a_2 = -(I_c/I_0)^2[1 + (I_c/I_0)^2]/3,$$

$$b_2 = -\tau_0/2 + \sin(4\tau_0)/8 + 4\cos(\tau_0)\sin^3(\tau_0)/3,$$

$$a_{2k} = \frac{3}{2(k^2-1)(4k^2-1)} + \frac{\cos\{2(k+1)\tau_0\}}{4(k+1)} - \frac{\cos\{2(k-1)\tau_0\}}{4(k-1)} + \cos(\tau_0)\left\{\frac{\cos\{(2k-1)\tau_0\}}{2k-1} - \frac{\cos\{(2k+1)\tau_0\}}{2k+1}\right\}, \quad (14)$$

$$b_{2k} = \frac{\sin\{2(k+1)\tau_0\}}{4(k+1)} - \frac{\sin\{2(k-1)\tau_0\}}{4(k-1)} + \cos(\tau_0)\left\{\frac{\sin\{(2k-1)\tau_0\}}{2k-1} - \frac{\sin\{(2k+1)\tau_0\}}{2k+1}\right\}.$$

Substituting this Fourier expansion of the function $I(t)z(t)$ into (4), we obtain an expression for the emf in the coil:

$$V_\varphi(t) = \frac{64M_s\cos^2\theta\sin\theta}{9c^2l\beta}I_0^2 \times\left[\tilde{\mu}''\frac{a_0}{2} + |\tilde{\mu}|\sum_{k=1}^{\infty} c_{2k}\sin(2k\omega t + \varphi_{2k} + \delta)\right], \quad (15)$$

where $\tilde{\mu}''$ is the imaginary part of the effective magnetic permeability, $c_{2k} = (a_{2k}^2 + b_{2k}^2)^{1/2}$, and $\varphi_{2k} = \arctan(a_{2k}/b_{2k})$.

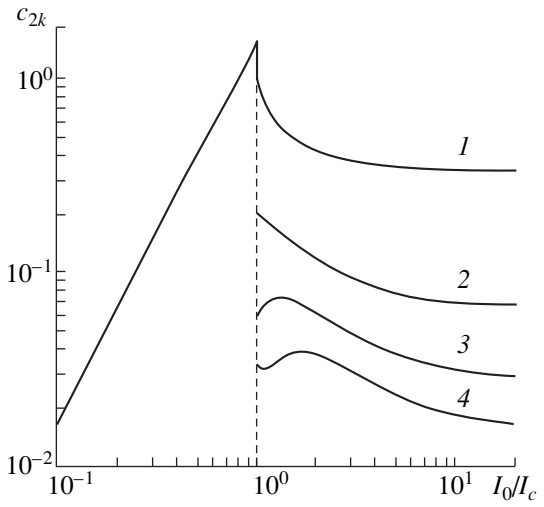


Fig. 3. Amplitudes of the first harmonics c_{2k} in the emf signal as functions of the ac current amplitude I_0: $k = 1$ (*1*), 2 (*2*), 3 (*3*), and 4 (*4*).

Figure 3 shows the plots of dimensionless amplitudes c_{2k} versus the current amplitude for the first four even harmonics of the emf signal. As seen from Fig. 3 for the model under discussion, the signal amplitude is independent of the current amplitude for sufficiently high values of I_0. At $I_0 \gg I_c$ ($\tau_0 \longrightarrow 0$), we obtain from (14) and (15) after simple transformations

$$V_\varphi(t) = \frac{64M_s\cos^2\theta\sin\theta}{9c^3l\beta}I_c^2 \times\left[\frac{\tilde{\mu}''}{2} + |\tilde{\mu}|\sum_{k=1}^{\infty}\frac{1}{4k^2-1}\sin(2k\omega t + \delta - \pi/2)\right]. \quad (16)$$

It is obvious that all even harmonics also appear in the frequency spectrum of V_φ at any arbitrary value of the restoring force $f(z)$. The magnitude of these harmonics becomes appreciable if the ac current amplitude is sufficiently large.

CONCLUSION

Thus, the domain wall motion in a ferromagnet with circular anisotropy produces a pronounced effect on the nondiagonal impedance component and on the emf induced in the pick-up coil wound around the wire. The signal in the coil arises if the applied constant magnetic field strength H_e is smaller than that of the anisotropic field H_A. The domain wall motion leads to qualitative changes in the emf frequency spectrum. At small amplitudes I_0 of the ac current, the emf spectrum contains only the second harmonic with an amplitude proportional to I_0^2. At larger ac current amplitudes, all even harmonics appear in the emf spectrum (Fig. 3). Within the whole range of the ac current amplitudes, the emf

strongly depends on the applied magnetic field strength H_e.

As was mentioned in [6], the signal V_φ induced in the coil cannot vanish at $H_e > H_A$ if the anisotropy axis deviates from the circular direction. This deviation can arise either due to the strains induced in the course of processing and twisting of the wire, or due to the stray magnetic field of the wire core (which produces an additional magnetization of the shell). This deviation of the anisotropy axis from the circular direction leads to smoothening of the $V_\varphi(H_e)$ curve at H_e close to H_A and to the hysteresis at $H_e > H_A$, which is observed both in a single-domain sample [6] and in a multidomain wire.

In conclusion, let us extend the obtained results to the case of samples with an arbitrary number of domain walls. Let N be the number of these walls. The circular emf calculated above is related to the motion of a single domain wall. It is easy to understand that, within the framework of approximations discussed above, the emf produced by the motion of N domain walls can be determined simply by multiplication of V_φ by N in equations (4), (8), (15), and (16).

ACKNOWLEDGMENTS

The authors are grateful to I.T. Yakubov for helpful discussions.

The work was supported in by the Russian Foundation for Basic Research, project no. 98-02-16827.

REFERENCES

1. K. Mohri, F. B. Humphrey, K. Kawashima, *et al*., IEEE Trans. Magn. **26**, 1789 (1990).
2. K. Mohri, T. Kohzawa, K. Kawashima, *et al*., IEEE Trans. Magn. **28**, 3150 (1992).
3. R. S. Beach and A. E. Berkowitz, Appl. Phys. Lett. **64**, 3652 (1994).
4. M. Vázquez and A. Hernando, J. Phys. D: Appl. Phys. **29**, 939 (1996).
5. A. Antonov, I. Iakubov, and A. Lagarkov, IEEE Trans. Magn. **33**, 3367 (1997).
6. A. S. Antonov, I. T. Iakubov, and A. N. Lagarkov, J. Magn. Magn. Mater. **187**, 252 (1998).
7. N. A. Usov, A. S. Antonov, and A. N. Lagarkov, J. Magn. Magn. Mater. **185**, 159 (1998).
8. K. Mohri, K. Kawashima, T. Kohzawa, *et al*., IEEE Trans. Magn. **29**, 1245 (1993).
9. A. G. Gurevich and G. A. Melkov, *Magnetic Oscillations and Waves* (Nauka, Moscow, 1994).
10. S. Chikazumi, *Physics of Ferromagnetism. Magnetic Characteristics and Applications* (Kabushiki Kaisha Syokabo, Tokyo, 1984; Mir, Moscow, 1987).
11. N. Usov, A. Antonov, A. Dykhne, *et al*., J. Magn. Magn. Mater. **174**, 127 (1997).

Translated by K.I. Kugel

Technical Physics, Vol. 45, No. 2, 2000, pp. 189–193. Translated from Zhurnal Tekhnicheskoĭ Fiziki, Vol. 70, No. 2, 2000, pp. 48–52.

Parametric Interaction of Space-Charge Waves in Thin-Film Semiconductor Structures

A. A. Barybin[1] and A. I. Mikhaĭlov[2]

[1] *St. Petersburg University of Electrical Engineering, St. Petersburg, 197376 Russia*

[2] *Saratov State University, Saratov, 410071 Russia*

Received July 15, 1998

Abstract—The general theory of parametric coupling between space-charge waves and drifting charge carriers in thin-film semiconductor structures has been worked out. This theory is applicable, in particular, to *n*-GaAs and *n*-InP semiconductors with negative differential conductance due to intervalley electron transitions under high electric fields. We started from the electrodynamic theory of waveguide excitation by extraneous currents, which was extended for arbitrary waveguide structures with composite active media. Our theory makes it possible to study parametric interaction between space-charge waves in semiconductor films with regard for boundary conditions, diffusion, the anisotropy and the frequency dispersion of the differential electron mobility, as well as the multifrequency and multimode nature of a wave process in thin-film structures. © *2000 MAIK "Nauka/Interperiodica".*

INTRODUCTION

Among thin-film semiconductor structures (TFSS), the ones with negative differential conductance (NDC), which occurs in semiconductors such as *n*-GaAs and *n*-InP, are of particular practical interest. This is due to the specific features of space charge wave (SCW) propagation in these semiconductors when electrons are heated in high electric fields. These structures can be employed in different integrated signal-processing devices with up to a mm-wave band [1–6]. For instance, a thin-film traveling-wave amplifier (TFTWA) based on an *n*-GaAs film fulfills functions such as amplification, generation, signal delay and phase shifting, channel switching, etc., in the microwave range [1–3]. More sophisticated designs where the parametric modes (including that with low-frequency pumping) are used make possible controllable filtering, frequency conversion, and frequency synthesis [4–6].

The propagation of eigenwaves in TFSS with NDC has been extensively studied [2, 3, 7, 8]. However, the parametric interaction of SCWs has not been adequately covered (see, e.g., [4, 6, 9–13]). Because of a number of starting simplifying assumptions made in these publications, the obtained results cannot be considered general. For instance, a one-dimensional description [9–11, 13] or one that can be called two-dimensional by convention [4, 6] were used, the diffusion component of the current and the frequency dispersion of the differential conductance of a medium [11, 12] were ignored, etc. In all of these works, analysis of the parametric interaction of SCWs has been carried out only for a single pair of frequencies (signal and idler) and, therefore, is of limited application.

In this paper, we elaborated the general theory of SCW parametric interaction in semiconducting thin films with consideration for the multifrequency and multimode nature of the wave process, the diffusion component of the current, the frequency dispersion of the hot-electron differential mobility, and the NDC conditions.

MAXWELL EQUATIONS AND PARAMETRICALLY EXCITING CURRENT

In macroscopic electrodynamics, the wave processes in a semiconductor plasma are described by the usual Maxwell equations

$$\nabla \times \tilde{\mathbf{E}} = -\mu_0 \frac{\partial \tilde{\mathbf{H}}}{\partial t}, \quad \nabla \times \tilde{\mathbf{H}} = \varepsilon \frac{\partial \tilde{\mathbf{E}}}{\partial t} + \tilde{\mathbf{J}}, \tag{1}$$

where the current density is given in the form [7, 14]

$$\tilde{\mathbf{J}} = \tilde{\rho}\tilde{\mathbf{v}} - D\nabla\tilde{\rho}, \tag{2}$$

which takes into consideration the drift ($\tilde{\rho}\tilde{\mathbf{v}}$) and diffusion ($-D\nabla\tilde{\rho}$) components. Hereafter, a tilde over the physical quantities stands for their real instantaneous values in contrast to the complex amplitudes, which are free of this symbol. Any physical quantity $\tilde{\Phi}(\mathbf{r}, t)$ will have the subscripts 0 and 1 to separate its steady and variable components:

$$\tilde{\Phi}(\mathbf{r}, t) = \Phi_0(\mathbf{r}) + \tilde{\Phi}_1(\mathbf{r}, t), \tag{3}$$

where $\mathbf{r}$ is a radius vector, which locates a point in the three-dimensional coordinate system, and t is time.

1063-7842/00/4502-0189$20.00

Since we consider the parametric interaction, the component $\tilde{\Phi}_1(\mathbf{r}, t)$ contains the pumping and signal contributions, marked, respectively, by the subscripts p and s:

$$\tilde{\Phi}_1(\mathbf{r}, t) = \tilde{\Phi}_p(\mathbf{r}, t) + \tilde{\Phi}_s(\mathbf{r}, t). \quad (4)$$

Within a linear approximation, as usual, the variable quantities and the signal are assumed to be small compared, respectively, to the steady quantities and the pump.

Let a pump SCW with a frequency ω_p propagate in a stream of electrons that drift with a velocity v_0; this results in spatial and time modulations of the semiconductor parameters. The excitation of a weak (relative to the pump wave) SCW signal with a frequency ω_s gives rise to a multifrequency spectrum (a wave packet of parametrically coupled SCWs) with combination frequencies $\omega_\nu = \omega_s + \nu\omega_p$ ($\nu = 0, \pm1, \pm2, \ldots$). As known [7], their related waves are of the drift nature and propagate with the same phase velocity, equal to the electron drift velocity $\mathbf{v}_0$.

Introducing complex amplitudes, one can represent the physical parameters of the one-frequency pump and the multifrequency signal spectrum as

$$\tilde{\Phi}_p(\mathbf{r}, t) = 2\,\mathrm{Re}\{\Phi_p(\mathbf{r})e^{i\omega_p t}\}$$

and

$$\tilde{\Phi}_s(\mathbf{r}, t) = \mathrm{Re}\sum_\nu\{\Phi_1^{(\nu)}(\mathbf{r})e^{i\omega_\nu t}\}, \quad (5)$$

where $2\Phi_p$ is the complex amplitude of $\tilde{\Phi}_p$ for the pump SCW and $\Phi_1^{(\nu)}$ is that of $\tilde{\Phi}_s$ for the signal component with a frequency ω_ν (the pump amplitude is doubled to simplify subsequent expressions).

In view of (3), equations (1) for the variable components take the form

$$\nabla\times\tilde{\mathbf{E}}_1 = -\mu_0\frac{\partial\tilde{\mathbf{H}}_1}{\partial t}, \quad \nabla\times\tilde{\mathbf{H}}_1 = \varepsilon\frac{\partial\tilde{\mathbf{E}}_1}{\partial t} + \tilde{\mathbf{J}}_1 + \tilde{\mathbf{J}}_b, \quad (6)$$

where $\tilde{\mathbf{R}}_1$, $\tilde{\mathbf{H}}_1$, and $\tilde{\mathbf{J}}_1$ have the form of (4).

From general expression (2) for the current density, it follows that $\tilde{\mathbf{J}}_1$ in (6) should be supplemented by a parametrically exciting current

$$\tilde{\mathbf{J}}_b = \mathrm{Re}\sum_\nu\mathbf{J}_b^{(\nu)}e^{i\omega_\nu t} \quad (7)$$

with the complex amplitudes

$$\mathbf{J}_b^{(\nu)} = (\rho_p\mathbf{v}_1^{(\nu-1)} + \rho_1^{(\nu-1)}\mathbf{v}_p) + (\rho_p^*\mathbf{v}_1^{(\nu+1)}\rho_1^{(\nu+1)}\mathbf{v}_p^*) \quad (8)$$

and frequencies ω_ν.

Taking into account (5), (7), and the orthogonality of the frequency components, we separate from (6) the equations

$$\nabla\times\mathbf{E}_p = -i\omega_p\mu_0\mathbf{H}_p, \quad \nabla\times\mathbf{H}_p = i\omega_p\varepsilon\mathbf{E}_p + \mathbf{J}_p \quad (9)$$

for the pump complex amplitudes with a frequency ω_p, where

$$\mathbf{J}_p = \rho_0\mathbf{v}_p + \rho_p\mathbf{v}_0 - D\nabla\rho_p \quad (10)$$

and

$$\nabla\times\mathbf{E}_1^{(\nu)} = -i\omega_\nu\mu_0\mathbf{H}_p^{(\nu)},$$
$$\nabla\times\mathbf{H}_1^{(\nu)} = i\omega_\nu\varepsilon\mathbf{E}_p^{(\nu)} + \mathbf{J}_1^{(\nu)} + \mathbf{J}_b^{(\nu)} \quad (11)$$

for the complex amplitudes of the signal components with frequencies ω_ν, where

$$\mathbf{J}_1^{(\nu)} = \rho_0\mathbf{v}_1^{(\nu)} + \rho_1^{(\nu)}\mathbf{v}_0 - D\nabla\rho_1^{(\nu)}. \quad (12)$$

As is seen from (9) and (10), the equations for the pump SCW have the conventional form, which is free of the exciting current and is applied, in particular, for analysis of the eigenwave spectrum in TFSS [7]. Because of this, we think of the boundary problem for the pump wave as being solved and will apply, when needed, the known results [7].

Equations (11) and (12) describe the parametric excitation of the SCW modes by $\mathbf{J}_b^{(\nu)}$, which acts as an extraneous current. At a frequency ω_ν, the parametrically exciting current $\mathbf{J}_b^{(\nu)}$, written as (8) and entering into equation (11), is defined both by the pump wave (ρ_p, $\mathbf{v}_p$) and by the adjacent frequency components ($\rho_1^{(\nu\pm1)}$, $\mathbf{v}_1^{(\nu\pm1)}$). This provides the parametric relation between the SCW modes with different frequencies. In the absence of pumping ($\mathbf{J}_b^{(\nu)} = 1$), the problem of excitation by an extraneous current [see (11) and (12)] transforms into the problem of eigenvalues. The solution of the latter is also known [7], as in the case of a pump SCW.

EQUATIONS FOR THE AMPLITUDES OF PARAMETRICALLY COUPLED SCW MODES

Parametric interaction of SCW modes in TFSS will be considered within the well-known electrodynamic theory of waveguide excitation by given currents [15]. This theory was generalized in [16, 17] for thin-film waveguide structures with different active media, including TFSS with NDC.

Let a semiconductor film with a lattice permittivity ε be sandwiched in insulators with permittivities ε_b and ε_H. In the case of multilayer structures, these values are known [7, 18] to characterize the effective (generally frequency-dependent) permittivities of the layers above and under the semiconductor film. Let the y-axis of our

coordinate system be normal to the film surface and the *xz* plane be parallel to the film surface and bisect the film. Then, the *z*-axis is directed lengthwise of the thin-film waveguide structure and coincides with the direction of the electron drift, which depends on the static electric field; that is, $\mathbf{v}_0 = \mathbf{e}_z v_0$, where $\mathbf{e}_z$ is the unit vector of the *z*-axis. We will consider that the film is of a finite thickness (along the *y*-axis) and is unbounded along the *x*- and *z*-axes.

In the semiconductor film, the electron drift velocity v depends on the electric field strength E. For the given TFSS geometry, a high electric field E_0 aligned with the *z*-axis results in the anisotropy of electron differential mobility [7]. Namely, in the drift direction, the electron small-signal mobility is defined by the differential mobility $\mu_d = (dv/dE)|_{E_0}$, while in the transverse directions, by the static mobility $\mu_e = v_0/E_0$. Mathematically, this can be described by the small-signal mobility tensor

$$\overleftrightarrow{\mu}_d = \mu_e \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & \kappa \end{pmatrix}, \qquad (13)$$

where $\kappa = \mu_d/\mu_e$ is the anisotropy factor.

In the case of *n*-GaAs and *n*-InP subjected to high electric fields, the field dependence of the electron drift velocity has a descending portion. It is such a behavior that provides NDC for the longitudinal (along the field) motion of electrons ($\kappa < 0$). It is also known that, at high frequencies (in mm- and sub-mm-wave bands), the anisotropy depends on the frequency [19–22]. Hence, for the pump and signal frequencies, one should discriminate between the anisotropy factors $\kappa^{(p)} \equiv \kappa(\omega_p)$ and $\kappa^{(v)} \equiv \kappa(\omega_v)$, as well as between the differential-mobility tensors $\overleftrightarrow{\mu}_d^{(p)} \equiv \overleftrightarrow{\mu}_d(\omega_p)$ and $\overleftrightarrow{\mu}_d^{(v)} \equiv \overleftrightarrow{\mu}_d(\omega_v)$.

The quasi-static eigen-SCWs propagating in a TFSS has an infinitely large number of modes, which differ in longitudinal propagation constants and transverse wavenumbers [7]. As shown in [7], under the NDC conditions (that is, when $\kappa < 0$), only the trigonometric modes, both building up and damped, exist in a TFSS. The influence of electrophysical parameters of the semiconductor film, as well as of the anisotropy of differential mobility and electron diffusion, on the structure and dispersion of SCW eigenmodes has been analyzed in detail in [7]. The effect of frequency dispersion of the electron differential mobility in *n*-GaAs and *n*-InP on the propagation characteristics of the eigen-SCWs in the TFSS has been investigated in [23]. Thus, the structure and dispersion of the eigenmodes of SCWs in TFSS with NDC are known.

Let only the SCW principal mode propagate at a pumping frequency in the longitudinal direction. Assume that we know all the physical parameters of the structure, including

$$\rho_p(\mathbf{r}_t, z) = \hat{\rho}_p(\mathbf{r}_t) e^{-\gamma_p z}$$

and

$$\mathbf{v}_p(\mathbf{r}_t, z) = \hat{\mathbf{v}}_p(\mathbf{r}_t) e^{-\gamma_p z}, \qquad (14)$$

where $\gamma_p = \alpha_p + i\beta_p \approx \alpha_p + i\omega_p/v_0$ is the longitudinal propagation constant of the principal mode of the pump SCW; so-called membrane functions, descriptive of the physical parameter traverse with a radius vector $\mathbf{r}_t$, are marked by the cap.

For any *k*th mode of the SCW signal component with a frequency ω_v, its propagation constant

$$\gamma_k^{(v)} = \alpha_k^{(v)} + i\beta_k^{(v)} \approx \alpha_k^{(v)} + i\omega_v/v_0 \qquad (15)$$

and the related membrane functions $\tilde{\Phi}_k^{(v)}(\mathbf{r}_t)$ are considered to be known; that is,

$$\Phi_k^{(v)}(\mathbf{r}_t, z) = \tilde{\Phi}_k^{(v)}(\mathbf{r}_t) e^{-\gamma_k^{(\mu)} z}; \quad k = 0, 1, 2, \ldots, \qquad (16)$$

where $\Phi_k^{(v)}$ stands for dynamic variables like $\rho_k^{(v)}$, $\mathbf{v}_k^{(v)}$, $\mathbf{E}_k^{(v)}$, $\mathbf{H}_k^{(v)}$, etc.

Since the eddy electric field in the SCW modes is negligibly small, $\mathbf{E}_k^{(v)} \approx -\nabla\varphi_k^{(v)}$, where $\varphi_k^{(v)}$ is the quasi-static potential of the *k*th mode; at the same time, the eddy magnetic field $\mathbf{H}_k^{(v)} \neq 0$ [7]. With regard to (16), the sought physical quantities in the region of parametric coupling between the SCW modes can be represented as expansions in modes [16, 17]:

$$\Phi_1^{(v)}(\mathbf{r}_t, z) = \sum_k A_k^{(v)}(z)\Phi_k^{(v)}(\mathbf{r}_t, z) = \sum_k A_k^{(v)}(z)\hat{\Phi}_k^{(v)}(\mathbf{r}_t) e^{-\gamma_k^{(v)} z}, \qquad (17)$$

where the excitation amplitudes $A_k^{(v)}(z)$ are yet to be known. To find them, we will take advantage of equations of excitation written as [17]

$$\sum_m N_{km}^{(v)} \frac{dA_m^{(v)}(z)}{dz} e^{-\gamma_m^{(v)} z} = -\int_S \mathbf{J}_b^{(v)} \cdot \hat{\mathbf{E}}_k^{(v)*} dS, \qquad (18)$$

where $m = 0, 1, 2, \ldots$ and $\mathbf{J}_b^{(v)}$ is the complex amplitude of the parametrically exciting extraneous current [see (8)] with a frequency ω_v.

At $k = m$, the normalization factors $N_{km}^{(v)}$ determine the power eigenvalue of the *k*th mode and at $k \neq m$, the power transferred by the *k*th and *m*th modes in a dissi-

pative system which a semiconductor plasma is [17]. Within the weak-diffusion approximation [7], where the diffusion mechanism of energy transfer by electrons can be neglected, the normalization factors

$$N_{km}^{(\nu)} \approx \int_S \hat{\mathbf{E}}_k^{(\nu)} \times \hat{H}_m^{(\nu)} + \hat{\mathbf{E}}_m^{(\nu)} \times \mathbf{H}_k^{(\nu)*}) \cdot \mathbf{e}_z dS \qquad (19)$$

were calculated in [24] for highly asymmetric TFSS with $|\varepsilon_b| \ll |\varepsilon_H|$ and $|\varepsilon_b| \gg |\varepsilon_H|$. The excitation integral on the right of equations (18) can be transformed with the use of expression (8) for the excitation current $\mathbf{J}_b^{(\nu)}$ and the mode expansions similar to (17) for $\rho_1^{(\nu \pm 1)}$ and $\mathbf{v}_1^{(\nu \pm 1)}$. This transformation yields the desired set of equations for the excitation amplitudes of the parametrically coupled SCW modes with frequencies ω_ν and $\omega_{\nu \pm 1}$:

$$\sum_m N_{km}^{(\nu)} \frac{dA_m^{(\nu)}(z)}{dz} e^{-\gamma_m^{(\nu)} z} = \sum_n C_{kn}^{(\nu, \nu-1)} e^{-(\gamma_n^{(\nu-1)} + \gamma_p) z} \times A_n^{(\nu-1)}(z) + \sum_n C_{kn}^{(\nu, \nu+1)} e^{-(\gamma_n^{(\nu+1)} + \gamma_p^*) z} A_n^{(\nu+1)}(z), \qquad (20)$$

where the coupling coefficient is defined as

$$C_{kn}^{(\nu, \nu-1)} = -\int_S \hat{E}_k^{(\nu)*} (\hat{\rho}_p \hat{\mathbf{v}}_n^{(\nu-1)} + \hat{\rho}_n^{(\nu-1)} \hat{\mathbf{v}}_p) dS \qquad (21)$$

and

$$C_{kn}^{(\nu, \nu+1)} = -\int_S \hat{E}_k^{(\nu)*} (\hat{\rho}_p^* \hat{\mathbf{v}}_n^{(\nu+1)} + \hat{\rho}_n^{(\nu+1)} \hat{\mathbf{v}}_p^*) dS, \qquad (22)$$

where $n = 0, 1, 2, \ldots$.

Equation (20) represents an infinite set of coupled equations for the excitation amplitudes of the modes $A_m^{(\nu)}$ and $A_n^{(\nu \pm 1)}$. Here, known quantities are the capped membrane functions for all of the physical quantities appearing in the integrands in expressions (21) and (22) for the coupling coefficients, as well as the longitudinal propagation constants [see (15)]; the procedure for determining the latter has been detailed in [7, 23], where the SCW eigenmodes in a TFSS under different boundary conditions were analyzed. Basically, with the coupling coefficients calculated, the set of equations (20) makes it possible to study the parametric interaction between the SCW modes at the signal (ω_ν) and idler ($\omega_{\nu \pm 1}$) frequencies under different boundary conditions with regard for the drift and diffusion, as well as the frequency dispersion and the anisotropy of the hot-electron differential mobility.

Special cases may significantly simplify the problem. For instance, if parametric interaction of only the principal modes ($k = m = n = 0$) at the signal and idler frequencies for $\nu = 0$ is considered, (20) turns into a set of three equations for the excitation amplitudes A_s, A_-, and A_+ for the principal SCW modes with the frequencies ω_s, $\omega_- = \omega_s - \omega_p$, and $\omega_+ = \omega_s + \omega_p$, respectively. If, from the physical grounds, one can leave the mode with one of the frequencies $\omega_\pm$ out of consideration and draw on a model of the rigid or quasi-free boundary of a carrier stream [7], system (20) can be brought to the conventional form of the equations for the amplitudes of two coupled waves [25] with specific analytical expressions for the coupling coefficients dependent on the pump amplitude.

CONCLUSION

In this paper, the general theory of the parametric interaction between SCWs in TFSS is presented. The TFSS may have NDC due to intervalley electron transitions under a high electrical field (like in *n*-GaAs and *n*-InP semiconductors). Our approach is based on the electrodynamic theory of waveguide excitation by extraneous currents, which was extended to arbitrary thin-film waveguide structures with composite active media. The elaborated theory will help to study the parametric interaction of SCWs in TFSS with NDC and take into account actual conditions at the film boundaries and diffusion, as well as the anisotropy and the frequency dispersion of the differential mobility of charge carriers.

REFERENCES

1. A. A. Barybin, *et al.*, Microélektronika **8**, 3 (1979).
2. R. H. Dean and R. J. Matarese, IEEE Trans. Electron Devices **60**, 1486 (1972).
3. K. Kumade and H. Kanbe, Int. J. Electron. **58**, 587 (1985).
4. A. I. Mikhaĭlov and S. A. Sergeev, Izv. Vyssh. Uchebn. Zaved., Radioélektron. **38** (10), 43 (1995).
5. G. L. Gurevich, M. A. Kitaev, A. L. Kogan, *et al.*, Radiotekh. Électron. **33**, 1272 (1988).
6. A. I. Mikhaĭlov and S. A. Sergeev, Pis'ma Zh. Tekh. Fiz. **22** (24), 75 (1996) [Tech. Phys. Lett. **23**, 1 (1997)].
7. A. A. Barybin, *Waves in Thin-Film Semiconductor Structures with Hot Electrons* (Nauka, Moscow, 1986).
8. A. A. Barybin, J. Appl. Phys. **46**, 1697 (1975).
9. V. A. Ivanchenko, B. N. Klimov, and A. I. Mikhaĭlov, Fiz. Tekh. Poluprovodn. (Leningrad) **12**, 601 (1978) [Sov. Phys. Semicond. **12**, 349 (1978)].
10. V. A. Ivanchenko, B. N. Klimov, and A. I. Mikhaĭlov, Fiz. Tekh. Poluprovodn. (Leningrad) **13**, 1172 (1979) [Sov. Phys. Semicond. **13**, 687 (1979)].
11. G. L. Gurevich and A. L. Kogan, Fiz. Tekh. Poluprovodn. (Leningrad) **12**, 1518 (1978) [Sov. Phys. Semicond. **12**, 898 (1978)].
12. G. L. Gurevich, A. L. Kogan, and G. M. Korobkov, Radiotekh. Électron. **29**, 333 (1984).
13. Yu. M. Ignat'ev and A. I. Mikhaĭlov, Izv. Vyssh. Uchebn. Zaved., Radioélektron. **33** (10), 76 (1990).

14. Yu. K. Pozhela, *Plasma and Current Instability in Semiconductors* (Nauka, Moscow, 1977).
15. L. A. Vaĭnshteĭn, *Electromagnetic Waves* (Radio i Svyaz', Moscow, 1988).
16. A. A. Barybin, J. Appl. Phys. **46**, 1707 (1975).
17. A. A. Barybin, *Progress in Electromagnetics Research*, Ed. by J. A. Kong (EMW Publishing, Cambridge, 1998), pp. 241–300.
18. A. A. Barybin, Izv. Vyssh. Uchebn. Zaved., Radioélektron. **20** (9), 118 (1977).
19. H. D. Rees, Solid State Commun. **7**, 267 (1969).
20. N. P. Belousov, E. I. Martynenko, and V. E. Chaĭka, Radiotekh. Élektron. **27**, 186 (1982).
21. N. P. Belousov and V. E. Chaĭka, Ukr. Fiz. Zh. **29**, 627 (1984).
22. E. V. Starikov and P. N. Shiktorov, Lietuvos Fizikos Rinkinys **32**, 471 (1992).
23. A. I. Mikhaĭlov, Pis'ma Zh. Tekh. Fiz. **21** (21), 89 (1995) [Tech. Phys. Lett. **21**, 902 (1995)].
24. A. A. Barybin and M. G. Stepanova, Izv. Leningrad. Élektrotekh. Inst., No. 437, 61 (1991).
25. W. H. Louisell, *Coupled Mode and Parametric Electronics* (Wiley, New York, 1960; Inostrannaya Literatura, Moscow, 1963).

Translated by B. Malyukov

Technical Physics, Vol. 45, No. 2, 2000, pp. 194–198. Translated from Zhurnal Tekhnicheskoĭ Fiziki, Vol. 70, No. 2, 2000, pp. 53–57.

The Generation of the Third Harmonic in Highly Anisotropic Media near the Percolation Threshold

A. A. Snarskiĭ, K. V. Slipchenko, and A. M. Satanin

Lobachevskiĭ State University, pr. Gagarina 23, Nizhni Novgorod, 630600 Russia

Received February 16, 1999

Abstract—The amplitude of the third harmonic in highly anisotropic media was calculated with a percolation model. Inclusion-shape and local-conductivity anisotropies were considered. Associated critical indexes and crossover values were obtained. © *2000 MAIK "Nauka/Interperiodica".*

The generation of the third harmonic is a promising method for nondestructive testing and inspection of composites. Over the last years, this method has been the subject of much theoretical and experimental investigation [1–6]. The third-harmonic amplitude increases with the inhomogeneity of the medium and is particularly large near the percolation threshold, where it grows according to a power law. The generation mechanism is as follows. When an alternative current with a frequency ω passes through an inhomogeneous medium, harmonics with frequencies $3\omega, 5\omega, \ldots$ appear in the voltage spectrum because of a slight nonlinearity in the local conductivity of the phases, for example, due to local overheating. Their amplitudes depend on the composite microstructure and may provide additional information about it. In [1, 2], the nonlinear conductivity was related to the $1/f$ noise factor and third-harmonic generation in weakly nonlinear media. In [4], an expression for the amplitude of the nth harmonic in a two-phase medium was derived under the assumption that the ratio between the phase conductivities is zero, i.e., that one of the phases becomes nonconductive above the percolation threshold. The amplitude above and below the percolation threshold, as well as near the percolation threshold (in the smear region), was calculated in [6] for a finite ratio of the phase conductivities. Related experimental data are given in [2–5].

The aim of this work is to calculate the third-harmonic amplitude in two-phase media with high electrical anisotropy near the percolation threshold. Such media can be subdivided into two groups: (1) composites with high anisotropy of the phase conductivities but with isotropic shapes of inclusions (the principal axes of the local conductivity tensor are parallel to each other) and (2) composites with high shape anisotropy of inclusions that are oriented in the same direction and are locally isotropic. The effective conductivity of such media near the percolation threshold was studied in [7–9].

We shall concentrate on geometrically anisotropic two-phase media with high anisotropy of the resistivity of inclusions with regard for the finiteness of the resistivity ratio. Expressions for the third-harmonic amplitude above and below the percolation threshold, as well as in the smear region, will be deduced and analyzed. Shape anisotropy of inclusions and the extreme cases for associated expressions will also be considered.

THIRD-HARMONIC GENERATION IN A TWO-PHASE MEDIUM WITH A HIGHLY ANISOTROPIC LOCAL CONDUCTIVITY TENSOR

The third-harmonic amplitude is calculated within a node–link–blob percolation model [10]. Consider first a medium from the first group. It may consist of an electrically anisotropic matrix and electrically anisotropic spherical inclusions with the equally oriented principal axes of the tensor. We consider the three-dimensional case and assume that the principal components of the resistivity tensor meet the conditions $\rho_{ix} \gg \rho_{iy} = \rho_{iz}$ and $\partial\rho_{iy}/\partial T = \partial\rho_{iz}/\partial T$ ($i = 1, 2$ is the phase no.) Thus, the medium as a whole is symmetric with respect to the X-axis and has two fundamentally different directions, X and Y. Another case of high anisotropy, $\rho_{ix} \ll \rho_{iy} = \rho_{iz}$, is basically the same, and expressions for the third-harmonic amplitude differ by a factor close to unity.

Let the model be represented as a conductive bridge and a poorly conductive spacer (Fig. 1). Taking into account the high anisotropy of the resistivity tensor, we split the bridge and the spacer into parts where the current passes along the directions with the least ($\Lambda_x S_x$) and the highest ($\Lambda_y S_y$) conductivity (Figs. 1a, 1b). To be strict, it is necessary, for the three-dimensional case, to consider also a "pin" in the Z direction and the part of the spacer that is perpendicular to Z. However, since the Y and Z directions are identical, we can concern only the elements shown in Fig. 1. The bridge and the spacer are connected in parallel at $p > p_c$ (Fig. 1c) and in series for $p < p_c$ (Fig. 1d). Here, p is the concentration of the conductive phase and p_c is the percolation threshold. The sizes of the bridges and the spacers are determined, to a first approximation, from the phase conductivity

1063-7842/00/4502-0194$20.00

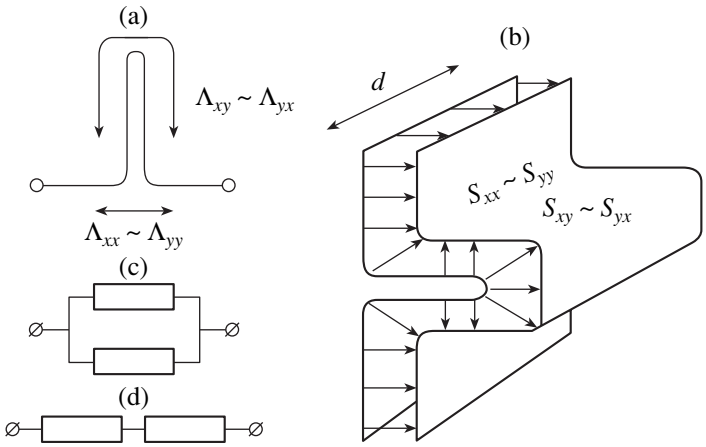


Fig. 1. Percolation model. (a) A conductive bridge in an anisotropic medium. The parts Λ_{xx} and Λ_{xy} of the bridge are aligned with the high- and low-resistivity directions. (b) A poorly conductive spacer in the anisotropic medium. The parts S_{xx} and S_{xy} are oriented in the same way as Λ_{xx} and Λ_{xy}. Arrows indicate the current passing through the spacer; d is the size of an inclusion. In (a) and (b), the conductivity tensor components for the different directions are given. Connection of the elements (c) above and (d) below the percolation threshold is also shown (R_Λ and R_s are the resistances of the bridge and the spacer, respectively).

ratio h [10]. The effective conductivity of such a medium is weakly anisotropic [8] in spite of the high electrical anisotropy of the inclusions. This is due to the fact that the local direction of the current depends on the percolation structure. As the concentration approaches the percolation threshold, the effective conductivity becomes progressively isotropic. It is taken into account here that, in our case, the nonuniformity of the local conductivity far exceeds the anisotropy of the inclusions. The diagonal components of the effective conductivity tensor in the principal axes are

$$\sigma_\perp^e = \sigma_{1x}\tau^t(1 + k_\perp\tau^{\lambda_1}), \quad \sigma_\parallel^e = \sigma_{1x}\tau^t(1 + k_\parallel\tau^{\lambda_1}),$$
$$p > p_c,$$
$$\sigma_\perp^e = \sigma_{2y}|\tau|^{-q}(1 + k_\perp|\tau|^{\lambda_2}), \qquad (1)$$
$$\sigma_\parallel^e = \sigma_{2y}|\tau|^{-q}(1 + k_\parallel|\tau|^{\lambda_2}), \quad p < p_c,$$

where $\tau = (p - p_c)/p_c$ is the nearness to the percolation threshold and t and q are the critical conductivity indexes above and below p_c, respectively.

Equating the resistivity of the correlation volume to that of the bridge (for $p > p_c$) or the spacer (for $p < p_c$), we obtain for the structure elements

$$\Lambda_{xx} = \Lambda_{yy} = a|\tau|^{-1}(1 - k_\parallel\tau^{\lambda_1}),$$
$$\Lambda_{xy} = \Lambda_{yx} = a|\tau|^{-1}(1 - k_\perp\tau^{\lambda_1}),$$
$$S_{xy} = S_{yx} = a|\tau|^{-q}(1 + k_\parallel\tau^{\lambda_2}), \qquad (2)$$
$$S_{xx} = S_{yy} = a|\tau|^{-q}(1 - k_\perp\tau^{\lambda_2}),$$

where a is the characteristic size of the inclusions and $k_\parallel$ and $k_\perp$ are constants.

The quantities $k_{\parallel,\perp}|\tau|^{\lambda_{1,2}} \ll 1$, which imply weak anisotropy of the effective conductivity [8], are much less than unity and subsequently will be neglected. Then, $\Lambda_{xx} = \Lambda_{xy} = \Lambda$ and $S_{xy} = S_{xy} = S$. We will assume that the current dependence of the resistivity is associated with local overheating; then, a resistivity variation in the structure elements, to an accuracy of the first term for Joule heat density $\rho_0 j^2$, is given by

$$\rho(T) = \rho_0(1 + m\rho_0 j^2\beta), \quad \beta = \frac{1}{\rho_0}\frac{\delta\rho}{\delta T}, \qquad (3)$$

where m is a constant that is a function of the mean temperature and the frequency (it relates a local rise in the temperature over its mean value to the released heat density and characterizes heat removal: $\delta T = m\rho_0 j^2$ and β is the temperature coefficient of resistance (TCR).

The resistances of the elements (Fig. 1) are

$$R_{\Lambda_{xx}} = \frac{\Lambda}{a}\rho_{1x}\left(1 + \frac{\partial\rho_{1x}}{\partial T}m\rho_{1x}j^2_{\Lambda_{xx}}\right),$$
$$R_{\Lambda_{xy}} = \frac{\Lambda}{a}\rho_{1y}\left(1 + \frac{\partial\rho_{1y}}{\partial T}m\rho_{1y}j^2_{\Lambda_{xy}}\right),$$
$$R_{S_{xx}} = \frac{a}{S}\rho_{2x}\left(1 + \frac{\partial\rho_{2x}}{\partial T}m\tilde{\rho}_{2x}j^2_{S_{xx}}\right), \qquad (4)$$
$$R_{S_{xy}} = \frac{a}{S}\rho_{2y}\left(1 + \frac{\partial\rho_{2y}}{\partial T}m\tilde{\rho}_{2y}j^2_{S_{xy}}\right),$$

where ρ_{1x} and ρ_{1y} are the components of the resistivity tensor of the conductive phase in the principal axes; ρ_{2x} and ρ_{2y}, the same for the poorly conductive phase; and $j_{\Lambda_{xx}}$, $j_{\Lambda_{yx}}$, $j_{S_{xx}}$, and $j_{S_{xy}}$, the current densities in the related elements.

Introduce the quantities

$$\rho_{1x}/\rho_{2x} = h_x \ll 1, \quad \rho_{1y}/\rho_{2y} = h_y \ll 1,$$
$$\rho_{1x}/\rho_{1y} = \alpha_1^2 \gg 1, \quad \rho_{2x}/\rho_{2y} = \alpha_2^2 \gg 1, \tag{5}$$

where h_i and α_i characterize the inhomogeneity and the anisotropy, respectively.

From the percolation model for $p > p_c$ (Fig. 1c), a current-dependent correction to the resistance of the correlation volume can be written as

$$\delta R = \frac{m\frac{a}{S}(\rho_{2y}^2\beta_{2y}j_{S_{xy}}^2 + \alpha_2^{-4}\rho_{2x}^2\beta_{2x}j_{S_{xx}}^2)}{\left(1 + \frac{1}{\alpha_2^2 h_x}\frac{a^2}{S\Lambda}\right)^2}$$
$$+ \frac{m\frac{\Lambda}{a}(\rho_{1x}^2\beta_{1x}j_{\Lambda_{xx}}^2 + \rho_{1y}^2\beta_{1y}j_{\Lambda_{xy}}^2)}{\left(1 + \alpha_2^2 h_x\frac{S\Lambda}{a^2}\right)^2},$$

where

$$\beta_{ij} = \frac{1}{\rho_{ij}}\frac{\partial\rho_{ij}}{\partial T}.$$

Taking into account that the correlation length in such a medium is defined as $\xi = a\tau^{\nu}$ and expressing the local current densities through the mean current density, we obtain for a correction to the effective resistivity

$$\delta\rho_e = \frac{m\frac{a\xi^2}{S^3}(\rho_{2y}^2\beta_{2y} + \alpha_2^{-8}\rho_{2x}^2\beta_{2x})}{\left(1 + \frac{1}{\alpha_2^2 h_x}\frac{a^2}{S\Lambda}\right)^4}\langle j\rangle^2$$
$$+ \frac{m\frac{\Lambda\xi^2}{a^3}(\rho_{1x}^2\beta_{1x} + \rho_{1y}^2\beta_{1y})}{\left(1 + \alpha_2^2 h_x\frac{S\Lambda}{a^2}\right)^4}\langle j\rangle^2. \tag{6}$$

Then, the addition to the mean field $\langle E\rangle$ has the form

$$\delta\langle E\rangle = \delta\rho_e\langle j\rangle. \tag{7}$$

Substituting $\delta\rho_e$ into (7) in view of (2) and (5), assuming that the external current follows the harmonic law $\langle j\rangle = \langle j_0\rangle\cos\omega t$, and writing the identity

$$\cos^3\omega t = \frac{3}{4}\cos(\omega t) + \frac{1}{4}\cos(3\omega t),$$

we obtain the expression for the third harmonic V_{3f}, which is convenient to normalize by $\langle j_0\rangle^3$:

$$B_{3f} = \frac{V_{3f}}{\langle j_0\rangle^3} \approx \rho_{1x}^2\left(\beta_{1x}\frac{m}{4}\right)\tau^{-(t+2\nu)}\left(1 + \alpha_1^{-4}\frac{\beta_{1y}}{\beta_{1x}}\right)$$
$$+ \rho_{2y}^2(h_x\alpha_2^2)^4\left(\beta_{2y}\frac{m}{4}\right)\tau^{-2\nu-4t-q}\left(1 + \alpha_2^{-4}\frac{\beta_{2x}}{\beta_{2y}}\right). \tag{8}$$

Similarly, using the model for $p < p_c$ (Fig. 1d), we come to

$$B_{3f} \approx \rho_{2y}^2\left(\beta_{2y}\frac{m}{4}\right)|\tau|^{3q-2\nu}\left(1 + \alpha_2^{-4}\frac{\beta_{2x}}{\beta_{2y}}\right)$$
$$+ \rho_{1x}^2\left(\beta_{1x}\frac{m}{4}\right)|\tau|^{-t-2\nu}\left(1 + \alpha_1^{-4}\frac{\beta_{1y}}{\beta_{1x}}\right). \tag{9}$$

Consider first two extreme cases of high inhomogeneity: $p > p_c$ and $p < p_c$. The former implies that the second phase is totally nonconductive; i.e., it is a perfect insulator:

$$B_{3f}(p > p_c) = \frac{V_{3f}}{\langle j\rangle^3}$$
$$= \rho_{1x}^2\left(\beta_{1x}\frac{m}{4}\right)\tau^{-t-2\nu}\left(1 + \alpha_1^{-4}\frac{\beta_{1y}}{\beta_{1x}}\right). \tag{10}$$

In the latter case, the metal is a perfect conductor; hence, the voltage drop can be neglected:

$$B_{3f}(p < p_c) = \rho_{2y}^2\left(\beta_{2y}\frac{m}{4}\right)|\tau|^{3q-2\nu}\left(1 + \alpha_2^{-4}\frac{\beta_{2x}}{\beta_{2y}}\right). \tag{11}$$

As follows from (10), if $\beta_{1y}/\beta_{1x} \sim \alpha_1^4 \gg 1$ or

$$\frac{\partial\rho_{1y}}{\partial T}/\frac{\partial\rho_{1x}}{\partial T} \sim \rho_{1x}/\rho_{1y},$$

the anisotropy parameter α_1 may significantly change the third-harmonic amplitude. Otherwise, α_1 has a negligible effect on B_{3f}. For $p < p_c$, the above conditions are read as $\beta_{2x}/\beta_{2y} \sim \alpha_2^4 \gg 1$ or

$$\frac{\partial\rho_{2x}}{\partial T}/\frac{\partial\rho_{2y}}{\partial T} \sim (\rho_{2x}/\rho_{2y})^3,$$

as follows from (11).

In deriving (8) and (9), it was meant that, along with (5), the conditions $(\alpha_1^2 h_y) = (\alpha_2^2 h_x) = \rho_{1x}/\rho_{2y} \ll 1$ are fulfilled; in other words, the maximum resistivity of the bridge is much less than the minimum resistivity of the spacer, which is the necessary condition for the percolation approach to apply [8].

From (8) and (9), it follows that, for $p > p_c$, the crossover conditions (when the critical behavior

changes) are set when the first and second terms in (8) equal each other:

$$\tau_c = \left[(\alpha_2^2 h_x)^2 \frac{\beta_{2y}}{\beta_{1x}} \left(\frac{1+\alpha_2^{-4}\beta_{2x}/\beta_{2y}}{1+\alpha_1^{-4}\beta_{1y}/\beta_{1x}}\right)\right]^{\frac{1}{3t+q}}. \quad (12)$$

$$p > p_c.$$

Similarly, for $p < p_c$ [expression (9)],

$$|\tau_c| = \left[(\alpha_1^2 h_y)^2 \frac{\beta_{1x}}{\beta_{2y}} \left(\frac{1+\alpha_1^{-4}\beta_{1y}/\beta_{1x}}{1+\alpha_2^{-4}\beta_{2x}/\beta_{2y}}\right)\right]^{\frac{1}{t+3q}}. \quad (13)$$

$$p < p_c.$$

If

$$\alpha_1^{-4}\frac{\beta_{1y}}{\beta_{1x}} \ll 1 \quad \text{and} \quad \alpha_2^{-4}\frac{\beta_{2x}}{\beta_{2y}} \ll 1$$

(12) and (13) are simplified and take the form

$$\tau_c = \left[(\alpha_2^2 h_x)^2 \frac{\beta_{2y}}{\beta_{1x}}\right]^{\frac{1}{3t+q}} = \left((\alpha_2^2 h_x)^3 \frac{\partial\rho_{1x}}{\partial T}/\frac{\partial\rho_{2y}}{\partial T}\right)^{\frac{1}{3t+q}}, \quad (14)$$

$$p > p_c$$

for $p > p_c$ and

$$|\tau_c| = \left[(\alpha_1^2 h_y)^2 \frac{\beta_{1x}}{\beta_{2y}}\right]^{\frac{1}{t+3q}} = \left(\alpha_1^2 h_y \frac{\partial\rho_{1x}}{\partial T}/\frac{\partial\rho_{2y}}{\partial T}\right)^{\frac{1}{t+3q}}, \quad (15)$$

$$p < p_c$$

for $p < p_c$.

The smear region in this case is

$$\Delta = \left(\frac{\rho_{1x}}{\rho_{2y}}\right)^{\frac{1}{t+q}} = (h_x\alpha_2^2)^{\frac{1}{t+q}} \quad [8].$$

Substituting this expression into (8) or (9) yields the third-harmonic amplitude in the smear region of the metal-to-insulator transition:

$$B_{3f} = \rho_{2y}^2\left(\beta_{2y}\frac{m}{4}\right)(h_x\alpha_2^2)^{\frac{3q-2\nu}{t+q}}\left(1+\alpha_2^{-4}\frac{\beta_{2x}}{\beta_{2y}}\right) + \rho_{1x}^2\left(\beta_{1x}\frac{m}{4}\right)(h_x\alpha_2^2)^{\frac{t+2\nu}{t+q}}\left(1+\alpha_1^{-4}\frac{\beta_{1y}}{\beta_{1x}}\right). \quad (16)$$

THIRD-HARMONIC GENERATION IN A GEOMETRICALLY ANISOTROPIC MEDIUM

Consider the case when the electrically isotropic inclusions have a highly anisotropic shape. Let them be ellipses with the minor and major axes a and b (Fig. 2) satisfying the condition $a/b = \alpha \ll 1$. Shklovskiĭ [7] was

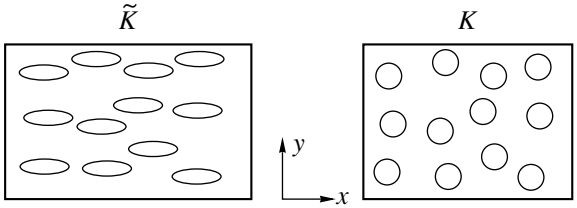


Fig. 2. The transition from a medium with locally anisotropic resistivity ($\rho_{xx} \neq \rho_{yy}$) and geometrically isotropic inclusions to that with isotropic resistivity ($\tilde{\rho}_{xx} = \tilde{\rho}_{yy} = \rho$) and geometrically anisotropic inclusions.

first to show that, using the coordinate transformation

$$x = \frac{\tilde{x}}{\sqrt{\alpha}}, \quad y = \tilde{y}\sqrt{\alpha}, \quad (17)$$

one can pass from a geometrically anisotropic medium with an isotropic local phase conductivity to a medium with a symmetric shape of inclusions and a highly anisotropic conductivity: $\rho_{1x}/\rho_{1y} = \rho_{2x}/\rho_{2y} = \alpha^2 \ll 1$. Such a transformation conserves the volume of the inclusions and, hence, the value of τ. It is such a medium that was considered above; therefore, using the already known parameters of the bridges and the spacers [see (2)] and taking into account (15), we can immediately find these parameters before the transformation:

$$\tilde{\Lambda}_{xx} = \tilde{\Lambda}_{xy} = \sqrt{\alpha}\Lambda_{xx}, \quad \tilde{\Lambda}_{yx} = \tilde{\Lambda}_{yy} = \frac{1}{\sqrt{\alpha}}\Lambda_{xx},$$
$$\tilde{S}_{xx} = \tilde{S}_{xy} = \frac{1}{\sqrt{\alpha}}S_{xx}, \quad \tilde{S}_{yx} = \tilde{S}_{yy} = \sqrt{\alpha}S_{xx}. \quad (18)$$

Here again, we discard small quantities of the order of $k_{\|,\perp}\tau^{\lambda_{1,2}}$ in the expressions for the correlation dimensions [see (2) and further], which now become direction-dependent (anisotropic), and obtain $\tilde{\xi}_x = \sqrt{\alpha}\xi = a\tau^{-\nu}$ and $\tilde{\xi}_y = (1/\sqrt{\alpha})\xi = b\tau^{-\nu}$. Using (16) and the same line of reasoning as in the previous section, we deduce the amplitudes when the external current is directed along the x- or y-axis:

$$B_x = \rho_1^2\left(\beta_1\frac{m}{4}\right)\tau^{-(t+2\nu)}\left(1+(\alpha^2 h)^2\frac{\beta_2}{\beta_1}\tau^{-(3t+q)}\right),$$
$$B_y = \alpha^4 B_x, \quad p > p_c \quad (19)$$

for $p > p_c$,

$$B_x = \frac{\rho_2^2}{\alpha^4}\left(\beta_2\frac{m}{4}\right)|\tau|^{3q-2\nu}\left(1+(\alpha^2 h)^2\frac{\beta_1}{\beta_2}|\tau|^{-(t+3q)}\right),$$
$$B_y = \alpha^4 B_x, \quad p < p_c \quad (20)$$

for $p < p_c$.

The size of the smear region is $\Delta = (h\alpha^2)^{\frac{1}{t+q}}$. Substituting this expression into (19) or (20) for τ yields the third-harmonic amplitude in the smear region:

$$B_x = \rho_1^2\left(\beta_1\frac{m}{4}\right)(\alpha^2 h)^{\frac{-t-2\nu}{\varphi}} + \alpha^{-4}\rho_2^2\left(\beta_2\frac{m}{4}\right)(\alpha^2 h)^{\frac{3q-2\nu}{\varphi}},$$
$$B_y = \alpha^4 B_x, \tag{21}$$

where $\varphi = t + q$.

High anisotropy of the third-harmonic amplitude, as well as the conductivity, is observed in the case of inclusion-shape anisotropy. Note that the amplitude is almost direction-independent when the local conductivity is anisotropic. The crossover values are

$$\tau_c = \left((\alpha^2 h)^2\frac{\beta_2}{\beta_1}\right)^{\frac{1}{3t+q}}, \quad p > p_c;$$
$$\tau_c = \left((\alpha^2 h)^2\frac{\beta_1}{\beta_2}\right)^{\frac{1}{t+3q}}, \quad p < p_c. \tag{22}$$

In both cases, the crossover is direction-independent.

Expressions (19) and (20) show that the only difference between the amplitudes for the different directions is the factor α. Hence, in this approximation, the anisotropy of the third harmonic depends on the shape anisotropy and does not depend on the nearness to the percolation threshold. We can also predict the critical behavior of the third-harmonic amplitude near the threshold. The value of V_{3f} drastically diverges at $p \longrightarrow p_c$. In a medium with local electrical anisotropy, the third-harmonic amplitude becomes isotropic, as does the effective conductivity. The expressions for the third-harmonic amplitude in both media incorporate the TCR and also the ratio of the resistivity tensor components. It can be thus concluded that additional information on local properties of a highly inhomogeneous medium can be extracted from third-harmonic measurements.

ACKNOWLEDGMENTS

This work was partially supported by the Russian Foundation for Basic Research (grant nos. 97-02-16923a and 97-02-18397).

REFERENCES

1. M. A. Dubson, Y. C. Hui, M. B. Weissman, *et al.*, Phys. Rev. B: Condens. Matter **39**, 6807 (1989).
2. Y. Yagil and G. Deutscher, Phys. Rev. B: Condens. Matter **46**, 16115 (1992).
3. Y. Yagil, G. Deutscher, and D. J. Bergman, Phys. Rev. Lett. **69**, 1423 (1992).
4. Y. Yagil, G. Deutscher, and D. J. Bergman, Int. J. Mod. Phys. B **7**, 19 (1993).
5. G. X. Ye, J. S. Wang, Y. Q. Xu, *et al.*, Phys. Rev. B: Condens. Matter **50**, 13163 (1994).
6. A. A. Snarskiĭ, Pis'ma Zh. Tekh. Fiz. **21** (1), 1 (1995) [Tech. Phys. Lett. **21**, 1 (1995)].
7. B. I. Shklovskiĭ, Pis'ma Zh. Tekh. Fiz. **7** (5), 21 (1981).
8. B. I. Shklovskiĭ, Phys. Status Solidi B **85**, K111 (1978).
9. F. Carmona and A. E. Amarti, Phys. Rev. B: Condens. Matter **35**, 3284 (1987).
10. A. Snarskiĭ, M. Setenko, and I. Bezsoudnov, Int. J. Electron. **77**, 1 (1994).

Translated by V.A. Isaakyan

Technical Physics, Vol. 45, No. 2, 2000, pp. 199–204. Translated from Zhurnal Tekhnicheskoĭ Fiziki, Vol. 70, No. 2, 2000, pp. 58–63.

Amplitudes and Shapes of Fabry–Perot Resonances in Laser Interference Thermometry of Semiconductors and Dielectrics

A. N. Magunov and S. V. Merkulov

Institute of Microelectronics, Russian Academy of Sciences, Yaroslavl, 150007 Russia

Received June 16, 1998; in final form, February 18, 1999

Abstract—The possibilities of the most sensitive method of laser thermometry of transparent and semitransparent plates have been determined for the varying optical properties of the surface and the conditions of the surface probing. The contrast of light interference in the plates was studied as a function of the diameter of the probing laser beam, the angle of the light incidence onto the surface and its roughness, and the presence of antireflecting and reflecting films on the surface. It is established that the contrast also depends on the transverse temperature gradient in the plate. The measurements were made on silicon single crystals and the optical K-8 glass. *© 2000 MAIK "Nauka/Interperiodica".*

INTRODUCTION

The solution of a number of scientific and technological problems requires the knowledge of the surface temperature preferably measured by contact-free methods. Recently, some new methods of active thermometry have been developed which are based on measuring the temperature-dependent parameters of a solid with the aid of a probing optical beam. The most sensitive low-noise method is the laser interferometric thermometry of transparent or semitransparent plates playing the part of the Fabry and Perot etalon for the probing radiation [1]. Heating the plate, one changes the refractive index and the geometrical thickness of the plate and determines the current temperature from the number of optical resonances (periods of the interference pattern) with respect to the known initial temperature. However, this method found no wide use because in some cases it is impossible to obtain the interference pattern for a plate heated by an irradiating probing laser beam. Moreover, it is taken *a priori* that the method can be applied only to thoroughly polished surfaces [2, 3].

The present article is devoted to the study of the Fabry–Perot resonances observed when heating glass and silicon plates. The interference patterns were obtained from industrially manufactured single crystals (including those with rough surfaces) and glasses with surfaces not subjected to any additional treatment.

EXPERIMENTAL SETUP

The schematic of a setup is shown in Fig. 1 and is described in detail elsewhere [4]. The probing radiation was that of an LGN-113 He–Ne laser (the wavelengths 0.63 and 1.15 μm). The beam was focused on the surface by a lens with the focal distance f = 8–30 cm; the probing-beam diameter on the surface was $d \approx f\Theta \approx$ 0.1–0.3 mm (where Θ is the divergence angle of the beam). Without a lens, the beam diameter on the specimen surface was 2 mm. The specimen temperature was increased at a step of 20–30 K, which was sufficient for recording several periods of the interference pattern used for the contrast determination. The signal from a photodetector (a silicon or a germanium photodiode) was recorded and then computer-processed. The photodetector can be tuned either to a reflected or a transmitted radiation. To study the contrast statistics, we recorded about 30 points spaced by 0.5 mm (for d = 0.1 and 0.3 mm), 1 mm (for d = 1 mm), and 2 mm (for d = 2 mm) along an arbitrarily chosen line. The dependence of the contrast on the angle of light incidence onto the surface was measured from the specimen on a rotating table, with the rotation angle around the vertical axis being set on the graduated circle with a scale value of 1°.

Figure 2 illustrates the sequence of resonances (an interference pattern) obtained in the light reflected from the heated 0.5-mm-thick silicon crystal. The shift by one fringe (i.e., the variation in the path difference by 2π) observed for a silicon single crystal at λ = 1.15 μm corresponded to the change in the optical thickness of

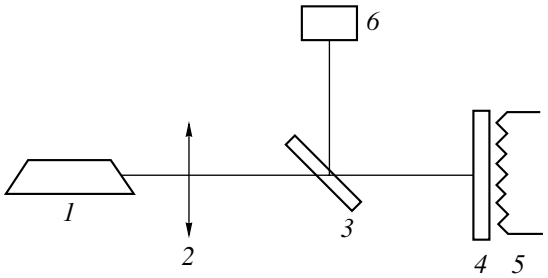


Fig. 1. Schematic of the setup for recording optical resonances in the plates: (*1*) laser, (*2*) lens, (*3*) semitransparent mirror, (*4*) plate under study, (*5*) heater, (*6*) photodetector.

1063-7842/00/4502-0199$20.00 © 2000 MAIK "Nauka/Interperiodica"

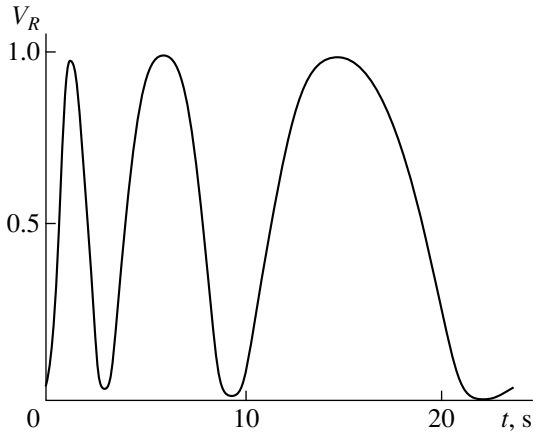


Fig. 2. A part of the interference pattern obtained in the reflected light from a heated silicon single crystal with two polished surfaces. The beam diameter 0.1 mm. Contrast $V_R \approx 1$.

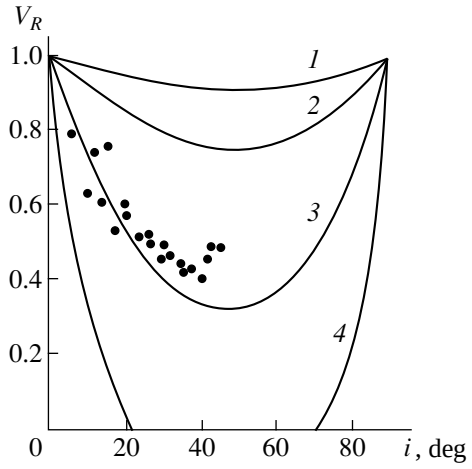


Fig. 3. Contrast of the interference pattern obtained in the reflected light (λ = 1.15 μm) as a function of the angle of beam incidence onto the surface of a 0.5 mm-thick silicon crystal. The curves indicate the data calculated for various D values: (*1*) 3, (*2*) 1, (*3*) 0.3, and (*4*) 0.1 mm; points indicate the data measured at D = 0.3 mm.

the crystal by $\Delta(nh) \approx 0.16$ μm or the change in the temperature by $\Delta T \approx 5.2$ K.

EFFECT OF THE ANGLE OF LIGHT INCIDENCE

The contrast (visibility) V of the interference fringes characterizes the amplitude of the intensity oscillations (I) of the transmitted or the reflected light during the change of the plate temperature, $V = (I_{max} - I_{min})/(I_{max} + I_{min})$. If the contrast is very low (e.g., $V \ll 0.01$), the calculation of the number of the fringes against the background of unavoidable noise becomes a rather difficult problem, and, therefore, the thermometric method under consideration becomes unreliable. A finite value of the laser-beam diameter imposes certain limitations on the thermometry method if the plate is irradiated at a certain angle to the surface normal. In order to eliminate the signal from the parasitic interference in the optical window, which is in contact with the plasma and, thus, is also heated. The substrate probing at angles ranging within 5°–20° to the surface normal is sometimes performed in a plasmochemical reactor [5, 6]. Moreover, some setups used in microtechnology allow the use of only large angles (40°–50° or larger) of the light incidence onto the surface. In this case, the reflected beams of different orders only partly overlap in the photodetector plane. In the region of nonoverlapping beams, a photodetector records the constant light intensity irrespectively of the thickness variations in the plate, and the interference contrast does not attain its maximum value. Consider the simplest case of a two-beam interference in a nonabsorbing plate. Using the expressions for the partial beam overlap [7, 8], we arrive at the dependence of V_R on the angle of light incidence onto the surface for different beam diameters (Fig. 3). The contrast varies in a nonmonotonic way: it decreases in the range of 0°–45° and increases in the range of 45°–90°. Moreover, at $D/h \le (\sin 2i)/(n^2 - \sin^2 i)^{1/2}$ (where D is the beam diameter, h is the plate thickness, i is the angle of incidence, and n is the refractive index), a specific "dead zone" of the incidence angles is observed, because the beams are spatially resolved in the photodetector plane, and, thus, there is no interference at all. An increase of the contrast observed at large angle of incidence onto the surface is very interesting for thermometry of the substrates for a number of microtechnological processes, because it allows the thermometric measurements through optical windows used in the ellipsometric measurements (usually, at the incidence angle of 70°).

EFFECT OF NONIDEAL PLANE-PARALLEL SUBSTRATE

In order to study the optical-resonance properties, we used industrially manufactured (100)-oriented KÉF-4.5 silicon single crystals with diameters of 76–100 mm. The working surface of each crystal was polished until the microrelief height attained the value of $R_z \le 0.05$ μm. According to the properties of the back surfaces, the specimens were divided into two types: type-1 specimens with back surfaces polished more roughly (with the microrelief height 0.1 μm) than the front face and type-2 specimens with rough back surfaces with the average relief height of about 1–2 μm. The shapes and the dimensions of the microfaces were ordered with respect to both crystallographic axes and the slope of the microfaces with respect to the basal plane [4].

The distribution of the local angles formed by two polished crystal surfaces was studied by the known method based on the use of a microphotometer with

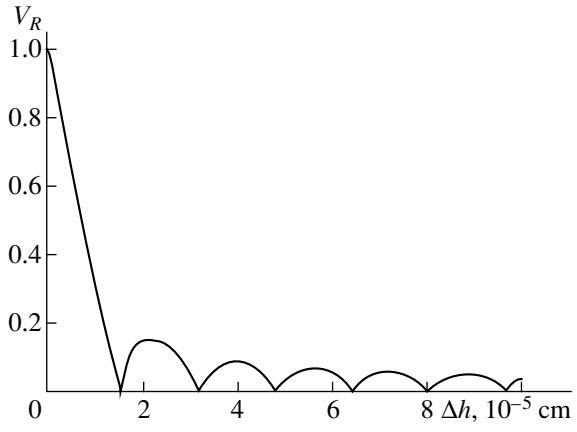


Fig. 4. Contrast as a function of the thickness variation in a Si crystal in the section of the light beam. The wavelength 1.15 μm, normal light incidence.

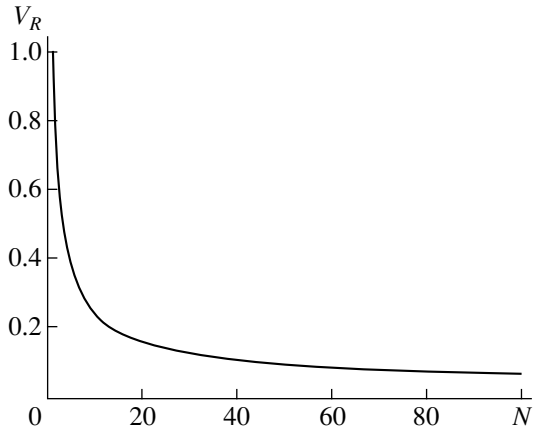


Fig. 5. Dependence of the average contrast on the number of microfaces in the cross section of the light beam in the case of its normal incidence onto the surface. All the microfaces have the same areas in the crystal plane.

three-dimensional scanning [6]. For a 0.35–0.38 mm-thick crystals of type 1 having a diameter of 76 mm, the most probable angle equals (1–3) × 10^{-4} rad, for 0.45–0.48 mm-thick crystals of the diameter 100 mm, the most probable angle is (3–10) × 10^{-5} rad. The main cause of the reduced interference contrast for such crystals is the variation of the thickness over the crystal. If the crystal thickness varies along the beam path, the photodetector yields the averaged partly or complete compensated interference maxima and minima. It is clear that with an increase of the probing-beam diameter, the contrast should increase if the angle formed by the plate surfaces has a nonzero value.

The table shows the values of the average contrast and the mean-square deviations from it for the specimens studied. The contrast decreases with an increase of the beam diameter. The theoretically possible contrast in the reflected light is $V_R = 1$ (for Si and glass) and in the transmitted light, $V_T = 0.57$ (for Si). The experimental values of V_R for type-1 Si at $D \approx 0.1$ mm are close to unity, whereas the experimental V_T values for the same specimens coincide with the theoretically predicted values. Figure 4 shows the calculated dependence of the contrast on thickness variation along the light-beam cross section for polished silicon plates. It is seen that the probability of recording a pattern with the zero contrast is negligibly low; the probability of recording the contrast $V_R \le 0.01$ does not exceed 0.01–0.03 and increases with an increase in the thickness variation. The average angle between the crystal surfaces dh/dx (where x is the coordinate along the surface) can be determined by two methods: (a) by comparing the plot $V(\Delta h)$ with the tabulated data, which yields $dh/dx \approx 1.2 \times 10^{-4}$, and (b) by scanning the crystals with a probing light beam (λ = 1.15 μm, d = 0.2 mm); the average distance between the neighboring interference extrema in the reflected light is about 1.5 mm, whence $dh/dx \approx 1.1 \times 10^{-4}$. Thus, the local measurements of the interference contrast provide the information on the thickness gradient of the plate.

Figure 5 shows the calculated dependence of the contrast as a function of the number N of microfaces in the section of the probing beam for a 0.5-mm-thick Si crystal with a rough surface. This dependence can be approximated by a power function. Thus, if $N \ge 2$, we have $V_R \approx 0.77 \times N^{-0.53}$; if $N \ge 4$ we have $V_T \approx 0.61 \times N^{-0.52}$ (in both cases, the correlation coefficient equals 0.995). The model of a rough surface approximately corresponds to one of the types of the surface roughness of single crystals studied earlier [9]. The height of the parallelepipeds forming the roughness was set within 0–2 μm with respect to the basal plane with the aid of a generator of random numbers. For each Fabry–Perot microresonator, we calculated the transmission and the reflection coefficients in the geometrical optics approximation. These coefficients were averaged over an ensemble of N microresonators of the same area. Then, the values of V_R and V_T were calculated for the thus-obtained microresonator with a rough surface. For each N value, the contrast was averaged over 300 tests. For any N, the distribution density of the contrast over

Contrast of the interference patterns obtained from a heated silicon single crystals of type 1 (Si I) and type 2 (Si II) at different diameters of the probing beam

D, mm	V_R (Si I)	V_T (Si I)	V_R (Si II)	V_T (Si II)
0.1	0.91 ± 0.04	0.57 ± 0.02	0.12 ± 0.05	0.14 ± 0.06
0.3	0.78 ± 0.05	0.56 ± 0.01	0.08 ± 0.04	0.10 ± 0.06
1	0.63 ± 0.15	0.34 ± 0.10	0.05 ± 0.02	0.03 ± 0.01
2	0.55 ± 0.12	0.2 ± 0.2	0.02 ± 0.001	0.015 ± 0.004

Note: The average contrast and the mean-square deviations were obtained using 20 values for Si I and 30 values for Si II.

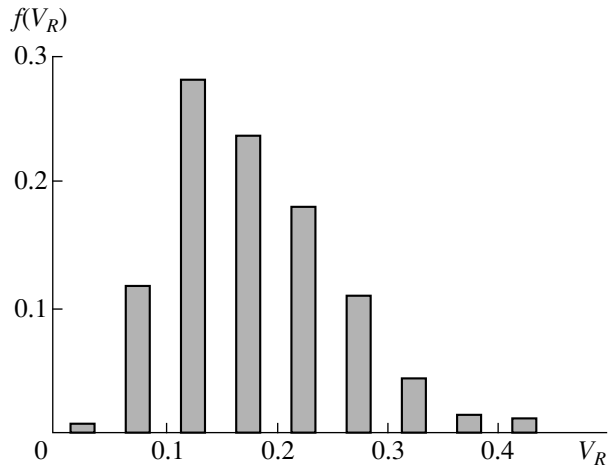


Fig. 6. Histogram of the contrast distribution $V_R = 0.18 \pm 0.09$ for 350 tests at $N = 16$.

a series of tests had the form of a curve with the maximum (Fig. 6) observed at a nonzero V value. The experimental data listed in table are consistent with the curve calculated in the range $N \approx 10$–3000. Earlier [4], similar distributions were also obtained experimentally in a scheme with "fiber-optical transport" of the light beam for $N \le 4$. Thus, even at high N values, the probability of recording an interference pattern with the zero contrast is close to zero. This amazing fact was repeatedly confirmed experimentally [4, 10, 11].

The interference patterns from type-1 crystals show a much higher contrast than those from type-2 crystals. Nevertheless, the use of a beam with a diameter 0.1–0.3 mm allows one to record the interference patterns from heated crystals with rough surfaces. This fact is important for the microtechnology, because the major part of the crystals used in this technology (substrates for integral circuits) have rough back surfaces. In our experiments, we irradiated the polished surfaces of the plates, because the illumination of the rough surface yields too weak mirror-reflected radiation already at a distance of 30 cm from the plate (however, using the collecting lens, it is possible to record the interference patterns in this case as well). The effect of roughness on light reflection from the surface can be reduced by using a more long-wavelength radiation (e.g., with the wavelength of 3.39 or 10.6 μm). This also promotes a higher contrast.

The measurements were performed, as well, on specimens of the optical K-8 glass (2.54 mm in thickness) at the wavelength 0.63 μm. Both glass surfaces were polished, and the average height of the microrelief was 0.01 μm. The angle formed by the surfaces was about 5×10^{-5} rad. Under a normal light incidence and a beam diameter equal to 0.3 mm, the contrast was $V_R = 0.97 \pm 0.02$. However, in this case the intensity of the reflected light was lower than for silicon (in the interference maximum, $R_{max} = [(n^2 - 1)/(n^2 + 1)]^2$; in other words, for glass with a refractive index of $n = 1.516$, we have $R_{max} \approx 0.16$; for silicon at $n = 3.535$, we have $R_{max} \approx 0.73$).

EFFECT OF TRANSVERSE TEMPERATURE GRADIENT

It was established experimentally that the temperature inhomogeneity of the plate along the diameter of the probing beam results in a reduction of the interference contrast. The temperature dependences $n(T)$ and $h(T)$ lead to the transformation of the temperature inhomogeneity into the nonuniform optical thickness nh along the beam diameter. This, in turn, leads to partial or complete compensation of the maxima and minima during recording of the beam intensity. In this case, the mechanism of the contrast reduction is the same as in the case of nonuniform geometrical thickness. The condition under which the temperature gradient can affect the contrast is written in the following way:

$$4Dh\left(\frac{\partial n}{\partial T}\frac{\partial T}{\partial x}\right)\lambda^{-1} \ge 1.$$

This condition signifies that more than one interference half-period fits the beam diameter, which is accompanied by the contrast reduction. Earlier [6], an analogous condition was formulated for the thickness gradient of the plate. For a beam of diameter 2 mm and the glass plate of the thickness of 1 mm, this condition is met if $\partial T/\partial x \ge 7$ K/cm for Si and $\partial T/\partial x \ge 270$ K/mm. Such a pronounced difference is explained by the fact that the $\partial n/\partial T$ values for semiconductors and dielectrics differ by almost two orders of magnitudes. We observed the reduction of the interference contrast in the case of the transverse temperature inhomogeneity in thermometric measurements on Si single crystals in a gas-discharge plasma if the crystal had a nonsymmetric position with respect to the axis of a cylindrical reactor. The dependence of the contrast on the temperature difference along the diameter of the light beam is described by the dependence similar to that shown in Fig. 4.

The contrast can also be reduced by displacing the reflected first-order beam with respect to the zero-order beam due to refraction from the transverse temperature gradient. In this case, the distance between the beam centers is determined by the expression

$$L \approx 2Hh\frac{\partial n}{\partial T}\frac{\partial T}{\partial x},$$

where H is the plate–detector distance.

In this case, the pronounced temperature gradients can give rise to complete spatial resolution of the beams on the photodetector.

EFFECT OF FILMS COATING THE SURFACE

The effect of antireflecting films on the interference contrast was studied on Si single crystals with the surfaces coated with 0.2–0.4 μm-thick SiO_2 or Si_3N_4 films. The effect of reflecting coatings was studied on Si single crystals coated with 0.3 μm-thick aluminum or copper films. The presence of reflecting and antireflecting films on the surface of Si single crystals usually results in the reduction of the contrast. The only exception is transparent films having the thickness multiple to $\lambda/2n_f$ (the subscript *f* indicates the film): the half-wave films on the surface produce no effect on the contrast. For the crystals with one surface coated with partly antireflecting film of silicon dioxide and the other surface, coated with the reflecting aluminum film, the contrast of the interference pattern does not exceed 0.03–0.05. However, the thermometric measurements of such films are still possible, because the electronic processing of the signal allows the identification of the extrema even on the interference patterns with the contrast 30–50 times lower [12]. Recording of interference patterns with the contrast $V \geq 0.001$ is, in fact, a sufficient condition for thermometric measurements.

RESONANCE SHAPES

The resonance properties of the plates are characterized by the fringe finesse *F* showing how many times the distance between the neighboring extrema exceeds the width of the reflection minimum. One can readily show that the contrast in the transmitted light V_T and the fringe finesse are related as

$$E = \pi(\arccos V_T)^{-1},$$

where

$$V_T = \frac{2\sqrt{R_1R_2}e^{-\alpha h}}{1 + R_1R_2e^{-2\alpha h}},$$

α is the linear absorption coefficient of light, and *h* is the film thickness.

It is seen that with an increase in V_T, the reflection minimum becomes narrower. This relationship shows in what way the variation in the surface properties affects the shapes of the Fabry–Perot resonances. Thus, applying a strongly reflecting film onto the back surface ($R_2 \sim 1$), one can make the reflection minima narrower. With an increase of the light absorption in the plate, the reflection minima broaden. Using the data listed in table, one can determine that $F \approx 3.26$ for type-1 Si and $F \approx 2.2$ for type-2 Si. Roughness reduces the effect of high-order reflection beams because of light scattering and, thus, makes the maximum and the minimum widths almost equal. The possibility of controlling the fringe shapes is of great importance for the continuous temperature determination from the interference pattern. For real Si, Ge, GaAs, and GaP crystals (with the nonzero deviation of the surfaces from parallelism, i.e., at $F \approx 3.3$–3.7), the temperature is determined only at the moments at which the interference pattern corresponds to the minimum reflection. Quasicontinuous temperature measurements (e.g., made at a step of 10^2–10^3 counts per period) of such crystals cannot be performed because of the complicated character of the overlap of the beams of various orders at the photodetector (that is why it is impossible to construct an analytical model adequately describing the interference pattern at all its points). If $F = 2$, such a description gives rise to no difficulties. A decrease of the fringe finesse from the specimens with both surfaces being partly covered with antireflecting films allows one to reduce the contribution of high-order beams and describe the interference pattern with considerable precision in the two-beam approximation. Thus, contrary to spectroscopy, the large number of interfering beams in the Fabry–Perot resonator is not the necessary condition—it is rather an undesirable factor for thermometric measurements.

CONCLUSION

In most cases, plates with polished and rough surfaces slightly deviating from the plane-parallel condition and those coated with thin films provide the formation of an interference contrast sufficient for the remote temperature measurements. The influence of differently modified surfaces on the shape and the amplitude of the optical resonances can preliminarily be estimated using rather simple relationships. In fact, there are almost no restrictions on the angles of light incidence onto the surface, which makes the method under consideration quite efficient for the studies and control of the processes occurring in the technological setups.

At present, laser interference thermometry is widely used for the study and control of the plasmochemical treatment, epitaxial growth, and film deposition from the gas phase. The sensitivity of the method exceeds the sensitivities of other thermometric methods by two to four orders of magnitude and is characterized by a wide temperature range (for silicon, from the cryogenic temperatures up to ~1000 K; for broader-band semiconductors and dielectrics, the upper limit is even higher). The use of the semiconductor diode lasers [13] and the fiber-optic communication links [14] in the optical scheme makes the method highly reliable, so that it can be used in the technological setups for measuring substrate temperatures.

ACKNOWLEDGMENTS

This study was supported by the Ministry of Science and Technologies of the Russian Federation (the program *Perspective Technologies and Devices in Micro- and Nanoelectronics*) and the Russian Foundation for Basic Research, project no. 96-02-18164.

REFERENCES

1. A. N. Magunov, Prib. Tekh. Exp., No. 3, 6 (1998).
2. R. L. Anderson, Proc. SPIE **1392**, 437 (1990).
3. K. L. Saenger, E. Tong, J. S. Logan, *et al.*, Rev. Sci. Instr. **63**, 3862 (1992).
4. A.N. Magunov and A.Yu. Gasilov, Mikroélektronika **26**, 384 (1997).
5. R. A. Bond, S. Dzioba, and H. M. Naguib, J. Vac. Sci. Technol. **18** (2), 35 (1981).
6. A. N. Magunov and E. V. Mudrov, Teplofiz. Vys. Temp. **30**, 372 (1992).
7. A. N. Magunov, Proc. SPIE **2208**, 103 (1995).
8. A. N. Magunov, Zavod. Lab, No. 9, 27 (1995).
9. A. N. Magunov, P. G. Buyanovskaya, A. Yu. Gasilov, *et al.*, Opt. Spektrosk. **84**, 68 (1998).
10. A.V. Belinskiĭ, Usp. Fiz. Nauk **165**, 691 (1995).
11. A. Yu. Gasilov and A. N. Magunov, Zavod. Lab **63** (8), 32 (1997).
12. A. N. Golubev, Opt. Tekh., No. 4, 40 (1995).
13. R. Kurosaki, J. Kikuchi, and Ya. Kobayashi, Proc. SPIE **2635**, 224 (1995).
14. O. V. Lukin and A. N. Magunov, Mikroélektronika **24**, 119 (1995).

Translated by L. Man

Technical Physics, Vol. 45, No. 2, 2000, pp. 205–208. Translated from Zhurnal Tekhnicheskoĭ Fiziki, Vol. 70, No. 2, 2000, pp. 64–68.

Periodic Structure Created by a CO_2-Laser Pulse Incident on an Open Water Surface

S. V. Gribin, I. I. Komissarova, G. V. Ostrovskaya, B. I. Spesivtsev, V. N. Filippov, and E. N. Shedova

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

Received May 14, 1998

Abstract—Experiments on the effect of a CO_2-laser pulse on an open water surface are reported. They testify that the multilayer periodic structure of the resultant scattering region is due to the intensity modulation at the trailing edge of the pulse. The generation of optoacoustic pulses is studied at two different values of laser-pulse length. The shorter laser pulse is obtained through the trailing-edge absorption in a laser-spark plasma produced by focusing the radiation. *© 2000 MAIK "Nauka/Interperiodica".*

We have reported [1, 2] the formation of a scattering region beneath an open water surface exposed to a focused pulse from a CO_2 laser at about 15 J/cm^2. Appearing on shadowgrams, the region first expands downwards (for a time less than 0.5 μs) and then acquires a multilayer periodic structure. The velocity of the lower boundary during the initial phase of evolution is twice as large as that of sound in water. This paper presents a closer examination of the phenomenon.

Further experiments with CO_2 laser radiation revealed that a similar structure can be created by a nonfocused pulse as well. The evaporative mechanism of acoustic wave generation will not do in this case because of a low-density energy flux. Extending to the leading edge of an optoacoustic pulse, the scattering region touches the water surface within initial 3 μs; then it separates from the surface and follows the leading edge, moving downwards at the velocity of sound. Figure 1 shows a shadowgram taken at 2 μs. It can easily be seen that the region has a periodic structure. The average period of the structure (about 0.1 mm) was measured with an illuminating pulse, which passed through the region and a lense and produced a diffraction pattern in its back focal plane (Fig. 1b). The period is virtually independent of the energy flux from the excitation laser.

In the case of evaporative generation, the periodic structure was detected by other workers on shadowgrams and interferograms [3, 4] or on oscilloscope traces from piezoelectric sensors [3, 5]. In [3], mechanisms that may underlie the high-frequency acoustic vibrations are covered; however, none of them has received firm experimental support. It was suggested [5] that the generation of sound at high excitation intensities, typical of the evaporative regime, may stem from the instability of an interfacial layer and its destruction by a low-pressure pulse (cavitation) due to the reflection of a shock wave from the water surface. With regard to this hypothesis, the questions arise as to how the destruction of the interfacial layer leads to the acoustic vibrations and what determines their frequency. By contrast, in [6, 7], the generation of ultrasonic waves under the action of solid-state mode-locked and solid-state free-running lasers was reported. The behavior of the high-frequency optoacoustic component here is associated with the intensity modulation of the excitation pulse.

In this study, we share the conjecture [6–8] that the acoustic wave train is produced by the intensity modulation of the excitation pulses. The view is supported by the facts that the periodic structure is observed also at relatively low energy fluxes ($E < 1$ J/cm^2), typical of thermal mechanisms, and that its spatial period is independent of laser intensity. Furthermore, the length of the acoustic wave train is equal to the total length of the excitation pulse, about 3 μs (Fig. 2a).

The waveform of an excitation pulse was picked up with a photosensor offering a time resolution of the order of 1 ns. The photosensor exploits the drag effect in a *p*-type germanium. A typical waveform is shown in Fig. 2b, where the weak periodic oscillation is observed (curve *1*). Its amplitude is 4 ± 2% of the pulse height, and its frequency, equal to about 125 MHz, correlates with that of acoustic waves, $\nu = a/d \approx 15$ MHz. Here, $a = 1.5 \times 10^5$ cm / s is the velocity of sound in water and $d = 0.1$ mm is the spatial period of the structure.

In order to verify our conjecture, we conducted the following experiment, depicted by Fig. 3. An excitation pulse from a CO_2 laser is focused by spherical mirror M_1 so that a laser spark is produced at point S in air. The spark cuts off the trailing edge of the pulse, since the radiation is absorbed by the spark plasma. The waveform of the plasma-shortened excitation pulse is shown in Fig. 2b (curve *2*). Then the pulse is directed to the

1063-7842/00/4502-0205$20.00

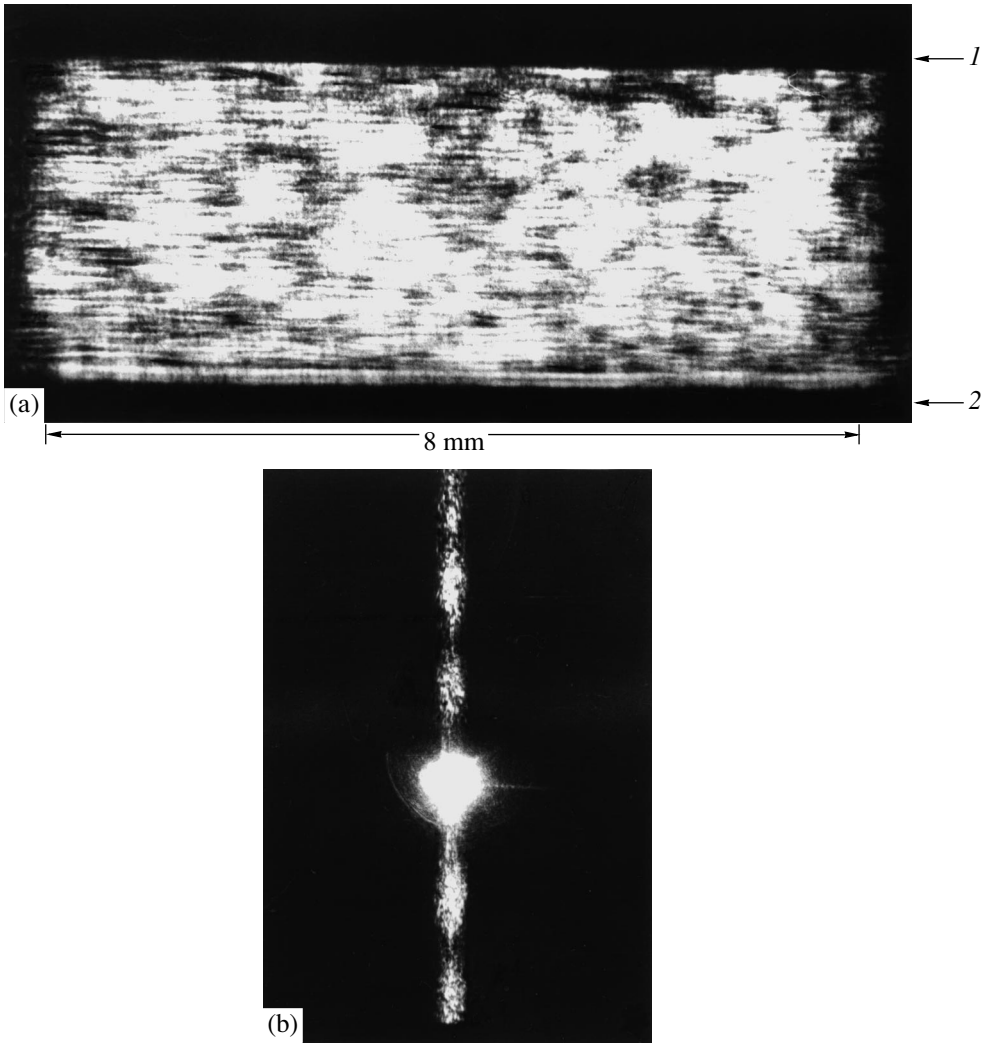


Fig. 1. Images of the scattering region: (a) shadowgram at 2 μs (*1*, water surface; *2*, leading edge of an optoacoustic pulse) and (b) diffractogram of an illuminating pulse in the back focal plane of the lens.

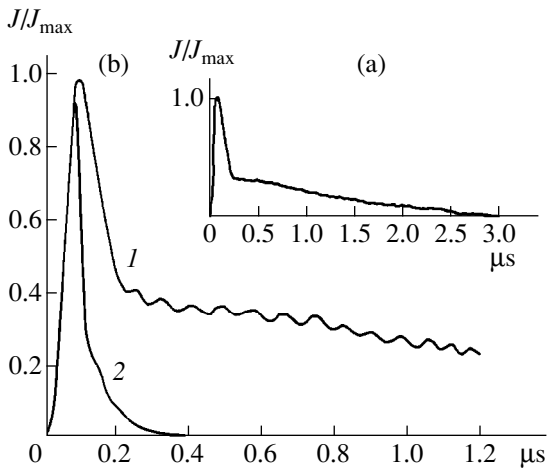


Fig. 2. Waveforms of the excitation pulses for an oscilloscope resolution of (a) 0.5 and (b) 0.2 μs/division. Curves *1* and *2* depict the longer and the shorter pulse, respectively.

water surface by spherical mirror M_2 and flat mirrors M_3 and M_4. The incident intensity is varied by adjusting the distance between M_3 and M_4. The region of hydrodynamic excitation is illuminated by a pulse from a ruby laser in the direction perpendicular to the plane of the drawing. A shadowgram and a holographic interfer-

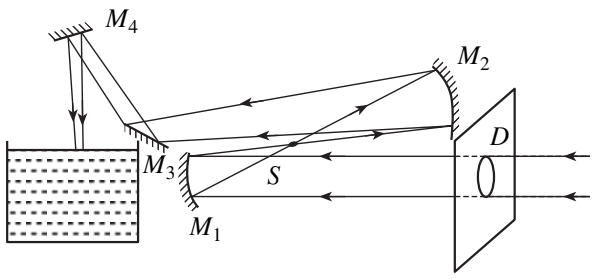


Fig. 3. Schematic drawing of the excitation pulse shaper.

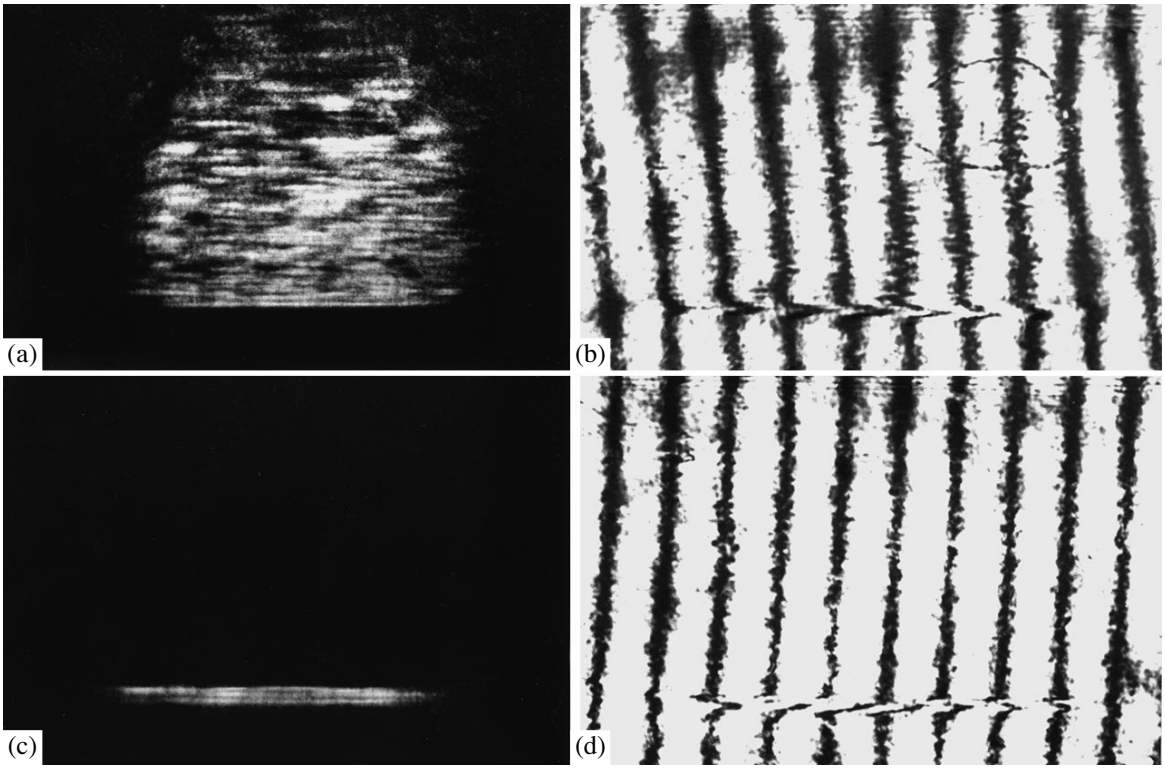


Fig. 4. Shadowgrams and holographic interferograms of the optoacoustic pulse that is generated by (a, b) the longer and (c, d) the shorter excitation pulse at I_{max} = 2.5 MW/cm^2 and τ = 2.2 μs.

ogram are recorded simultaneously during a pulse; this is achieved by means of a mirror with an opening that serves as the visualizing diaphragm of the schlieren system [7]. Figure 4 presents sample patterns obtained with the (a, b) longer or (c, d) the shorter excitation pulse. In the former case, the pulse is limited by diaphragm D so as to attenuate the intensity at the focus of M_1 to a value below the air breakdown threshold. The energy flux density on the water surface was ~1.5 or ~0.1 J/cm^2 for the longer and the shorter pulses, respectively, the maximum intensity I_{max} being ~2.5 MW/cm^2 in both cases. As indicated by Fig. 4, the shorter excitation pulse cannot generate an acoustic wave train and only a short bipolar optoacoustic pulse is present. Figure 5 shows the waveforms of the optoacoustic pressure pulses generated by (a) the longer or (b) the shorter excitation pulse. Waveforms (a) and (b) are obtained from holographic interferograms. Notice that the leading (positive) peaks are nearly the same in magnitude, whereas the negative peak (valley) is materially larger for the shorter excitation pulse. Panels (c) and (d) display waveforms computed from the notion that the waveform of an optoacoustic pressure pulse is the derivative of the excitation waveform if $\alpha a\tau \gg 1$, where α is the absorption coefficient, a is the velocity of sound, and τ is the excitation pulse length [9]. Good agreement between the measured and the computed data for the longer excitation pulse, as well as the absence of the high-frequency oscillation for the shorter excitation pulse, reinforce the conjecture that the creation of the periodic structure accompanying the optoacoustic pulse stems from the intensity modulation at the trailing edge of the excitation pulse. It is pertinent to note that the authors of [3] considered this process as

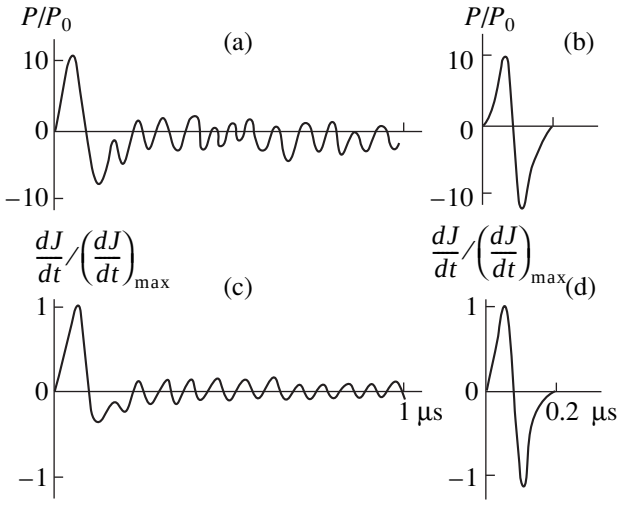


Fig. 5. Waveforms of the optoacoustic pressure pulses that are generated by (a, c) the longer and (b, d) the shorter excitation pulse. (a) and (b), experiment; (c) and (d), computation.

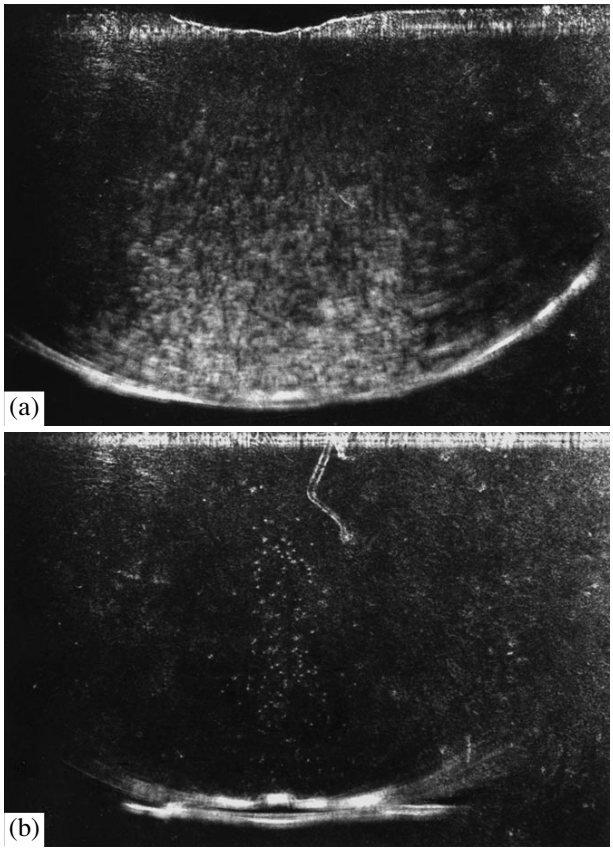


Fig. 6. Shadowgrams of the optoacoustic pulse that is generated by (a) the longer and (b) the shorter excitation pulse at I_{max} = 7.5 MW/cm^2 and τ = 5.5 μs.

one of the possible mechanisms underlying the ultrasonic generation by a CO_2-laser pulse but ultimately rejected it, failing to detect trailing edge modulation.

The generation of optoacoustic pulses by shortened excitation pulses has the following features. The transition from a thermal to an evaporative generation mechanism occurs at a significantly higher I_{max} value as compared with the longer-pulse case. This property enables one to excite bipolar optoacoustic pulses of a large magnitude without evaporation. Most dramatical is the improvement in the magnitude of the valley: it may appreciably surpass both the preceding (leading) peak and the valley that is generated by the longer excitation pulse, since the trailing edge of the shorter excitation pulse has a larger slope. The valleys, in turn, may create cavitation bubbles if the tensile strength of water is exceeded. Figure 6 presents shadowgrams for excitation pulses of both types when the energy flux density is five times larger than that in Fig. 4. The valley magnitude was estimated under the assumption that the pressure in an optoacoustic pulse is linearly dependent on the energy flux density of an excitation pulse. We obtained 35 atm for the longer excitation pulse and 60 atm for the shorter one. Figure 6 indicates that the cavitation bubbles are generated far more efficiently in the latter case in spite of the fact that the absorbed energy is 15 times smaller. Also, the bubbles are easier to observe in the absence of the periodic structure, which arises with the longer pulse.

To sum up, this study revealed a wealth of complicated physical processes accompanying the interaction between the radiation of a CO_2 laser and an open water surface. It has been found that the multilayer periodic structure, which follows an optoacoustic pulse, is caused by the intensity modulation at the trailing edge of an excitation pulse. It would be worthwhile to undertake further experimental investigations into the nature of the periodic structure.

REFERENCES

1. I. I. Komissarova, G. V. Ostrovskaya, V. N. Philippov, *et al.*, Proc. SPIE, No. 3099, 146 (1996).
2. I. I. Komissarova, G. V. Ostrovskaya, V. N. Filippov, *et al.*, Zh. Tekh. Fiz. **67**, 138 (1997) [Tech. Phys. **42**, 247 (1997)].
3. D. C. Emmony, I. Geerken, and H. Klein-Baltink, J. Acoust. Soc. Am. **73**, 220 (1983).
4. D. C. Emmony, Infrared Phys. **25**, 133 (1985).
5. S. V. Egerev, L. M. Lyamishev, K. A. Naugol'nykh, *et al.*, Akust. Zh. **31**, 277 (1985) [Sov. Phys.–Acoustics **31**, 163 (1985)].
6. M. W. Sigrist and F. K. Kneubuhl, J. Acoust. Soc. Am. **64**, 1652 (1978).
7. T. A. Dunina, S. V. Egerev, L. M. Lyamishev, *et al.*, Akust. Zh. **25**, 60 (1979) [Sov. Phys.–Acoustics **25**, 32 (1979)].
8. G. V. Dreĭden, Yu. I. Ostrovskiĭ, A. M. Samsonov, *et al.*, Zh. Tekh. Fiz. **59**, 203 (1989) [Sov. Phys.–Tech. Phys. **34**, 122 (1989)].
9. V. É. Gusev and A. A. Karabotov, *Laser Optoacoustics* (Nauka, Moscow, 1991).

Translated by A. A. Sharshakov

Technical Physics, Vol. 45, No. 2, 2000, pp. 209–213. Translated from Zhurnal Tekhnicheskoĭ Fiziki, Vol. 70, No. 2, 2000, pp. 69–73.

OPTICS, QUANTUM ELECTRONICS

Spontaneous Switching in Cavityless Optically Bistable Systems

T. M. Lysak and V. A. Trofimov

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

Received July 1, 1998

Abstract—Relaxation cavityless optical bistability is studied when a semiconductor is exposed to light quanta with an energy exceeding the transition energy and the free-electron relaxation time depends on temperature. The numerical analysis of dynamic equations for temperature and free-electron density demonstrates that 11 regimes of light–semiconductor interaction are possible. The regimes are classified in terms of the features of temperature–intensity characteristic curves, namely, by the number and arrangement of instability and quasi-stability segments and by the presence or absence of bistability. It is found that both the upper and the lower branch of a bistable characteristic curve may be unstable, which results in spontaneous switchings between the respective states. Similarly, if the upper state is unstable and the lower state is stable, a small perturbation of the parameters (free-electron temperature and density) near the former causes the transition to the oscillatory regime. If the upper state is unstable and a small perturbation of the lower (stable) state excites a damped oscillation, spontaneous switchings result if the fluctuating temperature exceeds some critical value for the lower state. © *2000 MAIK "Nauka/Interperiodica".*

INTRODUCTION

Optical bistability has remained a focus of interest for many years [1, 2]. It is regarded as an avenue for optical computing and bulk data storage. However, such systems are still unable to provide desired switching performance, which urges the quest for novel ways to implement optical bistability. It is also worth bearing in mind that the simultaneous action of multiple physical factors may result in a tricky evolution of switching waves. In particular, the process may be fraught with spatial and temporal instabilities [1–6]. Some of them stem from diffraction [3, 4], whereas others are produced by intrinsic fluctuations. The stability of stationary states is essential for the reliability of a data storage. A spontaneous switching to another state destroys useful information.

We have studied cavityless optical bistability in a semiconductor exposed to a light beam whose aperture was wide to avoid diffraction. The relaxation time of photogenerated carriers was considered as a function of temperature. It has been revealed that the system may be unstable. In those studies, the instability was found only on the upper branch of a bistable temperature vs. the input intensity curve. By contrast, this theoretical study demonstrates that the loss of stability may occur on both branches. Furthermore, due to temperature fluctuations, oscillations may develop even if the characteristic curve has such segments where only damped oscillations are observed. These findings were obtained within a lumped-parameter model of the semiconductor.

FORMULATION OF THE PROBLEM AND THE METHOD OF COMPUTATION

The photogenerated carriers are electrons excited by photons whose energy exceeds the corresponding transition energy. The excitation is followed by relaxation with which heat is released. The relaxation time is a function of temperature. The processes are described by the dimensionless equations

$$\begin{aligned} \frac{dn}{dt} &= I(t) - n\left(\frac{1}{\tau_r(T)} + \frac{1}{\tau_D}\right), \\ \frac{dT}{dt} &= q\frac{n}{\tau_r(T)} - T, \quad t > 0 \end{aligned} \tag{1}$$

subject to the initial conditions for the temperature and density of the free electrons

$$n|_{t=0} = n_0, \quad T|_{t=0} = T_0. \tag{2}$$

Here, n is the density of the excited electrons; T is temperature; t is time; $I(t)$ is the input light intensity as a function of time; τ_r is the relaxation time; q indicates what fraction of the absorbed light energy is converted into heat during the relaxation; and τ_D represents the diffusion of the excited electrons. In (1), n normalized to the maximum possible value for given ambient temperature, t is measured in terms of the time for heat removal from the exposed region due to heat conduc-

1063-7842/00/4502-0209$20.00

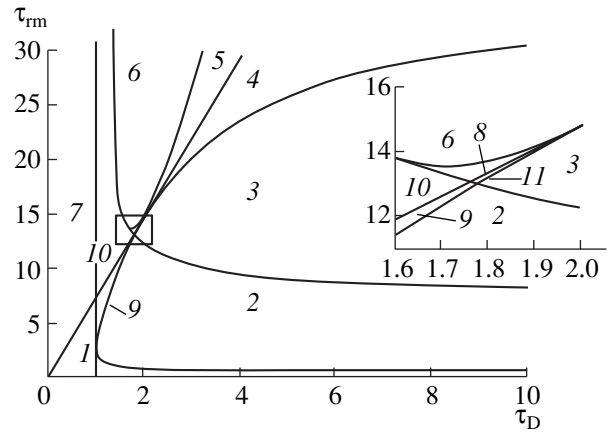


Fig. 1. Regions (*1–11*) of qualitatively different shapes of the *T–I* characteristic curve on the τ_D–τ_{rm} plane, which are distinguished by the number and arrangement of instability and quasi-stability segments and by the presence or absence of bistability. The *T–I* characteristics are shown in respective panels of Fig. 2.

tion, and the drain term represents heat conduction and carrier diffusion. The function $I(t)$ is specified as follows:

$$I(t) = I_0\left(1 - e^{-\frac{1}{\tau}}\right), \qquad (3)$$

where I_0 is the stationary value and τ_D provides the smooth and fast transition to the stationary state. Normally, $\tau = 0.1$.

The dependence of the relaxation time on temperature is approximated as

$$\tau_r(T) = \tau_{rm}\exp(-T/\beta), \qquad (4)$$

where β is set to 1. In any case, the value of β can be reduced to 1 by temperature renormalization.

System (1)–(4) was solved numerically with an implicit scheme. An iterative technique was applied to concomitant nonlinear equations. Each iteration was performed under the constant of energy conservation (hence, the scheme is conservative). The implicit scheme was chosen due to a considerable number of steps required (t may be as large as several thousands). Such a decision insures the stability of computation and prevents the accumulation of rounding errors upon integration.

We preceded the computation with the parametric analysis of stationary solutions. This enabled us to delineate regions on the τ_D–τ_{rm} plane, each related to a specific regime of light–semiconductor interaction. The regimes are distinguished by the presence or absence of bistability and by the stability types of the stationary states.

RESULTS AND DISCUSSION

Figures 1 and 2 present the results of the parametric analysis of stationary solutions to system (1)–(4). Remarkably, the conclusions do not depend on q. The τ_D–τ_{rm} plane in Fig. 1 is divided into regions that give qualitatively different *T–I* characteristics shown in respective panels of Fig. 2. The *T–I* characteristics are distinguished by the presence or absence of bistability and by the number and arrangement of segments showing instability or damped oscillations (stable foci).

Depending on τ_{rm} and τ_D, a *T–I* characteristic may be single-valued or bistable. In Fig. 1, the former alternative is realized in regions *1–4*, *9*, and *11*, whereas the latter, in regions *5–8* and *10*. Furthermore, the upper and the lower branch of a characteristic may have various instability segments (regions *3–6*, *8*, and *11* in Fig. 1). The corresponding unstable stationary states are nodes or foci. If they are perturbed, the system evolves in an exponential or oscillatory fashion, respectively. As far as information storage is concerned, quasi-stability segments (damped oscillations) also deserve attention, since the system may switch from a quasi-stable stationary state to an oscillatory regime if the perturbation exceeds a certain level and information will be lost. Characteristics with quasi-stability segments stem from regions *2–6* and *8–11* in Fig. 1.

It should be emphasized that instability occurs whether or not the system possesses bistability. For example, each of the bistable characteristics (regions *5–8* and *10* in Fig. 1) includes an instability segment lying between the two branches (Fig. 2), and some of the single-valued characteristics also have instability segments (see panels *1–4*, *9*, and *11* in Fig. 2). In what follows, we avoid considering the processes on the intermediate instability segment.

Now, we briefly discuss the features of the regions in Fig. 1. For regions *1*, *2*, *7*, *9*, and *10*, the stationary states are stable. However, the *T–I* characteristics for regions *2*, *9*, and *10* include quasi-stability segments (Fig. 2).

For regions *3–6*, *8*, and *11*, the corresponding *T–I* characteristics have instability segments as well as quasi-stability segments. If a characteristic is bistable (panels 5, 6, and 8 in Fig. 2), instability is normally confined to the upper branch. An exception is the characteristic for region *5*: there is a small instability segment in its lower branch as well. This segment lies between an intermediate instability segment and a quasi-stability segment. If we pass on to region *6*, the lower branch becomes free from instability, but a quasi-stability segment remains. For region *8*, instability

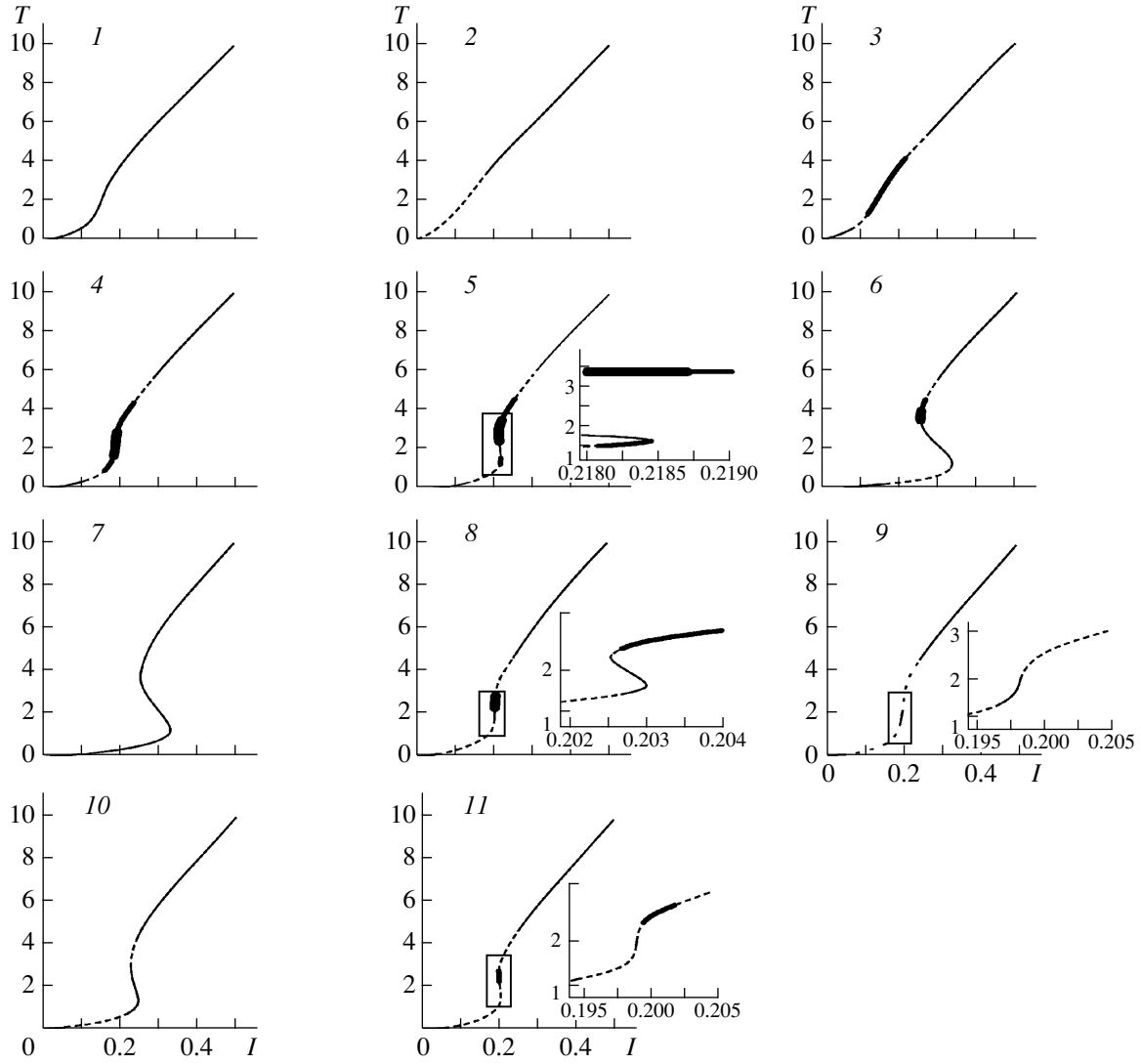


Fig. 2. Temperature T as a single-valued or bistable function of input light intensity I. The panels are labeled as the respective regions in Fig. 1. The thickest lines represent unstable nodes; medium lines, unstable foci; broken lines, stable foci; and thin solid lines, stable states. The curves are computed at $q = 20$. The values of τ_{rm} and τ_D, respectively, are (*1*) 2 and 0.5, (*2*) 2 and 5.4, (*3*) 2 and 20.6, (*4*) 2 and 25.4, (*5*) 30 and 3.5, (*6*) 30 and 2, (*7*) 15 and 0.5, (*8*) 13.3 and 1.75, (*9*) 11.7 and 1.61, (*10*) 12.5 and 1.25, and (*11*) 13.1 and 1.78. If perturbed, an unstable focus produces an oscillatory regime, whereas a stable focus yields damped oscillations.

occurs only on the upper branch. The instability segment here is separated from an unattainable portion in the middle by a quasi-stability segment.

Figure 2 indicates that regions *4*, *5*, *6*, *8*, and *11* of the τ_D–τ_{rm} plane may provide both spontaneously developing oscillations and spontaneous switchings between stable states. Regions *5* and *8* will be the subject of our investigation in what follows, since they were not discovered in our previous studies.

For region *5*, there is an interval of I values within which either (1) both of the stationary states are unstable or (2) only the upper stationary state is unstable, the lower state being stable. Alternative 1 is exemplified by $I = 0.21825$ with $\tau_{rm} = 30$, $\tau_D = 3.5$, and $q = 20$. Alternative 2 is exemplified by $I = 0.2180$ with the same τ_{rm}, τ_D, and q.

Consider alternative 1. A random perturbation of the stationary temperature value initiates an oscillatory regime, which has a pronounced relaxation character. Note that the oscillation develops with any perturbation, however small it is. An example is shown in Fig. 3a. The system evolves from $T_0 = 1.45$, which is close to a stationary value. Before the system goes to the oscillatory regime, it stays in the vicinity of the unstable stationary state for about ten periods. Clearly, the transition time would be comparable with the

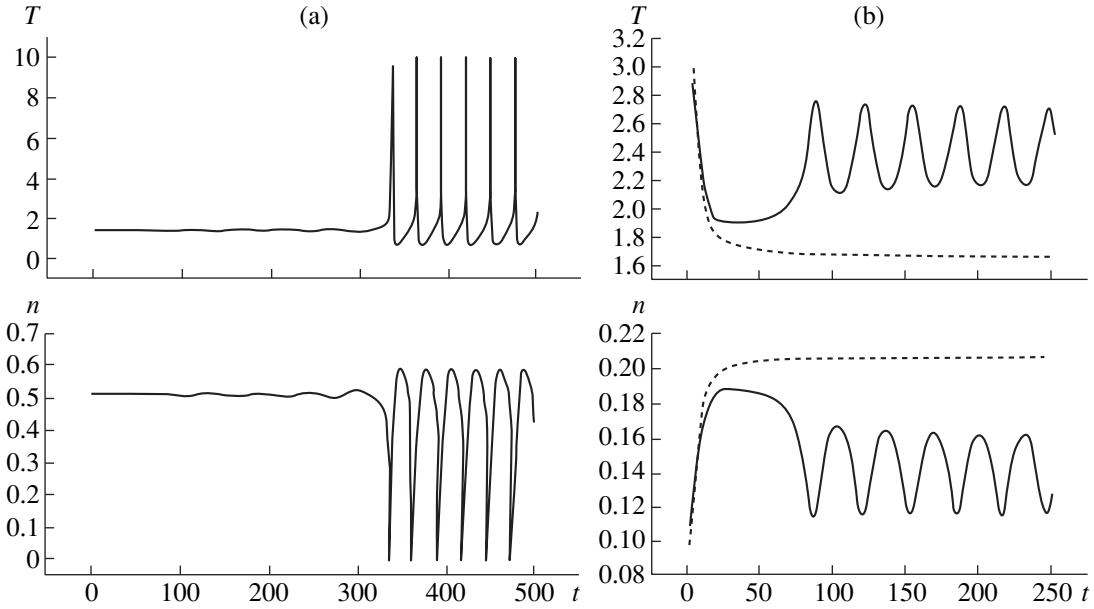


Fig. 3. Development of oscillatory regimes at $q = 20$ for τ_{rm} and τ_D from (a) region *5* and (b) region *8*. Panel (a) is computed at $\tau_{rm} = 30$, $\tau_D = 3.5$, and $I = 0.21825$. Panel (b) is computed at $\tau_{rm} = 13.3$, $\tau_D = 1.75$, and $I = 0.2030$. In (b), $T_0 = 2.9$ (solid curves) or 3.0 (broken curves).

period of the oscillation if the magnitude of the perturbation were sufficiently large.

Consider alternative 2. If the system is in the lower stationary state (which is stable), oscillation develops only from a perturbation that exceeds a certain threshold. For the specified example ($I = 0.2180$ and the above values of the parameters), the system returns to the stationary state if $1.300 \le T_0 \le 1.579$; otherwise, the system performs either periodic switchings between the states (if the amplitude is larger than the switching temperature difference) or the transition to an oscillation around the upper stationary state.

Now, proceed to region *8*. There also exists an interval of I values within which the upper stationary state is unstable. For example, this is the case at $I = 0.2030$ with $\tau_{rm} = 13.3$, $\tau_D = 1.75$, and $q = 20$ (panel *8* in Fig. 2). In this case, the lower stationary state is always stable. Therefore, the system behaves in a way resembling alternative 2 for region *5*: it changes from the lower stationary state to an oscillatory regime if the magnitude of a temperature perturbation is large and brings the temperature above a certain threshold. However, when the system is in the upper stationary state, oscillations arise only with modest perturbations (the region of attraction of oscillation is small); otherwise, the system goes to the lower state; i.e., spontaneous switching occurs. At $I = 0.2030$, the system comes to an oscillatory regime if $1.89 \le T_0 \le 2.90$ (see the solid curves in Fig. 3b) or switches to the lower state if $T_0 > 2.90$ (see the broken curves in Fig. 3b).

If both of the states are quasi-stable (panels *8* and *10* in Fig. 2), spontaneous switching depends on perturbation magnitude. It is not difficult to understand that the system may offer a single oscillatory regime or two such regimes. Which of the two stationary states will be attained can be determined by comparing their regions of attraction.

CONCLUSION

Our analysis of the bistable T–I characteristics has revealed that both of their branches may be unstable. Also, it has been found that an unstable upper branch may coexist with a quasi-stable lower branch. For the first-named case, it has been demonstrated that an oscillatory regime develops from any random perturbation applied to any of the two states. In particular, this causes spontaneous switchings between the states.

If the upper branch is unstable and the lower branch is stable, the system may go to the lower state or to an oscillatory regime whose amplitude may be either comparable to the switching temperature difference or much lower than this value. The evolution of the system is determined by an initial value of temperature, which is prone to fluctuation. If the lower stationary state is perturbed slightly, the system returns to this state (region *8* in Fig. 1). If the magnitude of a perturbation surpasses a certain threshold, the system switches to the upper state and then spontaneous oscillations develop according to the above scenario.

ACKNOWLEDGMENTS

This work was financially supported in part by the Russian Foundation for Basic Research and the pro-

gram *Universities of Russia: Basic Research Projects*, grant no. 5142.

REFERENCES

1. H. M. Gibbs, *Optical Bistability: Controlling Light with Light* (Academic Press, Orlando, Fla., 1985; Mir, Moscow, 1988).
2. N. N. Rozanov, *Optical Bistability and Hysteresis in Distributed Nonlinear Systems* (Nauka, Moscow, 1997).
3. I. G. Zakharova and V. A. Trofimov, Pis'ma Zh. Tekh. Fiz. **21** (16), 43 (1995) [Tech. Phys. Lett. **21**, 603 (1995)].
4. O. A. Gunaze and V. A. Trofimov, Pis'ma Zh. Tekh. Fiz. **22** (16), 1 (1996) [Tech. Phys. Lett. **22**, 642 (1996)].
5. O. S. Bondarenko and V. A. Trofimov, Pis'ma Zh. Tekh. Fiz. **22** (19), 6 (1996) [Tech. Phys. Lett. **22**, 779 (1996)].
6. Yu. N. Karamzin, S. V. Polyakov, and V. A. Trofimov, Pis'ma Zh. Tekh. Fiz. **18** (24), 38 (1992) [Sov. Tech. Phys. Lett. **18**, 809 (1992)].

Translated by A. A. Sharshakov

Technical Physics, Vol. 45, No. 2, 2000, pp. 214–220. Translated from Zhurnal Tekhnicheskoĭ Fiziki, Vol. 70, No. 2, 2000, pp. 74–80.

RADIOPHYSICS

Intermodulation in a Microwave Resonator with a High-Temperature Superconductor

E. A. Vopilkin, A. E. Parafin, and A. N. Reznik

Institute of Microstructure Physics, Russian Academy of Sciences, Nizhniĭ Novgorod, 603600 Russia

Received June 23, 1998; in final form, January 22, 1999

Abstract—Intermodulation, or the generation of a combination-frequency signal, is studied theoretically in the case where the nonlinearity is provided by a high-temperature superconductor placed in a microwave resonator. A general formula that relates the power of an intermodulation signal to the electromagnetic performance of the resonator, the physical parameters of the superconductor, and the intensity of pump waves is derived. A parameter that characterizes the nonlinearity is evaluated from data available from the literature. The potentialities of various resonators for the microwave characterization of the nonlinearities in high-temperature superconductors are assessed. © *2000 MAIK "Nauka/Interperiodica".*

INTRODUCTION

For several years, we have seen markedly increasing interest in the nonlinear properties of high-temperature superconductors (HTSCs) at microwave frequencies. The driving force comes from microwave electronic design, quality control of HTSC films, and basic research in superconductor physics. Nonlinear effects are often evaluated by means of intermodulation,[1] or the generation of a signal at the combination frequency $2\omega_1 - \omega_2$ if an HTSC is pumped at ω_1 and ω_2 [1–3]. Intermodulation tests are very sensitive to variable magnetic fields [2], which favors using them at low microwave powers. Intermodulation measurement involves stripline [1, 3] or cavity [2] microwave resonators, and HTSC nonlinearity is characterized by the combination-frequency output power P_{IM} as a function of the pump-frequency input power P_{in} (usually, $P_{in}^{\omega_1} = P_{in}^{\omega_2} = P_{in}$). For weak signals, $P_{IM} \sim P_{in}^3$. Obviously, the $P_{IM}(P_{in})$ dependence is essentially governed by the electromagnetic performance of the resonator as well as by the intrinsic properties of the HTSC. This fact hampers examination aimed at ascertaining the nature of HTSC nonlinearity or to contrast the performance of HTSC films tested by means of different resonators. It would therefore be desirable to develop an intermodulation theory that could extract [from a measured $P_{IM}(P_{in})$ curve] the true value of a parameter that characterizes microwave nonlinearity and is independent of the performance of a test system. For a stripline resonator, similar problems are addressed in [6, 7]. This paper tackles the problem regardless of resonator specifics. Also, we present a comparative study on the efficiency of various resonators in terms of intermodulation characterization of HTSC films.

Recently, microwave near-field microscopes have been introduced as an instrument to examine HTSCs [8–10]. For example, such a device may be a coaxial resonator whose inner conductor is in contact with an HTSC film [10]. The most important advantage of the microscopes is that their resolution is independent of the wavelength λ. For example, a resolution of ~1 μm at $\lambda \approx 25$ cm was attained [10]. Microwave near-field microscopes are known to be useful in studying the linear microwave properties of HTSCs [8–10]. We will briefly discuss their potential application to HTSC nonlinearity measurements.

ELECTRODYNAMIC PARAMETERS OF RESONATORS

Our investigation will be based on certain general concepts and formulas from the theory of resonators. Figure 1 shows two models: a one-port resonator and a two-port resonator.

One-Port Resonator (Fig. 1a)

An input radiation comes from a waveguide via a coupling element and excites a specific resonator mode. A portion of the input energy is reflected back into the waveguide, whereas another portion is absorbed by resonator walls. The loaded Q-factor of the resonator is

$$Q_L = \frac{Q_U}{1 + \beta}, \quad (1)$$

where Q_U is the unloaded Q-factor, due to the ohmic loss in the walls, and β is the coupling factor such that critical coupling occurs at $\beta = 1$.

[1] Physically, this phenomenon is similar to third-harmonic generation, which was also observed by some workers studying HTSC nonlinearity [4, 5]. We however restrict ourselves to intermodulation.

The electromagnetic energy W stored in the resonator is tied to the input power at the resonance frequency ω in the following way:

$$W = \omega^{-1} P_{in} Q_U \frac{4\beta}{(1+\beta)^2}. \quad (2)$$

The reflection and absorption factors of the resonator are respectively defined in terms of reflected power P_r and dissipated power P_d:

$$\frac{P_r}{P_{in}} = \frac{(1-\beta)^2}{(1+\beta)^2}, \quad \frac{P_d}{P_{in}} = \frac{4\beta}{(1+\beta)^2}. \quad (3)$$

Critical coupling ($\beta = 1$) implies zero-reflection excitation, where W is maximum for a given P_{in} and the input power is absorbed totally.

The unloaded Q-factor is related to the surface resistance R_S of the walls as follows:

$$Q_U = GR_S^{-1}. \quad (4)$$

The maximum magnetic amplitude H_m on the wall is related to W as follows:

$$H_m = \eta_H \sqrt{\omega W} = \eta_H \frac{2\sqrt{3}}{1+\beta} \sqrt{P_{in} Q_U}. \quad (5)$$

Formulas (4) and (5) include a geometric factor G and a field factor η_H. These factors are determined by the distribution of the excited mode and can be expressed in terms of the geometric parameters of a particular resonator.

Two-Port Resonator (Fig. 1b)

Input power is transmitted to the output waveguide, absorbed in the resonator, and reflected into the input waveguide in a certain proportion. The performance of the resonator can be expressed through S parameters:

$$S_{11} = 1 - \frac{2\beta_1}{1+\beta_1+\beta_2}, \quad S_{12} = \frac{2\sqrt{\beta_1\beta_2}}{1+\beta_1+\beta_2}, \quad (6)$$

where $\beta_{1,2}$ are the input and the output coupling factors, respectively [11].

Consider the symmetric case, where $\beta_1 = \beta_2 = (1/2)\beta$. According to [1, 12], the reflection, the transmission, and the absorption factor are respectively expressed as

$$\frac{P_r}{P_{in}} = \frac{1}{(1+\beta)^2} = (1-r_v)^2,$$

$$\frac{P_t}{P_{in}} = \frac{\beta^2}{(1+\beta)^2} = r_v^2, \quad (7)$$

$$\frac{P_d}{P_{in}} = \frac{2\beta}{(1+\beta)^2} = 2r_v(1-r_v),$$

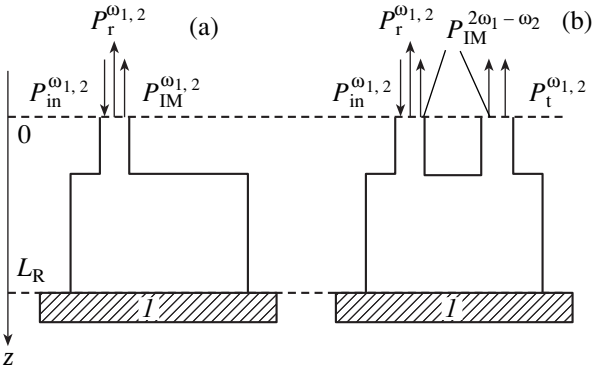


Fig. 1. Schematic diagram of intermodulation measurements with (a) the one-port resonator and (b) the two-port resonator.

where r_v is the voltage insertion ratio, which is used instead of β.

Formula (4) holds as well, whereas (5) is replaced by

$$H_m = \eta_H \frac{\sqrt{2\beta}}{1+\beta} \sqrt{Q_U P_{in}} = \eta_H \sqrt{2r_v(1-r_v) Q_U P_{in}}. \quad (8)$$

Thus, critical coupling ($\beta = 1$ and $r_v = 1/2$) maximizes H_m and distributes P_{in} so that 50% are absorbed in the resonator, 25% are reflected, and 25% are transmitted.

NONLINEAR RESPONSE OF A RESONATOR

Consider the intermodulation characterization of an HTSC film with a one-port resonator (Fig. 1a). Being a conducting wall of the resonator, the film is the only source of nonlinear waves. The resonator is pumped at ω_1 and ω_2 with equal powers, P_{in}. The frequencies ω_1, ω_2, and $2\omega_1 - \omega_2$ lie within the same passband of a mode: $2(\omega_1 - \omega_2)/\omega_1 \ll Q_L^{-1}$. The magnetic and the electric field excited on the film surface at ω_1 and ω_2 have the tangential components

$$H_R(\mathbf{r}) = H_R^m \varphi_R^H(\mathbf{r}),$$
$$E_R(\mathbf{r}) = Z_S H_R(\mathbf{r}) = Z_S H_R^m \varphi_R^H(\mathbf{r}), \quad (9)$$

where H_R and E_R are the magnetic and the electric amplitude, respectively; H_R^m is the maximum value of H_R; Z_S is the surface impedance of the HTSC film; and $\varphi_R^H(\mathbf{r})$ is the eigenfunction of the resonator mode, normalized so that $(\varphi_R^H)_{max} = 1$.

In (9), we assume without loss of generality that each field is aligned with an axis of an orthogonal basis and that H_R is a real number. Fields (9) with the fre-

quencies ω_1 and ω_2 induce a nonlinear intermodulation electric current, whose density is $j^{e}_{IM}(z, \mathbf{r})\exp(i(2\omega_1 - \omega_2)t)$. We assume that the current is aligned with $\mathbf{E}_R$. The complex amplitude of a surface intermodulation current is defined as $\xi^{e}_{IM}(\mathbf{r}) = \int_0^t j^{e}_{IM}(z, \mathbf{r})dz$, where t is the thickness of the HTSC film. The current ξ^{e}_{IM} is related to the pump magnetic fields at ω_1 and ω_2 as follows:

$$\begin{aligned}\xi^{e}_{IM}(\mathbf{r}) &= \alpha[H_R^{\omega_1}(\mathbf{r})]^2[H_R^{\omega_2}(\mathbf{r})]^* \\ &= \alpha(H_R^{m})^3[\varphi_R^{H}(\mathbf{r})]^3.\end{aligned} \tag{10}$$

Formula (10) is independent of nonlinearity mechanisms in a well-conducting medium if the medium satisfies the following constraints: (1) The conductivity tensor is diagonal, the medium being isotropic in the horizontal plane. (2) For H_R, the nonuniformity scale in the transverse direction, governed by $\varphi_R^{H}(\mathbf{r})$, is much larger than that in the z-direction: namely, the thickness of the skin layer [hence, (9) holds at each of the frequencies]. (3) The nonlinearity is local; i.e., the constitutive relation states that the current density at any point depends on the field at this point only. The constraints imply that α in (10) is determined by no more than the electromagnetic properties of the HTSC film and the type of its nonlinearity (the latter will not be specified in this study). Due to (10), we define the characteristic intermodulation magnetic field in the microwave region as $H_c^{IM} = |\alpha|^{-1/2}$. The field is tied to the microwave-current density $j_c^{IM} \approx H_c^{IM}/t$ if the London penetration depth is greater than t. According to our approach, measured values of P_{IM} serve as the basis for evaluating H_c^{IM} or j_c^{IM}. The intermodulation magnetic field and the intermodulation current density play the same role as the respective critical values in the stationary case (H_c and j_c): they characterize the physical features of the HTSC film and are independent of measuring-instrument specifics. We hope that this avenue of research will make it possible to contrast the nonlinear properties of HTSC films examined with different instruments. The evaluation of H_c^{IM} or j_c^{IM} would also promote the understanding of nonlinearity mechanisms in the microwave region.

Assume the nonlinearity to be weak so that four-wave mixing dominates in the intermodulation. Then $\xi^{e}_{IM} \ll H_R$ and the nonlinear current can be found with the help of a perturbation technique. In this approach, the intermodulation power P_{IM} (see Fig. 1a) is computed in terms of the linear theory of radiation for a given impressed current in the resonators. Specifically, we apply the theorem of reciprocity to the intermodulation fields, pump waves, and related currents in a waveguide cross section ($z = 0$) and on the HTSC film surface ($z = L_R$). At $z = 0$, the electric and the magnetic tangential component are respectively expressed as

$$\begin{aligned}E(\mathbf{r}) &= Z_W H_{in}(1 + r_0)\varphi_W^{E}(\mathbf{r}), \\ H(\mathbf{r}) &= H_{in}(1 - r_0)\varphi_W^{H}(\mathbf{r}),\end{aligned} \tag{11}$$

where H_{in} and $E_{in} = Z_W H_{in}$ are the maximum amplitudes of the input magnetic and electric field, respectively; Z_W is the wave impedance of the waveguide; r_0 is the amplitude reflection ratio; and $\varphi_W^{E,H}(\mathbf{r})$ are the normalized eigenfunctions of the waveguide mode. Each of fields (11) is assumed to be aligned with an axis from an orthogonal basis. According to the general theory of radiation [13], the fields $E(\mathbf{r})$ and $H(\mathbf{r})$ at $z = 0$ are respectively associated with a magnetic (ξ_m) or an electric (ξ_e) impressed current sheet:

$$\xi_m(\mathbf{r}) = E(\mathbf{r}), \quad \xi_e(\mathbf{r}) = H(\mathbf{r}). \tag{12}$$

Each of current sheets (12) makes a contribution to total electric field (9) on the HTSC film surface:

$$E_R(\mathbf{r}) = E_R^{e}(\mathbf{r}) + E_R^{m}(\mathbf{r}). \tag{13}$$

The theorem of reciprocity implies that

$$\begin{aligned}\int_{S_W} \xi_m(\mathbf{r})H_{IM}(\mathbf{r})ds &= \int_{S_R} \xi^{e}_{IM}(\mathbf{r})E_R^{m}(\mathbf{r})ds, \\ \int_{S_W} \xi_e(\mathbf{r})E_{IM}(\mathbf{r})ds &= \int_{S_R} \xi^{e}_{IM}(\mathbf{r})E_R^{e}(\mathbf{r})ds,\end{aligned} \tag{14}$$

where S_W and S_R are the areas of the waveguide cross section and the HTSC wall, respectively.

Add up the terms of (14), taking into account that

$$\begin{aligned}H_{IM}(\mathbf{r}) &= H_{IM}^{m}\varphi_W^{H}(\mathbf{r}), \\ E_{IM}(\mathbf{r}) &= E_{IM}^{m}\varphi_W^{E}(\mathbf{r}) = Z_V H_{IM}^{m}\varphi_W^{E}(\mathbf{r}).\end{aligned} \tag{15}$$

Based on (9)–(13), we obtain

$$\begin{aligned}&\alpha Z_S(H_R^{m})^4\int_{S_R}[\varphi_R^{H}(\mathbf{r})]^4 ds \\ &= 2Z_W H_{in}H_{IM}^{m}\int_{S_W}\varphi_W^{E}(\mathbf{r})\varphi_W^{H}(\mathbf{r})ds.\end{aligned} \tag{16}$$

Now, we introduce the following parameters of the waveguide and the resonator:

the effective waveguide cross-sectional area,

$$S_V^{ef} = \int_{S_W}\varphi_W^{E}(\mathbf{r})\varphi_W^{H}(\mathbf{r})ds; \tag{17}$$

Table 1. Expressions for resonator parameters ($u = (d/\pi w)|\ln(2\pi\zeta/d)|$ and $\mu_0 = 4\pi \times 10^{-7}$ H m^{-1})

Resonator type and excited mode	S_R^{ef}	G	η_H	γ
Cylindrical, TE_{011}	$\pi R^2 \dfrac{J_0^2(A_{01})}{(J_1^m)^2}$	$\dfrac{\omega\mu_0 L}{2}\dfrac{1+\left(\dfrac{A_{01}L}{\pi R}\right)^2}{1+\dfrac{L}{R}\left(\dfrac{A_{01}L}{\pi R}\right)^2}$	$\sqrt{\dfrac{4\left[1+\left(\dfrac{A_{01}L}{\pi R}\right)^2\right]^{-1}}{\omega\mu_0 L S_R^{ef}}}$	0.252
Coaxial, *TEM*	$2\pi\rho^2\ln(R/\rho)$	$\omega\mu_0\rho\ln(R/\rho)$	$\sqrt{\dfrac{4}{\omega\mu_0 L S_R^{ef}}}$	$\dfrac{1}{2\ln(R/\rho)}$
Stripline, *TEM*	$\dfrac{\zeta L w}{\pi d}(1+u)$	$\dfrac{\omega\mu_0 d}{2(1+u)}$	$\sqrt{\dfrac{1+u}{\pi^2\omega\mu_0 d S_R^{ef}}}$	$\dfrac{3d}{4\pi w}\dfrac{1+\dfrac{2\pi^2\zeta w}{d^2}(1+u)}{(1+u)}$

the effective HTSC-wall area,

$$S_R^{ef} = \int_{S_R} [\varphi_R^H(\mathbf{r})]^2 ds; \tag{18}$$

and the resonator-mode factor,

$$\gamma = [S_R^{ef}]^{-1} \int_{S_R} [\varphi_R^H(\mathbf{r})]^4 ds. \tag{19}$$

Notice that $\gamma < 1$, since current distribution (10) deviates from $\varphi_R^H(\mathbf{r})$.

Finally, H_{IM}^m and H_{in} appearing in (16) also enter into the relations

$$P_{in} = (1/2) S_W^{ef} Z_W H_{in}^2, \quad P_{IM} = (1/2) S_W^{ef} Z_W |H_{IM}|^2. \tag{20}$$

Due to the relation between H_R^m and P_{in} [see (5)], equation (16) gives the intermodulation power radiated from the resonator into the waveguide:

$$P_{IM} = (2H_c^{IM})^{-4} \gamma^2 (S_R^{ef})^2 |Z_S|^2 \eta_H^8 \left(\frac{2\sqrt{\beta}}{1+\beta}\right)^8 Q_C^4 P_{in}^3. \tag{21}$$

The two-port resonator (Fig. 1b) can be treated similarly. A final expression for P_{IM} is identical to (21) with $2\sqrt{\beta}/(1+\beta) \longrightarrow \sqrt{2\beta}/(1+\beta) = \sqrt{2r_v(1-r_v)}$.

Thus, formula (21) provides a solution to the problem: it expresses P_{IM} in terms of the resonator parameters (γ, S_R^{ef}, η_H, β, and Q_U) and the intrinsic characteristics of the HTSC film ($|Z_S$ and H_c^{IM}). With a measured value of P_{IM} at hand [see (21)], one can evaluate H_c^{IM} or j_c^{IM}, which characterize the HTSC nonlinearity irrespective of the instrument employed. The approach would be helpful in the quality control of HTSC films and in the examination of the nonlinearity mechanism that manifests itself in the values of H_c^{IM} or j_c^{IM} and in their temperature and frequency dependence. It is pertinent to note that there remains disagreement on this mechanism, as exemplified by [14]. Formula (21) clearly suggests that intermodulation power should be measured at $\beta = 1$, since the power is maximum in this case. Also note that, under the same conditions, the two-port resonator yields a value which is 2^4 times lower than for the one-port device.

EXPRESSIONS FOR RESONATOR PARAMETERS

In this section, we give some expressions for the resonator parameters appearing in (21). The relations are written for resonator types most frequently used in HTSC characterization. They are collected in Table 1.

Cylindrical Resonator with TE_{011}

Such resonators, as well as dielectric resonators (which are similar in field distribution), were used in [2, 15–17]. In cylindrical coordinates, the magnetic components are

$$H_r(r,z) = H_m \frac{J_1(\chi r)}{J_1^m} \cos\left(\frac{\pi z}{L}\right) \tag{22}$$

and

$$H_z(r,z) = \frac{A_{01}L}{\pi R} H_m \frac{J_0(\chi r)}{J_1^m} \sin\left(\frac{\pi z}{L}\right), \tag{23}$$

where J_0 and J_1 are Bessel functions of order 0 or 1, respectively; J_1^m is the maximum value of $J_1(x)$; L and R are the height and radius of the cylinder, respectively; $\chi = A_{01}/R$; and $A_{01} \approx 3.832$ is a solution of $J_1(x) = 0$.

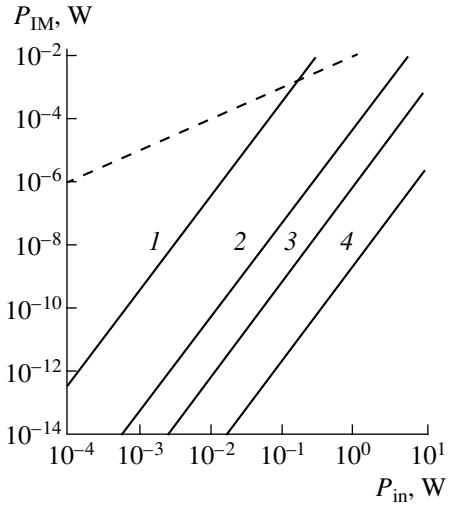


Fig. 2. $P_{IM}(P_{in})$ curves for the (*1*) stripline, (*2*) dielectric [2, 5], (*3*) coaxial, and (*4*) cylindrical HTSC resonator. The dashed line depicts the case $P_r = 10^{-2}P_{in}$.

The expression for G (see Table 1) is derived under the assumption that the resonator is made of a normal metal except for one end face, which is an HTSC film with a vanishing ohmic loss.

Coaxial Resonator with TEM

These resonators are used in the linear case as a part of near-field microscopes [9, 10]. Nevertheless, they seem to have great potential for nonlinear studies. Consider the configuration from [10]. One end face is an HTSC film, which is in a conducting contact with a thin inner conductor made of a normal metal (the contact may be ohmic, capacitive, or inductive). The magnetic field is

$$H_\alpha(r, z) = H_m\left(\frac{\rho}{r}\right)\cos\left(\frac{\pi}{L}z\right), \qquad (24)$$

where ρ and R are respectively the radii of the inner and the outer conducting surface ($\rho \le r \le R$) and L is the resonator length. The corresponding expressions in Table 1 are derived under the assumption that $\rho \ll R$. Consequently, resonator loss is due to absorption in the inner conductor only.

Stripline Resonator with TEM

Consider a configuration similar to that studied in [1, 12, 18]. The resonator is built around a half-wave section of a symmetric stripline whose conductors—a narrow strip and wide shields—are made of an HTSC. In Cartesian coordinates with the origin at the center of the strip, the magnetic field is

$$\mathbf{H}(x, y, z) = [H_x(x, y)\mathbf{x}^0 + H_y(x, y)\mathbf{y}^0]\cos\left(\frac{\pi}{L}z\right), \qquad (25)$$

where L is the longitudinal dimension of the stripline (along the z-axis).

The computation of the transverse field $H_{x,y}(x, y)$ faces the challenge that the field sharply peaks near the edges of the strip. A rigorous approach to the edge effect requires using complicated algorithms [18]. In this paper, we give only approximate formulas for the resonator parameters. We assume that $d \ll w$, where d and w are the thickness and the width of the line, respectively. Also, we tackle the edge effect with the approximation suggested in [19], which allows representing the field on the strip surface as follows [20]:

$$H_y(x, y = 0) = 0,$$
$$H_x(x, y = 0) = \begin{cases} H_m\varphi_H(x), & 0 \le |x| \le w/2 - \zeta \\ H_m\varphi_H(w/2 - \zeta), & w/2 - \zeta < |x| \le w/2, \end{cases} \qquad (26)$$

where $\zeta = 2\Lambda^2/t$ (Λ is the London penetration depth of the HTSC, and t is the strip thickness) and

$$\varphi_H(x) = \sqrt{\frac{1 - \exp(-2\pi\zeta/d)}{1 - \exp(-2\pi(w/2 - |x|)/d)}}. \qquad (27)$$

INTERMODULATION PERFORMANCE OF VARIOUS HTSC RESONATORS

We begin with evaluating H_c^{IM}, which appears in (21). The computation will be based on experimental data available from the literature. In [2, 15], a dielectric two-port resonator with the TE_{011} mode in the frequency band of 19 GHz was studied. It had the parameters $\eta_H = 11.6\ A/(m\sqrt{W})$, $Q_U = 2 \times 10^6$ at a temperature of ~4 K, and $S_R^{ef} \approx 5.3 \times 10^{-6}$ m^2. The value of $P_{IM} \approx 1.2 \times 10^{-9}$ W was obtained at $P_{in} \approx 3 \times 10^{-2}$ W, which belongs to a segment where $P_{IM} \sim P_{in}^3$ [2]. The corresponding magnetic field H_m attained a level close to 5.6×10^2 A/m; hence, from (5), $\beta \approx r_v \approx 0.02$. Recall that the imaginary part dominates in $|Z_S|$ if the temperature lies between the points of liquid helium and liquid nitrogen and if $f = \omega/2\pi < 50$ GHz. We therefore have $|Z_S| \approx \mu_0\omega\Lambda(T) \approx 1.5 \times 10^{-2}\ \Omega$. Taking $\gamma = 0.252$ from Table 1 and inserting it into (21), we obtain $H_c^{IM} \approx 1.6 \times 10^4$ A/m. Hence, $j_c^{IM} = H_c^{IM}/t \approx 10^7$ A/cm^2 at the HTSC-film thickness $t \approx 1.6 \times 10^{-7}$ m. Note the good agreement between the estimated value of j_c^{IM} and the critical values j_c of the stationary current density for high-quality HTSC films. This fact suggests that the mechanisms underlying resistance and microwave nonlinearity in HTSCs may have something in common.

Based on the estimate of H_c^{IM}, we computed $P_{IM}(P_{in})$ for various resonators. According to (21),

$$P_{IM} = P_{in}^3/P_c^2.$$

Here,

$$P_c = \left[(2H_c^{IM})^{-2}\gamma S_R^{ef}|Z_S|\eta_H^4\left(\frac{2\sqrt{\beta}}{1+\beta}\right)^4 Q_C^2\right]^{-1} \quad (28)$$

is a critical P_{in} value: $P_{in} = P_c$ yields $P_{IM} = P_{in}$. This value of P_{in} is specific for a resonator. Obviously, the equality $P_{IM} = P_{in}$ must be regarded as a requirement, since (21) is obtained under the condition $P_{IM} \ll P_{in}$. The computed data are depicted in Fig. 2, where curve *2*, taken from [2] is given for comparison. The computation was performed for the frequency band 10 GHz, where $|Z_S| \approx 1.6 \times 10^{-2}$ Ω (T = 77 K) and the surface resistance of copper, R_S^{Cu}, and HTSC, R_S^{SC}, are about 10^{-2} and 10^{-3} Ω, respectively. (R_S^{Cu} was employed for evaluating Q_U of the cylindrical and the coaxial resonator, and R_S^{SC}, of the stripline resonator.) We assumed $\beta = 1$ in each case. The one-port model was applied to the cylindrical and the coaxial resonator, and the two-port model, to the stripline resonator. The respective values of P_c and other parameters appearing in (28) are given in Table 2. In what follows, we briefly assess the potentialities of the resonators for the characterization of HTSC nonlinearity.

The performance of the stripline resonator was computed at $d = 10^{-4}$ m, $w = 3 \times 10^{-4}$ m, $L = 10^{-2}$ m, $t = \Lambda = 2 \times 10^{-7}$ m, and the permittivity of the dielectric $\varepsilon = 10$. Figure 2 (curve *1*) and Table 2 testify that the stripline resonators offer the smallest P_c; i.e., they provide the highest intermodulation power at the same P_{in}. This stems from the edge effect: an increase in field magnitude near the strip edges sharply reduces S_R^{ef} and amplifies η_H (Table 2). This conclusion is supported by experiments [1, 3], where intermodulation was observed even at P_{in} between 10^{-8} and 10^{-5} W. However, stripline resonators cannot be regarded as an appropriate tool to examine the nonlinear properties of HTSCs. The handicap is that the HTSC film is subjected to treatment during the fabrication of the resonator. Furthermore, only the areas where the edge effect occurs are available for measurement, and their performance heavily depends on the configuration of a resonator. Also, the edges of an HTSC film seem to be the site that is most affected by structuring.

Now, let us proceed to the cylindrical resonator. Note that there exists an optimal value of $L/R = \nu$ such that P_{IM} is maximum. According to (21) and Table 1, $P_{IM} \sim \Psi^4(\nu)$, where $\Psi(\nu) = \nu/[\sqrt{1 + k^2\nu^2}\,(1 + k^2\nu^3)]$ with $k = A_{01}/\pi \approx 1.22$. The function $\Psi(\nu)$ has a maximum at $\nu \approx 0.6$. Our computation was performed for a dielectric-unloaded resonator under the optimal condition at $L = 1.8 \times 10^{-2}$ m and $R = 3 \times 10^{-2}$ m. The results are displayed by curve *4* in Fig. 2. They suggest that the measurement of P_{IM} in a cylindrical resonator made of a normal metal requires a larger input power P_{in}. Bear in mind that the reflection of pump waves $P_r^{\omega_{1,2}}$ typically remains at an appreciable level even if β is in close proximity to 1 (matched resonator). Also, the high value of Q_L (>10^4, Table 2) implies that the frequencies ω_1, ω_2, and $2\omega_1 - \omega_2$ must lie in a narrow passband: $\Delta\omega < 10^{-5}\omega_{1,2}$. It is therefore almost impossible to filter out $P_r^{\omega_{1,2}}$ without suppressing the intermodulation signal P_{IM}. In this case, the intermodulation sensitivity of an instrument is limited by a certain fixed level with respect to P_r. Typically, the level is roughly equal to –70 dB [1, 2]. Based on the data in Fig. 2 and assuming that $P_r = 10^{-2}P_{in}$, we arrive at the conclusion that intermodulation measurements with cylindrical resonators require $P_{in} > 1$ W. In addition to a high input power, such resonators call for a large-area HTSC film. Under the specified conditions, the HTSC film must have a diameter of ~6 cm. This value cannot be decreased by means of increasing L/R without violating the optimality condition, which, in turn, increases P_c. For example, at a diameter of ~4 cm, P_c rises more than 50 times. Nevertheless, the capability of such resonators for HTSC characterization can be improved [2, 15]. Specifically, special dielectric fillers can decrease S_R^{ef} and increase η_H. Another technique consists in using a sheath made of a low-temperature superconductor and cryocooling the system. This raises the Q-factor by 2×10^6. These solutions lower the $P_c \sim 1.4 \times 10^2$ W (Fig. 2, curve *2*) but materially add to the cost of the system and make its operation and maintenance much more laborious.

It seems promising to examine HTSC nonlinearity with coaxial resonators that serve as near-field micro-

Table 2. Resonator parameters and their values used in the computations

Parameters	Stripline resonator	Cylindrical resonator	Coaxial resonator
S_R^{ef}, m^2	5.3×10^{-9}	1.3×10^{-3}	2.9×10^{-7}
G, Ω	2.8	7.1×10^2	36
Q_C	2.8×10^3	7.1×10^4	3.6×10^3
γ	7.6×10^{-2}	0.25	0.11
η_H, A/(m$\sqrt{\text{W}}$)	1.8×10^3	1.2	1.1×10^2
β	1	1	1
P_c, W	1.8	1.9×10^4	1.1×10^3

wave microscopes [9, 10]. They provide high resolution, since their S_R^{ef} is governed mainly by the radius ρ of the inner conductor (see Table 1), which can be reduced down to several micrometers. Also, the system is free from the above-mentioned disadvantages of stripline resonators. Importantly, its P_{IM} is independent of the geometric parameters, ρ and R, if $\rho \ll R$, as follows from (21) and Table 1. The data for Fig. 2 (curve *3*) and Table 2 are computed at $\rho = 10^{-4}$ m, $R/\rho = 10^2$, and $L = 1.5 \times 10^{-2}$ m. Notice that a coaxial resonator ranks between stripline and cylindrical resonators in P_c. Furthermore, coaxial resonators offer larger P_{IM} than cylindrical resonators at a Q-factor that is smaller by more than one order of magnitude. This fact makes it possible to change the P_{IM} measurement scheme by filtering out reflected pump waves, since the passband can be expanded to $\Delta\omega > 10^{-4}\omega_{1,2}$ if $Q_L < 10^3$. After the reflected waves are suppressed, intermodulation power can be measured with a sensitive detector operating at $P_{IM} \geq 10^{-13}$ W. According to Fig. 2, this technique allows measurements at $P_{in} > 5$ mW.

CONCLUSION

We presented a theoretical analysis of intermodulation in microwave resonators exploited for examining the nonlinear electromagnetic properties of HTSCs. It has been suggested to describe HTSC nonlinearity by the characteristic value of the intermodulation magnetic field H_c^{IM} or the intermodulation current density j_c^{IM}. These parameters are independent of resonator electromagnetic performance, being functions of HTSC physical properties alone. A general formula for the power P_{IM} of an intermodulation signal has been derived. This formula was used as a tool to evaluate H_c^{IM} or j_c^{IM} for an arbitrarily chosen resonator. The case of an HTSC film was scrutinized on the basis of experimental data available from the literature. The value of j_c^{IM} was estimated at ~10^7 A/cm^2. Remarkably, this value is close to the critical stationary current density j_c. This fact suggests the similarity of mechanisms underlying resistance and microwave nonlinearity in HTSCs. Our formalism was applied to the comparative performance evaluation of various microwave resonators serving as tools to examine HTSCs by means of intermodulation. Stripline, cylindrical, and coaxial resonators were covered. It has been concluded that coaxial resonators used as near-field microscopes are the most promising in this respect. We hope that they will enable a high-sensitive local examination of the nonlinearity in HTSC films.

ACKNOWLEDGMENTS

We are grateful to S.A. Pavlov for valuable discussion.

This work was financially supported within the program "Topical Problems of Condensed-Matter Physics: Superconductivity," project no. 96129, and by the Russian Foundation for Basic Research, project no. 96-02-16997.

REFERENCES

1. D. E. Oates, A. C. Anderson, D. M. Sheen, *et al.*, IEEE Trans. Microwave Theory Tech. **39**, 1522 (1991).
2. W. Diete, M. Getta, M. Hein, *et al.*, IEEE Trans. Appl. Supercond. **7**, 1236 (1997).
3. B. A. Willemsen, T. Dahm, and D. J. Scalapino, Appl. Phys. Lett. **71**, 3898 (1997).
4. C. Wilker, Z.-Y. Shen, P. Pang, *et al.*, IEEE Trans. Appl. Supercond. **5**, 1665 (1995).
5. E. S. Borovitskaya, V. M. Genkin, and G. I. Leviev, Zh. Éksp. Teor. Fiz. **10**, 1081 (1996) [JETP **83**, 597 (1996)].
6. O. G. Vendik, I. B. Vendik, and T. B. Samoilova, IEEE Trans. Microwave Theory Tech. **45**, 173 (1997).
7. T. Dahm and D. J. Scalapino, J. Appl. Phys. **81**, 2002 (1997).
8. M. Golosovsky and D. Davidov, Appl. Phys. Lett. **68**, 1579 (1996).
9. S. M. Anlage, C. P. Vlahacos, S. Dutta, *et al.*, IEEE Trans. Appl. Supercond. **7**, 3686 (1997).
10. I. Takeushi, T. Wei, F. Duewer, *et al.*, Appl. Phys. Lett. **71**, 2026 (1997).
11. V. I. Vol'man, *Handbook on Microwave Stripline Design* (Radio i Svyaz', Moscow, 1982).
12. D. E. Oates, A. C. Anderson, and P. M. Mankiewich, J. Supercond. **3**, 251 (1990).
13. V. V. Nikol'skiĭ, *Electrodynamics and Radio-Wave Propagation* (Nauka, Moscow, 1973).
14. M. Hein, W. Diete, M. Getta, *et al.*, IEEE Trans. Appl. Supercond. **7**, 1264 (1997).
15. D. W. Cooke, E. R. Gray, and P. N. Arendt, J. Supercond. **3**, 261 (1990).
16. W. Diete, B. Aschermann, H. Chaloupka, *et al.*, in *Applied Superconductivity*, Ed. by D. Dew-Hughes (IOP Publishing, Bristol, 1995), Vol. 148, pp. 1107–1110.
17. Z.-Y. Shen, C. Wilker, P. Pang, *et al.*, IEEE Trans. Microwave Theory Tech. **40**, 2424 (1992).
18. D. M. Sheen, S. M. Ali, D. E. Oates, *et al.*, IEEE Trans. Appl. Supercond. **1**, 108 (1991).
19. O. G. Vendik and A. Yu. Popov, Zh. Tekh. Fiz. **63** (7), 1 (1993) [Tech. Phys. **38**, 535 (1993)].
20. K. P. Gaĭkovich and A. N. Reznik, Zh. Tekh. Fiz. **68** (8), 78 (1998) [Tech. Phys. **43**, 948 (1998)].

Translated by A.A. Sharshakov

Technical Physics, Vol. 45, No. 2, 2000, pp. 221–227. Translated from Zhurnal Tekhnicheskoĭ Fiziki, Vol. 70, No. 2, 2000, pp. 81–86.

Angular Characteristics of Processes of Electron Detachment from Negative Ions and Hydrogen Atoms in Gases

G. D. Ved'manov, Yu. G. Lazarev, and V. I. Radchenko

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

Received February 23, 1998; in final form, October 19, 1998

Abstract—The results of experimental measurement of spatial–angular distributions of hydrogen particles (H^-, H^0, H^+) obtained in scattering of a collimated ribbon beam of H^- ions and $H^0(1s)$ atoms in He, Ar, Kr, Xe, H_2, O_2, and CO_2 gas targets for certain values of energy in the range from 0.6 to 15 MeV are reported. The experimental setup and the measurement procedure with an angular resolution of 5×10^{-6} rad are described. The angular characteristics of measured distributions, i.e., full width at half maximum and standard deviation, were determined. It is shown that the shape of distribution for a beam of hydrogen atoms produced by neutralization of H^- ions in a gas target varies with the type and thickness of the target, and the angular spread is smallest for the H_2 target. The variations in the shape of distribution are due to the contribution of scattering processes without changing the charge of particles. © *2000 MAIK "Nauka/Interperiodica".*

INTRODUCTION

There are a number of extensive scientific–technical projects involving the creation of storage rings, tandem accelerators, and meson factories: the beam-heating of thermonuclear plasma, etc., that require information about the variation of the transverse size (phase volume) and the charge composition of an ion beam, which occasionally traverses considerable distances, with the conditions of beam propagation [1]. Scattering of fast particles colliding with target atoms is completely defined by a set of differential cross sections (DCS) for various processes which are or are not accompanied by changes in the charge state of the incident particles (summed over all the final states of the target). These cross sections provide the basis for calculation of all other quantities that are used in particular applications.

The experimental measurement of DCS, especially in the range of collision energies $E > 0.5$ MeV, involves a number of stringent requirements for the angular divergence and the intensity of the initial beam, the resolution of the monitoring system, etc., which makes measurements of DCS extremely time consuming. At the same time, characteristics obtained on the basis of DCS, e.g., total scattering cross section $\sigma(E)$ as a function of collision energy E or characteristic scattering angle $\Theta_{1/2}(E)$, are often used in theoretical and experimental practice. In experimental investigations and accelerator engineering, a characteristic scattering angle $\Theta_{1/2}$ means a full width of angular distribution of the beam particles at half maximum (FWHM). In theory, the value Θ_m at which the product $\sin\Theta(d\sigma(\Theta)/d\Omega)$ reaches its peak is often used as a characteristic scattering angle. As a rule, the values of the angles $\Theta_{1/2}$ and $2\Theta_m$ coincide to within a few percent. The requirements of experiments aimed at measuring total cross sections and characteristic angles are less stringent. In particular, for angular measurements, it is possible to use ribbon ion beams (however, in doing so, the obtained values do not give adequate insight into the form and characteristic angles of DCS).

A beam of hydrogen particles (H^-, H^0, and H^+) is of interest from the standpoint of the aforementioned practical applications and is convenient for theoretical description. DCS of H^- ions and H^0 atoms with and without charge changes at $E > 50$ keV in gaseous targets have not been adequately investigated; primarily, this concerns experimental investigations. In [2], for comparatively low energies $E = 50$–150 keV, angular distributions and characteristic scattering angles of hydrogen atoms produced by neutralization of H^- ions in H_2, He, and Li targets were measured and the dependence of the characteristic scattering angle on target thickness was studied for circular beams. In [3], a method was developed, and differential (with respect to angle) cross sections of neutralization and elastic scattering of H^- ions with energy of 100 and 200 keV in H_2 and He targets were measured with a ribbon beam. In the range of higher collision energies, similar experimental data are absent. Theoretical results of calculations of DCS and characteristic scattering angles in collisions accompanied by neutralization of H^- ions are reported in [4–8] for $E \geq 100$ keV.

In this paper, which continues a series of works [6–9] performed earlier by us using the same experimental complex, we report the results of measurements of spatial–angular distributions of scattered hydrogen particles in processes of one- and two-electron detachment

1063-7842/00/4502-0221$20.00

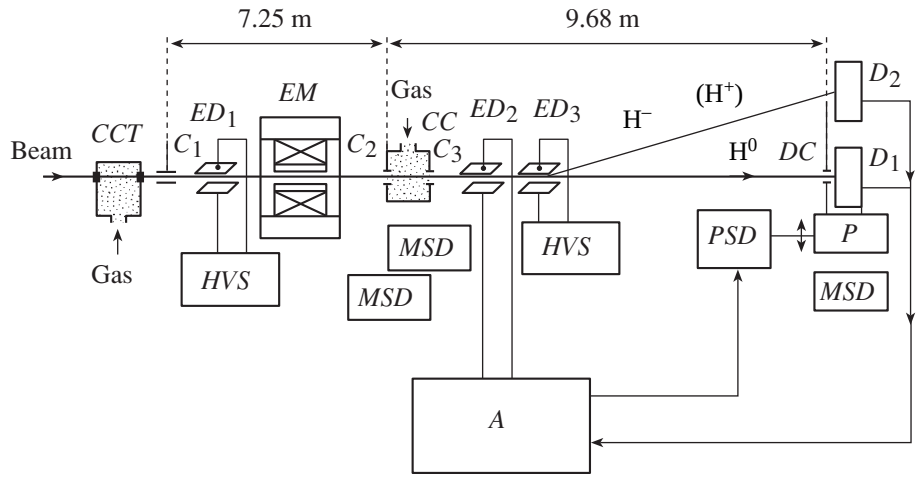


Fig. 1. The experimental setup: CCT, charge changing gas target; C_1–C_3 and *DC*, collimators; $ED_1 \ldots ED_3$, electric deflectors; *HVS*, high voltage source; *EM*, electromagnet; *MSD*, micrometric screw drive; CC, collision chamber; *A*, instrumentation complex; *PSD*, platform step drive; *P*, platform; and D_1 and D_2, detectors.

from H^- ions and $H^0(1s)$ atoms in He, Ar, Kr, Xe, H_2, O_2, and CO_2 gas targets for several energies in the range of 0.6–15 MeV. If the spatial distribution of the flux density of beam particles is determined not only by DCS but also by geometrical conditions of collimation of the initial beam and detection of the scattered beam, this is referred to as spatial–angular distribution.

EXPERIMENTAL

The experimental setup used for the measurement of spatial–angular distributions (SAD) of scattered beams of hydrogen particles and the measurement procedure are described elsewhere [6–10]. In the subsequent presentation, we describe the basic parameters of the setup as applied to the measurement of characteristic scattering angles of a ribbon beam and characteristic features of the measurement procedure for a beam of hydrogen atoms produced by electron detachment from H^- ion in a gas target, which were considered in more detail in [11].

A ribbon beam of hydrogen particles was formed in the setup [6] (schematic diagram is shown in Fig. 1) with the aid of two slit collimators (C_1, C_2) that had a vertical size of 20 μm and a horizontal size of up to 8 mm and that were spaced 7.25 m from each other along the beam path. A beam of H^- ions was directed into the research channel from the accelerator, whereas other charge states of hydrogen particles were produced by one- or two-electron detachment from H^- ions, which occurred in a charge exchange gas target (CCT) placed in the path of the beam upstream of the system of collimators. In the region between C_1 and C_2, the path of the particle beam was determined by electric and magnetic deflectors ED_1 and EM both upstream (not shown in Fig. 1) and downstream of the entrance collimator C_1. A gas target of adjustable thickness and composition was confined by the slits of collimators C_2 and C_3 of collision chamber (CC); these slits limited the gas flux coming out of the CC and provided passage of scattered particles with an angle of deflection in the vertical plane up to 0.4×10^{-3} rad. Vertical movement and rotation of collimators C_2 and C_3 across the beam were performed with the use of micrometric screw drives (MSD). Because of the large path length of the beam inside the CC (390 mm), the latter was shielded from the external magnetic fields by a three-layer cylindrical shield with a reduction factor of more than ten of the transverse (with respect to beam direction) component of the magnetic field. Shielding the beam path from magnetic fields by means of sheets of anisotropic cold-roll transformer steel was also used on most sections of the particle beam path between C_1 and C_2 as well as from C_3 to detector D_1.

SAD of the initial beam and the charge components of the particle beam scattered in the CC were measured in relation to the charge and the energy of particles by three methods.

By the method of scanning the beam with an electrical deflector, which is described in detail in [10], SAD of charged beam components downstream of the CC were measured. Deflector ED_2 was used for scanning the beam by a varying electric field, while deflector ED_3 was used for separating charge components of the beam and directing one of them to the center of detector D_1 with the use of collimator DC. Silicon detectors of the DKDPs-350 type with a sensitive layer up to 1.5 mm thick, which made it possible to carry out the energy analysis of hydrogen particles in a beam with energy up to 15 MeV, were used for counting the particles.

SAD of hydrogen atoms were measured by mechanical scanning, i.e., by means of vertical movement of

the platform P with detector D_1 and the composite collimator DC mounted upstream of the detector; both were attached to the platform [11]. The platform was moved with the aid of platform step drive (PSD) in a range of 9 mm with a step of 1.6 μm. The composite collimator DC of the detector D_1 consisted of two thin plates of equal thickness screening the detector window and forming a horizontal slit 10 μm high in the center of the window. The thickness of the plates was chosen such that the energy of particles traveling through the plates into the detector window was optimal for separating these particles in the spectrum from those of initial energy arriving at the detector through the slit. The vertical size of the ribbon of the chosen component of the scattered beam was invariably much less than the detector window height, and, therefore, the fraction of the particle beam that did not reach the detector due to the displacement of the beam during the process of SAD measurements was negligible in all the measurements. The sum of detected particles (the monitoring number) was used for normalization of the results of measurements in various points of the measured profile. For low energies of hydrogen atoms (less than 1 MeV), the normalization of results of measurements at various points of the measured profile were performed with respect to the number of particles detected with the aid of additional detector D_2, with one of the charge components of the beam directed at the center of this detector by the deflector ED_3.

In the third method used to measure SAD, which is a modification of the first method, the plates of the entrance collimator C_2 of the CC served as a scanning electrical deflector. Polished plates 500 μm thick with a clearance of 20 μm between them are insulated both from one another and from the housing of the CC; they are connected to electronic scanning system *A* (Fig. 1) [10]. This method is more effective than the previous one, but it is applicable only to the charged initial components of the beam.

The results of the measurements of SAD reported in this paper were obtained largely by moving detector D_1 with the composite collimator. Some measurements were performed under similar conditions by all three methods and gave identical results. On evidence derived from numerous measurements, the FWHM angular divergence of the beam in the vertical plane is equal to 3×10^{-6} rad, whereas the FWHM angular resolution of the spread function is 5×10^{-6} rad.

PROCESSING OF EXPERIMENTAL DATA

SAD measured with the above methods are presented in the form of spectra, i.e., as the dependence of the number of particles $Y(\beta)$ reaching the detector through the slit of the collimator DC for a given value of the monitoring number (the number of counts) on the value of the angular shift of collimator β from the center of the distribution (examples are shown in Figs. 2 and 3). The spectra were processed mathematically with the aim of selection of a function approximating these spectra and calculation of the parameters of this function. From ample numerical evidence, it was determined that the spectrum of hydrogen particles is described quite satisfactorily by a function in the form

$$Y(\beta) = Y_1/(1 + \beta^2/\omega^2)^{1/2} + Y_2/(1 + \beta^2 + \omega^2)^{3/2}, \quad (1)$$

whose parameters, i.e., Y_1, Y_2, and ω, are determined by adapting to the experimental data with the method of least squares.

The full width of the peak at half maximum (FWHM) $\beta_{1/2}$ is most often used as an integral angular characteristic of the peaked distributions. The advan-

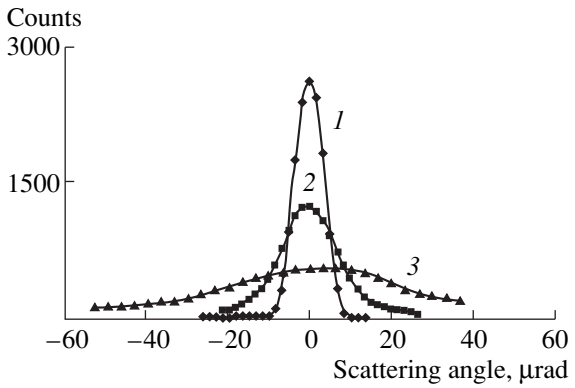


Fig. 2. SAD of hydrogen particles with an energy of 10.4 MeV in scattering of H^- ions without charge changes $(1^-, 1^-)$ attended by one-electron detachment $(1^-, 0)$ and two-electron detachment $(1^-, 1)$: (*1*) the $(1^-, 1^-)$ process, residual gas, and 10^4 counts; (*2*) the $(1^-, 0)$ process, the CO_2 target, $t/T_{max} = 0.09$, and 10^4 counts; and (*3*) the $(1^-, 1)$ process, the CO_2 target, $t/T_{max} = 0.003$, and 3×10^4 counts.

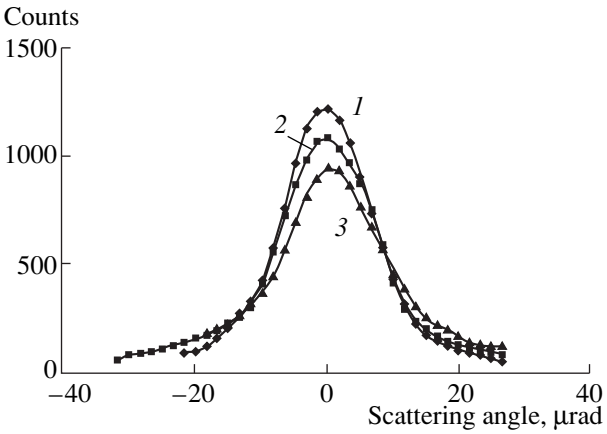


Fig. 3. SAD of H^0 atoms created in scattering of H^- ions with an energy of 10.4 MeV in CO_2 target with thickness t/T_{max} = (*1*) 0.09, (*2*) 0.4, and (*3*) 0.9; 10^4 counts.

tages of this characteristic are the informativity and simplicity of its determination. We use also the FWHM for characterizing SAD. However, the FWHM is only slightly sensitive to changes of the peripheral region of SAD, i.e., to changes of the distribution profile.

It is worth noting how scattering angles are interpreted for different measurement configurations. For the so-called "point" measurement configuration, a detector with the infinitesimal sizes Δx and Δy moving along the vertical axis y, or one with infinitesimal transverse sizes, detects ions of the beam that are scattered in the target at the angle Θ that has both horizontal and vertical components, Θ_x and Θ_y, respectively, with Θ_y equal to the angle β of the detector displacement with respect to the trajectory of the particles that are not scattered, and Θ_x equal to zero (if both the source and the detector are in the same vertical plane). Therefore, $\Theta = \Theta_y = \beta$.

In the case of the "ribbon" configuration, which we use, $\Delta x \gg \Delta y$ and particles with $\Theta_y = \beta$ and all possible values of Θ_x arrive at the detector; therefore, the result of the measurement should be associated with a certain effective scattering angle that is equal to (if the probability of scattering σt is small)

$$\Theta^*(\beta) = \int (\Theta_x^2 + \beta^2)^{1/2} \times \frac{d\sigma(\Theta_x, \beta)}{d\Omega} d\Theta_x \Big/ \int \frac{d\sigma(\Theta_x, \beta)}{d\Omega} d\Theta_x, \quad (2)$$

where $d\sigma(\Theta_x, \beta)/d\Omega$ is the differential cross section, σ is the total cross section, and t is the target thickness.

The value Θ^* virtually coincides with the angle β only for $\beta \gg$ FWHM; therefore, the FWHM of $\beta_{1/2}$ distribution measured in the ribbon configuration turns out to be incorrect (excessively large) in comparison with the FWHM of the differential cross section.

Variance σ^2 (or standard deviation σ) is a statistical characteristic accounting for the shape of the spectrum close to the center of the peak as well as far from it:

$$\sigma^2 = \int Y(\beta)\beta^2 d\beta \Big/ \int Y(\beta) d\beta \quad (3)$$

with the integration performed between the limits from $-\beta_{max}/2$ to $+\beta_{max}/2$.

The function $Y(\beta)$ of the form (1) for $\beta \gg \omega$ slowly decreases with increasing angle β as $Y_1\Omega\omega/\beta$ in the measured range of scattering angles ($\leq$100 μrad); therefore, the standard deviation of this function is $\delta \sim \beta_{max}/\ln(\beta_{max})$ and, thus, is not an invariant characteristic of the scattering process. This fact causes β_{max} to be assigned an arbitrary value in some sense. It seems convenient for us to use the criterion $Y(\beta_{max}/2) \approx Y(0)/10$, having averaged β_{max} over all the spectra that were measured at the same energy, irrespective of the kind and thickness of the target. This angle range corresponds to the range covered by the experiment. The values of β_{max} obtained by the measurement of SAD in the (–1, 0) scattering process are listed in Table 1. If $\beta_{max} \approx 3\beta_{1/2}$, $\beta_{1/2} \approx 2\delta$.

The variance of the spectrum of particles that were not scattered (so-called spread line measured for the residual gas in the CC) was used as the variance of the spread function that is subtracted from the variance of the measured distribution.

As it is shown in the next section, as the target thickness increases, the shape of the spectrum changes; in particular, the fraction of the peripheral regions of SAD ($\beta \geq$ FWHM) increases. For description of this process, it is convenient to use the quantity that is equal to the fraction of the "beam core":

$$W_i(t) = S(\langle\beta_{1/2}\rangle)/S(\beta_{max}), \quad (4)$$

where $S(\beta)$ is the area under the spectrum in the limits from $-\beta/2$ to $+\beta/2$; t is the target thickness; the quantity $\langle\beta_{1/2}\rangle$ is equal to $\beta_{1/2}$ averaged over all the spectra corresponding to the same energy of the beam, as in the case of β_{max}; and i is the charge of the particles.

The processed experimental data of the characteristic scattering angles are listed in Tables 1 and 2. The values listed in these tables are obtained by extrapolating the data to "zero" target thickness; in a number of cases, the measurements were performed only at a target thickness close to t_{max} (see the next section). According to our estimations, in doing so, the standard deviation increases by 15–20% compared to the case of $t = 0$.

EXPERIMENTAL RESULTS AND DISCUSSION

In Fig. 2, typical SAD of the beam of initial H^- ions, hydrogen atoms and protons, generated in the processes of one- and two-electron detachment in CO_2 target are shown. The target thickness is close to a single-collision mode for (–1, 0) and (–1, 1) processes. The measurements of the SAD were performed for a fixed number of particles counted by the detector at each point of the shown profiles. The results shown in Figs. 2–5 are obtained for a scattering of H^- ions with energy of particles of 10.4 MeV.

In Fig. 3, samples of SAD for hydrogen atoms created in the processes of one-electron detachment from H^- ions in CO_2 targets of various thicknesses are shown. It is convenient to use the quantity t_{max} determined individually for each target and energy value as a scale of the target thickness. In the process of neutralization of H^- ions, the fraction $\Phi_0(t)$ of hydrogen atoms peaks at $t = t_{max}$ and then gradually decreases [8]. The values of t_{max} for CO_2 target are given in Table 1, and for other targets, in [8].

In Figs. 4 and 5, the angular characteristics, i.e., standard deviation $\delta(t)$ and beam core fraction $W_i(t)$, are shown as a function of thickness of various targets. The measurement error of these characteristics is equal to

Table 1. Characteristic scattering angles of hydrogen particles in the CO_2 target

Energy, MeV	For the (–1, 0) process		Standard deviation for the process, μrad			t_{max}, 10^{14}/cm^2
	FWHM, μrad	β_{max}, μrad	(–1, 0)	(0, 1)	(–1, 1)	
0.6	49.4	286	56.5			19.2
1.15	36.4	117	22.2			32.6
1.67	28.0	101	18.3	91.4*	185.5*	44.2
5.15	20.0	63	11.3	24.5	30.9	110.4
7.0	15.6	65	9.9	19.8	25.9	141.7
10.4	12.6	43	7.0		20.6	195.3
15.0	9.6	30	5.0	20.5*	21.3*	262.7

* Measured for $t \approx t_{max}$.

5–7%. The data analysis leads to the conclusion that, within the measurement error, the angles of single scattering (in the limit $t \longrightarrow 0$) are the same for all the examined targets; therefore, in Table 1, the data for only a single target (CO_2) are presented. In the thickness region $t \le t_{max}$, a linear growth of $\delta(t)$ with a constant rate that is approximately equal for all the targets, except for H_2, and a corresponding decrease in $W_i(t)$ are seen.

The noted SAD behavior is the result of the fact that the cross sections of (–1, –1) and (0, 0) scattering processes without charge changes accompanying neutralization of H^- ions are comparable in magnitude to the cross section of one-electron detachment (–1, 0) [6], whereas the characteristic angles (FWHM $\Theta_{1/2}$) of these processes, for instance, for H, H_2, and He targets in the energy region E = 0.5–15 MeV satisfy the well-known condition [4, 6, 7]

$$\Theta_{1/2}^{(0,0)} \gg \Theta_{1/2}^{(-1,1)} > \Theta_{1/2}^{(-1,0)}. \quad (5)$$

As t increases, the fraction increases of hydrogen atoms scattered not only in the (–1, 0) transition but also in a certain sequence of (–1, –1), (–1, 0), and (0, 0) collisions and, hence, scattered at a larger angle. The resulting SAD of hydrogen atoms represents a t-dependent superposition of some distributions with various combinations of scattering processes without charge changes that complement the base transition accompanied by electron detachment:

$$(-1, 0);\ (-1, -1) + (-1, 0);\ (-1, 0) + (0, 0);\ (-1, -1) + (-1, 0) + (0, 0). \quad (6)$$

In this case, as t increases, the contribution of the combinations with (0, 0) atom scattering and Φ_0, the fraction of hydrogen atoms, appreciably increase; as a result, a peculiar pedestal in the measured SAD is formed. Further increase of $t > t_{max}$ involves an increase in the angular size of SAD and a gradual decrease in a small-angle part of SAD due to an increase in the multiplicity of the (0, 0) scattering process.

Changing the kind of particles of the neutralizing target results not only in inserting various fractions Φ_0^{max} of hydrogen atoms in the beam [8], but also in changing the shape of SAD. A change to a heavier target is accompanied by an increase in the ratio $\sigma(0, 0)/\sigma(-1, 0)$ and a corresponding increase in the pedestal contribution to the observed SAD at $t = t_{max}$. This pedestal is minimal for a hydrogen target, which reveals itself in the fact that the values of characteristic scattering angles are smaller than those for other targets.

Table 2. Characteristic scattering angles of H^- ions with electron detachment in various targets

Energy, MeV	Standard deviation for the target, μrad						
	H_2	He	O_2	CO_2	Ar	Kr	Xe
0.6		57.6*		56.5*			
1.15		21.9	23.1*	22.2			
1.67	18.1	19.7	18.9	18.3	19.1	17.9	19.1
5.15				11.3			
7.0				9.9			
10.4	7.1	7.0	7.2	7.0	7.2		
15.0	5.1			5.0			
	FWHM for the target, μrad						
	H_2	He	O_2	CO_2	Ar	Kr	Xe
0.6		57.2		49.4			
1.15		33.8	35.6	36.4			
1.67	27.6	29.8	28.0	28.0	29.4	27.4	28.4
5.15				19.8			
7.0				15.6			
10.4	13.2	13.8	13.4	12.6	14.2		
15.0	10.2			9.6			

* Measured for $t \approx t_{max}$.

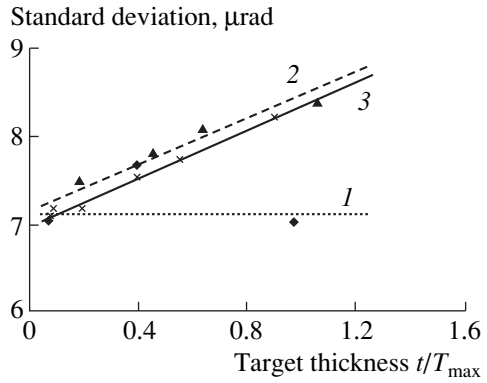


Fig. 4. Standard deviation of SAD of H^0 atoms created in scattering of H^- ions with an energy of 10.4 MeV in (*1*) H_2, (*2*) O_2, and (*3*) CO_2 targets as a function of target thickness; the lines are the result of linear approximation with determination of the coefficients by the method of least squares.

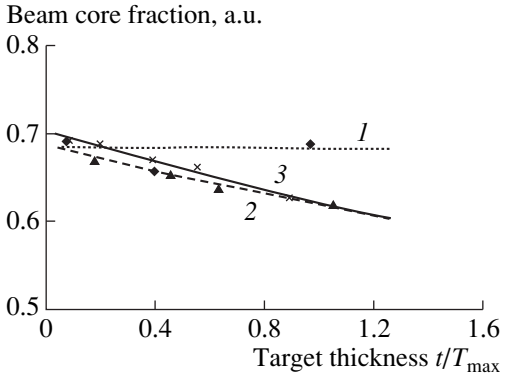


Fig. 5. Fraction of the central region of SAD [see (4)] of H^0 atoms created in scattering of H^- ions with an energy of 10.4 MeV in (*1*) H_2, (*2*) O_2, and (*3*) CO_2 targets as a function of target thickness. The lines are the result of linear approximation with determination of the coefficients by the method of least squares.

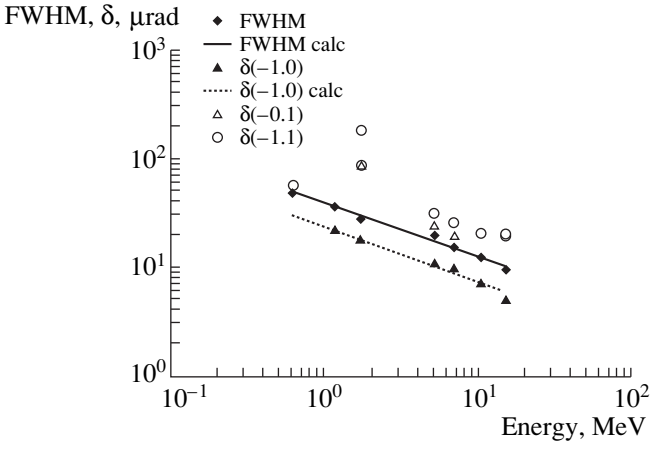


Fig. 6. Energy dependence of the full width of the spectrum at half maximum (FWHM) and standard deviation (δ) of SAD extrapolated to zero thickness of the CO_2 target. The values of δ measured at $t \approx t_{max}$ are encircled. The lines correspond to calculation by (7).

The data presented in Table 1 and Fig. 6 for scattering of H^- ions in a CO_2 target with electron detachment indicate that, if FWHM $\beta_{1/2}$ and standard deviation δ are extrapolated to zero target thickness, we have the same energy dependence:

$$\beta_{1/2} = C_\beta E^{-0.5}, \quad \delta = C_\delta E^{-0.5}, \qquad (7)$$

where $C_\beta = 40 \pm 2$ μrad, $C_\delta = 24 \pm 4$ μrad, and energy E is expressed in MeV.

In Table 1, the characteristic angles ($\beta_{1/2}$ and δ) for various scattering processes with one- or two-electron detachment are compared. In view of the experimental problems of forming beams of a required charge state, some measurements were performed at only one value of target thickness and the corresponding results should be regarded as estimates. Compared to the standard deviation of SAD for one-electron detachment from H^- ions, which for two-electron detachment is three times larger, the standard deviation of one-electron detachment for H^0 atoms is two times larger. The values of the constants C_δ in (7) for (0, 1) and (–1, 1) processes are equal to 54 and 68 μrad, respectively, while the values of the constants C_β for the same processes are 100 and 120 μrad (the values for $E = 1.67$ and 15 MeV, which were measured only in the region $t \approx t_{max}$, were not taken into account in calculation).

In Table 2, the characteristic angles $\beta_{1/2}$ and δ of the scattering spectra of H^- ions with electron detachment (the angles were reduced to zero target thickness) are compared for various targets. From the data listed in Table 2, it follows that, within the measurement error, the values of these angles are independent of the kind of investigated targets for the same ion energy.

CONCLUSIONS

1. Within the experimental error, the characteristic angles δ and $\beta_{1/2}$ for (–1, 0), (–1, 1), and (0, 1) processes determined here for the same value of energy under the conditions of single scattering (reduced to "zero" target thickness) are independent of the kind of target used, which verifies the similar conclusion made in [4, 8] for the (–1, 0) process. The obtained data on the characteristic angles for (–1, 1) and (0, 1) processes confirm the above inference for these processes too. The measured angles depend on the energy as $E^{-1/2}$.

2. Of all the examined targets, a neutralizing hydrogen target is the most preferable with respect to decreasing the angular size of a neutral hydrogen beam [owing to the smaller contribution of the additional (0, 0) scattering, which is of particular importance in the region of target thickness $t \approx t_{max}$].

REFERENCES

1. A. M. M. Todd, D. L. Bryhwiler, and M. F. Reusch, in *Proceedings of First International Workshop "Beam*

Dynamics and Optimization," St. Petersburg, Russia 1994, p. 169.

2. B. A. D'yachkov, V. I. Zinenko, and G. V. Kazantsev, Zh. Tekh. Fiz. **47**, 416 (1977).
3. R. N. Il'in, V. I. Sakharov, and I. T. Serenkov, Zh. Tekh. Fiz. **59**, 124 (1989).
4. Y. T. Lee and J. C. Y. Chen, Phys. Rev. A: Gen. Phys. **19**, 526 (1979).
5. J. A. Johnstone, Nucl. Instrum. Meth. Phys. Res., Sect. B **52**, 1 (1990).
6. V. I. Radchenko, Zh. Éksp. Teor. Fiz. **103**, 40 (1993) [JETP **76**, 22 (1993)].
7. V. I. Radchenko, Zh. Éksp. Teor. Fiz. **105**, 834 (1994) [JETP **78**, 445 (1994)].
8. V. I. Radchenko and G. D. Ved'manov, Zh. Éksp. Teor. Fiz. **107**, 3 (1995) [JETP **80**, 1 (1995)].
9. V. I. Radchenko and G. D. Ved'manov, Zh. Éksp. Teor. Fiz. **107**, 1204 (1995) [JETP **80**, 670 (1995)].
10. G. D. Ved'manov, V. P. Kozlov, and V. N. Kudryavtsev, Prib. Tekh. Éksp. **2**, 47 (1989).
11. G. D. Ved'manov, Yu. G. Lazarev, K. O. Khokhlov, *et al.*, Prib. Tekh. Éksp. **3**, 119 (1997) [Instrum. Exp. Tech. **40**, 402 (1997)].

Translated by M. Fofanov

Technical Physics, Vol. 45, No. 2, 2000, pp. 228–231. Translated from Zhurnal Tekhnicheskoĭ Fiziki, Vol. 70, No. 2, 2000, pp. 87–90.

ELECTRON AND ION BEAMS, ACCELERATORS

Stability of High-current Beams of Relativistic Electrons in Cyclic Systems

V. V. Dolgopolov and Yu. B. Kirichenko

National Science Center, Kharkov Institute of Physics and Technology, 310108 Kharkov, Ukraine

Received 17 July, 1998

Abstract—Stability conditions of high-current thin beams of relativistic electrons against excitation of long-wave oscillations in a stellatron and modified betatron have been investigated theoretically. The influence of self electric and magnetic fields obtained from electron beam delayed potentials has been taken into consideration. The correspondent dispersion relation has been found. The electron beam of the modified betatron has been shown to be always unstable against excitation of the oscillations considered. Necessary and sufficient conditions for the electron beam confinement in a stellatron have been found. © *2000 MAIK "Nauka/Interperiodica".*

INTRODUCTION

A toroidal magnetic field is applied to overcome the Coulomb repulsion of electrons in high-current noncompensated cyclic electron beams. A betatron supplied with a strong magnetic field is named the modified betatron [1]. However, there exist substantial losses of electrons in a such device provided that a mismatch takes place between energy of electrons and the betatron magnetic field. To solve the problem, it was proposed in [2] to supply the modified betatron with the coils of stellarator type. Such a device was called a stellatron.

The behavior of accelerated electrons in cyclic systems was studied theoretically in [3–9]. The stability of narrow electron beams in a stellatron was investigated in [5]. However, the self fields were not taken into consideration completely in [5]. The self fields were considered more consequently in the course of investigation of long-wave oscillations in [9]. But the last work deals with a number of incorrect simplifications.

A more correct approach to describing the relativistic high-current electron beams has been undertaken in the present work. Unlike [9], self fields produced by an electron beam are calculated with the help of delayed potentials; a drift approximation is not used to describe the electron motion.

SELF FIELDS PRODUCED BY ELECTRON BEAM

After the authors of [7–9], we use a system of axes x, y, θ, which is connected with the pseudotoroidal coordinates r, ϑ, θ by the relations

$$x = r\cos\vartheta, \quad y = r\sin\vartheta, \quad R = R_0 - x, \qquad (1)$$

where r and R are the short and long radii, respectively, R_0 is the radius of the magnetic axis of the stellarator field, and ϑ and θ are short and long azimuth angles, respectively.

The electron beams of small transversal sizes will be considered below, the density of electric charge $\rho(x, y, \theta, t)$ and electron beam density $\mathbf{j}(x, y, \theta, t)$ of which can be written in the following form:

$$\rho(x, y, \theta, t) = -en(x, y, \theta, t),$$

$$\mathbf{j}(x, y, \theta, t) = -en(x, y, \theta, t)\mathbf{v}(\theta, t), \qquad (2)$$

$$n(x, y, \theta, t) = N(\theta, t)\delta(x - x(\theta, t))\delta(y - y(\theta, t)),$$

where $n(x, y, \theta, t)$ and $N(\theta, t)$ are volume and linear electron densities; $\mathbf{v}(\theta, t)$ is the speed of the electrons; $x(\theta, t)$ and $y(\theta, t)$ are the coordinates in the cross section of the electron beam; and e is the electronic charge ($e < 0$).

Taking into account relation (2), we obtain the following formulas for delayed potentials produced by the electron beam:

$$\varphi(\mathbf{r}, t) = -e\int d\theta' \frac{F(\theta', \tau(t, \mathbf{r}))}{G(\mathbf{r}, \theta', \tau(t, \mathbf{r}))}, \qquad (3)$$

$$\mathbf{A}(\mathbf{r}, t) = -\frac{e}{c}\int d\theta' \frac{F(\theta', \tau(t, \mathbf{r}))\mathbf{v}(\theta', \tau(t, \mathbf{r}))}{G(\mathbf{r}, \theta', \tau(t, \mathbf{r}))}, \qquad (4)$$

where

$$F(\theta', \tau(t, \mathbf{r})) = R(\theta', \tau)N(\theta', \tau), \qquad (5)$$

$$G(\mathbf{r}, \theta', \tau(t, \mathbf{r})) = g - \frac{1}{c}\dot{x}(\theta', \tau)\Big[x(\theta, t) - x(\theta', \tau) + 2R(\theta, t)\sin^2\frac{\theta - \theta'}{2}\Big] - \frac{1}{c}\dot{y}(\theta', \tau)[y(\theta, t) - y(\theta', \tau)], \qquad (6)$$

$$g = \Big[(x(\theta, t) - x(\theta', \tau))^2 + (y(\theta, t) - y(\theta', \tau))^2$$

1063-7842/00/4502-0228$20.00

$$+ 4R(\theta, t)R(\theta', \tau)\sin^2\frac{\theta - \theta'}{2}\Big]^{-1/2},$$

where $\tau(t, \mathbf{r}) = t - g/c$; $x(\theta, t)$, $y(\theta, t)$, θ are the coordinates of the observation point; $x(\theta', \tau)$, $y(\theta', \tau)$, θ' are the coordinates of the point of the electron beam at the time τ, $R(\theta, t) = R_0 - x(\theta, t)$, $R(\theta', \tau) = R_0 - x(\theta', \tau)$; and the partial derivative with respect to τ is denoted by the points above the letters in (6).

In the equilibrium, the electrons move along a circle with the radius $\bar{R} = R_0 - \bar{x}$ (where x is the equilibrium value of the coordinate x, which will be found below; and $x \ll R_0$), the speed of electrons $\bar{v}$ in absolute value being constant. The equilibrium linear electron density $\bar{N}$ does not depend on t and θ. Electric and magnetic fields, $\mathbf{E}(\mathbf{r}, t)$ and $\mathbf{H}(\mathbf{r}, t)$, which correspond to delayed potentials (3) and (4), were calculated by us for zero- and first-order approximations on the basis of the deviations of the electron beam from the equilibrium. It can be shown that, because of geometric relations and, also, as a result of the screening effect produced by the oscillations of the electron beam, the integrals should be calculated in the limits from $-\theta_m$ up to θ_m, where

$$\theta_m = \begin{cases} (8r_c/R_0)^{1/2}, & \lambda > (8r_cR_0)^{1/2} \\ \lambda/R_0, & \lambda < (8r_cR_0)^{1/2}, \end{cases} \tag{7}$$

where λ is the wavelength of the oscillations, and r_c is the radius of toroidal chamber ($r_c \ll R_0$).

It is natural to consider that $\theta_m \ll 1$. In the course of expansion of the expressions under integral sign in terms of the small disturbance, there arises an integral of the function $1/|\theta|$, which obeys the following relation:

$$\int_{-\theta_m}^{\theta_m}\frac{d\theta}{|\theta|} \simeq \int_{-\theta_m}^{-\theta_\varepsilon}\frac{d\theta}{|\theta|} + \int_{\theta_\varepsilon}^{\theta_m}\frac{d\theta}{|\theta|} = 2\ln\frac{\theta_m}{\theta_\varepsilon} \equiv 2\Lambda \gg 1. \tag{8}$$

It was taken into account in (8) that the part of the electron beam $-\theta_\varepsilon < \theta < \theta_\varepsilon$ gives a negligible contribution to the field in the observation point chosen, which belongs to the electron beam, provided that $\theta_\varepsilon < \bar{r}_b/R_0$, where $\bar{r}_b$ is the average radius of the cross section of the electron beam. It should be noted that $\theta_m \gg \theta_\varepsilon$, and consequently, uncertainties in the values of θ_m and θ_ε have an insignificant influence on the values of Λ.

Let $\tilde{f}$ denotes a small deviation of a certain value f from its equilibrium value $\bar{f}$:

$$f = \bar{f} + \tilde{f}, \quad |\tilde{f}| \ll |\bar{f}|. \tag{9}$$

Taking into account the assumptions accepted above, we derive from formulas (3) and (4) the following expressions for electric and magnetic fields $\mathbf{E}(\mathbf{r}, t)$ and $\mathbf{H}(\mathbf{r}, t)$ at the observation points which belong to the electron beam:

$$\begin{aligned} E_x &= \bar{E}_x + \tilde{E}_x, \quad E_y = \tilde{E}_y, \quad E_\theta = \tilde{E}_\theta, \\ H_x &= \tilde{H}_x, \quad H_y = \bar{H}_y + \tilde{H}_y, \quad H_\theta = \tilde{H}_\theta, \end{aligned} \tag{10}$$

where

$$\bar{E}_x = \frac{e\bar{N}}{\bar{R}}\Lambda, \quad \bar{H}_y = -\frac{e\bar{N}\bar{v}}{c\bar{R}}\Lambda, \tag{11}$$

$$\tilde{E}_x = \frac{e\bar{N}}{\bar{R}}\Lambda\left(\frac{\tilde{N}}{\bar{N}} + \frac{\tilde{x}}{\bar{R}} + \frac{\ddot{\tilde{x}}}{\bar{R}} + \frac{2\bar{R}}{c^2}\dot{\tilde{v}}_x\right), \tag{12}$$

$$\tilde{E}_y = \frac{e\bar{N}}{\bar{R}}\Lambda\left(\frac{\hat{\tilde{y}}}{\bar{R}} + \frac{2\bar{R}}{c^2}\dot{\tilde{v}}_y\right), \tag{13}$$

$$\tilde{E}_\theta = \frac{2e}{\bar{R}}\Lambda\left(\frac{\partial\tilde{N}}{\partial\theta} - \frac{\tilde{N}}{2\bar{R}}\frac{\partial\tilde{x}}{\partial\theta} + \frac{\bar{v}}{c^2}\left(\dot{\tilde{N}}\bar{R} - \frac{1}{2}\bar{N}\dot{\tilde{x}}\right) + \frac{\bar{N}\bar{R}}{c^2}\dot{\tilde{v}}_\theta\right), \tag{14}$$

$$\tilde{H}_x = \frac{e\bar{N}}{c\bar{R}}\Lambda\left(2\frac{\partial\tilde{v}_y}{\partial\theta} - \ddot{\tilde{y}}\frac{\bar{v}}{\bar{R}}\right), \tag{15}$$

$$\tilde{H}_y = -\frac{e\bar{N}}{c\bar{R}}\Lambda\left(\tilde{v}_\theta + \bar{v}\left(\frac{\tilde{N}}{\bar{N}} + \frac{\tilde{x}}{\bar{R}}\right) + \frac{\partial\tilde{v}_x}{\partial\theta} - \frac{\bar{v}}{\bar{R}}\ddot{\tilde{x}}\right), \tag{16}$$

$$\tilde{H}_\theta = -\frac{e\bar{N}}{c\bar{R}}\Lambda\tilde{v}_y, \tag{17}$$

where $\hat{f} = \partial^2 f/\partial\theta^2 - (\bar{R}^2/c^2)\partial^2 f/\partial t^2$; $\bar{E}_x$, $\bar{H}_y$ are the equilibrium values of the fields; $\mathbf{E}$, $\mathbf{H}$ are the disturbances of the fields; and the partial derivatives with respect to t are denoted by the points above the letters.

DISPERSION RELATION

In the case of long-wave oscillations under consideration, the following relations should be satisfied:

$$\lambda \gg \bar{r}_b, \quad \lambda \gg \frac{2\pi R_0}{m}, \tag{18}$$

where $2\pi R_0/m$ is the period of the stellarator field, and m is an integer.

The second relation in (18) allows us to substitute the expression for the magnetic field of the doubly injected stellarator [2] by a more simple one, which produces the same rotational conversion as the stellarator one [9]. As a result, the magnetic field confining the electron beam will appear in the following form:

$$\begin{aligned} \mathbf{B} &= \mathbf{e}_\theta B_t - \mathbf{e}_\vartheta B_t ms^2 r/R_0 \\ &+ \mathbf{e}_x B_0\beta ny/R_0 + \mathbf{e}_y B_0\beta(1 + nx/R_0), \end{aligned} \tag{19}$$

where $B_t = B_0/(1 - x/R_0)$, $|s| < 1/2$, $|\beta| \ll 1$, $B_0 > 0$, $0 < n < 1$, and $\mathbf{e}_x$, $\mathbf{e}_y$, $\mathbf{e}_\theta$, $\mathbf{e}_\vartheta$ are the unit vectors of the corresponding directions.

The first summand in (19) represents the toroidal magnetic field, the second one answers for the stellarator field, and the third and fourth ones describe the betatron field. When $s = 0$, formula (19) defines the magnetic field of the modified betatron. The Lorenz equation for an electron moving in the fields described by (10) - (17) and (19) takes the following form:

$$\frac{d\mathbf{v}}{dt} = -\frac{e}{m_e}\left(1 - \frac{v^2}{c^2}\right)^{1/2} \times\left\{\mathbf{E} + \frac{1}{c}[\mathbf{v}, \mathbf{H} + \mathbf{B}] - \frac{1}{c^2}\mathbf{v}(\mathbf{vE})\right\}, \quad (20)$$

where $\mathbf{v} = \mathbf{e}_\theta \bar{v} + \tilde{\mathbf{v}}$, m_e is the rest mass of the electron.

Equation (20) should be completed with the continuity equation

$$\frac{\partial}{\partial t}(R(t, \theta)N(t, \theta)) + \frac{\partial}{\partial \theta}(N(t, \theta)v_\theta(t, \theta)) = 0. \quad (21)$$

In the zero-order approximation, equation (20) links the equilibrium values

$$\frac{\bar{x}}{R_0} = \frac{1}{\sigma_-}\left(\frac{\bar{v}}{\omega_0 \bar{R}} - \beta + \frac{2\gamma^2 - 1}{\gamma^2}\frac{c^2\eta}{\omega_0 \bar{v}\bar{R}}\right), \quad (22)$$

where $\eta = e^2\bar{N}\Lambda/(m_e\gamma c^2)$, $\omega_0 = eB_0/(m_e\gamma c)$, $\sigma_- = \beta n - \alpha m s^2$, and $\gamma = (1 - \bar{v}^2/c^2)^{1/2}$.

Having linearized equations (20), (21), we obtain a set of four simultaneous differential equations for the values of $\tilde{x}$, $\tilde{y}$, $\tilde{N}$, $\tilde{v}_\theta$:

$$(1 + 2\eta)\frac{d^2\tilde{x}}{dt^2} + \left(\frac{\bar{v}^2}{\bar{R}^2}\left(1 + \frac{2\gamma^2 - 1}{\gamma^2 - 1}\eta\right) - \frac{\bar{v}\bar{\sigma}_-\omega_0}{R_0}\right)\tilde{x} + \frac{c^2\eta}{\gamma^2\bar{R}^2}\hat{\tilde{x}} + \alpha\omega_0\frac{d\tilde{y}}{dt} + \left(\frac{2\bar{v}}{\bar{R}} + (3 - 2\gamma^2)\frac{\bar{v}}{\bar{R}}\eta + (\gamma^2 - 2)\right. \quad (23)$$
$$\left.\times\left(\beta + \sigma_-\frac{\bar{x}}{R_0}\right)\omega_0\right)\tilde{v}_\theta + \frac{c^2\eta(2\gamma^2 - 1)}{\bar{N}\bar{R}\gamma^2}\tilde{N} = 0,$$

$$(1 + 2\eta)\frac{d^2\tilde{y}}{dt^2} + \frac{\eta\hat{\tilde{y}}c^2}{\gamma^2\bar{R}^2} + \frac{\bar{v}\sigma_+}{R_0}\omega_0\tilde{y} - \alpha\omega_0\frac{d\tilde{x}}{dt} = 0, \quad (24)$$

$$\left(1 + \frac{2\eta}{\gamma^2}\right)\frac{\partial\tilde{v}_\theta}{\partial t} + \frac{\bar{v}}{\bar{R}}\frac{\partial\tilde{v}_\theta}{\partial\theta} + \left(\omega_0\left(\beta + \frac{\bar{x}\sigma_-}{R_0}\right)\right. \quad (25)$$
$$\left.-1 + 2\eta\right)\frac{\bar{v}}{\bar{R}}\frac{d\tilde{x}}{dt} - \frac{c^2\eta}{\bar{R}^2\gamma^2}\frac{\partial\tilde{x}}{\partial\theta} - \frac{\bar{v}\eta}{\bar{R}\gamma^2}\frac{\partial\tilde{x}}{\partial t}$$
$$+ \frac{2c^2\eta}{\bar{R}\bar{N}\gamma^2}\frac{\partial\tilde{N}}{\partial\theta} + \frac{2\bar{v}\eta}{\gamma^2\bar{N}}\frac{\partial\tilde{N}}{\partial t} = 0,$$

$$\frac{\partial\tilde{N}}{\partial t} - \frac{\bar{N}}{\bar{R}}\frac{\partial\tilde{x}}{\partial t} + \frac{\bar{N}}{\bar{R}}\frac{\partial\tilde{v}_\theta}{\partial\theta} + \frac{\bar{v}}{\bar{R}}\frac{\partial\tilde{N}}{\partial\theta} = 0, \quad (26)$$

where $\sigma_+ = \beta n + \alpha m s^2$, $\bar{\sigma}_- = \beta n - \alpha^2 m s^2$, $\alpha = 1/(1 - \bar{x}/R_0)$, and $d/dt \simeq \partial/\partial t + (\bar{v}/\bar{R})\partial/\partial\theta$ in the linear approximation.

Assuming that the dependence of the values of $\tilde{x}$, $\tilde{y}$, $\tilde{v}_\theta$, $\tilde{N}$ on the variables θ and t is defined by the multiplier $\exp[i(\mu\theta - \omega t)]$, we obtain the dispersion equation from the set of simultaneous equations (23)–(26):

$$\sum_{k=0}^{6} A_k \omega_\mu^k = 0, \quad (27)$$

where $\omega_\mu = \omega - \mu\bar{v}/\bar{R}$.

Coefficients A_k depend on the parameters of both the external magnetic field and the electron beam. General expressions are too awkward to be presented here.

ANALYSIS OF THE RESULTS

The stability of the electron beam against excitation of long-wave oscillations, i.e., the conditions for the roots of equation (27) to be real, has been studied by us under some simplifying assumptions, the main ones of which are the following: the values of the parameter η are small; the values of the relativistic factor γ are high; and the frequency of oscillations is relatively small. The aforesaid assumptions impose the following restrictions on the parameters of the problem:

$$2|\beta m|s^2 \ll 1, \quad (28)$$

$$\frac{\mu^2\beta^2\eta}{\gamma^4} \ll 1, \quad (29)$$

$$|\xi| > 1, \quad (30)$$

$$\gamma^2\beta^2 \gtrsim 1, \quad (31)$$

$$\gamma^2 \gg \max\left(4|\xi|, \mu^2\eta, \frac{2\mu^2\eta}{|n + \xi||\xi + 1 - n|}, \right. \quad (32)$$
$$\left.\frac{2\mu^2\eta}{|n + \xi||3 + 2(\xi - n)|}, \frac{2\mu\eta^2}{\gamma^2\beta^2|n + \xi||\xi + 1 - n|}\right),$$

$$\eta \ll \min\left(\frac{1}{12}, \frac{\gamma^2\beta^2|n + \xi|(\xi + 1 - n)^2}{24\mu^2}\right), \quad (33)$$

$$|\omega_\mu| \ll \omega_0, \quad (34)$$

where $\xi = \alpha m s^2/\beta$.

Under conditions of (28)–(34) the coefficients of dispersion relation (27) take the form

$$A_6 = 1, \quad A_5 = 8\omega_0\frac{\mu\beta\eta^2}{\gamma^2}, \quad A_4 = -\omega_0^2,$$
$$A_3 = -\frac{4\mu\beta\eta}{\gamma^2}\omega_0^3, \quad A_2 = \omega_0^4\beta^4(n+\xi)(1-n+\xi), \tag{35}$$
$$A_1 = \frac{2\mu\eta}{\gamma^2}\beta^5\omega_0^5(n+\xi)[3+2(\xi-n)],$$
$$A_0 = \frac{2\mu^2\eta}{\gamma^2}\beta^6\omega_0^6(n+\xi).$$

The following inequalities can be derived from formulae (35) and conditions (28)–(34):

$$\left|A_6\omega_\mu^6\right| \ll \left|A_4\omega_\mu^4\right|, \quad \left|A_5\omega_\mu^5\right| \ll \left|A_4\omega_\mu^4\right|. \tag{36}$$

The inequalities (36) allow us to reduce equation (27) of the sixth degree to the following equation of the fourth degree:

$$\omega_\mu^4 + a_3\omega_\mu^3 + a_2\omega_\mu^2 + a_1\omega_\mu + a_0 = 0, \tag{37}$$

where $a_k = -A_k/\omega_0^2$; $k = 0, 1, 2, 3$.

It follows from formula (22) that, if inequality (33) is fulfilled, the following relation takes place:

$$\frac{\overline{V}}{\overline{R}} \simeq \beta\omega_0. \tag{38}$$

The necessary and sufficient conditions of stability of the electron beam in a stellatron against excitation of the oscillations under consideration [the conditions for the roots of equation (37) to be real] take the form

$$\xi < 0, \tag{39}$$

$$8\mu^2\eta < \beta^2\gamma^2|\xi+n|(\xi+1-n)^2, \tag{40}$$

$$\frac{\mu^2\eta^2}{\gamma^2} < \beta^2\gamma^2\frac{|\xi+n||\xi+1-n|^3}{27(\xi+2-n)^2}. \tag{41}$$

To study the stability conditions for a modified betatron, it should be required that $\xi = 0$ in the formulas for coefficients a_k. The modified betatron appeared to be always unstable against excitation of the long-wave oscillations. In this case the simplifying assumptions are the same as that ones for a stellatron, except relation (32) which should be replaced by

$$\gamma^2 \gg \max\left(3, \frac{2\mu^2\eta}{n(1-n)}, \frac{2\mu^2\eta}{\gamma^2\beta^2 n(1-n)}\right). \tag{42}$$

CONCLUSION

It follows from the results obtained that the beam of relativistic electrons in the modified betatron is always unstable against excitation of the oscillations considered. The necessary and sufficient conditions of stability in the stellatron are inequalities (39)–(41) provided that conditions (28)–(34) are fulfilled. As follows from (39), to stabilize the electron beam, the averaged stellarator and betatron magnetic fields should be oriented in space in a such way that the following condition is fulfilled:

$$m\beta < 0. \tag{43}$$

This means that the component of the vector of the stellarator field averaged over θ and projected onto direction of betatron field should be positive when $R < R_0$.

REFERENCES

1. N. Rostoker, Part. Accel. **5** (2), 93 (1973).
2. C. W. Robertson and A. Mondelli, Phys. Rev. Lett. **50**, 507 (1983).
3. A. A. Kolomensky and A. N. Lebedev, Atomic Energy **7**, 549 (1959).
4. C. A. Kapetanakos, D. Dialetis and S. J. Marsh, Part. Accel., **21** (1), 1 (1987).
5. D. Chernin, Phys. Fluids **29**, 556 (1986).
6. C. A. Kapetanakos and S. J. Marsh, Phys. Fluids **28**, 2263 (1985).
7. V. V. Dolgopolov, Yu. V. Kirichenko, Ya. F. Leleko, *et al.*, Zh. Tekh. Fiz. **65** (6), 141 (1995) [Tech. Phys. **40**, 599 (1995)].
8. V. V. Dolgopolov, Yu. V. Kirichenko, S. S. Romanov, *et al.*, Pis'ma Zh. Tekh. Fiz. **20** (20), 67 (1994) [Tech. Phys. Lett. **20**, 840 (1994)].
9. V. V. Dolgopolov, Yu. V. Kirichenko, S. S. Romanov, *et al.*, Ukr. Fiz. Zh. **39**, 161 (1994).

Translated by A. Lytkin

Technical Physics, Vol. 45, No. 2, 2000, pp. 232–235. Translated from Zhurnal Tekhnicheskoĭ Fiziki, Vol. 70, No. 2, 2000, pp. 91–94.

The Use of Donkin's Formula in the Theory of Energy Analyzers. I

P. G. Gabdullin, Yu. K. Golikov, N. K. Krasnova, and S. N. Davydov

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

Received December 1, 1998

Abstract—Synthesis of field structures with conic equipotentials is considered in terms of Donkin's analytical representation. The hierarchy of these structures is given, examples of equipotential portraits are offered, and the Cauchy problem for symmetric fields is solved. In part II, electronic and optical properties of some such systems promising for energy analyzers and monochromators will be studied. © *2000 MAIK "Nauka/Interperiodica".*

Over the last years, electrostatic fields with equipotentials converging to a point, so-called cone-shaped fields, have attracted much attention. Energy analyzers built around coaxial circular cones [1], wedge prisms [2], sectional circular wedge prisms [3], and dihedral energy analyzers [4] have been reported. All of these specific designs, however, demonstrate the wide potentialities of such systems for corpuscular optics only slightly. Interest in these fields has been aroused for their specific feature, a telescopic effect. Its essence is as follows. Charged isoenergetic particles moving in the cone-shaped fields trace out geometrically similar trajectories. According to the Helmholtz–Lagrange law, when entering into the field, a particle beam parallel in the plane of motion retains parallelism but changes its geometrical sizes. For example, moving toward the common vertex of the cone-shaped potentials, the beam converges; conversely, it diverges when moving in the opposite direction.

Geometrically, the cone-shaped fields can be viewed as the nearest generalization of planar fields; hence, the theory of the former can easily be reduced to the complex variable theory. Donkin's formula [5] provides a simple and elegant relation between these theories. With its help, the general solution of the Laplace equation for Euler-uniform zero-fold potentials can be written through a specific function of complex variable $\omega(x, y, z) = u + iv$ in the form

$$\Phi(x, y, z) = F(\omega) = F\left(\frac{x + iy}{z + \rho}\right) = \phi_1 = i\phi_2, \quad (1)$$

where

$$\omega = \frac{x + iy}{z + \rho}, \quad \rho = \sqrt{x^2 + y^2 + z^2}, \quad (2)$$

and $F(\omega)$ is an arbitrary analytic function.

Here, the relation with the planar fields is obvious. Practically, the cone-shaped fields are usually described in polar coordinates, where a potential Φ is expressed via conventional angular coordinates θ and φ. However, analytically, representation (1) is also very useful, which will be demonstrated in the most obvious cases.

CLASSIFICATION OF ANALYTIC STRUCTURES

Expression (1) can be viewed as the generalization of the "complex potential" concept for planar fields to the case of cone-shaped ones. The representation of the planar fields with a single function of simple complex variable $\omega = x + iy$ makes it possible to conveniently and hierarchically classify field structures by their analytic attributes. The attributes may primarily be integer positive or negative exponents (ω^n), resulting in direct or reciprocal multipoles; an exponential or logarithmic function of ω (e^ω or $\ln\omega$); and a number of combinations from this basis. An extension of this principle to the more complex Donkin argument (2) allows the classification of the cone-shaped fields with specific geometrical and physical interpretations. In the case of the planar fields, a complex potential in the form of a linear function of argument $\omega = x + iy$ is known to generate the field of a parallel-plate capacitor. For the cone-shaped fields, a linear function of Donkin's argument gives rise to a new, nontrivial type of field. Let for simplicity

$$F = \omega = \frac{x + iy}{z + \sqrt{x^2 + y^2 + z^2}}. \quad (3)$$

As a scalar potential, we take the real part of this expression and put

$$\Phi = \frac{x}{z + \sqrt{x^2 + y^2 + z^2}}. \quad (4)$$

Assuming that $x = 0$, we discover that the plane yz is everywhere a zero equipotential except for the semiaxis

1063-7842/00/4502-0232$20.00

$x = y = 0$, $z < 0$, in the vicinity of which multivalence is already possible. To have a better idea of the shape (cross section) of the conic equipotentials, it is appropriate, in this case, to cut their set by the sliding plane $x = x_0$. If $\Phi = \Phi_0$ is the number of an equipotential, its cross section is described by the simple expression

$$z = \frac{x_0}{2}\left(\frac{1}{\Phi_0} - \Phi_0\right) - \frac{\Phi_0}{2x_0}y^2, \quad (5)$$

from which it follows that the equipotentials have a parabolic cross section convex upwards (Fig. 1). As $x_0 \longrightarrow 0$, all of the cross sections degenerate (turn into triangles) and merge into the ray $z < 0$, $x = y = 0$. The equipotential portrait as a whole is symmetric with respect to the plane yz, since function (4) is even in y and odd in x. Since the equipotentials merge not only at the point $x = y = z = 0$ but also on the ray $z < 0$, the potential becomes akin to a specific cone-shaped dipole. Its electronic and optical characteristics will be described in part II. The form $F = (a + ib)\omega$ of the complex potential mixes the real and imaginary parts of (3); however, a rotation about the axis z and the renormalization of the potentials bring these cases into coincidence, and we do not obtain any new data for the field structure.

Consider now the reciprocal potential

$$F = \frac{1}{\omega} = \frac{z + \sqrt{x^2 + y^2 + z^2}}{x + iy} = \frac{x - iy}{z - \sqrt{x^2 + y^2 + z^2}}. \quad (6)$$

Changing the signs of x and z in (3) yields (6); hence, cases (3) and (4) differ only by mirror reflection in the planes $x = 0$ and $z = 0$ and have equally shaped equipotential portraits; in the planar fields, however, the complex potentials $F = (x + iy)\omega$ and $F = 1/(x + iy)$ provide essentially different equipotential portraits.

Integer exponents in the expression for cone-shaped potentials in the form

$$F(\omega) = \left(\frac{x + iy}{z + \sqrt{x^2 + y^2 + z^2}}\right)^n, \quad n = \pm 1, \pm 2, \pm 3, \ldots \quad (7)$$

are another hierarchic elements, both positive and negative values of n providing structures of the same shape. For example, at $n = 2$, we obtain the real-valued potential of the form

$$\Phi = \frac{x^2 + y^2}{\left(z + \sqrt{x^2 + y^2 + z^2}\right)^2}. \quad (8)$$

It is even in x and y and has a singular line: the ray $x = y = 0$, $z < 0$, around which the field goes into infinity and the equipotentials merge. Since the three-dimensional form of the equipotentials is hard to depict for this field, we will show only its sections obtained by cutting with the planes x = const, y = const, and z = const (Fig. 2). Here, zero equipotentials appear as the orthogonal cross formed by the planes $x = y$ and $x = -y$ with the ray $z < 0$ at their intersection excluded.

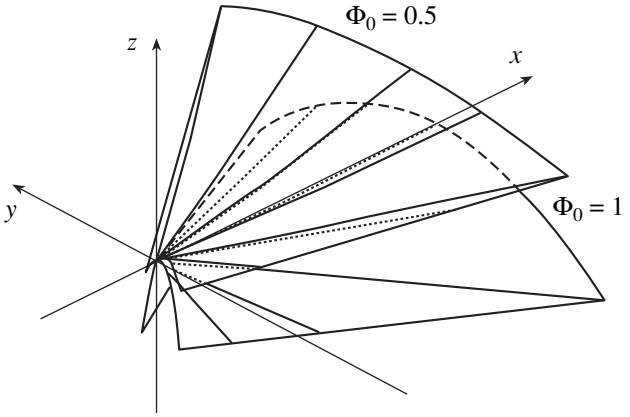


Fig. 1. Equipotentials of field (4) for $\Phi_0 = 1$ and $\Phi_0 = 0.5$.

A remarkable difference of power functions (7) from "planar" potentials in the form $(x + iy)^n$ is that each "conic" power function in the equipotential portrait simultaneously has the analog $(x + iy)^{+|n|}$ for $z > 0$ and $(x + iy)^{-|n|}$ for $z < 0$. This is of importance also in considering series including (7).

The potential in the form

$$F = \ln\omega = \ln\frac{\sqrt{x^2 + y^2}}{z + \sqrt{x^2 + y^2 + z^2}} i\arctan\frac{y}{x} = -\ln\left(\frac{z}{r} + \sqrt{1 + \frac{z^2}{r^2}}\right) + i\arctan\frac{y}{x}, \quad (9)$$

where $r = \sqrt{x^2 + y^2}$, is one more transcendental generalization. Here, the real and imaginary parts are apparently the potential of coaxial cones and that of a dihedral angle, respectively. Such cases are well known [6–8].

A more nontrivial case appears for the "conic" exponential

$$F = \exp(\omega) = \exp\left(\frac{x + iy}{x + \rho}\right) = \exp\left(\frac{x}{z + \rho}\right)\cos\left(\frac{y}{z + \rho}\right) + i\exp\left(\frac{x}{z + \rho}\right)\sin\left(\frac{y}{z + \rho}\right). \quad (10)$$

Here, the real-valued potential

$$\Phi = \exp\left(\frac{x}{z + \sqrt{x^2 + y^2 + z^2}}\right) \times \cos\left(\frac{y}{z + \sqrt{x^2 + y^2 + z^2}}\right) \quad (11)$$

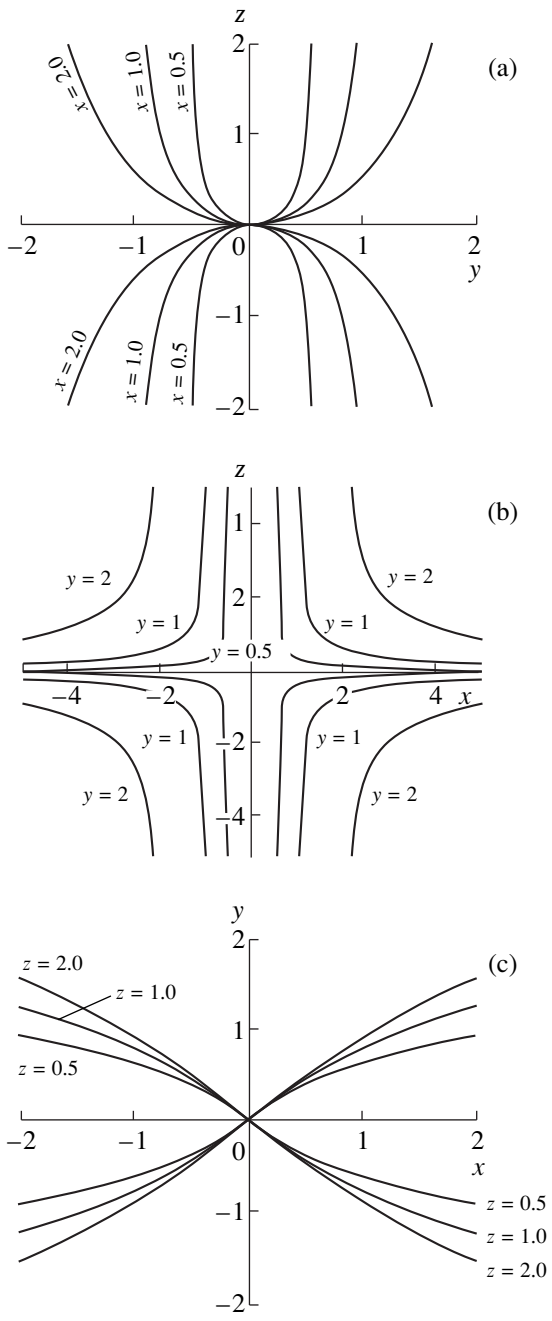


Fig. 2. Projections of the equipotential of field (8) for $\Phi = 1$ onto the planes (a) $x = \text{const}$, (b) $y = 0$, and (c) $z = 0$.

has a set of surfaces

$$\frac{y}{z + \sqrt{x^2 + y^2 + z^2}} = \frac{\pi}{2} + \pi n, \quad n = 1, 2, 3, \ldots \quad (12)$$

as zero equipotentials. These surfaces are parabolic cones that merge into the singular ray $x = y = 0$, $z < 0$. Of practical importance, in this case, may be field fragments between neighboring zero equipotentials. The configuration of the field is very complex and needs further investigation. The same applies to conic potentials $\sin\omega$, $\tan\omega$, etc.

CAUCHY PROBLEM FOR SYMMETRIC FIELDS

Assume that the function F in (1) is nowhere imaginary except in the argument ω. Let its real part be the scalar physical potential of a field symmetric about the y-axis and this field have the x/z-dependent distribution

$$\Phi|_{y=0} = f\left(\frac{x}{z}\right), \quad (13)$$

where f is a given function.

Our aim is to restore the spatial distribution of the potential, specifically, the cone-shaped Laplace potential. Donkin's formula is best suited to solve the problem. Indeed, if it is assumed that the real part of (1) gives the desired potential, then, provided that $y = 0$ in (13), the unknown function F and the given function f are related as

$$F\left(\frac{x}{z + \sqrt{x^2 + z^2}}\right) = f\left(\frac{x}{z}\right). \quad (14)$$

From the equality (at $y = 0$)

$$\omega = \frac{x}{z + \sqrt{x^2 + z^2}} = \frac{x/z}{1 + \sqrt{1 + (x/z)^2}}, \quad (15)$$

we easily find

$$\frac{x}{z} = \frac{2\omega}{1 - \omega^2}. \quad (16)$$

To explicitly determine the function $F(\omega)$, one needs only to substitute (16) into (14) and then replace ω by full complex expression (2). Thus, the complex potential that is the solution to the Cauchy problem for symmetric fields is given by

$$F(\omega) = f\left(\frac{2\omega}{1 - \omega^2}\right). \quad (17)$$

The above algorithm is a direct extension of the Cauchy problem for planar fields [9]. It is very convenient for synthesis of fields with given corpuscular–optical properties in the plane of symmetry, since the function f can be sought using reverse problems in mechanics of particles.

In part II, we will consider in detail electronic and optical properties of cone-shaped potential (4) and its application in designing particular electrostatic devices.

REFERENCES

1. L. A. Baranova, G. N. D'yakova, and S. Ya. Yavor, Zh. Tekh. Fiz. **58**, 207 (1988) [Sov. Phys. Tech. Phys. **33**, 128 (1988)].

2. V. M. Kel'man, S. P. Karetskaya, L. V. Fedulina, and E. M. Yakushev, *Electronic–Optical Elements of Prism Spectrometers of Charged Particles* (Nauka, Alma-Ata, 1979).

3. L. A. Baranova and S. Ya. Yavor, *Electrostatic Electron Lens* (Nauka, Moscow, 1986).

4. S. N. Davydov, S. N. Romanov, Yu. K. Golikov, *et al.*, Proc. SPIE—Int. Soc. Opt. Eng. **3345**, 136 (1998).

5. E. T. Whittaker and J. N. Watson, *A Course of Modern Analysis: An Introduction to the General Theory of Infinite Processes and of Analytic Functions with an Account of the Principal Transcendental Functions* (Cambridge Univ. Press, Cambridge, 1927).

6. N. K. Krasnova, S. N. Davydov, Yu. K. Golikov, *et al.*, J. Electron Spectrosc. Relat. Phenom. **72**, 323 (1995).

7. S. N. Davydov, S. N. Romanov, and N. K. Krasnova, J. Electron Spectrosc. Relat. Phenom., No. 97, 209 (1998).

8. P. G. Gabdullin, S. N. Davydov, and Yu. K. Golikov, Nucl. Instrum. Methods Phys. Res. A, No. 427, 145 (1999).

9. Yu. K. Golikov, K. G. Utkin, and V. V. Cheparukhin, *Design of Components for Electrostatic Electron–Optic Systems* (Leningrad Politekh. Inst., Leningrad, 1984).

Translated by V.A. Isaakyan

Technical Physics, Vol. 45, No. 2, 2000, pp. 236–242. Translated from Zhurnal Tekhnicheskoĭ Fiziki, Vol. 70, No. 2, 2000, pp. 95–101.

Conditions for the Formation of an H^-/D^- Ion Beam in a Volume Plasma Source with a Magnetic Filter

O. L. Veresov, S. V. Grigorenko, and S. Yu. Udovichenko

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

Received October 14, 1998

Abstract—A one-dimensional model for the transport of a plasma with two ion species across the magnetic field in a steady-state volume negative-ion source is proposed. The conditions in the magnetic filter adjacent to the plasma electrode optimum for the generation, formation, and extraction of an H^-/D^- ion beam are found. The theoretical results are in good agreement with the experimental data. © *2000 MAIK "Nauka/Interperiodica".*

INTRODUCTION

Plasma sources with a volume generation of H^-/D^- ions are widely used to produce ion beams with a current up to 10 mA for external injection into cyclotron accelerators [1] and intense neutral beams for injection into nuclear-fusion reactors [2, 3]. In the former case, the steady-state current density attains 10 mA/cm^2. In the latter case, the current density in quasi-steady sources exceeds 50 A/cm^2.

A standard device for the production of ion beams is a two-chamber multicusp plasma source. A non-self-sustaining arc discharge excited in the first chamber between a hot cathode and a cylindrical anode serves as the source of a partially ionized plasma in which the gas molecules are excited by high-energy electrons. In the second chamber, which is an extension of the first, there is a magnetic filter that decreases the density of primary electrons, cools the thermal plasma electrons, and provides conditions for generating negative ions due to dissociative attachment of slow electrons to the excited molecules. The magnetic filter occupies the entire volume of the second chamber; its length should be minimum in order to prevent a substantial decrease in the density of the excited hydrogen molecules as they move toward the emission aperture.

Simulations of the elementary processes in the plasma sources of H^- ions and a comparison between the calculation and experimental results were carried out in [4–8]. The effect of the magnetic-field magnitude in the magnetic filter and the plasma-electrode potential on the emission characteristics of a multipole ion source was investigated. However, the interpretation of the experimental results did not advance beyond theoretical estimates. In order to calculate the density of negative ions as a function of the gas pressure, steady-state plasma parameters in the arc discharge and a numerical model for the transport of the plasma electrons across the magnetic filter detached from the plasma electrode were used. However, the plasma transport across a strong magnetic field in the filters of H^- ion sources is governed by the heavy-ion plasma component rather than the electrons [9]. Note also that the negative-ion balance in the region behind the magnetic filter where the beam is formed differs from the corresponding balance in the region of the magnetic filter adjacent to the plasma electrode.

In this paper, using the time-independent diffusion equations for the plasma components, we analytically found the distributions of the electric field and charged-particle density along the magnetic filter adjacent to the plasma electrode in the multipole source of negative ions. The plasma parameters in the extraction region are optimized to provide the maximum ion current. The theoretical results are compared with the experimental ones.

PLASMA TRANSPORT IN THE MAGNETIC-FILTER CHAMBER

Analysis of the plasma transport across the magnetic field and the calculation of the steady-state parameters in the region of the beam extraction were carried out for a classical multicusp two-chamber source (Fig. 1) used in the systems of external injection in cyclotron accelerators [10].

The discharge chamber is a 7.5-cm-diameter copper cylinder with length L = 10 cm. Outside the chamber, there are $SmCo_5$ magnets arranged in 12 rows, which produce approximately 0.2-cm-wide and 10-cm-long cusps on the cylinder surface. Each row consists of five elements; the polarity of the magnetic field in the row

1063-7842/00/4502-0236$20.00

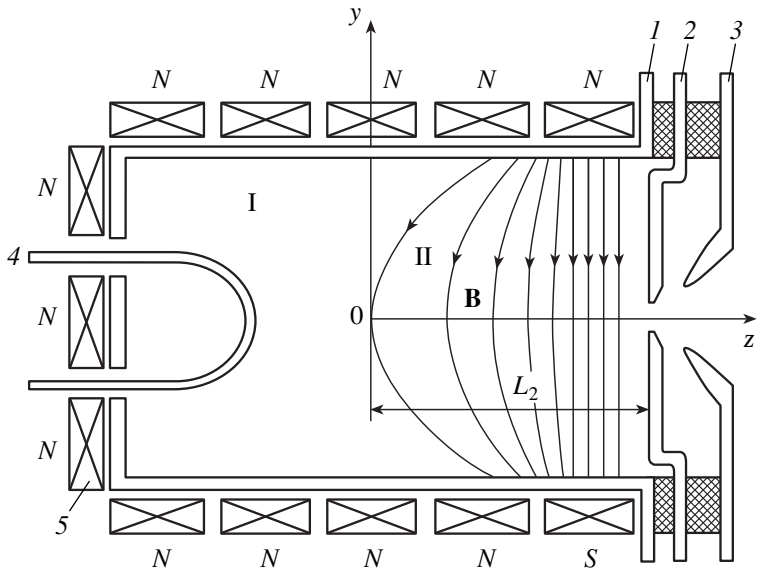


Fig. 1. Schematic of a multicusp ion source with a transverse magnetic filter: (I) plasma generation region, (II) magnetic-filter region, (*1*) anode, (*2*) plasma electrode, (*3*) extracting electrode, (*4*) hot cathode, and (*5*) magnetic element.

closest to the plasma electrode is reversed in order to produce a filter with the magnetic field perpendicular to the cylinder axis. The length of the first chamber with an arc discharge is L_1, and the length of the magnetic filter is L_2.

The steady-state diffusion across the magnetic field in a partially ionized plasma containing positive and negative ions, as well as fast and slow electrons, is described by the following set of equations:

$$-\frac{d}{dz}\left(D_+\frac{dn_+}{dz} - b_+En_+\right) = \langle\sigma v\rangle_{IZ}n_{\mathrm{H}}n_e - \langle\sigma v\rangle_{MN}n_-n_+ - \langle\sigma v\rangle_{RV}n_en_+, \quad (1)$$

$$-\frac{d}{dz}\left(D_-\frac{dn_-}{dz} + b_-En_-\right) = \langle\sigma v\rangle_{DA}n_{\mathrm{H}_2}(v'')n_e - \langle\sigma\nu\rangle_{AD}n_{\mathrm{H}}n_- - \langle\sigma\nu\rangle_{MN}n_-n_+ - \langle\sigma v\rangle_{ED}n_-n_e, \quad (2)$$

$$-\frac{d}{dz}\left(D_e\frac{dn_e}{dz} + b_eEn_e\right) = -\langle\sigma v\rangle_{DA}n_{\mathrm{H}_2}(v'')n_e + \langle\sigma v\rangle_{AD}n_{\mathrm{H}}n_- + \langle\sigma v\rangle_{ED}n_-n_e + \langle\sigma v\rangle_{IZ}n_{\mathrm{H}}n_e - \langle\sigma v\rangle_{RV}n_en_+, \quad (3)$$

where $n_+ = n_- + n_e$ (the plasma quasineutrality condition); $D_j = T_j\nu_{jo}\mu_{jo}/m_j^2\omega_{Bj}^2$; $b_j = D_j/T_j$; n_j, T_j, D_j, and b_j are the density, temperature, diffusivity, and mobility of the charged plasma components, respectively; $j = +, -$, or e; $\mu_{jo} = m_jm_0/(m_j + m_0)$; m_j and m_0 are the masses of charged particles and neutral H_2 molecules; $\omega_{Bj} = eB/m_jc$ is the cyclotron frequency; $\nu_{j0} \approx \sigma_{j0}v_{Tj}n_{\mathrm{H}_2}$; v_{Tj} is the thermal velocity; σ_{j0} is the cross section for elastic collisions between charged and neutral particles; n_{H_2} and $n_{\mathrm{H}_2}(v'')$ are the densities of H_2 molecules in the ground and excited ($v'' = 6$) states, respectively; $\langle\sigma v\rangle_{DA}$ is the rate of dissociative attachment of electrons to the excited hydrogen molecules; $\langle\sigma v\rangle_{AD}$ is the rate of associative electron detachment due to collisions between the negative ions and hydrogen atoms; $\langle\sigma v\rangle_{MN}$ and $\langle\sigma v\rangle_{RV}$ are the ion–ion and electron–ion recombination rates, respectively; $\langle\sigma v\rangle_{IZ}$ is the rate of ionization of H atoms by electrons; and $\langle\sigma v\rangle_{ED}$ is the rate of electron detachment due to collisions between the primary electrons and negative ions.

The set of equations (1)–(3) for a plasma in the magnetic field is valid when b_+ and $b_- \gg b_e$; the ions can be unmagnetized. The degree of plasma ionization is assumed to be low enough so that $m_e\nu_{ei} \ll \mu_{i0}\nu_{i0}$, where ν_{ei} is the electron–ion collision frequency.

We will assume that the density of H_2 molecules is fixed and the atomic hydrogen density and the density of the excited molecules in the vibrational state $H_2(v'')$ are determined by the steady-state parameters of the plasma in the first chamber [2]. The excited molecules are produced in collisions of neutral molecules with the fast electrons arriving from the hot cathode. The estimates show that the loss of the excited molecules on the walls of both chambers occurs more efficiently than their volume loss in the $\langle\sigma v\rangle_{DA}$ and $\langle\sigma v\rangle_{IZ}$ reactions. The density of the excited molecules is determined by the balance between their loss and gain,

$$n_{\mathrm{H}_2}n_{fe}\langle\sigma v\rangle_{EV} = n_{\mathrm{H}_2}(v'')v_{\mathrm{H}_2}/bR, \quad (4)$$

where $\langle\sigma v\rangle_{EV}$ is the rate of production of the excited molecules, v_{H_2} is the average velocity of the $H_2(v'')$ molecules, the fast-electron density $n_{fe} \approx (10^{-1}–10^{-2})n_e$ is taken from experiment [11], R is the radius of the source chamber, and b = 5–10 is the mean number of the molecule–wall collisions required for the molecule deexcitation.

If the density of the ion current to the probe placed in the emission plane of the source is $j_+ \geq 15$ mA/cm^2, then the most abundant positive ions in the plasma are the H^+ ions [2]. For $n_+ \simeq 3n_-$, this corresponds to the current density of the negative ions $j_- \geq 5$ mA/cm^2 (or $I_- \geq 1$ mA for an emission area of 0.2 cm^2). The ion transport coefficients D_j and b_j can be considered fixed, because the heavy-ion temperature varies slowly along the chamber axis (z-axis) and the magnetic field is assumed to be constant in magnitude. The nonuniformity of the magnetic field at the boundary between the chambers is neglected. In the first chamber, the high-energy electrons arriving from the hot cathode are thermalized and the average temperature of thermal electrons is less than 3 eV. In the magnetic filter, the thermal electrons are cooled and their temperature decreases by one order of magnitude. It follows from calculations [12] that the temperature of negative ions remains almost unchanged, $T_- \approx 0.2$–0.3 eV, and the temperature of the positive ion is the same or somewhat lower.

The boundary condition for the H^+ ion density in the magnetic filter is determined by the plasma density n_{+0} in the first chamber, which can be found from the equation of continuity for the electron flux on the anode [4]:

$$\frac{I_d}{eS_{\text{eff}}} = 0.25n_{e0}\left(\frac{8T_e}{\pi m_e}\right)^{1/2} \times \exp\left(\frac{e\Delta\varphi}{T_e}\right) - 0.76n_{+0}\left(\frac{2T_e}{m_i}\right)^{1/2}. \quad (5)$$

Here, I_d is the current of high-energy electrons from the hot cathode, n_{e0} and n_{+0} are the densities of thermal electrons and ions, $n_{e0} \approx n_{+0}$ (the plasma quasineutrality condition), $\Delta\varphi$ is the potential drop across the anode sheath, $S_{\text{eff}} = NLh$ is the effective area of the anode through which the plasma escapes from the source, and N, L, and h are the number, length, and width of cusps, respectively. In the absence of the discharge current through the plasma, the value of the potential drop $\Delta\varphi = (T_e/2e)\ln(0.14m_i/m_e)$ is determined from the balance of the electron and ion fluxes onto the chamber wall. When the current flows through the plasma, $\Delta\varphi$ is somewhat lower than in the absence of the current and the flux of thermal electrons is greater than the ion flux onto the anode. In this case, in order to find $\Delta\varphi$, it is necessary to use the energy balance equation for the primary and secondary charged plasma particles [4].

In the case of a magnetized plasma, we neglect the electron mobility and diffusion across the magnetic field. Then, from (3) and the quasineutrality condition $n_e + n_- = n_+$, we obtain the expression for the steady-state density of H^- ions in the second chamber

$$n_- \approx n_+\left(1 + \frac{\beta}{\gamma}\right)^{-1}, \quad (6)$$

where $\beta = \langle\sigma v\rangle_{AD}n_H$ and $\gamma = \langle\sigma v\rangle_{DA}n_{H_2}(v'')$.

In deducing (6), we assumed that $\gamma \gg \langle\sigma v\rangle_{ED}n_-$, $\langle\sigma v\rangle_{RV}n_+$, $\langle\sigma v\rangle_{IZ}n_H$.

Expression (6) differs from the frequently used expression [2, 5] obtained from the balance of negative ions (2), whose mobility and diffusion are neglected. Such an approach is valid only in the absence of the magnetic field.

The electric-field distribution in the second chamber is found from equations (1) and (2) and condition (6):

$$E(z) = E_0\frac{b_+n_{+0} + b_-n_{-0}}{b_+n_+ + b_-n_-} + \frac{1}{e}\frac{T_+b_+(1 + \beta/\gamma) - T_-b_-}{b_+n_+ + b_-n_-}\frac{dn_-}{dz}, \quad (7)$$

where E_0, n_{-0}, and n_{+0} are the electric field and ion densities at the boundary between two chambers. The first term on the right-hand side of (7) corresponds to the field associated with the external current flowing through the plasma and satisfies the equation $dE(b_+n_+ + b_-n_-)/dz = 0$; the second term corresponds to the ambipolar field in the plasma.

Note that the ambipolar field in the second chamber is proportional to T_+ and T_-, whereas in the first chamber, where the magnetic field is absent, the ambipolar field is determined by the more mobile electrons and is proportional to T_e.

In the presence of a strong electric field $E \gg T_\pm/eL_2$ (where L_2 is the length of the second chamber) associated with the external current, we can neglect the diffusion terms in (1) and (2) and use the drift approximation. Taking into account that $E(z) \approx E_0n_{+0}/n_+(z)$, from (1) or (2) we find the equation for the plasma-ions density

$$\frac{dn_-}{dz} + \frac{\gamma^2}{\beta}\frac{1 + (1 + \beta/\gamma)b_+/b_-}{E_0n_{-0}b_+} \times \left[\langle\sigma v\rangle_{ED} - \langle\sigma v\rangle_{RV}\left(1 + \frac{\beta}{\gamma}\right)\right]^{-1} \times \left[\left(1 + \frac{\beta}{\gamma}\right)\left(\langle\sigma v\rangle_{MN} + \frac{\beta}{\gamma}\langle\sigma v\rangle_{RV}\right)n_-^2\right. \quad (8)$$

$$-\frac{\beta}{\gamma}\langle\sigma v\rangle_{IZ} n_{\mathrm{H}} n_{-}\Big] = 0.$$

In this equation, we can neglect the last term associated with ionization of the gas atoms by slow and fast electrons, because the corresponding rates are relatively low. According to (8), the plasma density and the electric potential in the second chamber are distributed as follows:

$$\begin{aligned} n_{-}z &= n_{-0}\left(1+\frac{\alpha z}{L_2}\right)^{-1}, \\ \varphi(z) &= \varphi_0 + \frac{E_0 L_2}{2\alpha}\left[1-\left(1+\frac{\alpha z}{L_2}\right)^2\right], \\ \alpha &= \left(1+\frac{\beta}{\gamma}\right)\frac{\gamma^2}{\beta}\frac{1+(1+\beta/\gamma)b_+ b_-}{b_+ E_0 L_2^-} \\ &\times \frac{\langle\sigma v\rangle_{MN} + \langle\sigma v\rangle_{RV}\beta/\gamma}{\langle\sigma v\rangle_{ED} - \langle\sigma v\rangle_{RV}(1+\beta/\gamma)}, \end{aligned} \tag{9}$$

where $\alpha > 0$ (because $\langle\sigma v\rangle_{ED} \gg \langle\sigma v\rangle_{RV}$) and φ_0 is the plasma potential at the boundary between two chambers. For $\alpha \ll 1$, the electric field at this boundary is $E_0 = (\varphi_0 - \varphi_e)/L_2$, where φ_e is the potential of the plasma electrode.

ANALYSIS OF THE EXPERIMENTAL RESULTS

To compare the theoretical and experimental results, we will use the following steady-state parameters of the negative-ion source. The schematic of the source is presented in Fig. 1. The effective length of the magnetic filter is $\int B\,dz$ = 100–600 G cm; the gas density is $n_{\mathrm{H}_2} \geq 3.3\times10^{13}$ cm^{-3} ($P \geq 10^{-3}$ torr); and the discharge current and voltage are $I_d \leq 10$ A and $U_d \leq 160$ V, respectively. For the gas density in the source $n_{\mathrm{H}_2} \approx 10^{14}$ cm^{-3}, the magnetic-filter length $L_2 \approx 4$ cm, and the average magnetic field $B \approx 150$ G, the average values of the plasma parameters in the second chamber were close to the corresponding data from [11]: $n_+ \approx 6\times10^{11}$ cm^{-3}, $n_e \approx (2/3)n_+$, $n_- \approx (1/3)n_+$, $\beta/\gamma \approx 2$, $n_{fe} \approx 2\times10^9$ cm^{-3}, and $n_{\mathrm{H}} \approx 10^{13}$ cm^{-3}. All the data for the rates of elementary processes in a gas-discharge plasma are taken from [5, 13].

When the plasma electrode is at the floating potential, the electric field in the magnetic-filter chamber is determined by the second term in expression (7). There exists a critical gas density at which the ambipolar field changes its sign: $T_+ b_+(1+\beta/\gamma) \approx T_- b_-$. The ratio between the ion mobilities in a gas is $b_+/b_- = (m_+ m_-)^{3/2}(\sigma_{+0}/\sigma_{-0})(T_+/T_-)^{1/2} < 1$, where $T_+ < T_-$ and σ_{+0} and σ_{-0} are the cross sections for elastic collisions of ions with the gas molecules. Experimental results [5] show that the atomic hydrogen density increases more rapidly with gas pressure than the density of excited molecules; i.e., the ratio β/γ increases. As the gas pressure increases, the electric field first decreases to zero and then increases again, but with an opposite sign.

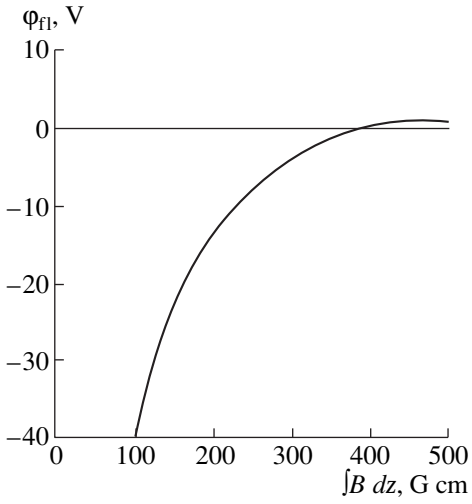


Fig. 2. Floating potential of the plasma electrode as a function of $\int B\,dz$ for I_d = 5 A, U_d = 60 V, and the gas pressure P = 2 mtorr.

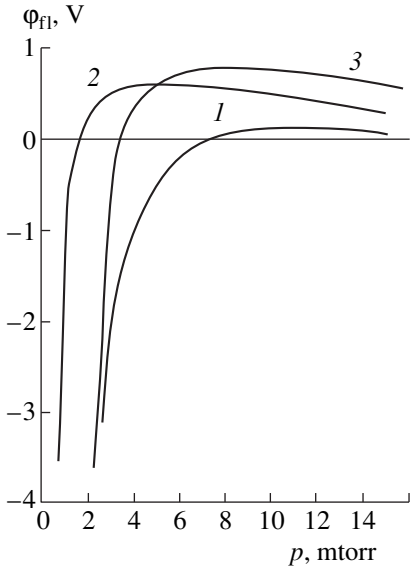


Fig. 3. Floating potential of the plasma electrode as a function of the gas pressure for I_d = 3 A; U_d = 80 V; and $\int B\,dz$ = (*1*) 350, (*2*) 500, and (*3*) 650 G cm.

The experimental dependences of the plasma-electrode floating potential on the effective length of the magnetic filter and on the gas pressure are presented in Figs. 2 and 3, respectively. The first dependence illustrates the well-known effect of the change of the sign of the floating potential when the magnetic field attains

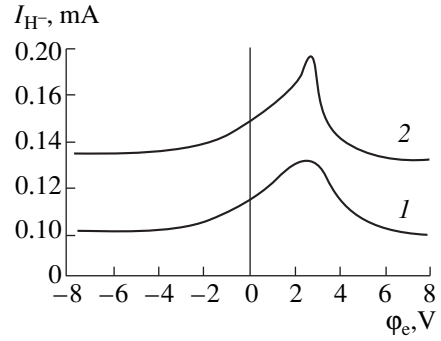


Fig. 4. H^- ion current as a function of the potential bias on the plasma electrode for $I_d = 3$ A, $U_d = 80$ V, $\int B\,dz = 600$ G cm, and P = (*1*) 3 and (*2*) 9 mtorr.

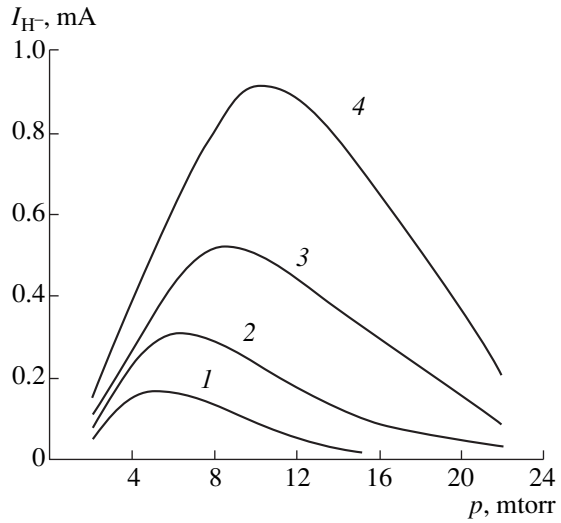


Fig. 5. H^- ion current as a function of the gas pressure for different discharge parameters at $\int B\,dz = 500$ G cm: I_d = (*1*) 3, (*2*) 5.5, (*3*) 8, and (*4*) 10 A; U_d = (*1, 3*) 80, (*2*) 70, and (*4*) 160 V; and φ_e = (*1*) 2.8, (*2*) 2.7, (*3*) 2.4, and (*4*) 1.8 V.

the critical value. As the magnetic field increases and the plasma becomes magnetized, the plasma electrode acquires a positive charge with respect to the plasma bulk. Without a magnetic field, the electrode is charged negatively with respect to the plasma. The second dependence is obtained for the first time and is evidence for the presence of two oppositely charged ion species in the plasma. At low gas pressures, the self-consistent electric field in the plasma is determined by the diffusion of negative ions. As the pressure increases, the diffusion of positive ions becomes dominating.

When a positive (with respect to the anode) bias potential is applied to the plasma electrode, the negative-ion current from the source can be increased compared to the current extracted at the floating positive potential. As is seen from Fig. 4, there is a critical value of the bias potential at which the extracted H^- ion current is maximum. This potential weakly depends on the gas pressure.

The critical value φ_e of the bias potential is found as follows. Let us determine the negative-ion density at the plasma boundary near the emission aperture, taking into account that the current does not flow across this boundary. In the currentless plasma near the emission boundary, the field is described by the second term in expression (7) and is determined by the diffusion of negative ions, because the pressure is reduced in this region of the source. An increase in the potential $\varphi \approx (T_-/e)\ln(n_-)$ + const on the left from the plasma boundary, as well as its increase on the right from the boundary (in the accelerating gap), correctly describes the boundary condition on the steady-state plasma surface $E = -d\varphi/dz = 0$. The potential drop ($\Delta\varphi \approx T_-/e$) in the currentless region near the plasma emitter is substantially lower than the potential of the plasma electrode (several volts). Setting $\varphi = 0$ at the plasma boundary and assuming the distance between this boundary and the equipotential surface in the plasma under the electrode potential φ_e to be much less than the filter length L_2, from (9) we find

$$n_-(L_2) \approx \left(\frac{2e\varphi_e}{T_-}\right)^{1/2} n_{-0}\left(1 - \frac{\alpha_1 L_2}{\varphi_0 - \varphi_e}\right). \qquad (10)$$

When deducing (10), we took into account that a thermal current produced by negative ions penetrates from the boundary deep into the plasma; we also used the condition $\alpha = \alpha_1/E_0 \ll 1$. Let us estimate the value of α for the following plasma parameters: $\langle\sigma v\rangle_{DA} = 6.5 \times 10^{-9}$ cm^3/s, $\langle\sigma v\rangle_{AD} = 10^{-9}$ cm^3/s, $\langle\sigma v\rangle_{ED} = 3 \times 10^{-8}$ cm^3/s, $\langle\sigma v\rangle_{EV} = 3 \times 10^{-10}$ cm^3/s, $\langle\sigma v\rangle_{MN} = 5 \times 10^{-7}$ cm^3/s, $\langle\sigma v\rangle_{Vt} = 8 \times 10^{-11}$ cm^3/s, $n_{H_2}(v'') \approx 3 \times 10^{11}$ cm^{-3}, $\nu_{+0} = (32\sigma_{+0}/3\pi^{1/2})(T_+/m_+)^{1/2} n_{H_2} \approx 8.4 \times 10^5$ s^{-1}, $\sigma_{+0}(H^+ + H_2) = 1.6 \times 10^{-15}$ cm^2, $b_+ \approx 8 \times 10^7$ cm^2/(s V), $\varphi_0 \approx 4$ V, $\varphi_e \approx 2.5$ V, and the magnetic field $B = 125$ G. The value φ_0 is determined by the potential drop $\Delta\varphi$ used in (5), assuming that $\varphi = 0$ on the anode. Under these conditions, $\alpha \approx 0.3$.

By equating the derivative of the current density of negative ions $f_-(\varphi_e)$ with respect to the potential φ_e at the plasma-emitter boundary to zero, $[dj_-(\varphi_e)/d\varphi_e]|_{z = L_2} = 0$, we obtain the critical value of the bias potential $(\varphi_e)_{cr}$. From (10), taking into account that the current extracted from the emitter is the thermal H^- ion current, we find

$$\frac{[\varphi_0 - (\varphi_e)_{cr}]^2}{(\varphi_e)_{cr}} = \left(1 + \frac{\beta}{\gamma}\right)\frac{2\gamma^2 L_2^2}{\beta b_+}\frac{\langle\sigma v\rangle_{MN}}{\langle\sigma v\rangle_{EV}}. \qquad (11)$$

From (11), it follows that, for $\varphi_0 \approx 4$ V, the plasma-electrode potential is $(\varphi_e)_{cr} \approx 2.5$ V, which is in good agreement with the experimental value (Fig. 4). If

$\beta/\gamma \gg 1$, then $(\varphi_e)_{cr}$ is independent of the gas pressure, because $\gamma \sim n_{H_2}(v'')$ and $b_+ \sim n_{H_2}$ and the density of the excited molecules is a linear function of n_{H_2} for the gas pressure up to P = 9 mtorr [6]. As the ratio β/γ decreases, a weak dependence of $(\varphi_e)_{cr}$ on the gas pressure shows up.

The H^- ion density as a function of the pressure in the discharge chamber with the magnetic filter for $\int B\,dz$ = 500 G cm and different discharge currents is presented in Fig. 5. The ion beam was extracted from a 5-mm-diameter emission aperture in the plasma electrode. It is seen that the optimum gas pressure at which the ion-beam current I_{H^-} is maximum depends on the discharge current. The plasma source emits the thermal ion current (saturation current); therefore, the dependence (10) of the density $n_-(L_2)$ on the gas pressure reflects the current behavior in Fig. 5. The negative-ion density at the plasma-emitter boundary and the plasma density in the first chamber are related by (6); the plasma density in the first chamber is proportional to the density of the primary fast electrons n_{ef} [i.e., it is proportional to the discharge current (5)]. According to (10), the H^- ion density increases with increasing gas pressure; at a certain pressure, the H^- ion density saturates. For an excessive gas supply into the source, the deexcitation of $H_2(v'')$ molecules in collisions with neutral molecules, which was neglected in (6) and (10), becomes important. In addition, the H^- ion beam is attenuated because of the loss of negative ions due to stripping by the H_2 molecules in the extraction region. Different rates of these processes at different gas pressures determine the different inclination angles of the dropping regions in curves *1–4* in Fig. 5.

CONCLUSION

We have proposed a model describing the steady-state density distribution of negative ions in the region of formation and extraction of an ion beam from the multipole two-chamber source with the magnetic filter adjacent to the plasma electrode. The model is based on the equations for the transport of components of a plasma with two ion species across the magnetic field.

The use of a one-dimensional model to describe the plasma transport in the magnetic filter is justified because the plasma-electrode diameter is substantially greater than the filter length, or more exactly, the length of the region where the magnetic field lines are closed on the anode. The magnetic field has no effect on the motion of the charged particles along **B**. The ambipolar plasma motion toward the anode along the magnetic field lines is determined by the light plasma component (electrons). The potential drop between the central plasma region and the anode falls at a narrow sheath near the chamber wall, whereas, in the plasma volume, the electric field is close to zero.

The distributions of the electric field and the densities of the plasma components found by using the model proposed allowed us to determine the optimum conditions for the formation of an H^- ion beam. The results of calculations are in good agreement with the experimental data.

It is shown that, as the gas pressure in the source increases, the floating potential of the plasma electrode changes its sign from negative to positive, which is accompanied by an increase in the negative-ion current extracted from the source. If there is no need to obtain the extreme value of the H^- ion current, it is not necessary to apply a positive potential bias to the plasma electrode with respect to the anode. In sources in which the magnetic filter is detached from the plasma electrode, it is required to apply a positive potential to the plasma electrode in order to compensate for the negative potential drop in plasma near the electrode.

We have found the expression for the plasma-electrode positive potential bias at which the extracted current of negative ions is maximum. The value of this bias is a function of the magnetic-filter length, magnitude of the magnetic field, plasma density, and the discharge current in the first chamber and depends only slightly on the gas pressure in the plasma source.

We have determined the steady-state profile of the negative-ion density along the magnetic-filter chamber. By varying this profile, it is possible to obtain the maximum negative-ion current. The density of negative ions is determined from the equations for the charged-particle balance with taking into account the plasma quasineutrality and assuming that only heavy particles are transported across the magnetic field. Note that, if the diffusion and mobility of negative ions in the magnetic filter detached from the plasma electrode are not incorporated into the gain–loss balance for the particles [2, 5], then the calculations give a lower value of the H^- ion density as compared to the experimental data.

REFERENCES

1. G. Dutto, K. Jayavanna, T. Kuo, *et al.*, Rev. Sci. Instrum. **65**, 1189 (1994).
2. A. J. T. Holmes, L. M. Lea, A. F. Newman, *et. al.*, Rev. Sci. Instrum. **58**, 223 (1987).
3. P. A. Pincosy, K. W. Ehlers, and A. F. Lietzke, Rev. Sci. Instrum. **57**, 2387 (1986).
4. A. J. T. Holmes, in *Proceedings of International Symposium on the Production and Neutralization of Negative Ions and Beams, Brookhaven, NY, 1992*, p. 101.
5. P. Berlemont, D. A. Skinner, and M. Bacal, in *Proceedings of International Symposium on the Production and Neutralization of Negative Ions and Beams, Brookhaven, NY, 1992*, p. 76.

6. D. A. Skinner, P. Berlemont, and M. Bacal, in *Proceedings of International Symposium on the Production and Neutralization of Negative Ions and Beams, Brookhaven, NY, 1990*, p. 557.
7. M. Ogasawara, T. Yamakawa, F. Sato, *et al.*, in *Proceedings of International Symposium on the Production and Neutralization of Negative Ions and Beams, Brookhaven, NY, 1990*, p. 596.
8. K. N. Leung, C. A. Hauck, W. B. Kunkel, *et al.*, Rev. Sci. Instrum. **60**, 531 (1989).
9. V. A. Rozhanskiĭ and L. D. Tsendin, *Collisional Transport in a Partially Ionized Plasma* (Énergoizdat, Moscow, 1988).
10. Yu. V. Afanas'ev, M. F. Vorogushin, S.V. Grigorenko, *et al.*, in *Proceedings of Eighth Conference on the Industrial and Medical Applications of the Charged-Particle Accelerators, St. Petersburg, Russia, 1995*, p. 150.
11. K. Jayananna, M. McDonald, D. H. Yuan, *et al.*, in *Proceedings of European Particle Accelerator Conference, Nice, France, 1990*, p. 647.
12. J. M. Wadehra, Phys. Rev. A **29**, 106 (1984).
13. O. Fukumasa, J. Phys. D **22**, 1668 (1989).

Translated by A.D. Smirnova

Technical Physics, Vol. 45, No. 2, 2000, pp. 243–246. Translated from Zhurnal Tekhnicheskoĭ Fiziki, Vol. 70, No. 2, 2000, pp. 102–105.

Radiation-Assisted Formation of Microprobes for Scanning Tunnel Microscopes

T. I. Mazilova

Kharkov Institute of Physics and Technology, Ukrainian Scientific Center, Akademicheskaya ul. 1, Kharkov, 310108 Ukraine

Received April 21, 1998; in final form, November 11, 1998

Abstract—The formation of tipped tungsten microprobes by ion bombardment in high electric fields with subsequent low-temperature field evaporation is reported. During irradiation, the current density was shown to rise nonmonotonically because of heavy particles present in the bombarding beam due to emitter erosion. It was found that the initially hemispheric working part of the probes turns into the axially symmetric complex-shaped surface. A correlation of these effects with inert gas ionization and tungsten sputtering under the action of super-high-density electron beams is discussed. The atomically smooth microprobes obtained by ion bombardment and field evaporation offer high stability and an atomic-level resolution. *© 2000 MAIK "Nauka/Interperiodica".*

INTRODUCTION

The microprobes of scanning tunnel microscopes (STMs), atomic force microscopes (AFMs), and other modifications of scanning probe instruments are responsible for their resolution, as well as measurement accuracy and reliability [1, 2]. Microtips ion- or electrochemically etched on the top of the probes are of small radius of curvature and provide high, up to an atomic level, probing locality [1, 3]. However, the position of the tips cannot usually be strictly controlled, which makes it difficult to unambiguously treat obtained images and reduces the parameter reproducibility of the nanodevices. In particular, the special requirements are placed on the geometry of microprobes that are used in STM-based devices operating in ultra-high electric fields or under conditions of tight atomic-force scanning [4]. In such regimes, the microprobes exert a severe and irreversible effect on the surface being processed. These STM-based instruments have proven to be effective tools for local surface modification and are promising for fabricating ultrasmall electronic components and high-density storage media [1, 5]. However, the axial asymmetry of the working part of the probes causes nonuniform probe–surface contact, resulting in uncontrollable differences in the pattern on gently and steeply inclined surface regions. Therefore, atomically smooth axially symmetric microprobes are imperative for adequate interpretation of the images and reproducible local modification of the surface.

In this work, we report the fabrication of the tipped tungsten microprobes by ion bombardment in high electric fields.

EXPERIMENTAL PROCEDURE

Tipped tungsten probes with an initial radius of curvature ranging from 10 to 100 nm and a taper angle of 2–10° were prepared by electrochemically etching axially textured wires [10]. The samples cooled to 21–80 K were treated by a field-emission microscope operating in the electron and ion regimes. Helium at a pressure between 10^{-2} and 10^{-3} Pa was used as an image gas. In the autoelectron regime, helium was pumped off to a 10^{-8}–10^{-7} Pa level and the chamber was filled by neon to a pressure of 10^{-1}–10^{-4} Pa. The negatively biased tips were irradiated by neon ions that were generated by the autoelectronic current. For autoelectronic current densities of less than 10^6 A/cm^2, the ion bombardment rate was determined from relationships given in [6]. If the ion beam originated at distances $r < 10r_0$ from the tip, where r_0 is the tip radius, the tip was bombarded by ions from the conic part of the beam; for $r > 10r_0$, the ions impinged from the cylindrical part coaxial with the sample had a greater effect. In our work, the field strength E in the autoelectron regime was $(4–5) \times 10^7$ V/cm. For $r_0 < 100$ nm, the ions produced at $r < 10r_0$ had an energy below the tungsten sputtering threshold and the surface was modified primarily by those generated in the remote, cylindrical part of the beam. In this part, the ion energies obey the Maxwell distribution [7] with the mean energy eEr_0, where e is the charge of an electron. The mean energy of the ion beam incident on the surface lay within the 50–500 eV range. It was thus demonstrated that directional ion fluxes with rates between 10^{13} and 10^{16} cm^2 s^{-1} can be obtained by passing the autoelectronic current in neon at pressures 10^{-1}–10^{-3} Pa. The integrated neon fluxes were 10^{16}–5×10^{17} ion/cm^2. During the bom-

1063-7842/00/4502-0243$20.00 © 2000 MAIK "Nauka/Interperiodica"

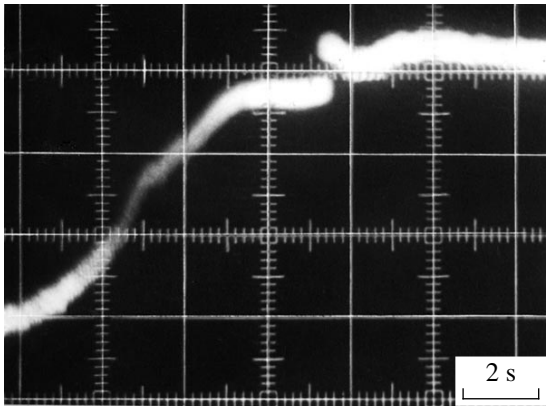


Fig. 1. Oscillogram of the emission current from the tungsten probe subjected to neon ion bombardment.

bardment of the surface by the neon ions, emission current oscillograms were taken and autoelectron images were video recorded. Measurements were performed at a constant voltage of 1–5 kV kept with an accuracy of 0.25%.

RESULTS AND DISCUSSION

Figure 1 shows a typical current oscillogram (voltage 800 V, neon pressure 2×10^{-3} Pa) for the tipped probe with a radius of 26 nm. As follows from the probe topographs obtained in the autoion regime, the rise in the current is associated with emitter sharpening, controlled by the radiation-enhanced mobility of the surface atoms. The current rises first slowly and then rapidly, which is accompanied by qualitative changes in the autoelectron images (Fig. 2). The random arrangement of bright emitting centers, which is typically observed for current densities below 10^8 A/cm^2 (Fig. 2a), is changed by {100} faceting (Fig. 2b) at $J = 2 \times 10^6$ A/cm^2, and then, at $J = 10^7$ A/cm^2, the entire surface is faceted by low-index planes (Fig. 2c). In the last case, new atomic layers are nucleated and grow, which is accompanied by synchronous current oscillations. As the radius of curvature of the tips decreases, the relative amplitude of the oscillations increases, reaching 10% for tips with a radius of curvature of 10 nm.

A typical $\frac{1}{J}\frac{dJ}{dt}$ vs. electron current density J dependence is plotted in Fig. 3 for a neon pressure of 4×10^{-3} Pa. This curve was constructed from the time variation of the current for the tip with an initial radius of curvature of 60 nm to which a voltage of 1800 V was applied. The relative rate of change of the current goes over the peak. Its amplitude grows with neon pressure, while its position remains nearly the same (between 5×10^6 and 10^7 A/cm^2) from sample to sample. The presence of the peak indicates that the energy that is imparted to the surface by the incident ions and is responsible for the radiation-assisted sharpening of the emitter nonmonotonically depends on the current density of autoelectrons that ionize atoms in the interelectrode gap. The run of the $\frac{1}{J}\frac{dJ}{dt}$ vs. J curve can be explained if it is assumed that heavy particles sputtered from the emitter are present in the bombarding ion beam. Then, as the electron current density rises, so does the ionization probability of tungsten atoms knocked out of the emitter surface and, in proportion, the energy released on the surface. Above some density of the electron current at which most of the knocked-out atoms have been ionized, sputtered atoms ionize more and more closely to the surface and, accordingly, heat release at the field-emission cathode drops. Within the concept of ion bombardment of cold emitters [6, 7], it can be shown that tungsten ionization becomes significant at electron current densities on the order of 10^7 A/cm^2. At higher densities, the majority of atoms knocked out of the tip surface is ionized by autoelectrons and bombards the tip. A further increase in the current density causes tungsten to ionize in the immediate vicinity of the surface; hence, the mean energy of the bombarding ions drops. Thus, the peak in the $\frac{1}{J}\frac{dJ}{dt}$ vs. J curve at some critical value of electron current density may indicate specific features of the

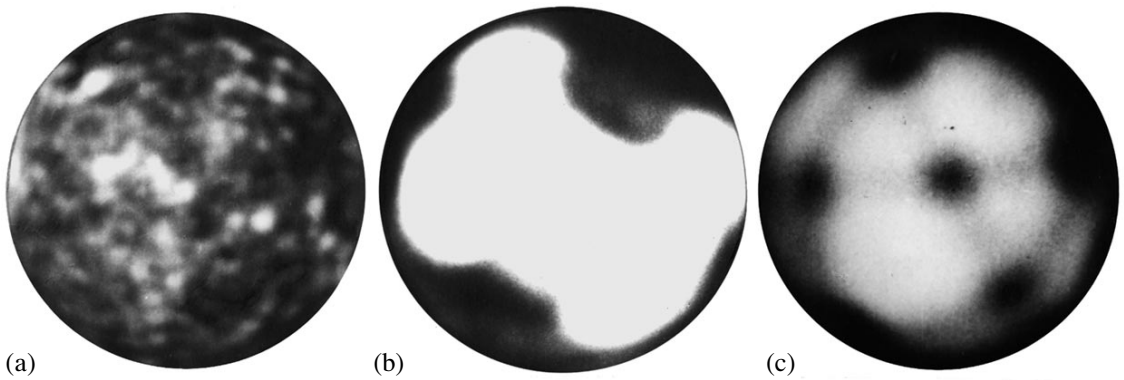


Fig. 2. Evolution of the field-emission images during the formation of the tips. J = (a) 8×10^5, (b) 2×10^6, and (c) 10^7 A/cm^2.

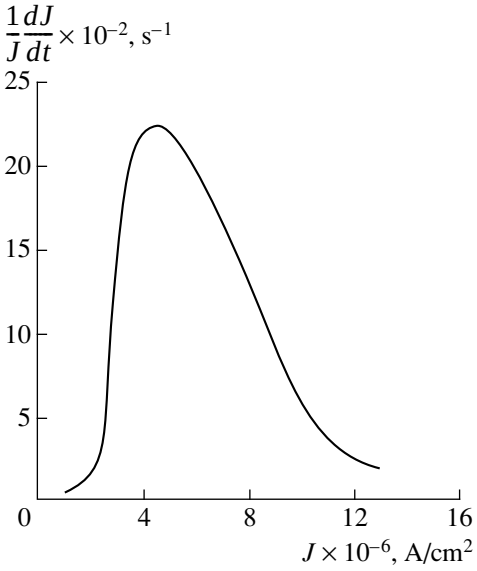


Fig. 3. Rate of change of the emission current vs. current density during ion bombardment.

processes of inert gas ionization and tungsten sputtering in super-high-density electron beams that are formed near the tipped microprobes in the autoelectron regime.

Ion microscopic studies revealed that radiation-assisted sharpening of the microprobes changes also the shape of their tips. The surface topography evolution observed by means of layer-by-layer evaporation of the samples (Fig. 4) showed that, during bombardment, the working (semispherical) part of the probes transforms into the axially symmetric surface that can be represented as a paraboloid adjoint to a hemisphere of radius 4–6 nm, which is much less than the principal radius of curvature at the paraboloid vertex. The center of the hemisphere formed by ion bombardment was always located on a ⟨110⟩ crystallographic axis aligned with the geometric axis of the probe. Thus, forming due to ion bombardment improves the probing locality by one order of magnitude. The use of controllable layer-by-layer field evaporation [8] allowed us to obtain, with a high level of reproducibility, the atomically smooth surface of the semispherical tips with a given radius of curvature within the resolution of the field emission microscope (0.27 nm). Such microprobes are of paramount importance when nanodevices are made under the tight scanning regime.

The atomic-scale resolution of STMs is usually provided by passing the tunnel current through microprotrusions on the surface of atomic-size probes [1]. It is believed that individual atoms or ultrasmall clusters on the top of the microprotrusions are responsible for emission. In this context, the reproducibility of the STM atomic resolution upon using the microprobes with the atomically smooth tips seems to be somewhat surprising. By way of example, Fig. 5 demonstrates the STM image taken of the graphite test object using an

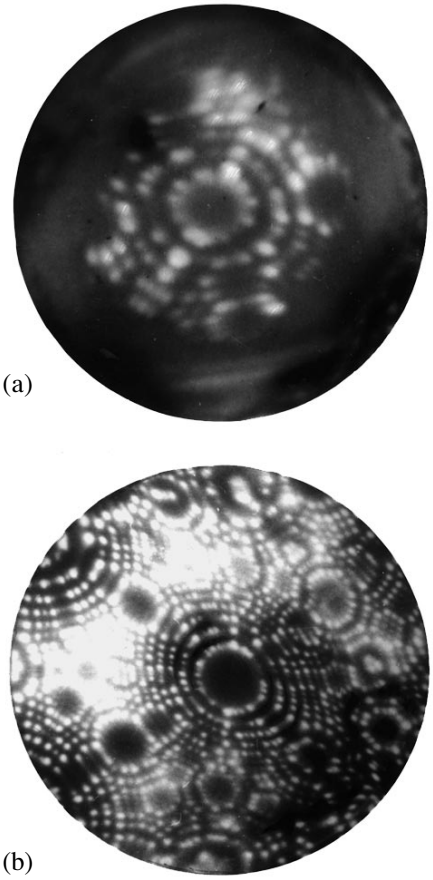


Fig. 4. Ion micrographs of the probe working part (a) immediately after radiation-assisted forming and (b) after the removal of the semispherical protrusion.

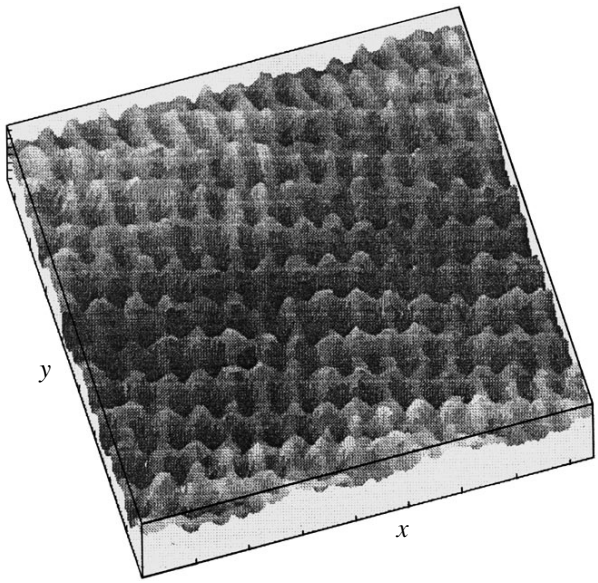


Fig. 5. Graphite surface image obtained in the STM with an atomically smooth microprobe.

atomically smooth microprobe. Such probes, unlike those with irregularly shaped microprotrusions obtained by electrochemical etching, featured high stability and provided atomic resolution during many-hour continuous operation.

ACKNOWLEDGMENTS

The author is indebted to V. Gurin, I. Mikhaĭlovskiĭ, V. Solonovich, and V. Shulaev for valuable discussion.

REFERENCES

1. S. N. Magonov and M.-H. Whangbo, *Surface Analysis with STM and AFM* (Springer, Berlin, 1996).
2. V. A. Ksenofontov, I. M. Mikhaĭlovskiĭ, V. M. Shulaev, *et al.*, *Physics, Chemistry, and Application of Nanostructures*, Ed. by V. E. Borisenko *et al.* (Belarussian State Univ. of Informatics and Radioelectronics, Minsk, 1995), pp. 244–245.
3. L. Libioulle, Rev. Sci. Instrum. **66**, 97 (1995).
4. H. Rohrer, Jpn. J. Appl. Phys. **32**, 1335 (1993).
5. S. Kondo, S. Heike, M. Lutwyche, *et al.*, Appl. Phys. **78**, 155 (1995).
6. P. A. Bereznyak and V. V. Slezov, Radiotekh. Électron. **17**, 354 (1972).
7. P. A. Bereznyak, O. A. Velikodnaya, V. I. Gerasimenko, *et al.*, Vopr. At. Nauki Tekh., Ser. Fiz. Radiatsionnykh Povrezhdeniĭ Radiatsionnoe Materialovedenie, No. 1, 41 (1994).
8. M. K. Miller and G. D. W. Smith, *Atom Probe Microanalysis: Principles and Applications to Materials Problems* (Oak Ridge National Lab., 1987; Mir, Moscow, 1993).

Translated by V.A. Isaakyan

Technical Physics, Vol. 45, No. 2, 2000, pp. 247–250. Translated from Zhurnal Tekhnicheskoĭ Fiziki, Vol. 70, No. 2, 2000, pp. 106–109.

Scattering of Plasma-Jet Ions by a Tantalum Target

V. I. Batkin

Institute of Chemical Kinetics and Combustion, Russian Academy of Sciences, Siberian Division, Novosibirsk, 630090 Russia

Received August 11, 1999

Abstract—The possibility of recording a chemical reaction occurring on the surface of a metal target in a collisionless plasma jet by analyzing the longitudinal-velocity spectra of jet-plasma ions scattered by the target is shown. The target is a Ta strip with a 3×0.1 mm cross section. The ion energy is 30 eV. It is found that, after rapid heating of the target to 1000 K, the current produced by scattered ions decreases and is then restored in a time of 40 s for nitrogen ions and 60 s for hydrogen ions. The relaxation of the current correlates with the accumulation of an intermediate reagent (probably, TaO_2Na^+ ions) on the target surface; the reagent participates further in the formation of experimentally observed ions with a mass of 245 ± 1 (probably, $TaO_2Na^+ \cdot H_2O$).

Low-energy ion beams are used in film-deposition technology either directly [1] or as auxiliary beams (in systems employing higher energy ion flows) [2]. For ion energies below 100 eV, the beam-current density is limited by the space charge in the region of the ion extractor and is less than 1 mA/cm^2. The plasma jet produced by an arc plasma source [3] is similar in its properties to a compensated ion beam [4], but has a substantially higher density. After the expander was improved, the velocity of the jet-plasma ions could be varied within the range from 2 to 7.6 eV$^{0.5}$ with a velocity spread of 1 eV$^{0.5}$ [5]. With such a plasma jet, it is possible not only to act selectively upon the surface, but also to monitor its state by analyzing the ion scattering. Here, we analyze the longitudinal-velocity and mass spectra of the ions scattered by the target. The relaxation of a tantalum target after heating is studied.

Figure 1 shows the schematic of the device [5]. An arc plasma source *1* with an electrode *2* for controlling the ion velocity ejects a plasma jet falling onto the target *3* supplied with a drive mechanism for rotation and translation. A flat grid diode *4* and a collimator *5* collect ions for their analysis. The ion velocity component parallel to the diode does not change. This results in the displacement of the ions, which is recorded by profilometers *6* and *7*. The ion-beam profiles reflect the longitudinal-velocity distribution of the jet-plasma ions. Electrodes *5* and *8*, creating the transverse electric field that compensates the displacement; a collimator *9*; and a mass-spectrometer *10* make it possible to analyze the mass composition and velocities of the ions simultaneously. This technique is described in more detail in [4–6]. The target was a strip made of tantalum of the T brand; it had a 3×0.1-mm cross section and was located 25 mm from the analyzer entrance slit *5*.

A pulsed plasma jet was produced during 100 μs, and the interval between pulses was no less than 10 s. In step with the plasma generation, we performed the gas puffing of ~10^{17} molecules, so that, in a time of 1 ms, the pressure in the vicinity of the target increased from an initial level of ~10^{-4} Pa to ~10^{-1} Pa. The pumping of the vacuum chamber was performed by three NMDO-01-1 (Nord-100) magnetic-discharge pumps. The target was heated by a 4-s current pulse to a temperature of 300–1200 K. The current pulse was applied before the generation of the plasma jet. The characteristic target-cooling time was 10 s. To study the relaxation of the target after the heating pulse, a sequence of 30 plasma-jet pulses was produced. Experiments were carried out with hydrogen and nitrogen plasma jets. The ion kinetic energy was 30 eV for hydrogen ions and about 32 eV for nitrogen ions. According to [4], the plasma potential φ_0 with respect to the plasma-source anode and the ion kinetic energy are related by the equality $\varphi \approx -mv^2/2e$.

Figure 2 shows a typical longitudinal-velocity spectrum of the ions. The right peak with a maximum at 5.5 eV$^{0.5}$ and a FWHM of 1.1 eV$^{0.5}$ corresponds to the ions incident on the target. The left, 1-eV$^{0.5}$-wide peak, whose position varies from –1.84 eV$^{0.5}$ (when the target potential with respect to the plasma-source anode is $U_t > -16$ V) to –1.13 eV$^{0.5}$ (when the potential is $U_t = -30$ V), corresponds to the ions arriving from the target. For $U_t < -30$ V, these ions were absent. The ion flux from the target had two components: the ion emission stimulated by the electron bombardment, which increased as the target potential increased, and the jet-plasma ions scattered by the target. The current produced by the latter ions was maximum at $U_t = -mv_\parallel^2/2e$, when the ion-velocity component $v_\perp$ perpendicular to the target surface vanished under the action of the electric field and the ions did not reach the surface. The ratio between the numbers of the emitted

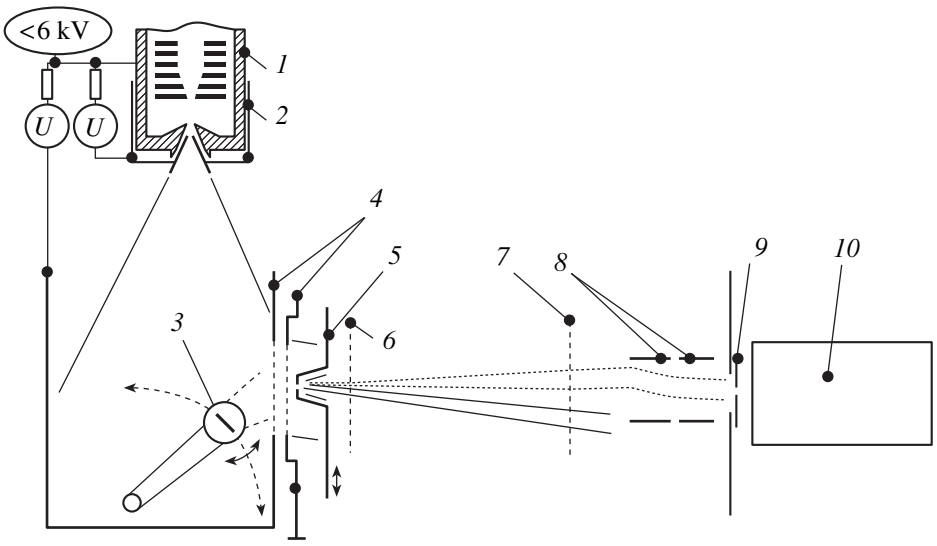


Fig. 1. Schematic of the experimental device.

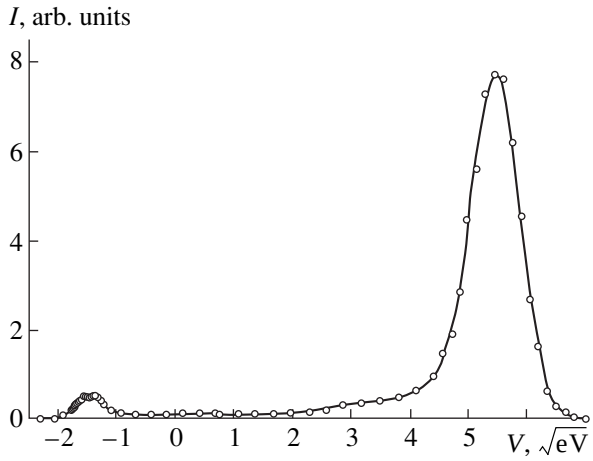


Fig. 2. Ion longitudinal-velocity spectrum; the right peak corresponds to the ions incident on the target, and the left peak corresponds to the ions arriving from the target.

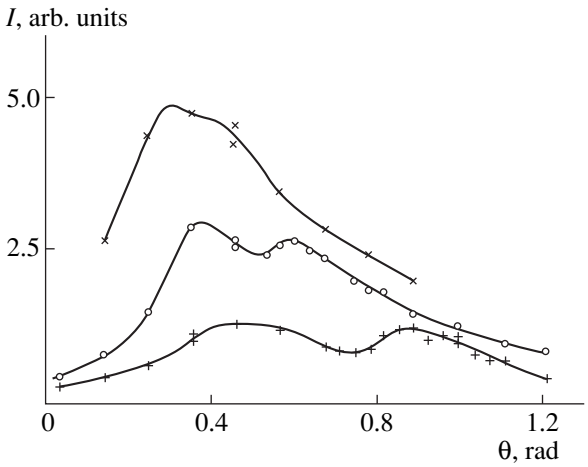


Fig. 3. Dependence of the flux of H_2^+ ions arriving from the target on the angle θ for the target potential U_t = (×) 0, (○) – 8, and (+) –19 V.

and scattered ions could be varied by varying the angle θ between the plasma jet and the normal to the target surface.

Figure 3 shows the flux of H_2^+ ions arriving from the target as a function of the angle θ for different values of the target potential U_t. The experiment was conducted without target heating. For the negative potential and small angles, the ions reach the target and are scattered by its surface. For large angles, the ions are reflected by the electric field and do not reach the target. If we assume that the electric field is localized in a thin layer, then, for U_t = –8 V, the transition from scattering to reflection must occur at the critical angle $\theta_c \approx 0.55$ rad and, for U_t = –19 V, we have $\theta_c \approx 0.9$ rad. A small value of the current jump near θ_c means that a significant fraction of ions is scattered by the surface without charge exchange. The situation is different from that observed for higher ion energies [7]. For H^+ and H_3^+ ions, the current jump during the transition from reflection to scattering is also small.

The current of the emitted ions is independent of the target temperature and remains unchanged upon heating. The number of ions reflected from the target decreases upon heating and then returns to its initial value in a time of $\tau \approx 60$ s. The relaxation time τ is substantially longer than both the target-cooling time and the time (on the order of 1 s) associated with the adsorption of a monolayer of residual-gas molecules. The absence of the relaxation processes in the flux of ions emitted by the target also disagrees with the possi-

ble explanation of the relaxation of the reflected-ion current by the gas adsorption. At a relatively low vacuum, the gas adsorption does not manifest itself in the experiments with a tantalum target, because Ta easily dissolves oxygen and nitrogen [8], thus inhibiting the formation of a dense adsorbate layer.

Figure 4 shows the relative decrease $\Delta I/I$ in the current of H^+ ions arriving from the target upon its heating as a function of the target potential U_t for θ = (*1*) 26.5 and (*2*) 39°. The target temperature T = 1000 K attained the values at which the dependence $\Delta I(T)$ saturates. Ions with longitudinal velocities close to the central velocity of the ion velocity spectrum were recorded. The effect was absent for positive U_t, when the jet-plasma ions did not reach the target. As U_t decreased, the relative decrease in the ion current attained almost 100% when the ions reached the target. The monotonic increase in the effect of the target heating on the number of ions scattered by the target when the ion velocity component is normal to the target surface is likely related to the decrease in the elastic scattering of ions by the clean tantalum surface. The mass spectrum of ions scattered by the target differs from the incident-ion spectrum: the H_3^+ content in the former spectrum is lower by a factor of 2–3. For the target potentials from –16 to –12 V, the dependence of $\Delta I/I$ on U_t is nonmonotonic and is different for different incident-ion species. For H^+ ions, this function has a well-reproducible peak, whose position is independent of the target orientation. This peak is characteristic of the elastically scattered ions reaching the analyzer with a high longitudinal velocity. Since the position of this peak is independent of the angle of incidence of ions, it is likely produced due to a single elastic scattering by local scattering sites. A relatively high energy corresponding to the cross section maximum indicates the presence of a substantial space potential in the vicinity of these sites.

Upon the target heating, the current produced by the ions arriving from the target relaxes by nearly the exponential law. This is seen from Fig. 5, in which the data on the hydrogen-ion current are fitted by the function $A - B\exp(-t/\tau)$. The measurements were carried out at U_t = –19 V upon heating the target to 1000 K. The relaxation times for H^+, H_2^+, H_3^+ ions are almost the same and equal 58 ± 3, 59 ± 5, and 63 ± 4 s, respectively. It was found that the irradiation and vacuum conditions had little or no effect on the time τ. Relaxation upon heating depends on the target material only. In the experiments with a nitrogen plasma jet, the relaxation time was τ = 40 ± 2 s for N^+ and τ = 40 ± 8 s for N_2^+. The longer time τ in the experiments with a hydrogen plasma jet can be attributed to the fact that hydrogen penetrates into the surface layer of the metal and slows down the diffusion processes there.

In addition to the scattering of jet-plasma ions, we studied the emission of ions from the heated target in

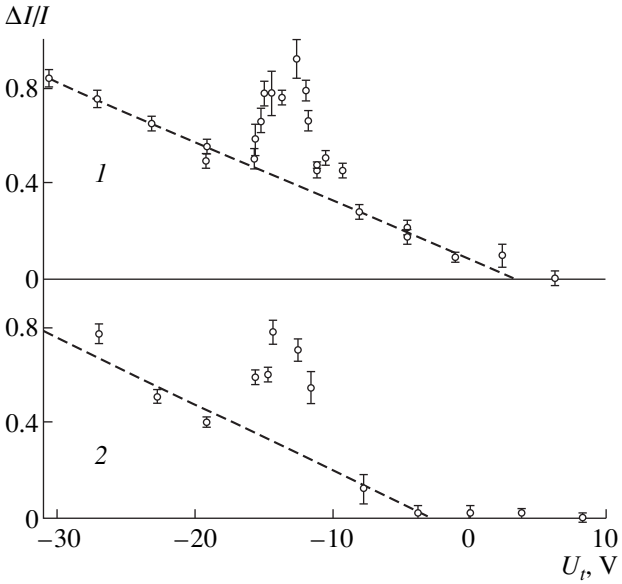


Fig. 4. Relative decrease $\Delta I/I$ in the current of H^+ ions arriving from the target upon its heating as a function of the target potential U_t for θ = (*1*) 26.5 and (*2*) 39°.

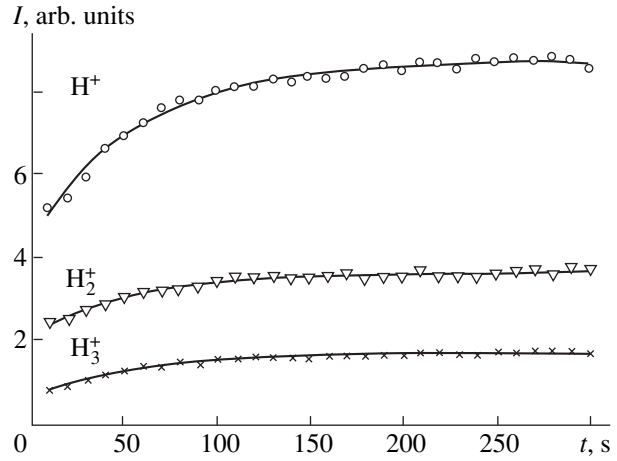


Fig. 5. Relaxation of the current of ions scattered by the target upon its heating.

the absence of plasma. At temperatures T > 950 K, the emission of K^+ and Na^+ ions with currents higher than 10^{-11} A was observed. At temperatures above 650 K, for a time of 0.25 s, the emission of the ions with a mass of 245 ± 1 (probably, $TaO_2Na^+ \cdot H_2O$ ions) with a small addition of the heavier ions (including $TaO_2K^+ \cdot H_2O$ ions) was observed. We studied the dependence of the ion emission on the time interval between two successive heatings. It was observed that the current produced by K^+ and Na^+ ions was independent of t. The heavy-ion current was absent for small t and restored by the law $\sim t^2$ as the time interval increased, which is typical of two-stage chemical reactions. Presumably, in the first stage of the reaction, the TaO_2Na^+ ions are formed

from the TaO_2 oxide and, in the second stage, the complex $TaO_2Na^+ \cdot H_2O$ ions are produced. The relaxation time of the concentration of the product of the first reaction stage (probably, TaO_2Na^+ ions) is about $\tau_e = 40 \pm 5$ s when the gas is not puffed into the vacuum chamber. The periodic puffing of hydrogen through the valve of the ion source with a 10-s interval increases the time τ_e to 60 ± 8 s. The puffing of nitrogen does not affect the relaxation time. The second stage lasts nearly 300 s. The participation of water molecules in this stage is confirmed indirectly: the admission of O_2 or H_2 (at a pressure of 0.1 Pa for 1 s) to the heated target increases the heavy-ion emission observed upon the next heating.

The relaxation time τ_e of the ion emission from the tantalum target is close to the relaxation time τ of the current of the jet-plasma ions scattered by the target and depends in the same fashion on the gaseous medium in which the target is placed. This points to the relation between the ion scattering by the tantalum target and the chemical reaction in which the ions with a mass of 245 ± 1 (probably, $TaO_2Na^+ \cdot H_2O$ ions) are produced on the target surface.

The results of this paper show that, by placing a metal target into a plasma jet, it is possible to record a chemical reaction on a solid surface by analyzing the longitudinal-velocity spectra of the jet-plasma ions scattered by the metal surface.

ACKNOWLEDGMENTS

This work was supported by the Ministry of Education of the Russian Federation (Grant on Basic Research in Nuclear Engineering and Physics of Ionizing-Radiation Beams).

REFERENCES

1. M. D. Gabovich, N. V. Pleshchivtsev, and N. N. Semashko, *Ion and Atom Beams for Controlled Nuclear Fusion and Technological Purposes* (Énergoatomizdat, Moscow, 1986).
2. W. Ensinger, Rev. Sci. Instrum. **63**, 5217 (1992).
3. G. I. Dimov and G. V. Roslyakov, Vopr. At. Nauki Tekh., Ser.: Termoyad. Sintez **3/16**, 3 (1984).
4. V. I. Batkin, Zh. Tekh. Fiz. **66** (6), 37 (1996).
5. V. I. Batkin and O. Ya. Savchenko, in *Contemporary Electric Technology in Engineering* (TGU, Tula, 1997) p. 211.
6. V. I. Batkin and O. Ya. Savchenko, Prikl. Mech. Tech. Fiz. **36** (6), 35 (1995).
7. C. F. Barnett and M. F. A. Harrison, *Applied Atomic Collisions Physics Plasmas* (Academic, New York, 1984; Énergoatomizdat, Moscow, 1987).
8. *Niobium and Tantalum*, Ed. by O. P. Kolchin (Inostrannaya Literatura, Moscow, 1960), p. 485.

Translated by N.F. Larionova

Technical Physics, Vol. 45, No. 2, 2000, pp. 251–254. Translated from Zhurnal Tekhnicheskoĭ Fiziki, Vol. 70, No. 2, 2000, pp. 110–113.

EXPERIMENTAL INSTRUMENTS AND TECHNIQUES

An Automated System for the Image Processing and Reconstruction of the Density Field of an Aerodynamic Object

N. P. Mende, A. B. Podlaskin, and V. A. Sakharov

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

Received July 2, 1998

Abstract—A software package is developed that automatically processes interferograms represented in electronic format and provides the tomographic reconstruction of the density field of an aerodynamic object. This software installed on a personal computer equipped with a scanner comprises a workstation for processing results of gas-dynamic experiments. Efficient mathematical methods combined with modern engineering tools improve the reliability, accuracy, and speed of obtaining quantitative experimental results. © *2000 MAIK "Nauka/Interperiodica".*

Optical tomography as a method for studying the spatial structure of aerodynamic objects has been known since the early 1960s. It relies on mathematical methods of interference image processing and on the principle of tomographic synthesis of information collected at different object illumination directions [1]. Therefore, the problem of reconstructing the internal structure of an object can conventionally be divided into two stages: the image (interferogram) processing and the density field reconstruction.

The goal of the first stage, interferogram processing, is to obtain the optical path difference (OPD) distribution of the interfering beams. When two waves interact, one of which has passed through the studied inhomogeneity and the other through a constant-density medium, illumination of the interference field (*xy*-plane) varies as a function of the phase difference (optical path difference) between the interfering waves. This path length difference $S(x, y)$ is related to the refractive index $n(x, y, z)$ of the studied object by the following equation:

$$S(x, y) = \frac{1}{\lambda}\int_L (n(x, y, z) - n_0)\,dz,$$

where n_0 is the constant refractive index of the unperturbed medium, L is the geometric path of the light rays through the inhomogeneity, λ is the wavelength of the sounding radiation, and z is the coordinate along the propagation path.

A relationship exists between refractive index n of a gaseous medium and its density ρ, which is usually described by the Gladstone–Dale formula

$$\frac{n-1}{\rho} = \text{const}.$$

The interferogram maps isolines of the wave surface of the sounding beam distorted by the inhomogeneity with respect to the reference beam surface. When one passes from one illumination extremum on the image to an adjacent one of like shade (i.e., from one light interference fringe to another or between dark fringes), the corresponding optical path difference changes by one wavelength λ. Thus, the interferograms can be processed; i.e., the shape of the wave surface can be reconstructed, by fixing the fringe numbers at the measurement points and the coordinates of these points in the chosen reference system.

The second stage, reconstruction of the gas density, consists in solving the above integral equation (or a system of such equations) by one of the appropriate techniques. A great number of approaches to solving this problem exist, and the choice of a specific method is governed, in particular, by properties of the object under study. For example, when a body moves in a gaseous medium at an ultrasonic speed, gas flow regions appear at whose boundaries the gas-dynamic parameters have a step discontinuity. In this situation, approximating these parameters by a function continuous in the entire cross section is physically meaningless because this would blur the discontinuities. In our opinion, it is advantageous in this case to represent the OPD distribution in the cross section by a piecewise-continuous function. This representation also determines the choice of the method for reconstructing the density field. We have selected the Shardin zone method that was first chronologically. It can easily perform the layerwise splitting of the studied cross section into

1063-7842/00/4502-0251$20.00

domains bounded by singularities of the gas density distribution function [2]. In this method, each region (layer) of the flow is divided into annular zones representing the OPD distribution in each zone by a simple function. The density distribution in the external zone can be obtained by processing the OPD accumulated over the rays that cross this zone alone. Then the contribution of the external region is subtracted from the OPD obtained above (compensation for the external zone); the object's dimensions apparently decrease by the width of the external zone, and the problem is reduced to the application of the above procedure to the residual OPD.

The computer advancements achieved in recent decades have facilitated the development and improvement of mathematical models and methods of the reconstructive tomography. Against this background, the low degree of automation in receiving raw data, in particular, in measuring photographic images, looks like an oppressive anachronism. Application of optomechanical instruments (this trend was being developed some time ago) requires a comparatively long amount of time for processing photographic images and is the main source of errors in the process of creating raw data files, primarily because a number of operations are performed manually.

At the Physical Gas Dynamic Laboratory of the Ioffe Physicotechnical Institute, Russian Academy of Sciences, an automated system was developed that processes electronic images and subsequently reconstructs the density field of a three-dimensional or an axially symmetric flow around a body of revolution and visually represents the results of calculations. The software package is designed for the Windows environment and consists of three independent modules: a program for automated image processing (*Device*) and two programs for reconstructing the density distribution of the three-dimensional (*Tomograph*) and axisymmetric (*Axisym*) flow around a body of revolution. The user controls the programs with the help of a keyboard and a mouse.

To start working with the *Device* program (henceforth referred to as the meter), one should have an image, in the form of a bitmap file, of the interference pattern of the object under study (one for the axisymmetric case or several for three-dimensional flows). Operating principles of the meter comply with requirements imposed by the utilized density reconstruction algorithm [3] and provide raw data for the *Tomograph* or *Axisym* program.

The image processing program includes graphic and text editors. The figure illustrates a display of the *Device* program with an open window of the text editor (*Edit*), which can be moved to any place of the screen or can be closed. The meter offers the following features:

(1) The measurement point visualized by cross *1* can be moved throughout the image in two directions of the orthogonal coordinate axes.

(2) Using the graphic editor, one can set an arbitrarily oriented rectangular coordinate system and scale the image in chosen units in order to use it in measurements (the *Set Axis* drop-down menu option open in the figure).

(3) The meter can draw a point, a segment, or a circle on the image; can move the cross to the center of the segment (the *Draw* option of the menu); and can measure the segment length and its angle of inclination. This is oftentimes necessary when finding characteristic points on the image to determine its scale and adjust the coordinate system.

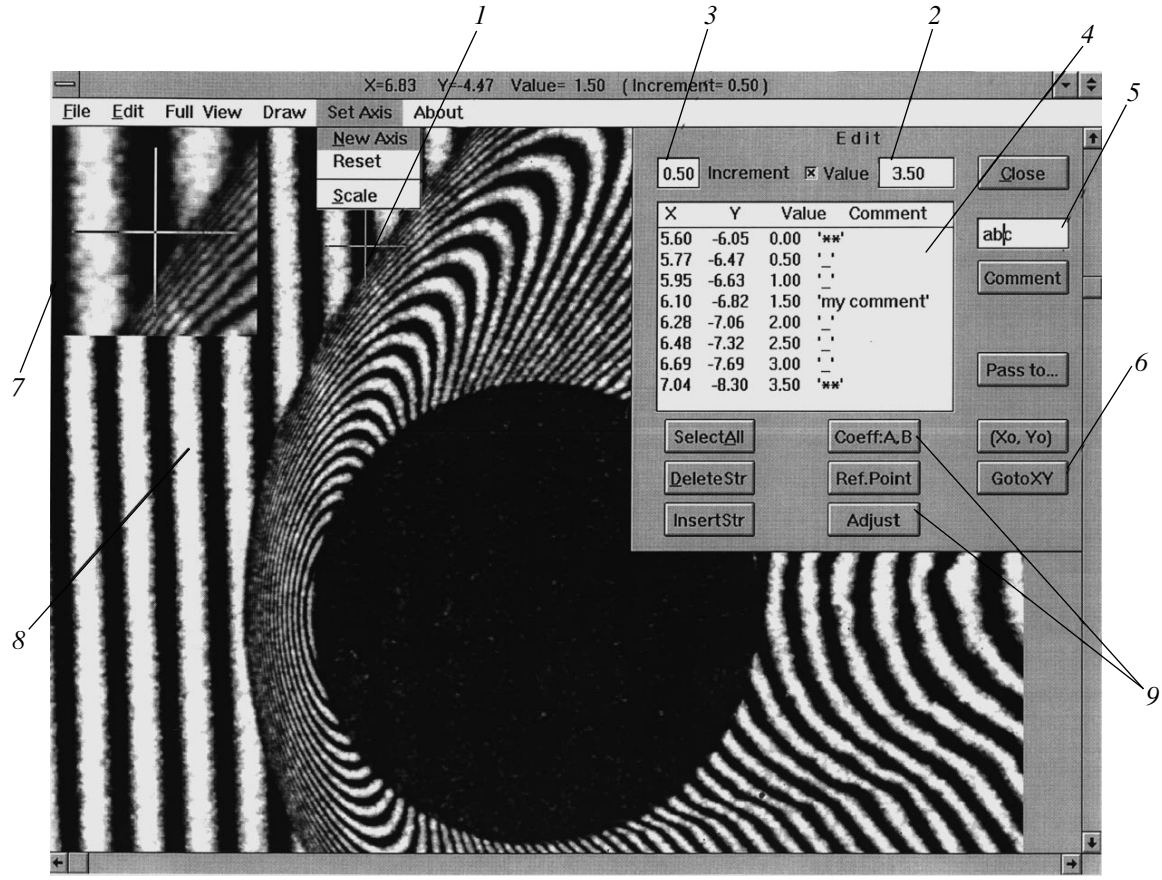
(4) The *x*- and *y*-coordinates of the measurement point corresponding to the position of the cross on the image and the number of the interference band *2* can be simultaneously written to the computer. The fringe number automatically changes by a specified value *3* in response to pressing the enter key: by 1 when passing to an adjacent fringe of the same shade and by 0.5 when passing to a fringe of the opposite shade.

The OPD readings are taken from the interferograms by counting the fringes with measurements being performed only at the illumination maxima and minima, which fall at centers of the interference fringes. This technique eliminates errors associated with nonlinearity of the photographic material, with nonuniformity of the interference pattern contrast, and with similar factors that distort the actual illumination distribution between the extrema.

(5) The built-in text editor can be used to browse and change list of points *4* and to supply this list with comments *5*. A special comment is introduced to outline boundaries of the weak and strong gas-dynamic discontinuities in order to identify boundaries of the flow regions, which is necessary for the density reconstruction programs.

For convenience, the editor relates position *6* of the cross on the screen to the chosen point on the list and, in addition to the full-screen image, its magnified fragment containing the cross can be displayed in a separate window *7* in order to position the cross more accurately. One can also display the entire scanned image on a smaller scale (the *Full View* menu option) and quickly output the necessary part of it to the full screen.

The final result of interference processing is a set of OPD values. These values are equal to the difference between the fixed fringe numbers and contribution to the OPD due to alignment of the interferometer. The interference pattern of the unperturbed space that is in the field of view of the instrument referred to as the alignment field is either an almost uniformly illuminated field (when the interferometer is tuned to an infinitely wide fringe) or almost equispaced straight lines *8* (when the interferometer is tuned to fringes of a finite width), which allows one to approximate them by a lin-



Display of the image meter.

ear two-dimensional function. In the first case, the contribution to the OPD is constant over the entire field; in the second case, it depends on the position of the measurement point. In the latter case, to account for the contribution to the OPD introduced by the instrument, it is necessary first to make measurements in the region of the interferogram unperturbed by the flow and then, after calculating coefficients of the approximating function, to compute the refined OPD in the studied cross section. These procedures are implemented in the meter *9* and their results can be displayed using tools offered by the editor. Afterwards, results of the measurements are stored in the raw data file and are ready to be used by the programs that reconstruct the density field.

Both program modules that implement the density field reconstruction procedure use files of a standard format created by the meter. Since the *Axisym* program, which reconstructs the density of the axisymmetric flow, is a particular case of the more general *Tomograph* program intended for three-dimensional flows, we consider the features offered by the latter program for generality of presentation.

The *Tomograph* density field reconstruction program performs the following operations:

(1) Selects and lists raw data (measurement) files subject to further processing. In order to prevent failures in the program, it verifies mutual correspondence of the special comments that specify the internal gas-dynamic structure of the object between different files of the list. With the built-in text editor, one can view the content of any of these files, add or remove a file, and reverify the files on the list.

(2) Enters the physical parameters at which the images were taken, the parameters that specify calculation conditions, and dialog options that control the current calculations.

(3) Reconstructs the density field by the zone method, which divides the processed cross section into separate layers. The employed algorithm [3] allows for internal gas-dynamic discontinuities (no more than four, including the front step discontinuity and the surface of the object immersed in the flow). It contains procedures that use a low-degree algebraic polynomial to smooth the radial OPD distribution when passing from one zone to another and approximates the azimuth distribution by a truncated Fourier series. The accuracy of this density reconstruction method was analyzed in [4].

(4) Monitors operation of the program by displaying intermediate results for any preset zone of the studied cross section. If the calculations are interrupted, it generates a message at the stage at which this occurred and takes possible measures to correct the error.

(5) When the calculations are completed, the results are displayed on the screen as plots of dimensionless gas density versus azimuth and radial coordinates in the inhomogeneity cross section.

(6) In order to analyze the process of calculations after it terminates, the program creates a data file containing information about the main computation stages. Optionally, the information content of this file can be changed.

(7) After the calculations are completed, the tables of the reconstructed density can be written to files, which can be displayed by external graphic editors.

The variety and complex internal structure of the studied objects prevent one from presetting control parameters of the reconstruction algorithm so as to provide an acceptable accuracy. Among these parameters are the number of flow regions, the number of zones in the cross section and distribution of these over the regions, the degree of the approximating polynomials, etc. As a result, the reconstruction procedure implies that the above parameters should be chosen by the trial-and-error method, which, of course, requires the researcher to know the specifics of the studied object and the essence of the calculation procedure.

The accuracy of the reconstructed density is determined, in particular, by the error in interferometric measurements, which directly depends on the quality of the electronic copy. Therefore, when creating it, one should provide the minimal necessary scanner resolution, which, as a rule, is determined by the quality of the optical system. In particular, the resolution of a shearing interferometer constructed based on an IAB-451 optical instrument is ten lines per millimeter and the minimal resolution necessary for scanning a 1 : 1 image must be no lower than 500 dpi. In this case, the error in determining coordinates of the fringe center is about 0.1 mm; in the region of the flow with the maximum fringe density (a 0.1-mm-long interfringe distance), the error in the OPD reconstructed in the cross section can be as large as one bringe. Naturally, increasing the scanner resolution improves the accuracy of determining the point coordinates from the electronic copy and, therefore, reduces the error in reconstructing the shape of the wave surface (values of the OPD). However, there is a limit, beyond which it is inefficient to further improve the resolution. For example, if the error in interferometric measurements is about a 0.1 fringe, a 2500-dpi resolution is sufficient to provide this measurement error in the flow region with the highest fringe density. Clearly, in the regions where the fringe density is lower, the required accuracy of reconstructing the OPD can be provided at a lower scanner resolution. Also note that the requirements imposed on the scanner resolution can be relaxed proportionally to an increase in the dimensions of the scanned photographic image.

Our experience has shown that the OPDs measured from the interferogram and from its 600-dpi electronic copy almost coincide. Nevertheless, when creating an electronic image, one should take into account in each particular case not only the quality of the optical system and features of the available electronic equipment, but also the specific properties of the object under study and goals at hand.

In order to demonstrate the accuracy of the density field reconstruction algorithm, we consider solutions to problems of supersonic (at Mach's number $M = 2$) motion of an acute cone at a nonzero angle of incidence [5] and of supersonic (at $M = 19.3$) motion of a sphere [6] in the atmosphere. In the first case, the density field was reconstructed from five projections divided into eight zones. In the cross sections that intersect the cone's surface, the density reconstruction error was no higher than 10%. In the axially symmetric problem of motion of the sphere, it is sufficient to use only one projection with the number of zones in the cross section set equal to the number of points at which the OPD is measured. In this case, the density was determined with an error of no higher than 2%.

The automated image processing system substantially simplifies measurements and the creation of raw data files, which are the most time-consuming stages in the tomographic reconstruction of the density field of an aerodynamic object.

REFERENCES

1. V. S. Tatarenchik, in *Proceedings of the Zhukovsky Air Force Engineering Academy: Investigation of Spatial Gas-Dynamic Flows Based on Optical Methods* (Moscow, 1964), No. 1059, p. 11.
2. V. A. Komissaruk and N. P. Mende, in *Optical Research Methods in Ballistic Experiments*, Ed. by G. I. Mishin (Nauka, Leningrad, 1979), p. 178.
3. V. A. Komissaruk, N. P. Mende, and L. N. Popov, Preprint No. 1349, FTI im. A. F. Ioffe AN SSSR (Ioffe Physicotechnical Institute, Academy of Sciences of the USSR, 1989).
4. N. P. Mende, Preprint No. 1350, FTI im. A. F. Ioffe AN SSSR (Ioffe Physicotechnical Institute, Academy of Sciences of the USSR, 1989).
5. N. P. Mende and A. B. Podlaskin, Zh. Tekh. Fiz. **64** (7), 9 (1994) [Tech. Phys. **39**, 637 (1994)].
6. N. P. Mende and Yu. É. Shtengel', Preprint No. 1564, FTI im. A. F. Ioffe AN SSSR (Ioffe Physicotechnical Institute, Academy of Sciences of the USSR, 1991).

Translated by A.D. Khzmalyan

Technical Physics, Vol. 45, No. 2, 2000, pp. 255–259. Translated from Zhurnal Tekhnicheskoĭ Fiziki, Vol. 70, No. 2, 2000, pp. 114–117.

EXPERIMENTAL INSTRUMENTS AND TECHNIQUES

Photosensitivity of Semiconductor–Protein Systems

V. Yu. Rud'[1], Yu. V. Rud'[2], and V. Kh. Shpunt[2]

[1] *St.-Petersburg State University, St.-Petersburg, 195251 Russia*

[2] *Ioffe Physicotechnical Institute, Russian Academy of Sciences, St. Petersburg, 194021 Russia*

Received July 13, 1998

Abstract—Photovoltaic effects have been revealed in a new class of semiconductor/B heterojunctions. Considering various semiconductor materials (Si, GaAs, InSe, $CdSiAs_2$, $ZnGeP_2$, and $CuGaS_2$) and native proteins, we found that it is possible to create photosensitive structures on their basis and managed to prepare them. The photoelectric parameters of the semiconductor/B systems are measured for the first time. It is shows that these structures have a photosensitivity whose level is characteristic of solid-state photoconverters. The spectral dependences of the quantum efficiency of photoconvertion are studied. The window effect for such structures has been established: the long-wavelength boundary of photosensitivity is determined by the energy gap of a semiconductor, whereas the short-wavelength boundary in the vicinity of 3.55 is explained by the quasi-interband transitions in the wide-gap component common for all the heterojunctions—the protein. The conclusion is drawn that the structures of this new class based on crystals of anisotropic semiconductors can be used as broad-band photoconverters of the natural radiation and photoanalyzers of a linearly polarized radiation.

The development of semiconductor optoelectronics provided the creation of a large group of heterophotoelements for various purposes [1–5]. Recently, we have revealed the appearance of the photovoltaic effect in the contact between a semiconductor and a native protein. Below, we consider the first results of the experimental studies of the photoelectric phenomena observed in the heterojunction of diamondlike semiconductors of various crystal classes with proteins, which are of great interest for both fundamental and applied sciences.

1. Photosensitive structures were prepared by the application of a native protein onto the surface of $5 \times 5 \times 0.5$-mm large electrically homogeneous semiconductor wafers. The table lists the semiconductors used, their energy-gap values E_G, and the concentrations of free charge carriers 1/Re determined from the measurements of the Hall coefficients. The surfaces of Si, GaAs, $CdSiAs_2$, and $CuInS_2$ wafers were preliminarily polished: first mechanically and then chemically. The InSe wafers had natural mirror-smooth cleavages. The as-grown $ZnGeP_2$ and $CuGaS_2$ crystals had mirror-smooth (112) faces, which were used in the experiments without any additional treatment. The electric contacts applied to the protein were thin (0.5–1 μm) Ni- or Mo-layers on a glass platelet. The metal-coated surface of the glass plate was brought into contact with the protein, which acquired the shape of a 50–100 μm plane-parallel protein layer depending on contact type. The position of the semiconductor crystal with respect to the metallized glass was fixed with the aid of a dielectric lacquer, which also prevented protein drying and, at the same time, the direct contact of the metal with the semiconductor.

2. All the heterojunctions obtained by the above method had the rectifying current–voltage characteristics illustrated by Fig. 1 for the *n*-InSe/protein structure (hereafter, the protein is denoted by B). The positive polarity of the external voltage applied to the metallized glass surface corresponds to the forward direction. It should also be emphasized that in the course of the further experiments, the current–voltage character-

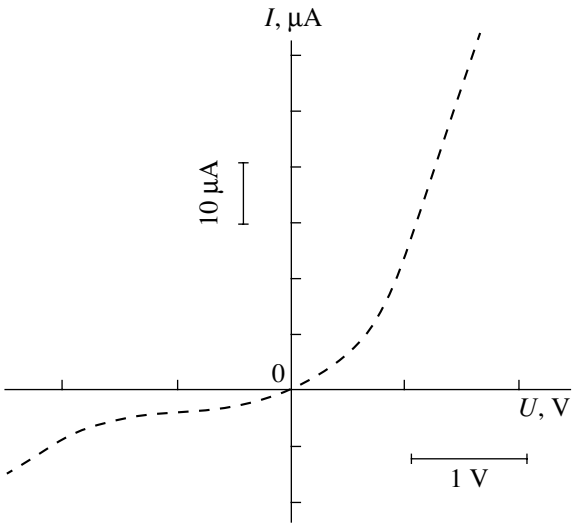


Fig. 1. Stationary current–voltage characteristic for the *n*-InSe/B structure at $T = 300$ K (the contact area ≈ 0.1 cm^2).

1063-7842/00/4502-0255$20.00 © 2000 MAIK "Nauka/Interperiodica"

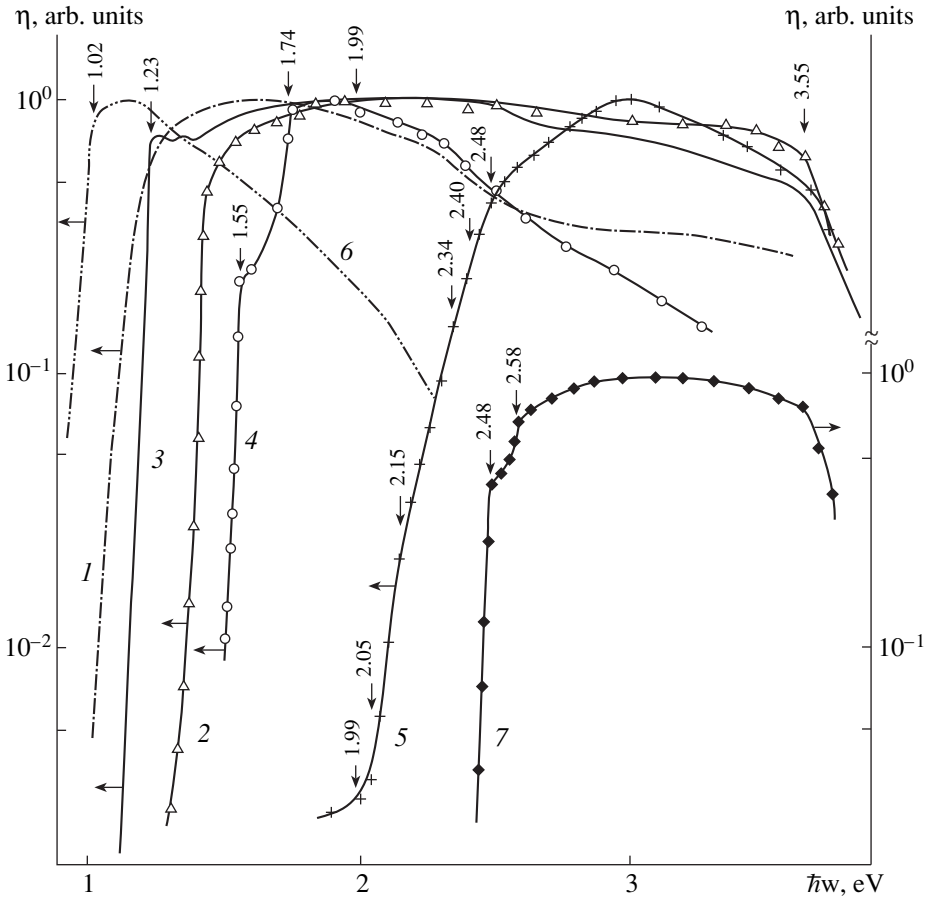


Fig. 2. Spectral dependences of the relative quantum efficiency of photoconversion η in the semiconductor/B structures illuminated with natural light at $\hbar\omega$ = 300 K. Semiconductors used: (*1*) Si, (*2*) GaAs, (*3*) InSe, (*4*) $CdSiAs_2$, (*5*) $ZnGeP_2$, (*6*) $CuInSe_2$, and (*7*) $CuGaS_2$; illumination from the protein side along the normal to the photodetector plane, spectral resolution ~1 meV.

istics of the semiconductor/B structure showed no degradation effects. The rectification coefficient of the structures studied was usually low and, at the voltages of the order of 2 V, ranged within 2–4, whereas the residual resistance was determined mainly by the parameters of the semiconductors used.

3. The illumination of the semiconductor/B heterojunctions gave rise to the appearance of photovoltage whose sign was independent of the energy of incident photons and the position of the light probe (0.2 mm in diameter) on the photosensitive surface of the structures. This leads to the assumption that the photovoltaic effect is associated with the separation of the photoinduced carriers by the only active region formed at the heteroboundary between the semiconductor surface and the protein "wetting" this surface.

In all the structures based on the group of semiconductors with the energy gap E_G ranging within 1.0–2.5 eV and a protein (the same in all the cases), the most pronounced photosensitivity was observed at their illumination through the protein layer. This leads to the assumption that in these structures, the protein plays the role of a wide-gap component. The table lists the maximum voltages, S_U, and the current photosensitivities of the heteropairs used, S_I. The maximum photosensitivity was observed for the *n*-InSe/B structures. On the whole, it was also noticed that photosensitivity of the heterojunctions based on ternary semiconductors was not worse than the photosensitivity of other types of photoconverters prepared earlier using the crystals of analogous quality [5–9]. Taking into account the simplicity of the structure preparation and the widespread occurrence of proteins in nature, the use of this new class of photoconverters seems to be very promising.

4. Typical spectral dependences of the relative quantum efficiency of photoconversion η determined as the ratio of the short-circuit photocurrent to the number of incident photons for heterostructures based on various semiconductor materials illuminated with a natural radiation from the side of the protein are shown in Fig. 2. It is seen that the energy position and the structure of the long-wavelength edge of the heterostructure photosensitivity is determined by interband absorption in the energy spectrum of the semiconductors used.

Indeed, the long-wavelength edge of photosensitivity in the coordinates corresponding to the character of the interband absorption becomes linear, whereas the extrapolation of these dependences to the zero photon energy yields the E-values close to E_G-value for these semiconductors. Thus, the long-wavelength edge η for the *n*-Si/B structures follows the root law characteristic of indirect junctions (Fig. 3, *1*), whereas η for the *n*-InSe/B and *n*-GaAs/B, to the parabolic law characteristic of the direct junctions (Fig. 3, curves *2* and *3*).

In the structures based on the anisotropic $CdSiAs_2$ and $CuGaS_2$ semiconductors (Fig. 2, curves *4* and *7*), a fine structure is observed, which is formed due to the junctions from the subbands of the valence band split by the anisotropic crystal field [6]. Earlier, similar features were observed for the spectral dependences of η for the solid structures based on the same crystals [4–6].

The main distinctive feature of the photosensitivity of the semiconductor/B heterojunctions is the fact that the η maximum is attained in the depth of the fundamental absorption band in all the structures of these semiconductors ($\hbar\omega > E_G$). Therefore, the spectra of η show the clearly pronounced window effect, and the high photosensitivity is observed over a wide spectral range. In the table, the broad-band character of the photosensitivity spectra is seen from the values of the total width of the η spectra at their half-width $\delta_{1/2}$ and the spectral range of the maximum photosensitivity $\Delta\hbar\omega_m$. It should be emphasized that the window effect in most of the semiconductor/B structures studied is of the same level as in the ideal solid heterojunctions [1]. However, in the latter heterojunctions, the rigid conditions imposed on possible heteropartners should be rigorously fulfilled [1, 3], whereas in the new class of heterojunctions, no such conditions exist at all, which, in fact, lifts all the limitations on the choice of a semiconductor to be used in the structure.

Figure 2 also shows that photosensitivity in the Si-, InSe-, GaAs-, and $CuGaS_2$-based structures shows no drastic short-wavelength decrease up to a value of 3.5 eV. This indicates that the optical absorption by the protein in this range is not strong enough, and the decrease in η at $\hbar\omega > 3.55$ eV can be caused by the beginning of the quasi-interband absorption in the energy spectrum of the protein (Fig. 2, curves *2*, *3*, *5*, and *7*). The decrease of η in the range $\hbar\omega > E_G$ observed for the $CuInSe_2$-, $CdSiAs_2$-, $ZnGeP_2$-based structures (Fig. 2, curves *4*–*6*) can be explained by the properties of the crystals used, because such a decrease is also observed in their photoconductivity spectra, and also in the photosensitivity of the diode structures prepared from these crystals [6, 10]. Thus, the photosensitivity of the $ZnGeP_2$/B structure is determined by the pseudodirect and direct interband junctions and correlates with the data on the surface-barrier structures [10]. In this connection, one of the promising applica-

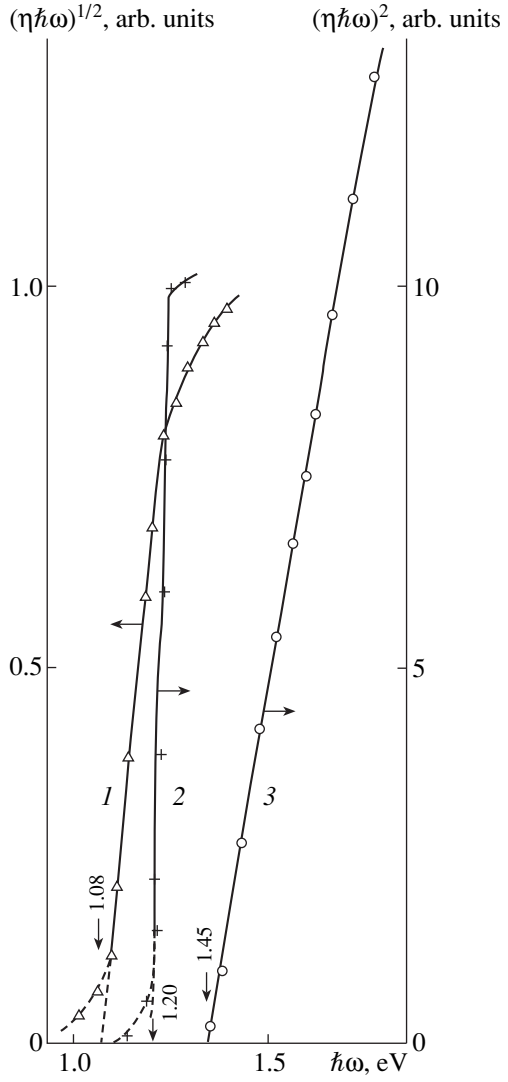


Fig. 3. Spectral dependences η for the semiconductor/protein structures at T = 300 K. Semiconductors: (*1*) Si, (*2*) InSe, (*3*) GaAs.

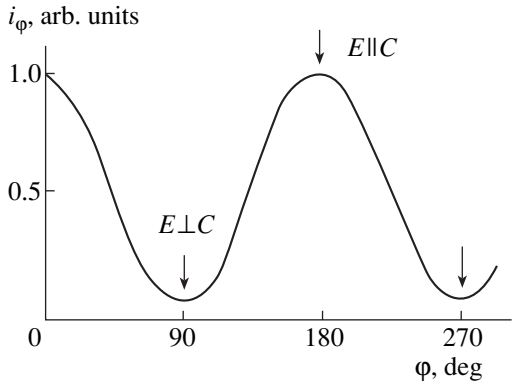


Fig. 4. Polarization indicatrix of the short-circuit photocurrent for the *p*-$CdSiAs_2$/B structure at T = 300 K ($\hbar\omega$ = 1.58 eV, (110)-oriented $CdSiAs_2$ wafer illuminated along the normal to the photodetector plane from the protein side).

Photoelectric properties of some semiconductor/protein systems at $T = 300$ K

Semiconductor	E_G, eV	1/Re, cm^{-3}	S_U^m, mA/W	S_i^m, mA/W	$\delta_{1/2}$, eV	$\Delta\hbar\omega_m$, eV
n-Si	1.1	10^{15}	10	0.4	1.2	1.5–1.6
n-GaAs	1.43	10^{17}	230	1	2.3	1.9–2.5
n-InSe	1.25	10^{15}	10^4	2	2.3	2.0–2.5
p-$CdSiAs_2$	1.55	10^{16}	610		0.85	1.75–2.0
p-$ZnGeP_2$	1.99	10^{10}	920		1.07	3.0
p-CuInSe	1.02	10^{18}	20		≈1.8	1.1–1.2
p-$CuGaS_2$	2.46	10^8*	50		0.54	2.9–3.3

* Resistivity is given in Ω cm.

tion of the $ZnGeP_2$/B heterojunctions, which does not require any special treatment of the semiconductor, is the diagnostics of the near-edge absorption in $ZnGeP_2$, which is of great importance for its use in highly efficient transducers of the radiation frequency [11].

5. In the structures based on the anisotropic oriented semiconductors illuminated from the protein side along the normal to the photodetector plane, the short-circuit photocurrent follows the Malus law (Fig. 4) as in the case of the solid structures [6]. In this case, the photocurrent with the polarization **E** ∥ **c** (where *c* is the unit vector along the fourfold axis of the tetragonal crystal), the value of i^p is much more higher than the value of i^s (**E** ⊥ **c**) in the whole range of the polarization photosensitivity. The inequality $i^p > i^s$ is determined in accordance with the selection rules for interband optical *A*-transitions in $CdSiAs_2$ [6].

The spectral dependence of the coefficient of natural photopleochroism P_N of the *p*-$CdSiAs_2$/B structure is shown in Fig. 5. This dependence is also similar to the corresponding dependences characteristic of solid photoconverters based on this semiconductor [6]. In accordance with the selection rules, photopleochroism is positive and attains the maximum value in the range of the interband *A*, whereas its decrease in the short-wavelength range is determined by the transitions from the valence-band subbands split due to the crystal-field effect and the spin–orbital interactions allowed mainly for the polarization **E** ⊥ **c**. We should like to emphasize that the sign of P_N in the depth of the fundamental absorption band in $CdSiAs_2$ remains positive. This allows one to state that the heteroboundary in the new class of the structures is almost perfect with respect to the recombination processes.

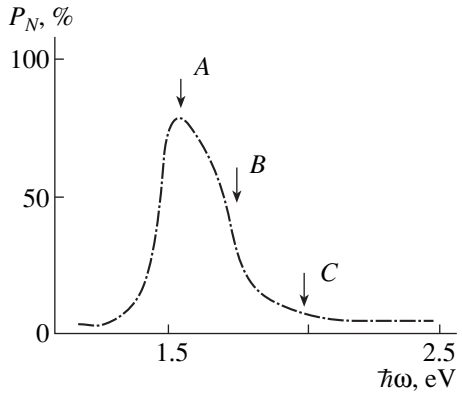


Fig. 5. Spectral dependence of the coefficient of natural pleochroism of the B/$CdSiAs_2$ structures. The arrows indicate the energies of the interband transitions in $CdSiAs_2$ [10].

Comparing the polarization properties of the $CdSiAs_2$/B structures and their solid analogues [6, 10], one can state that the penetration of low-polarized radiation through the protein layer into the active region of such heterojunctions does not change the polarization characteristics of the radiation.

6. Thus, the semiconductor–protein contact shows the photovoltaic and polarimetric (in the case of anisotropic semiconductors) effects, which can be used in the design of photosensors of the optical radiation and in the studies of protein metabolism.

REFERENCES

1. Zh. I. Alferov, Fiz. Tekhn. Poluprovodn. (St. Petersburg) **11**, 2072 (1977).
2. Zh. I. Alferov, Fiz. Tekhn. Poluprovodn. (St. Petersburg) **32** (2), 3 (1998).
3. A. Milnes and D. Feucht, *Heterojunctions and Metal–Semiconductor Junctions* (Academic, New York, 1972; Mir, Moscow, 1975).

4. A. A. Abdurakhimov, Yu. V. Rud', K. V. Sanin, *et al.*, Zh. Tekh. Fiz. **53**, 325 (1983).
5. V. Yu. Rud', Yu. V. Rud', and V. Kh. Shpunt, Fiz. Tekh. Poluprovodn. (St. Petersburg) **29**, 1178 (1995) [Semicond. **29**, 608 (1995)].
6. V. Yu. Rud', Yu. V. Rud', M. Serginov, *et al.*, Phys. Status Solidi A **113** (1), 207 (1989).
7. A. A. Abdurakhimov, L. B. Karlina, Yu. V. Rud', *et al.*, Pis'ma Zh. Tekh. Fiz. **24**, 1477 (1982).
8. Yu. V. Rud' and M. A. Tairov, Fiz. Tekh. Poluprovodn. (Leningrad) **21**, 615 (1987) [Sov. Phys. Semicond. **21**, 377 (1987)].
9. A. A. Abdurakhimov, A. V. Lunev, Yu. V. Rud', *et al.*, Izv. Vyssh. Uchebn. Zaved., Fiz. **25** (7), 7 (1985).
10. Yu. V. Rud', Izv. Vyssh. Uchebn. Zaved., Fiz. **26** (8), 68 (1986).
11. V. Yu. Rud'. Yu. V. Rud', M. C. Ohmer, *et al.*, in *Proceedings of Materials Research Society*, Ed. by M. Monarech, T. Mayers, and F. Julich (Boston, 1996), Vol. 450, pp. 309–315.

Translated by L. Man

Technical Physics, Vol. 45, No. 2, 2000, pp. 260–262. Translated from Zhurnal Tekhnicheskoĭ Fiziki, Vol. 70, No. 2, 2000, pp. 118–120.

EXPERIMENTAL INSTRUMENTS AND TECHNIQUES

Microdoping of Subsurface Gallium Arsenide Layers with Hydrogen Ions

V. V. Anisimov[1], V. P. Demkin[1], I. A. Kvint[1], S. V. Mel'nichuk[2], and B. S. Semukhin[1]

[1] *Tomsk State University, Tomsk, 634050 Russia*

[2] *Institute of Physics of Strength and Materials Science, Siberian Division of the Russian Academy of Sciences, Toms, 634021 Russia*

Received July 29, 1998

Abstract—A method of microdoping subsurface semiconductor layers with hydrogen ions (protons) with the use of a plasma-beam discharge is suggested. The method was tested on gallium arsenide layers and was proven to be more efficient than other well-known methods used for modifying subsurface layers. © *2000 MAIK "Nauka/Interperiodica".*

Subsurface ion microdoping of semiconductors is used for the creation of materials with a "microcomposite" structure possessing new properties. Low-temperature ion saturation of the subsurface layers of a material is accompanied by migration of a considerable number of ions into the material bulk and their low-temperature incorporation into the subsurface region [1]. The dopants can be various ions of which most widespread are hydrogen ions or protons (H^+). There are several methods for introducing protons into the material including hydrogenation, implantation, and microdoping (protonation) [2]. Among the methods of microdoping, the most widespread is the use of powerful ion beams [2]. However this method is very labor-consuming and expensive. Moreover, in some instances, where the experimental conditions (high vacuum, absence of impurities, etc.) should be strictly maintained, this method fails to yield the satisfactory results. We suggest a new method of the treatment of the material surface with plasma. Plasma is a good source of ions formed in the plasmochemical processes, and, especially, it is a plasmochemical reactor [3] based on the use of a high-voltage discharge.

It is well known that a high-voltage pulse discharge is characterized by the presence in the plasma of a group of fast electrons with the energies comparable with the voltage applied to the electrodes [4]. Creating a narrow discharge gap between the cathode and the grid anode, it is possible to obtain a high-energy electron beam providing a high ion concentration in the plasma. We used the He + H_2 mixture as the working gaseous mixture for a plasmochemical reactor. The presence of the buffer gaseous helium provided the presence of high-energy electrons in the beam, which, in turn, provided the transport of H^+-ions from the anode over macroscopic distances (of the order of several centimeters) [5]. Protons interact with the target surface and, if their concentration is high enough, can efficiently saturate the layers of the target material. In order to create the necessary proton concentration in the vicinity of the target surface, one has to select the appropriate discharge parameters and the He : H_2 ratio in the gaseous mixture. Our aim was to study the parameters of the high-voltage plasma-beam discharge in order to determine the range of the experimental conditions necessary for the optimum saturation of the subsurface layers of a semiconductor material (gallium arsenide, GaAs). The He : H_2 concentration ratio in the gas mixture was selected proceeding from the necessity of attaining the maximum luminescence intensity of the ionic-helium and atomic-hydrogen lines due to transitions from the highly excited states.

The experiments showed that the optimum conditions for protonation in the gaseous He + H_2 mixture with the total pressure P = 9 torr and the concentration ratio He : H_2 = 8 : 1 correspond to the following discharge parameters: the amplitude of the voltage pulse supplied to the accelerating gap d = 0.9 mm is U_0 = 2 kV; the pulse duration is τ = 2 μs; and the pulse frequency is f = 2.5 kHz.

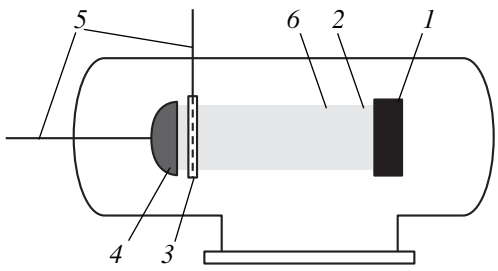


Fig. 1. Gas-discharge source: (*1*) specimen, (*2*) stage, (*3*) grid anode, (*4*) cathode, (*5*) current-supplying electrodes, and (*6*) plasma flare.

1063-7842/00/4502-0260$20.00 © 2000 MAIK "Nauka/Interperiodica"

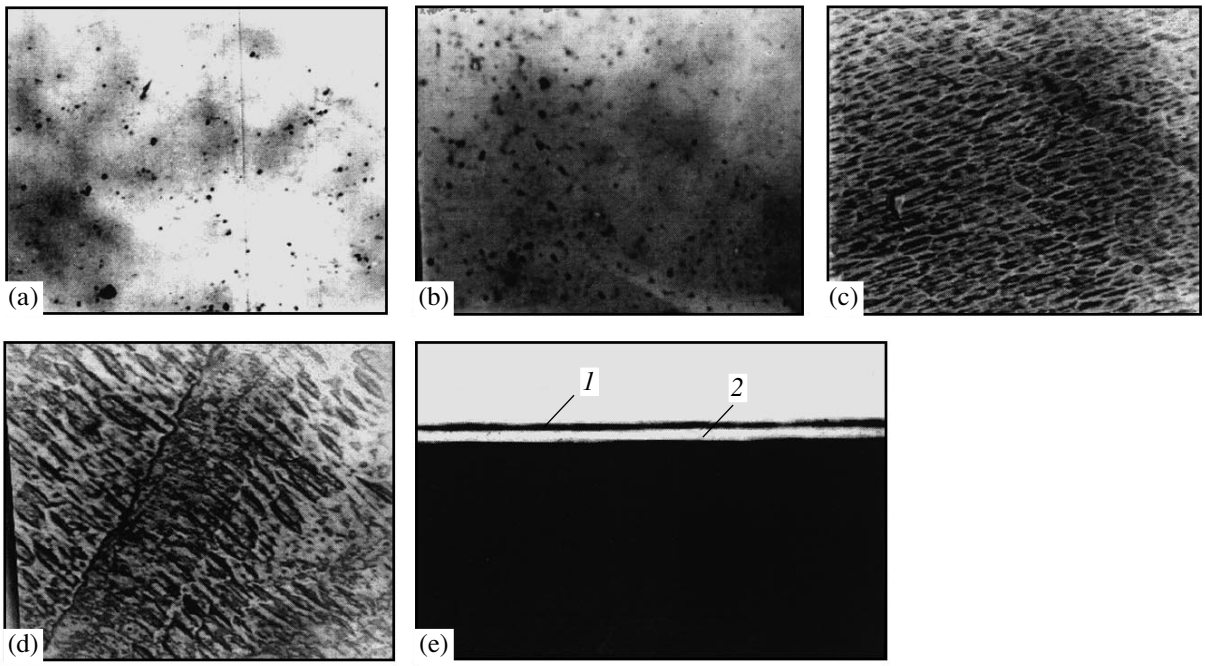


Fig. 2. Surfaces of (a) untreated specimen and (b, c, d) specimen treated for Δt = (b) 1, (c) 2, and (d) 7 min (1000 times magnification); (e) view from side (negative): (*1*) layer consisting of blocks formed due to protonation, (*2*) GaAs matrix.

The proton (H^+) concentration was estimated from the spectroscopic intensities of the He and He^+ lines at different distances from the grid anode and their comparison with the data calculated for the degradation spectra of an electron beam in the He–H_2 plasma [6]. It was shown that the proton concentration in the flare under the given discharge conditions ranged within 10^{13}–10^{14} cm^{-3}, thus providing the ion flux onto the target surface ranging within 10^{17}–10^{18} cm^{-2} s^{-1}.

The protonation experiments were performed as follows. A GaAs specimen (*1*) fixed on a subject table (*2*) was placed into the gas-discharge source at a distance of 15 mm from the grid anode (*3*) (Fig. 1). The GaAs specimens were treated with protons at three different exposures, Δt = 1, 2, and 7 min. The thus treated GaAs surface was studied by the methods of optical microscopy and X-ray diffractometry.

Figure 2 shows the surfaces of (a) untreated and (b, c, d) plasma-treated GaAs specimen at the exposures (b) 1, (c) 2, and (d) 7 min. It is seen from Fig. 2b that 1-min plasma treatment gave rise to no changes in the optical contrast. Figures 2 (c and d) show that 2- and 7-min protonation results in the formation of blocks uniformly distributed over the specimen surface. The block dimensions depend on the exposure—at 2-min treatment, the block dimensions do not exceed 5 μm, whereas at 7-min treatment, the blocks attain the dimensions of 20 μm. Figure 2e shows the side surface of the GaAs specimen. It is established that an approximately 6-μm-long layer *1* consisting of the blocks formed upon protonation is separated from the matrix *2* by a sharp invariant boundary. Figure 3 shows the dependence of the natural logarithm of the block dimension (L) on the treatment time. It is seen that the dependence tends to the equilibrium value already upon 2 min of GaAs treatment.

Figure 4 shows the fragments of the X-ray diffraction patterns from GaAs [the dependence of the integrated intensity of the (200) Bragg reflection on the scattering angle 2Θ] for various times of treatment). It is seen that the integrated intensity of the (200) reflec-

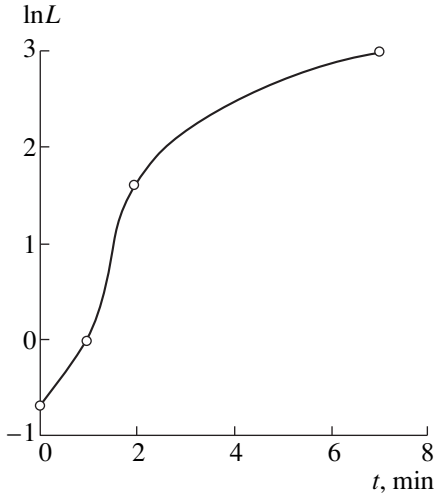


Fig. 3. Temporal dependence of the natural logarithm of the block dimension (L) in (100)GaAs.

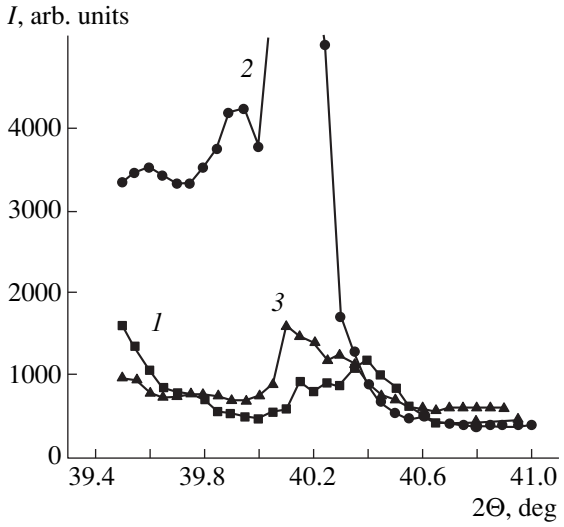


Fig. 4. Dependence of the integrated intensity I of the (200) Bragg reflection on the scattering angle 2Θ, (*1*) initial specimen and (*2*, *3*) the specimen protonated for Δt = (*2*) 2 and (*3*) 7 min.

tion from the specimen treated for 2 min increases four to five times. The initial reflection and the reflection observed from the GaAs specimen treated for 7 min have complicated structures, whereas the reflection observed from the specimen treated for 2 min is rather smooth and can well be described by the Gaussian function. The results obtained indicate that 2-min protonation of the specimen improves the subsurface (10–15 μm) crystallinity in comparison with the crystallinity of the initial specimen, despite the fact that the initial surface of the specimen was the polished single-crystal (100) surface.

Such intensity redistribution indicates the formation of a number of coherently scattering D-regions formed by blocks of different dimensions formed during protonation. Table lists the block dimensions (D) calculated by the well-known Warren method [7]. It is seen that 2-min treatment results in the formation of 400-Å-large blocks, whereas the 7-min treatment results in the formation of blocks with the dimensions ranging from 250 to 340 Å. The treated structure is more homogeneous and is closer to the ideal structure than the initial one. The measurements showed that the structure modification is seen from the formation of the subsurface layer with a homogeneous block structure at various structural levels.

Block dimensions in protonated GaAs

Specimen	D, Å
Initial	344
	420
	631
Treated for 2 min	420
Treated for 7 min	344
	473

To confirm the modification of the material due to its irradiation with protons, we also performed similar experiments in pure He and H_2 gases. The microscopy studies showed that no structure modification occurred upon specimen treatment with He, whereas the specimen treatment with H_2 yielded the positive effect. However, the specimen treatment in pure hydrogen plasma considerably deteriorates the "beam properties" of electrons, which, in turn, requires longer exposures of the material to radiation.

Thus, the use of the He–H_2 plasma beam as a source of protons and their transport to the surface of the materials treated with this plasma does not require any complicated technological solutions and, at the same time, provides the development of intense physical processes during ion formation. These properties allow one to use this type of plasma for designing simple devices for modifying various materials.

REFERENCES

1. A. N. Sergeev, G. E. Remnev, and S. V. Rudnev, *Thin-Film Proton–Ion Modification* (Tomsk University, Tomsk, 1993).
2. H. A. Davis, G. E. Remnev, R. W. Stinnett, *et al.*, MRS Bulletin, p. 58 (1996).
3. A. V. Karelin, V. P. Demkin, and S. V. Melnichuk, in *Proceedings of International Conference on Lasers'96, 1996*, p. 653.
4. Yu. P. Raizer, *Physics of Gas Discharge* (Nauka, Moscow, 1997).
5. V. P. Demkin, S. V. Mel'nichuk, and I. I. Murav'ev, Opt. Atmosfery Okeana, No. 3, 253 (1993).
6. V. P. Demkin and N. L. Kupchinskiĭ, *Processes of Ionization with Allowance for Excited Atoms* (Leningrad, 1988), p. 168.
7. B. E. Warren, Phys. Rev. **59**, 693 (1941).

Translated by L. Man

Technical Physics, Vol. 45, No. 2, 2000, pp. 263–266. Translated from Zhurnal Tekhnicheskoĭ Fiziki, Vol. 70, No. 2, 2000, pp. 121–124.

EXPERIMENTAL INSTRUMENTS AND TECHNIQUES

Raising the Sensitivity of the Electron-Paramagnetic-Resonance Spectrometer Using a Ferroelectric Resonator

I. N. Geĭfman*, I. S. Golovina*, E. R. Zusmanov*, and V. I. Kofman**

**Institute of Semiconductor Physics, National Academy of Sciences of Ukraine, Kiev, 252028 Ukraine*

***North Western University, Evanston, USA*

Received August 26, 1998; in final form, February 8, 1999

Abstract—In order to raise the sensitivity of microwave electron-paramagnetic-resonance (EPR) spectrometers, it is proposed to use a piece of ferroelectric material as an additional resonator. The method has been tested using the RE-1307 microwave EPR spectrometer and a pulsed microwave spectrometer. The possibility of raising the signal-to-noise ratio when using ferroelectric resonators of rectangular-parallelepiped and spherical shape has been considered. For a potassium-tantalate ferroelectric resonator of rectangular-parallelepiped shape, the signal-to-noise ratio has been raised by a factor of 16 at 331 K and by a factor of 10 at 292 K. In the pulse experiment, the presence of the ferroelectric resonator permits a reduction in microwave power, required for sample saturation, by a factor of 50 at 50 K. © *2000 MAIK "Nauka/Interperiodica".*

The electron-paramagnetic-resonance (EPR) method is widely used in physics, chemistry, medicine, biology, and other areas of science. At the initial stage of the development of this method, investigations were mainly performed using specially prepared samples, into which an impurity substance was introduced as a paramagnetic probe. At present, the EPR method is implemented in investigating natural (nondoped) objects, in which the concentration of paramagnetic centers is frequently insufficient for the EPR signal to be observed. For this reason, efforts are being made to increase microwave spectrometer sensitivity. One of the simplest ways to improve the signal-to-noise ratio is to introduce dielectric material into the resonator. In [1], a quartz plate was used for this purpose. Locating the sample near the plate resulted in an increase in the signal-to-noise ratio by a factor of 4.5. TiO_2 [2, 3]; $SrTiO_3$ [4] ferroelectric resonators were also used, but their shape did not allow us to investigate the EPR of other materials.

Extensive investigation of the influence of dielectric liquid on both continuous (CW) and pulsed EPRs was performed in [5]. The signal was found to vary with the square of the microwave-field intensity in the sample if the microwave-field variation was connected with the location of the sample in the resonator.

When recording electron-spin-echo (ESE) signals, it is necessary to use microwave-power amplifiers. This is due to the fact that the product of the incident microwave power and ESE pulse duration determines the magnetization rotation ($\pi/2$ or π). The pulses used are selective if their duration exceeds 7–10 ns. For nonselective π pulses employed, e.g., in the Hyperfine Sublevel Correlation Experiment (HYSCORE), high power is required. The ferroelectric material proposed in this work allows one either to do without a microwave-power amplifier or to use an amplifier with a smaller output power.

In this work, we studied the influence of a ferroelectric resonator present in a cylindrical cavity resonator on the intensity of CW (ordinary) and ESE EPR signals.

1. THE INFLUENCE OF THE FERROELECTRIC RESONATOR ON THE CW EPR SIGNAL

Measurements were performed in the 3-cm (*X*) band in a temperature range from 220 to 370 K using the RE-1307 microwave EPR spectrometer with rf modulation of the magnetic field. The ferroelectric resonator was located at the center of a TE_{011} cylindrical cavity resonator. Two ferroelectric resonators were made of sin-

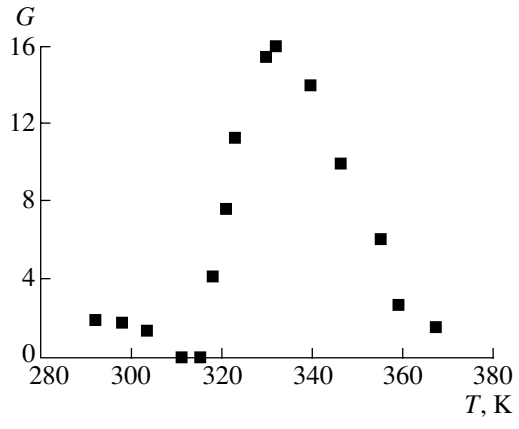


Fig. 1. Gain in the signal-to-noise ratio *G* as a function of temperature; the sample is located in ferroelectric resonator I.

1063-7842/00/4502-0263$20.00 © 2000 MAIK "Nauka/Interperiodica"

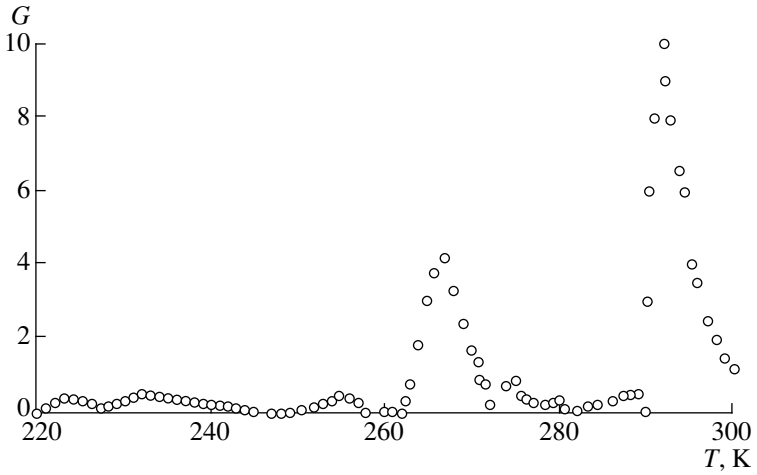


Fig. 2. Gain in the signal-to-noise ratio G as a function of temperature; the sample is located in ferroelectric resonator II.

gle-crystal potassium tantalate ($KTaO_3$), representing tetrahedral prisms measuring $2.85 \times 2.6 \times 3.4$ mm (I) and $2.75 \times 3.5 \times 4.6$ mm (II), respectively. Potassium tantalate was chosen as the ferroelectric-resonator material due to its unique properties; in particular, it is the only material whose loss decreases with an increase in its permittivity [6]. Through the center of the prism, holes were drilled with the radius $R = 0.9$ mm and depth $h = 2$ and 4 mm for resonators I and II, respectively. Into the hole, a sample, diphenil picril hydrazine [DPPH], was placed enclosed in a quartz ampule 1 mm in diameter. The ferroelectric resonator was located in the cavity resonator so that the axis of the drilled hole coincided with the lines of force of the magnetic microwave-field component.

The presence of the ferroelectric resonator in the cavity resonator improves the signal-to-noise ratio. For example, placing a sample in resonator I results in an increase in the signal-to-noise ratio by a factor of 16 at $T = 331$ K (Fig. 1); and placing the sample in resonator II, by a factor of 10 at 292 K (Fig. 2). As seen from Figs. 1 and 2, in the presence of a ferroelectric resonator, the variation of the signal-to-noise ratio with temperature is pronouncedly nonmonotonic. In Fig. 2, several maxima can be observed that correspond to the oscillation modes excited in the ferroelectric resonator. With variation in temperature, both the resonant magnetic field (a jump process) and the resonator quality factor Q vary. A drop in Q corresponds to a zero of the temperature-dependent gain in the signal-to-noise ratio. Note that when a DPPH sample was placed outside the ferroelectric resonator (but close to it), no increase in sensitivity had been observed.

2. THE INFLUENCE OF THE FERROELECTRIC RESONATOR ON THE SPIN ECHO

Measurements were performed using a pulsed microwave spectrometer (a modified *Varian* spectrometer) with a solid-state preamplifier (100 mW) and a power amplifier (1 kW) operating in the 3-cm wavelength band. A loop–gap resonator was used. A ferroelectric resonator in the shape of a trihedral prism that was assembled from three $KTaO_3$ plates, each one measuring $0.5 \times 1.8 \times 4.0$ mm, was placed in this resonator at a temperature of 50 K. A sample (carbon) was placed in the gap. With the help of an attenuator, a 31-dB microwave-power attenuation was introduced instead of the typical 14-dB attenuation in order to avoid saturation when the sample was in the ferroelectric resonator. This corresponds to a decrease in the power exerted on the resonator by a factor of 50. The latter follows from the formula

$$z = 10\log(P_0/P_n). \tag{1}$$

Here z is the attenuator reading, P_0 is the power corresponding to the turned-out attenuator ($z = 0$), P_n is the power corresponding to the turned-in attenuator. For measurements made with the ferroelectric resonator,

$$z_{31} = 10\log(P_0/P_{KTaO_3}) \tag{2}$$

and

$$10^{3.1} = P_0/P_{KTaO_3} \tag{2a}$$

without the resonator, we have

$$z_{14} = 10\log(P_0/P), \tag{3}$$

and

$$10^{1.4} = P_0/P. \tag{3a}$$

Dividing (2a) by (3a), we find that the power required for recording the signal is reduced by a factor of 50.

3. INTERPRETATION OF THE RESULTS OBTAINED

A method is proposed for using the calculations of a regular-shaped dielectric resonator to select the shape and estimate the dimensions of a ferroelectric resonator, with the aim of using the resonator in EPR measurements. In essence, the method consists in determining the dimensions of a solid ferroelectric resonator that has the same resonant frequency as a ferroelectric resonator made of the same material with a hole to receive a sample.

(a) *Calculation of the dimensions of a rectangular-parallelepiped ferroelectric resonator aimed at increasing the signal-to-noise ratio*. Let us calculate the resonance frequency of the above-described ferroelectric resonators I and II cut from a potassium tantalate single crystal. Then, let us mentally replace the cylindrical volume $V_c = \pi R^2 h$ drilled out of the ferroelectric resonator with a volume of the same value but of rectangular-parallelepiped shape $V_p = d^2 h$ and decrease the resonator length and width by $d = \sqrt{\pi r^2}$, where $r^2 = R^2 h/L$ (L is the resonator length). As a result, the dimensions of the ferroelectric resonator I will be $1.63 \times 1.38 \times 3.4$ mm instead of $2.85 \times 2.6 \times 3.4$ mm and the dimensions of resonator II will be $1.26 \times 2.01 \times 4.6$ mm instead of $2.75 \times 3.5 \times 4.6$ mm.

When a dielectric rectangular-parallelepiped resonator is present in the cavity resonator, the resonant frequency of the latter is calculated from the following system of equations [7]:

$$\begin{aligned} f &= (\beta_x^2 + \beta_y^2 + \beta_z^2)^{1/2} c/(2\pi\varepsilon^{1/2}), \\ \beta_z \tan(L\beta_z/2) &= (\beta_x^2 + \beta_y^2 + \beta_0^2)^{1/2}, \end{aligned} \tag{4}$$

where $\beta_x = m\pi/A$; $\beta_y = n\pi/B$; $\beta_z = \pi\delta/L$; $\beta_0 = 2\pi f/c$; A, B, and L are the resonator width, height and length, respectively; $m = n = 1$ are integers corresponding to the $H_{11\delta}$ mode; δ is a part of a half-wave in the resonator measured along its height; c is the speed of light; and ε is the resonator permittivity.

The calculation has shown that the frequencies are equal at the dimensions $1.69 \times 1.439 \times 3.4$ mm and $\delta = 0.833$ for resonator I ($\nu = 9150$ MHz) and at the dimensions $1.2 \times 1.958 \times 4.6$ mm and $\delta = 0.877$ for resonator II ($\nu = 9124$ MHz). It is precisely at these dimensions that the increase in the EPR signal is maximal.

The calculation results show that it is possible to use the proposed method in estimating the dimensions of the ferroelectric resonator to be used in the EPR measurements.

(b) *The influence of a spherical ferroelectric resonator on the intensity of the EPR signal*. An increase in the intensity of the EPR signal can be described by expression [A4] from [5], which illustrates the variation of the magnetic-field intensity at the sample placed inside a spherical ferroelectric resonator

$$B_{ir}/B = Cj_1(kr)\cos\theta/r. \tag{5}$$

Here B_{ir} is the magnetic-field intensity in the presence of a ferroelectric resonator inside the cavity resonator, B is the magnetic-field intensity in the absence of a ferromagnetic resonator, r is the distance from the center of the ferroelectric resonator to the sample, $j_1(x)$ is the first-order spherical Bessel function, and C is determined from the formula

$$C = (a + D/a^2)/j_1(ka), \tag{6}$$

where $D = a^3[((2\mu + 1)j_1(ka) - kaj_0(ka))/((\mu - 1)j_1(ka) + kaj_0(ka))]$; a is the ferroelectric-resonator radius; $j_0(x)$ is the zero-order spherical Bessel function; μ is the magnetic susceptibility of the sample (we assume $\mu = 1$ as in the case of a nonmagnetic object); the wave number $k = (\mu\mu_0\varepsilon\varepsilon_0\omega^2 - i\sigma\mu\mu_0\omega)^{1/2}$, where ε is the permittivity of the ferroelectric resonator, μ_0 and ε_0 are the vacuum permeability and permittivity, σ is the electric conductivity of the ferroelectric resonator, and ω is the circular frequency of the incident microwave power.

Taking into account that our ferroelectric resonator has a very low conductivity, we assume $\sigma = 0$. Substituting the known values of μ_0, ε_0, and $\omega = 2\pi\nu$ ($\nu = 9124$ MHz for the ferroelectric resonator II) into the expression for k, we obtain $k = 192\sqrt{\varepsilon}\ \mathrm{m}^{-1}$.

We calculated the spherical Bessel functions (of the nth order) using the expansion in series

$$j_n(ka) = \sum_{i=0}^{\infty} \frac{(-1)^i (ka/2)^{n+2i}}{i!\Gamma(n+i+1)}, \tag{7}$$

where $\Gamma(n + i + 1)$ is the gamma function; $\Gamma(n + i + 1) = (n + i)!$.

Taking into account the summands of up to the 12th order, we obtained a good agreement of the calculated values of the Bessel function with the standard tabulated values. Thus, for $j_1(kr)$, we have

$$j_1(kr) = \frac{kr}{2\times 1}\left(1 - \frac{k^2 r^2}{2\times 4} + \ldots\right). \tag{8}$$

It can be seen that, when (8) is substituted into (5), r is canceled.

For the $KTaO_3$ ferroelectric resonator, the permittivity is a complex quantity $\varepsilon = \varepsilon' - i\varepsilon''$. The real part ε' is temperature dependent and has the form $\varepsilon' = 45 + 64000/(T - T_c)$, $T_c = 4$ K [6]. The imaginary part $\varepsilon'' = \varepsilon'\tan\delta$, where $\tan\delta = 0.032$.

The calculation of the B_{ir}/B for $a = 0.77 \times 10^{-3}$ m shows that B_{ir}/B has its maximum at $T = 288$ K. At the maximum, $B_{ir}/B = 70$. This calculation gives grounds to expect that a further increase in the intensity of the

EPR signal may be attained through varying the shape of the ferroelectric resonator.

4. CONCLUSION

The use of a ferroelectric resonator in the pulsed EPR allows one to decrease the microwave-power level necessary for the saturation to occur and to substantially increase the signal-to-noise ratio in the CW EPR. Note that the same increase in the signal-to-noise ratio has been attained using a dielectric sapphire resonator [8] through the optimization of the resonator filling factor. In our case, there is still enough room for further increase in the signal-to-noise ratio by increasing the resonator filling factor.

ACKNOWLEDGMENTS

This work was supported in part by the Foundation for Basic Research of the Ministry of Science and Technology of Ukraine, project no. 2.4/516.

REFERENCES

1. P. Hedvig, Acta Phys. Hung. **10** (1), 115 (1959).
2. D. L. Carter and A. Okaya, Phys. Rev. **118**, 1485 (1960).
3. A. Okaya and L. F. Barash, Proc. IRE **50**, 2081 (1962).
4. H. Y. Yee, IEEE Trans. Microwave Theory Tech. **13**, 256 (1965).
5. M. Sueki, G. A. Rinard, S. S. Eaton, *et al.*, J. Magn. Reson. Ser. A **118**, 173 (1996).
6. I. M. Buzin, I. V. Ivanov, N. N. Moiseev, *et al.*, Fiz. Tverd. Tela **22**, 2057 (1980).
7. *Dielectric Resonators*, Ed. by M. E. Il'chenko (Radio i Svyaz', Moscow, 1989).
8. R. Biehl, Bruker Report, No. 1, 45 (1986).

Translated by I.N. Nikishin

Technical Physics, Vol. 45, No. 2, 2000, pp. 267–271. Translated from Zhurnal Tekhnicheskoĭ Fiziki, Vol. 70, No. 2, 2000, pp. 125–129.

EXPERIMENTAL INSTRUMENTS AND TECHNIQUES

Reconstruction of the Form of Complex Motion of an Object from the Autodyne Detection Signal of a Semiconductor Laser

D. A. Usanov, A. V. Skripal', and M. Yu. Kalinkin

Saratov State University, Saratov, 410026 Russia

Received November 3, 1998

Abstract—The possibility of reconstructing the shape of an interference signal on the basis of the autodyne detection signal of a semiconductor laser is demonstrated in cases of harmonic and nonharmonic vibrations of the object of study. The shape of the interference signal coincides with the shape of the signal of an interference system decoupled from the radiation source. It is shown that the form of nonharmonic vibrations can be reconstructed from the values of the autodyne signal function at four varying distances from the external vibrating reflector. Numerical simulation of the procedure of reconstruction of the shape of the interference signal was performed. The main limitations of the method under consideration are discussed. © *2000 MAIK "Nauka/Interperiodica".*

INTRODUCTION

Reconstruction of the form of the complex motion of an object from the signal spectrum is a classical problem of vibration theory [1, 2]. For an optical homodyne interference system decoupled from the radiation source, this problem can be solved by simultaneous expansion of the detected signal into Fourier and Bessel series [3]. Use of autodyne measuring systems stimulated the development of the homodyne interferometry as a method of analysis of the parameters of motion of an object. Optical autodyne systems based on the effect of the autodyne detection in semiconductor lasers combine functions of a generator and detector of electromagnetic waves [4–6]. However, in contrast to interference systems decoupled from the radiation source, these systems do not necessarily allow the conventional methods of reconstruction of the form of motion of an object to be applied [7–10]. This is caused by the fact that the shape of the signal at the output of a photodetector integrated in the semiconductor laser in an autodyne system may differ significantly from the shape of the interference signal in a system decoupled from the source of coherent radiation [11].

EXPERIMENTAL STUDY OF THE AUTODYNE SIGNAL SHAPE

The effect of the parameters of the external optical feedback on the shape of the autodyne signal of a semiconductor laser was studied in the preceding work [11]. The diagram of the experimental setup is shown in Fig. 1. Radiation of an ILPN-206 semiconductor laser (*1*) is directed to an object (*3*) attached to piezoelectric ceramics (*4*). The laser radiation is stabilized with a current source (*2*). A generator of acoustic vibrations (*5*) excites vibrations in the piezoelectric ceramics. A micrometric gear (*6*) moves the vibrating object relative to the plane of the output end of the laser diode. The radiation reflected by the object is partly fed back into the semiconductor laser cavity. Changes in the output power of the laser are detected with a photodetector (*7*). The output signal of the photodetector is applied via an amplifier (*8*) to an analog-to-digital converter (*9*) of a computer (*10*).

To study the effect of the external optical feedback on the shape of the signal at the output of the integrated photodetector of the semiconductor laser, the laser radiation reflector was set in harmonic motion and the distance to the external reflector was varied. The dependence of the normalized power of the autodyne signal P on time t normalized to the period T for various values of the steady-state phase incursion was measured (Fig. 2). Various values of the steady-state phase incursion correspond to various distances to the external reflector: $\varphi_0 = 1.5\pi$ (Fig. 2a) and 0.5π (Fig. 2b). It can be seen from Fig. 2a that the shape of the autodyne signal for $\varphi_0 = 1.5\pi$ is close to sinusoidal, whereas the shape of the signal shown in Fig. 2b differs substantially from the sinusoidal oscillations of the autodyne signal power. The causes of such a difference are discussed in [11]. It was shown in [11] that, at certain values of the steady-state phase incursion and given feed-

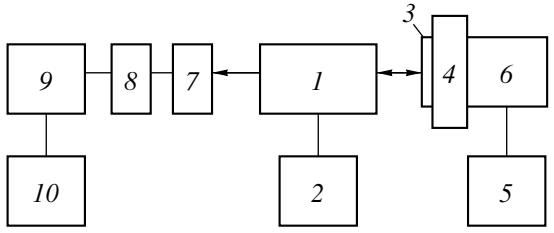


Fig. 1. Schematic diagram of autodyne detection of vibrations of an object.

1063-7842/00/4502-0267$20.00 © 2000 MAIK "Nauka/Interperiodica"

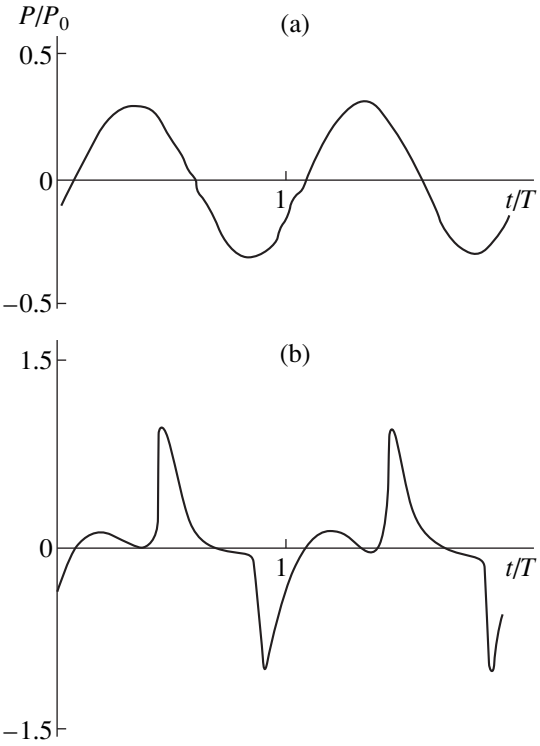


Fig. 2. Experimentally measured curves of the time dependence of the autodyne signal. P_0 is the maximum power of the autodyne signal within the entire range of steady-state phase incursions (0–2π).

back level, the shape of the autodyne signal is significantly distorted in comparison with the signal in the decoupled homodyne interference system. This effect is observed only if the amplitude of the reflector vibration exceeds a certain threshold. The distortions are caused by the laser radiation self-detection nonlinearity. Thus, if the effect of the external optical feedback on the shape of the detected signal is disregarded, the harmonic vibrations of the object can be misinterpreted as complex mechanical vibrations.

In the general case, the shape of the autodyne signal depends on the feedback level, the line broadening coefficient of the semiconductor laser radiation, the amplitude of the object vibrations, and the steady-state phase incursion. Thus, the shape of the autodyne signal can be rather intricate [6, 11–13]. Therefore, the procedure of reconstruction of the form of the object motion should be carried out in two stages. At the first stage, the shape of the detected signal coinciding with the shape of the signal of the homodyne interference system decoupled from the radiation source is reconstructed from the autodyne detection signal. This allows the conventional methods of reconstruction of the form of mechanical motion of an object [3, 7] to be used at the second stage of the procedure. These methods provide reconstruction of both sinusoidal and nonharmonic motion of the object from inteferograms obtained using decoupled homodyne systems.

In this work, we demonstrate the possibility of reconstructing the shape of the interference signal of harmonic and nonharmonic vibrations of an object in a homodyne system decoupled from the radiation source. The reconstruction was performed from the autodyne signals detected with an integrated photodetector of the semiconductor laser.

THEORETICAL ANALYSIS

The composite resonator model described in [14] is used. The steady-state values of the radiation power P at the inner face of the diode, the difference in the charge carrier density in the active region $(N - N_{th})$, and the difference in the phase incursion $(\varphi - \varphi_0)$ between a laser with an external reflector and a solitary laser are given by

$$P = \frac{J - N/\tau_s}{G_N(N - N_0)}, \tag{1}$$

$$G_N(N - N_{th}) = -2(k/\tau_{in})\cos(\varphi), \tag{2}$$

$$\varphi_0 = \varphi + X\sin(\varphi + \psi), \tag{3}$$

where $\varphi = \omega\tau$ is the phase incursion for a laser with feedback; ω is the laser radiation frequency; τ is the external-cavity round-trip travel time; $\varphi_0 = \omega_0\tau$ is the phase incursion for a solitary laser; ω_0 is the frequency of radiation of a single-mode laser diode; $X = (\tau/\tau_{in})k(1 + \alpha^2)^{1/2}$, where $k = (1 - R)\sqrt{r/R}$ is the coefficient of the external optical feedback; r and R are the power reflection factors for external and laser mirrors, respectively; $\psi = \arctan(\alpha)$, where α is the lasing line broadening coefficient depending on the laser type; τ_{in} is the diode-cavity round-trip travel time; $G_N = (\partial G/\partial N)(N_0)$, where G is the active region gain; N_0 is the charge carrier density for which $G(N_0) = 0$; J is the injection current; and τ_s is the carrier lifetime.

We shall restrict our consideration to the feedback level $X < 1$. At this level, the laser operates in the single-mode external cavity regime [12, 13]. In this case, equation (3) is uniquely solved for the phase incursion φ.

Analysis of (1) and (2) shows that, in the feedback range under consideration, the power P depends on the phase incursion $\varphi(t)$ in the external cavity as $\cos(\varphi)$:

$$P = P_1 I + P_2, \tag{4}$$

where

$$I = \cos[\varphi(\varphi_0(t))],$$
$$P_1 = 2k(J\tau_s - N_0)/(G_N^2\tau_s\tau_{in}(N_{th} - N_0)^2), \tag{5}$$
$$P_2 = (J - N_{th}/\tau_s)/(G_N(N_{th} - N_0)).$$

Thus, the absolute value of the ratio between dI/dt

and sin(φ) is

$$\frac{|dI/dt|}{|\sin(\varphi)|} = \left|\frac{d\varphi}{dt}\right|. \quad (6)$$

Taking into account that $|\sin(\varphi)| = \sqrt{1-\cos^2(\varphi)} = \sqrt{1-I^2}$ and

$$\frac{d\varphi}{dt} = \frac{d\varphi}{d\varphi_0} - \frac{d\varphi_0}{dt},$$

we rewrite (6) as

$$\frac{|dI/dt|}{\sqrt{1-I^2}} = \left|\frac{d\varphi}{d\varphi_0} - \frac{d\varphi_0}{dt}\right|. \quad (7)$$

On differentiating equation (3) with respect to φ_0, we obtain

$$\frac{d\varphi}{d\varphi_0} = (1 + X\cos(\varphi + \psi))^{-1}. \quad (8)$$

As can be seen from (8), $d\varphi/d\varphi_0$ is non-negative at the feedback level $X < 1$. Therefore, equation (7) can be written as

$$\frac{|dI/dt|}{\sqrt{1-I^2}} = \frac{d}{d\varphi_0}\varphi(\varphi_0)\left|\frac{d\varphi_0}{dt}\right|. \quad (9)$$

We now consider an autodyne signal obtained by reflection of laser radiation with another phase from the same object:

$$I_\Delta(\varphi_0) = I(\varphi(\varphi_0 + \Delta\varphi_0)).$$

Similarly to (9), the following equation can be obtained:

$$\frac{|dI_\Delta/dt|}{\sqrt{1-I_\Delta^2}} = \frac{d}{d\varphi_0}\varphi(\varphi_0 + \Delta\varphi_0)\left|\frac{d\varphi_0}{dt}\right|. \quad (10)$$

As can be seen from (9) and (10), both autodyne signals depend on the absolute value of $|d\varphi_0/dt|$ in the same manner. We introduce the function z

$$z = \frac{dI/dt}{\sqrt{1-I^2}}\frac{\sqrt{1-I_\Delta^2}}{|dI_\Delta/dt|}, \quad (11)$$

which is independent of $d\varphi_0/dt$, so that

$$z = \frac{d}{d\varphi_0}\varphi(\varphi_0)/\frac{d}{d\varphi_0}\varphi(\varphi_0 + \Delta\varphi_0). \quad (12)$$

Substituting (8) into (12), we have

$$z(\varphi_0) = \frac{1 + X\cos(\varphi(\varphi_0 + \Delta\varphi_0) + \psi)}{1 + X\cos(\varphi(\varphi_0) + \psi)}, \quad (13)$$

where $\varphi(\varphi_0)$ and $\varphi(\varphi_0 + \Delta\varphi_0)$ are determined from (3).

Consider equation (13) for $\Delta\varphi_0 = \pi$. In the approximation of $X \ll 1$, equation (3) yields

$$\varphi = \varphi_0 - X\sin(\varphi_0 + \psi), \quad (14)$$

and equation (13) takes the form

$$z = \frac{1 - XI_0 - X^2(1 - I_0^2)}{1 + XI_0 - X^2(1 - I_0^2)}, \quad (15)$$

where

$$I_0 = \cos(\varphi_0 + \psi). \quad (16)$$

Calculations show that only one of the roots of quadratic equation (15) satisfies the condition $-1 \le I_0 \le 1$. Therefore, the amplitude I_0 of the signal of the interference system decoupled from the radiation source can be uniquely determined from the value of z obtained by substituting the measured amplitudes of two autodyne signals I and I_Δ into (11). Numerical simulation shows that there is a one-to-one relationship between I_0 and z not only for $X \ll 1$, but also within the $0 < X < 1$ range.

Parameters X and ψ necessary for calculations can be determined experimentally. If the autodyne signal phase $\varphi = \pi n - \psi$, where n is an integer, equation (3) yields $\varphi = \varphi_0 = \varphi_{0n} = \pi n - \psi$, whereas equation (13) takes the form

$$z(\varphi_{0n}) = \frac{1 + X}{1 - X}$$

for even n and

$$z(\varphi_{0n}) = \frac{1 - X}{1 + X}$$

for odd n.

It can be seen that, at these points, z attains its maximum and minimum values. Therefore, X can be determined from the maximum or minimum value of z:

$$X = (z_{\max} - 1)/(z_{\max} + 1) \quad \text{or} \quad X = (1 - z_{\min})/(1 + z_{\min}). \quad (17)$$

Parameter ψ can be calculated from the amplitudes of signals $I(\varphi(\varphi_0))$ and $I_\Delta(\varphi(\varphi_0))$ for $\varphi = \varphi_0 = \varphi_{0n} = \pi n - \psi$:

$$\psi = \arccos\left(\frac{|I(\varphi_{0n}) - I_\Delta(\varphi_{0n})|}{2}\right). \quad (18)$$

METHOD OF SIGNAL-SHAPE RECONSTRUCTION

Reconstruction of the shape of an interference signal can be performed by the following algorithm. Parameters X and ψ are determined from the experimentally measured time dependence of the autodyne signal using equations (17) and (18). The values of $z(t_m)$ at the instants of time t_m are determined from the $I(t_m)$

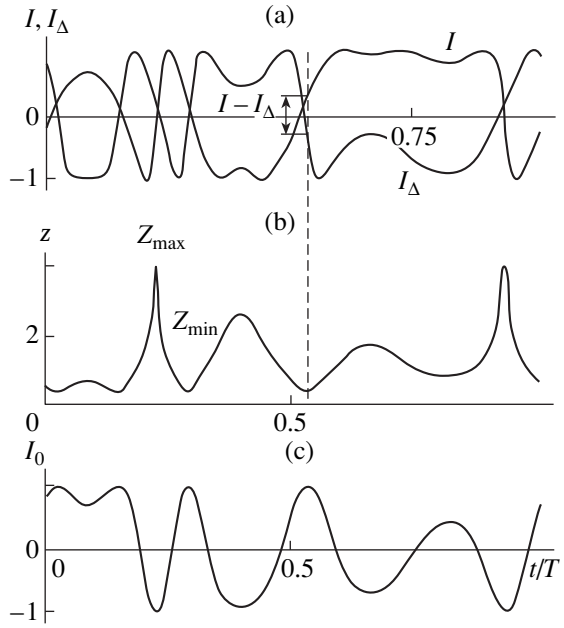


Fig. 3. Signal processing sequence.

and $I_\Delta(t_m)$ samples and the calculated values of $I'(t_m)$ and $I'_\Delta(t_m)$ using equation (11). Then, the amplitude of the signal of the interference system decoupled from the radiation source $I_0(t_m)$ at the instant of time t_m is determined by numerical solution of the set of transcendental equations (3) and (13). The value of $\varphi_0(t_m)$ to be substituted into these equations is determined from (16) as

$$\varphi_0(t_m) = \arccos(I_0(t_m)) - \psi. \quad (19)$$

Numerical simulation is used to test the theoretical conclusions. The time function of phase incursion caused by the reflector motion is determined by its harmonic expansion

$$\varphi_0(t) = \sum_{m=1}^{\infty} c_m \cos(m\Omega t + \varepsilon_m),$$

with coefficients $c_1 = 0.5$, $c_2 = 0.4$, $\varepsilon_1 = 1.5$, and $\varepsilon_2 = 0.2$, which corresponds to nonharmonic periodic vibrations of the reflector.

The values of phase incursion in the autodyne system $\varphi(\varphi_0)$ and $\varphi(\varphi_0 + \Delta\varphi_0)$ are determined from (3) for given values of φ_0, $\Delta\varphi_0 = \pi$, $X = 0.6$, and $\alpha = 5$. Simulated autodyne signals of complex motion of an object for the phase shift $\Delta\varphi_0 = \pi$ are shown in Fig. 3a. Simulation was performed using equation (5).

The interference signal of the decoupled system is reconstructed from the obtained autodyne signals. The derivative of $I(t)$ was calculated using direct and inverse Fourier transformation performed by special fast Fourier transform software. The function $I(t)$ is expanded into a series with coefficients $\{a_m\}$, $\{b_m\}$. The derivative of $I(t)$ was then calculated using the following formula:

$$I'(t) = \left(\frac{a_0}{2} + \sum_{m=1}^{\infty} a_m \cos(m\Omega t) + b_m \sin(m\Omega t)\right)' = \sum_{m=1}^{\infty} m b_m \cos(m\Omega t) - m a_m \sin(m\Omega t). \quad (20)$$

The function z (Fig. 3b) was determined from (11) using the calculated values of I, I_Δ, I', and I'_Δ. The time dependence of the amplitude I_0 (16) of the signal of the interference system was determined when equations (13) and (3) had been solved numerically for φ_0 (Fig. 3c).

The reconstructed time function of the signal of the interference system decoupled from the radiation source (Fig. 3c) is used at the second stage of the procedure for reconstructing the amplitude and form of the mechanical motion of the object. However, it is shown in [3] that reconstruction of the form of nonharmonic vibrations requires the interference signal function (16) to be determined for two phase values φ_0. Taking into account that determination of function (16) also requires two values of function (4) to be known, the total number of the autodyne power functions was found to be four. These functions are determined for phase values $\varphi_0 = 0$, $\pi/2$, π, and $3\pi/2$, corresponding to four distances from the external vibrating reflector. The method for reconstruction of the amplitude and form of complex motion of an object from an interferogram coinciding with the decoupled system interferogram and examples of implementation of this method are described in [3].

CONCLUSION

The method described above can be used for $X < 1$, when the autodyne operates in a single-mode regime. As shown in (11), the noise in the signal $\sqrt{1 - I^2}$ or its derivative $|dI/dt|$ introduces errors into calculation of the function z. As can be seen from (1) and (2), if the feedback level k is too low, both the signal amplitude P and the signal-to-noise ratio P/P_n are decreased. The minimum permissible value of the coefficient k and, therefore, of the parameter X is determined by the following condition:

$$P_n \ll 2k[G_N^2 \tau_s \tau_{in}(N_{th} - N_0)]^{-1}. \quad (21)$$

As follows from (8), for the feedback level X close to 1, the derivative $d\varphi/d\varphi_0$ tends to infinity for $\varphi = \varphi_{0n}$. This cannot be attained in experiments, because the parasitic RC circuits are present in the experimental setup. Thus, the upper limit of the experimentally attained feedback level is $X < 0.95$. Condition (21) imposed on the noise level in the derivative implies that

the signal should not contain rapidly varying interference $dP_n/dt \ll X$.

Thus, a method for solution of the inverse problem of reconstruction of the shape of the interference signal from the autodyne signal of a semiconductor laser with an external reflector is proposed. The shape of the interference signal coincides with the shape of the signal of an interference system decoupled from the radiation source. The autodyne signal is obtained for two values of a phase corresponding to two distances from the external vibrating reflector. In a homodyne interference system decoupled from the radiation source, the reconstruction of the form of mechanical motion of an object from the interference signal is performed using conventional methods applicable to homodyne systems. The procedure described above allows the amplitude and form of the mechanical motion of an object in the autodyne system to be calculated by the methods that had initially been developed for interference systems decoupled from the radiation source. This allows the range of the vibration amplitudes reconstructed from the autodyne system interferograms to be extended.

REFERENCES

1. A. A. Andronov, A. A. Vitt, and S. É. Khaĭkin, *Theory of Oscillations* (Moscow, 1981).
2. Yu. I. Neĭmark and P. S. Landa, *Stochastic and Chaotic Oscillations* (Nauka, Moscow, 1987).
3. D. A. Usanov and A. V. Skripal', in *Problems of Optical Physics* (Saratov Univ., Saratov, 1997), pp. 19–23.
4. E. M. Gershenzon, B. N. Tumanov, and B. I. Levit, Izv. Vyssh. Uchebn. Zaved., Radiofiz. **23**, 535 (1980).
5. V. Van Luc, P. G. Eliseev, M. A. Man'ko, *et al.*, in *Injection Lasers and Their Applications* (Nauka, Moscow, 1992), p. 144.
6. P. De Groot, G. M. Gallatian, and S. H. Macomber, Appl. Opt. **27**, 4475 (1988).
7. D. A. Usanov, A. V. Skripal', V. A. Vagarin, *et al.*, Zarub. Radioélektron., No. 6, 43 (1995).
8. V. S. Sudarshanam and K. Srinivasan, Opt. Lett. **14**, 140 (1989).
9. W. Jin, D. Uttamchandani, and B. Culshaw, Appl. Opt. **31**, 7253 (1992).
10. V. A. Vagarin, A. V. Skripal', and D. A. Usanov, Avtometriya, No. 3, 103 (1995).
11. D. A. Usanov, A. V. Skripal', and M. Yu. Kalinkin, Izv. Vyssh. Uchebn. Zaved., Prikl. Nelineĭnaya Dinamika **6**, 3 (1998).
12. B. Tromborg, J. Mork, and V. Velichansky, J. Eur. Opt. Soc. B. Quant. Semiclass. Opt. **9**, 831 (1997).
13. L. Goldberg, H. F. Taylor, A. Dandridge, *et al.*, IEEE J. Quantum Electron. **18**, 555 (1982).
14. B. Tromborg, J. H. Osmundsen, and H. Olesen, IEEE J. Quantum Electron. **20**, 1023 (1984).
15. H. Olesen, J. H. Osmundsen, and B. Tromborg, IEEE J. Quantum Electron. **22**, 762 (1986).

Translated by K.S. Chamorovskiĭ

Technical Physics, Vol. 45, No. 2, 2000, pp. 272–273. Translated from Zhurnal Tekhnicheskoĭ Fiziki, Vol. 70, No. 2, 2000, pp. 130–132.

BRIEF COMMUNICATIONS

Summation of Signals during Energy Extraction from 3-cm Range Oversized Cavities

V. A. Avgustinovich, S. N. Artemenko, V. L. Kaminskiĭ, and Yu. G. Yushkov

Research Institute of Nuclear Physics, Tomsk Polytechnical University, Tomsk, 634050 Russia

Received April 20, 1998

Abstract—Data on the summation of signals from two synchronously operating resonant microwave (3-cm) compressors with oversized cavities are described. The energy was extracted with interference switches. The summation was performed in a waveguide tee and a 3-dB three-slot bridge. For a signal duration of ~30 ns at a level of 0.5, a peak power of ~1 MW, and an efficiency of ~30%, a gain was found to approach ~11 dB. © *2000 MAIK "Nauka/Interperiodica".*

(1) It is known [1] that the energy of signals coming from microwave resonant compressors depends on the energy density accumulated in the cavity and on its volume. An exceedingly high energy density causes cavity breakdown, and a too large volume results in the high-density oscillation spectrum. For this reason, a single compressor sometimes cannot completely utilize the energy carried by pulses to be compressed. Of interest is therefore to look for ways of increasing the energy of compressed pulses, for example, by concurrently accumulating the energy in several cavities with the subsequent summation of their outputs. This was first accomplished in [2], where single-mode 10-cm-range cavities were used. Such an approach may appear to be particularly promising in the short-microwave range. Cavities designed for this range are of relatively small weight and size and thus can serve as a basis for simple and compact sets of compressors to generate high-energy microwave pulses. To the best of our knowledge, data on signal summing in the short-microwave range are lacking.

In this work, experimental results obtained upon summing the outputs of microwave (3-cm-range) compressors with oversized cavities and energy extraction through interference switches are reported.

(2) Two compressors with 210-mm-long cylindrical copper cavities with a diameter of 90 mm were used. Their operating frequency for the $H_{01(12)}$ oscillation mode was 9.42 GHz, and the intrinsic Q-factor, ~10^5. The energy was extracted through interference switches based on rectangular waveguide *H* tees made of conventional waveguides with a cross section of 23×10 mm^2.

The compressors were powered from a magnetron with an output of ~100 kW. The pulse duration was ~1 μs. The magnetron pulse energy was divided with a matched waveguide tee. Changeover of the cavities from the accumulation to the extraction mode was effected through glow-discharge breaks of the interference switches by initiating a discharge in argon or helium under an overpressure of ~1 atm. For comparable field strengths in the breaks, the discharge time was controlled by UV illumination of the spark gap and by adjusting the gas pressure in each of the breaks. As in [2], synchronous illumination of the gaps was provided by simultaneously applying high-voltage pulses from one source to the ignitor electrodes of each of the breaks. Summing was performed in a waveguide *H* tee or in a 3-dB narrow-wall-coupled slotted bridge.

An experimental setup is shown in Fig. 1, where *1* is a magnetron generator; *2*, ferrite valve; *3*, directional couplers; *4*, phase shifters; *5*, sealing windows; *6*, input matched tee (microwave divider); *7*, storing cavities; *8*, interference switches; *9*, breaks; *10*, summing device (tee or 3-dB bridge); *11*, detectors; *12*, matched load; and *13*, driver of the magnetron and the breaks. The setup was equipped with a system for separate gas delivery to the cavities under a controlled overpressure.

(3) Figure 2 shows (a, b) oscillograms at the output of either compressor, (c) the same signals separated in time, (d) the net pulse with the delay time minimized, and (e) the difference signal. The results of summing in the tee and the waveguide bridge are basically the same. The only difference is that the added in the tee is more sensitive to a difference between the electrical lengths of the output microwave channels of the compressors. Most likely, the decay in the input arms of the tee is smaller as compared with that in the bridge. Hence, the interaction between the cavities at the instant of energy extraction is stronger.

The oscillogram of the time-separated pulses coming from both of the compressors (Fig. 2c) shows the mutual time spread due to the statistical nature of discharge evolution. If the time delay between the pulses is on the order of the pulse duration, the spread is 10 ns or more. As the pulses approach each other, the spread diminishes and becomes minimum (several nanoseconds) when the leading edges of the pulses coincide. Long-term stability (within given limits) of the net sig-

1063-7842/00/4502-0272$20.00

nal is provided by keeping the operating pressure difference between the breaks within a range found experimentally.

As follows from the oscillograms, when two relatively long microwave pulses are added, their envelopes transform into the envelope of the net signal and high-frequency fluctuations of the envelopes of the signals are averaged. The former feature makes it possible to correct the net envelope, and the latter smoothes it out. In experiments on summing pulses with heavily distorted envelopes, the net signals decayed nearly exponentially and their envelopes were close to rectangular. The smeared decay of the net pulse is most probably associated with fluctuating decay of either initial pulse due to plasma instabilities in the microwave discharges of the breaks. The oscillograms shown in Fig. 1 were obtained by superimposing ~500 pulses.

For our system, measurements showed that, if the initial pulses are to a certain extent identical, ~90% of their total energy are fed to the load. For the pulses of width about 30 ns at a level of 0.5, the maximum gain and the efficiency of the system were ~11 dB and ~30%, respectively. The rated values were, respectively, 11.5 dB and 35%. The performance of the system is thus lower than that of each of the compressors. This may be caused by incomplete identity of the initial signals, because of which a part of the energy is lost either in the cavities (summation in the tee) or in the difference arm (summation in the bridge). The ambient atmosphere (argon or helium) does not influence the results.

Thus, we may suppose that, with the cavities excited simultaneously, the interference switch provides the summation of the outputs of the microwave compressors in any case. Neither mode interaction in a multimode cavity nor an adiabatic frequency drift during energy extraction is a barrier to the summation. The physical foundation for the summation in our system is the synchronous excitation of the cavities and the coherence of their outputs with generator signals [3]. This means that small and light 3-cm-range compressors can be viewed as a basis for creating systems, for example, of eight or sixteen compressors to form ~50-ns-wide pulses with a peak power of 10–100 MW. Such systems can be powered by magnetrons with an output power of ~1 MW or by higher-power klystrons or magnecons.

(4) Thus, we demonstrated the feasibility of increasing the output power (energy) of resonant microwave (3-cm range) compressors with oversized cavities and energy extraction through interference switches. This is provided by adding signals from several compressors operating in parallel. The process can be accomplished both in a waveguide tee and in a slotted bridge. Our results offer possibilities of raising, in a similar way, the power of shorter-microwave (e.g., mm-wave) compressors provided that the input and output signals of the compressors are coherent, the energy is extracted through interference switches, and the storing cavities are excited simultaneously. The summation of signals

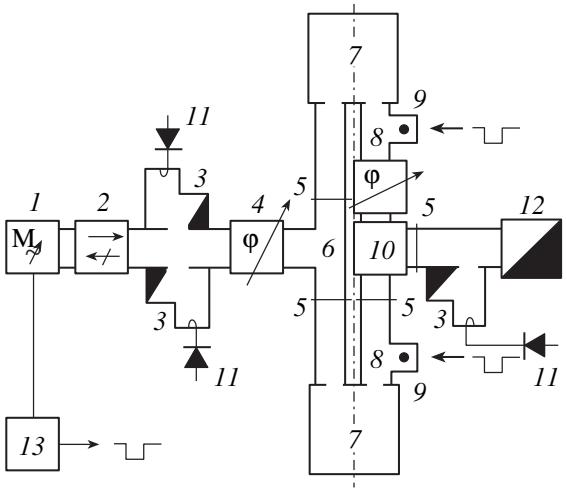


Fig. 1. Experimental setup.

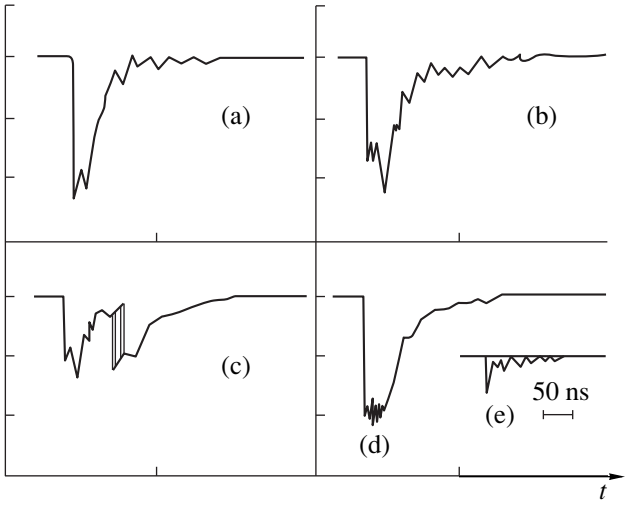


Fig. 2. Oscillograms of the output microwave pulses.

when the energy is extracted from oversized cavities not only increases the power (energy) of generated microwave pulses but also makes possible the correction and smoothing of their envelopes.

ACKNOWLEDGMENTS

This work was supported by the Russian Foundation for Basic Research (grant no. 97-02-16461).

REFERENCES

1. A. N. Didenko and Yu. G. Yuskov, *High-Power Microwave Pulses of Nanosecond Width* [in Russian] (Énergoatomizdat, Moscow, 1984).
2. S. A. Novikov, S. V. Razin, P. Yu. Chumerin, *et al.*, Pis'ma Zh. Tekh. Fiz. **16** (20), 46 (1990).
3. D. Birx, G. J. Dick, W. A. Little, *et al.*, Appl. Phys. Lett. **33**, 466 (1978).

Translated by V.A. Isaakyan

Technical Physics, Vol. 45, No. 2, 2000, pp. 274–276. Translated from Zhurnal Tekhnicheskoĭ Fiziki, Vol. 70, No. 2, 2000, pp. 133–135.

BRIEF COMMUNICATIONS

Electromagnetic Wave Backscattering from a Cloud of Alpha-Emitting Radioactive Elements in Air

V. A. Polyanskiĭ

Naval Institute of Radioelectronics, St. Petersburg, 198903 Russia

Received August 19, 1998

Abstract—A track model is proposed for the electromagnetic wave scattering from a cloud of radioactive elements in air. Relationships are derived for estimating the radar cross section of the cloud of alpha-emitting radioactive elements as a function of the cloud activity and the wavelength of the electromagnetic wave. The activity of a cloud detected by a standard radar is estimated. *© 2000 MAIK "Nauka/Interperiodica".*

INTRODUCTION

The injection of a cloud of radioactive elements into the air is accompanied by their radioactive decay, which causes alpha, beta, and gamma emissions. The ratio of the intensities of these emissions depends on the content of radioactive elements in the cloud, but each of the emissions ionizes the air, thus changing the dielectric properties of the air in the cloud. The latter effect enables one to detect a cloud of radioactive elements in the air, for example, by registering a change of characteristics of electromagnetic waves scattered from or propagating through the cloud. In this paper, the possibility to detect the electromagnetic wave scattered by a cloud of radioactive elements in the air (radar detection) using only the alpha emission is evaluated. The radar detection of continuous atmospheric ionized regions is widely covered in the literature [1, 2]. Being an ensemble of discrete ordered structures called tracks, a cloud of radioactive elements differs from a continuum. In view of this, in this paper, the electromagnetic wave scattering from a cloud of radioactive elements in air is modeled as the scattering from a set of tracks. Using this model, the problem formulated in this work is reduced to analyzing the possibility of radar detection of a set of plasma cylinders.

THE MODEL OF A CLOUD OF ALPHA EMITTING RADIOACTIVE ELEMENTS IN AIR

In the process of alpha emission, radioactive elements emit alpha particles moving in air, which causes the formation of excited and charged particles (pairs of free charges, i.e., electrons and ions) along the way. When electrons receive a sufficiently large energy from an alpha particle, these electrons (called δ-electrons), in turn, also ionize and excite air molecules until their energy becomes lower than a fixed value E. The result is that, in the vicinity of the alpha particle path, a region is formed containing electrons with an energy higher than the thermal energy but lower than energy E necessary for ionization of air. In addition, this region contains both primary and secondary electrons formed, respectively, by alpha particles and δ-electrons exciting the air molecules. The number of the primary electrons is about one-third of the total number of electrons formed in the process of alpha emission [3]. Thus, in the vicinity of the alpha particle path, there appear free (hot) charge carriers, whose energy is lower than the ionization energy but higher than the thermal energy, and bound excited states. Then, the hot carriers are thermalized; i.e., their temperature decreases to the thermal level. The region in the vicinity of the alpha particle path containing primary and secondary charges is called the track of this alpha particle [4]. The period from the moment when the track is formed until the moment of thermalization is considered to be the time of the track existence [4].

Thus, the alpha particle track can be regarded as a cylinder consisting of pairs of primary and secondary free charges (electrons and ions). The radius of the cylinder is determined by the δ-electron path length, which depends on the energy of δ-electrons and the air density. Track radius r is assumed to be equal to $0.84R$, where R is the maximum path length of the δ-electron [3]. The length of the cylinder (track) is governed by the alpha particle energy, which lies in the interval of 4–10 MeV [4] and depends on the sort of a radioactive element. Under standard conditions in air, the length of the alpha particle track is in the range of ~3–8 cm [5].

The number N of pairs of free charges (formed during the radioactive emission) per unit of track length is called the specific ionization. Under standard conditions in air, alpha particle tracks are characterized by a value of N equal to 5×10^4–10^5 cm^{-1} [5]. Below, the minimum value of N is used.

As mentioned above, the track radius is governed by the maximum path length of a δ-electron depending on its energy. Energy E_δ of a δ-electron is determined by the expression $E_\delta E_0 \cos^2\varphi$, where E_0 is the maximum energy of the δ-electron, φ is the angle between the

1063-7842/00/4502-0274$20.00 © 2000 MAIK "Nauka/Interperiodica"

directions of the motions of the δ-electron leaving the track and the alpha particle. Value E_0 depends on the alpha particle energy. As the alpha particle energy varies from 4 to 10 MeV, E_0 varies from 2.2 to 5.3 keV [5]. Since the formation of one pair of charges needs an energy of 35 eV, a δ-electron produces 30–70 pairs.

Air molecules are ionized and excited by a δ-electron in a cylinder, whose radius (the capture radius) increases with the δ-electron energy. Supposing that 3–7 particles are ionized in the region with the capture radius [6], one can assume that the air atoms are ionized by a δ-electron after a number of collisions less by the corresponding factor. Then, the δ-electron covers a distance R equal to ten free path lengths. Since, under standard conditions in gases, the electron free path is $\sim 10^{-5}$ cm [4], $R = 10^{-4}$ cm and $r = 0.84 \times 10^{-4}$ cm.

The charge density in the track can be determined by the formula

$$\rho = \frac{N}{\pi r^2}.$$

Hence, in the alpha track, the charge density is not less than 2×10^{12} cm^{-3}. Provided the condition $6.9\pi\sqrt{T/N} < 1$, where T is the air temperature measured in K, holds in the track [4], a plasma is formed in this region. One can see that, under standard conditions in the alpha track, this inequality is valid; i.e., plasma is formed in the track.

Thus, the alpha track is a plasma cylinder several centimeters (3–8 cm) in length with a radius of $\sim 10^{-4}$ cm. The spatial orientation of this cylinder is supposed as random.

As mentioned above, the track lifetime equals the electron thermalization time. On the average, this time is 4×10^{-7} s in gases [4]. Suppose that, during this time, the charge density in the track is constant and equals 2×10^{12} cm^{-3}. Thus, the cloud of alpha emitting radioactive elements is an ensemble of plasma cylinders with limited lifetime which are chaotically oriented in space.

ESTIMATION OF THE RADAR CROSS SECTION OF THE CLOUD OF RADIOACTIVE ELEMENTS IN AIR BY THEIR ALPHA EMISSION

First, consider the radar cross section (RCS) of one track as the RCS of a plasma cylinder. Determining the characteristics of the electromagnetic wave scattered by a plasma cylinder is substantially simplified when this cylinder is supposed to be conducting [7]. As usual, assume that the cylinder is a conductor illuminated by a harmonic field with frequency f if the inequality $\alpha = \gamma/2\pi f\varepsilon \gg 1$ is valid, where γ and ε are, respectively, the conductivity and permittivity of the cylinder. Taking into account particle collisions in plasma yields after some algebra

$$\alpha = \frac{\omega_p^2 \nu}{\omega(\omega^2 + \nu^2 - \omega_p^2)},$$

where $\omega_p = 2\pi\sqrt{81N}$ is the plasma frequency, ν is the average number of electron collisions with neutral molecules and ions per unit time, and $\omega = 2\pi f$.

When determining ν for the alpha track, only electron collisions with neutral molecules should be taken into account because the density of these molecules $N_0 = 2.7 \times 10^{19}$ cm^{-3} is much greater than the density of ions $N_+ \approx 2 \times 10^{12}$ cm^{-3}. Then, taking into consideration only elastic electron scattering, one can find ν [1]

$$\nu = 8.3 \times 10^5 \pi a^2 \sqrt{T} N_0,$$

where a is the radius (cm) of a molecule and T is the air temperature (K).

Assuming that, under normal conditions, $a \approx 10^{-8}$ cm [4] yields $\nu \approx 10^{11}$ s^{-1}. In this case, $\omega_p \approx 8 \times 10^{10}$ s^{-1}. Taking $f = 3 \times 10^9$ Hz, one obtains $\alpha \approx 15 \gg 1$; i.e., when $f \le 3 \times 10^9$ Hz, the track (the plasma cylinder) can be considered a conductor. Therefore, below, in the case when $f \le 3 \times 10^9$ Hz, the RCS of the track is estimated as the RCS of a conducting cylinder of radius r and length L ($r \ll \lambda$, where λ is the wavelength of the illuminating electromagnetic wave, L is comparable with λ, and $r \ll L$). Then, in the case of backscattering when receiving and transmitting antennas are identically polarized, the average RCS σ_1 of a randomly oriented conducting cylinder is about $(0.02–0.2)\lambda^2$ depending on L [8]. For example, the minimum value of σ_1 at $\lambda = 10$ cm is $\sigma_1 = 2$ cm^2 $= 2 \times 10^{-4}$ m^2.

When the alpha emission activity of the cloud of radioactive elements is Q (Ci), $3.7 \times 10^{10} Q$ tracks are formed per second. In this case, the cloud permanently contains, on the average, $\sim 1.5 \times 10^4 Q$ tracks if the track lifetime is taken into account. Then, the RCS σ of the cloud of radioactive elements is $\sigma = 1.5 \times 10^4 Q\sigma_1$ for alpha emission activity Q[Ci]. For example, when the alpha activity of the cloud is 10^{-3} Ci, its RCS is about 3×10^{-3} m^2 at $\lambda = 10$ cm. A radar target with this RCS can by detected by a standard radar (it is supposed that the body of the injection does not exceed the space of the radar resolution cell) [9].

CONCLUSION

In this paper, a model is proposed for the electromagnetic wave scattering from a cloud of alpha-emitting radioactive elements. Using this model, the RCS of the cloud of radioactive elements is estimated, and it is demonstrated that, in practice, this cloud can be detected by a standard radar by alpha emission with the activity of, at least, thousandths of a curie.

REFERENCES

1. V. L. Ginzburg, *Propagation of Electromagnetic Waves in Plasma* (Nauka, Moscow, 1967).
2. K. A. Boyarchuk, G. A. Lyakhov, and M. S. Suyazov, Zh. Tekh. Fiz. **67**, 76 (1997) [Tech. Phys. **42**, 190 (1997)].
3. V. I. Kalashnikova and M. S. Kozodoev, *Detectors of Elementary Particles* (Nauka, Moscow, 1966).
4. V. K. Lyapidevskiĭ, *Methods of Radiation Detection* (Énergoatomizdat, Moscow, 1987).
5. S. Glasstone, *Sourcebook on Atomic Energy*, 2nd ed. (Van Nostrand, Princeton, New York, 1958; Inostrannaya Li-teratura, Moscow, 1961).
6. M. N. Medvedev, *Scintillation Detectors* (Atomizdat, Moscow, 1977).
7. *Radar Cross Section Handbook*, Ed. by G. T. Ruck, 2 vols. (Plenum, New York, 1970).
8. J. H. Van Vleck, F. Blanch, and M. Hamermech, J. Appl. Phys. **18**, 274 (1947).
9. *Radar Device*, Ed. by V. V. Grigorin-Ryabov (Sovetskoe Radio, Moscow, 1970).

Translated by I.G. Efimova

Technical Physics, Vol. 45, No. 2, 2000, pp. 277–280. Translated from Zhurnal Tekhnicheskoĭ Fiziki, Vol. 70, No. 2, 2000, pp. 136–138.

BRIEF COMMUNICATIONS

On Amplification of the Spin Wave Envelope Solitons in Ferromagnetic Films

M. P. Kostylev and B. A. Kalinikos

St. Petersburg State Electrotechnical University, St. Petersburg, 197376 Russia

Received October 19, 1998

Abstract—The process of parametric amplification of the spin wave envelope solitons in ferromagnetic films with the use of parallel magnetic pumping was theoretically studied. Solutions obtained by numerical methods show that the value of the amplification coefficient depends on the relationship between the initial phases of the signal and the pumping carrier waves. © *2000 MAIK "Nauka/Interperiodica".*

INTRODUCTION

Experimental observations of the parametric amplification of solitons of the envelope of reverse volume spin waves by spatially inhomogeneous magnetic pumping in ferromagnetic films were recently reported in [1–3]. Theoretical estimates of the efficiency of amplification were obtained in [1, 2]. The analysis was performed on the basis of equations for the envelope of the interacting packets of spin waves [4]. Unfortunately, the existence of a gap in the spectrum of spin waves of the film does not allow the spectral characteristics of the parametric process to be correctly taken into account. As a result it is difficult to consider concurrent propagation of the co- and counter-directed idle pulses. That is why the codirected idle packet was neglected in [4] and the counterpropagating one was neglected in [1].

The purpose of this work was to obtain equations in the spectral domain that would strictly describe the parametric process of interaction of a nonlinear packet of spin waves excited by microwave current of a transducer (antenna) with a spatially inhomogeneous magnetic pumping field generated by another (external) microwave current. These equations are analyzed with the use of numerical methods. Thus, the peculiarities of excitation of the spatially bounded packets of spin waves by microstrip transducers and the peculiarities of configuration of the microwave magnetic field generated by a microstrip pumping electrode are taken into account in the model calculations for the first time. The spectral approach makes it possible to take into account the existence of a gap in the spectrum and the presence of both idle pulses. Special attention is paid to the influence of the initial phase difference of the signal and pumping carrier waves ϕ on the efficiency of amplification.

MODEL EQUATIONS

Consider a model of the spin wave delay line (Fig. 1a) based on a ferromagnetic film of thickness L. The input and output microstrip antennas with width W are situated on the surface of the film at a distance l from each other. The film is magnetized to the saturation level by a constant field with the strength H applied in the plane of the film perpendicularly to the longitudinal axis of antennas.

A microwave current in the input antenna is $j(z, t) = J(z, t)\cos(\omega_0 t)$, where $J(t, z)$ is the envelope of the linear current density and ω_0 is the carrier frequency. Assume that the current is uniformly distributed over the length of the strip. A magnetic field of this current excites reverse volume spin waves in the film, which propagate as wave packets in both directions from the antenna along the direction of the constant magnetic field. An additional microstrip electrode with the thickness W_p is located between the input and output transducers at the distance l_p from the input antenna. A microwave current in this electrode $j_p(z, t) = J_p(z, t)\cos(\omega_p t)$ generates spatially inhomogeneous magnetic pumping field $\mathbf{h}_p(t, z)$ inside the film. Here, ω_p is the carrier frequency of the pumping current, which is approximately two times the carrier frequency of the input signal:

$$\omega_p/2 = \omega_0 + \Delta\omega. \quad (1)$$

The analysis of propagation of a pulse of spin waves in such a structure can be performed in two stages: (i) the analysis of excitation of a pulse of spin waves by the microwave current of the strip transducer and (ii) the analysis of interaction of the pulse of spin waves with a pulse of the spatially inhomogeneous magnetic pumping field. In solving these problems, we have used the results presented in [5–8]. According to the approach proposed in these works, proceeding from an

1063-7842/00/4502-0277$20.00

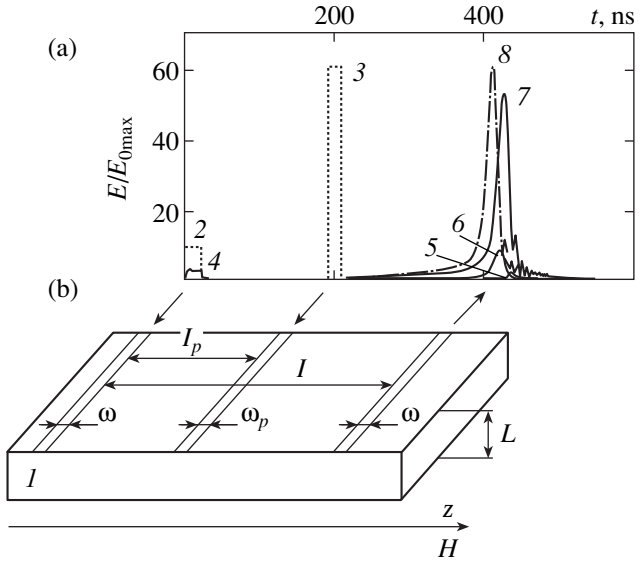


Fig. 1. An example of calculations: (a) input and output signals and (b) the model structure: (*1*) ferromagnetic film; (*2*) input current pulse (the vertical scale is arbitrary); (*3*) pumping current pulse (the vertical scale is arbitrary); (*4*) envelope of the pulse of linear electromotive force induced by variable magnetization in the input antenna generated by the input current pulse (the value is averaged over the antenna width); (*5*) output pulse in the absence of pumping; (*6*) output pulse for a supercritical factor $\nu = 12$ and a pumping pulse duration of $\tau_p = 15$ ns; (*7*) output pulse for a supercritical factor $\nu = 12$ and a pumping pulse duration of $\tau_p = 20$ ns; (*8*) pulse induced in the input antenna by the counterpropagating idle pulse of spin waves for a supercritical factor $\nu = 12$ and a pumping pulse duration of $\tau_p = 20$ ns. The envelopes of the pulses of electromotive force (*4*–*8*) are normalized to the amplitude of the input signal in the absence of pumping $E_{0\max}$.

expression for the energy of a ferromagnetic film, we obtain a system of nonlinear parametric equations for the spatial Fourier harmonics of the alternating magnetization excited by the external microwave current and interacting parametrically with the magnetic field of the longitudinal pumping:

$$\frac{\partial m_k^1}{\partial t} + i\left[\omega_H + \frac{\omega_M}{2}(1 - P_k)\right]m_k^1$$

$$+ i\int_{-\infty}^{\infty} dk_1 \int_{-\infty}^{\infty} dk_2 \int_{-\infty}^{\infty} dk_3 \left[\sum_{p_1, p_2, p_3 = 1, 2} T_{k_1 k_2 k_3 k}^{p_1, p_2, p_3} m_{k_1}^{p_1} m_{k_2}^{p_2} m_{k_3}^{p_3}\right.$$

$$\left.\times \delta(p_1 + p_2 + p_3 - 4)\delta(k_1 + k_2 + k_3 - k)\right] \qquad (2)$$

$$+ i|g|\mu_0 \int_{-\infty}^{\infty} dk_1 \int_{-\infty}^{\infty} dk_2 J_{pk_1}(t)\cos(\omega_p t) Y_{k_1} m_{k_2}^1 \delta(k_1 + k_2 - k)$$

$$= |g|\mu_0 J_{sk}(t)\cos(\omega_0 t)\frac{k}{|k|}Y_k,$$

$$\frac{\partial m_k^2}{\partial t} - i\left[\omega_H + \frac{\omega_M}{2}(1 - P_k)\right]m_k^2$$

$$- i\int_{-\infty}^{\infty} dk_1 \int_{-\infty}^{\infty} dk_2 \int_{-\infty}^{\infty} dk_3 \left[\sum_{p_1, p_2, p_3 = 1, 2} T_{k_1 k_2 k_3 k}^{p_1, p_2, p_3} m_{k_1}^{p_1} m_{k_2}^{p_2} m_{k_3}^{p_3}\right.$$

$$\left.\times \delta(p_1 + p_2 + p_3 - 5)\delta(k_1 + k_2 + k_3 - k)\right] \qquad (3)$$

$$- |g|\mu_0 \int_{-\infty}^{\infty} dk_1 \int_{-\infty}^{\infty} dk_2 J_{k_1}(t)\cos(\omega_p t) Y_{pk_1} m_{k_2}^2 \delta(k_1 + k_2 - k)$$

$$= -|g|\mu_0 J_{sk}(t)\cos(\omega_0 t)\frac{k}{|k|}Y_k.$$

Here, P_k is the matrix element of the dipole–dipole interaction [5]:

$$P_k = 1 - [1 - \exp(-|k|L)]/(|k|L), \qquad (4)$$

Y_k is the overlap integral between the Fourier component of the magnetic field of the current and the "membrane function" for the lowest mode of the reverse volume spin waves [5]:

$$Y_k = 1/\sqrt{(2\pi)^{3/2}}[1 - \exp(-|k|L)]/2L, \qquad (5)$$

$T_{k_1 k_2 k_3 k}^{p_1, p_2, p_3}$ is the coefficient of nonlinear four-wave interaction, $\omega_H = |g|\mu_0 H$, and $\omega_M = |g|\mu_0 M_0$.

The system of equations (2) and (3) was solved numerically. We simulated excitation of a signal packet of spin waves by a rectangular current pulse of the input antenna j_s with the duration τ_s. The packet interacted with a magnetic field generated by a rectangular pulse of current j_p with the duration τ_p applied to the pumping electrode. We also set a certain phase shift ϕ between the carrier waves of the signal pulse and pumping.

RESULTS AND DISCUSSION

The results of our numerical calculations showed that the main peculiarities of the process of parametric amplification can be studied with neglect of the nonlinear four-wave interaction in the spin system of the film ($T_{k_1 k_2 k_3 k}^{p_1, p_2, p_3} = 0$). That is why we restrict the discussion of results to this case. The influence of the four-wave nonlinearity will be discussed in brief at the end of the paper, where the results of the numerical calculation are compared to experimental data.

One of the results of our simulation is presented in Fig. 1b. The initial data for calculations corresponded

to the experimental parameters from [2]. We present the pulses of the input current, pumping current, and linear electromotive force E induced in the output antenna by a pulse of spin waves. The pulses of E are shown for the absence and presence of pumping.

The calculations showed that the shape of the output pulse depends strongly on the amplitude of the pumping pulse. If the amplitude of pumping is lower than a threshold for the parametric generation of spin waves J_{th}, the pumping pulse entering at the moment when the signal packet of spin waves passes under the electrode results in partial compensation of the losses related with propagation of the packet in an absorbing medium. The duration and amplitude of the output pulse appear to be independent on time of termination of the pumping pulse, provided that the pulse duration is longer than the time required for the signal pulse to escape from the zone of pumping. Under this regime, with a sufficiently small duration of the input pulse such as in [2] (22 ns), the compensation of the losses in course of propagation does not exceed a few tenths of dB.

If the amplitude of pumping is higher than the threshold J_{th} of the parametric generation of spin waves by a spatially localized field of the current in microstrip antenna, the amplification occurs in a qualitatively different regime analogous to the regime of mode amplification [9]. Under this regime, the shape and duration of the output pulse are primarily determined by the moment of termination of the pumping pulse and depend weakly on the amplitude and duration of the current pulse in the input antenna (Fig. 1b). In the case of infinite duration of pumping, the system exhibits the transition to the parametric generation of spin waves and the signal amplitude increases infinitely with time.

Figure 2 shows a plot of the coefficient of amplification K (calculated as the ratio of amplitudes of the output signals with and without pumping) versus supercritical factor $\nu = (J_p/J_{th})^2$. The measurements of K [1, 2] revealed a strong influence of the initial phases of the signal and pumping carrier waves on the efficiency of amplification. That is why in simulations we paid special attention to the influence of the parameter ϕ on the behavior of $K(\nu)$.

The calculations showed that the dependence of K on ϕ is explained by interference of the signal pulse and the idle pulse propagating in the same direction in the output transducer. An example of the calculated dependence of K on ϕ for the case of pumping exceeding the threshold value J_{th} is shown in the inset to Fig. 2. In this particular case, a change in the initial phase difference by 90° results in the change of the K value by 15 dB.

Figure 2 shows maximum and minimum possible values of K for the corresponding values of ϕ. The open circles correspond to the experimental data from [2] (only the mean value of the amplification coefficient was measured in that work). There is a good agreement between the results of calculations and experimental data in the initial part of the curve. A saturation of the amplification coefficient is observed in experiment for large amplitudes of pumping. Theoretical analysis suggests that the saturation can be related to the four-wave processes of nonlinear interactions of the volume reverse spin waves, such as the self-action and the scattering on short spin waves.

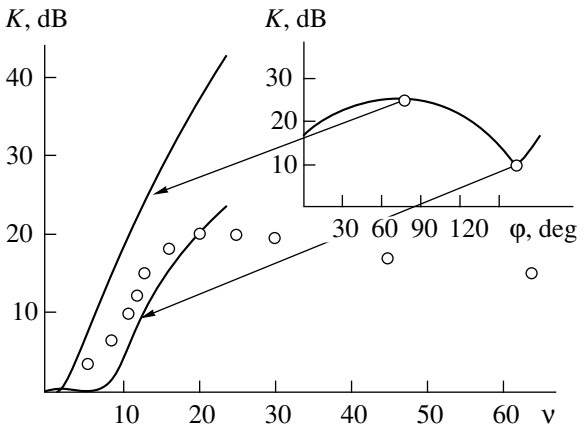


Fig. 2. The plot of amplification coefficient K versus supercritical factor ν. Solid lines correspond to the results of calculations and circles correspond to the experimental data from [2]. Maximum and minimum calculated amplification coefficients are shown. Inset gives an example of the calculated dependence of K on the initial phase difference of the pumping and signal carrier waves: the width of the input and output antennas is $\omega = 50$ μm, the width of the pumping strip is $\omega_p = 0.25$ mm, the distance between the input and output antennas is $l = 8.5$ mm, the ferromagnetic film thickness is $L = 4.9$ μm, the parameter of magnetic losses is $2\Delta H = 0.5$ Oe, the input pulse duration is $\tau = 22$ ns, the pumping pulse duration is $\tau_p = 15$ ns, the pumping pulse is delayed by 193 ns relative to the leading edge of the input pulse, the detuning of the carrier frequency of the signal from the half of the carrier frequency of the pumping pulse is $\Delta\omega/2\pi = 15$ MHz.

CONCLUSION

We used numerical methods to study the parametric amplification of nonlinear packets of reverse volume spin waves in ferromagnetic films by microwave parallel magnetic pumping. It is demonstrated that a relationship between the initial phases of signal and pumping carrier waves strongly influences the amplification coefficient because of a spatial interference of the signal wave with an idle wave propagating in the same direction.

ACKNOWLEDGMENTS

This work was supported in part by the Russian Foundation for Basic Research (project no. 96-02-19515), Russian Ministry of Education (project no. 97-8.3-13), and German Scientific Society (DFG project no. 436 RUS 113/27/0).

REFERENCES

1. B. A. Kalinikos, N. G. Kovshikov, M. P. Kostylev, *et al.*, Pis'ma Zh. Éksp. Teor. Fiz. **66**, 346 (1997).
2. P. A. Kolodin, P. Kabos, C. E. Patton, *et al.*, Phys. Rev. Lett. **80** (9), 1976 (1998).
3. A. V. Bagada, G. A. Melkov, A. A. Serga, *et al*., J. Appl. Phys. **81**, 5081 (1997).
4. B. A. Kalinikos and M. P. Kostylev, IEEE Trans. Magn. **33**, 3445 (1997).
5. V. S. L'vov, *Nonlinear Spin Waves* (Nauka, Moscow, 1987).
6. B. A. Kalinikos, Izv. Vyssh. Uchebn. Zaved., Fiz. **24** (8), 42 (1981).
7. M. P. Kostylev, B. A. Kalinikos, and H. Dötsch, J. Magn. Mater. **145**, 93 (1995).
8. A. N. Slavin and I. N. Rojdestvenskii, IEEE Trans. Magn. **30** (1), 37 (1994).
9. A. P. Sukhorukov, *Nonlinear Wave Interactions in Optics and Radiophysics* (Nauka, Moscow, 1988).

Translated by A. Chikishev

Technical Physics, Vol. 45, No. 2, 2000, pp. 281–284. Translated from Zhurnal Tekhnicheskoĭ Fiziki, Vol. 70, No. 2, 2000, pp. 139–142.

BRIEF COMMUNICATIONS

The Investigation of Noise in Hard Radiation Detectors by Pulse-Amplitude Analysis

A. M. Ivanov and N. B. Strokan

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

Received December 1, 1998

Abstract—The potentialities of pulse-amplitude analysis for noise measurements are demonstrated with p^+–n silicon detectors. It is suggested to use the detector current as a parameter and vary it by illuminating the samples. The instrument was calibrated by the shot noise of the photocurrent. The criteria for shot noise are the linearity of the noise squared vs. current dependence and its slope. It is shown that conventional instrumentation for pulse-amplitude analysis provides accurate yet rapid noise investigation. For the detectors studied, flicker noise was absent even when the trapped charge in the field oxide increases by one order of magnitude. © *2000 MAIK "Nauka/Interperiodica".*

INTRODUCTION

Device noise is routinely measured by selective amplification techniques (see, e.g., [1]), which directly record the noise spectral density. Another approach is suggested in this work. It has been shown that the noise nature can be disclosed with conventional instrumentation for taking amplitude spectra, consisting of three main units: a charge-sensitive preamplifier, a passband-forming amplifier, and a spectrum analyzer. The amplifier employs *RC* differentiating–integrating filters, which provide a relatively wide passband.

METHODS OF MEASUREMENT

Analysis of noise at the output of the above system for a single *RC* signal formation goes back to the 1950s [2]. Subsequently, more sophisticated filters, including n-fold integration, were studied. For $n = 2$, the formula for noise expressed in units of charge has the form [3]

$$\bar{q}_n^2 = (1/4)\{(\lambda/\lambda + 1)^2[2kT(0.66/g_m)(C^2/T_{\text{int}}) + (1 + 2\lambda)(eI + 2kT/R)T_{\text{int}}] + C^2\bar{V}_f^2\}. \quad (1)$$

Here, e is the charge of an electron, k is the Boltzmann constant, T is absolute temperature, C is the total input capacitance, I is the sum of the detector current and the gate current of a FET with a transconductance g_m, R is the load resistance, λ is the *RC* differentiation-to-integration constant ratio ($T_{\text{diff}}/T_{\text{int}}$), and V_f is the flicker noise voltage. It is worthy to note that mathematically [see (1)] various noise sources make a different contribution to the passband formed by *RC* networks. The shot noise of the FET source–drain current turns out to be proportional to C^2/T_{int}. The input noise, related to the current I and the thermal noise of the load, varies as T_{int}. Finally, the flicker noise does not depend on integration time. This allows one to distinguish between noise sources of various nature.

Under our experimental conditions, formula (1) is greatly simplified. First, the parameters of a p^+–n detector are its capacitance and the reverse current of its current–voltage characteristic. These, as a rule, specify the values of C and I in the input circuit. The resistance R is taken sufficiently large so as to minimize its noise. The value of λ is close to unity. Second, the nonequilibrium charge is transferred most completely if a high electric field is produced in the working zone of the detector. At the same time, the high voltage decreases the detector capacitance and thus depresses the noise level. Therefore, in practice, it only remains to determine the value of the voltage (maximum admissible voltage) that does not generate excessive noise. This noise may change the second term in (1), which depends on the input current I, and also appear as flicker noise. Whether or not excessive noise is present can be judged from the value of the derivative $d\bar{q}_n^2/dI$. With regard for the above simplifications, it can be expressed as

$$d\bar{q}_n^2/dI = (3/16)eT_{\text{int}} + C^2/4(d\bar{V}_f^2/dI), \quad (2)$$

where $d\bar{V}_f^2/dI$ appears in the implicit form.

It follows that, in the absence of flicker noise (or when $d\bar{V}_f^2/dI$ is small), the derivative $d\bar{q}_n^2/dI$ equals (in units of energy) 1.16 $(\text{keV})^2/\text{nA}$ for $T_{\text{int}} = 1$ μs.

Light sensitivity of the detectors gives one a chance to vary the current noise by sample illumination. Also, illumination serves to calibrate the instrument by the photocurrent noise, which is known to be shot.

1063-7842/00/4502-0281$20.00

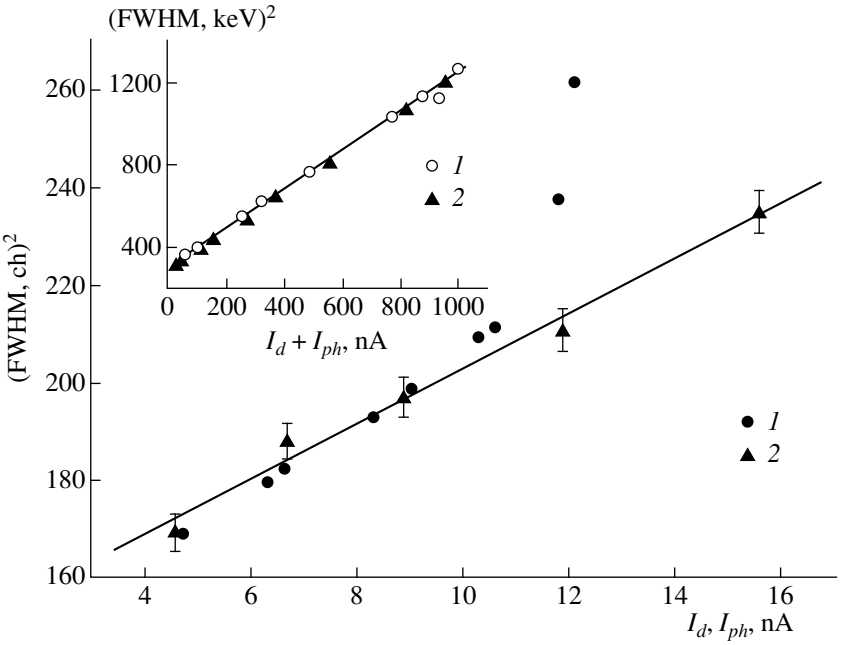


Fig. 1. Noise vs. current passing through detector *N*1. The current varies (*1*) with increasing voltage (50–225 V) and (*2*) under illumination at 50 V. The insert: noise vs. photocurrent in the absence of excessive noise for V = (circles) 50 and (triangles) 190 V.

Associated measurements were performed on Si detectors of accelerated ions. The device characteristics were described in detail elsewhere [4, 5]. Samples used were p^+–n junctions obtained by shallow diffusion of boron into high-resistivity n = Si. The contact to the p^+ layer was made by the so-called "extended electrode" technique.

The detectors and a standard LED (the source of calibrating photocurrent) were placed in a shielded chamber connected to the input of the preamplifier. Simultaneously, constant-amplitude pulses were applied to the input through a ~1-pF capacitance. The amplitude of the pulses was smeared by noise, and the resulting spectrum was displayed with the pulse-amplitude analyzer. The FWHM value, equal to $2.35\sqrt{\bar{q}_\mu^2}$, was measured. A pulse repetition rate of 10^4–10^5 Hz provided the desired count accuracy for a spectrum-formation time of about a minute. The scale division of the analyzer channel in units of energy was calibrated by seven alpha-decay lines of a ^{228}Th preparation.

RESULTS OF NOISE MEASUREMENTS

Figure 1 shows current dependences of the detector noise for two regimes. Curve *1* was obtained when the bias voltage was increased between 50 and 225 V. For curve *2* (approximated by the straight line), the voltage was kept constant (U = 50 V) and the current was varied by illuminating the detector (LED emission). With T_{int} = 3 μs and the small capacitance of the detector, the current noise was dominant.

From Fig. 1 it follows that the dark current noise ($U \le$ 200 V) and the noise of the photocurrent (U = 50 V) are nearly equal up to $I \approx$ 10 nA. The slope of the curve is, as expected, 1.19 (keV)^2/(nA μs). The noise can thus be thought of as being shot, with its level independent of the current nature. At U > 200 V, excessive noise arises, most probably because of local breakdown of the p^+–n junction.

The insert in Fig. 1 shows the run of the photocurrent noise up to 10^3 nA for two voltages: U_1 = 50 and near-breakdown voltage U_2 = 190 V. In this case, the slope (1.04) is somewhat less than the calculated value.

Figure 2 shows the current dependences of the noise for another illuminated sample for U_1 = 50 and U_2 = 90 V (here, the latter value falls into the region where the dark current noise steeply rises). It is remarkable that the lines are parallel to each other. The constancy of the derivative $d(FWHM)^2/dI$ means that, first, the noise of I_d and that of I_{ph} are summed algebraically without influencing each other. In other words, the noise depends on the voltage rather than the current. Second, the density of I_{ph} (in those regions of the junction where U_2 = 90 V becomes a critical value) is apparently far less than that of I_d, which causes the excessive noise. That is, only a small fraction of the total detector area is responsible for the excessive noise. Finally, the coincidence of $d(FWHM)^2/dI$ with its calculated value suggests that the flicker noise is insignificant. The last conclusion is of practical importance, since it implies that the periphery of the p^+–n junction is well protected.

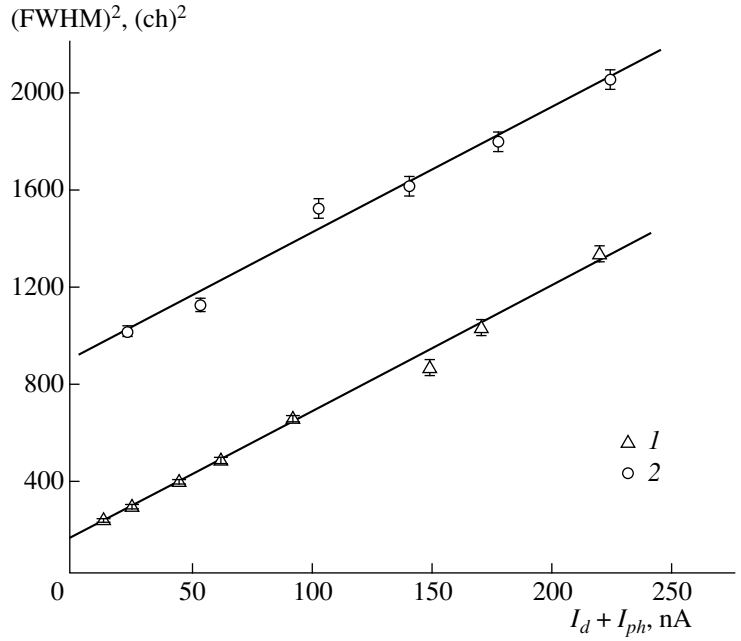


Fig. 2. Noise vs. photocurrent in the presence of excessive noise of the dark current (detector *N*2). *V* = (*1*) 50 and (*2*) 90 V. The value of $d(FWHM)^2/dI$ is constant.

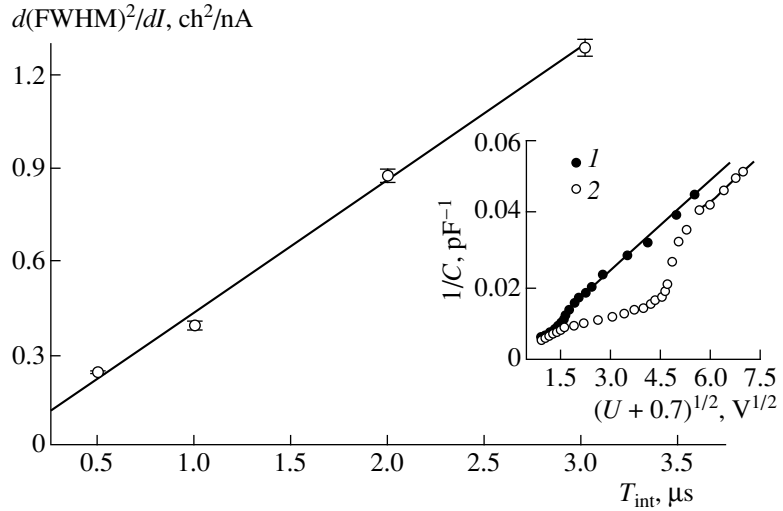


Fig. 3. Current derivative of noise plotted as a function of T_{int} to show the absence of flicker noise (detector *N*2). Unfavorable factors: additional charge built in the field oxide and voltage taken early in the sharp rise in the dark current. The insert: capacitance of the detector against bias voltage (*1*) before and (*2*) after the incorporation of the charge.

To confirm the high quality of the periphery, an additional charge was introduced into the field oxide. For this purpose, the detector was exposed to 12-keV X-ray radiation for 60 min. The radiation-induced current of the detector was 0.33 μA. The irradiation substantially changed the capacitance–voltage curve (insert in Fig. 3). After the irradiation, the built-in charge increased by one order of magnitude. However, the increase in the charge had no effect on the values of $d(FWHM)^2/dI$, which were determined at different T_{int}'s to obtain more reliable results (see table).

In the absence of the flicker noise, the dependence $d\bar{q}_n^2/dI = f(T_{int})$ passes through the origin, as follows from formula (2). This was checked by intentionally raising the bias to the near-breakdown voltage (for the given sample) U = 180 V. Figure 3, where $d(FWHM)^2/dI$ is plotted on the ordinate axis, support the expected run. The data points fall on the straight line $y = 0.002 + 0.43T_{int}$. Its slope is 1.10 (keV)^2/(nA μs), which is close to the calculated value 1.10 (keV)^2/(nA μs).

Values of the derivative $d(FWHM)^2/dI$ in the analyzer channels [ch^2/nA] before and after X-ray irradiation of the specimen

T_{int}, μs	As-prepared specimen		Irradiated specimen	
0.5	0.81	0.85	0.82	0.83
1.0	1.68	1.70	1.54	1.55
2.0	3.47	3.56	3.49	3.42
3.0	5.10	5.08	5.10	5.02

Note: Bias voltage is 50 V for the second and fourth columns and 90 V for the third and fifth columns.

To conclude, our approach, where the passing current varies, seems to be promising for noise measurement. The choice of current (instead of, say, capacitance) as a parameter is physically more appropriate, since it provides a more penetrating insight into the noise character. Also, the detectors were found to be virtually free of flicker noise, which suggests high quality of the field oxide.

ACKNOWLEDGMENTS

The authors thank M. E. Boiko for the irradiation experiments.

REFERENCES

1. A. van der Ziel, *Noise: Sources, Characterization, Measurement* (New York, 1969; Sov. Radio, Moscow, 1973).
2. A. B. Gillespie, *Signal, Noise, and Resolution in Nuclear Counter Amplifiers* (Pergamon, London, 1953), p. 155.
3. M. Tsukuda, Nucl. Instrum. Methods Phys. Res., Sect. B. **14**, 241 (1961).
4. E. M. Verbitskaya, V. K. Eremin, A. M. Malyarenko, *et al.*, Fiz. Tekh. Poluprovodn. (St. Petersburg) **27**, 2052 (1993) [Semicond. **27**, 1127 (1993)].
5. V. Eremin, I. Ilyashenko, A. Malyarenko, *et al.*, in *Electrochem. Soc. Ser. High Purity Silicon IV,* Ed. by C. L. Claeys (Pennington), Vol. 96-13, pp. 395–406.

Translated by V.A. Isaakyan