Technical Physics, Vol. 49, No. 1, 2004, pp. 1–7. Translated from Zhurnal Tekhnicheskoĭ Fiziki, Vol. 74, No. 1, 2004, pp. 3–9.

THEORETICAL AND MATHEMATICAL PHYSICS

The Second Painlevé Equation in the Electrostatic Probe Theory: Numerical Solutions for the Partial Absorption of Charged Particles by the Surface

A. V. Kashevarov

Zhukovsky Central Institute of Aerohydrodynamics, Zhukovskiĭ, Moscow Oblast, 140180 Russia

e-mail: kash@dept.aerocentr.msk.su

Received January 31, 2000; in final form, May 30, 2002

Abstract—We continue the study of the second Painlevé equation within the framework of the electrostatic probe theory. The integrability conditions for the equation are found for the partial absorption of charged particles by the probe surface. A sets of solutions with the asymptotics $y \sim \nu/x$ for $x \longrightarrow +\infty$ is constructed numerically in a wide range of the free parameter ν. Also, solutions (related to those mentioned above) for half-integer and integer ν, including solutions representable in asymptotic form at $x \longrightarrow +\infty$ through the Airy function $y \sim c\mathrm{Ai}(x)$ in the limit $\nu \longrightarrow 0$, are found. The results are discussed from the standpoint of the isomonodromic deformation method. © *2004 MAIK "Nauka/Interperiodica".*

INTRODUCTION

The second Painlevé equation (P2)

$$d^2y/dx^2 = 2y^3 + xy - \nu, \tag{1}$$

which is a nonlinear analogue of the Airy equation, has appeared from the evolution of the theory of differential equations and has been of pure mathematical interest for a long time [1]. Primary attention was focused on the asymptotic behavior of the P2-generated function at large values of its argument and on a relationship between the asymptotics of a particular solution at $\pm\infty$. The problem was solved by using the isomonodromic deformation method [2–5] developed recently, which is an analogue of the Laplace integral transformation method as applied to the Painlevé equations.

Nearly at the same time, it was found that the P2 equation occurs in a number of problems of modern theoretical and mathematical physics. Eventually, many of the researchers stated their belief that the P2 equation, along with five other Painlevé equations, must play the same role in nonlinear theoretical physics as classical special functions in linear problems [4, 5]. Therefore, the problem of calculating the values of this nonlinear special function, which is often referred to as the Painlevé function of the second kind, becomes topical.

This function is expressed via rational and known special functions (Airy functions) only in two particular cases, namely, when the free parameter ν equals 0 and 1/2. In the general case, a mathematical description of this function, e.g., in the form of a series or integral representation, that would allow one to calculate its values is absent. The only method is numerical integration.

In integrating the P2 equation, the problem of choosing initial conditions providing a given asymptotics of particular solutions arises. Moreover, a general solution to the P2 equation is known to be a discontinuous function with movable singularities in the form of first-order poles; i.e., the positions of singularities are dependent on initial data. Only two attempts to numerically integrate the P2 equation with $\nu = 0$ have met with success [6, 7]. These attempts have culminated in solutions with the asymptotics $y \sim c\mathrm{Ai}(x)$ for $x \longrightarrow +\infty$, where c is a parameter and $\mathrm{Ai}(x)$ is the Airy function.

The application of the P2 equation to the electrostatic probe problem [8] has remained practically unknown. In [9], it was shown how the properties of the P2 equation and its solutions show up in the electrostatic probe theory. It was found, in particular, that the well-known relationship [1] between solutions to the P2 equation at $\nu = 0$ and 1/2 is given by the equation

$$u_{xx} = -u_x^2 - x/2 + e^{2u}/2. \tag{2}$$

The derivative u_x satisfies (1) with $\nu = 1/2$, and the function $\eta(\zeta) = 2^{-2/3}e^{u(x)}$ with $\zeta = -2^{-1/3}x$ is a solution to the P2 at $\nu = 0$. Using (2), the form of the monodromic matrix [2] at $\nu = 1/2$ has been refined.

In the probe theory, solutions to the P2 equation that are regular at $+\infty$ with the asymptotics $y \sim \nu/x$ ($x \longrightarrow +\infty$) appear. The detailed asymptotic description of these functions in the form of a series is given in [1]. In [9], initial conditions that these solutions must meet were found and the solutions were constructed numerically. The initial conditions found resulted from the analysis of a singularly perturbed set of differential equations that describes the operation of a probe when charged particles are totally absorbed by its surface. In

1063-7842/04/4901-0001$26.00 © 2004 MAIK "Nauka/Interperiodica"

this case, for any $\nu < 1/2$, only a single solution with the asymptotics mentioned above can be constructed. However, it is known [1] that there exists a set of such solutions at any ν. In this paper, the physically justified case of partial absorption of charged particles is studied. This makes it possible to perform numerical integration and find solutions from this set.

INTEGRABILITY CONDITIONS AND NUMERICAL PROCEDURE

The operation of a spherical probe in a collisional plasma can be described by the equations [9]

$$dn_+/d\xi + n_+E = -I_+, \quad dn_-/d\xi - n_-E = -I_-,$$

$$\alpha^2\xi^4 dE/d\xi = n_- - n_+, \quad E = -d\psi/d\xi, \tag{3}$$

$$\xi \in [0, 1], \quad \psi(0) = 0, \quad n_\pm(0) = 1.$$

Here, n_+ and n_- are the dimensionless numerical concentrations of positively and negatively charged particles, respectively; I_+ and I_- are their currents toward the probe; E is the electric field strength; ψ is the dimensionless electric potential; $\xi = 1/r$, where r is the radial coordinate of the spherical coordinate system; and α is the ratio of the Debye length to the probe radius.

Detailed asymptotic analysis of singularly perturbed set (3) at $\alpha \ll 1$ is presented in [9]. Canonical form (1) of the P2 equation can be obtained from set (3) by reducing it to a single equation for the field strength E and applying the transformations [9]

$$x = 2^{1/3}\alpha^{-2/3}\xi_s^{-5/3}(\xi_s - \xi),$$
$$E(\xi) = 2^{4/3}\alpha^{-2/3}\xi_s^{-5/3}y(x) \tag{4}$$

to this equation. Here, $\xi_s = 2/(I_+ + I_-)$ is the singularity point of an external solution to (3). The parameter ν in (1) is related to the probe currents as $\nu = (I_- - I_+)/(I_- + I_+)/2$. Since the dimensionless currents I_+ and I_- are always positive in the absence of particle emission from the probe surface, the free parameter ν in (1) may vary (in the framework of the probe theory) only within the interval $|\nu| \le 1/2$. However, solutions to the P2 for any ν from this interval, excluding $\nu = -1/2$, are known [1] to define solutions for arbitrary ν by recurrence relations. Therefore, in [9], solutions to the P2 were found in a wide range of ν.

The integrability conditions for the P2 at $\nu < 1/2$ were found [9] for the case when the electric potential $\psi(1) = \psi_0$ at the probe surface is given and the concentrations of both positively and negatively charged particles are $n_+(1) = n_-(1) = 0$ (boundary conditions for (3)). Let us find the integrability conditions in the general case, where $n_+(1) = n_{0+} > 0$ and/or $n_-(1) = n_{0-} > 0$ at the probe surface.

From set (3), it follows as an intermediate result that

$$n_- = \frac{\alpha^2(\xi^4E')' + \alpha^2\xi^4EE' + I_- - I_+}{2E}, \tag{5}$$

where the prime denotes differentiation with respect to ξ.

Let us apply transformations (4) to (5) and put

$$n_- = 2^{-1/3}\alpha^{2/3}\xi_s^{2/3}\eta_-. \tag{6}$$

Then, at a point x_0 (unknown beforehand) on the probe surface, the relationship

$$y_0'' = 2y_0y_0' + y_0\eta_{0-} - \nu \tag{7}$$

is valid. In (7), the subscript 0 designates quantities found at $x = x_0$ and the prime means differentiation with respect to x. Applying transformations (5) and (6) to the third equation of (2), we arrive at

$$y_0' = (\eta_{0+} - \eta_{0-})/4 \tag{8}$$

at $\alpha \longrightarrow 0$. Here, η_{0+} and n_{0+} are related by the same relationship (6) as η_{0-} and n_{0-}. Substituting (8) into (7) and taking into account (1) yields

$$y_0 = \pm\sqrt{-x_0/2 + (\eta_{0+} + \eta_{0-})/4}. \tag{9}$$

Upon numerically integrating (1) at $|\nu| < 1/2$, conditions (8) and (9) provide solutions to the P2 with the asymptotics $y \sim \nu/x$ for $x \longrightarrow +\infty$. If charged particles are absorbed completely, when $\eta_{0+} = \eta_{0-} = 0$, conditions (8) and (9) coincide with those found in [9]. Note that, by physical considerations, the plus sign was chosen in [9].

In the case $\nu = 1/2$, which corresponds to a large negative probe potential ($\psi_0 \ll -1$), so that one can assume $I_+ = 0$, it is convenient to use Eq. (2) for calculating numerical solutions to the P2. In order that the boundary-value problem for (3) be solvable, it is necessary that $n_+(1) = n_{0+} = \exp(\psi_0) > 0$ [9]. For $\nu = 1/2$, a set of regular solutions to the P2 with the same asymptotics at $x \longrightarrow +\infty$ was found [9] under the condition that negatively charged particles are totally absorbed: $n_-(1) = n_{0-} = 0$.

In the case of partial absorption of negatively charged particles, the following integrability conditions for Eq. (2) can be obtained:

$$u_0' = \pm\sqrt{-x_0/2 + (\eta_{0+} + \eta_{0-})/4}, \quad u(x_0) = u_0, \tag{10}$$

where $\eta_{0+} = \exp(2u_0)$. Note that, in contrast to η_{0-}, η_{0+} cannot take an arbitrary value at a given u_0.

The computational procedure is described in detail elsewhere [9]. Note only that, after the transition to equivalent sets of first-order differential equations, we will use difference schemes [10] to integrate (1) and

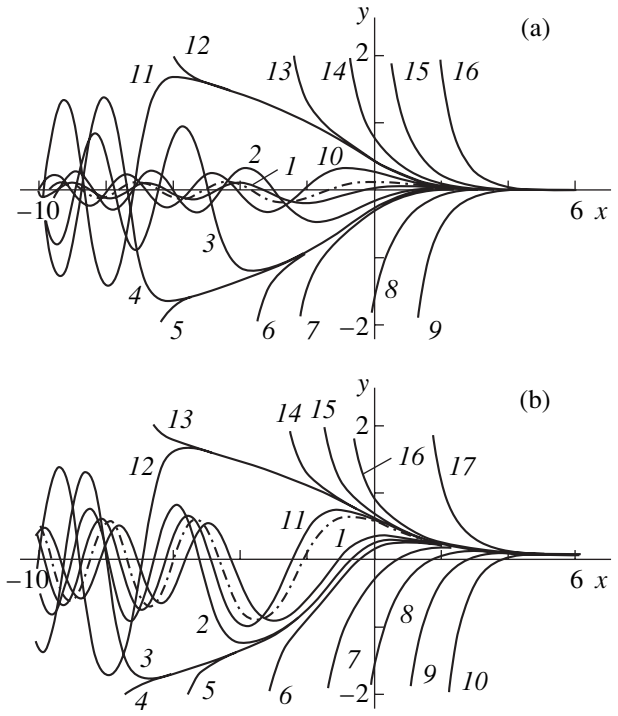

Fig. 1.

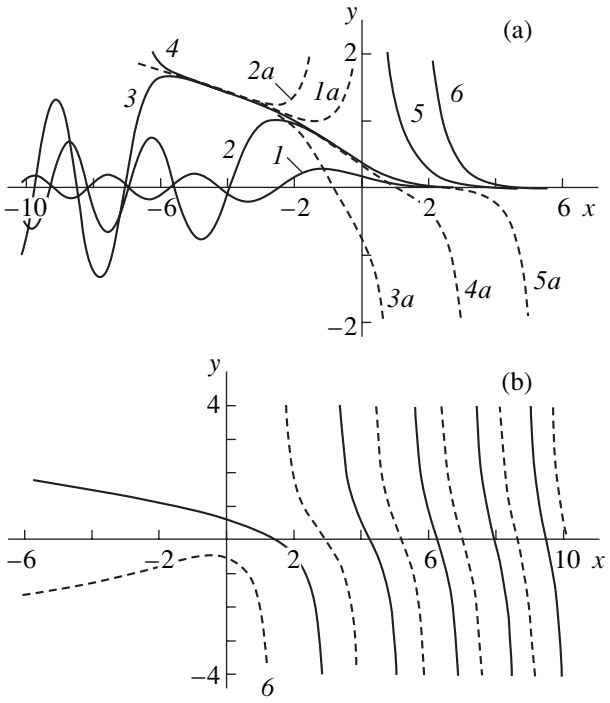

Fig. 2.

(2). For example, for Eq. (1), the scheme is the following:

$$y_{i+1} = y_i + h(az_{i+1}/b + z_i/b),$$

$$z_{i+1} = z_i + \frac{h}{12}[bf(x_{i+1}, y_{i+1})]/a + bf(x_i, y_i).$$

Here, $a = 5 - \sqrt{24}$, $b = 6 - \sqrt{24}$, $z = y'$, and $f(x, y)$ is the right of Eq. (1). The convergence of the difference scheme was verified [9] by splitting a step h, which was finally taken to be $h = 10^{-4}$. The major problem here is to select the starting point of integration x_0.

NUMERICAL RESULTS

Figure 1 shows the results of numerical integration of the P2 at $\nu = 0.1$ and 0.4 for various values of the parameters η_{0+} and η_{0-}. The dash-and-dot lines are the solutions at $\eta_{0+} = \eta_{0-} = 0$ [9].

Curves *1–9* in Fig. 1a were calculated for $\eta_{0+} = \eta_{0-} = 1$, 2.5, 2.87, 2.870990, 2.870991, 2.88, 3, 4, and 5, respectively. Curves *10–15* were obtained at $\eta_{0+} = 0$ and $\eta_{0-} = 0.5$, 1.049210, 1.049211, 1.1, 2, and 6, respectively. Curve *16* corresponds to $\eta_{0+} = 1$ and $\eta_{0-} = 6$. From Fig. 1a, it follows that there are regular solutions with the same asymptotics $y \sim \nu/x$ at $x \longrightarrow +\infty$ and those with different asymptotics at $-\infty$. Among them are oscillating solutions *1–3*, *11*, and *12* and also singular solutions *4–10* and *13–17*. The curves of the two types are separated by two separatrix solutions, which are unstable against weak perturbations of the initial data. The separatrix solutions have the asymptotics

$$y \sim \pm\sqrt{-x/2}. \qquad (11)$$

Note that the solutions to the P2 with the asymptotics $y \sim -\sqrt{-x/2}$ at $x \longrightarrow -\infty$ were known previously for $\nu < 0$ [1].

Similar results were obtained at $\nu = 0.4$ (Fig. 1b). Curves *1–10* correspond to $\eta_{0+} = \eta_{0-} =$ (*1*) 0.5, (*2*) 0.77, (*3*) 0.779663, (*4*) 0.779664, (*5*) 0.78, (*6*) 1, (*7*) 2, (*8*) 3, (*9*) 4, and (*10*) 5. Curves *11–16* were calculated for $\eta_{0+} = 0$ and $\eta_{0-} =$ (*11*) 0.5, (*12*) 1.009605, (*13*) 1.009606, (*14*) 1.1, (*15*) 2, and (*16*) 6; curve *17*, at $\eta_{0+} = 1$ and $\eta_{0-} = 6$.

The results of calculating the starting point of integration x_0 for various values of ν, η_{0+}, and η_{0-} are listed in the table. The table also includes the numbers n of terms involved in the asymptotic series [1] that describes the asymptotics of the regular (at $x \longrightarrow +\infty$) solutions to the P2. In [9], four terms of the series were taken into account; this number is, as a rule, insufficient to obtain a desired accuracy of determining x_0 when charged particles are partially absorbed.

The behavior of solutions to the P2 at small $\nu \longrightarrow 0$ is of interest. For this case, it was found [1] that the coordinate x_0 and, together with it, $y_0 \longrightarrow 0$ and the solution to the P2 tends to the trivial solution $y = 0$ for the complete absorption of charged particles. In our

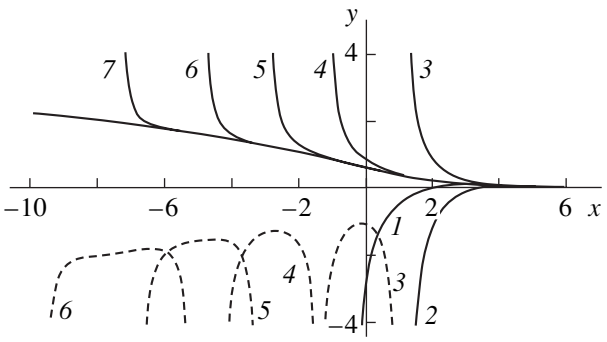

Fig. 3.

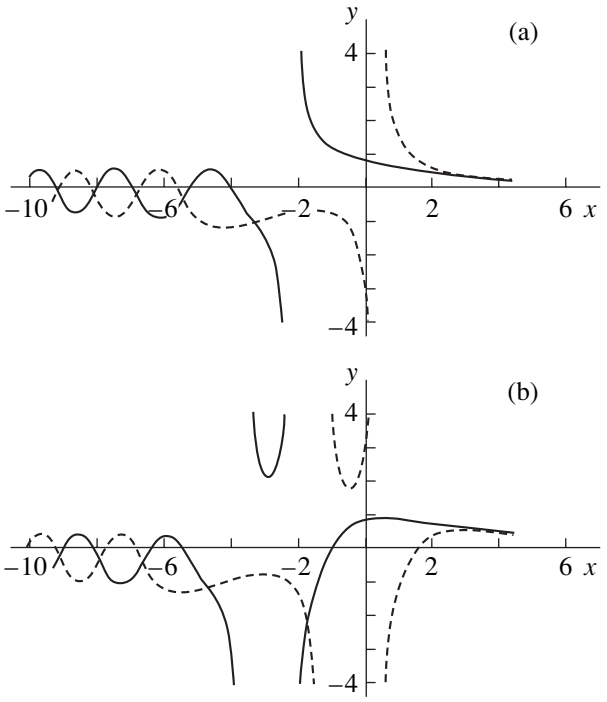

Fig. 4.

case (partial absorption), x_0 and y_0 tend to certain limits other than zero.

The results at $\nu = 10^{-11}$ are shown in Fig. 2a (solid lines) and in the table. Curves *1–5* were constructed at $\eta_{0+} = 0$ and $\eta_{0-} = 0.5$, 1, 1.009623, 1.009624, and 6, respectively; curve *6*, at $\eta_{0+} = 1$ and $\eta_{0-} = 6$. The curves symmetric with respect to the x axis are also solutions that appear when η_{0-} and η_{0+} are substituted for η_{0+} and η_{0-}.

The solid lines in Fig. 2b show a solution to the P2 at $\nu = 0.5$, which can be derived from solution *1* (Fig. 2a) with the well-known formula relating the solutions for $\nu = 0.5$ and 0 (see the Introduction). It is seen that continuous solution *1* for $\nu = 0$ gives rise to the singular solution for $\nu = 0.5$. The latter turns out to be close to the solution

$$y = -2^{-1/3}(\mathrm{Ai}'(\zeta) + C\mathrm{Bi}'(\zeta))/(\mathrm{Ai}(\zeta) + C\mathrm{Bi}(\zeta)), \quad \zeta = -2^{-1/3}x \tag{12}$$

to the one-parameter set [2] at $C = 0$. As is known, solutions to this set are not related to the solutions to the P2 at $\nu = 0$. For comparison, the dashed lines in Fig. 2b show another solution to this set in the limit $C \longrightarrow \infty$; i.e., $y = -2^{-1/3}\mathrm{Bi}'(\zeta)/\mathrm{Bi}(\zeta)$.

The solid lines in Fig. 3 demonstrate solutions to the P2 for $\nu = 0.5$ that were found (as in [9]) by numerically integrating auxiliary equation (2) (see also the table). First, the results of [9] for $\eta_{0-} = 0$ and $\eta_{0+} > 0$ were supplemented. In [9], attempts to find solutions to (2) and, correspondingly, to the P2 for $u_0 > 0.203$, i.e., $\eta_{0+} > e^{0.406}$, failed. It turned out that, at such values of η_{0+}, one has to choose the negative value of the root in condition (10). In view of this fact, the solutions to Eqs. (2) and P2 for $u_0 = 0.5$ and 1 were found (curves *1, 2*). Curves *3–7*, which belong to a different set of solutions than curves *1* and *2*, were found in the case of partial absorption of both positively and negatively charged particles at $\eta_{0+} = 1$ and $\eta_{0-} = 6$, 2, 1.76, 1.7542, and 1.75417557, respectively.

The integral curves for the P2 equation at $\nu = 0$ that correspond to some of the curves discussed above are shown by the dashed lines (marked by the letter *a*) in Fig. 2a. The parts of curves *3a–5a* for $y < 0$ were found by directly integrating Eq. (1) with $\nu = 0$. Then, the solutions to the P2 for $\nu = 0.5$ cm (the dashed curves in Fig. 3) were completed using the well-known relationship between the solutions.

The solid curves in Figs. 4a and 4b show the solutions $y(x, \nu)$ to the P2 equation for $\nu = 1$ and 2 that were constructed with the recurrent formula [1] from solution *1* (Fig. 2a); the dashed curves are similar solutions derived from the symmetric solution $-y(x, 0)$.

Figures 5a and 5b show the typical form of other solutions, $y(x, 1)$ and $y(x, 2)$, that were derived from solution *5a* (Fig. 2a). Here, the poles of the functions at the right extreme of the range of integration (breaks in the y axis) are noteworthy.

Figures 6a and 6b illustrate the solutions $y(x, 1.5)$ and $y(x, 2.5)$, which were derived from the corresponding solutions $y(x, 0.5)$ (Fig. 3a). Note the different number of the poles of solution *5* for $\nu = 2.5$.

Finally, Figs. 7a and 7b exhibit the functions $y(x, 1.5)$ and $y(x, 2.5)$, which were derived from the corresponding solutions $y(x, 0.5)$ in Fig. 2b. Here, it is worthy to note the shift of the zeros and poles of the function $y(x, 2.5)$ with respect to those of $y(x, 0.5)$ and $y(x, 1.5)$ on the right of the plots.

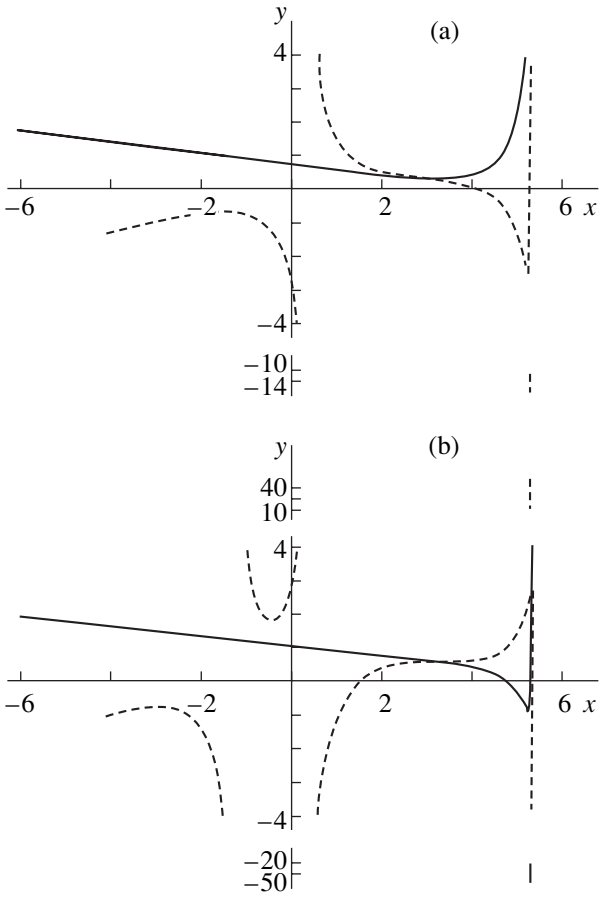

Fig. 5.

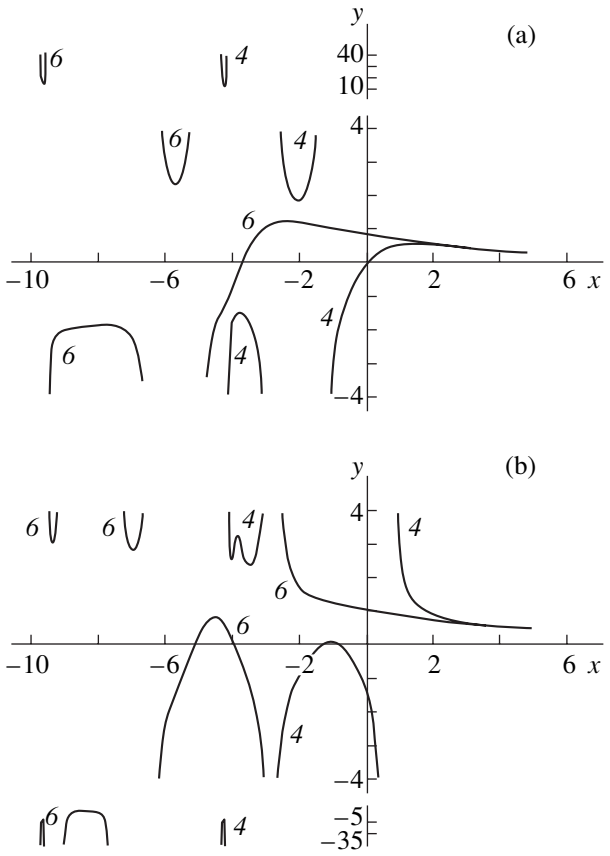

Fig. 6.

DISCUSSION OF THE SOLUTIONS IN TERMS OF THE ISOMONODROMIC DEFORMATION METHOD

Let us discuss the results found in terms of the isomonodromic deformation method [2]. This method was used both in [3] to describe the asymptotic behavior of real solutions to the P2 at $\nu = 0$ and in [5] to study the case of arbitrary ν. The asymptotics are represented through the Stokes multipliers, which are the nontrivial components of the Stokes matrices for the set of linear differential equations

$$\begin{aligned}\Psi_{1\zeta} &= -i(4\zeta^2 + x + 2y^2)\Psi_1 (4\zeta y + \nu/\zeta + 2iy_x)\Psi_2,\\ \Psi_{2\zeta} &= (4\zeta y + \nu/\zeta - 2iy_x)\Psi_1 + i(4\zeta^2 + x + 2y^2)\Psi_2\end{aligned} \tag{13}$$

that is associated with (1). Here, Ψ_1 and Ψ_2 are functions of the complex variable ζ. The Stokes matrices relate the formal asymptotic expansions of two linearly independent solutions to (13) at $\zeta \longrightarrow \infty$ in adjacent sectors of the complex plane. It is well known that y in (13) satisfies (1) if the Stokes matrices are independent of the deformation parameter x and vice versa. For the Stokes multipliers s_1, s_2, and s_3, we have

$$s_1 - s_2 + s_3 + s_1 s_2 s_3 = -2\sin\pi\nu. \tag{14}$$

In the case of real x and y, $s_1 = s_3^*$ and $s_2 = s_2^*$.

In [5], the complete description of the solutions $y \sim \nu/x$ ($x \longrightarrow +\infty$), which are studied here numerically, is lacking; however, it was noted that for them $s_2 = 0$. Then, putting $s_1 = \beta + i\gamma$, we find from (14) that for these solutions

$$\beta = -\sin\pi\nu. \tag{15}$$

As was found in [5], any solution to the P2 has one of three asymptotics for $x \longrightarrow -\infty$. A solution may have the form of damped oscillations, be a singular function of the cosecant type, or represent either an increasing or decreasing function with asymptotics (11). The plots presented above show that any of these asymptotics is realizable for those solutions to the P2 corresponding to the case $s_2 = 0$ at all but half-integer ν. The type of asymptotics depends on the value of the imaginary part γ of the Stokes multiplier s_1.

After transformations, one of the formulas in [5] yields

$$y \sim |-x|^{-1/4}\rho\sin(2|x|^{3/2}/3 - (3\rho^2/4)\ln|x| + \varphi),$$
$$\rho^2 = -\pi^{-1}\ln(1-|s_1|^2), \quad (16)$$
$$\varphi = -(3\rho^2\ln 2)/2 - \pi/4 + \arg[is_3\Gamma(i\rho^2/2)]$$

for real solutions to the P2 at $x \longrightarrow -\infty$.

Calculations show that, in view of (15), the asymptotics of the solutions to the P2 that are plotted in Fig. 1a satisfy (16) if $\gamma \approx 0$ (the dash-and-dot line), $\gamma \approx 0.15$ (curve *1*), 0.6 (*2*), 0.95 (*3*), 0.951056 (*4*), $\gamma \approx -0.5$ (*10*), and 0.951056 (*11*). The asymptotics of the solutions to the P2 that are shown in Fig. 1b are described by (16) if $\gamma \approx -0.2$ (the dash-and-dot line), $\gamma \approx 0.15$ (*1*), 0.3 (*2*), 0.3090155 (*3*), $\gamma \approx -0.26$ (*11*), and –0.3090167 (*12*).

Formula (16) is invalid for $|s_1|^2 > 1$, i.e., when $\gamma^2 > 1 - \sin^2\pi\nu$. In this case, the asymptotic behavior of the solutions to the P2 at $-\infty$ must be found by the formula [5]

$$y \sim |x|^{-1/2}\mathrm{cosec}(2|x|^{3/2}/3 - (3\rho^2/4)\ln|x| + \theta),$$
$$\rho^2 = \pi^{-1}\ln(|s_1|^2 - 1),$$
$$\theta = (3\rho^2/2)\ln 2 - \arg[s_1\Gamma(1/2 + i\rho^2/2)].$$

Accordingly, for solutions *5–9* and *12–16* in Fig. 1a, which have this asymptotics, as well as for solutions *4–10* and *13–17* in Fig. 1b, we have $|\gamma| > 0.951056$ and $|\gamma| > 0.309017$, respectively.

Separatrix solutions (11) meet the condition $1 - s_1s_3 = 0$ [5]; therefore, with $s_2 = 0$, we find $\gamma = \pm(1 - \sin^2\pi\nu)^{1/2}$. At $\nu = 1/2$, we have $\gamma = \pm 1$, $\beta = 0$, and $|s_1| = 1$. This implies that the solutions to the P2 that are regular at $+\infty$ cannot have oscillating asymptotics (16) in this case. This is confirmed by the results of numerical integration (Fig. 3).

If $\nu = 0$, then $\beta = 0$ and the Stokes multiplier s_1 is a pure imaginary. It is known [3] that, in this case, the behavior of the solutions to the P2 at $+\infty$ is described by the Airy function

$$y \sim c\mathrm{Ai}(x) \sim 2^{-1}\pi^{-1/2}cx^{-1/4}\exp(-2x^{3/2}/3). \quad (17)$$

At $x \longrightarrow -\infty$ and $c < 1$, the asymptotics of these solutions is represented by formula (16), which, at $\nu = 0$, is reduced to the similar formula in [3]. As follows from calculations, the asymptotics of solutions *1–3* in Fig. 2a satisfies (16) with $\gamma = -0.5$, –0.99, and –0.999999, respectively. In other words, having numerically constructed the solutions to the P2 for small $\nu \longrightarrow 0$ with the asymptotics $y \longrightarrow \nu/x$ at $x \longrightarrow +\infty$, we arrive at the solutions with asymptotics (17), which were found earlier [6, 7]. In [3], it was argued [3] that $c = \gamma$ in (16) at $\nu = 0$. However, the results presented here show that $c = -\gamma$.

Note that formula (17) can be derived by linearizing the P2 at $\nu = 0$. If $\nu \neq 0$, the linearization of Eq. (1) yields an inhomogeneous Airy equation; therefore, the

Table

Curve no.	Fig. 1a		Fig. 1b		Fig. 2a		Fig. 3	
	x_0	n	x_0	n	x_0	n	x_0	n
1	0.99256	4	0.26350	4	0.20631	14	1.2181	4
2	2.4978	5	0.60967	4	0.35894	15	2.1211	4
3	2.8682	6	0.62133	4	0.36154	16	2.0078	9
4	2.8692	6	0.62133	4	0.36154	16	1.1042	4
5	2.8692	6	0.62173	4	1.1196	17	1.0370	4
6	2.8782	6	0.87861	4	2.5531	20	1.0353	4
7	2.9984	6	1.9507	4			1.0353	4
8	3.9980	9	2.9737	5				
9	4.9993	12	3.9838	8				
10	0.11835	5	4.9891	11				
11	0.24812	6	–0.72004	4				
12	0.24812	6	–0.67894	4				
13	0.25879	6	–0.67894	4				
14	0.42248	6	–0.67210	4				
15	0.87682	9	–0.60981	4				
16	2.46000	10	–0.41046	5				
17			2.1342	9				

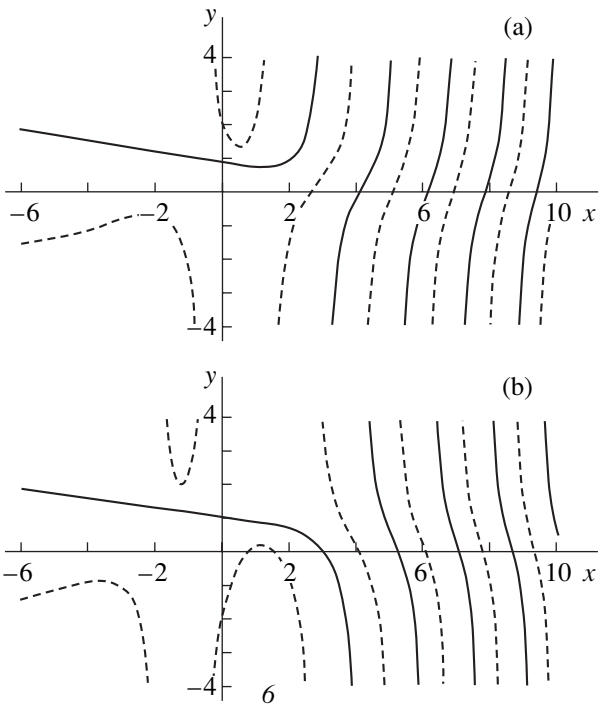

Fig. 7.

asymptotics of the solutions to the P2, $y \longrightarrow \nu/x$ at $x \longrightarrow +\infty$, can be represented in the form

$$y \sim c\mathrm{Ai}(x) + \nu\pi\mathrm{Gi}(x).$$

Here, $\mathrm{Gi}(x) \sim \pi^{-1}x^{-1}$ at $x \longrightarrow +\infty$ [11] is a solution of the inhomogeneous Airy equation $y'' = xy - \pi^{-1}$.

Curves *1a–5a* (Fig. 2a) are also separatrix curves at $x \longrightarrow -\infty$ with asymptotics (11). They correspond to the same values of β and γ ($\beta = 0$ and $\gamma = -1$) but different s_2.

If $s_2 \neq 0$, the asymptotics of the solutions to the P2 at $x \longrightarrow +\infty$ is given by [5]

$$y(x) \sim \sigma\sqrt{x/2}\cot[2^{1/2}x^{3/2}/3 + (3\rho/4)\ln x + (5\rho/4)\ln 2 - X/2], \quad (18)$$

where $\sigma = \mathrm{sgn}(s_2)$, $\rho = \pi^{-1}\ln|s_2|$, and $X = \arg[\Gamma(3/4 + \sigma\nu/2 + i\rho/2)\Gamma(1/4 - \sigma\nu/2 + i\rho/2) - \arg(1 + s_1s_2)$.

Solutions *1a* and *2a* correspond to $s_2 < 0$; *3a–5a*, to $s_2 > 0$.

For the solutions to the P2 at $\nu = 1/2$ that are shown in Fig. 2b, $\beta = -1$ and $\gamma = 0$. The asymptotics of the solution that is depicted by the solid lines (Fig. 2b) satisfies (18) at $s_2 \approx 0.72$. Solution (12) at $C \longrightarrow \infty$ shown by the dashed lines corresponds to $s_2 = 1$, when formula (18) becomes invalid. It is of interest that, using the asymptotic representations of the Airy functions Ai($-x$) and Bi($-x$) and of their derivatives for $x \longrightarrow +\infty$ [11], we find asymptotics (18) for solution (12) with $C = 0$ ($C \longrightarrow \infty$) if the Stokes multiplier s_2 tends to unity from the left (right): $s_2 = 1 - 0$ and $s_2 = 1 + 0$, respectively.

CONCLUSIONS

It is shown that basically the P2 equation may be numerically integrated. This allows one to evaluate the Painlevé function of the second kind in several particular cases. The question of how to denote this function to single out a specific solution arises. Since any specific solution may be uniquely characterized by the value of the free parameter ν, the imaginary part γ of the Stokes multiplier s_1, and the real multiplier s_2, we suggest that the Painlevé function of the second kind be denoted as P2(x; ν, γ, s_2). For example, solutions with the asymptotics $y \sim \nu/x$ for $x \longrightarrow +\infty$ obtained in this paper may be denoted as P2(x; ν, γ, 0).

ACKNOWLEDGMENTS

The author is grateful to S.Yu. Slavyanov, A.I. Kirillov, and T.V. Rogozkina for their assistance in publication.

REFERENCES

1. V. I. Gromak, Differentsial'nye Uravneniya **18**, 753 (1982).
2. H. Flaschka and A. C. Newell, Commun. Math. Phys. **76**, 65 (1980).
3. A. A. Kapaev and V. Yu. Novokshenov, Dokl. Akad. Nauk SSSR **290**, 590 (1986) [Sov. Phys. Dokl. **31**, 719 (1986)].
4. A. P. Its and A. A. Kapaev, Izv. Akad. Nauk SSSR, Ser. Mat. **51**, 878 (1987).
5. A. A. Kapaev, Teor. Mat. Fiz. **77**, 323 (1988).
6. R. R. Rosales, Proc. R. Soc. London, Ser. A **361** (1706), 265 (1978).
7. J. Miles, Proc. R. Soc. London, Ser. A **361** (1706), 277 (1978).
8. I. M. Cohen, Phys. Fluids **6**, 1492 (1963).
9. A. V. Kashevarov, Zh. Vychisl. Mat. Mat. Fiz. **38**, 992 (1998).
10. I. Babuška, M. Práger, and E. Vitásek, *Numerical Processes in Differential Equations* (Wiley, New York, 1966; Mir, Moscow, 1969).
11. *Handbook of Mathematical Functions*, Ed. by M. Abramowitz and I. A. Stegun (Dover, New York, 1971; Nauka, Moscow, 1979).

Translated by M. Fofanov

Technical Physics, Vol. 49, No. 1, 2004, pp. 104–107. Translated from Zhurnal Tekhnicheskoĭ Fiziki, Vol. 74, No. 1, 2004, pp. 104–107.

ELECTRON AND ION BEAMS, ACCELERATORS

Specific Features of the Formation of a Uniform Ribbon Electron Beam by a Plasma Source in the Forevacuum Pressure Range

V. A. Burdovitsin, Yu. A. Burachevsky, E. M. Oks, and M. V. Fedorov

Tomsk State University of Control Systems and Radio Electronics, Tomsk, 634050 Russia

e-mail: burdov@fet.tusur.ru

Received February 21, 2003

Abstract—Results are presented from experimental studies of the formation of a ribbon electron beam during the extraction of electrons from the plasma of a steady-state hollow-cathode discharge in the forevacuum pressure range. It is shown that the main reason for the nonuniformity of the current density is the increase in the local nonuniformity of the emission plasma density caused by the return flow of ions from the accelerating gap. Taking this feature into account when developing a system of beam extraction provides for the generation of a ribbon beam with a nonuniformity of the current density along the beam of less than 10%. © *2004 MAIK "Nauka/Interperiodica".*

INTRODUCTION

One of the promising applications of ribbon electron beams is the formation of large-area (up to 1 m^2) "plasma sheets" that can be used in various technological processes (such as plasmochemical and ion etching and the deposition of various coatings in the course of decomposition and fusion reactions in the generated plasma) and as moving microwave mirrors [1]. A fairly high gas pressure (10–100 Pa) is required to generate such a plasma. As a result, it is very difficult to use sources with a thermionic cathode for this purpose, so that there is, in fact, no alternative to the use of plasma electron sources based on the extraction of electrons from the plasma of low-pressure discharges with nonincandescent electrodes [2–4]. In forming a ribbon electron beam, one of the most important problems is that of attaining a highly uniform current density. This problem was considered by Bugaev *et al.* [5] who analyzed the main reasons for the nonuniformity of the emission current in large-cross-section beams and suggested ways of eliminating these reasons. At the same time, Bugaev *et al.* [5] largely treated "standard" plasma sources of electrons, whose working pressure range was, as a rule, below 0.1 Pa. At such pressures, the degree of uniformity of the current density of an electron beam extracted from the plasma is largely determined by the uniformity of the emission plasma. The transition to the forevacuum pressure range results in an increase in the effect of the return flow of ions formed in the accelerating gap and the electron-beam transport region on the emissive properties of the plasma [6]. Therefore, for plasma sources of electrons operating in the forevacuum pressure range, it seems insufficient to achieve the initial uniformity of the emission plasma. It was shown in [6, 7] that an increase in the working pressure makes it necessary to take into account the ionization processes in the accelerating gap and attendant phenomena.

In this paper, we describe the results of investigations of the formation of a ribbon electron beam with a highly uniform current density in a forevacuum plasma source of electrons based on a discharge with an extended hollow cathode [8].

EXPERIMENTAL SETUP

The experimental forevacuum plasma source of electrons for the generation of a ribbon beam comprised the same basic elements as the source of a cylindrical electron beam described by us in [6], namely, a hollow cathode, a flat anode with an emission opening, an accelerating electrode, and a collector. Rectangular hollow cathode *1* (Fig. 1) 300 × 80 × 40 mm in size maintained a steady-state discharge with a current of up to 1.5 A. The beam size was determined by a 250 × 10 mm emission slot in anode *2*. The slot was overlapped by fine-mesh metal grid *3*. Cathode *1*, anode *2*, and accelerating electrode *4* were electrically separated from one another by caprolan insulators *5* and *6*. The parameters of the emitting plasma were measured with cylindrical probes *7* introduced into the plasma via channels in insulator *5*. The probes were arranged so that fast ions from the accelerating gap could not fall on their collecting surfaces. The working gas was air. Figure 1 also shows how the sources of the discharge and accelerating voltages (U_d aqd U_a, respectively) were connected to the electrodes. The electron current distribution along the electron beam was measured using movable molybdenum collector *8* located behind a grounded grid having a slot with a width of 1 mm and

1063-7842/04/4901-0104$26.00 © 2004 MAIK "Nauka/Interperiodica"

length exceeding the beam size. The distance from the emission grid to the collector was 15 cm.

EXPERIMENTAL RESULTS

In order to determine the distribution of the plasma density along the hollow cathode, the currents to probes *7* (Fig. 1) were measured in the ion segment of the current–voltage characteristic. In the absence of electron emission, the nonuniformity of the plasma density along the hollow does not exceed 5–10%, except for the density maxima at the edges. At the same time, the electron beam extracted when applying the accelerating voltage is significantly nonuniform and, as visual observations demonstrate, turns out to consist of at least ten fine beams (jets). The distribution of the current density $i(x)$ at the movable collector under different experimental conditions is given in Fig. 2. In this case, a much smaller number of experimentally recorded maxima is due to the effect of individual jets in the region where the distribution of the electron beam current is measured. A decrease in the pressure resulted in the disappearance of the beam nonuniformities. A significant decrease in nonuniformities was also observed when the grid cell size was reduced.

We performed a special experiment with a composite grid in order to more clearly establish a correlation between the beam current density and the density of the emitting plasma. The middle 6-cm-long part of the composite grid was a grid with a 0.8×0.8 mm mesh, and the remainder was a grid with a 0.4×0.4 mm mesh. The effect of the electron emission on the plasma density distribution in the hollow for this situation is illustrated by Fig. 3. Figure 4 gives the corresponding distributions of the current density $i(x)$ along the beam. It can be seen that, in the absence of emission, the nonuniformity of the plasma density along the hollow does not exceed 20%. At the same time, the electron emission results in a several-fold increase in the plasma density in the middle part of the hollow. The results presented in this figure clearly demonstrate the agreement between the positions of the maximum of the beam current density and the maximum of the plasma density in the hollow. We also note that, as the pressure decreases, the nonuniformities of the beam current density and the plasma density are smoothed out.

ANALYSIS OF THE RESULTS

The basic experimental results can be formulated as follows. In the absence of electron emission, the nonuniformity of the plasma density along the hollow does not exceed 10%. The extraction of electrons from the plasma at elevated pressures leads to the emergence of more significant nonuniformities in both the plasma and the electron beam. In this case, the spatial positions of the maxima of the emission current density and the density of the emitting plasma coincide. The nonuniformity of the current density of the electron beam exceeds

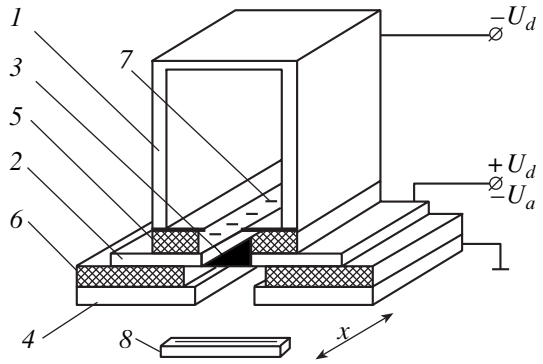

Fig. 1. Schematic of an electron source.

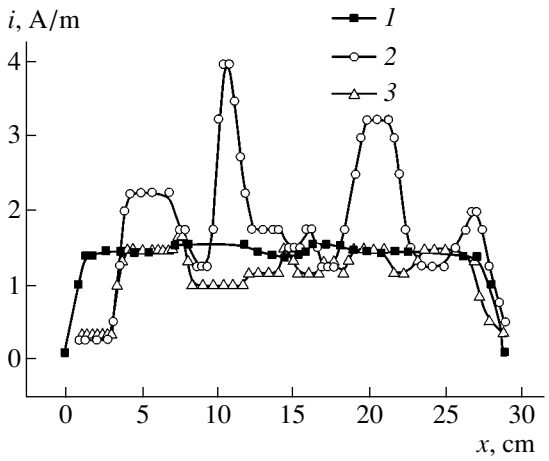

Fig. 2. Distribution of the electron current along the beam for pressures of p = (*1*, *3*) 4 and (*2*) 1 Pa and grid cell sizes of (*1*, *2*) 0.8×0.8 and (*3*) 0.4×0.4 mm.

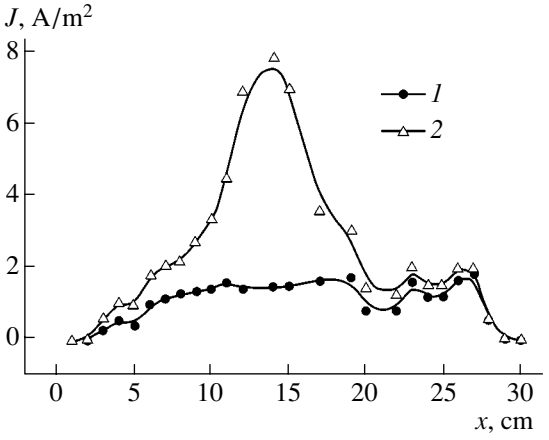

Fig. 3. Distribution of the probe current density in the anode plasma for accelerating voltages of U_a = (*1*) 0 and (*2*) 3 kV at a pressure of p = 4 Pa.

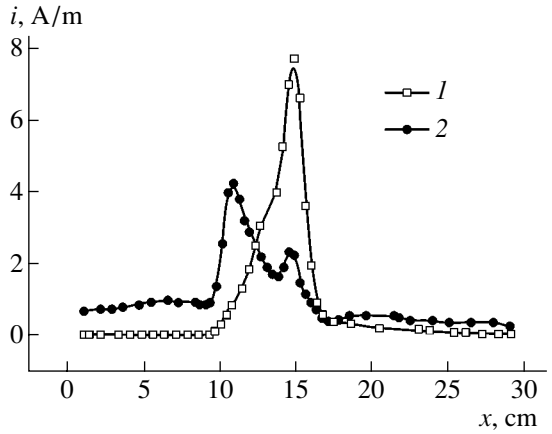

Fig. 4. Distribution of the current density along the beam for pressures of p = (*1*) 4 and (*2*) 2.4 Pa at U_a = 3 kV.

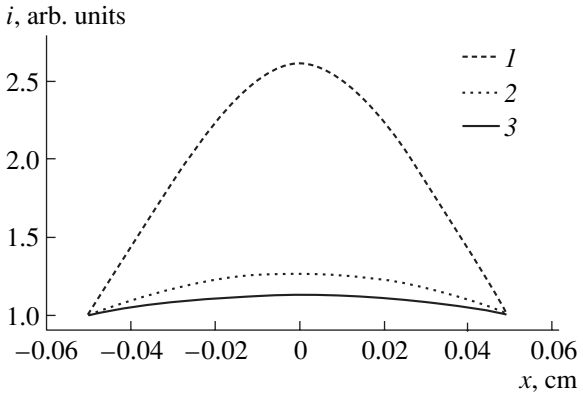

Fig. 5. Calculated dependence of the electron current density on the x coordinate in the perturbation region for pressures of p = (*1*, *2*) 6.6 and (*3*) 4 Pa and grid cell sizes of h = (*1*, *3*) 0.8×0.8 and (*2*) 0.4×0.4 mm.

the corresponding nonuniformity of the plasma density. The effect of emission on the nonuniformity of the current and plasma densities significantly reduces as the gas pressure and the size of the emission hole (grid cell) decrease. Based on the experimental results, we can assume the following mechanism for increasing the nonuniformity of the electron emission current. In the initial stage of the extraction of electrons from the plasma, the nonuniformity of the current density is mainly due to the nonuniformity of the plasma density in the hollow. The current density distribution may also be affected by the nonuniformity of the emission grid, namely, by the differences in the local curvature and the scatter in the sizes of elementary cells. The ionization of the residual gas in the accelerating gap and in the region of electron beam transport, which is significant in the forevacuum pressure range, leads to the emergence of a significant return flow of ions. Since the ionization rate is proportional to the electron current density, the density profile of the return ion flow must correspond to the initial distribution of the current density of the electrons emitted by the plasma. Fast ions, which get into the plasma and exchange charges with gas molecules, bring with them a positive space charge, which is neutralized by the plasma electrons. This results in a local increase in the nonuniformity of the plasma density and in the corresponding increase in the nonuniformity of the emission current. The increase in the emission current density with increasing plasma density is also due to the increase in the area of the open plasma surface within each cell of the anode grid because of the narrowing of the space charge layer separating the plasma from the grid. Therefore, a minor local variation in the plasma density results in a disproportionate increase in the local density of the electron emission current. The return ion flow associated with the electron current causes a further local increase in the plasma density and the corresponding further disproportionate increase in the density of the electron emission current at this site. This positive feedback reaches saturation and the plasma density ceases to increase when the formation rate of slow ions is balanced by their diffusion from the perturbation region.

In order to qualitatively estimate the possibility of the existence of a local maximum in accordance with the above mechanism, we will write balance equations for slow ions formed in the cathode hollow due to charge exchange of fast ions arriving at the plasma from the accelerating gap. The production of ions is balanced by their departure from the perturbed plasma region due to diffusion. In the one-dimensional case, the balance equation has the following form:

$$\frac{1}{4}\sqrt{\frac{8kT_e}{\pi m}}n_n^2 Q_e Q_r d\int_0^l \exp(-Q_r n_n y)dy \times \int_{-X_b}^{X_b} n(x)dx = -D_i \frac{dn(x)}{dx}\bigg|_{X_b} l, \quad (1)$$

where D_i is the ionic diffusion coefficient; $n(x)$ is the plasma density in the perturbation region; $\pm X_b$ are the coordinates of the boundaries of the perturbation region; $dn(x)/dx|_{X_b}$ is the gradient of the plasma density at the edge of the perturbation region; l is the depth of the hollow, which is less than the mean free path of charge-exchange ions, whereby a one-dimensional model may be employed; n_n is the density of neutral particles; Q_e is the effective cross section for the ionization of gas particles by electrons; d is the accelerating gap length; and Q_r is the effective cross section for the charge exchange of ions.

In order to take into account the variation in the area of the emitting surface, the coefficient $K(x)$ is intro-

duced in the left-hand side of Eq. (1),

$$K(x) = \frac{(h - 2l_s(x))^2}{h^2}, \quad (2)$$

where h is the inner size of the grid cell, $l_s(x) = 2\sqrt{\varepsilon_0 U^{3/2}(n(x)\sqrt{ekT_e})^{-1}}$ is the thickness of the ion layer separating the plasma from the grid [2], and U is the plasma potential relative to the anode.

It was assumed in the calculations that the dependence of the plasma density on the x coordinate within the perturbation region has the form of a Gaussian function. It follows from experiment that the value of X_b remains constant. The value of $dn(x)/dx|_{X_b}$ is determined from the condition of equality of the perturbed (n) and unperturbed (n_0) densities at $x = X_b$.

The numerical results presented in Fig. 5 demonstrate a local increase in the emission current density $i(x)$ with increasing both gas pressure and cell size of the emission grid. Therefore, the results of our calculations qualitatively confirm the possibility of the existence of local nonuniformities of the emission current due to the above physical mechanism.

CONCLUSIONS

A specific feature of the formation of a ribbon electron beam by a plasma source in the forevacuum pressure range is the high probability of the emergence of a nonuniformity in the distribution of the emission current along the beam. The results of our investigations indicate that this nonuniformity is caused by the return ion flow from the accelerating gap into the emitting plasma. The positive feedback that arises between the electron emission current and the return ion flow causes a disproportionate increase in the primary nonuniformity of the beam due to local nonuniformities of the plasma density, as well as to the local bending of the cells and difference in their sizes. The effect of the return ion flow may be reduced, e.g., by its defocusing accomplished by varying the grid shape in the emission electrode. This provides the generation of a ribbon electron beam with a nonuniformity of 10% or less. The investigation results provide one with more assurance in developing plasma electron sources generating electron beams in the forevacuum pressure range.

REFERENCES

1. W. M. Manheimer, R. F. Fersner, M. Lampe, *et al.*, Plasma Sources Sci. Technol. **9**, 370 (2000).
2. Yu. E. Kreĭndel', *Plasma Electron Sources* (Atomizdat, Moscow, 1977).
3. M. A. Zav'yalov, Yu. E. Kreĭndel', A. A. Novikov, and L. P. Shanturin, *Plasma Processes in Electron Guns* (Énergoatomizdat, Moscow, 1989).
4. E. M. Oks and P. M. Schanin, Phys. Plasmas **6**, 1649 (1999).
5. S. P. Bugaev, Yu. E. Kreĭndel', and P. M. Shchanin, *Wide Electron Beams* (Énergoatomizdat, Moscow, 1984).
6. Yu. A. Burachevskiĭ, V. A. Burdovitsin, E. M. Oks, *et al.*, Zh. Tekh. Fiz. **71** (2), 48 (2001) [Tech. Phys. **46**, 179 (2001)].
7. V. A. Burdovitsin, M. N. Kuzemchenko, and E. M. Oks, Zh. Tekh. Fiz. **72** (7), 134 (2002) [Tech. Phys. **47**, 926 (2002)].
8. V. A. Burdovitsin, Yu. A. Burachevskiĭ, E. M. Oks, *et al.*, Prib. Tekh. Éksp., No. 2, 1 (2003).

Translated by A. Bronshteĭn

Technical Physics, Vol. 49, No. 1, 2004, pp. 108–113. Translated from Zhurnal Tekhnicheskoĭ Fiziki, Vol. 74, No. 1, 2004, pp. 108–114.

**SURFACES,
ELECTRON AND ION EMISSION**

Electron Flow Enhancement with a Diamond Membrane

S. A. Gavrilov*, N. N. Dzbanovsky, É. A. Il'ichev*, P. V. Minakov**, É. A. Poltoratsky, G. S. Rychkov*, and N. V. Suetin****

** Lukin Research Institute of Physical Problems, Moscow, 124460 Russia*

e-mail: polt@niifp.ru

*** Nuclear Physics Institute, Moscow State University, Vorob'evy Gory, Moscow, 119899 Russia*

Received February 20, 2003

Abstract—Secondary electron emission from 2.5- to 5.0-μm thick diamond films (membranes) is considered. The process is studied in the reflection regime, where secondary electrons leave the front surface of the membrane exposed to primary electrons, and in the transmission regime, where primary electrons cause secondary emission from the opposite surface. The secondary emission coefficient is determined based on the behavior of 0.1- to 30-keV electrons in the solid. In the reflection regime, the secondary emission coefficient may be higher than 100 for electron energies of about 3 keV; in the transmission regime, it is no more than 5 even for 30-keV electrons. The emissivity of the membranes in the transmission regime can be improved, specifically, by using porous membranes, which allow one to obtain characteristics similar to those in the reflection regime. Experimental data obtained agree with calculations. The production of diamond films, including porous membranes, is described. © *2004 MAIK "Nauka/Interperiodica".*

INTRODUCTION

Bulk diamond and diamond films are promising materials for high-temperature and radiation-resistant electronics, as well as vacuum microelectronics. This is because diamond offers a unique combination of physicochemical properties: a high chemical and radiation stability plus an extremely high thermal conductivity and carrier mobility. Moreover, some of the diamond faces have a negative electron affinity (NEA); i.e., the energy of an electron in a vacuum is lower than its energy at the conduction band bottom. Because of this, electrons may be emitted from the conduction band into a vacuum in a thresholdless manner [1]. The discovery of NEA has given impetus to extensive research aimed at creating efficient cold (field) emitters based on diamond films. A large body of experimental data available today confirms the possibility of electron emission from polycrystalline diamond films; however, the emission is more likely to stem from intergranular regions than being related to the NEA effect [2]. The basic reason why the NEA-based concept of creating electron emitters has failed is the absence of electrons in the conduction band of diamond, since it has a wide energy gap and no appropriate donors have been found to date.

The NEA of diamond also shows up in an extremely high secondary electron emission (SEE) coefficient [3]. This effect may give rise to the development of high-efficiency vacuum microelectronic devices.

Such devices may be built on polycrystalline diamond membranes, which are capable of enhancing an electron flow with a given density distribution in the plane normal to the direction of electron motion.

It has been shown in many publications that, with the diamond surface processed properly, the SEE coefficient in the reflection regime may reach a high value, for example, 120 for 3-keV electrons (when the diamond film is covered by a nanometer Cs layer) [4]. Unfortunately, encouraging results have been obtained for only the reflection regime, where incident (primary) and secondary electrons are on the same side of the target. In one of the pioneering works concerned with electron emission in diamond films in the transmission regime [5], it was shown that a high SEE coefficient in this case is basically possible but practically hard to achieve. Physically, this may be related to the recombination of electron–hole pairs in the bulk of the film when they diffuse toward its rear.

To fabricate devices for 2D image processing like electron flow enhancers, it is of importance to preserve the electron density distribution in the plane normal to the flow direction. This is a challenge when electrons are multiplied in the reflection configuration. Therefore, the problem of the devices operating in the transmission regime remains quite topical. There are other difficulties associated with the fabrication of the devices, namely, diamond films patterning, the need for doping, and making low-resistance contacts. These three obstacles (and especially the absence of low-resistance contacts) prevent the production of industrial field emitters based on diamond films.

The aim of this work is to generate secondary emission in diamond films with a multiplication ratio much higher than unity in the transmission configuration. In Section 1, we consider the dynamics of fast electrons in membranes and estimate the SEE coefficient in both

regimes. In Section 2, the technology of diamond films and membranes is described. Section 3 is devoted to investigation techniques and presents experimental data obtained for void-free diamond films. Finally, in Section 4, we discuss ways of enhancing secondary emission in the transmission configuration.

1. DYNAMICS OF FAST ELECTRONS IN MEMBRANES

The basic element of a hypothetical electron flow enhancer is a membrane, which is exposed to fast primary electrons and emits secondary electrons from its surface layer. The operation of a solid-state electron flow enhancer may be elucidated by analyzing the electron dynamics in a solid. In similar devices, the energy of electrons usually lies in the range 0.1–40 keV. In this case, the electron velocity is much lower than the speed of light and the energy loss of an electron with an energy E per 1 m of electron travel can be expressed as [6]

$$-\frac{dE}{dl} = \frac{Ne^4 Z \ln 2 \ln E_1}{8\pi\varepsilon_0^2 E}. \quad (1)$$

Here, N is the number of scattering atoms in 1 m^3, e is the electron charge, Z is the nuclear charge, ε_0 is the dielectric constant of vacuum, E is the electron energy, $E_1 = E/\bar{E}$, and $\bar{E}$ is the mean energy spent on ionization. If a fast electron strikes the membrane at a right angle, the initial portion of its path in the material may be considered as rectilinear at least until its velocity drops markedly (by a factor of $\sqrt{2}$, as follows from estimates). From (1), one can estimate the maximal path length l_m of the electron:

$$l_\mathrm{m}[\mu\mathrm{m}] = \alpha\frac{3E_0^2[\mathrm{keV}]}{2\ln E_{10}}, \quad E_{10} = \frac{E_0}{\bar{E}},$$
$$\alpha = 6.2\times10^{-2}\frac{A[\mathrm{g}]}{Z\rho[\mathrm{g/cm^3}]}, \quad (2)$$

where E_0 is the electron initial energy, A is the atomic weight, and ρ is the density of the material.

For diamond, $\alpha = 3.54\times10^{-2}$. If it is assumed that the mean energy of ionization equals the second ionization potential E_2 (the latter is 16.3 and 24.4 keV for silicon and carbon, respectively, [7]), the maximal penetration depth for $E_0 = 20$ keV will be 3.2 and 4.5 μm in diamond and silicon, respectively. These estimates agree well with data in [8]. As the collision rate increases, the electron will progressively deviate from the initial direction. The probability that the electron will be scattered by an angle Θ is given by the error function [6]

$$p(\Theta) = \frac{1}{\sqrt{2\pi\lambda^2}}\exp(-\Theta^2/2\lambda^2), \quad \Theta \geq 0, \quad (3)$$

where

$$\lambda = \frac{\beta_0\sqrt{x[\mu\mathrm{m}]}}{\varphi[\mathrm{kV}]}, \quad \beta_0 = 4Z\sqrt{\frac{\rho[\mathrm{g/cm^3}]}{A[\mathrm{g}]}}. \quad (4)$$

Here, x is the thickness of the layer and φ is the accelerating voltage. For diamond, $\beta_0 = 13$.Using (3) and (4) and making several relevant simplifications, we find that, having entered into the membrane, the electron is inside a cone that is generated by rotating a curve $y(x)$ about the initial direction of the electron. The curve $y(x)$ is defined by the differential equation

$$\frac{dy}{dx} = \tan(\lambda_0\Phi^{-1}(\eta)), \quad \lambda_0 = \frac{\beta_0\sqrt{x[\mu\mathrm{m}]}}{E_0[\mathrm{keV}]}, \quad (5)$$

where $\Phi^{-1}(\eta)$ is the function reciprocal to $\Phi(x)$, $\Phi(x)$ is the probability integral, and η is a given probability that the electron is inside the cone (usually $\eta = 0.7$–0.8). The length of the curve $y(x)$ is roughly equal to l_m. Letting η tend to zero, we obtain a continuous set of cones whose bases lie on a convex surface bounding the initial cone. The initial cone ($\eta = 0.7$–0.8) and the convex surface bound the space where primary electrons generate most of the secondary electrons. The analysis may be simplified by replacing the cone by a sphere of radius $l_\mathrm{m}/3$ that is centered at a distance of $2l_\mathrm{m}/3$ from the membrane surface. With $\eta = 0.7$–0.8, this sphere will contain roughly half the number of secondary electrons. The enhancement coefficient will be the highest if secondary electrons arise as close to the emitting surface as possible. Clearly, the thickness of the membrane in the transmission configuration must not exceed l_m. In this case, the secondary electrons can approach the surface (by diffusion or drift), overcome the barrier (if any), and escape into a vacuum. Let us consider the dynamics of secondary electrons in greater detail. A secondary electron arises simultaneously with a hole. The time electron–hole recombination in diamond is $\tau_n = 10^{-9}$–10^{-8} s. Therefore, the electrons must leave the membrane for the time τ_n. If an internal field is absent, only electrons that are in a surface layer of thickness roughly equal to the electron diffusion length L_n may reach the surface. In diamond films, the mobility and lifetime of electrons are, respectively, $\mu_n = 0.1$–1.0 cm^2/(V s) and 10^{-9}–10^{-8} s. Hence, $0.016 \leq L_n \leq 0.16$ μm. Then, if an internal field is absent and $l_\mathrm{m} \gg L_n$, the emitting surface is bound to lie near the center of the sphere mentioned above and the thickness of the membrane should be taken to be roughly equal to $2l_\mathrm{m}/3$. If the secondary electrons are uniformly distributed over the sphere and the thickness of the membrane is $h = 2l_\mathrm{m}/3$, the enhancement coefficient upon transmission is $K_D \approx 3L_nE_0/4h\bar{E}$. With $h = 3$ μm and $E_0 = 20$ keV, K_D is no higher than 33 for the best films with $L_n = 0.16$ μm, whereas for typical films with $L_n = 0.016$ μm, K_D is smaller than 3.3. At the same time, if secondary electrons are emitted by the surface exposed to primary

electrons, the enhancement coefficient upon reflection is estimated as $K_D \approx E_0 \min(3L_n/2l_m, 1)/\bar{E}$. For an electron energy much lower than needed in the transmission configuration, this estimate yields a much higher value of K_D. For example, with $E_0 = 3$ keV, we find that K_D may reach 120 for the best films and remains sufficiently high for typical films. At energies below 3 keV, it amounts to 30. A significant advantage of the reflection configuration is that no fundamental restrictions are imposed on the membrane thickness.

2. DIAMOND MEMBRANE TECHNOLOGY

According to the above estimates for the transmission configuration, the operating voltage must be as high as 30 kV for the films 3 to 4 μm thick. The films must be prepared in the form of void-free or porous membranes. The membranes are formed on the silicon substrate surface. The technology used in this work allows for the local nucleation and growth of a diamond film on the silicon surface covered with a photoresist containing nanodisperse diamond particles. Substrates were phosphorus-doped single-crystal (100)Si wafers with a resistivity of 2.5–4.5 Ω cm. Prior to diamond film formation, the wafers were subjected to standard cleaning in organic and inorganic solvents. Then, a photoresist film was applied on the substrate by spinning. The mean size of diamond grains contained in the photoresist was less than 0.1 μm. The volume content of the diamond powder in the photoresist was varied from 10 to 50%. After drying at 80°C for 10 min, the resist was exposed to the radiation from a mercury-vapor lamp and developed in a 0.5% aqueous solution of KOH. In the case of void-free films, exposure and development were not carried out. Diamond films were deposited by microwave-plasma-assisted CVD in several stages.

At the preparatory stage, the resonator was evacuated to a pressure of 10^3 Pa and then filled with hydrogen to $(5–6) \times 10^3$ Pa. The hydrogen flow rate was 10 l/h

Before deposition, the sample surface was cleaned of organic contaminants. For this purpose, the sample was kept in a hydrogen plasma for about 15 min at $(5–6) \times 10^3$ Pa and annealed 650–700°C.

The annealing stage graded into the stage of deposition. Good diamond films were obtained when the microwave power was 1–2 kW; the pressure in the reactor, $(9–10) \times 10^3$ Pa; the sample temperature, 750–800°C; and the ethanol vapor content in the hydrogen atmosphere 10–15%. Good doped films were obtained under the same conditions except that a trimethylborate–boric alcohol mixture, instead of ethanol, was poured into the evaporator. The percentage of trimethylborate is usually varied from 0.2 to 0.8% depending on the degree of doping of the diamond film.

The sample with the film deposited was annealed again under nearly the same conditions as for the predeposition annealing.

The annealed sample with the film was covered by a chemically inactive material. A part of the rear side of the substrate that specified the membrane dimensions was left unprotected. Then, the silicon was chemically etched in an $HF : HNO_3 = 2 : 1$ (by volume) solution for 60 min until the diamond films appeared, as detected with an optical microscope. At the final stage, the protective mask was dissolved in a boiling solution of trichloroethylene and the entire structure was cleaned in the isopropyl alcohol vapor for 10 min.

3. SECONDARY EMISSION IN VOID-FREE FILMS

1. Measuring bench and techniques for studying the electrophysical parameters of the films. A measuring bench was built around a CamScan scanning electron microscope [9] and comprised (i) an electron gun with set of electromagnetic lenses, which make it possible to form collinear electron beams with a diameter of less than 1000 Å on the film surface and scan a 5×5- to 1000×1000-μm area with a frequency of 0–100 Hz at beam energies from 0.5 to 30.0 keV; (ii) a sample holder (arm), which allows one to change the sample position relative to the electron beam; (iii) a contactor, a switch, and power supplies, which control measuring conditions and apply an electric potential in the range 0–1500 V to the sample; (iv) an evaporating system, which includes a Cs evaporator and power supply; (v) a two-stage differential current amplifier, which handles currents from 10^{-12} to 10^{-7} A at frequencies varying between 0 and 100 Hz; and (vi) a recording and display system, which consists of an electrometric amplifier, digital voltmeter, plotter, oscillograph, and storage oscilloscope.

Preliminary studies of polycrystalline diamond films, graphite films, and amorphous carbon films have shown that noticeable secondary emission in these materials, unlike those intended for field-emission cathodes, takes place when the films have mostly sp^3 bonds. That is the reason why CVD was the method of choice for depositing diamond films. The phase composition was confirmed by Raman spectroscopy (sp^3 bonds prevail) and electron diffraction in a transmission electron microscope (phase microanalysis). To suppress charge effects during measurements, the growing films were doped by boron. Hall measurements on free diamond films showed that the hole concentration in the boron-doped films might reach 5×10^{18} cm^{-3}. The free carrier mobility was no higher than 1 cm^2/(V s).

2. Secondary emission measurement in the transmission and reflection geometry. Typical dependences of the secondary emission coefficient on the primary electron energy for void-free diamond films upon transmission and reflection are given in Figs. 1 and 2.

The secondary emission coefficient K_D is defined as the ratio between secondary and primary electron currents, which are measured with a Faraday cup. The SEE current may also include elastically and inelastically scattered particles.

For films not subjected to any additional processing, the maximum of K_D in the reflection configuration was, as a rule, in the primary energy range 800–900 eV (Fig. 1). In this range, K_D varied between 5 and 20 and depended significantly on the membrane surface condition.

In the transmission geometry, the maximum shifts to 25–35 keV, equals 2–3, and also depends on the surface condition. These observations were made for the case when the secondary electron yield was measured on the front side of the film, which did not contact with the substrate during growth. When primary electrons were injected from the opposite side, the SEE coefficient was one order of magnitude lower because of a large density of defects and grain boundaries on that surface. This strengthens the supposition that a mechanism of electron transport toward the membrane surface is of fundamental importance.

To gain a better insight into the effect of surface condition on the SEE coefficient, the films were subjected to additional annealing and cesium deposition. The surface of some of the diamond films grown was additionally hydrogenated by means of hydrogen annealing. The vacuum annealing of these samples at 700°C led to a decrease in K_D. This detrimental effect was enhanced upon subsequent annealings probably because of the diamond surface dehydrogenation. These findings are well consistent with published data [10].

Cesium applied on the diamond films hydrogenated raised the secondary electron yield two- or threefold.

3. Discussion. Comparing our experimental data, as well data obtained by other authors, with the model (see Section 1) shows that the model provides an adequate description of the secondary electron dynamics. Figure 3 plots the dependences of K_D on the primary electron energy [10] and our data for the reflection geometry. Peaks in our curves suggest that the maximal penetration depth of primary electrons is roughly equal to the diffusion length, and the energy shift of the maximum after cesium deposition indicates that the work function decreases compared with the hydrogenated films.

Since the values of K_D for our films and those studied in [10] are close to each other with primary electron energies up to 1 keV, it may be argued that the surface condition is also nearly the same. The discrepancy at energies above 1 keV may be explained by the fact that the diffusion lengths in the films used in [10] were almost three times longer than those in our films. Accordingly, the electrons were delivered to the surface from greater depths. Such a reason seems plausible, since films in [10] had a thickness of ≈20 μm and, hence, a grain size greater than in our films (3 to 5 μm thick). A lesser effect of grain boundaries may, in our opinion, be responsible for the wider energy range of linear growth of K_D in single crystals compared to the linear range in polycrystalline films [10].

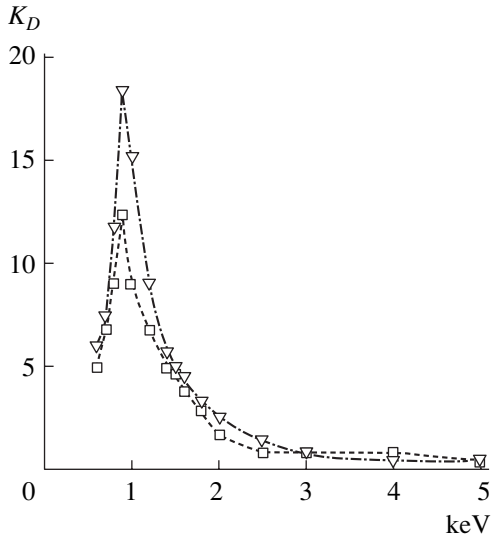

Fig. 1. SEE coefficient vs. the primary electron energy in the reflection configuration for two samples.

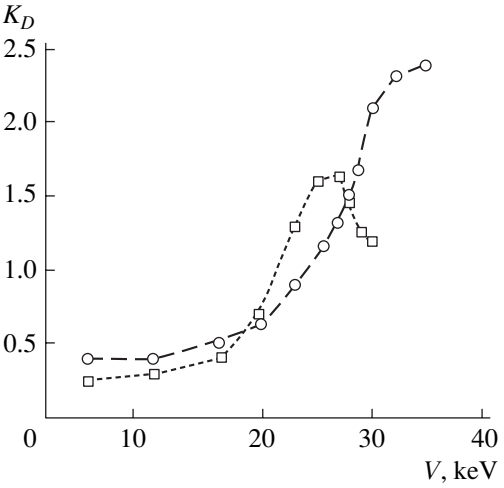

Fig. 2. SEE coefficient vs. the primary electron energy in the transmission configuration for the (□) 2.5-μm- and (○) 5-μm-thick membranes.

In the transmission configuration, K_D is much smaller (than in the reflection geometry), because the diffusion mechanism of secondary electron delivery to the surface is inefficient, as follows from the estimates in Section 1.

4. POTENTIALITIES FOR SECONDARY EMISSION ENHANCEMENT IN THE TRANSMISSION GEOMETRY

1. Secondary emission in void-free membranes in the presence of an internal sweeping field. We have

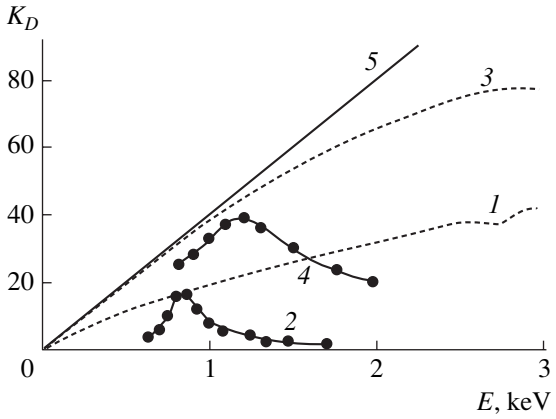

Fig. 3. Results of this work versus those obtained in [10]: (*1*, *3*) hydrogenated and cesium-covered films from [10], (*2*, *4*) hydrogenated and cesium-covered membranes from this work, and (*5*) theoretical curve (Section 1).

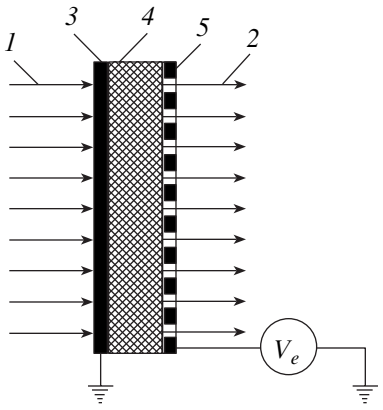

Fig. 4. Optimized design of the membrane for the transmission configuration: (*1*, *2*) incident and outgoing electron flows, respectively; (*3*) thin heavily doped layer of p^+-Si; (*4*) diamond membrane; and (*5*) metallic grid applied on the membrane.

analyzed the case when a field in the membrane was absent. An internal field may be generated with an external electrode (shield) that visualizes the electron flow enhanced. Consider an undoped diamond film. In undoped diamond structures, the field may be considered as coordinate independent. Using an appropriate continuity equation and assuming that the generation rate k of electrons is uniform across the film, we find for the electron distribution

$$n = k\tau_n(1 - \exp(-x/\mu_n E_s \tau_n)), \quad (6)$$

where E_s is the strength of the field generated by the external electrode to which a voltage V_e is applied. The $0x$ axis is normal to the film, and its origin is on that surface subjected to primary electrons. Electron recombination is practically absent if

$$\frac{h}{\mu_n \tau_n} \ll E_s, \quad (7)$$

where h is the membrane thickness; then, $n \approx kx/\mu_n E_s$.

It is natural that (6) is valid if the concentration of nonequilibrium carriers may be neglected. It may be neglected if (7) is met and mobile carriers are effectively removed from the ends of the film. For the electrons, the latter condition is readily satisfied because of negative electron affinity (the electrons are free to escape into a vacuum). The holes can be removed if a good ohmic contact for them is formed (e.g., by applying a thin layer of p^+-Si on the membrane surface). The concentration p^+ of equilibrium holes must many times exceed the concentration p of nonequilibrium holes in the diamond. If the external electrode is at a distance d_e from the membrane surface, we have

$$E_s = \frac{V_e}{(h + d_e\varepsilon)}, \quad (8)$$

where ε is the permittivity of diamond.

Since $\varepsilon d_e \gg h$, condition (7) may be recast as

$$V_e \gg \frac{\varepsilon h d_e}{\mu_n \tau_n}. \quad (9)$$

With $\mu_n = 1$ cm^2/(V s), $\tau_n = 10^{-8}$ s, $h = 2$ μm, $d_e = 1$ mm, and $\varepsilon = 4.5$, we have $V_e \gg 10^4$ V. Such voltages are unrealistic in real devices. Therefore, a shield cannot be used as an electrode for generating a field in the membrane. It is necessary that an accelerating electrode be placed in the immediate vicinity of the membrane ($d_e \le 10$ μm). The optimal design of an electron flow enhancer is that where the electrode is formed in the form of a grid on the membrane surface that does not degrade the resolving power of the enhances (Fig. 4). Then, the condition $V_e \gg h^2/\mu_n\tau_n$ would suffice. With $\mu_n = 1$ cm^2/(V s), $\tau_n = 10^{-8}$ s, and $h = 2$ μm, we find that $V_e \gg 4$ V. This condition is readily satisfied. As an electrode, one can also use a thin (0.2–0.5 μm) p^+ diamond layer. Experimental implementation of such a structure will be possible if the problems of diamond film surface planarization and good ohmic contacts are resolved.

2. Secondary emission in porous membranes. Another way of enhancing secondary emission is the use of a porous membrane. In this case, the secondary electrons are not emitted from the surface exposed to primary electrons, as in the reflection configuration; instead, if the energy of secondary electrons is sufficiently low, they travel through the voids toward the opposite surface under the action of an external field. Certainly, the multiplication ratio K_n will be somewhat smaller than in the reflection geometry, since some of the primary electrons pass through the voids without generating secondary particles. In this case, $K_n = (1 - S_n/S)K_D$, where S_n and S are the total area occupied by voids and the membrane operating area, respectively, and K_D is the SEE coefficient in the reflection geometry.

To check this idea, we estimated the secondary electron energy and fabricated porous membranes. The sec-

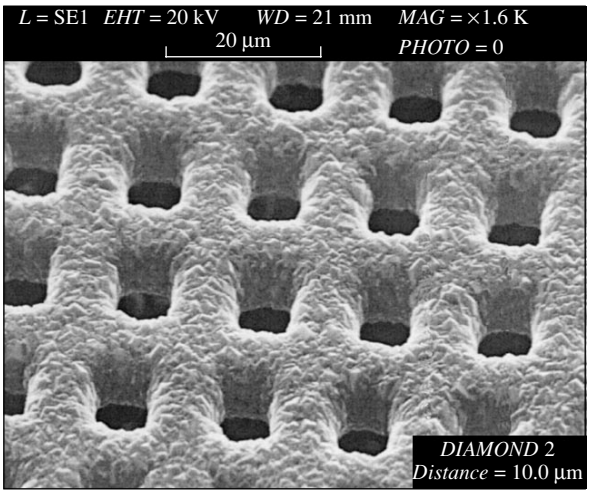

Fig. 5. SEM image of a porous diamond film.

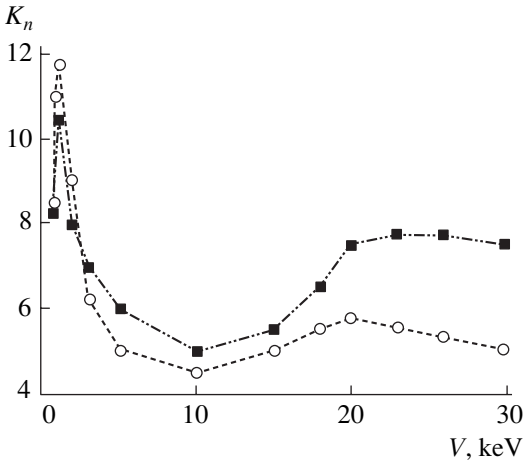

Fig. 6. SEE coefficient vs. the primary electron energy for the porous film in the (○) reflection and (■) transmission configurations.

ondary electron energy was directly measured to be 3 to 5 eV.

Using the technology described in Section 2, we fabricated porous diamond membranes with voids of size 8 μm and a void spacing of 8 μm (Fig. 5). Figure 6 plots the SEE coefficient throughout the energy range for the same sample in both configurations. Unlike Fig. 2, here two peaks corresponding to the reflection and transmission geometries are observed. The increase in K_D upon transmission is probably due to enhanced electron multiplication in the voids; however, this assumption needs further corroboration.

The design suggested is of great practical importance. It loosens drastically the dependence of the multiplication ratio on the membrane thickness; makes it possible to use low-energy electrons in the transmission configuration; and raises the mechanical strength of the membranes, removing stresses that arise in the diamond film growing on the silicon substrate.

CONCLUSIONS

In boron-doped polycrystalline diamond films, the SEE coefficient much higher than unity in the transmission configuration is obtained for the first time. The effect of processing of the emitting surface (the application of cesium, hydrogenation, and annealing) on the SEE coefficient is studied. It is shown that the size and structure of grains are of fundamental importance for achieving high SEE coefficients. The model of secondary electron behavior suggested in this work fits well the experimental data. The transmission configuration with porous membranes is first implemented. With the porous membrane, the electrons generated on the front surface can be effectively directed toward the opposite surface of the membrane.

REFERENCES

1. F. J. Himpsel, J. A. Knapp, J. A. van Vechten, *et al.*, Phys. Rev. B **20**, 624 (1979).
2. A. T. Rakhimon, N. V. Suetin, E. S. Soldaton, *et al.*, J. Vac. Sci. Technol. B **18** (1), 76 (2000).
3. G. T. Mearini, I. L. Krainsky, and J. A. Dayton, Surf. Interface Anal. **21**, 138 (1994).
4. G. T. Mearini, I. L. Krainsky, J. A. Dayton, *et al.*, Appl. Phys. Lett. **66**, 242 (1995).
5. N. N. Dzbanovsky, P. A. Minakov, A. F. Pal, *et al.*, in *Proceedings of the 3rd International Conference on Vacuum Electron Sources "IVESC–2000," Orlando, 2000*, p. 33.
6. *Physics Pocket Book*, Ed. by H. Ebert (Oliver & Boyd, Edinburgh, 1967; Nauka, Moscow, 1973).
7. P. Atkins, *Quanta: A Handbook of Concepts*, 2nd ed. (Oxford Univ. Press, Oxford, 1991; Mir, Moscow, 1977).
8. T. E. Evehart and P. H. Hoff, J. Appl. Phys. **42**, 5837 (1971).
9. N. N. Dzbanovsky, P. A. Minakov, A. F. Pal, *et al.*, in *Proceedings of the International Topical Meeting on Field Electron Emission from Carbon Materials, Moscow, 2001*.
10. J. E. Yater, A. Shih, and R. H. Abrams, J. Vac. Sci. Technol. A **16**, 913 (1998).

Translated by V. Isaakyan

Technical Physics, Vol. 49, No. 1, 2004, pp. 114–118. Translated from Zhurnal Tekhnicheskoĭ Fiziki, Vol. 74, No. 1, 2004, pp. 115–119.

EXPERIMENTAL INSTRUMENTS AND TECHNIQUES

Controlling the Diffusion Process Via Time-Variable Diffusion Coefficient

E. L. Pankratov

Lobachevsky State University, pr. Gagarina 23, Nizhni Novgorod, 603950 Russia

e-mail: elp@rf.unn.ru

Received January 28, 2003

Abstract—The establishment of the steady-state dopant profile in a medium with a time-variable diffusion coefficient is considered within the approach proposed previously for estimating mass- and heat-transfer time characteristics. It is shown that the time it takes for the equilibrium concentration to set in may be increased or decreased by appropriately choosing the law of variation of the diffusion coefficient. © *2004 MAIK "Nauka/Interperiodica".*

High-temperature diffusion of doping impurities is a basic step in the production of semiconductor devices [1]. However, studies of how the variation of the diffusion coefficient during heating, cooling, and other processes that take place in semiconductor devices [1, 2] affects the diffusion process are virtually lacking. The aim of this study is to estimate quantitatively the effect of this variation on the establishment of the steady-state doping profile in the semiconductor technology. Such an analysis is intended to improve the reproducibility of the parameters of devices subjected to multiple annealing. In addition, whether or not the time characteristics of diffusion can be described adequately in terms of the averaged parameters of the sample is discussed.

STATEMENT OF THE PROBLEM

Let us consider a one-dimensional homogeneous sample of thickness L with a time-variable diffusion coefficient. An impurity of unit mass, $\int_0^L C(x, t)dx = 1$, with an initial concentration profile $C(x, 0) = f(x)$ starts diffusing into the medium at a time t (which is taken to be the zero time $t = 0$ for simplicity). With time, the distribution of the impurity becomes stationary, $C(x, \infty) = 1/L$. Our aim is to determine the time taken to establish the steady-state concentration at a given point $x \in [0, L]$ (point of observation).

THE SOLUTION TECHNIQUE

The space–time dopant distribution $C(x, t)$ is described by the diffusion equation [1, 3]

$$\frac{\partial C(x, t)}{\partial t} = D(t)\frac{\partial^2 C(x, t)}{\partial x^2} = -\frac{\partial G(x, t)}{\partial x}, \qquad (1)$$

where $G(x, t)$ is the dopant flow. The diffusion equation should be complemented by the initial, $C(x, 0) = f(x)$, and boundary, $G(0, t) = G(L, t) = 0$, conditions. Since the time dependence of the transient described by Eq. (1) is rather complicated, it is, in general, impossible to estimate quantitatively the transient period by directly solving this equation. In the situation of most practical interest when the dopant is initially concentrated at one of the boundaries, $f(x) = \delta(x - 0)$, and the transient period is determined at a point on the opposite boundary, the concentration varies with time roughly by the exponential law. The initial delta profile of doping corresponds to an impurity source deposited on a sample wall [3, 4]. Note that, in transient period calculation, the delta function should be considered as one-sided, $\delta(x - 0)$ [5]. If the concentration varies exponentially, the transient period can be conveniently determined from the well-known [6–9] asymptotically optimal [10] integral criterion

$$\Theta(x) = [C(x, 0) - C(x, \infty)]^{-1} \times \int_0^\infty [C(x, t) - C(x, \infty)]dt. \qquad (2)$$

To calculate the transient period, it is necessary to find the space–time dopant distribution $C(x, t)$ satisfying Eq. (1). Let us represent a solution to the diffusion equation as an expansion in eigenfunctions [11]:

$$C(x, t) = \frac{1}{L} + \frac{2}{L}\sum_{n=1}^{\infty} \cos\frac{\pi n x}{L} \times \int_0^L f(\nu)\cos\frac{\pi n \nu}{L}d\nu \exp\left\{-\left(\frac{\pi n}{L}\right)^2 \int_0^t D(u)du\right\}. \qquad (3)$$

Substituting solution (3) into Eq. (2) yields an exact expression for the transient period (the time the concen-

1063-7842/04/4901-0114$26.00 © 2004 MAIK "Nauka/Interperiodica"

tration takes to reach the steady-state distribution):

$$\Theta(x) = \frac{2\sum_{n=1}^{\infty}\cos\frac{\pi n x}{L}\int_0^L f(\nu)\cos\frac{\pi n\nu}{L}d\nu\int_0^{\infty}\exp\left\{-\left(\frac{\pi n}{L}\right)^2\int_0^t D(u)du\right\}dt}{Lf(x)-1}. \quad (4)$$

The fact that transient period (4) depends on the law of variation of the diffusion coefficient complicates the study of the diffusion dynamics. Let us first restrict our analysis to the case when the diffusion coefficient varies insignificantly. Such an approach makes the analysis of diffusion clearer and, at the same time, allows one to calculate the asymptotic transient period in the case of a rapidly varying diffusion coefficient.

TRANSIENT PERIOD AT SMALL VARIATION OF THE DIFFUSION COEFFICIENT

Under the assumption that the diffusion coefficient varies insignificantly, the diffusion process is almost completely characterized by the mean value D_0 of the diffusion coefficient. Following [9, 12], we represent the time variation of the diffusion coefficient $D(t)$ as the sum of its mean value and a small deviation from this value: $D(t) = D_0[1 + \mu\nu(t)]$, $0 \le \mu \ll 1$. Since the parameter μ, which characterizes the amplitude of variation of the diffusion coefficient, is small and the values of $\nu(t)$ are limited, $|\nu(t)| \le 1$, we may apply the Poincaré method and seek for the solution $C(x, t)$ in the form of an expansion in powers of μ:

$$C(x,t) = \sum_{k=0}^{\infty}\mu^k C_k(x,t). \quad (5)$$

As criterion (2) is linear in impurity concentration $C(x, t)$, the transient period can be calculated by using the principle of superposition. In this case, the transient period may be represented as a series in powers of the parameter μ:

$$\Theta(x) = \Theta_0(x)\left(1 + \sum_{k=1}^{\infty}\mu^k\tau_k(x)\right), \quad (6)$$

where the zeroth-order approximation of the transient period is given by

$$\Theta_0(x) = [C(x,0) - C(x,\infty)]^{-1} \times \int_0^{\infty}[C_0(x,t) - C(x,\infty)]dt. \quad (7)$$

The factors multiplying a kth power ($k \ge 1$) of the small parameter μ in Eq. (6),

$$\tau_k(x) = \left[\int_0^{\infty}[C_0(x,t) - C(x,\infty)]dt\right]^{-1}\int_0^{\infty}C_k(x,t)dt \quad (8)$$

are normalized corrections to the transient period.

The zeroth-order approximation of the transient period, which corresponds to a time-invariable diffusion coefficient D_0, and the first normalized correction to it are easy to find:

$$\Theta_0(x) = \frac{2}{D_0}\left(\frac{L}{\pi}\right)^2\frac{\sum_{n=1}^{\infty}n^{-2}\cos\left(\frac{\pi n x}{L}\right)\int_0^L(\nu)\cos\left(\frac{\pi n\nu}{L}\right)d\nu}{Lf(x)-1},$$

$$\tau_1(x) = -D_0\left(\frac{\pi}{L}\right)^2\frac{\sum_{n=1}^{\infty}n^{-2}\cos\left(\frac{\pi n x}{L}\right)\int_0^L f(u)\cos\left(\frac{\pi n u}{L}\right)du\int_0^{\infty}\nu(u)\exp\left\{-\left(\frac{\pi n u}{L}\right)D_0 u\right\}du}{\sum_{n=1}^{\infty}n^{-2}\cos\left(\frac{\pi n x}{L}\right)\int_0^L f(u)\cos\left(\frac{\pi n u}{L}\right)du}. \quad (9)$$

Since μ is a small parameter and the modulus of $\nu(t)$ is no more than unity, the terms of the second and higher orders in series (5) and (6) may be neglected (linear approximation). The linear approximation is necessary for the visualization of the main features of the diffusion dynamics in a medium with time-varying parameters.

When calculating the correction $\tau_1(x)$, we, for simplicity, approximate the diffusion coefficient law $\nu(t)$ by using the functions $\cos\omega t$ and $\sin\omega t$. If the initial concentration profile has the form of the delta function $f(x) = \delta(x - 0)$ and the point of observation is located at the opposite boundary $x = L$, the first relative corrections $\tau_{1\cos}(L)$ and $\tau_{1\sin}(L)$ to the transient period, which

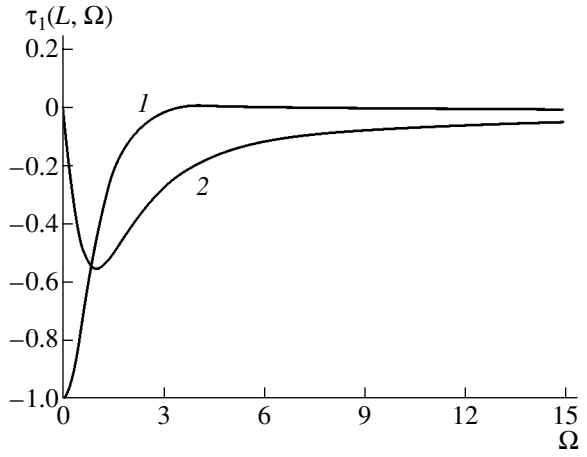

Fig. 1. Frequency dependences of the corrections (*1*) $\tau_{1\cos}(L)$ and (*2*) $\tau_{1\sin}(L)$.

correspond to the harmonic variation of the diffusion coefficient, are given by

$$\tau_{1\cos}(L, \Omega) = \frac{12}{\pi^2}\sum_{n=1}^{\infty}\frac{(-1)^n n^2}{\Omega^2 + n^4}, \quad (10)$$

$$\tau_{1\sin}(L, \Omega) = \frac{12}{\pi^2}\sum_{n=1}^{\infty}\frac{(-1)^n \Omega}{\Omega^2 + n^4}, \quad (11)$$

where $\Omega = \omega L^2/\pi^2 D_0$ is the dimensionless frequency of variation of the diffusion coefficient.

The calculation of the corrections creates no difficulties; the alternating series obtained converge strongly. The first corrections versus the dimensionless frequency of variation of the diffusion coefficient are presented in Fig. 1.

When the diffusion coefficient varies by the cosine law, the frequency dependence is nearly monotonic and the modulus of the function reaches a maximum at the zero frequency (Fig. 1, curve *1*). This corresponds to an increase in the mean diffusion coefficient by μD_0. Negative sign of the correction indicates that the transient period shrinks as the diffusion coefficient grows. The frequency range where the variation of the diffusion coefficient has a noticeable effect on the transient period at a level of 3 dB is limited from above by the value $\Omega_c = 0.636$. The correction decreases with increasing frequency of variation of the diffusion coefficient and becomes sufficiently small in the range $\Omega >$ 3–4, so that the averaged diffusion coefficient can describe the process with a reasonable accuracy.

When the diffusion coefficient varies by the sine law, the frequency dependence has a resonant shape (Fig. 1, curve *2*). A similar curve for the transient period was obtained in a constant-diffusion-coefficient medium subjected to a periodic external action [13]. The absolute value of the correction $\tau_{1\sin}(L, \Omega)$ peaks at $\Omega_0 = 0.977$. An appreciable gain in time is only possible if the stationary profile is established within the positive half-period of variation of the diffusion coefficient. The change of sign in the diffusion coefficient law slows down the process and the transient period grows. For frequencies $\Omega > 1$, the effect of the correction is weak and the description of the diffusion dynamics in terms of the averaged diffusion coefficient is adequate.

The linearity of Eq. (9) in correction function $\nu(t)$ allows us to apply spectral analysis methods. The expansion of the diffusion coefficient function in the Fourier series in harmonic functions,

$$\nu(t) = \sum_{k=1}^{\infty} a_k \cos\omega_k t + \sum_{k=1}^{\infty} b_k \sin\omega_k t \quad (12)$$

makes it possible to represent the total correction as a sum of partial corrections:

$$\tau_1(x) = \sum_{k=1}^{\infty} a_k \tau_{1\cos}(x, \omega_k) + \sum_{k=1}^{\infty} b_k \tau_{1\sin}(x, \omega_k),$$

where $\tau_{1\cos}(L, \omega_k)$ and $\tau_{1\sin}(L, \omega_k)$ are the normalized corrections to the transient period when the diffusion coefficient varies harmonically as $\cos\omega_k t$ and $\sin\omega_k t$, respectively.

The constant component of the function $\nu(t)$ is zero by virtue of the condition imposed previously. As is evident from the above discussion, the additivity of the corrections facilitates substantially the calculation of the transient period $\tau_1(x)$.

EXAMPLE OF CALCULATION

The diffusion coefficient in a solid depends exponentially on the temperature [3, 14]:

$$D(T) = D\exp[-E/kT(t)],$$

where D and E are material constants, $T(t)$ is the temperature, and k is the Boltzmann constant.

Upon heating or cooling, the sample temperature varies exponentially provided that the thermal conductivity is constant. With this restriction, the time variation of the diffusion coefficient in a sample subjected to single doping via high-temperature diffusion is shown in Fig. 2 (curve *1*). Taking into account that the thermal conductivity is a time-variable parameter leads to somewhat more complicated time dependences of the temperature and diffusion coefficient.

For long-term annealing (the annealing time t_0 exceeds the time constant of heating the sample several-fold), the diffusion coefficient law $\nu(t)$ can be closely approximated by the second-order Walsh function wal(2, t) [15] (Fig. 2, curve *2*). The corresponding

correction $\tau_{1\text{wal}(2,\,t)}(L)$ has the form

$$\tau_{1\text{wal}(2,\,t)}(L) = \frac{12}{\pi^2} \times \sum_{n=1}^{\infty} \frac{(-1)^n}{n^2} e^{-\frac{\pi n^2}{2\Omega}} \left\{ \tanh\left[\frac{\pi n^2}{2\Omega}\right] + 1 \right\} - 1. \quad (13)$$

The frequency dependence of such a correction is shown in Fig. 3 (curve *6*). An advantage of approximation by symmetric or near-symmetric functions is evident: the diffusion coefficient can be expanded in an alternating strongly convergent series in cosines of multiple frequencies. Even the first term in the Fourier expansion of the Walsh function in cosines gives a fairly accurate approximation of the frequency dependence of correction (13) to the transient period (Fig. 3, curve *1*). Curves *2–5* show the successive improvement of the approximation accuracy as the second, third, fourth, and fifth term, respectively, in the Fourier expansion of the Walsh function is taken into account. The contribution of the harmonics drops sharply with increasing frequency, so that the first spectral component alone approximates the transient period with an acceptable accuracy for $\Omega > 0.5$. The slight asymmetry of curve *1* in Fig. 2 is accounted for by the sine components of the expansion; their relative contribution is small and decreases rapidly with an increase in annealing time.

TRANSIENT PERIOD AT LARGE VARIATION OF THE DIFFUSI

The exact expression for transient period (4) with the initial dopant profile $f(x) = \delta(x - 0)$ and the point of observation $x = L$ considered above has the form

$$\Theta(L) = 2\sum_{n=1}^{\infty} (-1)^{n+1} \int_0^{\infty} \exp\left\{ -\left(\frac{\pi n}{L}\right)^2 \int_0^t D(u)\,du \right\} dt. \quad (14)$$

From Eq. (14), we can determine the applicability range of the linear approximation. For the diffusion coefficient law adopted, it was found that the linear approximation of the transient period and concentration is valid up to $\mu \approx 0.1$ and 0.15, respectively. With μ taken at the upper limit of the applicability range, the maximal variation of the transient period is ≈10% of $\Theta_0(L)$. For stronger variations, $\mu \approx 1.0$, one should take into consideration the terms nonlinear in μ in expansion (6) and use the exact expression for the transient period. Let us calculate the transient period for the annealing process (which is of most practical interest) in the case when the diffusion coefficient law has the form of the second-order Walsh function. Then, transient period (14) is recast as

$$\Theta(L) = \frac{L^2}{6D_0(1+\mu)} + \frac{2L^2}{\pi^2 D_0} \sum_{n=1}^{\infty} \frac{(-1)^n}{n^2} \left\{ \frac{1}{1+\mu} \exp\left[-(1+\mu)\frac{\pi n^2}{2\Omega}\right] - \frac{1}{1-\mu} \exp\left[-(1-\mu)\frac{\pi n^2}{2\Omega}\right] \Big/ \left[1 + \exp\left(-(1-\mu)\frac{\pi n^2}{2\Omega}\right)\right] - \frac{1}{1+\mu} \exp\left[-(3-\mu)\frac{\pi n^2}{2\Omega}\right] \Big/ \left[1 + \exp\left(-(1+\mu)\frac{\pi n^2}{2\Omega}\right)\right] \right\}. \quad (15)$$

It is of interest to see how the transient period depends on the degree of variation μ of the diffusion coefficient. According to our analysis, one may use the

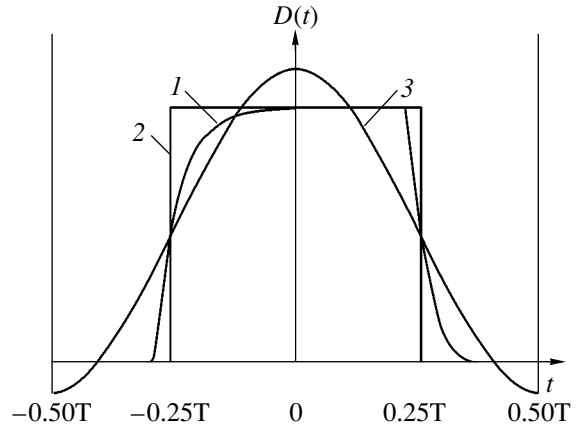

Fig. 2. (*1*) Time dependence of the diffusion coefficient. The approximation of the diffusion coefficient law by the sum of its mean value and the first spectral component of the function $v(t)$: (*2*) the second-order Walsh function and (*3*) a harmonic function.

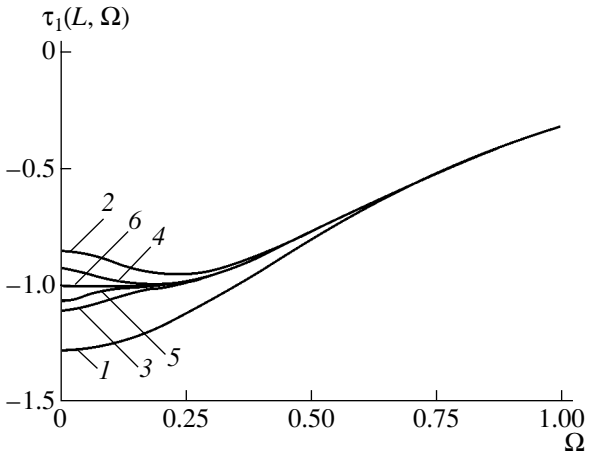

Fig. 3. (*1*) Frequency dependence of the corrections $\tau_1(L)$ corresponding to the first harmonic spectral component. (*2–5*) Successive improvement of the approximation accuracy as the second, third, fourth, and fifth harmonic spectral components of the second-order Walsh function, respectively, are taken into account. (*6*) Dependence of the correction $\tau_{1\text{wal}(2,\,t)}(L)$ calculated by (13).

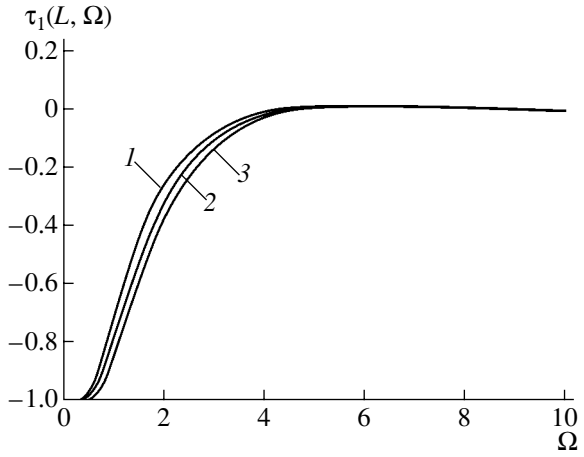

Fig. 4. Frequency dependence of the correction $\tau_{1\text{eff}}(L)$ when the diffusion coefficient is approximated by the second-order Walsh function for the degree of variation μ = (*1*) 0.01, (*2*) 0.3, and (*3*) 0.999.

simpler functional relationship

$$\Theta(L) = \Theta_0(L)\left(1 + \frac{\mu}{1+\mu}\tau_{\text{eff}}(L)\right), \qquad (16)$$

where $\tau_{\text{eff}}(L)$ is an effective correction to the transient period. At small μ ($0 \le \mu < 0.1$), relationship (16) coincides with the linear approximation. As μ approaches unity, the frequency dependence of $\tau_{\text{eff}}(L)$ remains nearly unchanged (Fig. 4). Thus, the calculation may be performed for the case when the diffusion coefficient varies weakly.

CONCLUSIONS

Studying the time it takes for the impurity concentration to reach the steady-state distribution in a homogeneous sample in relation to the law of variation of the diffusion coefficient, we found that the appropriate choice of this law may accelerate or slow down the transient. If the diffusion coefficient varies rapidly, the transient can be described in terms of the averaged diffusion coefficient; otherwise, one should handle its instantaneous values. The linear approximation is shown to have the advantage that it allows one to use the spectral approach and the principle of superposition in describing the process and calculate the total transient period as the sum of partial transient periods for individual spectral components of the diffusion coefficient law.

ACKNOWLEDGMENTS

This work was supported by the Russian Foundation for Basic Research (grant nos. 02-02-17517 and 03-02-06158), Scientific School (grant no. 1729.2003.2), and INTAS (grant no. 2001-0450).

REFERENCES

1. I. P. Stepanenko, *Foundations of Microelectronics* (Sov. Radio, Moscow, 1980).
2. *Field-Effect Transistors: Physics, Technology, and Applications*, Ed. by J. T. Wallmark and H. Jonson, (Prentice-Hall, Englewood Cliffs, 1970).
3. A. E. Dolbak, R. A. Zhachuk, and B. A. Ol'shanetskiĭ, Fiz. Tekh. Poluprovodn. (St. Petersburg) **36**, 1031 (2002) [Semiconductors **36**, 958 (2002)].
4. B. A. Zon, S. B. Ledovskiĭ, and A. N. Likholet, Zh. Tekh. Fiz. **70** (4), 38 (2000) [Tech. Phys. **45**, 419 (2000)].
5. G. A. Korn and T. M. Korn, *Mathematical Handbook for Scientists and Engineers* (McGraw-Hill, New York, 1968; Nauka, Moscow, 1977).
6. D. A. Garanin, Phys. Rev. E **54**, 3250 (1996).
7. A. N. Malakhov, Chaos **7**, 488 (1997).
8. V. T. Coffey, D. S. F. Crothers, and Yu. P. Kalmykov, Phys. Rev. E **55**, 4812 (1997).
9. A. A. Dubkov, A. A. Mal'tsev, and E. L. Pankratov, Zh. Tekh. Fiz. **72** (11), 14 (2002) [Tech. Phys. **47**, 1359 (2002)].
10. A. N. Malakhov and E. L. Pankratov, Izv. Vyssh. Uchebn. Zaved. Radiofiz. **44**, 367 (2001) [Radiophys. Quant. Electron. **44**, 339 (2001)].
11. A. I. Raĭchenko, *The Mathematical Theory of Diffusion in Applications* (Naukova Dumka, Kiev, 1981).
12. A. A. Mal'tsev and E. L. Pankratov, in *Proceedings of the 5th Scientific Conference on Radiophysics, Nizhni Novgorod, 2001*, Ed. by A. V. Yakimov (Nizhegorodsk. Gos. Univ., Nizhni Novgorod, 2001), pp. 211–212.
13. A. N. Malakhov and A. L. Pankratov, Adv. Chem. Phys. **121**, 357 (2002).
14. W. P. Gillin and D. J. Dunstan, Phys. Rev. B **50**, 7495 (1994).
15. I. S. Gonorovskiĭ, *Circuits and Signals* (Nauka, Moscow, 1977).

Translated by A. Sidorova

Technical Physics, Vol. 49, No. 1, 2004, pp. 119–122. Translated from Zhurnal Tekhnicheskoĭ Fiziki, Vol. 74, No. 1, 2004, pp. 120–122.

EXPERIMENTAL INSTRUMENTS AND TECHNIQUES

Effect of Resonance Charge Exchange between Argon Ions on the Effective Sputtering Rate in a Magnetron Discharge

I. Yu. Burmakinskiĭ and A. V. Rogov

Russian Research Centre Kurchatov Institute, pl. Kurchatova 1, Moscow, 123182 Russia

e-mail: alex-rogov@yandex.ru

Received May 29, 2003

Abstract—The sputtering rates of various materials in a magnetron ion sputterer are compared. The effective sputtering yields obtained when argon is used as a working gas are given. The difference between the sputtering yields found in this work and those obtained with monoenergetic ion beams is shown to be associated with resonance charge exchange between argon ions and neutral argon gas. © *2004 MAIK "Nauka/Interperiodica".*

INTRODUCTION

The method of monoenergetic ion beam is a standard technique for determining the sputtering yield of a material [1–3]. When analyzing plasma–ion sputtering processes, one usually needs to determine the rate of material erosion (consumption). This problem is related to an energy spread in the sputtering beam, its component composition, and processes occurring in the discharge plasma [4, 5]. In this work, we report experimental data for the effective sputtering rates of materials used in magnetron sputtering. The experiments were performed with a dc replaceable-cathode magnetron sputterer. For all samples, the sputtering process was carried out under the same vacuum conditions and discharge current.

EXPERIMENTAL

The sputtering process was conducted in a 250-mm-high cylindrical chamber with a diameter of 400 mm. A planar magnetron with cathodic inserts (targets) was used (Fig. 1). The diameter of the targets was 21 mm, and the maximal diameter of the sputtering area was 16 mm. The water-cooled magnetic system of the magnetron was placed in a hermetic housing. The wall thickness over the magnetic system surface was 1 mm. Heat transfer from the target made of a material to be tested was provided with a heat-conducting paste. The thickness Δ of the target was varied according to the material; accordingly, the magnetic field B_0 on the target surface varied from 0.110 to 0.155 T. The target thicknesses and the corresponding values of the magnetic field are listed in the table. The magnetic field over the cathode surface has an arc configuration, which is typical of dc ion-beam magnetron sputterers [1, 2].

The system was evacuated to a residual gas (nitrogen and oxygen) pressure of 1×10^{-4} torr by means of a turbomolecular pump. The process started at a working argon pressure $P \approx 5 \times 10^{-3}$ torr for all the materials under test. Thus, the concentration of residual impurities was no more than 2%.

The discharge current was the same for all the materials: I_d = 100 mA.

The sputtering time τ_{sp} was varied from 10 to 20 min depending on the material to provide a desired measurement accuracy.

During the process, the discharge voltage decreased insignificantly; therefore, when processing experimental data, we used the mean voltage value given by

$$\langle U_d \rangle = \frac{1}{\tau_{sp}} \int_{\tau} U_d(t)dt. \quad (1)$$

The amount of the material sputtered was found as the difference between the cathodic insert mass before, M^0_{cath}, and after, M^{end}_{cath}, the process. The sputtering parameters for each of the test materials are listed in the table.

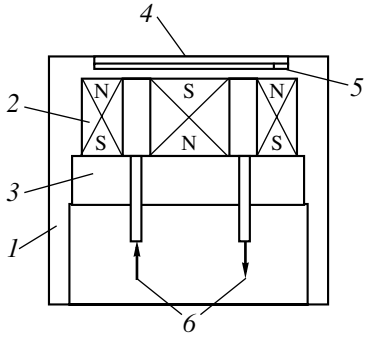

Fig. 1. Design of the cathode unit: *1*, magnetron housing; *2*, magnetic system; *3*, magnetic circuit; *4*, cathodic insert (target); *5*, heat-conducting paste; and *6*, water cooling of the magnetic system and cathode.

1063-7842/04/4901-0119$26.00 © 2004 MAIK "Nauka/Interperiodica"

Table

No.	Cathode material	Δ, mm	$\langle B_0 \rangle$, T	$\langle U_d \rangle$, V	$\langle \varepsilon_i \rangle$, eV	S_y(tm), atom/ion(atom)	S_y(Ar), atom/ion(atom)	S_y^{eff}, atom/ion(atom)
1	Al(26.28)	0.91	0.115	330	260	0.40	0.42	0.54
2	Ti(47.9)	1.00	0.110	305	245	0.22		0.19
3	Fe*(55.84)	0.23	*	320	255	0.65		0.51
4	Ni*(58.7)	0.07	*	360	290	0.8	0.8	0.63
5	Cu(63.54)	0.24	0.150	312	250	0.8	0.9	1.15
6	Zr(91.22)	0.40	0.135	260	210			0.21
7	Nb(92.91)	0.50	0.130	262	211			0.22
8	Mo(95.94)	0.32	0.14	277	222	0.21	0.32	0.36
9	Ag(107.87)	0.57	0.125	361	290		1.2	1.81
10	Ta(180.95)	0.20	0.155	280	225		0.22	0.28
11	W(183.8)	0.20	0.155	292	235	0.15	0.25	0.33
12	Pt(195.0)	0.27	0.145	373	300			1.11
13	C(12)	0.81	0.120	374	300	0.09	0.1	0.16
14	Ge(72.5)	0.81	0.120	355	285		0.65	0.71

* Magnetic material. Δ, the target thickness; $\langle B_0 \rangle$, the mean magnetic field value on the cathode surface; $\langle U_d \rangle$, the mean discharge value (given by (1)); $\langle \varepsilon_i \rangle$, the mean energy of a sputtering ion; S_y(tm), the sputtering yield when the target material is sputtered by its own ions (atoms); S_y(Ar), the sputtering yield when the target is sputtered by mean-energy argon ions (atoms); S_y^{eff}, the measured (effective) yield of target sputtering.

RESULTS AND DISCUSSION

The discharge current measured in the experiment is the sum of the ion current I_i toward the cathode and the secondary electron current $I_e^{sec} = \gamma_e I_i$ due to electron bombardment [2, 3, 6]:

$$I_d = I_i + I_e^{sec} = I_i(1 + \gamma_e), \quad (2)$$

where γ_e is the secondary electron emission coefficient.

The effective (or apparent [6]) sputtering yield S_y^{eff} was calculated for each of the elements from the formula

$$S_y^{eff} = \frac{(M_{cath}^0 - M_{cath}^{end})}{\tau_{sp} M_t} \frac{e(1 + \gamma_e)}{I_d}, \quad (3)$$

where e is the ion charge; M_t is the atomic mass of the target; I_d is the measured discharge current; τ_{sp} is the sputtering time; and M_{cath}^0 and M_{cath}^{end} are the masses of the target before and after the process, respectively.

From [5, 7], it follows that the energy $\langle \varepsilon_i \rangle$ (eV) of a sputtering ion averaged over the ion spectrum is 0.7–0.8 of the discharge voltage U_d expressed in energy units (eV). For comparison, the table lists the sputtering yields for the materials tested that were obtained by the method of monoenergetic beam when the energy of working gas (argon) ions and that of target material ions are $\langle \varepsilon_i \rangle = 0.8U_d$ (eV) [1]. Also given is the mean value of the discharge voltage (given by (1)) measured during the process for each of the target materials.

It turned out that, for most of the materials, the effective sputtering yield calculated by (3) exceeds the tabulated value $S_y^{tab}(\langle \varepsilon_i \rangle)$ [1, 2] by $\delta(S_y) = 15$–25% on average:

$$\delta(S_y) = \frac{\left| S_y^{eff} - S_y^{tab}(\langle \varepsilon_i \rangle) \right|}{S_y^{tab}(\langle \varepsilon_i \rangle)} \times 100\%.$$

One of the most plausible reasons for the increase in the effective sputtering yield is that the sputtering flux exceeds the purely ionic flux (current). This may happen if fast neutrals take part in the sputtering process, which is provided when argon ions exchange charge with the argon gas [7]. The sputtering yield for a neutral argon atom is nearly equal to that for an argon ion if their energies are the same [2, 3, 6].

The energy spectrum $f_i(\varepsilon)$ of ions bombarding the cathode [8] is shown in Fig. 2b.

The integral of the function $f_i(\varepsilon, 0)$ over the energy yields the ion current toward the cathode (up to a nor-

malizing factor and the cathode surface area):

$$I_i \propto \int_\varepsilon f_i(\varepsilon, 0)d\varepsilon,$$
$$I_{sp} \propto \int_\varepsilon f_i(\varepsilon, L_0)d\varepsilon. \quad (4)$$

Here, I_{sp} is the total flux of sputtering particles under the assumption that charge exchange takes place in a layer of extension L_0 that is localized between the ion flux formation region [9] and the cathode.

Then, for the functions $f_i(\varepsilon)$ and $f_i(\varepsilon, x)$, one can write

$$\frac{df_i(\varepsilon, x)}{dx} = f_i(\varepsilon, x)\sigma_{red}(\varepsilon)N_0^{Ar},$$
$$f_i(\varepsilon, 0) = f_i(\varepsilon) \quad (5)$$

and

$$f_i(\varepsilon, L_0) = f_i(\varepsilon)\exp\{L_0/\lambda(\varepsilon)\},$$
$$\lambda(\varepsilon) = \frac{1}{\sigma_{red}(\varepsilon)N_0^{Ar}}. \quad (6)$$

Thus, since the sputtering yields for an argon atom and an argon ion are equal to each other, expression (4) can be represented in the form

$$\delta(S_y) = \frac{\left|\int_\varepsilon f_i(\varepsilon, L_0)S_y(\varepsilon)d\varepsilon - \int_\varepsilon f_i(\varepsilon)S_y(\varepsilon)d\varepsilon\right|}{\int_\varepsilon f_i(\varepsilon)S_y(\varepsilon)d\varepsilon} \times 100\%$$
$$= \frac{\left|\langle S_y\rangle\int_\varepsilon f_i(\varepsilon, L_0)d\varepsilon - \langle S_y\rangle\int_\varepsilon f_i(\varepsilon)d\varepsilon\right|}{\langle S_y\rangle\int_\varepsilon f_i(\varepsilon)d\varepsilon} \times 100\%$$

or, after cancellation,

$$\delta(S_y) = \frac{\int_\varepsilon f_i(\varepsilon)[\exp\{L_0/\lambda(\varepsilon)\} - 1]d\varepsilon}{\int_\varepsilon f_i(\varepsilon)d\varepsilon} \times 100\%. \quad (7)$$

For a working gas pressure $P \approx 5 \times 10^{-3}$ torr and a charge exchange layer thickness L_0 = 3–5 mm, relationship (7) yields $\delta(S_y)$ = 15–25%, which agrees well with experimental data.

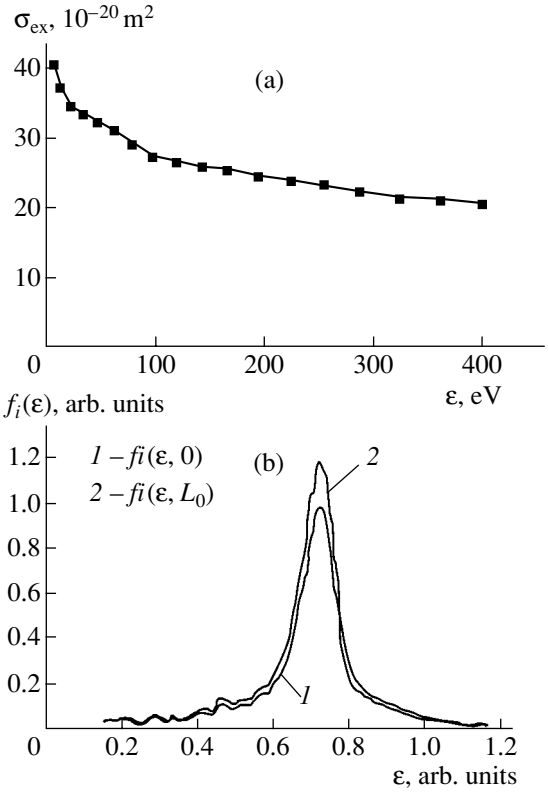

Fig. 2. (a) Energy dependence of the cross section of charge exchange between argon ions and neutral argon gas [7] and (b) ion energy spectrum (*1*) vs. total spectrum (*2*) of sputtering particles on the cathodic target.

CONCLUSIONS

Using an ion-beam magnetron sputterer, we measured the sputtering rates and effective sputtering yields for various target materials. For most of the materials, the sputtering yield exceeds that obtained by the method of monoenergetic ion beam [1–3] by 15–25% on average. The increase is related to lossless resonance charge exchange between some of the argon ions moving toward the cathode and the argon gas with the formation of fast neutrals participating in the target sputtering process. Thus, resonance charge exchange between argon ions and the argon gas in the magnetron discharge region is responsible for increased values of the effective sputtering yield.

For magnetic materials (iron and nickel), as well as for titanium, the decreased sputtering yield is obtained. In the former case, this is explained by the discharge parameters other than those for nonmagnetic materials because of a change in the magnetic field value and configuration over the cathode. In the case of titanium, this is because of a large amount of the reflected material, which is associated with the titanium-by-argon

scattering parameters and the geometry (dimensions) of the experimental system [10].

REFERENCES

1. W. Eckstеĭn, C. Garcia-Rosales, J. Roth, and W. Ottenberger, *Sputtering Data*, Tech. Rep. IPP 9/82, Max-Planck-Institut fur Plasmaphysik, Garching, 1993.
2. U. A. Arifov, *Interaction of Atomic Particles with a Solid Surface* (Nauka, Moscow, 1968; Consultants Bureau, New York, 1969).
3. M. Kaminsky, *Atomic and Ionic Impact Phenomena on Solid Surfaces* (Springer-Verlag, Berlin, 1965; Mir, Moscow, 1967).
4. P. J. Kelly and R. D. Arnell, Vacuum **56**, 159 (2000).
5. B. S. Danilin, *Application of Low-Temperature Plasma for Thin-Film Deposition* (Énergoatomizdat, Moscow, 1989).
6. *Electronic and Ionic Impact Phenomena*, Ed. by H. S. W. Massey and E. Burhop (Clarendon, Oxford, 1952; Inostrannaya Literatura, Moscow, 1958).
7. Yu. P. Raizer, *Gas Discharge Physics* (Nauka, Moscow, 1987; Springer-Verlag, Berlin, 1991).
8. M. Touzeau, M. Prioul, S. Roche, *et al.*, Plasma Phys. Controlled Fusion **42**, B323 (2000).
9. F. F. Elakshar, M. A. Hassouba, and A. A. Garamoon, Fizika (Zagreb) **9** (4), 177 (2000).
10. K. Macak, P. Macak, and U. Helmersson, Comput. Phys. Commun. **120**, 238 (1999).

Translated by V. Isaakyan

Technical Physics, Vol. 49, No. 1, 2004, pp. 123–127. Translated from Zhurnal Tekhnicheskoĭ Fiziki, Vol. 74, No. 1, 2004, pp. 123–127.

EXPERIMENTAL INSTRUMENTS AND TECHNIQUES

Pyrometer Unit for GaAs Substrate Temperature Control in an MBE System

S. E. Aleksandrov*, G. A. Gavrilov*, A. A. Kapralov*, G. Yu. Sotnikova*, D. F. Chernykh*, A. N. Alekseev, A. L. Dudin**, I. V. Kogan**, and A. P. Shkurko****

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

e-mail: gga_holo@mail.ru

*** ATC-Semiconductor Technologies & Equipment Joint-Stock Corporation, ul. Engel'sa 27, St. Petersburg, 194156 Russia*

e-mail: sapport@semiteq.ru

Received January 14, 2003; in final form, June 3, 2003

Abstract—An optical pyrometer designed for precision measurement of the GaAs substrate temperature during MBE growth is considered. The pyrometer can be calibrated against a certain characteristic absolute temperature that is visually determined from a change in the RHEED pattern. This enables one to calculate the absolute temperature of the substrate with regard to its radiant emissivity and minimize the inaccuracy of radiation temperature measurement. The inaccuracy is associated with the deposition of growth products on the pyrometer window. © *2004 MAIK "Nauka/Interperiodica".*

INTRODUCTION

Temperature determination on the substrate surface is a most challenging issue in MBE growth of semiconductor heterostructures. The need for rotating the sample during growth to provide structural homogeneity excludes the installation of a temperature-sensitive element on the substrate holder near the growth surface. A thermocouple is usually mounted in the substrate heater on the back side of the holder and is surrounded by a set of screens; thus, the environmental conditions for the thermocouple and substrate differ noticeably. The relative temperature measured by the thermocouple may differ from the actual temperature at the outer (growth) substrate surface by several tens of degrees and serves largely as a feedback signal in the substrate heating circuit.

The accuracy of measuring the absolute temperature at the surface of a growing epitaxial film specifies in many respects the quality of heterostructures and, ultimately, of semiconductor devices. As a rule, there are rather narrow (20–30°C) temperature intervals where an MBE-grown material offers an optimal optical or electric performance. A vivid example is the growth of ternary or quaternary semiconductor compounds. In particular, AlGaAs films with an aluminum content of 25–70%, which are used as waveguides or emitters in high-power semiconductor lasers, have optimal optical and crystalline quality when grown at temperature of 710–720°C [1]. At such high (for GaAs) temperatures, gallium desorption from the growing film surface becomes appreciable. To reproducibly grow several-micrometer-thick MBE layers of a given composition at a typical rate of 1 μm/h, it is necessary to determine the substrate temperature with an accuracy of no worse than ±3°C. Another problem is heteroepitaxial growth of compounds with substantially different optimal growth temperatures, e.g., InAlGaAs (550–570°C) on AlGaAs (710–720°C). Under these conditions, the fast variation of the substrate temperature cannot be detected by a thermocouple. As a result, the efficient control of the real temperature at the growing film surface in the course of epitaxial growth is virtually absent.

Optical pyrometry is widely used as a tool for additional substrate temperature monitoring [2–4]. Unlike thermocouples, a pyrometer, along with standard reflection high-energy electron diffraction, provides *in situ* data on the growth surface condition. Industrial pyrometers, calibrated against blackbody radiation, measure the so-called radiation temperature, which may differ considerably from the actual temperature of the object, since the exact value of its emissivity is usually unknown. The emissivity of a GaAs substrate is, as a rule, determined in special experiments, e.g., from the temperature of oxide evaporation, which is detected visually from a change in the RHEED pattern. When combined with RHEED, a pyrometer provides a necessary accuracy of substrate temperature measurement.

Our device for precision measurement of the GaAs substrate temperature can be calibrated against reference temperatures immediately during the MBE growth of the InAlGaAs/GaAs system. This enables one to determine the emissivity of the substrate and calculate its actual temperature from the radiation measured.

1063-7842/04/4901-0123$26.00 © 2004 MAIK "Nauka/Interperiodica"

SPECIFICATIONS

Temperature range—450–850°C

Design accuracy of temperature measurement—no worse than ±3°C (from 500 to 600°C) and ±2°C (from 600 to 850°C)

Number of measuring photodetectors—1

Rate of data display—no rarer than twice a second

Diameter of temperature measurement area—2–3 mm.

The optical scheme is designed for measuring the temperature of a substrate that is placed 65 ± 2 cm from the outer surface of the pyrometer window. The photodetector is tuned to the point of measurement with a special adjuster and eyepiece with cross. The device can be calibrated against a temperature preset by the user, which makes it possible to measure the actual temperature of the object with regard to its emissivity and the transmission coefficient of the optics. The actual surface temperature (in centigrade degrees) is displayed on a six-bit LED indicator and can be transferred to a computer.

In this paper, we discuss the operation and design of the pyrometer, as well as present the results obtained in an EP1203 MBE setup.

MEASUREMENT OF GaAs SUBSTRATE TEMPERATURE DURING GROWTH IN THE MBE SYSTEM

The radiative methods of temperature measurement are based on determining the emission intensity from an object in a narrow spectral range where its emissivity remains constant in an operating temperature interval. Within such an approach, the actual temperature is calculated using the Planck blackbody distribution law [5], which relates the spectral radiance distribution to the temperature of a heated body. With regard for the emissivity of real bodies, we have

$$R^0_{\lambda, T} = \varepsilon(\lambda, T)\frac{c_1}{\lambda^5}\frac{1}{\exp(c_2/\lambda T) - 1}, \qquad (1)$$

where c_1 and c_2 are emission constants, T is the actual temperature of a real body, and $\varepsilon(\lambda, T)$ is the emissivity of the real body.

The values of $\varepsilon(\lambda, T)$ for real bodies are known approximately. This parameter depends on the composition of a material and its surface condition. That is why pyrometry techniques measure not the actual temperature of a body but the so-called brightness temperature, which is the temperature of absolutely black body radiating at a given wavelength with a radiance equal to the radiance of a real object (in formula (1), $\varepsilon(\lambda, T)$ is assumed to be equal to unity). Obviously, the brightness temperature is lower than the real temperature. The correction ΔT is estimated as

$$\Delta T \approx T^2\frac{\lambda \ln(\varepsilon)}{c_2}. \qquad (2)$$

Note that the introduction of this correction requires that $\varepsilon(\lambda, T)$ be determined experimentally.

In an MBE setup, the GaAs substrate temperature is controlled by varying the current passing through a heating element, which is placed immediately behind the substrate. The heater temperature is, as a rule, 200–250°C higher than the temperature of the substrate surface (here, we consider an indium-free substrate holder, when the back side of the substrate is directly exposed to the heater). Calculations by formula (1) show that, in the temperature range specified (450–850°C), the spectral flux density of the heater is two to three orders of magnitude higher than the spectral flux density of the substrate.

The transmitted part of the heater radiation may be attenuated considerably if the measurements are taken in the spectral range where the radiation is attenuated by the sample as much as possible. Figure 1 presents the GaAs absorption spectra taken at T = 400 and 700°C [4]. The spectra have a sharp absorption edge, which shifts toward longer waves as the temperature increases. Experimental studies of absorption on substrates 350–400 μm thick that are used in heteroepitaxial growth have shown that the transmission coefficient of the substrates is on the order of 10^{-5} at a wavelength of 0.63 μm and may be expected to decrease as the wavelength increases up to 1 μm. Therefore, the radiation intensity from the GaAs substrate surface will exceed the transmitted radiation intensity from the heater 250 and 10^3 times at temperatures of 450 and 850°C, respectively. Thus, for the temperature interval of concern and in the spectral range of up to 1 μm, the GaAs substrate, on the one hand, is opaque to the heater radiation and, on the other hand, has (according to [4]) an almost constant emissivity $\varepsilon(\lambda, T)$.

In view of the above, the measurements were performed with a narrow-band interference filter designed for a wavelength of 0.88 μm (the half-width of its transmission spectrum is 50 nm; Fig. 1, curve *3*) that transmits the as yet fairly intense component of the heat radiation and a silicon photodiode with a maximal spectral sensitivity in this range (Fig. 1, curve *4*).

During the epitaxial process, the observation window is gradually covered by growth products, as a result of which its transmission coefficient varies. This is a serious problem in the pyrometer determination of the substrate temperature. It can be eliminated by periodically calibrating the temperature meter against a certain fixed (reference) point. As a reference point, we chose the temperature of oxide desorption from the GaAs surface. For MBE conditions, this temperature is well known and equals 580°C. Such an approach will be efficient if the GaAs surface is processed properly

(has a thin dense oxide layer) and the substrate is heated at a rate of no higher than 10°C/min, since the temperature of oxide desorption is detected visually from a change in the RHEED pattern [6].

As an additional calibration point that can be used immediately during the growth of the test structure, one may take the temperature boundary on the phase diagram for the AlAs surface under static conditions. According to [7], the (3×2)-to-(5×2) surface phase transition, which is observed with RHEED upon smoothly heating the substrate, occurs at a constant temperature (≈700°C) in the operating range of arsenic fluxes.

DESIGN OF THE DEVICE

The device consists of an optical module, where the radiation from the GaAs substrate is transformed into an electric signal, and a signal processing unit, which is connected to the optical module via a cable. The optical module may be mounted on the 40CF flange of the pyrometer window of an ÉP1203 setup. The optical scheme of the module is shown in Fig. 2. A lens 45 mm in diameter with a focal length of 183.5 mm images a part of the GaAs substrate in the photodetector plane with a twofold demagnification. The entrance diaphragm makes it possible to smoothly attenuate (up to 50 times) the intensity incident on the photodetector. The image is focused by moving the lens and diaphragm within the range –15…+26 mm relative to the photodetector. The photodetector can travel ±8 mm to the right (left) and ±8 mm up (down), thus making it possible to bring different parts of the substrate in sharp focus. An eyepiece serves to facilitate the fine adjustment of the photodetector.

The electronic circuit of the photodetector generates an electric signal, which varies in proportion to the photodetector illumination, and also preprocesses it, i.e., converts an analog signal to a time-modulated pulsed signal.

The signal processing unit receives the output signal from the photodetector, calculates and displays temperature values, and transfers them to a computer. A six-bit LED indicator on the front panel of the unit displays temperatures (in centigrade degrees) and other relevant information. The circuit measuring the pulse duration is built around a programmable logic array. A microprocessor calculates actual temperature values, evaluates calibration factors, and programs operating conditions (e.g., sets the accuracy of calculation, which is defined as the number of significant digits after the decimal point).

Data transfer to the computer is provided with the RS232 interface. The accuracy of data transferred is the same as displayed on the indicator. A program for graphically displaying the temperature vs. the number of time counts and for indicating running temperature values on the monitor screen is also available.

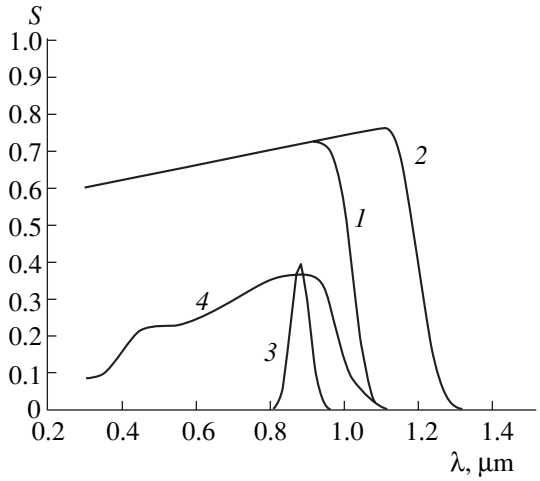

Fig. 1. Absorption spectra of GaAs at (*1*) 400 and (*2*) 700°C, respectively; (*3*) transmission spectrum of the interference filter; and (*4*) sensitivity spectrum of the photodetector.

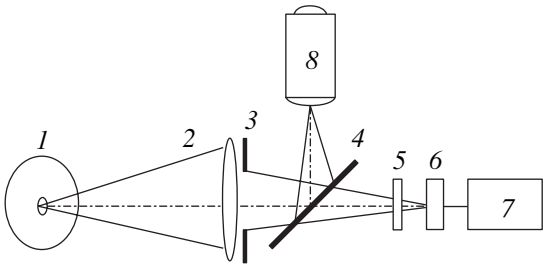

Fig. 2. Optical scheme of the device: (*1*) object, (*2*) lens, (*3*) diaphragm, (*4*) semitransparent mirror, (*5*) interference filter, (*6*) photodiode, (*7*) signal processing circuit, and (*8*) eyepiece.

CALCULATION OF ACTUAL TEMPERATURE AND CALIBRATION OF DEVICE

For the spectral line $\lambda = 0.88$ μm and the given temperature range, the condition $\lambda T < 3000$ μm K is met, so that one may use the approximation of Planck's formula (1) that is known as Wien's law [5] for spectral radiance distribution with regard to the emissivity of an object:

$$R_{\lambda, T} = \varepsilon(\lambda, T)\frac{c_1}{\lambda^5}\exp\left(-\frac{c_2}{\lambda T}\right). \quad (3)$$

From (3), one finds an analytical relationship between the temperature and radiance of a body:

$$1/T = -\frac{\lambda}{c_2}\ln\frac{R_{\lambda, T}\lambda^5}{c_1\varepsilon(\lambda, T)}. \quad (4)$$

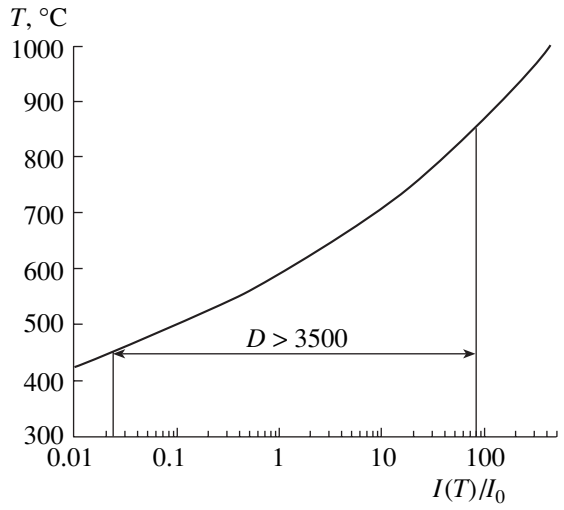

Fig. 3. Temperature vs. the photodiode relative signal.

In expression (4), the exact value of the substrate emissivity $\varepsilon(\lambda, T)$, as well as the proportionality factor between the true radiance of the spectral line and the measurement, are unknown. The calibration procedure enables us to eliminate the unknowns and obtain a final expression for the sample temperature.

The radiance of the object's spectral line is determined with the silicon photodiode, the output current of which is related to the desired value $R_{\lambda, T}$ as follows:

$$I(T) = kR_{\lambda, T}, \tag{5}$$

where k is the proportionality factor including the characteristics of the optics of the device and the conversion ratios of its electronic circuitry.

Substitution of (5) into (4) yields an expression for calculating the substrate temperature:

$$1/T = -\frac{\lambda}{c_2}\left[\ln\left(\frac{\lambda^5}{c_1\varepsilon(\lambda, T)k}\right) + \ln(I(T))\right]. \tag{6}$$

The first item in the square brackets is found upon calibration. It includes the optical properties of the object's surface ($\varepsilon(\lambda, T)$) and the value of k. The calibration process consists in determining the value of I_0 corresponding to a certain known temperature T_0:

$$\begin{aligned}\ln\left(\frac{\lambda^5}{c_1\varepsilon(\lambda, T_0)k}\right) &= -\frac{c_2}{\lambda T_0} - \ln(I_0) \\ &= -A(T_0) - \ln(I_0),\end{aligned} \tag{7}$$

where $A(T_0)$ is a constant that depends on calibration temperature.

The final formula for the actual temperature of the object that relates the photodiode current to the substrate temperature has the form

$$\begin{aligned}1/T &= \frac{\lambda}{c_2}[A(T_0) - \ln(I(T)) + \ln(I_0)] \\ &= \frac{\lambda}{c_2}\left[A(T_0) - \ln\left(\frac{I(T)}{I_0}\right)\right],\end{aligned} \tag{8}$$

where $I(T)$ is a current measurement.

Figure 3 plots the temperature measured against the ratio of photodiode currents that correspond to the measured and calibration temperatures (expression (8), λ = 0.88 μm). The plot implies that the detector must be capable of recording a thermal radiation signal from the substrate that varies by a factor of greater than 3500 as the substrate temperature varies from 450 to 850°C (i.e., the dynamic range D of the detector must exceed 3500).

CONCLUSIONS

Our device for pyrometric temperature control was tested upon the MBE growth of AlGaAs/GaAs laser heterostructures with an $In_{0.1}Ga_{0.9}As$ lasing layer [8]. Owing to the optimization of the growth process and doping profile, as well as the use of the device developed, the heterostructures had a low lasing threshold (J_{th}^0 = 70 A/cm^2), high internal quantum efficiency (η_{int} = 98%), and low optical losses (α_{int} = 1.5 cm^{-1}). Based on these heterostructures, high-power (1 W) laser diodes intended for pumping Yb^{3+}-ion solid-state lasers were made. They had a lasing wavelength of 945 nm, a high differential quantum efficiency (η > 65%), a low threshold current density (J_{th} = 100 A/cm^2), and a lifetime exceeding 10000 h. The temperature control device was applied in a series of experiments where three laser heterostructures were grown in succession. In these experiments, the lasing wavelength varied within ±2 nm from structure to structure, thereby demonstrating a high reproducibility of temperature indications. The growth temperature for the In-containing lasing layer was chosen at the beginning of the operating range, where indium desorption from the surface becomes noticeable. Accordingly, a considerable change in the growth temperature (>3°C) would change the active layer composition and, thus, the lasing wavelength.

The basic advantage of the pyrometer developed is that the device takes into account the features of a particular semiconductor compound (GaAs in our case) and makes it possible to measure *in situ* the actual surface temperature in a wide range (from 450 to 850°C) with a high accuracy. A high accuracy is provided by relating the actual temperature to the RHEED pattern. Thus, the device offers advantages over the existing analogues and can be widely used in MBE growth of

AlGaAs/GaAs heterostructures. Moreover, the basic design of the device can be adapted to any other substrate material in view of its emissivity.

REFERENCES

1. V. P. Chaly, D. M. Demidov, G. A. Fokin, *et al.*, J. Cryst. Growth **150**, 1350 (1995).
2. S. Strite, M. Kamp, and H. P. Meier, J. Vac. Sci. Technol. B **13**, 290 (1995).
3. F. G. Boebel, H. Moller, B. Hertel, *et al.*, J. Cryst. Growth **150**, 54 (1995).
4. S. Green, Eur. Semicond. **22** (1), S-35 (2000).
5. M. Born and E. Wolf, *Principles of Optics* (Pergamon Press, Oxford, 1969; Nauka, Moscow, 1970).
6. A. Y. Cho and J. C. Tracy, US Patent No. 3,969,164.
7. A. M. Dabrian and P. I. Cohen, J. Cryst. Growth **150**, 23 (1995).
8. S. V. Aleksandrov, A. N. Alekseev, D. M. Demidov, *et al.*, Pis'ma Zh. Tekh. Fiz. **28** (16), 71 (2002) [Tech. Phys. Lett. **28**, 696 (2002)].

Translated by M. Lebedev

Technical Physics, Vol. 49, No. 1, 2004, pp. 128–132. Translated from Zhurnal Tekhnicheskoĭ Fiziki, Vol. 74, No. 1, 2004, pp. 128–133.

EXPERIMENTAL INSTRUMENTS AND TECHNIQUES

Experimental Production of Long-Lived Luminous Organic Polymeric Objects

V. L. Bychkov, A. V. Bychkov, and I. B. Timofeev

Moscow State University, Vorob'evy Gory, Moscow, 119992 Russia

e-mail: bychvl@orc.ru

Received June 23, 2003

Abstract—Experiments on creating a plasma jet from capillaries made of polymeric materials, such as wax, ground wood, and rosin, as well as of their mixtures, are described. The plasma jet is used to ignite the wax vapor. The burning wax vapor forms ellipsoidal objects up to 10–15 cm in size, which lift. The time of their burning may reach 0.4 s. Long-lived (about 1.5 s) luminous objects (LLOs) with an apparent size of 1.5 cm are also produced in the experiments. As follows from scanning electron microscopy data, LLO residues consist of polymer fibers. The experiments allow one to suggest that some natural ball lightnings are of organic nature. © *2004 MAIK "Nauka/Interperiodica".*

INTRODUCTION

The creation of artificial ball lightnings or LLOs [1], as well as the analysis of polymeric ball lightnings and their relation with luminous objects (balls) in gas-discharge experiments [2–4], has raised the question as to whether LLOs are of polymeric nature and necessitated an experimental check of the available theories and hypotheses [4–6]. According to the theory developed in [3] and the hypotheses put forward in [2, 4], ball lightnings and spherical objects that are produced in gas-discharge experiments carried out in the organic atmosphere or in the presence of macromolecular components (in particular, wood components [2]) have a polymeric framework and appear via plasma-chemical processes involving macromolecules, dust particles, and silica particles. Due to the dielectric or polymeric nature of the materials involved, the recombination of charges accumulated in the objects takes a long time and the stability of these objects is temperature dependent; that is, the surrounding plasma must be kept at temperatures below the temperature at which the objects break down or ignite.

This work is an extension of works [5–7], which are concerned with long-lived luminous polymeric plasmoids and LLOs composed of cotton microfragments and coal particles produced in erosive discharges, and also of work [8], where the effect of plasma jets on various materials was studied.

EXPERIMENTAL

Since the objective of this work was to study polymeric LLOs, capillaries in the discharge chambers were made of organic polymers. Experiments were carried out with a standard capillary plasmatron described in detail in [1].

Its basic circuit is depicted in Fig. 1. It comprises capacitive energy storage system *1*, switching element (gap) *2*, and capillary plasmatron *3*. The parameters of the discharge circuit are as follows: the current pulse duration is 6 ms; the peak energy, 200 J; the voltage across the discharge gap, 300–340 V; and the peak current value, 100–150 A.

The *I–V* characteristic of this plasmatron with capillaries made of organic glass or a mixture of rosin, paraffin, and ground wood shows that the current and voltage pulses have a bell-shaped time waveform, which is typical of erosive plasmatrons [1]. For the organic glass capillary, the duration of the discharge pulse was ≈1.5 ms shorter than for the capillary made of the mixture.

The resistors incorporated into the voltage divider that was used to take the time dependence of the voltage across the discharge gap have the following values: R_1 = 62 kΩ and R_2 = 1.2 kΩ. The shunt resistance was

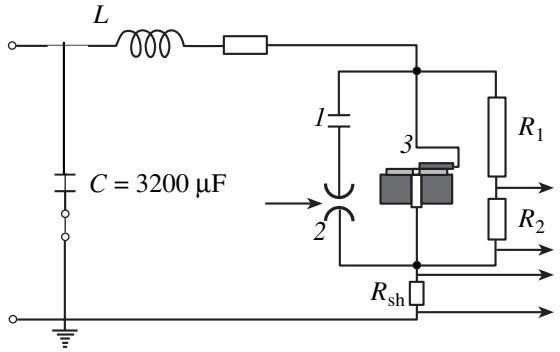

Fig. 1. Discharge circuit: *1*, capacitive energy storage system; *2*, gap; *3*, capillary plasmatron; R_1 and R_2, resistors of voltage divider; and R_{sh}, shunt resistor.

1063-7842/04/4901-0128$26.00 © 2004 MAIK "Nauka/Interperiodica"

$R_{sh} = 0.013\ \Omega$, and the total capacitance of the storage system was $C = 3200\ \mu F$.

The discharge was initiated with 3- to 4-mm-long capillaries of diameter from 1 to 2 mm (Fig. 2) made of different polymers. The lower electrode (*4* in Fig. 2) was made of VDPM tungsten–copper alloy, and the upper electrode was all-copper. Other elements of the plasmatron (except for the capillary) were made of organic glass.

The plasmatron and oscilloscope were triggered by sync pulses from a G5-15 pulser.

The discharge was photographed in the open-diaphragm integrated regime with the use of neutral filters. Video recording was performed in the continuous regime by means of cameras with frame durations of 33 and 41 ms. The current waveforms on the gap were recorded with an 0.013-Ω ohmic shunt.

The plasma was generated with capillaries made of PMMA, paraffin, a wax–rosin mixture, and a paraffin : rosin : ground wood (particle sizes of 100 to 300 μm) = 1 : 1 : 3 (by volume) mixture. Upon using various capillary (plasma-generating) materials, the outer plasma regions (i.e., those beyond the capillary) differed in shape and size.

The length of the luminous area was typically 10–12 mm for all the capillaries. The apparent diameter of the jet in the case of PMMA varied from 3 to 5 mm, while for capillaries made of the other materials, the diameter increased to 15–40 mm.

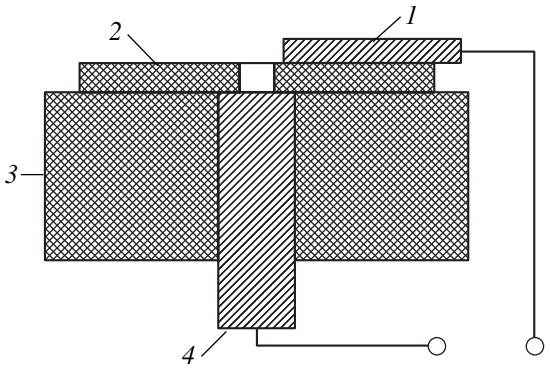

Fig. 2. Erosive plasmatron: *1* and *4*, electrodes; *2*, insulating plate with capillary (discharge chamber); and *3*, plasmatron base.

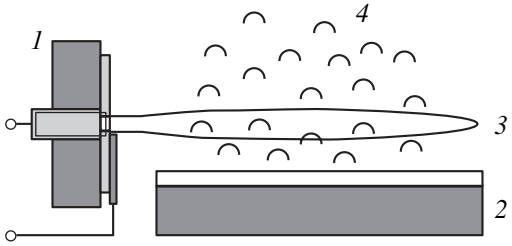

Fig. 3. Experimental scheme with the erosive plasmatron: *1*, plasmatron; *2*, cell with molten wax; *3*, plasma; and *4*, wax vapor.

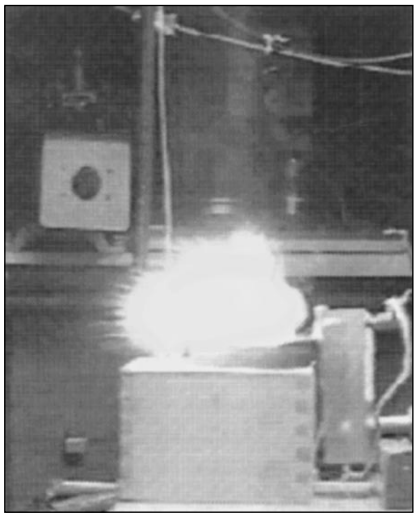

Fig. 4. Burning of the light fraction of the wax vapor as a result of interaction with the plasma jet. The plasmatron is to the right of the LLO.

EXPERIMENTS WITH THE CAPILLARY PLASMATRON

Production of luminous balls. The scheme of experiments with the capillary plasmatron is shown in Fig. 3. Wax in cell *2* was brought to the boiling point, and then the cell was placed at a distance of 1–2 cm from the plasmatron nozzle.

Figure 4 demonstrates the burning of the light fraction of the wax vapor as a result of interaction with the plasma jet from the PMMA capillary (3 mm long and 1 mm in diameter). After ignition, an LLO formed completely separates from the cell and rises 15–20 cm above its surface. This process lasts about 0.4 s. From the time and height of rise of the LLO, one can estimate the temperature of the gas inside it: $T \approx 2000$ K. Such a temperature means that the paraffin vapor burns out completely [9]. Assuming that, upon heating, the paraffin decomposes to methane, we can estimate the ratio of the molecular components over the heated paraffin surface: $CH_4 : O_2 : N_2 \approx 1 : 1 : 8$.

The results of this experiment deserve attention, since similar luminous balls with a lifetime of <1 s have been repeatedly observed under natural conditions, namely, during thunderstorms in forests under hot weather conditions [10]. Under these conditions, external factors, such as high temperature and calm, may cause a local increase in the concentration of organic vapors in the air layer near the Earth's surface. Since the rate of diffusion of organic polymeric particles of which these vapors consist is much lower than that of oxygen and nitrogen molecules, $D_{org} \sim D_{O_2}(M_{O_2}/M_{org})^{0.5}$, the particles may accumulate near the source (e.g., a tree). In essence, this experiment simulates the ignition

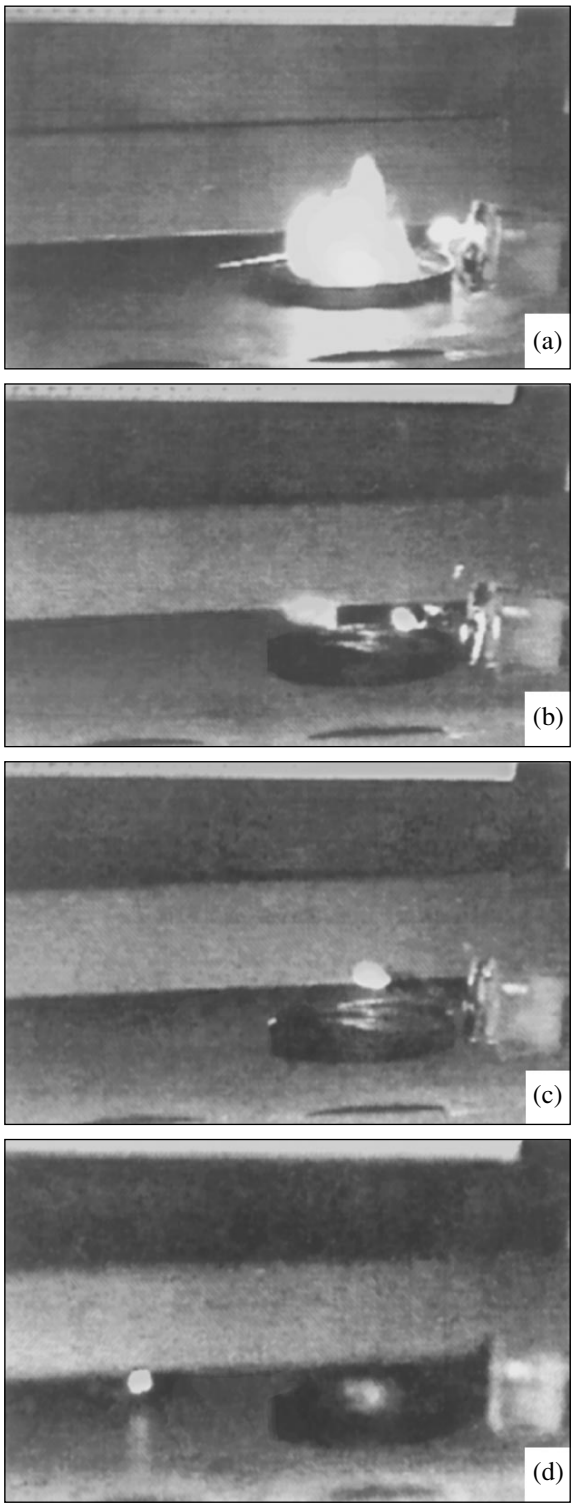

Fig. 5. Time evolution of a polymeric LLO produced with the capillary plasmatron. (a) LLO formation (within ≈41 ms). LLO after (b) ≈82, (c) ≈123, and (d) ≈533 ms.

of vapors interacting with the streak lightning discharge with the subsequent floating-up of the burning area. In our opinion, this experiment sheds light on the nature of the luminous balls.

Production of LLOs with a polymeric structure. With the capillary plasmatron, we conducted experiments aimed at producing LLOs with parameters as close to those of ball lightnings as possible, namely, with a lifetime of 1 s or more and a size of 1 cm or more.

These experiments followed two schemes. The first one is shown in Fig. 3. In the other, the cell with the wax vapor is absent. However, the presence of the hot vapor does not affect noticeably the LLO formation.

Figures 5a–5d show typical LLOs produced in these experiments. The images were made by a video camera with a frame duration of 41 ms. Knowing the number of frames, one can estimate the LLO lifetime. Here, frames 1–3 and 13 are shown (the total number of frames is 29 and the LLO lifetime, ≈1.2 s). Frame 0, which corresponds to the discharge pulse, is usually spoiled. With the energy deposit from the discharge mentioned above, LLOs appear from the plasmatron channel, as a rule, within 40–45 ms after the application of the discharge pulse. Another important finding is that the eye does not perceive small sizes of LLOs when they die out. In this specific case, the time of observation was ≈1 s.

A prerequisite for LLO appearance in these experiments was the proper choice of components of which the discharge chamber (capillary) is made. The sizes of the capillary remained the same: 3.0–3.5 mm in length and 1 mm in diameter. In this series of experiments, LLOs appeared only when the capillary was made of a mixture of pure medical paraffin (P), rosin (R), and ground wood (W) of mean "grain" size $1 \times 0.3 \times 0.3$ mm. The components were mixed in the ratio of 1 paraffin to 1 rosin to 3 wood by volume. We used pine wood in this work, since, according to [2], it contains a large amount of lignin. It was assumed that, when interacting with the plasma, lignin polymerizes with the formation of a new structure.

As a rule, LLOs were detected in three out of every four experiments. If a capillary was used many times, LLOs might not appear after each subsequent discharge. However, they often appear after 20 and 30 discharges. In this case, the diameter may reach 2.0 to 2.5 mm. If the rate of discharges was not too high (one discharge in 5 s), a capillary retained its initial diameter $d = 1$ mm after the first five to seven start-ups of the plasmatron.

It is remembered that LLOs appeared roughly 40 ms after the initiation of the discharge pulse, as readily demonstrated by video recording. However, we failed to estimate the energy consumed by the capillary over the time interval from the end of the current pulse to the appearance of an LLO from the plasmatron channel because of the specific design of the generator. Yet, a rather high LLO initial velocity (3–5 m/s) suggests that the energy evolution in the channel does take place over this time. From this velocity, the energy delivered to the capillary is estimated as $E = 0.4$–0.6 J.

During their lifetime, virtually each of the LLOs repeatedly collided with the surface of the experimental setup. The surface was covered by black paper. The video records showed that the objects experienced almost perfectly elastic collisions. The LLO velocity at the exit from the capillary varied from 3 to 5 m/s, as estimated from the video records. Some of the LLOs gradually shrank from 5–7 to 2–3 mm in diameter. Analyzing the records frame by frame, we sometimes observed how large LLOs disintegrate into smaller ones. The objects disappeared, as a rule, suddenly when colliding with the surface or in flight.

Once LLOs had died out, their residues (black particles up to 2 mm across) could be detected on the surface in a number of cases. The residues were examined under a SCAN-240-Cambridge scanning electron microscope at different magnifications. The most interesting micrographs are demonstrated in Fig. 6.

Figure 7 shows SEM images from fragments of the initial materials used to prepare the discharge chambers (capillaries). These fragments represent continuous polymeric structures with various inclusions. Some of the images exhibit coarse filaments or fibers up to 20 μm thick. Smooth surfaces correspond most likely to molten paraffin or rosin components.

The smooth surfaces are nearly completely absent on the LLO fragments (Fig. 6) unlike the initial material (Fig. 7). At a large magnification (×2980), all the surfaces in Fig. 6 appear porous and consist of interlaced filaments with distinct clusters in between. It may be assumed that the paraffin or rosin material has burned out, visualizing the LLO inner structure, which consists of cellulose and lignin (the components of ground wood).

The LLOs may be considered as the products of burning of coarse polymeric clusters, which arise when the wall material polymerizes in the channel or when material fragments separate from the wall and agglomerate inside the channel. In this case, heating the cluster surface may produce a burning film of gases leaving the surface. As was shown theoretically (by solving the heat conduction equation for thin samples) and confirmed experimentally in [11], the flame front for natural polymers may reach five to eight diameters of the sample. The lifetime of an organic object when it burns in a plasma is given by $t \sim 1.5d^2$ s (d in millimeters) [11]. Substituting d = 1–2 mm (the thickness of a polymeric cluster is limited by the plasmatron channel diameter, 1–2 mm) into this relationship, we find the theoretical lifetime of a luminous object, t = 1.5–6.0 s, and its apparent diameter, D = 5–16 mm. These values agree with experimental observations.

Note that clusters with a density of spruce, pine, or birch (the density of these materials is close to that of our mixture, $\rho = (0.3–0.5) \times 10^3$ kg/m^3)) and an apparent diameter of 10–16 mm, occupy a hot zone with a flame temperature T_f = 1400–1500 K (typical of polymer burning). Starting from Archimedes' principle, one can show that such clusters may soar owing to the buoyancy force of the cold air, which acts on the region occupied by the luminous object.

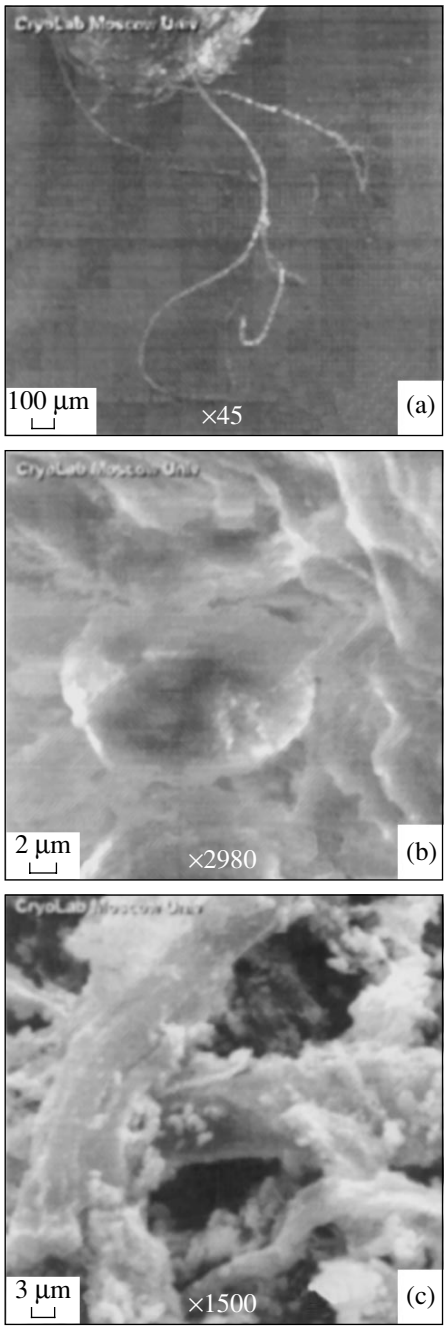

Fig. 6. SEM micrographs from polymeric LLO residues under different magnifications.

Thus, we may assume that burning polymeric particles that arise when streak lightning strikes a tree may generate ball lightning with a lifetime of 1.5–6.0 s and a diameter of 3–4 cm. In this case, the flame front of a size much larger than the particle size may be perceived by observers as the effective size of the ball lightning.

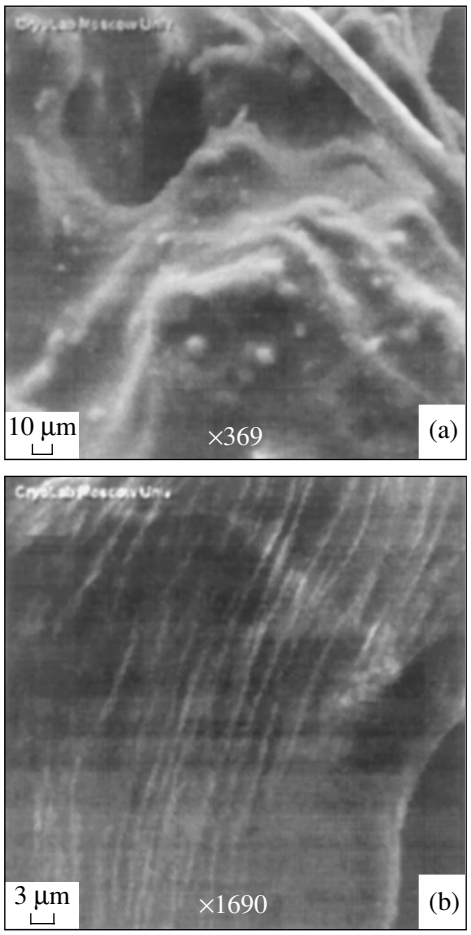

Fig. 7. SEM micrographs from the material of which the capillaries were made (different magnifications).

The flame front will also specify the ability of the particle to levitate owing to the buoyancy force.

Our results suggest that LLOs observed in [12, 13] were also of organic nature. They appeared upon closing the electrodes covered by pieces of charred wood [12] or because of the discharge from the Tesla generator high-voltage electrode covered by pieces of charred wood or wax [13].

CONCLUSIONS

Using a capillary plasmatron, we produced plasma jets from capillaries made of polymeric materials, such as wax, paraffin, ground wood, rosin, and their mixtures. The injection of the plasma jet ignites the wax vapor. From the rate of rise of areas occupied by the burning wax vapor (ellipsoids of characteristic size D = 10–15 cm) and the burning time (≈0.4 s), the temperature of the burning areas was estimated as $T \approx 2000$ K.

We also produced LLOs with a lifetime of up to 1.5 s and an apparent size of up to 1.5 cm. An optimal set of the capillary materials (paraffin, rosin, and ground wood) making it possible to form LLOs was found. LLO residues and the material of the capillary of which LLOs were formed were examined under a scanning electron microscope. The tentative analysis of the micrographs led us to assume that the materials studied consist largely of polymeric fibers.

Some ball lightnings observed in nature may be of organic character, i.e., represent burning areas occupied by heavy organic vapors or luminous structures with a polymeric framework.

REFERENCES

1. R. F. Avramenko, B. I. Bakhtin, V. I. Nikolaeva, and L. P. Poskacheeva, in *Collection of Articles*, Ed. by R. F. Avramenko *et al.* (Khimiya, Moscow, 1994), pp. 15–56.
2. N. M. Slyusarev, in *Ball Lightning*, Ed. by B. M. Smirnov (Inst. Vys. Temp., Akad. Nauk SSSR, Moscow, 1990), pp. 18–29.
3. V. L. Bychkov, A. V. Bychkov, and S. A. Stadnik, Phys. Scr. **53**, 749 (1996).
4. V. L. Bychkov, Phys. Scr. **50**, 591 (1994).
5. S. E. Emelin, V. S. Semenov, A. I. Éĭkhval'd, and A. K. Khassani, in *Collection of Articles*, Ed. by R. F. Avramenko (Khimiya, Moscow, 1994), pp. 87–95.
6. S. E. Emelin, V. S. Semenov, V. L. Bychkov, *et al.*, Zh. Tekh. Fiz. **67** (3), 19 (1997) [Tech. Phys. **42**, 269 (1997)].
7. V. L. Bychkov, A. F. Gridin, and A. I. Klimov, Teplofiz. Vys. Temp. **32**, 190 (1994).
8. A. P. Ershov, V. V. Rozanov, N. N. Sysoev, *et al.*, Preprint No. 8/1994 (Moscow State University, Moscow, 1994).
9. M. B. Ravich, *Simplified Methods of Thermal Analysis* (Akad. Nauk SSSR, Moscow, 1961).
10. A. R. Taylor, *Yearbook of Science and Technology* (McGraw-Hill, New York, 1974), pp. 201–203.
11. É. V. Konev, *Physical Grounds of Combustion of Plant Materials* (Nauka, Novosibirsk, 1977).
12. R. A. Ford, *Homemade Lightning* (McGraw-Hill, New York, 2001), pp. 227–229.
13. K. L. Korum and J. F. Korum, Usp. Fiz. Nauk **160** (4), 47 (1990).

Translated by V. Isaakyan

Technical Physics, Vol. 49, No. 1, 2004, pp. 133–134. Translated from Zhurnal Tekhnicheskoĭ Fiziki, Vol. 74, No. 1, 2004, pp. 134–135.

SHORT COMMUNICATIONS

Calculation of the Charge Composition of an Ion Beam from an E-Mevva Source

K. P. Artyomov and V. V. Ryzhov

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

e-mail: artyomov@to.hcei.tsc.ru

Received February 4, 2003

Abstract—The charge composition of an ion beam from an E-Mevva source is calculated within a model of sequential electron ionization of ions in a vacuum arc plasma. The effect of the plasma initial parameters on the charge composition of the ion beam is studied. Possible charge compositions for more than 30 elements omitted from this article are estimated. Also, the charge composition of the ions is calculated for the ultimate current density of the electron beam. © *2004 MAIK "Nauka/Interperiodica".*

The ionization of plasma ions by electrons is a way of generating multiply charged ions [1]. This method is used in an E-Mevva (Electron-beam Metal-vapor vacuum-arc) ion source, a modified version of the Mevva-V setup, to generate multiply charged metal ions. A vacuum arc initiated in the ionization chamber generates a plasma of density $n \approx 10^{11}$ cm^{-3}, and an electron beam (current density $j_b \approx 120$ A/cm^2, energy E_e = 20 keV, and pulse duration τ_p = 100 μs) is injected into the plasma through an opening in the cathode and transported in a magnetic field through the plasma. This device allowed the researchers to increase the charge of Cd, In, Sn, Sm, Pb, and Bi ions. For the two elements last mentioned, the record charges, Pb^{7+} and Bi^{8+}, were attained [2–4].

Calculations performed with the 2.5D PIC-code KARAT [5] showed that, when the electron beam is injected into the plasma, slow electrons of the plasma leave the transport region and the beam is transported under the conditions of complete charge neutralization. In these conditions, the processes of recombination and charge exchange may occur only with the participation of beam electrons and the characteristic times of these processes are much longer than the electron beam (pulse) duration τ_p. In this case, the basic process responsible for ion generation is the sequential ionization of plasma ions by beam electrons.

The mean path of a beam electron in the chamber exceeds the length of the chamber (40 cm) several times even if the effect of the external magnetic field is taken into account. Therefore, a plasma layer through which the electrons travel may be viewed as a one-collision layer and ion generation may be described by the set of kinetic equations [1]

$$\frac{dn_i}{dt} = v_e \sigma_{i-1 \to i} n_e n_{i-1}; \quad i = 1, 2, \ldots, m, \qquad (1)$$

where m is the maximal charge of ions generated, $\sigma_{i-1 \to i}$ is the cross section of ion ionization from an $(i-1)$th to ith state by 20-keV electrons, n_{i-1} is the concentration of ions in the $(i-1)$ state, n_i is the concentration of ions in the ith state, and v_e is the electron velocity.

For multiply charged ions, the ionization cross section may be calculated by the Lotz formula [6]

$$\sigma_i = 4.5 \times 10^{-14} \sum_{i=1}^{N} \frac{r_i}{E_e I_i} \ln\left(\frac{E_e}{I_i}\right), \qquad (2)$$

where σ_i is the cross section of ionization of ions in an ith state by electrons with an energy E_e, I_i is the ionization potential for an ith sublevel, r_i is the number of

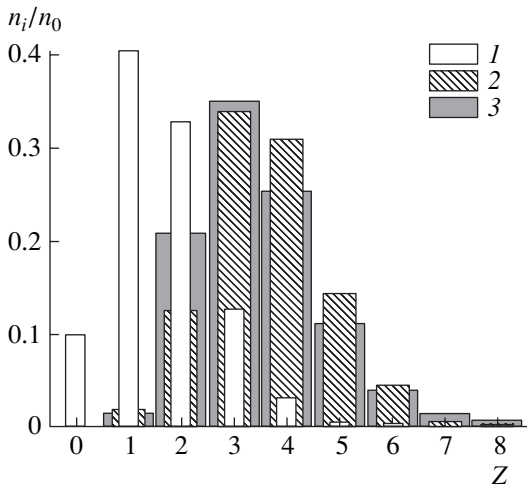

Fig. 1. Charge distribution for Bi ions upon the ionization of the (*1*) metal vapor and (*2*) vacuum arc plasma. (*3*) Experimental data.

1063-7842/04/4901-0133$26.00 © 2004 MAIK "Nauka/Interperiodica"

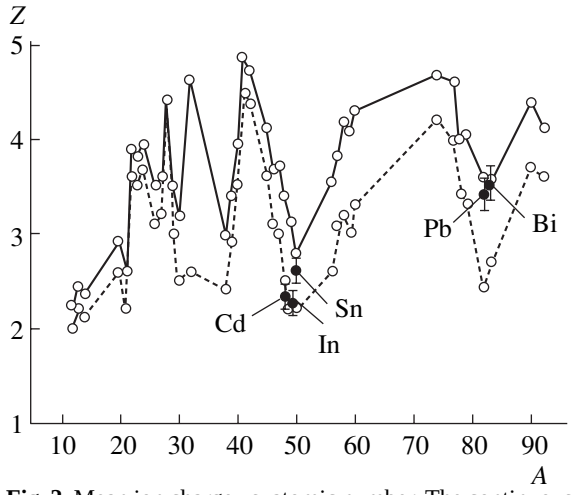

Fig. 2. Mean ion charge vs. atomic number. The continuous curve was calculated for the case when an electron beam acted on the plasma of a vacuum arc initiated in a high magnetic field. The dashed curve was calculated in the absence of the field. Symbols show the mean charges of Cd, In, Sn, Pb, and Bi ions.

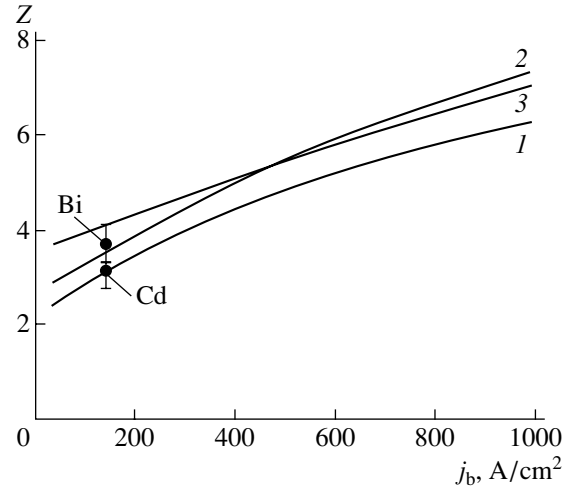

Fig. 3. Mean charge of (*1*) In, (*2*) Bi, and (*3*) U ions vs. j_b. Symbols show the experimental data for In and Bi at j_b = 120 A/cm^2.

electrons on the *i*th sublevel, and the *N* is the nuclear charge.

Set (1) was solved numerically. The data input was ion charge distributions obtained with the Mevva-V device [4] for various conditions of arc initiation. The calculations (Fig. 1) reinforced the statement that the use of the vacuum arc plasma (and not metal vapors) as an initial medium makes feasible an increase in the mean charge of the ions at a given beam parameter $j_b\tau_p$. In the case of metal vapors, a plasma of the same charge composition can be generated if the current density of an electron beam is three times higher than in the E-Mevva device [1].

Our model and ion charge distributions obtained with the Mevva-V device allow us to estimate the potentialities of an electron beam for generating multiply charge ions in the E-Mevva device [7]. Figure 2 plots the mean charge of ions against their atomic number. The continuous curve was calculated for the case when an electron beam acted on the plasma of a vacuum arc initiated in a high magnetic field (under these conditions, ion charge distributions were taken for Bi and Pb). The dashed curve was calculated in the absence of the field. Also shown are the mean charges of Cd, In, Sn, Pb, and Bi ions that were measured when the vacuum arc plasma was exposed to the electron beam in the E-Mevva device. The calculations indicate that the E-Mevva device makes it possible to generate Ni, Ge, Mo, W, Ir, and Th ions with a charge as high as 9 or 10.

To further increase the ion charge, it is necessary to raise the current density j_b of the electron beam. The ultimate current density tolerable for the E-Mevva ionization chamber depends on the plasma concentration and external magnetic field strength. Estimations with the code KARAT show that, with the plasma concentration $n = 10^{11}$ cm^{-3} and external magnetic field $B = 3$ kG, the maximal current density in the chamber may reach j_b = 700–800 A/cm^2. The mean charge of In, Bi, and U ions vs. j_b is plotted in Fig. 3. For the E-Mevva device, the maximal current density currently available is 350–400 A/cm^2. Such values may be reached by increasing either the total current of the beam or the magnetic field strength. As follows from calculations, U^{9+} ions may be generated in the E-Mevva device for j_b = 400 A/cm^2.

REFERENCES

1. E. Donets, in *The Physics and Technology of Ion Sources*, Ed. by I. G. Brown (Wiley, New York, 1989; Mir, Moscow, 1998), pp. 267–304.
2. A. Bugaev, V. Gushenets, and G. Yushkov, Appl. Phys. Lett. **79**, 919 (2001).
3. V. A. Batalin, A. S. Bugaev, and V. I. Gushenets, Rev. Sci. Instrum. **73**, 702 (2002).
4. A. Anders, G. Yushkov, and E. Oks, Rev. Sci. Instrum. **69**, 1332 (1998).
5. V. P. Tarakanov, *User's Manual for Code Karat* (Berkley Research Associate, Springfield, 1992).
6. D. H. Grandall, Phys. Scr. **23**, 153 (1981).
7. K. P. Artyomov and V. V. Ryzhov, in *Proceedings of the 6th International Conference on Modification of Materials with Particle Beams and Plasma Flows, Tomsk, 2002*, p. 635.

Translated by V. Isaakyan

Technical Physics, Vol. 49, No. 1, 2004, pp. 135–137. Translated from Zhurnal Tekhnicheskoĭ Fiziki, Vol. 74, No. 1, 2004, pp. 136–138.

SHORT COMMUNICATIONS

Stimulated Raman Scattering in Fiber-Optic Communication Lines

V. Yu. Golyshev, E. A. Zhukov, I. É. Samartsev, and D. G. Slepov

IRÉ-Polyus Research Association, Fryazino, Moscow Oblast, 141120 Russia

e-mail: mail@ntoire-polus.ru

Received April 29, 2003

Abstract—Stimulated Raman scattering (SRS) in a long-distance fiber-optic communication line with wavelength-division multiplexing (WDM) is studied theoretically at a high power of the signal transmitted. A new criterion for determining the SRS threshold is used to calculate the critical input power versus the number of optical channels and frequency separation between them. The theoretical model is verified experimentally. SRS interaction between two channels in a communication line with an SRS amplifier for which the optical waveguide of the line serves as a nonlinear medium is measured. © *2004 MAIK "Nauka/Interperiodica"*.

INTRODUCTION

To extend the distance over which information can be transmitted, modern fiber-optic communication lines usually employ erbium-doped amplifiers. However, as the power of the transmitted signal grows, a variety of nonlinear effects, such as stimulated Raman scattering, self-modulation and cross modulation, stimulated Brillouin scattering, and four-wave mixing, appear in the light guide. SRS limits the communication distance to the greatest extent [1, 2]. This effect is particularly severe in WDM systems. In this case, longer wave signals are enhanced at the expense of shorter wave ones [3, 4] and the SRS-induced power exchange between the channels increases with frequency separation within the bandwidth (15 THz) of Raman amplification. When the input exceeds a certain level, the output of shorter wave channels virtually stops growing.

In earlier works devoted to SRS, calculations were made either for short (no longer than 10 km) lines or for a low (no higher than 20 mW) power of the transmitted signal. However, today's long-distance fiber-optic WDM lines are longer than 200 km with typical signal powers of 100 mW or more. Therefore, one goal of this work is to clear up SRS-induced power limitations on modern optical fibers.

An exact analytical solution for the evolution of SRS channel interaction in WDM systems was found in [5]. Based on this result, we constructed a model for calculating the threshold power versus the number of spectral channels and frequency separation between them. In this model, a linear approximation of the gain spectrum with a slope $g' = 5.1 \times 10^{-18}$ m/(W GHz) is used.

EXPERIMENTAL VERIFICATION OF THE THEORETICAL MODEL

The theoretical model was verified with the experimental scheme shown in Fig. 1. Signals come from an erbium-doped fiber laser (EFL) with a spectral line width of less than 0.06 nm. Two optical signals with wavelengths $\lambda_1 = 1550.12$ nm and $\lambda_2 = 1561.42$ nm are applied to a 200-km-long fiber-optic line through a multiplexer (MUX). The power distribution between the two channels is measured by an optical spectrum analyzer over a wide (up to 250 mW) input range.

The experimental and theoretical results are shown to be in good agreement (Fig. 2). We thus may conclude that the model is valid in a wide range of powers of signals transmitted over long distances.

SIMULATIONS RESULTS

As the power of a "multichannel" signal at the input to a fiber-optic line grows, the output of the shortest wavelength channel first grows. Subsequently, however, its growth rate slows down because of SRS and the power reaches a maximum. As the input signal increases further, the output drops. We define the criti-

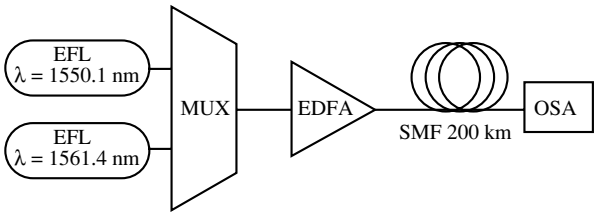

Fig. 1. Fiber-optic line with an erbium-doped laser amplifier at the input: EFL, erbium-doped fiber laser; MUX, multiplexer; EDFA + SMF, input device and fiber-optic line; and OSA, optical spectrum analyzer.

1063-7842/04/4901-0135$26.00 © 2004 MAIK "Nauka/Interperiodica"

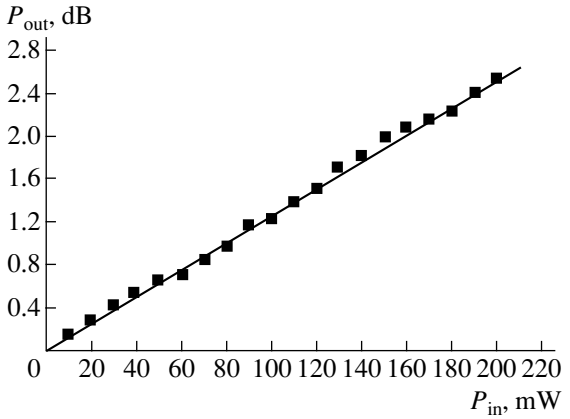

Fig. 2. Crosstalk between two channels for different channel output P_{out}. Solid line, theory; (■) data points. P_{in} is the channel input.

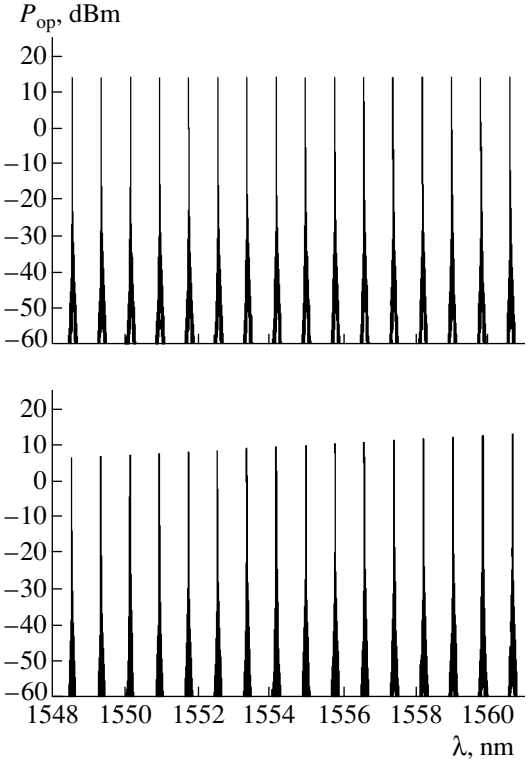

Fig. 4. Multichannel optical signal at the input (upper panel) and output (lower panel) of the 200-km-long line.

cal SRS intensity at the input as that which provides the maximum output of the shortest-wavelength channel. Simulation results are shown in Fig. 3. The number of channels was varied from 2 to 64. The frequency separation between the channels was either 100 or 200 GHz. In particular, for an eight-channel system with a 200-GHz channel separation, the critical power was 135 mW. For a smaller number of channels and smaller separation, the critical power was still greater.

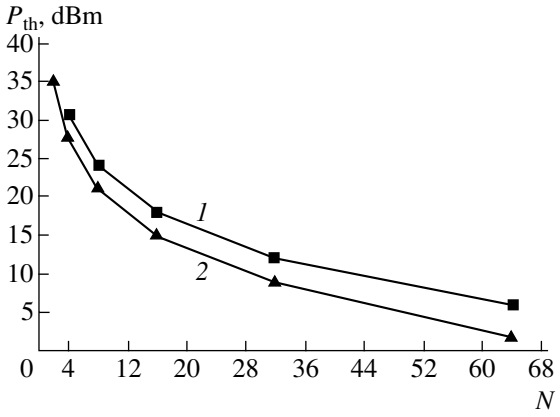

Fig. 3. Threshold power versus the number N of channels. The frequency separation is (*1*) 100 and (*2*) 200 GHz.

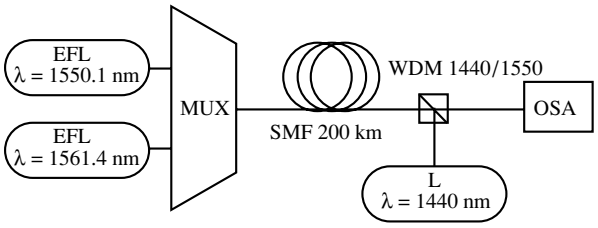

Fig. 5. Fiber Raman amplifier built in the fiber-optic line. L stands for pumping laser. The other designations are the same as in Fig. 1.

SRS causes a spectral nonuniformity at the output of the line. Figure 4 shows the spectra of a 16-channel signal at the input and output of a 200-km-long light guide. The power applied to the fiber was equal to the critical power, and the channel separation was 100 GHz.

With equal input powers, the maximal output difference $P_N - P_1$ will be that between the longest and shortest wavelength channels. Table lists the calculated changes in P_1 and P_N (the blank means the absence of the SRS).

The spectral nonuniformity mentioned above can be compensated for by using spectral filters mounted at the input and output of the line.

Thus, the SRS effect imposes significant limitations on the maximum transmitted power and, hence, on the communication distance. However, these limitations are not fundamental. If the passive segment of the line is long, they may be overcome with a fiber Raman amplifier (FRA) for which the light guide of the line serves as a nonlinear medium that accomplishes pump-to-signal conversion.

We measured the spectral nonuniformity in a 2-channel system with a distributed FRA pumped from the output of the line. The scheme of the experiment is shown in Fig. 5.

Table

Number of channels, N	Change in P_1 (P_N), dB	
	Δf = 100 GHz	Δf = 200 GHz
2	–	–3.6/+1.9
4	–3.8/+2.6	–3.8/+2.6
8	–3.9/+2.8	–3.9/+2.8
16	–3.9/+2.9	–3.9/+3.0
32	–3.9/+3.0	–3.9/+3.0
64	–3.9/+3.0	–

The FRA was pumped by a 1440-nm laser, whose power was varied from 0 to 1.2 W. The pumping radiation was fed into the fiber through a WDM 1440/1550 optical multiplexer. The optical input was 20 mW per channel, and the output varied between 0 and 85 μW. The difference in optical signals proved to be very small (less than 0.3 dB) throughout the range of measurement. The FRA reduces significantly the signal power losses due to SRS interaction between the channels compared with the standard design using an erbium-doped laser amplifier at the input of the line. A disadvantage of the FRA is that the power of the pumping laser must be one order of magnitude higher than the power of the erbium amplifier at the same output.

CONCLUSIONS

We simulated SRS interaction in a long-distance WDM system over a wide input range. Our results allow one to relate an optimum input power to the number of spectral channels and separation between them. The analytical model proposed was verified experimentally. The findings obtained in this work may be useful in designing advanced long-distance fiber-optic communication lines. SRS interaction between signals in a system with an FRA is found to be negligible.

ACKNOWLEDGMENTS

We thank O.G. Shkurikhin and A.G. Getman for their assistance.

REFERENCES

1. R. G. Smith, Appl. Opt. **11**, 2489 (1972).
2. R. H. Stolen, Proc. IEEE **68**, 1232 (1980).
3. A. R. Chraplyvy and P. S. Henry, Electron. Lett. **19**, 641 (1983).
4. A. R. Chraplyvy, J. Lightwave Technol. **8**, 1548 (1990).
5. D. N. Christodoulides and R. B. Janger, IEEE Photonics Technol. Lett. **8**, 1722 (1996).

Translated by A. Khzmalyan

Technical Physics, Vol. 49, No. 1, 2004, pp. 138–141. Translated from Zhurnal Tekhnicheskoĭ Fiziki, Vol. 74, No. 1, 2004, pp. 139–142.

SHORT COMMUNICATIONS

Dynamics of a Charged Particle Bunch in a Penning Trap

N. D. Naumov

Received May 5, 2003

Abstract—The Lagrangean equations for gas dynamics of a spherical bunch of charged particles in a Penning trap are solved. The solution describes the pulsation of an inhomogeneous particle bunch whose center behaves as a spatial oscillator in a coordinate system rotating with the Larmor frequency. *© 2004 MAIK "Nauka/Interperiodica".*

INTRODUCTION

The construction of analytical solutions to the equations of gas dynamics is a method of studying the properties of nonlinear systems [1]. A bunch of particles in a Penning trap is an example of space-bounded charged particle distributions for which one can obtain a nonstationary solution to the self-consistent problem.

A Penning trap is usually subjected to an external uniform magnetic field $\mathbf{B} = B\mathbf{e}_3$ and a nonuniform electric field whose potential in the cylindrical coordinates has the form

$$\Phi_0 = \kappa(2z^2 - \rho^2).$$

Such a field may be produced by hyperbolic electrodes: two surfaces $2z^2 = \rho^2 + 2d^2$ with a potential Φ and a surface $\rho^2 = 2z^2 + 2d^2$ with a potential $-\Phi$. Then, $\kappa = \Phi/2d^2$.

An exact solution to the equations of gas dynamics for a homogeneous spherical bunch of charged particles in a Penning trap was obtained earlier [2]. In this work, we construct a nonstationary solution to the self-consistent problem for a cold charged fluid,

$$\left[\frac{\partial}{\partial t} + (\mathbf{V}\cdot\nabla)\right]\mathbf{V} = \frac{e}{m}\left(\mathbf{E} + \mathbf{E}_{\text{ext}} + \frac{1}{c}[\mathbf{V}\times\mathbf{B}]\right), \quad (1)$$

$$\frac{\partial n}{\partial t} + n\,\text{div}\mathbf{V} = \mathbf{V}\cdot\text{grad}n = 0, \quad \text{div}\mathbf{E} = 4\pi en \quad (2)$$

that has the form of a spherical inhomogeneous bunch of particles in a Penning trap. Here, e and m are the charge and mass of the particle, respectively; n and $\mathbf{V}$ are the density and velocity of the fluid, respectively; $\mathbf{E}_{\text{ext}} = -\nabla\Phi_0$; and $\mathbf{E}$ is the intensity of the self-electric field of the bunch.

MOTION OF THE BUNCH CENTER

The motion of the bunch center will be described in terms of the vector $\mathbf{r} = x_i\mathbf{e}_i$, where $\mathbf{e}_i$ are the unit vectors of the coordinate system. In Eqs. (1) and (2), we pass to the variable $\mathbf{s} = \mathbf{x} - \mathbf{r}(t)$ assuming that $\mathbf{V}(\mathbf{x}, t) = \dot{\mathbf{r}} + \mathbf{U}(\mathbf{s}, t)$. The external field is a linear function of coordinates; hence, $\mathbf{E}_{\text{ext}}(\mathbf{x}) = \mathbf{E}_{\text{ext}}(\mathbf{r}) + \mathbf{E}_{\text{ext}}(\mathbf{s})$.

As a result, the motion of the fluid relative to the bunch center is described by the following equations:

$$\left(\frac{\partial}{\partial t} + \mathbf{U}\cdot\frac{\partial}{\partial\mathbf{s}}\right)\mathbf{U} = \frac{e}{m}\left(\mathbf{E} + \mathbf{E}_{\text{ext}}(\mathbf{s}) + \frac{1}{c}[\mathbf{U}\times\mathbf{B}]\right), \quad (3)$$

$$\frac{\partial n}{\partial t} + n\,\text{div}\mathbf{U} + \mathbf{U}\cdot\frac{\partial n}{\partial\mathbf{s}} = 0, \quad \text{div}\mathbf{E} = 4\pi en. \quad (4)$$

As we might expect, the dynamics of the bunch center is defined by equations of motion of a single particle in a Penning trap:

$$\ddot{x}_1 = \frac{1}{2}\omega^2 x_1 + 2\Omega\dot{x}_2, \quad \ddot{x}_2 = \frac{1}{2}\omega^2 x_2 - 2\Omega\dot{x}_1, \quad \ddot{x}_3 = -\omega^2 x_3, \quad (5)$$

where $\omega^2 = 4e\kappa/m$ and $\Omega = eB/2mc$.

The Hamiltonian equations for charged particle motion in a Penning trap were solved in [3].

It will be shown that a self-consistent solution for an inhomogeneous bunch may be obtained only if the external electric field gradient and the magnetic field are related as $\kappa = m\Omega^2/6e$. In this case, along with the initial coordinate system, it is convenient to consider another coordinate system where the unit vectors have the form

$$\mathbf{n}_1 = \mathbf{e}_1\cos\Omega t - \mathbf{e}_2\sin\Omega t,$$

$$\mathbf{n}_2 = \mathbf{e}_2\cos\Omega t + \mathbf{e}_1\sin\Omega t, \quad \mathbf{n}_3 = \mathbf{e}_3,$$

in other words, the new system rotates with an angular frequency $-\Omega$ relative to the initial coordinate system. The use of the rotating coordinate system makes it possible to more clearly demonstrate the motion of the bunch center.

1063-7842/04/4901-0138$26.00 © 2004 MAIK "Nauka/Interperiodica"

To describe the motion of the bunch center in the rotating coordinate system, we will use a vector $\mathbf{X} = X_j\mathbf{n}_j$. The initial conditions in this coordinate system are the initial positions x_{i0} and X_{i0} of the bunch center:

$$X_{j0} = x_{j0}, \quad \dot{X}_{10} = \dot{x}_{10} - \Omega x_{20},$$

$$\dot{X}_{20} = \dot{x}_{20} + \Omega x_{10}, \quad \dot{X}_{30} = \dot{x}_{30}.$$

Substituting $x_i = X_j\mathbf{n}_j\mathbf{e}_i$ into (5), we find that the motion of the bunch center in the new coordinate system is nothing but oscillations of a spatial oscillator:

$$X_j = X_{j0}\cos\omega t + \frac{1}{\omega}\dot{X}_{j0}\sin\omega t,$$

and, accordingly,

$$\ddot{X}_j = -\omega^2 X_j.$$

Thus, the motion of the bunch center in the initial coordinate system obeys the law

$$x_1 = X_1\cos\Omega t + X_2\sin\Omega t,$$

$$x_2 = X_2\cos\Omega t - X_1\sin\Omega t, \quad x_3 = X_3.$$

As is known, a spatial oscillator describes a circle, ellipse, or rectilinear segment depending on the initial conditions, the center of the path coinciding with the field center [4]. In our rotating coordinate system, the path of the bunch center lies in the plane that is normal to the vector $\mathbf{L} = m[\mathbf{X}_0 \times \mathbf{X}_0]$ and represents a circle if $E_0 = \omega L$ or an ellipse if $E_0 > \omega L$. The semiaxes a and b of the ellipse are given by

$$a = \frac{1}{\omega}\sqrt{\frac{E_0}{m}(1+D)}, \quad b = \frac{1}{\omega}\sqrt{\frac{E_0}{m}(1-D)},$$

where

$$D = \sqrt{1-\left(\frac{\omega L}{E_0}\right)^2}.$$

Here, $E_0 = m(\dot{\mathbf{X}}_0^2 + \omega^2\mathbf{X}_0^2)/2$, $\mathbf{X}_0 = X_{i0}\mathbf{n}_i$, and $\dot{\mathbf{X}}_0 = \dot{X}_{i0}\,\mathbf{n}_i$.

Taking into account that $X^2\dot{X}^2 = \omega^2(a^2 - X^2)(X^2 - b^2)$, one can easily check that, in the rotating coordinate system, the vector

$$\mathbf{N}_a = \sqrt{a^2 - X^2}[\mathbf{X}\times\mathbf{L}] - \frac{X\dot{X}[\dot{\mathbf{X}}\times\mathbf{L}]}{\omega^2\sqrt{a^2 - X^2}}$$

is also an integral of motion (along with the vector $\mathbf{L} = m[\mathbf{X}\times\dot{\mathbf{X}}]$).

If the bunch center is placed on the axis of symmetry of the ellipse that coincides with the semiaxis a, the vector $[\dot{\mathbf{X}}\times\mathbf{L}]$ is collinear with the vector $\mathbf{X}$ (since the vectors $\mathbf{X}$ and $\dot{\mathbf{X}}$ are mutually perpendicular at this point). Therefore, the vector $\mathbf{N}_a$ is directed along the major axis of symmetry of the ellipse (as the Laplace vector in the Kepler problem).

Similarly, one can check that the other vector that is retained in the rotating coordinate system,

$$\mathbf{N}_b = \sqrt{X^2 - b^2}[\mathbf{X}\times\mathbf{L}] + \frac{X\dot{X}[\dot{\mathbf{X}}\times\mathbf{L}]}{\omega^2\sqrt{X^2 - b^2}}$$

is directed along the minor axes of symmetry of the ellipse.

If the initial conditions for the bunch center are such that the vectors $\mathbf{X}_0$ and $\dot{\mathbf{X}}_0$ are collinear or one of them equals zero, the path of the center will be a rectilinear segment in the rotating coordinate system.

SOLUTION OF THE GAS DYNAMICS EQUATIONS

Instead of the Cartesian coordinates $\mathbf{s}_i = \mathbf{s}\mathbf{e}_i$, we will use the spherical coordinate system (s, η, ψ). The Lagrangean equations for fluid motion in this system appear as

$$\frac{\partial u_s}{\partial t} - \left[\frac{1}{s}(u_\eta^2 + u_\psi^2)\right]$$

$$= \frac{e}{m}[E + 2\kappa s(1 - 3\cos^2\eta)] + 2\Omega u_\psi\sin\eta,$$

$$\frac{\partial u_\eta}{\partial t} + \frac{1}{s}(u_s u_\eta - u_\psi^2\cot\eta) = 2\Omega u_\psi\cos\eta + 3\kappa s\frac{e}{m}\sin 2\eta,$$

$$\frac{\partial u_\psi}{\partial t} + \frac{1}{s}(u_s u_\psi + u_\eta u_\psi\cot\eta) = -2\Omega(u_s\sin\eta + u_\eta\cos\eta).$$

Here, u_s, u_η, and u_ψ are the coordinates of the velocity vector for a fluid element. It is easy to see that, if $\kappa = m\Omega^2/6e$, possible solutions to the two last equations have the form $u_\eta = 0$ and $u_\psi = -s\Omega\sin\eta$. In this case, the first equation simplifies to

$$\frac{\partial u_s}{\partial t} = \frac{e}{m}E - \omega^2 s. \tag{6}$$

Thus, under the conditions selected, a fluid element, as well as the bunch center, rotates with the Larmor frequency and executes radial oscillations. If fluid elements that are equidistant from the bunch center have the same initial velocities, they move identically. Therefore, if the form of the radial initial conditions for the fluid is spherically symmetric about the bunch center, we may consider the pulsation of a separate spherical layer of the fluid in the bunch rotating and executing translational motion as a whole. Then, the Lagrangean variable s is the radius of the layer. This variable depends on time t and the initial position s_0 of the layer: $s = s(t, s_0)$.

Eventually, the problem reduces to the analysis of the radial motion of a cold charged fluid subjected to a combination of external and collective fields. The tech-

nique for solving such a problem when particles move in the radial directions without overtaking was described in [5].

With this technique, we find for the particle density

$$n(t, s_0) = n_0 \frac{s_0^2 \omega(s_0)}{s^2 R(t, s_0)}; \quad R(t, s_0) = \frac{\partial s(t, s_0)}{\partial s_0}, \quad (7)$$

where the functions $s(t, s_0)$ and $R(t, s_0)$ are solutions to the equations

$$\ddot{s} = \frac{\omega_0^2}{s^2} W(s_0) - \omega^2 s, \quad (8)$$

$$\ddot{R} = \omega_0^2 \omega(s_0) \frac{s_0^2}{s^2} - \left[\omega^2 + 2W(s_0)\frac{\omega_0^2}{s^3}\right] R \quad (9)$$

(the time derivative is denoted by overcircles).

The initial conditions for Eqs. (8) and (9) are

$$s(0, s_0) = s_0, \quad \dot{s}(0, s_0) = v(s_0),$$

$$R(0, s_0) = 1, \quad \dot{R}(0, s_0) = \frac{dv(s_0)}{ds_0}.$$

Here, the functions $v(s)$ and $\omega(s)$ depend on the initial distributions of the particle density and gas velocity relative to the bunch center, $n(s, 0) = n_0\omega(s)$ and $U(s, 0) = v(s)$, and

$$\omega_0^2 = \frac{4\pi n_0 e^2}{m}, \quad W(s_0) = \int_0^{s_0} w(x) x^2 dx.$$

For a homogeneous bunch, $\omega(s) = H(1 - s/q)$, where $H(x)$ is the Heaviside step function and q is the initial radius of the bunch. If the initial velocity linearly varies with the distance to the center of symmetry, $v(s) = ks$ (k is a constant), we have $s(t, s_0) = s_0 R(t)$. Then, to calculate the pulsation of a homogeneous bunch, one needs solve one differential equation in the function R:

$$3\ddot{R} + 2\Omega^2 R - \frac{\omega_0^2}{R^2} = 0 \quad (10)$$

with the initial conditions $R_0 = 1$ and $\dot{R}_0 = k$.

For $s_0 \le q$, we obtain for the gas-dynamic parameters of a homogeneous bunch

$$n = \frac{n_0}{R^3}, \quad u_s = s_0 \dot{R},$$
$$u_\eta = 0, \quad u_\psi = -\Omega s_0 R \sin\eta. \quad (11)$$

In this case, the function R characterizes the time dependence of the bunch radius: $s(t, q) = qR(t)$.

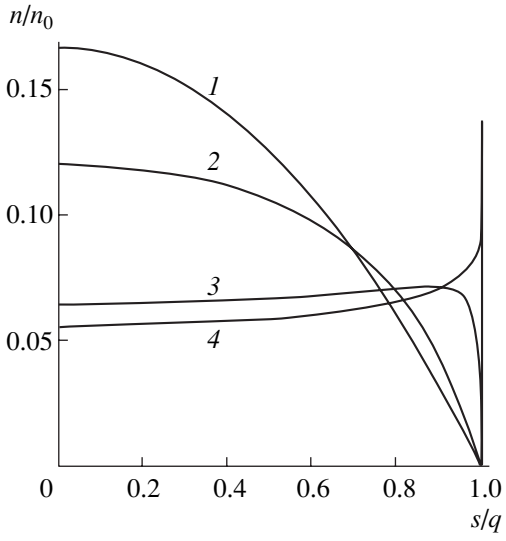

Fig. 1. Formation of the particle density peak.

NUMERICAL CALCULATION

The gas-dynamic properties of an inhomogeneous bunch can be calculated by varying s_0 in a step-type manner and solving Eqs. (8) and (9) at each step. Figure 1 shows the calculated particle density distributions in the bunch for $\tau = \Omega t$. Curve *1* corresponds to the initial particle density distribution

$$w(s) = \begin{cases} [1 - (s/q)^2]/6, & s \le q \\ 0, & s > q. \end{cases} \quad (12)$$

Calculation was performed for $v(s) = 0$ and $\omega_0^2 = 30\Omega^2$. As follows from Eq. (8), the former condition reflects a balance between space-charge and external-field forces in the surface layer of the fluid.

The formation of the particle density peak suggests that the assumption of overtaking absence is invalid. Overtaking makes itself evident in the fact that the condition $R(t_k, s_0) = 0$ is met at a certain time $t = t_k$ in a particular layer. At the time of overtaking, the particle density tends to infinity (so-called gradient catastrophes [1, 6]). In this case, a solution to Eqs. (8) and (9) describes the variation of the gas-dynamic properties of an inhomogeneous bunch till the time instant t_k for which $1.137 < \tau_k < 1.138$.

As follows from (10) and (11), overtaking is absent if the particle density distribution is uniform and the fluid velocity varies linearly with the distance to the bunch center. One might therefore expect that the applicability domain of the results will be different with initial distribution (12) changed appropriately.

Figure 2 shows the calculation results for

$$w(s) = \begin{cases} 1/2 - (s/q)^2/6, & s \le q \\ 0, & s > q. \end{cases}$$

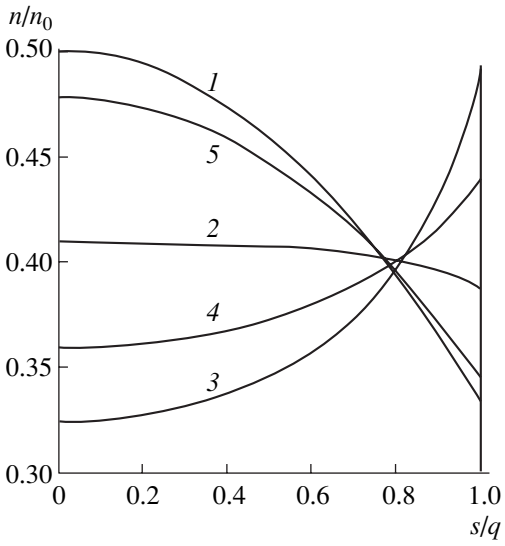

Fig. 2. Particle density oscillations.

In this case, space-charge and external-field forces balance at $\omega_0^2 = 5\Omega^2$. Here, τ = (*2*) 1, (*3*) 2, (*4*) 3, and (*5*) 4. Curve *1* corresponds to the initial particle density distribution. In this case, the particle density peak is of a finite height, indicating that the motion of the fluid is free of overtakings.

REFERENCES

1. A. A. Samarskii and A. P. Mikhailov, *Principles of Mathematical Modeling: Ideas, Methods, Examples* (Nauka, Moscow, 1997; Taylor & Francis, London, 2002).
2. N. D. Naumov, Fiz. Plazmy **19**, 1406 (1993) [Plasma Phys. Rep. **19**, 738 (1993)].
3. Yu. G. Pavlenko, *Hamiltonian Methods in Electrodynamics and Quantum Mechanics* (Mosk. Gos. Univ., Moscow, 1985).
4. L. D. Landau and E. M. Lifshitz, *Course of Theoretical Physics*, Vol. 1: *Mechanics* (Nauka, Moscow, 1982; Pergamon Press, New York, 1988).
5. N. D. Naumov, Zh. Tekh. Fiz. **69** (6), 84 (1999) [Tech. Phys. **44**, 682 (1999)].
6. V. P. Bykov, A. V. Gerasimov, and V. O. Turin, Usp. Fiz. Nauk **165**, 955 (1995) [Phys. Usp. **38**, 911 (1995)].

Translated by V. Isaakyan

Technical Physics, Vol. 49, No. 1, 2004, p. 142.

ERRATA

Erratum: "Photochromic Reactions in Silver Nanocomposites with a Fractal Structure and Their Comparative Characteristics" [Technical Physics 48 (6), 749 (2003)]

S. V. Karpov, A. K. Popov, and V. V. Slabko

"Asymmetric broadening (width)" and "symmetric broadening (width)" should be replaced by "inhomogeneous broadening (width)" and "homogeneous broadening (width)," respectively, throughout the text.

Technical Physics, Vol. 49, No. 1, 2004, pp. 19–21. Translated from Zhurnal Tekhnicheskoĭ Fiziki, Vol. 74, No. 1, 2004, pp. 21–23.

GASES AND LIQUIDS

On the Shear Strength of Structured Water

S. V. Stebnovskiĭ

Lavrent'ev Institute of Hydrodynamics, Siberian Division, Russian Academy of Sciences, pr. Lavrent'eva 15, Novosibirsk, 630090 Russia

e-mail: root@hydro.nsc.ru

Received April 30, 2003

Abstract—The structural characteristics of shear-stressed water are studied as a function of temperature in the range 0–20°C. At a temperature of 8°C, water exhibits a maximal shear strength (2×10^{-2} Pa) and a maximal recovery time ($\simeq 32$ s) of the locally disordered structure. With an aqueous sodium chloride solution as an example, it is shown that even a minor (0.08%) concentration of ions extends significantly the temperature range of water structure ordering. *© 2004 MAIK "Nauka/Interperiodica".*

INTRODUCTION

It is known [1] that some materials may be in the liquid-crystal state, i.e., have the fluidity of liquids, on the one hand, and exhibit molecular order and anisotropy of a number of physical parameters, which is typical of solids, on the other. A material with liquid-crystal (mesomorphic) properties is in the solid state at temperatures below some temperature T_*. When heated to $T > T_*$, it melts and passes into the mesomorphic state and, at a temperature $T_{**} > T_*$, it becomes a normal liquid. Thus, this material is in the mesomorphic state in the interval $T_* < T < T_{**}$. In addition, recent experiments have shown [2] that polar Newtonian liquids (water, glycerol, ethanol, etc.) acquire an ordered structure under certain conditions. This is consistent with data obtained in [3], where it was established experimentally that water and glycerol show a slight elasticity of shear (note that the Newtonian behavior of glycerol at a slow strain rate was observed in [4]). According to [2], water at rest has an ordered structure at $T < T_{**} = +19$°C. At $T > T_{**}$, molecular ordering breaks down and the structure of water becomes "chaotic" (i.e., $kT > \Delta U$, where kT is the mean energy of thermal motion of molecules and ΔU is the difference in the potential energies of the system at different molecular arrangements). Thus, one may formally assume that water has weak mesomorphic properties in the temperature interval $T_* = 0$°C $< T < T_{**} \simeq +19$°C.

In this work, we study the structure of tap water, distilled water, and snow melt subjected to local shear loads in the interval $T_* < T < T_{**}$.

RATE THRESHOLD OF WATER STRUCTURE ORDERING BREAKDOWN UNDER SHEAR

In an equilibrium ordered thixotropic structure, which water is at $T_* < T < T_{**}$ [2], the rate of breakdown of the supermolecular order equals the rate of its restoration [5]. It is therefore clear that the ordered structure of water in the mesomorphic state will break down if the rate of intermolecular (interassociate) bond breaking exceeds the rate of bond reconstruction. With this in mind, we estimated the threshold shear strain rate $\dot{\varepsilon}_\tau = \dot{\varepsilon}_\tau^*$ at which the water structure becomes disordered.

Experiments were carried out on a bench shown in Fig. 1. Here, *1* is a thermostatically controlled dish with transparent windows *2* filled with water *3*. Plate *4* is placed parallel to the axis of symmetry of the dish. Electric drive *5*, rotating the plate about the axis of symmetry of the dish, is equipped with a revolution counter and a reduction gear, which makes it possible

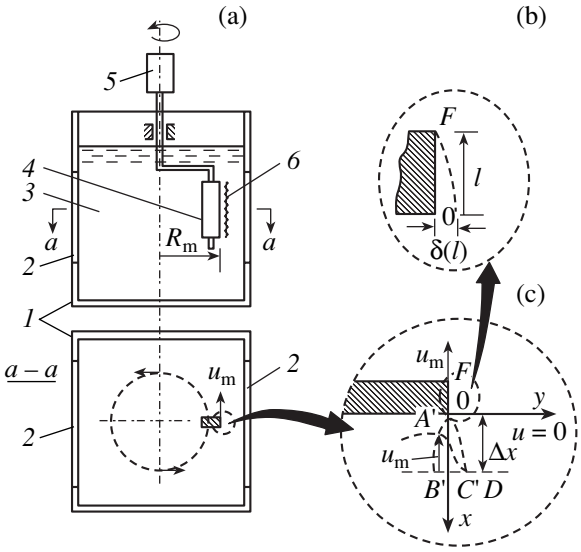

Fig. 1. Experimental scheme for the determination of the threshold shear strain rate at which the water structure becomes disordered.

1063-7842/04/4901-0019$26.00 © 2004 MAIK "Nauka/Interperiodica"

to vary the rate of rotation ω from 0.01 to 5 s^{-1}. During rotation, a zone with a maximal gradient of rate of rotation of the liquid forms near the plate's outer edge, which has a linear velocity $u_m = \omega R_m$ (Fig. 1a). At a certain value of u_m, waviness *6* (optical inhomogeneity due to water structure disordering [2]) is observed (Fig. 1a). Here, $R_m = 1$ cm is the distance from the outer edge of the plate to the axis of revolution. Optical inhomogeneity in the liquid was detected with schlieren and video cameras (the method of optical inhomogeneity detection similar to that described in [2]).

It is obvious that the shear strain rate $\dot{\varepsilon}_\tau$ near the outer edge of the plate is equal to $\nabla_y u$ (Fig. 1b), i.e., to the derivative of the velocity with respect to the coordinate normal to the direction of plate motion at a fixed time *t*. The derivative $\nabla_y u$ was estimated as follows. If the plate rotates slowly, one may assume that the liquid adjacent to the portion *OA* of the plate ($OA \ll R_m$) moves with a velocity roughly equal to u_m. Hence, the plane 0°C separates the moving and quiescent ($u = 0$) masses of the liquid. Then, according to [6], a transition layer *B*0*D* forms along the plane 0°C. In this layer, the velocity *u* drops from u_m to 0. The thickness of the transition layer is estimated by the formula [6]

$$\Delta y = 7(\nu \Delta x/u_m)^{1/2},$$

where ν is the kinematic viscosity coefficient of the liquid and Δ*x* is the distance from the edge of the plate (the point 0) to the point where the thickness $\Delta y = BD$ of the transition layer is determined. Therefore, we may set $\nabla_y u \simeq u_m/\Delta y$. However, this estimate yields $\nabla_y u \longrightarrow \infty$ at $\Delta x \longrightarrow 0$; therefore, one should take into account the formation of a boundary layer at the end face 0*F* of the plate upon moving (Fig. 1c). The thickness of the boundary layer at the point 0 can be found from the relationship [6] $\delta = 1.72(\nu l/u_m)^{1/2}$, where *l* is the thickness of the plate. Then, with $\Delta x \longrightarrow 0$, $\Delta y \longrightarrow \delta$ and the shear strain rate is given by

$$\dot{\varepsilon}_\tau = \nabla_y u \simeq \frac{u_m}{\delta} = \frac{u_m^{1/2}}{1.72(\nu l)^{1/2}}.$$

Curve *1* in Fig. 2 shows the experimental temperature dependence of the minimal shear strain rate $\dot{\varepsilon}_\tau^*$ at which the water structure becomes disordered (data for tap water). It is seen that the strength of the water ordered structure (which is expressed in terms of the ultimate shear stress $\tau^* = \mu \dot{\varepsilon}_\tau^*$, where μ is the shear viscosity of water) is the highest at 8°C ($\tau^* \simeq 2 \times 10^{-2}$ Pa) and then decreases with decreasing temperature (at least until $T \longrightarrow 3$°C). Such behavior is likely to be associated with the specific evolution of the water structure upon approaching the point of solidification.

In general, the threshold value of $\dot{\varepsilon}_\tau^*$ is sufficiently high. This fact may explain experimental observations that the ordered structure of water persists under the conditions of laminar jet flow (at least with Re below 10^3) in the temperature interval $T_* < T < T_{**}$.

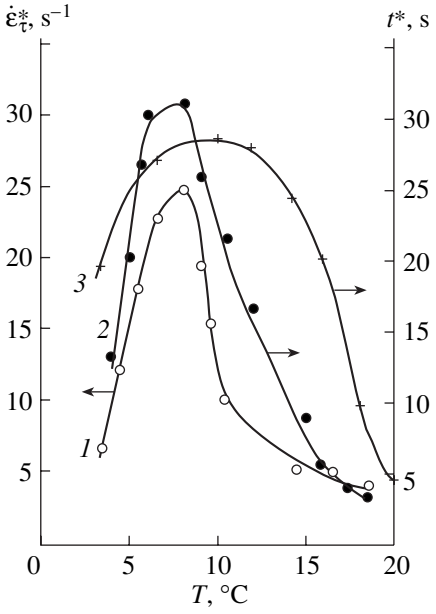

Fig. 2. Temperature dependence of (*1*) the shear strain rate $\dot{\varepsilon}_\tau^*$ at which the structure becomes disordered, (*2*) healing time t^* of disordered areas in the water sample, and (*3*) healing time t^*_{NaCl} of disordered areas in the 0.08% water solution of NaCl.

TEMPERATURE DEPENDENCE OF THE TIME OF DISORDER–ORDER TRANSITION FOR WATER IN THE MESOMORPHIC STATE

It has been shown [2] that the time taken to heal up disordering due to local shear strains in the water sample is $t^* = 3$–4 s at 17°C. If the entire sample is disordered, it is healed at least for more than an hour. In other words, the thixotropic properties of water at 17°C are weak. The question arises as to what time will be taken to heal up disordered areas of the sample when the temperature decreases, specifically, at $T \longrightarrow T_{**}$.

To perform relevant experiments, we followed the technique developed in [2] but the sample was placed in a thermostatically controlled dish where the temperature was kept constant in the interval from 1 to 20°C. The results for tap water are represented by curve *2* in Fig. 2. It is seen that the healing time has a maximum $t_m^* = 33$ s at 8°C (as for curve *1*).

Thus, water at 8°C has the highest structural strength and the longest time of healing of local disordered areas. Almost the same results were obtained for distilled water and snow melt (for all the samples, the spread of experimental data was no more than 7%).

It is worthy to note that even a minor concentration of ions affects drastically the structural order of water.

Curve *3* in Fig. 2 shows the temperature dependence of t^* for a 0.08% NaCl solution in distilled water (t^*_{NaCl}). This curve differs markedly from the same curve for the pure distillate (cf. curve *2*). It was found that the structural order in the 0.08% NaCl solution persists at least up to 50°C. Presumably, Na^+ and Cl^- ions, which have a high surface charge density, bind firmly water molecules in the hydration sheath [7]. This, in turn, raises the concentration of structured associates in the solution.

CONCLUSIONS

Our results can be summarized as follows. The temperature dependences of $\dot{\varepsilon}_\tau$ and t^* have maxima at 8°C. In other words, water exhibits the highest structural strength and the longest healing time of local disorder at this temperature. It may be assumed that the degree of order in water increases as the temperature decreases to 8°C; accordingly, a higher stress $\tau = \mu\dot{\varepsilon}_\tau$ is needed to destroy the structure.

However, the higher the degree of order before destruction, the longer the time taken to reconstruct the structure. This argument, however, does not explain why $\dot{\varepsilon}_\tau$ and t^* go on decreasing as the temperature declines below 8°C. It appears that a specific mechanism of reconstruction comes into play near the freezing point (before the onset of solidification).

A rather high threshold value of the shear strain rate ($\dot{\varepsilon}^*_\tau > 4\ s^{-1}$) accounts for the effect observed experimentally: the ordered structure of water persists up to $Re \simeq 10^3$ in the case of a laminar jet flow (i.e., when the velocity gradient along the jet radius is low) at $T < T_{**}$.

It was also found that even a minor concentration of ions affects drastically the degree of order, raising the temperature threshold T_{**} of water amorphization. For example, for a 0.08% NaCl solution, $T_{**} > +50$°C; therefore, blood plasma, containing about 0.9% NaCl, is bound to have mesomorphic properties at least up to 50°C.

REFERENCES

1. I. G. Chistyakov, *Liquid Crystals* (Nauka, Moscow, 1966; NTIS, Springfield, 1967).
2. S. V. Stebnovskiĭ, Zh. Tekh. Fiz. **72** (11), 24 (2002) [Tech. Phys. **47**, 1369 (2002)].
3. R. A. Apakashev and V. V. Pavlov, Izv. Ross. Akad. Nauk, Mekh. Zhidk. Gaza, No. 1, 3 (1997).
4. I. A. Dumanskiĭ and L. V. Khaĭlenko, Kolloidn. Zh. **15**, 426 (1953).
5. V. N. Zakharchenko, *Colloid Chemistry* (Vysshaya Shkola, Moscow, 1989).
6. G. K. Batchelor, *An Introduction to Fluid Dynamics* (Cambridge University Press, Cambridge, 1967; Mir, Moscow, 1973).
7. A. Ya. Samoĭlov, *Structure of Water Solutions of Electrolytes and Ionic Hydration* (Akad. Nauk SSSR, Moscow, 1957).

Translated by V. Isaakyan

Technical Physics, Vol. 49, No. 1, 2004, pp. 22–29. Translated from Zhurnal Tekhnicheskoĭ Fiziki, Vol. 74, No. 1, 2004, pp. 24–31.

GASES AND LIQUIDS

Internal Nonlinear Resonance of Charged Drop Capillary Oscillations in a Dielectric Medium at Multimode Initial Deformation of the Interface

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

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

e-mail: shir@uniyar.ac.ru

Received June 5, 2003

Abstract—An analytical expression for the generatrix of the shape of a nonlinearly vibrating charged drop of a perfect incompressible conducting fluid immersed in an ideal incompressible dielectric medium is found in the second order of smallness in terms of perturbation theory. The drop experiences multimode initial deformation. The expression contains resonant (small-denominator) terms. With the effect of the environment taken into account, the number of resonant situations becomes dependent on the drop-to-environment density ratio and the resonant self-charge of the drop changes. It is shown that nonlinear vibrations may be of resonant character even if the charge of the drop is far away from exact resonant values. This is because Rayleigh subcritical values of the self-charge affect the frequencies of higher vibration modes insignificantly. © *2004 MAIK "Nauka/Interperiodica".*

Vibration and stability of charged drops is of considerable interest in a variety of areas of science and technology (see, e.g., [1] and Refs. therein). In practice, one often deals with vibrating drops that are suspended or move in another immiscible fluid [2–9]. Such systems were repeatedly studied in an approximation linear in vibration amplitude [10–12]. Despite increased interest in nonlinear vibrations of a drop in general [13–16], vibration of a drop in a medium has come to the attention of researchers only recently [17].

(1) We will look for the shape of a drop of a perfect incompressible conducting fluid of density ρ_1 that executes nonlinear vibrations in a perfect incompressible dielectric medium of density ρ_2 and permittivity ε_*. The drop has a charge Q, equilibrium radius R, and interfacial tension coefficient σ. Analysis will be performed in the spherical coordinate system with the origin at the center of the drop in terms of dimensionless variables such that $R = \rho_1 = \sigma = 1$. We assume that the drop vibrates owing to a virtual axisymmetric multimode deformation imposed on the initially spherical shape. The deformation is defined by a finite spectrum of modes (with numbers $i \in \Xi$), each being described by an appropriate Legendre polynomial $P_i(\mu)$, where $\mu \equiv \cos\theta$.

Supposing that the vibrating drop remains axisymmetric, we write the interface equation in the form

$$r(\theta, t) = -1 + \xi(\theta, t); \quad |\xi| \ll 1, \tag{1}$$

where r and θ are the spherical coordinates and $\xi(\theta, t)$ is a function that describes the vibration-related deformation of the spherical interface.

The motion of the fluid in the drop and in the medium is assumed to be potential with velocity field potentials $\psi_1(\mathbf{r}, t)$ and $\psi_2(\mathbf{r}, t)$.

The mathematical statement of the problem includes the Laplace equations for hydrodynamic and electrostatic potentials

$$\Delta\psi_j(\mathbf{r}, t) = 0, \quad j = 1, 2; \quad \Delta\Phi(\mathbf{r}, t) = 0; \tag{2}$$

boundedness conditions

$$r \longrightarrow 0\colon \psi_1(\mathbf{r}, t) \longrightarrow 0, \tag{3}$$

$$r \longrightarrow \infty\colon \psi_2(\mathbf{r}, t) \longrightarrow 0, \quad |\mathrm{grad}\Phi(\mathbf{r}, t)| \longrightarrow 0; \tag{4}$$

and boundary conditions at the interface (see (1))

$$r = 1 + \xi(\theta, t)\colon \frac{\partial\xi}{\partial t} = \frac{\partial\psi_j}{\partial r} - \frac{1}{r^2}\frac{\partial\xi}{\partial\theta}\frac{\partial\psi_j}{\partial\theta} \ (j = 1, 2), \tag{5}$$

$$\frac{\partial\psi_1}{\partial r} - \frac{1}{r^2}\frac{\partial\xi}{\partial\theta}\frac{\partial\psi_1}{\partial\theta} = \frac{\partial\psi_2}{\partial r} - \frac{1}{r^2}\frac{\partial\xi}{\partial\theta}\frac{\partial\psi_2}{\partial\theta}, \tag{6}$$

$$\begin{aligned} &-\frac{\partial\psi_1}{\partial t} - \frac{1}{2}(\nabla\psi_1)^2 + P_0^{(1)} \\ &+ P_E - P_\sigma = -\rho\frac{\partial\psi_2}{\partial t} - \frac{\rho}{2}(\nabla\psi_2)^2 + P_0^{(2)}, \end{aligned} \tag{7}$$

$$\Phi(\mathbf{r}, t) = \Phi_s(t). \tag{8}$$

Also, we assume that the charge and volume of the

drop remain constant:

$$-\frac{1}{4\pi}\oint_S \varepsilon_*(\mathbf{n}\cdot\mathrm{grad}\Phi)r(\theta,t)^2\sin\theta d\theta d\phi = Q,$$

$$S = \begin{cases} r = 1+\xi(\theta,t) \\ 0\le\theta\le\pi \\ 0\le\phi\le 2\pi, \end{cases} \tag{9}$$

$$\int_V r^2 dr\sin\theta d\theta d\phi = \frac{4}{3}\pi, \quad V = \begin{cases} 0\le r\le 1+\xi(\theta,t) \\ 0\le\theta\le\pi \\ 0\le\phi\le 2\pi. \end{cases} \tag{10}$$

Finally, the initial conditions are

$$t = 0: \xi(\theta,t) = \xi_0 P_0(\mu) + \varepsilon\sum_{i\in\Xi} h_i P_i(\mu),$$
$$\frac{\partial\xi(\theta,t)}{\partial t} = 0. \tag{11}$$

In the above expressions, $P_0^{(j)}$ is the pressure in the drop ($j = 1$) and medium ($j = 2$) in equilibrium; $P_E = \varepsilon_*(\mathrm{grad}\Phi)^2/8\pi$ is the electric field pressure on the interface; $P_\sigma = \mathrm{div}\mathbf{n}$ is the Laplace pressure; $\mathbf{n}$ is the unit vector of the positive normal to the drop surface; $\Phi(\mathbf{r}, t)$ is the electrostatic potential induced by the drop self-charge in the medium; $\Phi_s(t)$ is the constant electrostatic potential on the drop surface; $\rho = \rho_2/\rho_1$; ε is the initial deformation amplitude (a small parameter of the problem); h_1 is the partial contribution of an ith vibrational mode to the initial perturbation:

$$\sum_{i\in\Xi} h_i = 1;$$

ξ_0 is a constant that is found from the drop volume constancy condition:

$$\varepsilon_0 = -\varepsilon^2\sum_{i\in\Xi}\frac{h_i^2}{(2i+1)} + O(\varepsilon^3);$$

and Δ is the Laplacian.

A complete set of equations that describes the problem must include the stationarity condition for the center-of-mass of the system. In the presence of an environment, this condition has the form

$$\frac{\int_\Omega\int_0^{r(\theta,t)}\mathbf{r}\cdot r^2 dr d\Omega + \rho\int_\Omega\int_{r(\theta,t)}^{L}\mathbf{r}\cdot r^2 dr d\Omega}{\int_\Omega\int_0^{r(\theta,t)} r^2 dr d\Omega + \rho\int_\Omega\int_{r(\theta,t)}^{L} r^2 dr d\Omega} = 0, \tag{12}$$

where L is the characteristic linear dimension of the space occupied by the environment ($L \gg 1$) and $d\Omega$ is a solid angle element.

In the problem of drop surface vibration in a vacuum ($\rho = 0$), conditions (10) and (12) impose additional restrictions upon the amplitude of the zero (volume) and first (translational) modes, respectively, in the expansion of the equation for the drop shape generatrix in Legendre polynomials. These restrictions are totally consistent with set (2)–(11) (i.e., an expression, e.g., for the translational mode amplitude that is derived from the stationarity condition for the center of mass coincides with that obtained from the set of boundary conditions). For a drop in an environment, the role of condition (10) persists (since it is assumed that both media are incompressible). At the same time, it is easy to check that, if the linear dimension L of the environment is sufficiently large, equality (12) can be made valid for an arbitrary function $r(\theta, t)$ with a however high accuracy. Thus, for the drop surface vibrating in an environment of large but finite volume, the stationarity condition for the center of mass may be omitted, since it is met automatically. As to the amplitude of the first (translational) mode, it is defined, as well as the amplitudes of all vibrational modes ($n > 2$), from boundary conditions (5)–(7). Note that, when excited, the translational mode compensates for the displacement of the center of mass due to vibrational surface modes.

(2) The problem stated by (2)–(11) was solved by the method of many scales, as in the case of a drop in a vacuum [13–16], in a second-order approximation. All unknown functions were expanded in the small parameter ε,

$$\psi_1(r,\theta,t) = \varepsilon\psi_1^{(1)}(r,\theta,t) + \varepsilon^2\psi_1^{(2)}(r,\theta,t) + O(\varepsilon^3),$$
$$\psi_2(r,\theta,t) = \varepsilon\psi_2^{(1)}(r,\theta,t) + \varepsilon^2\psi_2^{(2)}(r,\theta,t) + O(\varepsilon^3),$$
$$\Phi(r,\theta,t) = \Phi^{(0)}(r) + \varepsilon\Phi^{(1)}(r,\theta,t) + \varepsilon^2\Phi^{(2)}(r,\theta,t) + O(\varepsilon^3), \tag{13}$$
$$\xi(\theta,t) = \varepsilon\xi^{(1)}(\theta,t) + \varepsilon^2\xi^{(2)}(\theta,t) + O(\varepsilon^3),$$

under the assumption that they depend on two time scales: fast, $T_0 = t$, and slow, $T_1 = \varepsilon t$. The time derivatives were calculated as

$$\frac{\partial}{\partial t} = \frac{\partial}{\partial T_0} + \varepsilon\frac{\partial}{\partial T_1} + O(\varepsilon^2).$$

Substituting expansions (13) into (2)–(11) and collecting terms with the same powers of ε, we readily obtain boundary-value problems of different orders of smallness. A zero-order solution describes the electrostatic field distribution near the charged conducting drop,

$$\Phi^{(0)}(r) = \frac{Q}{\varepsilon_* r},$$

and pressure balance on its surface,

$$P_0^{(1)} - P_0^{(2)} = 2 - Q^2/8\pi\varepsilon_* r.$$

Clearly, the dependence of the functions $\psi_i^{(j)}(r, \theta, t)$ and $\Phi^{(j)}(r, \theta, t)$ $(i, j = 1, 2)$ on the spatial coordinates is defined by Laplace equations (2) with boundary condition (3) or (4):

$$\psi_1^{(j)}(r, \theta, t) = \sum_{n=1}^{\infty} D_n^{(j)}(T_0, T_1) r^n P_n(\cos\theta),$$

$$\psi_2^{(j)}(r, \theta, t) = \sum_{n=0}^{\infty} G_n^{(j)}(T_0, T_1) r^{-(n+1)} P_n(\cos\theta), \quad (14)$$

$$\Phi^{(j)}(r, \theta, t) = \sum_{n=0}^{\infty} F_n^{(j)}(T_0, T_1) r^{-(n+1)} P_n(\cos\theta).$$

The functions $\xi^{(j)}(\theta, t)$ $(j = 1, 2)$, which describe the shape of the oscillating drop surface, are also expanded in Legendre polynomials:

$$\xi^{(j)}(\theta, t) = \sum_{n=0}^{\infty} M_n^{(j)}(T_0, T_1) P_n(\cos\theta). \quad (15)$$

It should be noted that second-order analysis of the problem allows one to find the dependence of the first-order evolutionary coefficients on both time scales T_0 and T_1, $D_n^{(1)}(T_0, T_1)$, $G_n^{(1)}(T_0, T_1)$, $F_n^{(1)}(T_0, T_1)$, and $M_n^{(1)}(T_0, T_1)$, and the dependence of the second-order functions on the slow time T_0, $D_n^{(2)}(T_0)$, $G_n^{(2)}(T_0)$, $F_n^{(2)}(T_0)$, and $M_n^{(2)}(T_0)$.

From boundary and additional conditions (5)–(10), we find the set of first-order equations (in view of the zero-order solution)

$$r = 1: \frac{\partial\xi^{(1)}}{\partial T_0} = \frac{\partial\psi_1^{(1)}}{\partial r}; \quad \frac{\partial\psi_1^{(1)}}{\partial r} = \frac{\partial\psi_2^{(1)}}{\partial r},$$

$$-\frac{\partial\psi_1^{(1)}}{\partial T_0} + \rho\frac{\partial\psi_2^{(1)}}{\partial T_0}$$

$$-\frac{Q}{4\pi}\left(\frac{2Q}{\varepsilon_*}\xi^{(1)} + \frac{\partial\Phi^{(1)}}{\partial r}\right) = -(2 + \Delta_\Omega)\xi^{(1)},$$

$$\Phi^{(1)} = \frac{Q}{\varepsilon_*}\xi^{(1)} + \Phi_s^{(1)},$$

$$\int_0^\pi \frac{\partial\Phi^{(1)}}{\partial r}\sin\theta d\theta = 0; \quad \int_0^\pi \xi^{(1)}\sin\theta d\theta = 0. \quad (16)$$

The evolution of the vibrating drop shape is of most interest; therefore, substituting expansions (14) and (15) into (16) at $j = 1$, we obtain a differential equation for the coefficients $M_n^{(1)}$:

$$\frac{\partial^2 M_n^{(1)}}{\partial T_0^2} + \omega_n^2 M_n^{(1)} = 0. \quad (17)$$

Its general solution is written in the form

$$\begin{aligned} M_n^{(1)}(T_0, T_1) &= A_n^{(1)}(T_1)\exp(i\omega_n T_0) + (\text{c.c.}) \\ &= a_n^{(1)}(T_1)\exp(i\omega_n T_0 + b_n^{(1)}(T_1)) + (\text{c.c.}). \end{aligned} \quad (18)$$

Hereafter, the abbreviation c.c. means complex conjugates.

Solution (18) has either one arbitrary complex function $A_n^{(1)}(T_1)$ or two real arbitrary functions $a_n^{(1)}(T_1)$ and $b_n^{(1)}(T_1)$. The dependence of the latter on the time scale T_1 is found by solving the second-order problem. The quantities ω_n have the meaning of the eigenfrequencies of drop surface vibrations and are given by

$$\begin{aligned} \omega_n^2 &= \kappa_n n(n-1)[(n+2) - W]; \\ \kappa_n &\equiv \left(1 + \frac{n\rho}{n+1}\right)^{-1}, \end{aligned} \quad (19)$$

where $W \equiv Q^2/(4\pi\varepsilon_*)$ is the Rayleigh parameter characterizing the stability of the drop against its self-charge.

From (7)–(10), we found the set of second-order equations

$$r=1: \frac{\partial\xi^{(2)}}{\partial T_0} + \frac{\partial\xi^{(1)}}{\partial T_1} = \frac{\partial\psi_1^{(2)}}{\partial r} + \frac{\partial^2\psi_1^{(1)}}{\partial r^2}\xi^{(1)} - \frac{\partial\psi_1^{(1)}}{\partial\theta}\frac{\partial\xi^{(1)}}{\partial\theta},$$

$$\frac{\partial\psi_1^{(2)}}{\partial r} + \frac{\partial^2\psi_1^{(1)}}{\partial r^2}\xi^{(1)} - \frac{\partial\psi_1^{(1)}}{\partial\theta}\frac{\partial\xi^{(1)}}{\partial\theta}$$

$$= \frac{\partial\psi_2^{(2)}}{\partial r} + \frac{\partial^2\psi_2^{(1)}}{\partial r^2}\xi^{(1)} - \frac{\partial\psi_2^{(1)}}{\partial\theta}\frac{\partial\xi^{(1)}}{\partial\theta},$$

$$-\frac{\partial\psi_1^{(2)}}{\partial T_0} - \frac{\partial\psi_1^{(1)}}{\partial T_1} - \frac{\partial^2\psi_1^{(1)}}{\partial r\partial T_0}\xi^{(1)} - \frac{1}{2}\left(\frac{\partial\psi_1^{(1)}}{\partial r}\right)^2 - \frac{1}{2}\left(\frac{\partial\psi_1^{(1)}}{\partial\theta}\right)^2$$

$$+ \rho\left[\frac{\partial\psi_2^{(2)}}{\partial T_0} + \frac{\partial\psi_2^{(1)}}{\partial T_1} + \frac{\partial^2\psi_2^{(1)}}{\partial r\partial T_0}\xi^{(1)} + \frac{1}{2}\left(\frac{\partial\psi_2^{(1)}}{\partial r}\right)^2\right.$$

$$\left. + \frac{1}{2}\left(\frac{\partial\psi_2^{(1)}}{\partial\theta}\right)^2\right] - \frac{\varepsilon_*}{8\pi}\left\{\frac{2Q^2}{\varepsilon_*^2}[2\xi^{(2)} - 5(\varepsilon^{(1)})^2]\right.$$

$$\left. + \frac{2Q}{\varepsilon_*}\left[\frac{\partial\Phi^{(2)}}{\partial r} - \left(2\frac{\partial\Phi^{(1)}}{\partial r} - \frac{\partial^2\Phi^{(1)}}{\partial r^2}\right)\xi^{(1)}\right] - \left(\frac{\partial\Phi^{(1)}}{\partial r}\right)^2\right. \quad (20)$$

$$-\left(\frac{\partial\Phi^{(1)}}{\partial\theta}\right)^2\Bigg\} = -(2+\Delta_\Omega)\xi^{(2)} + 2\xi^{(1)}(1+\Delta_\Omega)\xi^{(1)},$$

$$\Phi^{(2)} = \frac{Q}{\varepsilon_*}\xi^{(2)} - \frac{Q}{\varepsilon_*}(\xi^{(1)})^2 - \frac{\partial\Phi^{(1)}}{\partial r}\xi^{(1)} + \Phi_s^{(2)},$$

$$\int_0^\pi\Bigg[\frac{\partial\Phi^{(2)}}{\partial r} + \left(2\frac{\partial\Phi^{(1)}}{\partial r} + \frac{\partial^2\Phi^{(1)}}{\partial r^2}\right)\xi^{(1)} - \frac{\partial\Phi^{(1)}}{\partial\theta}\frac{\partial\xi^{(1)}}{\partial\theta}\Bigg]\sin\theta d\theta = 0,$$

$$\int_0^\pi(\xi^{(2)} + (\xi^{(1)})^2)\sin\theta d\theta = 0.$$

Using (20), expansions (14) and (15), and solution (18), we can derive an inhomogeneous differential equation for the second-order evolutionary coefficients $M_n^{(2)}$:

$$\frac{\partial^2 M_n^{(2)}}{\partial T_0^2} + \omega_n^2 M_n^{(2)} = -2i\omega_n\frac{\partial A_n^{(1)}}{\partial T_1}\exp(i\omega_n T_0) + \sum_{k=0}^{\infty}\sum_{m=0}^{\infty}\{[\gamma_{kmn} + \omega_k\omega_m\eta_{kmn}]A_k^{(1)}A_m^{(1)}\exp(i(\omega_k+\omega_m)T_0) + [\gamma_{kmn} - \omega_k\omega_m\eta_{kmn}]A_k^{(1)}\overline{A_m^{(1)}}\exp(i(\omega_k-\omega_m)T_0)\}, \quad (21)$$

where

$$\gamma_{kmn} = \kappa_n\Bigg\{K_{kmn}\Bigg[\omega_k^2\left(n-k+1-\rho\frac{n(n-k-1)}{(n+1)}\right) + 2n(m^2+m-1) + W\frac{n}{2}(m(k+1)-k(2k-2n+7)+3)\Bigg] + \alpha_{kmn}\Bigg[\frac{\omega_k^2}{k}\left(1-\rho\frac{kn}{(k+1)(n+1)}\right) + W\frac{n}{2}\Bigg]\Bigg\};$$

$$\eta_{kmn} = \kappa_n\Bigg\{K_{kmn}\frac{1}{2}\Bigg[n-2k+2-\rho\frac{n(n-2k-3)}{(n+1)}\Bigg] + \alpha_{kmn}\Bigg[\frac{(n+2m)}{2km} - \rho\frac{n(n+2m+3)}{2(n+1)(k+1)(m+1)}\Bigg]\Bigg\};$$

$$K_{kmn} = (c_{k0m0}^{n0})^2;$$

$$\alpha_{kmn} = -\sqrt{k(k+1)m(m+1)}C_{k0m0}^{n0}C_{k(-1)m1}^{n0};$$

and C_{k0m0}^{n0} and $C_{k(-1)m1}^{n0}$ are the Clebsch–Gordan coefficients [18], which are nonzero if their subscripts satisfy the conditions

$$|k-m| \le n \le (k+m);$$

$$k+m+n = 2g \ (g \text{ is an integer}).$$

Omitting secular (linearly growing with time) terms from a solution to (21) yields the dependence of the functions $A_n^{(1)}$, $a_n^{(1)}$, and $b_n^{(1)}$ on the time scale T_1:

$$\frac{\partial A_n^{(1)}}{\partial T_1} = 0 \longrightarrow \frac{\partial a_n^{(1)}}{\partial T_1} = 0; \quad \frac{\partial b_n^{(1)}}{\partial T_1} = 0.$$

Thus, in the second-order approximation, the functions $A_n^{(1)}$, $a_n^{(1)}$, and $b_n^{(1)}$ are constants that equal their initial values.

Using a general solution to Eq. (21) and solution (18) and satisfying the first- and second-order initial conditions

$$t = 0: \xi^{(1)} = \sum_{i\in\Xi}h_i P_i(\mu); \quad \frac{\partial\xi^{(1)}}{\partial T_0} = 0;$$

$$\xi^{(2)} = -\sum_{i\in\Xi}\frac{h_i^2}{(2i+1)}; \quad \frac{\partial\xi^{(2)}}{\partial T_0} + \frac{\partial\xi^{(1)}}{\partial T_1} = 0,$$

which are obtained from (11), we can write final expressions for the evolutionary coefficients $M_n^{(1)}(t)$ and $M_n^{(2)}(t)$, which specify the shape of the vibrating drop (see (13) and (15)):

$$M_n^{(1)}(t) = \delta_{n,i}h_i\cos(\omega_i t) \ (i\in\Xi);$$

$$M_0^{(2)}(t) = -\sum_{i\in\Xi}\frac{h_i^2}{(2i+1)}\cos(\omega_i t)\cos(\omega_i t);$$

$$M_n^{(2)}(t) = \frac{1}{2}\sum_{i\in\Xi}\sum_{j\in\Xi}h_i h_j\{\lambda_{ijn}^{(+)}[\cos((\omega_i+\omega_j)t) - \cos(\omega_n t)] + \lambda_{ijn}^{(-)}[\cos((\omega_i-\omega_j)t) - \cos(\omega_n t)]\}; \quad (n>0); \quad (22)$$

where

$$\lambda_{ijn}^{(\pm)} \equiv \frac{[\gamma_{ijn} \pm \omega_i\omega_j\eta_{ijn}]}{[\omega_n^2 - (\omega_i \pm \omega_j)^2]},$$

and $\delta_{n,i}$ is the Kronecker delta.

(3) Figures 1a–1d show the time dependences of the second-order amplitudes of the second, fourth, sixth, and eighth modes (except for the zero one) that are excited via nonlinear interaction for various drop-to-medium density ratios ρ. Here, the initial deformation is specified by the virtual excitation of the fourth ($n = 4$) mode. It is seen that, as ρ increases, the eighth-mode

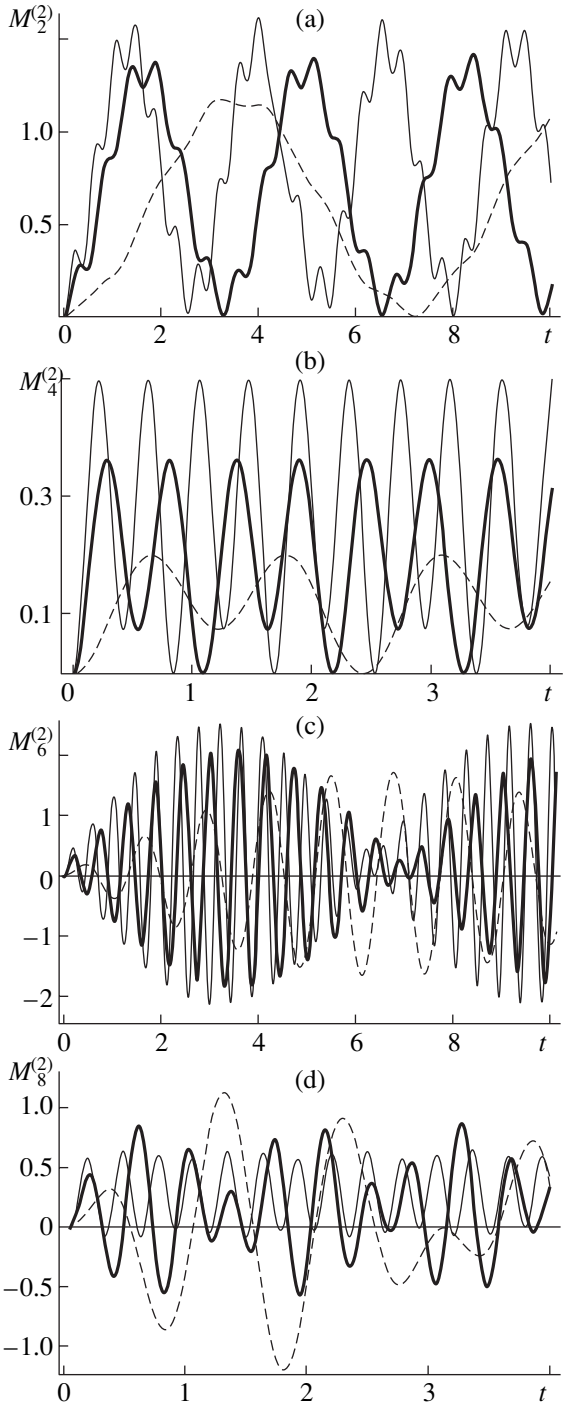

Fig. 1. Dimensionless second-order amplitudes $M_n^{(2)}(t)$ vs. dimensionless time for the initial deformation defined by the fourth mode at $W = 1$. $\rho = 0.1$ (thin lines), 1 (thick lines), and 10 (dashed lines). (a) Second mode, $n = 2$; (b) fourth mode, $n = 4$; (c) sixth mode, $n = 6$; and (d) eighth mode, $n = 8$.

amplitude grows, while those of the other modes decline (the zero-mode amplitude remains constant). The same tendency is also observed for modes with other numbers that specify the initial deformation of the drop. The zero-mode amplitude remains constant as ρ varies, because the zeroth mode does not participate in second-order nonlinear interaction. The dependence of its amplitude on time, which causes monopole acoustic radiation in the case of a compressible environment [19], is derived from the drop volume constancy condition and is defined by the square of the amplitude of a mode specifying the initial deformation. The ρ dependence of the zero-mode amplitude appears in third-order calculations.

The curves in Fig. 1 also demonstrate the well-known [10–12] effect of linear decrease (with increasing ρ) in the frequencies of all vibration modes that may be realized in the system considered.

Calculations by (22), which are illustrated in Fig. 1, were performed for the Rayleigh parameter $W = 1$, which is far away from the critical value $W_c = 4$ and the resonance value $W_r = 2.67$, at which the degenerate three-mode interaction between the fourth and sixth modes takes place [13, 15, 20, 21]. Nevertheless, it follows from Fig. 1c that the fourth and sixth modes resonate in the situation used in calculations: the amplitude of the sixth mode is substantially (several times) greater than those of the other modes excited via nonlinear interaction and grows in a resonant manner (linearly with time), although it seems that resonance energy transfer from the fourth to the sixth mode is impossible at $W = 1$. Of still greater interest is the fact that the resonant buildup of the sixth mode due to its interaction with the fourth mode is observed for any ρ used in calculations (Fig. 1c).

Note that, in our system, resonance situations arise because of small denominators in expression (22). Namely, if the frequencies of nonlinearly interacting modes satisfy the condition $\omega_n^2 = (\omega_i \pm \omega_j)^2$, one of the denominators in the coefficients $\lambda_{ijn}^{(\pm)}$, through which the second-order amplitudes are expressed, vanishes. The standard procedure of eliminating such situations consists in introducing a small offset of the frequency of one mode from the resonant value and then expanding in powers of this small offset and eliminating secular terms [21]. The condition for degenerate resonance between the sixth and fourth modes is $\omega_6^2 = 4\omega_4^2$ and is met at $\rho = 0$ ($W_r = 2.67$). If $W = 1$ and $\rho = 0.1$, it is easy to find that $\omega_6^2 \approx 193.4$ and $\omega_4^2 \approx 55.6$, so that $\omega_6^2 - 4\omega_4^2 \approx 28.8$, which is roughly eight times smaller than ω_6^2. Consequently, the ratio $(\omega_6^2 - 4\omega_4^2)/\omega_6^2$ may serve as a small parameter and the difference $\omega_6^2 - 4\omega_4^2$ may be considered as close to zero. In other words, with $W = 1$, the relationship between the frequencies of the sixth and fourth modes is close to the resonance condition and the resonance buildup of the sixth mode due to energy transfer from the initially excited fourth mode may show up in calculations. This fact is interesting in that the resonance occurs when W is far away from W_r.

This means that the drop charge may not be large for one mode to resonantly build up.

Figure 2 illustrates curves similar to those depicted in Fig. 1c. They show the resonant buildup of the sixth mode at the initially excited fourth mode with $W = 0$ (the zero charge of the drop). In this case, at $\rho = 0.1$, we have $\omega_6^2 \approx 221$, $\omega_4^2 \approx 66.7$, and, accordingly, $\omega_6^2 - 4\omega_4^2 \approx 45.7$. Here, the system is still farther from exact resonance than at $W = 1$ (see above), as also follows from the fact that the ratio of the resonantly growing amplitude of the sixth mode to those of other nonlinear excited modes is somewhat lower than at $W = 1$ (see Fig. 1).

It should be taken into account that, according to [21–23], the number of resonances where lower modes resonate along with higher ones is large (several hundred at $n, i, j \le 100$). Since the condition for nonlinear resonant energy exchange between modes depends on the self-charge of the drop (the parameter W) only slightly, we may expect that, under natural conditions (for example, in a storm cloud), all resonance situations allowable at a given set of initially excited modes will be realized in free-falling drops even if their charge is far from the resonant value. This circumstance is of importance for the simulation of an as yet unclear mechanism of lightning discharge origination in a storm cloud (according to the present-day concept, lightning originates from a corona initiated in the vicinity of a coarse free-falling hailstone). Since the electrical charges of drops in a cloud are no greater than one-third of the Rayleigh critical value, as determined in full-scale experiments, and intracloud fields are many times lower than those necessary for the initiation of a corona [24], a most plausible reason for a corona near a water-covered hailstone or drop is instability of its charged surface. The unstable surface emits a large amount of heavily charged fine droplets at the surface of which a corona discharge may be initiated [1, 25]. The resonant buildup of the fundamental mode ($n = 2$) renders a weakly charged (in terms of instability against the self-charge) drop in a cloud similar to a prolate spheroid. In this case, the surface of the drop may become unstable at the vertices of the spheroid, where the self-charge and polarization charge density increase (the charge is redistributed when the drop elongates) [1, 26, 27]. The problem is that, in second-order calculations for an ideal liquid drop nonlinearly vibrating in a vacuum, exact fundamental resonance is absent. It appears if the viscosity of the liquid is taken into account [23]. However, viscosity can today be considered only on a qualitative basis, since nonlinear vibrations of a viscous drop have not yet been explored adequately because of the complexity of the analysis. For an ideal liquid drop, fundamental resonance appears in third-order calculations (at four-mode interaction) [23]. However, the amplitude of the resonantly growing fundamental mode at four-mode interaction is small (of the second order of smallness), while, at three-mode resonant interaction, this effect has the first order of smallness [21, 28]. Therefore, it is of interest to study the effect of a nonzero-density insulating environment on the number of second-order three-mode resonances.

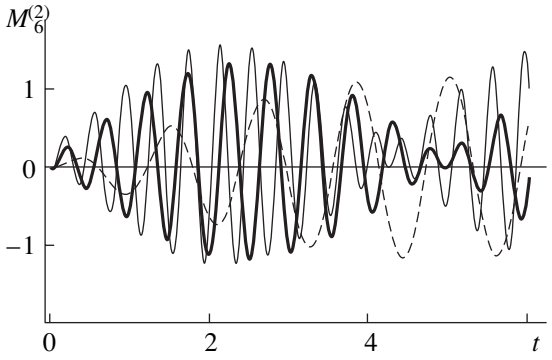

Fig. 2. Dependences $M_6^{(2)}(t)$ illustrating the nonlinear buildup of the sixth mode at the initially excited fourth mode with $W = 0$.

(4) As was noted, resonances appear when the frequencies of interacting modes of capillary vibration satisfy the condition $\omega_n^2 = (\omega_i \pm \omega_j)^2$. According to (22), when a charged drop vibrates in a medium, the frequency ω_n depends not only on the mode number n but also on the dimensionless parameter W and dimensionless density ρ. This means that the positions of resonances in the mode number space will depend on W and ρ, while, in a vacuum, they depend on W alone. Simple calculations show that an additional degree of freedom (associated with ρ) changes the resonant values of W (relative to those in a vacuum at $\rho = 0$). At $i, j \le 100$ and

Variation of degenerate resonance positions with the parameter W

W	n	m	ρ
0	8	5	0
	16	10	0
	24	15	0.28
	32	20	0.945
0.5	8	5	0.6352
	16	10	0.766
	24	15	1.965
	32	20	13.74
1	8	5	3.6
	16	10	6.68
2	30	19	0.3012
2.7	6	4	0.093
	14	9	0.608
	22	14	2.1675
	30	19	85.241

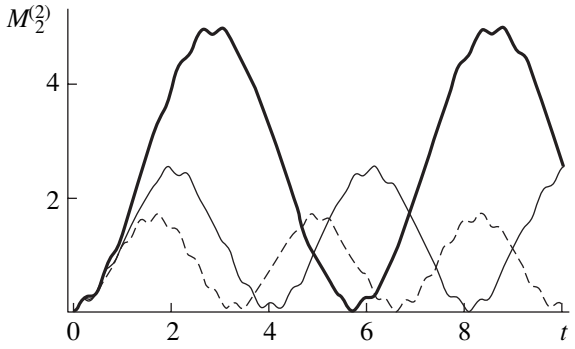

Fig. 3. Dimensionless second-order amplitude $M_2^{(2)}(t)$ of the fundamental mode vs. dimensionless time at the initial deformation defined by the fifth mode with $\rho = 1$. $W = 1$ (dashed line), 2 (thin line), and 3 (thick line).

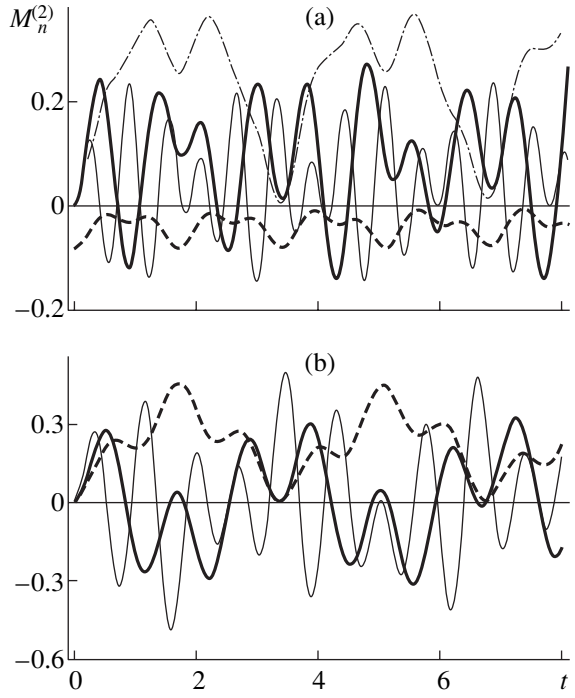

Fig. 4. Dimensionless second-order amplitudes $M_n^{(2)}(t)$ vs. dimensionless time at the initial deformation defined by the superposition of the second and third modes with $\rho = 1$ and $W = 1$. (a) Even modes: zeroth mode (dashed line), second mode (dash-and-dot line), fourth mode (thick line), and sixth mode (thin line). (b) Odd modes: first mode (dashed line), third mode (thick line), and fifth mode (thick line).

$W < 4$, the total number of resonances is several thousand; therefore, it is unreasonable to list them all. As an illustration, the table lists degenerate ($i = j$) resonant situations that may occur at a single-mode initial deformation with several fixed W.

(5) Calculations by (22) (Figs. 1, 2) were made for $\rho = 0.1$, 1, and 10. These values provide the most descriptive results, since the variation of ρ with $\rho < 0.1$ or >10 affects the capillary vibration frequencies insignificantly. In the limit $\rho \longrightarrow 0$, we are dealing with a liquid drop vibrating in a gas; in the limit $\rho \longrightarrow \infty$, with a gas bubble vibrating in a liquid. In practice, as was already noted, one often has to deal with charged drops suspended or moving in another immiscible dielectric fluid [2–9]. The most common situation in this case is characterized by $\rho \approx 1$, $\varepsilon_* \approx 2$, and the interfacial tension coefficient varying from $\sigma \approx 2.5$ dyn/cm (the paraffin oil–water interface [5]) to $\sigma \approx 50$ dyn/cm (the *n*-heptane–water [9] or silicone–water [4] interface). For vibration of charged drops in a gas, $\rho \approx 10^{-3}$, $\varepsilon_* \approx 1$, and $\sigma \approx 50$ dyn/cm [1]. Charged bubbles in a dielectric liquid are objects of extensive research as applied to charged helium surface stability [29], electrical discharges in liquids [30], and thermonuclear fusion [31]. For bubbles, we typically have $1 < \varepsilon_* < 80$, $\rho \sim 1000$, and $0.354 < \sigma < 70$ dyn/cm [29–31].

Figure 3 shows the time variation of the second-order fundamental mode amplitude at a fixed ρ and various W. In the dimensional form, the parameter $W = Q^2/4\pi\sigma R^3\varepsilon_*$ combines all physical quantities which are of importance for the phenomenon considered: interfacial tension coefficient, permittivity of the medium, drop charge, and drop radius. Note that the interfacial tension coefficient σ is related to the surface tension coefficients σ_1 and σ_2 of pure phases by the Antonov rule [32] $\sigma \approx |\sigma_1 - \sigma_2|$. In our case, σ_1 and σ_2 are the surface tension coefficients of phases contacting the gas. As a rule, σ is much smaller than σ_1 and σ_2; therefore, the interface is sometimes unstable against the self-charge [5–7]. As W grows (approaches the value $W_{cr} = 4$, which manifests instability of the drop against the self-charge [1]), so does the fundamental mode amplitude.

Figure 4 demonstrates the calculated results for an off-resonance situation where the initial deformation is defined by the superposition of the second and third modes with $h_2 = h_3 = 1/2$. The spectrum of second-order modes is seen to contain both even and odd modes with comparable amplitude factors (except for the zeroth mode).

CONCLUSIONS

In the case of an ideal incompressible conducting liquid drop nonlinearly vibrating in an ideal incompressible dielectric medium, the energy maximum in the spectrum of nonlinearly excited modes shifts toward the highest mode with increasing medium-to-drop density ratio, no matter which of the modes defines the initial deformation. With the effect of the environment taken into account, the number of resonant situations becomes dependent on the drop-to-environment density ratio and the resonant self-charge of the drop changes. It is shown that nonlinear vibrations may be of resonant character even if the charge of the drop is far away from exact resonant values. This is because Rayleigh subcritical values of the self-charge affect the frequencies of higher vibration modes insig-

nificantly. That is why the calculated amplitude of the nonlinearly excited fundamental mode is high even if the initial deformation is defined by higher modes.

ACKNOWLEDGMENTS

This work was supported by the Russian Foundation for Basic Research (grant no. 03-01-00760).

REFERENCES

1. A. I. Grigor'ev and S. O. Shiryaeva, Izv. Ross. Akad. Nauk, Mekh. Zhidk. Gaza, No. 3, 3 (1994).
2. A. I. Grigor'ev, Yu. V. Syshchikov, and S. O. Shiryaeva, Zh. Prikl. Khim. (Leningrad) **62**, 2020 (1989).
3. Ch. T. O'Konski and F. E. Harris, J. Phys. Chem. **61**, 1172 (1957).
4. C. G. Garton and Z. Krasucki, Trans. Faraday Soc. **60**, 211 (1964).
5. S. A. Ryce and R. R. Wyman, Can. J. Phys. **42**, 2185 (1964).
6. S. A. Ryce and D. A. Patriarche, Can. J. Phys. **43**, 2192 (1965).
7. M. N. Beĭtuganov, Meteorol. Gidrol., No. 9, 42 (1989).
8. N. C. Scott, O. I. Basaran, and C. H. Byers, Ind. Eng. Chem. Res. **2**, 901 (1990).
9. T. Mochizuki, Y. H. Mory, and N. Kaji, AIChE. J. **36**, 1039 (1990).
10. I. P. Stakhanov, Zh. Tekh. Fiz. **44**, 1373 (1974) [Sov. Phys. Tech. Phys. **19**, 861 (1974)].
11. A. I. Grigor'ev, S. O. Shiryaeva, and V. A. Koromyslov, Zh. Tekh. Fiz. **68** (9), 1 (1998) [Tech. Phys. **43**, 1011 (1998)].
12. S. O. Shiryaeva, A. I. Grigor'ev, and D. F. Belonozhko, Zh. Tekh. Fiz. **69** (10), 34 (1999) [Tech. Phys. **44**, 1159 (1999)].
13. J. A. Tsamopoulos and R. A. Brown, J. Fluid Mech. **147**, 373 (1984).
14. Z. C. Feng, J. Fluid Mech. **333**, 1 (1997).
15. S. O. Shiryaeva, Zh. Tekh. Fiz. **71** (2), 27 (2001) [Tech. Phys. **46**, 158 (2001)].
16. S. O. Shiryaeva, Zh. Tekh. Fiz. **72** (4), 15 (2002) [Tech. Phys. **47**, 389 (2002)].
17. A. R. Gaibov and A. I. Grigor'ev, Zh. Tekh. Fiz. **73** (7), 13 (2003) [Tech. Phys. **48**, 813 (2003)].
18. D. A. Varshalovich, A. N. Moskalev, and V. K. Khersonskiĭ, *Quantum Theory of Angular Momentum* (Nauka, Leningrad, 1975; World Sci., Singapore, 1988).
19. A. R. Gaibov, S. O. Shiryaeva, A. I. Grigor'ev, and D. F. Belonozhko, Pis'ma Zh. Tekh. Fiz. **29** (4), 22 (2003) [Tech. Phys. Lett. **29**, 138 (2003)].
20. S. O. Shiryaeva, Pis'ma Zh. Tekh. Fiz. **26** (22), 76 (2000) [Tech. Phys. Lett. **26**, 1016 (2000)].
21. S. O. Shiryaeva, Zh. Tekh. Fiz. **73** (2), 19 (2003) [Tech. Phys. **48**, 152 (2003)].
22. S. O. Shiryaeva, A. I. Grigor'ev, and D. F. Belonozhko, Pis'ma Zh. Tekh. Fiz. **28** (22), 45 (2002) [Tech. Phys. Lett. **28**, 945 (2002)].
23. S. O. Shiryaeva, A. I. Grigor'ev, and D. F. Belonozhko, Pis'ma Zh. Tekh. Fiz. **29** (6), 69 (2003) [Tech. Phys. Lett. **29**, 249 (2003)].
24. *Clouds and Cloudy Atmosphere: A Handbook*, Ed. by I. P. Mazin, A. Kh. Khrgian, and I. M. Imyanitov (Gidrometeoizdat, Leningrad, 1987).
25. A. I. Grigor'ev and S. O. Shiryaeva, J. Phys. D **23**, 1361 (1990).
26. A. I. Grigor'ev, S. O. Shiryaeva, and S. I. Shchukin, Zh. Tekh. Fiz. **68** (7), 33 (1998) [Tech. Phys. **43**, 774 (1998)].
27. S. I. Shchukin and A. I. Grigor'ev, Zh. Tekh. Fiz. **69** (8), 49 (1999) [Tech. Phys. **44**, 913 (1999)].
28. M. I. Rabinovich and D. I. Trubetskov, *An Introduction to the Theory of Vibrations and Waves* (Nauka, Moscow, 1984).
29. N. M. Zubarev, Zh. Éksp. Teor. Fiz. **121**, 624 (2002) [JETP **94**, 534 (2002)].
30. A. N. Zharov and S. O. Shiryaeva, Élektrokhim. Obrab. Met., No. 6, 9 (1999).
31. I. N. Didenkulov, D. A. Selivanovskiĭ, V. E. Semenov, and I. V. Sokolov, Izv. Vyssh. Uchebn. Zaved. Radiofiz. **42**, 183 (1999).
32. R. C. Reid and T. K. Sherwood, *Properties of Gases and Liquids* (McGraw-Hill, New York, 1966; Khimiya, Moscow, 1971).

Translated by V. Isaakyan

Technical Physics, Vol. 49, No. 1, 2004, pp. 30–38. Translated from Zhurnal Tekhnicheskoĭ Fiziki, Vol. 74, No. 1, 2004, pp. 32–39.

GASES AND LIQUIDS

Nonlinear Periodic Waves on the Charged Free Surface of a Perfect Fluid

A. V. Klimov, D. F. Belonozhko, and A. I. Grigor'ev

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

e-mail: grig@uniyar.ac.ru

Received May 26, 2003

Abstract—Analytical expressions for the profile of a nonlinear wave and for a nonlinear correction to its frequency are derived in the fourth-order approximation in amplitude of a periodic traveling wave on a uniformly charged free surface of an infinitely deep perfect incompressible fluid. It is found that corrections to the amplitude and frequency of the nonlinear wave are absent if the problem is solved under the initial condition that provides the constancy of the first-order amplitude and wavelength in time. Nonlinear analysis of conditions for instability of the fluid free surface against the surface charge shows that the critical charge density and wavenumber of the least stable wave are not constant (as in the linear theory) and decrease with growing amplitude of the wave. © *2004 MAIK "Nauka/Interperiodica".*

(1) Critical conditions for instability of the charged surface of a fluid are of great interest for science and applications [1, 2]. Yet, most of the theoretical studies concerning the stability of infinitesimal-amplitude capillary waves have been carried out in a linear approximation [2], although the obvious nonlinear nature of the phenomenon, which follows from the nonlinear equations of hydrodynamics, has been corroborated in multiple experiments [1, 3–5]. Methods of asymptotic analysis of nonlinear capillary gravitational waves on the neutral fluid surface are well known (see, e.g., [6–10] and Refs. therein) and may be applied to the problem posed. This is the subject of this investigation. Note that nonlinear waves on the charged surface of a perfect fluid have already been studied [11–14]; however, those works were aimed at finding soliton solutions and the methods used in them did not allow the researchers to obtain nonlinear corrections to the frequencies and critical conditions for instability.

(2) Let a perfect incompressible ideally conducting fluid of density ρ and surface tension coefficient γ occupy the half-space $z < 0$ of the Cartesian coordinate system in the gravitational field and let the unit vector $\mathbf{e}_z$ of this system be directed opposite to the gravitational acceleration, $\mathbf{e}_z \| \mathbf{g}$. It is also assumed that the free surface of the fluid is charged and a uniform electrostatic field E_0 parallel to the unit vector $\mathbf{e}_z$ is present near the surface.

We will consider a plane wave traveling over the free surface of a fluid making contact with a vacuum in the direction of the unit vector $\mathbf{e}_x$. Then, the distortion of the free fluid surface due to the traveling wave, the wave velocity field, and the pressure and electric field distributions near the distorted surface will depend on time t and the coordinates x and z. A motion equation for the free fluid surface distorted by small-amplitude wave motion will have the form $z = \xi(x, t)$.

Our goal is to find the time-invariable profile of the traveling wave and a frequency correction that is nonlinear in amplitude in the fourth-order approximation in wave amplitude, which is considered to be small compared with the wavelength. It should be noted that the purely sinusoidal profile persists only for infinitesimal-amplitude waves and becomes unsteady even in second-order calculations. Therefore, classical works on the theory of finite-amplitude waves in a perfect incompressible fluid [6–10, 15, 16] treat the second-order calculation of finite-amplitude waves as a problem of finding the time-invariable wave profile. It will be shown below that the appearance of a nonlinear frequency correction (quadratic in small parameter) in higher-than-second-order calculations will cause a difference between the phase velocities of the wave linear in small parameter and of nonlinear corrections to it. This means that the profile of finite-amplitude waves remains stationary only over limited time intervals.

(3) In view of the aforesaid, the phenomenon can be mathematically stated in the form of the boundary-value problem

$$z > \xi: \nabla^2\Phi = 0; \tag{1}$$

$$z < \xi: \nabla^2\varphi = 0; \tag{2}$$

$$p = p_0 - \rho g z - \rho\frac{\partial\varphi}{\partial t} - \frac{\rho}{2}\left(\left(\frac{\partial\varphi}{\partial x}\right)^2 + \left(\frac{\partial\varphi}{\partial z}\right)^2\right); \tag{3}$$

$$z = \xi: \frac{\partial\xi}{\partial t} + \frac{\partial\xi}{\partial x}\frac{\partial\varphi}{\partial x} = \frac{\partial\varphi}{\partial z}; \tag{4}$$

1063-7842/04/4901-0030$26.00 © 2004 MAIK "Nauka/Interperiodica"

$$p + \frac{(\nabla\Phi)^2}{8\pi} = -\gamma\frac{\partial^2\xi}{\partial x^2}\left(1 + \left(\frac{\partial\xi}{\partial x}\right)^2\right)^{-3/2}; \quad (5)$$

$$\Phi = 0; \quad (6)$$

$$z \longrightarrow \infty: \nabla\Phi \longrightarrow -E_0\mathbf{e}_z; \quad (7)$$

$$z \longrightarrow -\infty: \nabla\varphi \longrightarrow 0. \quad (8)$$

For the problem to be solved uniquely, it is necessary to set initial conditions. For problems of this type, the choice of initial conditions is a delicate subject, since arbitrary initial conditions set in advance may make the solution extremely awkward. In fact, the need for setting initial conditions is replaced by the search for a solution the least awkward in mathematical terms. Such an approach is used in this article.

We assume that, in a zeroth approximation, the free surface is undisturbed and is described by the equation $z = 0$, the fluid is at rest, and the electric field is uniform throughout the space:

$$\xi_0 \equiv 0; \quad \nabla\varphi_0 \equiv 0; \quad \nabla\Phi_0 \equiv -E_0\mathbf{e}_z; \quad p_0 = -\frac{E_0^2}{8\pi}.$$

Substituting these expressions into (1)–(8) yields

$$\Phi_0 \equiv -E_0 z.$$

(4) Let us separate different orders of smallness from our problem. Before doing so, we eliminate the pressure function $p(\mathbf{r}, t)$ from consideration by substituting expression (3) for pressure into dynamic condition (5). The latter then takes the form

$$-\frac{E_0^2}{8\pi} - \rho g\xi - \rho\frac{\partial\varphi}{\partial t} - \frac{\rho}{2}\left(\left(\frac{\partial\varphi}{\partial x}\right)^2 + \left(\frac{\partial\varphi}{\partial z}\right)^2\right) + \frac{1}{8\pi}$$

$$\times\left(\left(\frac{\partial\Phi}{\partial x}\right)^2 + \left(\frac{\partial\Phi}{\partial z}\right)^2\right) = -\gamma\frac{\partial^2\xi}{\partial x^2}\left(1 + \left(\frac{\partial\xi}{\partial x}\right)^2\right)^{-\frac{3}{2}}.$$

The unknown functions of the problem are the free surface disturbance ξ, the velocity potential φ, and the electric potential Φ. They will be sought as expansions in a small parameter:

$$\xi = \varepsilon\xi_1 + \varepsilon^2\xi_2 + \varepsilon^3\xi_3 + \varepsilon^4\xi_4 + O(\varepsilon^5); \quad (9)$$

$$\Phi = -E_0 z + \varepsilon\Phi_1 + \varepsilon^2\Phi_2 + \varepsilon^3\Phi_3 + \varepsilon^4\Phi_4 + O(\varepsilon^5); \quad (10)$$

$$\varphi = \varepsilon\varphi_1 + \varepsilon^2\varphi_2 + \varepsilon^3\varphi_3 + \varepsilon^4\varphi_4 + O(\varepsilon^5);$$
$$\xi_n \sim O(1); \quad \Phi_n \sim O(1); \quad \varphi_n \sim O(1). \quad (11)$$

Here, ε is a small dimensionless parameter defined as the wave amplitude a times the wavenumber k.

The problem will be solved by the method of many scales from the perturbation theory. We assume that the unknown functions ξ_n, φ_n, and Φ_n depend on the coordinates x and z and on various time scales: the basic scale ($T_0 = t$) and slower ones ($T_1 = \varepsilon t$, $T_2 = \varepsilon^2 t$, $T_3 = \varepsilon^3 t$, etc.); that is,

$$\xi_n = \xi_n(T_0, T_1, T_2, T_3, x); \quad \Phi_n = \Phi_n(T_0, T_1, T_2, T_3, x, z);$$
$$\varphi_n = \varphi_n(T_0, T_1, T_2, T_3, x, z).$$

Then, the operator of time differentiation takes the form

$$\frac{\partial}{\partial t} = \frac{\partial}{\partial T_0} + \varepsilon\frac{\partial}{\partial T_1} + \varepsilon^2\frac{\partial}{\partial T_2} + \varepsilon^3\frac{\partial}{\partial T_3} + O(\varepsilon^4). \quad (12)$$

Let us expand the boundary conditions on the free surface in deviation of the surface from the equilibrium plane form (i.e., in ε) in the vicinity of $z = 0$. In the stricter sense, all the z-dependent quantities entering into conditions (4)–(6) will be expanded into the Taylor series in the vicinity of $z = 0$. Also, we will expand $\partial\xi/\partial t$ and $(\partial^2\xi/\partial x^2)/(1 + (\partial\xi/\partial x)^2)^{3/2}$ in powers of the small parameter. When expanding the partial derivatives $(\partial\xi/\partial t)$ and $(\partial\varphi/\partial t)$, we will take into account expression (12). Substituting expansions (9)–(11) into (1), (2), (7), and (8), collecting terms with the same powers of ε, and equating them to zero, we split the problem into orders of smallness from the first to the fourth.

(5) Mathematically, the first-order problem is stated as follows:

$$z > 0: \nabla^2\Phi_1 = 0;$$

$$z < 0: \nabla^2\varphi_1 = 0;$$

$$z = 0: \frac{\partial\xi_1}{\partial T_0} - \frac{\partial\varphi_1}{\partial z} = 0;$$

$$-\rho g\xi_1 - \rho\frac{\partial\varphi_1}{\partial T_0} - \frac{E_0}{4\pi}\frac{\partial\Phi_1}{\partial z} + \gamma\frac{\partial^2\xi_1}{\partial x^2} = 0;$$

$$\Phi_1 - E_0\xi_1 = 0;$$

$$z \longrightarrow \infty: \nabla\Phi_1 \longrightarrow 0;$$

$$z \longrightarrow -\infty: \nabla\varphi_1 \longrightarrow 0.$$

The first-order problem is easily solved by the conventional methods [17, 18]:

$$\xi_1 = \frac{1}{2}\zeta\exp(i\theta) + \frac{1}{2}\bar{\zeta}\exp(-i\theta);$$

$$\varphi_1 = -\frac{i\omega}{2k}\zeta\exp(kz)\exp(i\theta) + \frac{i\omega}{2k}\bar{\zeta}\exp(kz)\exp(-i\theta);$$

$$\Phi_1 = \frac{E_0}{2}\zeta\exp(-kz)\exp(i\theta) + \frac{E_0}{2}\bar{\zeta}\exp(-kz)\exp(-i\theta);$$

$$\omega^2 \equiv gk + \frac{\gamma k^3}{\rho} - \frac{E_0^2 k^2}{4\pi\rho}; \quad \theta \equiv kx - \omega T_0.$$

Now, we can set the initial condition that will be used in the subsequent consideration. Let us assume that all

nonlinear corrections to the first-order wave profile (i.e., to the function $\zeta_1 = \zeta_1(T_0, x)$) depend on $m\theta$, where m is an integer greater than unity. In other words, we assume that the amplitude factors multiplying possible nonlinear corrections to the wave profile with the argument $\theta = 0$ equal zero. It turns out that such an initial condition provides a simple form of solutions [8, 9].

(6) The second-order problem is stated as

$$z > 0: \nabla^2\Phi_2 = 0;$$

$$z < 0: \nabla^2\varphi_2 = 0;$$

$$z = 0: \frac{\partial\xi_2}{\partial T_0} - \frac{\partial\varphi_2}{\partial z} = \xi_1\frac{\partial^2\varphi_1}{\partial z^2} - \frac{\partial\xi_1}{\partial T_1} - \frac{\partial\varphi_1}{\partial x}\frac{\partial\xi_1}{\partial x};$$

$$-\rho g\xi_2 - \rho\frac{\partial\varphi_2}{\partial T_0} - \frac{E_0}{4\pi}\frac{\partial\Phi_2}{\partial z} + \gamma\frac{\partial^2\xi_2}{\partial x^2} = \rho\frac{\partial\varphi_1}{\partial T_1}$$

$$+ \rho\xi_1\frac{\partial^2\varphi_1}{\partial z\partial T_0} + \frac{\rho}{2}\left(\frac{\partial\varphi_1}{\partial z}\right)^2 + \frac{\rho}{2}\left(\frac{\partial\varphi_1}{\partial x}\right)^2$$

$$-\frac{1}{8\pi}\left(\frac{\partial\Phi_1}{\partial z}\right)^2 + \frac{E_0}{4\pi}\xi_1\frac{\partial^2\Phi_1}{\partial z^2} - \frac{1}{8\pi}\left(\frac{\partial\Phi_1}{\partial x}\right)^2;$$

$$\Phi_2 - E_0\xi_2 = -\xi_1\frac{\partial\Phi_1}{\partial z};$$

$$z \longrightarrow \infty: \nabla\Phi_2 \longrightarrow 0;$$

$$z \longrightarrow -\infty: \nabla\varphi_2 \longrightarrow 0.$$

A solution to this problem is easy to find:

$$\xi_2 = \frac{k(2\pi(k^2\gamma + g\rho) - E_0^2k)}{8\pi(\rho g - 2\gamma k^2)}$$

$$\times(\zeta^2\exp(2i\theta) + \bar{\zeta}^2\exp(-2i\theta));$$

$$\varphi_2 = -\frac{i\omega k(6\pi k\gamma - E_0^2)}{8\pi(\rho g - 2\gamma k^2)}$$

$$\times(\zeta^2\exp(2i\theta) - \bar{\zeta}^2\exp(-2i\theta))\exp(2kz);$$

$$\Phi_2 = \frac{E_0k}{2}\zeta\bar{\zeta} + \frac{E_0k(2\pi(-k^2\gamma + 2g\rho) - E_0^2k)}{8\pi(\rho g - 2\gamma k^2)}$$

$$\times(\zeta^2\exp(2i\theta) + \bar{\zeta}^2\exp(-2i\theta))\exp(-2kz).$$

(7) The third-order problem is stated as

$$z > 0: \nabla^2\Phi_3 = 0;$$

$$z < 0: \nabla^2\varphi_3 = 0;$$

$$z = 0: \frac{\partial\xi_3}{\partial T_0} - \frac{\partial\varphi_3}{\partial z} = \Omega_{31}; \quad \Phi_3 - E_0\xi_3 = \Omega_{33};$$

$$-\rho g\xi_3 - \rho\frac{\partial^2\varphi_3}{\partial T_0} - \frac{E_0}{4\pi}\frac{\partial\Phi_3}{\partial z} + \gamma\frac{\partial^2\xi_3}{\partial x^2} = \Omega_{32};$$

$$z \longrightarrow \infty: \nabla\Phi_3 \longrightarrow 0;$$

$$z \longrightarrow -\infty: \nabla\varphi_3 \longrightarrow 0.$$

Expressions for the inhomogeneity functions Ω_{31}, Ω_{32}, and Ω_{33} on the rights of the initial conditions on the free surface, which depend on the first- and second-order solutions, are given in Appendix A.

A solution to the resultant inhomogeneous problem is obtained by applying tedious while mathematically straightforward calculations:

$$\xi_3 = X(\zeta^3\exp(3i\theta) + \bar{\zeta}^3\exp(-3i\theta));$$

$$X = \frac{k^2(32\alpha^2k^2W^2 + 32\alpha kW(1 - \alpha^2k^2W) + 6\alpha^4k^4 + 21\alpha^2k^2 + 6)}{32(1 - 2\alpha^2k^2)(1 - 3\alpha^2k^2)};$$

$$\varphi_3 = Y_{31}\zeta\bar{\zeta}(i\zeta\exp(i\theta) - i\bar{\zeta}\exp(-i\theta))\exp(kz)$$

$$+ Y_{33}(i\zeta^3\exp(3i\theta) - i\bar{\zeta}^3\exp(-3i\theta))\exp(3kz);$$

$$Y_{33} = -\frac{\alpha k^2\omega(32\alpha kW^2 - 104\alpha^2k^2W - 8W + 78\alpha^3k^3 + 15\alpha k)}{32(1 - 2\alpha^2k^2)(1 - 3\alpha^2k^2)};$$

$$Y_{31} = \frac{gk^2(2\alpha kW(8\alpha kW - 12\alpha^2k^2 - 9) + 14\alpha^4k^4 + 25\alpha^2k^2 + 2)}{32\omega(1 - 2\alpha^2k^2)};$$

$$\Phi_3 = \Theta_{31}\zeta\bar{\zeta}(\zeta\exp(i\theta) + \bar{\zeta}\exp(-i\theta))\exp(-kz)$$

$$+ \Theta_{33}(\zeta^3\exp(3i\theta) + \bar{\zeta}^3\exp(-3i\theta))\exp(-3kz);$$

$$\Theta_{33} = \frac{E_0 k^2 (8\alpha kW(4\alpha kW + 5\alpha^2 k^2 - 7) + 6\alpha^4 k^4 - 33\alpha^2 k^2 + 24)}{32(1 - 2\alpha^2 k^2)(1 - 3\alpha^2 k^2)};$$

$$\Theta_{31} = \frac{E_0 k^2 (4\alpha^2 k^2 + 7 - 12\alpha kW)}{16(1 - 2\alpha^2 k^2)};$$

$$W = \frac{E_0^2}{4\pi\rho g\gamma}; \quad \alpha \equiv \sqrt{\frac{\gamma}{\rho g}}.$$

Here, α is the capillary constant of the fluid and W is the Tonks–Frenkel parameter, which characterizes the stability of a flat uniformly charged free surface of a conducting liquid [1, 2].

A result of solving the third-order problem is the dependence of the wave amplitude ζ on the time scales T_2 and T_3:

$$\zeta = \zeta_1 \exp(i\beta_0)\exp(i\delta\zeta_1^2 T_2);$$

$$\delta = \frac{gk^3(\alpha^2 k^2(1 + 2\alpha^2 k^2) + 8 - 16\alpha kW(1 + \alpha^2 k^2 - \alpha kW))}{16\omega(1 - 2\alpha^2 k^2)},$$

where ζ_1 and β_0 are functions that depend only on T_3 and are found by solving higher order problems.

(8) The fourth-order problem is stated as

$$z > 0: \nabla^2 \Phi_4 = 0;$$

$$z < 0: \nabla^2 \varphi_4 = 0;$$

$$z = 0: \frac{\partial \xi_4}{\partial T_0} - \frac{\partial \varphi_4}{\partial z} = \Xi_{41};$$

$$-\rho g\xi_4 - \rho\frac{\partial \varphi_4}{\partial T_0} - \frac{E_0}{4\pi}\frac{\partial \Phi_4}{\partial z} + \gamma\frac{\partial^2 \xi_4}{\partial x^2} = \Xi_{42};$$

$$\Phi_4 - E_0\xi_4 = \Xi_{43};$$

$$z \longrightarrow \infty: \nabla\Phi_4 \longrightarrow 0;$$

$$z \longrightarrow -\infty: \nabla\varphi_4 \longrightarrow 0.$$

Expressions for the inhomogeneity functions Ξ_{41}, Ξ_{42}, and Ξ_{43} on the rights of the initial conditions, which depend on the first-, second-, and fourth-order solutions, are given in Appendix B.

With the initial condition adopted, the fourth-order solution has the form

$$\xi_4 = \varphi_4 = \Phi_4 \equiv 0; \quad \zeta_1 = \frac{1}{k}; \quad \beta_0 = 0.$$

Thus, a solution to the entire problem lacks fourth-order quantities.

(9) Final expressions for the profile of a nonlinear wave on the uniformly charged free surface of a fluid, velocity field potential, and electric field potential are as follows:

$$\xi = a\cos(\theta_*) + a^2\Lambda\cos(2\theta_*) + 2a^3 X\cos 3\theta;$$

$$\varphi = a\frac{\omega}{k}\exp(kz)\sin(\theta_*)$$

$$+ a^2\frac{\omega\alpha k(3\alpha k - 2W)}{2(1 - 2\alpha^2 k^2)}\exp(2kz)\sin(2\theta_*)$$

$$+ a^3(-2Y_{31}\exp(kz)\sin\theta - 2Y_{33}\exp(3kz)\sin 3\theta); \quad (13)$$

$$\Phi = -E_0 z + aE_0\exp(-kz)\cos(\theta_*)$$

$$+ a^2\frac{E_0 k}{2}\left(1 + \frac{2 - \alpha^2 k^2 - 2\alpha kW}{1 - 2\alpha^2 k^2}\exp(-2kz)\cos(2\theta_*)\right)$$

$$+ a^3(2\Theta_{31}\exp(-kz)\cos\theta + 2\Theta_{33}\exp(-3kz)\cos 3\theta),$$

where

$$\Lambda \equiv \frac{k(1 + \alpha^2 k^2 - 2\alpha kW)}{2(1 - 2\alpha^2 k^2)};$$

$$\theta_* \equiv \theta + \delta a^2 t \equiv kx - \omega t + \delta a^2 t.$$

(10) From the final solution and expression (13) for wave profile, it follows that the time-invariable profile of a finite-amplitude wave is a matter of convention. Indeed, the wave profile is stationary up to the second order of smallness. In this case, the nonlinear correction to the frequency is absent and the wave profile depends on the first two terms in (13) with $\delta = 0$ ($\theta_* \equiv \theta$). In third-order calculations, the frequency correction nonlinear in amplitude appears in (13) in the term linear in α and is absent in the term quadratic in a. It should be noted that in [8, 9] third-order nonlinear frequency corrections are involved in all the terms (from linear to cubic) of the expression describing the profile, which seems to be incorrect. In fact, if we expanded in (13) the cosine involved in the quadratic (in small parameter) term in powers of the small parameter in the vicinity of θ, the term proportional to δ would have at least the fourth order of smallness and it would be incorrect to give this term in third-order calculations. In the fourth-order calculations made above, the nonlinear correction

to the frequency appears legitimately in the quadratic term, while the undisturbed argument θ figures in the term cubic in small parameter.

As follows from the aforesaid, the phase velocities of various components of the wave profile that are calculated in both the third- and the fourth-order approximation will differ, causing the distortion of the profile in time. It only remains for us to obtain the time-invariable profile in the fifth-order approximation if a nonlinear correction to the profile turns out to be zero (as in the fourth-order approximation). In this case, however, a nonlinear correction to the frequency may be legitimately reduced to a cubic correction to the amplitude (the third term in (13)).

(11) The expression obtained for the profile of a nonlinear wave on the charged surface of a perfect fluid coincides, in the limit $W \longrightarrow 0$, with the known [8, 9] expression for the profile of a nonlinear capillary gravitational wave on the neutral surface of a perfect fluid (in view of the remark that corrections to the frequency cannot be included in the terms of the profile that are nonlinear in amplitude). From the expressions given above, one sees that the amplitude factor Λ multiplying the second-order correction grows resonantly at $k = k_2 \equiv 1/\alpha\sqrt{2}$. The amplitude factor X multiplying the third-order correction exhibits two resonances: at $k = k_2$ and $k = k_3 \equiv 1/\alpha\sqrt{3}$. Upon resonant interaction in the quadratic approximation (where only one resonance takes place), the energy is transferred from longer waves with the wavenumbers $k = k_2$ to shorter waves with the wavenumbers $k = 2k_2$ [19]. According to [13], the energy will be transferred in the same direction at the resonance $k = k_3$: from longer waves with the numbers $k = k_3$ to waves with $k = 3k_3$. However, the effect in this case will have a higher order of smallness.

From (13), it also follows that the nonlinear addition to the frequency is proportional to the wave amplitude *a* squared and negative at $k > k_2$ (the range of interest for investigating the stability of the fluid charged surface) and also has the third order of smallness. It is of interest that the nonlinear correction to the frequency, as well as the amplitude factors Λ and X, has resonance form. Therefore, expression (13) is of limited applicability near the wavenumbers $k = k_2$ and k_3, since both the amplitude factors and the correction to the frequency must be small compared with first-order quantities.

Now we take into account that Tonks–Frenkel instability occurs when the virtual wave frequency squared passes through zero and the first derivative of the frequency with respect to wavenumber vanishes (from the latter condition, one determines the wavenumber of the wave that has a maximal instability increment) [18, 20]. In the linear approximation, the critical values of the Tonks–Frenkel parameter $W = W_*$ and wavenumber $k = k_8$ are as follows [18]:

$$k_* = \alpha^{-1}; \quad W = k_*\alpha + k_*^{-1}\alpha^{-1} = 2.$$

In the nonlinear situation considered, a set of algebraic equations for *W* and *k* is generally very awkward. However, since of interest for us are order-of-magnitude estimates, we expand the Tonks–Frenkel parameter and wavenumber in the wave amplitude squared:

$$W \approx 2 - w\varepsilon^2; \quad k\alpha \approx 1 - \kappa\varepsilon^2.$$

Substituting these expansions into the set of equations for determining instability conditions, one easily finds that $w = 11/8$ and $\kappa = 23/16$.

Thus, the nonlinear analysis shows that the charge density at which the surface of the fluid becomes unstable and the wavenumber of the least stable mode are lower than those predicted from the linear theory.

We also recall that calculations in the fourth order of smallness in wave amplitude show that a term proportional to the third power of the amplitude is absent in the nonlinear correction to the frequency. Whether or not a frequency correction proportional to the amplitude in the fourth power is other than zero will be elucidated in fifth-order calculations. Such calculations would help in tracing the further variation of nonlinear corrections to the instability conditions for the charged surface of a fluid.

(12) Further analysis of expression (13) will be restricted to the first two terms, which provide the time-invariable profile of a finite-amplitude wave in the approximation used here.

As was noted, the amplitude factor Λ in the second-order term has resonant form: with $k = k_2$, the denominator of the expression for Λ vanishes and the correction tends to infinity. This phenomenon has been investigated in [8] as applied to nonlinear waves on the neutral surface of a fluid. Unlike the case of the fluid neutral surface, the numerator of the expression for Λ in (13) contains the parameter *W* with negative sign, which characterizes the stability of the charged surface against the pressure of the self-charge electric field. This means that, for a certain relationship between the physical parameters, Λ may remain finite even if its denominator vanishes. The denominator of Λ goes to zero at $k \longrightarrow k_2$; however, if simultaneously *W* tends to $W_x = 3/2\sqrt{2} \approx 1.06$, the numerator of Λ will also tend to zero. Eventually, in the limit $k \longrightarrow k_2$ with $W = W_x$, we have uncertainty of type 0/0, which is eliminated by the l'Hospital rule to yield $\Lambda = 1/8$. The dependence $\Lambda = \Lambda_*(\alpha k)$ becomes continuous. It is depicted by the dashed line in Fig. 1, which plots the dimensional factor Λ vs. dimensional wavenumber αk for various *W*.

Note that, as *k* approaches k_2, expansion (13) can no longer represent the solution irrespective of *W*, since the correction quadratic in the dimensionless amplitude *a*

becomes much greater than the term linear in a. The effect of viscosity, which plays an important part in resonance phenomena, is of interest in this respect and will be demonstrated with water.

According to [21, 22], where nonlinear waves on the neutral surface of a viscous fluid were investigated, the effect of viscosity on the wave profile for water is essential within the interval $\alpha k \in D \equiv [0.6, 0.7]$. At the extremities of this interval, the second-order amplitude correction calculated without considering the effect of viscosity turns out to be overestimated by several percent. However, at $k \longrightarrow k_2$, the overestimation becomes infinitely large. Beyond this interval, the models of viscous and inviscid water give similar results. Therefore, analysis which follows is performed for waves with $\alpha k \notin D$. The parts of the curves that fall into the domain D in Fig. 1 are physically meaningless, since they are constructed in the range where the initial inviscid model is invalid.

From Fig. 1, it is seen that the physically meaningful parts of the curves $\Lambda = \Lambda(\alpha k)$ tend to the curve $\Lambda = -\Lambda_*(\alpha k)$ in the limit $W \longrightarrow W_x$. It is natural to consider the value $W = W_*$ as a criterion for separating different wave motions. At this value of W, the asymptotic value of Λ in the curves analyzed changes sign in the limit $k \longrightarrow k_2$. It is also seen that the second-order addition to the linear part of the solution tends to zero at $W \longrightarrow W_x$.

Figure 1a shows a family of curves $\Lambda = \Lambda(\alpha k)$ constructed at various $W \le W_x$. In this range of W, the second-order amplitude addition varies with the wavenumber in the same manner as in the case of the neutral surface: for longer waves with $k < k_2$, the profiles have a pointed top; for shorter waves with $k > k_2$, the top is blunt (for details, see [8, 9]).

Figure 1b shows curves $\Lambda = \Lambda(\alpha k)$ calculated for various W from the range $W_x \le W \le 2$. It is easy to see that these curves are in a sense reversed relative to those shown in Fig. 1a: the right-hand physically meaningful parts of these curves, which correspond to shorter waves, are positive in an extended right-hand vicinity of the point $k = k_2$ and not negative as in Fig. 1a (branch *1* reaches the negative range away from $k = k_2$). Therefore, nonlinear waves associated with these curves can be naturally viewed as a previously unknown type of surface-charge-related periodic wave motion on the surface of a perfect fluid. Figure 2 compares the profiles of these waves with those of waves on the uncharged surface. The profiles of capillary gravitational waves on the charged and uncharged surface differ substantially. It is more appropriate to call waves discovered on the charged surface of a perfect fluid electrocapillary gravitational, or simply electrocapillary, waves, since the gravitational field may be neglected for the wavenumber range considered.

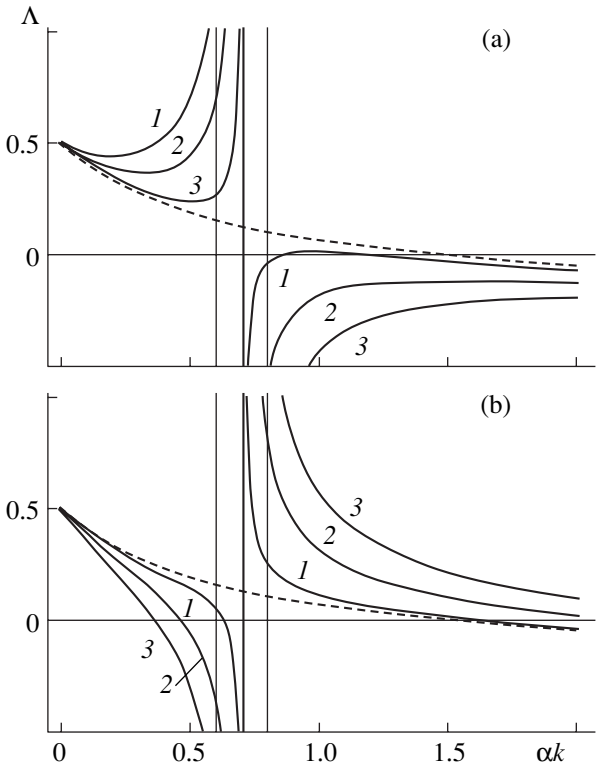

Fig. 1. Dimensionless amplitude factor Λ vs. dimensionless parameter αk for various W. (a) (*1*) $W = W_x - 0.05$, (*2*) $W = W_x - 0.25$, and (*3*) $W = W_x - 0.5$. (b) (*1*) $W = W_x + 0.05$, (*2*) $W = W_x + 0.25$, and (*3*) $W = W_x + 0.5$.

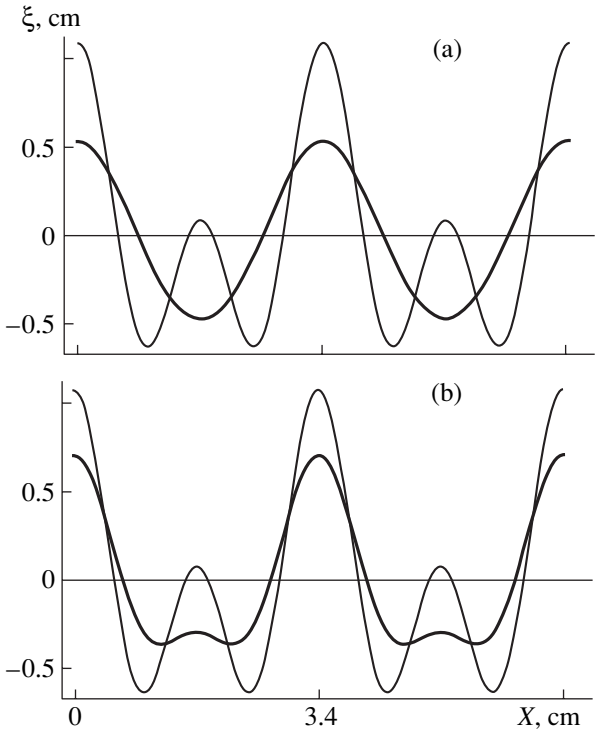

Fig. 2. Wave profiles on the water surface at $W = 0$ (thin curve) and 1.2 (thick curve) for αk = (a) 0.5 and (b) 0.9.

CONCLUSIONS

Thus, the profile of a periodic capillary gravitational wave traveling on the uniformly charged surface of a perfect incompressible fluid is found to be unsteady when calculated in asymptotic approximations of higher-than-second order: it diffuses because the phase velocities of corrections of different order of smallness to the profile differ.

A nonlinear correction to the wave frequency that depends on the amplitude squared appears in third-order calculations, has resonant form, and generates nonlinear corrections to the critical conditions under which the free surface becomes unstable against surface charge. In the fourth-order approximation, the critical charge density and the wavenumber of the least stable wave decrease in proportion to the wave amplitude squared.

Nonlinear waves on the charged surface differ substantially from those on the neutral surface. The curvature of the electrocapillary wave tops grows with the surface charge density (i.e., with the parameter W) for $W_x \lhd W \longrightarrow 2$. With $W \longrightarrow W_x$, the electric charge serves to decrease the efficiency of nonlinear interaction, since the second-order correction tends to zero in this case.

APPENDIX A

$$\Omega_{31} \equiv \xi_2\frac{\partial^2\varphi_1}{\partial z^2} + \xi_1\frac{\partial^2\varphi_2}{\partial z^2} + \frac{1}{2}\xi_1^2\frac{\partial^3\varphi_1}{\partial z^3} - \frac{\partial\xi_1}{\partial T_2} - \frac{\partial\xi_2}{\partial T_1}$$
$$- \frac{\partial\varphi_1}{\partial x}\frac{\partial\xi_2}{\partial x} - \frac{\partial\varphi_2}{\partial x}\frac{\partial\xi_1}{\partial x} - \xi_1\frac{\partial^2\varphi_1}{\partial x\partial z}\frac{\partial\xi_1}{\partial x};$$

$$\Omega_{32} \equiv \frac{3\gamma}{2}\left(\frac{\partial\xi_1}{\partial x}\right)^2\frac{\partial^2\xi_1}{\partial x^2} + \rho\frac{\partial\varphi_1}{\partial T_2} + \rho\frac{\partial\varphi_2}{\partial T_1} + \rho\xi_1\frac{\partial^2\varphi_1}{\partial z\partial T_1}$$
$$+ \rho\xi_2\frac{\partial^2\varphi_1}{\partial x\partial T_0} + \rho\xi_1\frac{\partial^2\varphi_2}{\partial z\partial T_0} + \frac{\rho}{2}\xi_1^2\frac{\partial^3\varphi_1}{\partial x^2\partial T_0}$$
$$+ \rho\frac{\partial\varphi_1}{\partial z}\frac{\partial\varphi_2}{\partial z} + \rho\xi_1\frac{\partial\varphi_1}{\partial z}\frac{\partial^2\varphi_1}{\partial z^2} + \rho\frac{\partial\varphi_1}{\partial x}\frac{\partial\varphi_2}{\partial x}$$
$$+ \rho\xi_1\frac{\partial\varphi_1}{\partial x}\frac{\partial^2\varphi_1}{\partial x\partial z} - \frac{1}{4\pi}\frac{\partial\Phi_1}{\partial x}\frac{\partial\Phi_2}{\partial z} + \frac{E_0}{4\pi}\xi_2\frac{\partial^2\Phi_1}{\partial z^2}$$
$$- \frac{1}{4\pi}\xi_1\frac{\partial\Phi_1}{\partial z}\frac{\partial^2\Phi_1}{\partial z^2} + \frac{E_0}{4\pi}\xi_1\frac{\partial^2\Phi_2}{\partial z^2} + \frac{E_0}{8\pi}\xi_1^2\frac{\partial^3\Phi_1}{\partial z^3}$$
$$- \frac{1}{4\pi}\frac{\partial\Phi_1}{\partial x}\frac{\partial\Phi_2}{\partial x} - \frac{1}{4\pi}\xi_1\frac{\partial\Phi_1}{\partial x}\frac{\partial^2\Phi_1}{\partial x\partial z};$$

$$\Omega_{33} \equiv -\xi_2\frac{\partial\Phi_1}{\partial z} - \xi_1\frac{\partial\Phi_2}{\partial z} - \frac{1}{2}\xi_1^2\frac{\partial^2\Phi_1}{\partial z^2}.$$

Substituting the first- and second-order solutions into these expressions yields

$$\Omega_{31} \equiv -\left(\frac{1}{2}\frac{\partial\zeta}{\partial T_2} + \Psi_{31}\zeta^2\bar{\zeta}\right)\exp(i\theta) - \left(\frac{1}{2}\frac{\partial\bar{\zeta}}{\partial T_2} - \Psi_{31}\zeta\bar{\zeta}^2\right)$$
$$\times \exp(-i\theta) + \Psi_{33}\zeta^3\exp(3i\theta) + \Psi_{33}\bar{\zeta}^3\exp(-3i\theta);$$

$$\Psi_{31} = \frac{ik^2\omega(\pi(8k^2\gamma + 5g\rho) - 3E_0^2k)}{16\pi(\rho g - 2\gamma k^2)};$$

$$\Psi_{32} = \frac{k(2\pi(k^2\gamma + g\rho) - E_0^2k)}{4\pi(\rho g - 2\gamma k^2)};$$

$$\Psi_{33} = \frac{9ik^2\omega(\pi(4k^2\gamma + g\rho) - E_0^2k)}{16\pi(\rho g - 2\gamma k^2)};$$

$$\Omega_{32} \equiv -\mathrm{K}_{33}(\zeta^3\exp(3i\theta) + \bar{\zeta}^3\exp(-3i\theta))$$
$$-\left(\frac{i\rho\omega}{2k}\frac{\partial\zeta}{\partial T_2} + \mathrm{K}_{31}\zeta^2\bar{\zeta}\right)\exp(i\theta)$$
$$+\left(\frac{i\rho\omega}{2k}\frac{\partial\bar{\zeta}}{\partial T_2} - \mathrm{K}_{31}\zeta\bar{\zeta}^2\right)\exp(-i\theta);$$

$$\mathrm{K}_{31} = \frac{k^2}{32\pi^2(\rho g - 2\gamma k^2)}(E_0^4k^2 - 4\pi E_0^2\gamma k^3 + 12\pi^2\gamma^2k^4$$
$$- 4\pi E_0^2kg\rho + 24\pi^2\gamma k^2g\rho - 6\pi^2g^2\rho^2);$$

$$\mathrm{K}_{33} = \frac{k^2}{32\pi^2(\rho g - 2\gamma k^2)}(5E_0^4k^2 - 20\pi E_0^2\gamma k^3 + 60\pi^2\gamma^2k^4$$
$$- 20\pi E_0^2kg\rho + 48\pi^2\gamma k^2g\rho + 6\pi^2g^2\rho^2);$$

$$\Omega_{33} \equiv \mathrm{M}_{31}\zeta^2\bar{\zeta}\exp(i\theta) + \mathrm{M}_{31}\zeta\bar{\zeta}^2\exp(-i\theta)$$
$$+ \mathrm{M}_{33}\zeta^3\exp(3i\theta) + \mathrm{M}_{33}\bar{\zeta}^3\exp(-3i\theta);$$

$$\mathrm{M}_{31} = \frac{E_0k^2(\pi(4k^2\gamma + 7g\rho) - 3E_0^2k)}{16\pi(\rho g - 2\gamma k^2)};$$

$$\mathrm{M}_{33} = \frac{3E_0k^2(3\pi g\rho - E_0^2k)}{16\pi(\rho g - 2\gamma k^2)}.$$

APPENDIX B

In the fourth-order problem, the inhomogeneity functions on the rights of the boundary conditions on the free surface $z = 0$ of a fluid are given by

$$\Xi_{41} = \xi_3\frac{\partial^2\varphi_1}{\partial z^2} + \xi_2\frac{\partial^2\varphi_2}{\partial z^2} + \xi_1\frac{\partial^2\varphi_3}{\partial z^2} + \xi_1\xi_2\frac{\partial^3\varphi_1}{\partial z^3}$$
$$+ \frac{1}{2}\xi_1^2\frac{\partial^3\varphi_2}{\partial z^3} + \frac{1}{6}\xi_1^3\frac{\partial^4\varphi_1}{\partial z^4} - \frac{\partial\xi_1}{\partial T_3} - \frac{\partial\xi_2}{\partial T_2} - \frac{\partial\xi_3}{\partial T_1}$$

$$-\frac{\partial\xi_3}{\partial x}\frac{\partial\varphi_1}{\partial x}-\frac{\partial\xi_2}{\partial x}\frac{\partial\varphi_2}{\partial x}-\frac{\partial\xi_1}{\partial x}\frac{\partial\varphi_3}{\partial x}-\xi_2\frac{\partial\xi_1}{\partial x}\frac{\partial^2\varphi_1}{\partial x\partial z}$$
$$-\xi_1\frac{\partial\xi_2}{\partial x}\frac{\partial^2\varphi_1}{\partial x\partial z}-\xi_1\frac{\partial\xi_1}{\partial x}\frac{\partial^2\varphi_2}{\partial x\partial z}-\frac{1}{2}\xi_1^2\frac{\partial\xi_1}{\partial x}\frac{\partial^3\varphi_1}{\partial x\partial z^2};$$

$$\Xi_{42} = 3\gamma\frac{\partial\xi_1}{\partial x}\frac{\partial\xi_2}{\partial x}\frac{\partial^2\xi_1}{\partial x^2}+\frac{3\gamma}{2}\left(\frac{\partial\xi_1}{\partial x}\right)^2\frac{\partial^2\xi_2}{\partial x^2}+\rho\frac{\partial\varphi_1}{\partial T_3}+\rho\frac{\partial\varphi_2}{\partial T_2}$$
$$+\rho\frac{\partial\varphi_3}{\partial T_1}+\rho\xi_1\frac{\partial^2\varphi_1}{\partial z\partial T_2}+\rho\xi_2\frac{\partial^2\varphi_1}{\partial z\partial T_1}+\rho\xi_1\frac{\partial^2\varphi_2}{\partial z\partial T_1}$$
$$+\rho\xi_3\frac{\partial^2\varphi_1}{\partial z\partial T_0}+\rho\xi_2\frac{\partial^2\varphi_2}{\partial z\partial T_0}+\rho\xi_1\frac{\partial^2\varphi_3}{\partial z\partial T_0}$$
$$+\frac{\rho}{2}\xi_1^2\left(\frac{\partial^3\varphi_1}{\partial T_1\partial z^2}\right)^2+\rho\xi_1\xi_2\frac{\partial^3\varphi_1}{\partial T_0\partial z^2}+\frac{\rho}{2}\xi_1^2\frac{\partial^3\varphi_2}{\partial T_0\partial z^2}$$
$$+\frac{\rho}{6}\xi_1^3\frac{\partial^4\varphi_1}{\partial T_0\partial z^3}+\frac{\rho}{2}\left(\frac{\partial\varphi_2}{\partial z}\right)^2+\rho\frac{\partial\varphi_1}{\partial z}\frac{\partial\varphi_3}{\partial z}$$
$$+\rho\xi_2\frac{\partial\varphi_1}{\partial z}\frac{\partial^2\varphi_1}{\partial z^2}+\rho\xi_1\frac{\partial\varphi_2}{\partial z}\frac{\partial^2\varphi_1}{\partial z^2}+\frac{\rho}{2}\xi_1^2\left(\frac{\partial^2\varphi_1}{\partial z^2}\right)^2$$
$$+\rho\xi_1\frac{\partial\varphi_1}{\partial z}\frac{\partial^2\varphi_2}{\partial z^2}+\frac{\rho}{2}\xi_1^2\frac{\partial\varphi_1}{\partial z}\frac{\partial^3\varphi_1}{\partial z^3}+\frac{\rho}{2}\left(\frac{\partial\varphi_2}{\partial x}\right)^2$$
$$+\rho\frac{\partial\varphi_1}{\partial x}\frac{\partial\varphi_3}{\partial x}+\rho\xi_2\frac{\partial\varphi_1}{\partial x}\frac{\partial^2\varphi_1}{\partial x\partial z}+\rho\xi_1\frac{\partial\varphi_2}{\partial x}\frac{\partial^2\varphi_1}{\partial x\partial z}$$
$$+\frac{\rho}{2}\xi_1^2\left(\frac{\partial^2\varphi_1}{\partial x\partial z}\right)^2+\rho\xi_1\frac{\partial\varphi_1}{\partial x}\frac{\partial^2\varphi_2}{\partial x\partial z}+\frac{\rho}{2}\xi_1^2\frac{\partial\varphi_1}{\partial x}\frac{\partial^3\varphi_1}{\partial x\partial z^2}$$
$$-\frac{1}{8\pi}\left(\frac{\partial\Phi_2}{\partial z}\right)^2-\frac{1}{4\pi}\frac{\partial\Phi_1}{\partial z}\frac{\partial\Phi_3}{\partial z}+\frac{E_0}{4\pi}\xi_3\frac{\partial^2\Phi_1}{\partial z^2}$$
$$-\frac{1}{4\pi}\xi_2\frac{\partial\Phi_1}{\partial z}\frac{\partial^2\Phi_1}{\partial z^2}-\frac{1}{4\pi}\xi_1\frac{\partial\Phi_2}{\partial z}\frac{\partial^2\Phi_1}{\partial z^2}$$
$$-\frac{1}{8\pi}\xi_1^2\left(\frac{\partial^2\Phi_1}{\partial z^2}\right)^2+\frac{E_0}{4\pi}\xi_2\frac{\partial^2\Phi_2}{\partial z^2}-\frac{1}{4\pi}\xi_1\frac{\partial\Phi_1}{\partial z}\frac{\partial^2\Phi_2}{\partial z^2}$$
$$+\frac{E_0}{4\pi}\xi_1\frac{\partial^2\Phi_3}{\partial z^2}+\frac{E_0}{4\pi}\xi_1\xi_2\frac{\partial^3\Phi_1}{\partial z^3}-\frac{1}{8\pi}\xi_1^2\frac{\partial\Phi_1}{\partial z}\frac{\partial^3\Phi_1}{\partial z^3}$$
$$+\frac{E_0}{8\pi}\xi_1^2\frac{\partial^3\Phi_2}{\partial z^3}+\frac{E_0}{24\pi}\xi_1^3\frac{\partial^4\Phi_1}{\partial z^4}-\frac{1}{8\pi}\left(\frac{\partial\Phi_2}{\partial x}\right)^2$$
$$-\frac{1}{4\pi}\frac{\partial\Phi_1}{\partial x}\frac{\partial\Phi_3}{\partial x}-\frac{1}{4\pi}\xi_2\frac{\partial\Phi_1}{\partial x}\frac{\partial^2\Phi_1}{\partial x\partial z}$$
$$-\frac{1}{4\pi}\xi_1\frac{\partial\Phi_2}{\partial x}\frac{\partial^2\Phi_2}{\partial x\partial z}-\frac{1}{8\pi}\xi_1^2\left(\frac{\partial^2\Phi_1}{\partial x\partial z}\right)^2$$
$$-\frac{1}{4\pi}\xi_1\frac{\partial\Phi_1}{\partial x}\frac{\partial^2\Phi_2}{\partial x\partial z}-\frac{1}{8\pi}\xi_1^2\frac{\partial\Phi_1}{\partial x}\frac{\partial^3\Phi_1}{\partial x\partial z^2};$$

$$\Xi_{43} = -\xi_3\frac{\partial\Phi_1}{\partial z}-\xi_2\frac{\partial\Phi_2}{\partial z}-\xi_1\frac{\partial\Phi_3}{\partial z}-\xi_1\xi_3\frac{\partial^2\Phi_1}{\partial z^2}$$
$$-\frac{1}{2}\xi_1^2\frac{\partial^2\Phi_2}{\partial z^2}-\frac{1}{6}\xi_1^3\frac{\partial^3\Phi_1}{\partial z^3}.$$

Substituting the first- and second-order solutions into these expressions yields

$$\Xi_{41} = A_{410}-\frac{1}{2}\frac{\partial\zeta}{\partial T_3}\exp(i\theta)-\frac{1}{2}\frac{\partial\bar{\zeta}}{\partial T_3}\exp(-i\theta)$$
$$+A_{412}\exp(2i\theta)+\overline{A_{412}}\exp(-2i\theta)+A_{413}\exp(3i\theta)$$
$$+\overline{A_{413}}\exp(-3i\theta)+A_{414}\exp(4i\theta)+\overline{A_{414}}\exp(-4i\theta);$$

$$\Xi_{42} = A_{420}-\frac{i\rho\omega}{2k}\frac{\partial\zeta}{\partial T_3}\exp(i\theta)-\frac{i\rho\omega}{2k}\frac{\partial\bar{\zeta}}{\partial T_3}\exp(-i\theta)$$
$$+A_{422}\exp(2i\theta)+\overline{A_{422}}\exp(-2i\theta)+A_{423}\exp(3i\theta)$$
$$+\overline{A_{423}}\exp(-3i\theta)+A_{424}\exp(4i\theta)+\overline{A_{424}}\exp(-4i\theta);$$

$$\Xi_{43} = A_{430}+A_{432}\exp(2i\theta)+\overline{A_{432}}\exp(-2i\theta)$$
$$+A_{433}\exp(3i\theta)+\overline{A_{433}}\exp(-3i\theta)$$
$$+A_{434}\exp(2i\theta)+\overline{A_{434}}\exp(-4i\theta),$$

where A_{410}, A_{412}, A_{413}, A_{414}, A_{420}, A_{422}, A_{423}, A_{424}, A_{430}, A_{432}, A_{433}, and A_{434} are functions of the time scales T_1, T_2, and T_3.

ACKNOWLEDGMENTS

This work was supported by the Russian Foundation for Basic Research (grant no. 03-01-00760) and by the President of the Russian Federation (grant no. MK-929.2003.01).

REFERENCES

1. M. D. Grabovich, Usp. Fiz. Nauk **140**, 137 (1983) [Sov. Phys. Usp. **26**, 447 (1983)].
2. A. I. Grigor'ev and S. O. Shiryaeva, Izv. Ross. Akad. Nauk, Mekh. Zhidk. Gaza, No. 3, 3 (1994).
3. G. I. Taylor and A. D. McEwan, J. Fluid Mech. **22**, 1 (1965).
4. A. L. Pregenzer and B. M. Marder, J. Appl. Phys. **60**, 3821 (1986).

5. A. H. Schooley, J. Geophys. Res. **65**, 4075 (1960).
6. J. H. Michell, Philos. Mag., Ser. 5 **36**, 430 (1893).
7. J. R. Wilton, Philos. Mag., Ser. 6 **29**, 688 (1915).
8. A. H. Nayfeh, Phys. Fluids **13**, 545 (1970).
9. A. H. Nayfeh and S. D. Hassan, J. Fluid Mech. **48**, 463 (1971).
10. I. Shugan and K. Voliak, J. Fluid Mech. **368**, 321 (1998).
11. A. I. Zhakin, Izv. Akad. Nauk SSSR, Mekh. Zhidk. Gaza, No. 3, 94 (1984).
12. A. Gonzales and A. Castellanos, Phys. Rev. E **49**, 2935 (1994).
13. N. M. Zubarev, Zh. Éksp. Teor. Fiz. **116**, 1990 (1999) [JETP **89**, 1078 (1999)].
14. N. M. Zubarev and O. V. Zubareva, Zh. Tekh. Fiz. **71** (7), 21 (2001) [Tech. Phys. **46**, 806 (2001)].
15. J. J. Stoker, *Water Waves* (Wiley, New York, 1957; Inostrannaya Literatura, Moscow, 1959).
16. G. B. Whitham, *Linear and Nonlinear Waves* (Wiley, New York, 1974; Mir, Moscow, 1977).
17. V. G. Levich, *Physicochemical Hydrodynamics* (Fizmatgiz, Moscow, 1959).
18. Ya. I. Frenkel, Zh. Éksp. Teor. Fiz. **6**, 348 (1936).
19. D. F. Belonozhko and A. I. Grigor'ev, Pis'ma Zh. Tekh. Fiz. **29** (8), 1 (2003) [Tech. Phys. Lett. **29**, 309 (2003)].
20. S. O. Shiryaeva, A. I. Grigor'ev, and V. V. Morozov, Zh. Tekh. Fiz. **73** (7), 21 (2003) [Tech. Phys. **48**, 822 (2003)].
21. D. F. Belonozhko, A. I. Grigor'ev, and Shiryaeva, Pis'ma Zh. Tekh. Fiz. **28** (19), 1 (2002) [Tech. Phys. Lett. **28**, 795 (2002)].
22. D. F. Belonozhko and A. I. Grigor'ev, Zh. Tekh. Fiz. **73** (4), 28 (2003) [Tech. Phys. **48**, 404 (2003)].

Translated by V. Isaakyan

Technical Physics, Vol. 49, No. 1, 2004, pp. 39–46. Translated from Zhurnal Tekhnicheskoĭ Fiziki, Vol. 74, No. 1, 2004, pp. 40–47.

GAS DISCHARGES, PLASMA

Effect of Magnetic Field on the Thermal Nonlinearity of Surface Waves in Plasma–Metal Structures

N. A. Azarenkov, Yu. A. Akimov, and V. P. Olefir

Institute of High Technologies, Karazin National University, pr. Kurchatova 31, Kharkov, 61108 Ukraine

e-mail: olefir@pht.univer.kharkov.ua

Received December 17, 2002; in final form, June 30, 2003

Abstract—A study is made into the effect of the nonlinear mechanism of plasma electron heating on the dispersion properties of potential surface waves propagating along the interface between a metal and finite-pressure magnetoactive plasma. An external steady magnetic field is directed normally to the interface. Different mechanisms of electron energy loss are treated in a weak heating approximation. The energy balance equation is used to determine the spatial distribution of the plasma electron temperature under conditions of nonlocal heating. The effect of the plasma parameters on the nonlinear shift of the wavenumber and on the spatial damping factor of surface waves is investigated. The results obtained are valid for both semiconductor and gaseous plasmas. © *2004 MAIK "Nauka/Interperiodica".*

INTRODUCTION

The properties of surface waves (SWs) in plasma–metal structures are presently subjected to intensive theoretical and experimental investigations. This is due to their numerous applications in plasma and semiconductor electronics, gas discharges, and plasma technologies [1]. The linear theory of SWs in such structures is rather well developed [2–4]. However, the SW behavior may become significantly nonlinear even in the case of fairly low wave amplitudes [5–7].

The scope of investigations of nonlinear effects determining the SW properties in plasma waveguide structures is rather wide. These investigations involve resonant second-harmonic generation, resulting in the transfer of SW energy from the first to second harmonic and vice versa [2]; a nonlinear SW damping caused by the volume second-harmonic generation [2, 8]; SW parametric excitation; and the interaction with low-frequency perturbations, which leads to SW instability [2]. We also mention the studies into the nonlinear SW interaction that are devoted to different mechanisms of SW self-interaction. For example, the SW self-interaction due to the nonlinearity of the set of quasi-hydrodynamic equations was treated in [2, 9], the ionization nonlinearity was treated in [10, 11], and the thermal nonlinearity was treated in [2, 12]. This interest in the SW self-interaction is primarily associated with the fact that these processes perturb the plasma parameters and lead to the dependence of the SW phase velocity on the SW amplitude. Note that the latter factor is of decisive importance in the case of SW excitation. In particular, in the case of SW parametric excitation [2, 13], the SW self-interaction results in a violation of the condition of spatiotemporal synchronism between the SW and the pump field; this, in turn, leads to SW saturation. In the case of SW excitation by charged-particle beams, the dependence of the phase velocity of the excited waves determines the efficiency of their interaction with the beam particles [8, 14, 15].

The objective of this study is to investigate the thermal-nonlinearity-induced SW self-interaction on the plasma–metal interface in the presence of a normal magnetic field. This configuration of the magnetic field is characteristic of RF and microwave discharges, magnetrons, Penning sources, magnetic-discharge pumps, and Hall devices, as well as of plasma processing of metal surfaces [4, 9, 14–16]. We note that the thermal nonlinearity on the plasma–metal interface is of interest from the standpoint of solving problems associated with controlled fusion devices. Since the SW energy is localized in the vicinity of the plasma boundary, the existence of wave perturbations of this type may cause undesirable heating of the plasma periphery and, as a result, lead to a stronger interaction of plasma particles with the structural material of the devices. Especially undesirable is the increase in the energy of plasma particles in the divertor region of fusion devices, because this may increase the flows of charged particles toward the device wall.

FORMULATION OF THE PROBLEM

We will treat the nonlinear self-interaction of high-frequency surface waves as a result of electron heating in the field of a finite-amplitude wave. We will assume that the wave propagates along the interface between a metal and a finite-pressure plasma across an external magnetic field directed normally to the interface. A warm magnetoactive plasma takes up the half-space $x > 0$ (Fig. 1) and, in the $x = 0$ plane, is bounded by a per-

1063-7842/04/4901-0039$26.00 © 2004 MAIK "Nauka/Interperiodica"

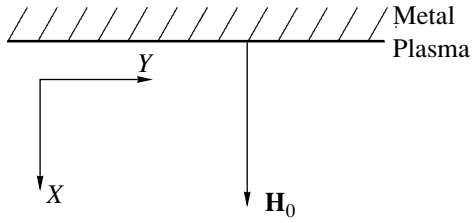

Fig. 1. Geometry of the problem.

fectly conducting metal surface. The external magnetic field $\mathbf{H}_0$ is directed along the *x* axis.

It is known that the SW properties in an inhomogeneous plasma substantially depend on the spatial distribution of the plasma density in the boundary layer. In the cases of strongly and weakly inhomogeneous plasma, the SW properties are determined by the integral characteristics of the plasma in the region where the wave field is localized [7]. In these cases, the plasma–metal interface may be considered sharp, assuming that the plasma is homogeneous with a density equal to its mean value in the region of SW localization. This approach demonstrated its efficiency and good agreement with experimental data, in particular, when investigating gas discharges maintained by SWs. In what follows, we will assume the plasma–metal interface to be sharp and the plasma to be homogeneous.

We assume the effective frequency of electron collisions with scattering centers $\nu = \nu_{col} + \nu_* + \nu_i$ (where ν_{col}, ν_*, and ν_i are the frequencies of elastic collisions, excitation, and ionization, respectively) to be much lower than the wave frequency ω. In the case of a gaseous plasma, ions and atoms of the working gas or impurities may serve as scattering centers. In the case of a semiconductor plasma, such centers may also be provided by optical and acoustic phonons [17–19].

It is known [2] that the thermal mechanism of SW self-interaction consists in that the plasma electrons receive additional energy from the electric field of the wave and then lose this energy in collisions with the scattering centers. This causes a change in the spatial distribution of the electron temperature, which defines the collision frequency and pressure of the plasma electrons. As a result, the electrodynamic properties of the plasma change, which, in turn, leads to the dependence of the parameters of an SW on its amplitude.

Note that the thermal mechanism of self-interaction is closely associated with the ionization nonlinearity [2, 10, 11]. For example, an increase in the amplitude of a high-frequency wave causes a variation in the spatial distribution of the electron temperature and, because of the temperature dependence of the coefficients of elementary processes in plasma, leads to a variation in the plasma density profile. As a result, the SW parameters change. In the case of weak nonlinearity, the SW amplitude is low and, accordingly, the perturbations of the plasma parameters (electron temperature, pressure, collision frequency, and so on) are much smaller than their unperturbed values. In this case, the effect of the thermal and ionization nonlinearities on the dispersion characteristics of SWs may be taken into account additively [2]. This enables one to study the effect of these self-interaction mechanisms independently.

LINEAR THEORY

The dispersion properties and spatial distribution of the electric field potential of high-frequency SWs propagating in a plasma–metal structure were previously studied in a linear (with respect to the field amplitude) approximation in [4]. It was demonstrated in [4] that, in the case of a collisionless gaseous plasma, the SWs under study exist in the $\omega^2 > \omega_{ce}^2$ frequency range (where ω_{ce} is the electron cyclotron frequency) and the necessary condition of their existence is the finiteness of the thermal velocity of plasma electrons $V_{Te} = \sqrt{2T/m_e}$, where T is the plasma electron temperature. In the general case of a collisional semiconductor plasma, the equation for the SW potential Ψ may be written in the form

$$\frac{\partial^4\Psi}{\partial x^4} + \frac{\partial^2\Psi}{\partial x^2}\left[\frac{\omega\omega'}{V_{Te}^2}(1-\alpha) - k_2^2(1+\beta)\right] + k_2^2\Psi\left[k_2^2\beta - \frac{\omega\omega'}{V_{Te}^2}(1-\alpha\beta)\right] = 0, \quad (1)$$

where $\alpha = \omega_{pe}^2/(\varepsilon_0\omega\omega')$, $\beta = \omega'^2/(\omega'^2 - \omega_{ce}^2)$, ε_0 is the semiconductor lattice permittivity (in the case of a gaseous plasma, $\varepsilon_0 = 1$), ω_{pe} is the electron plasma frequency, k_2 is the complex wavenumber of the SW, and $\omega' = \omega + i\nu$.

Assuming that the spatial distribution of the SW potential has the form [4]

$$\Psi(x) = A_1\exp(-\lambda_1 x) + A_2\exp(-\lambda_2 x), \quad (2)$$

where A_1 and A_2 are constants, one can derive expressions for the quantities $\lambda_{1,2}$, characterizing the SW penetration into the plasma:

$$\lambda_{1,2}^2 = \frac{1}{2}\frac{\omega\omega'}{V_{Te}^2}\left\{\frac{k_2^2V_{Te}^2}{\omega\omega'}(1+\beta) - (1-\alpha) \pm \sqrt{(1-\alpha)^2 + \frac{k_2^2V_{Te}^2}{\omega\omega'}(1-\beta)\left[\frac{k_2^2V_{Te}^2}{\omega\omega'}(1-\beta) + 2(1+\alpha)\right]}\right\}. \quad (3)$$

Using the condition that both the electric potential and the normal component of hydrodynamic velocity of plasma electrons vanish on the plasma–metal inter-

face, we derive the following dispersion relation:

$$1 + k_2^2 r_{de}^2 - r_{de}^2(\lambda_1^2 + \lambda_1\lambda_2 + \lambda_2^2) = 0,$$
$$r_{de} = \sqrt{\varepsilon_0} V_{Te}/\omega_{pe}, \tag{4}$$

from which, for the wavenumber k_2, we obtain

$$k_2^2 = \frac{1}{2V_{te}^2}\frac{\omega}{\omega + i\nu}\left(\frac{\omega + i\nu}{\omega_{ce}^2} - 1\right)$$
$$\times\left\{(\omega + i\nu)^2 + \omega_{ce}^2 + \frac{\omega + i\nu}{\omega}\frac{\omega_{pe}^2}{\varepsilon_0}\right. \tag{5}$$
$$\left. - \sqrt{\left((\omega + i\nu)^2 - \omega_{ce}^2 + \frac{\omega + i\nu}{\omega}\frac{\omega_{pe}^2}{\varepsilon_0}\right)^2 + 4\frac{\omega + i\nu}{\omega}\frac{\omega_{ce}^2\omega_{pe}^2}{\varepsilon_0}}\right\}.$$

Therefore, taking into account plasma electron collisions results in both the damping of SWs and the expansion of the range of their existence. It can be seen that SWs may also exist in the range of frequencies below the electron cyclotron frequency; however, in this case, they are strongly damped ($\mathrm{Im}k_2 > \mathrm{Re}k_2$).

We note that, when the thermal motion of plasma electrons is taken into account, the expressions for the wave potential and the wavenumber are rather cumbersome even in a linear approximation with respect to the wave field amplitude. Therefore, further investigation of the SW self-interaction will be performed for a fairly dense plasma for which the condition $\omega_{ce}^2 < \omega^2 \ll \omega_{pe}^2/\varepsilon_0$ is satisfied. In this case, expressions (3) and (5) are simplified to

$$k_2 = k_2' + ik_2''$$
$$= \frac{\omega}{V_{Te}}\sqrt{\varepsilon_0\frac{\omega^2 - \omega_{ce}^2}{\omega_{pe}^2}}\left(1 + i\frac{\nu}{\omega}\frac{\omega^2}{\omega^2 - \omega_{ce}^2}\right), \tag{6}$$

$$\lambda_1 = \lambda_1' + i\lambda_1'' = \frac{\omega_{pe}}{V_{Te}\sqrt{\varepsilon_0}}\left(1 - \frac{1}{2}\varepsilon_0\frac{\omega^2}{\omega_{pe}^2}\left(1 + i\frac{\nu}{\omega}\right)\right),$$
$$\lambda_2 = \lambda_2' + i\lambda_2'' = \frac{\omega}{V_{Te}}\sqrt{\varepsilon_0}\frac{\omega}{\omega_{pe}}\left(1 + i\frac{\nu}{\omega}\right). \tag{7}$$

Analysis of expression (6) shows that the SW phase velocity significantly exceeds the thermal velocity of electrons, which agrees with the applicability range of the hydrodynamic description of the plasma. At the same time, in the case of a dense plasma, the condition of wave potentiality imposes the following restriction on the thermal velocity of plasma electrons: $V_{Te} \ll c(\omega^2 - \omega_{ce}^2)^{1/2}/\omega_{pe}$, where c is the speed of light in vacuum.

SPATIAL DISTRIBUTION OF THE ELECTRON TEMPERATURE

Let us consider weak thermal nonlinearity, assuming that the change in the electron temperature δT due to electron heating in the SW field is much less than its equilibrium value T_0: $T = T_0 + \delta T$, $\delta T \ll T_0$. We also assume that the frequency variation $\delta\nu = \delta\nu_{\mathrm{col}} + \delta\nu_* + \delta\nu_i$ is small compared to its unperturbed value ν in the absence of SWs. In this case, the expressions for the electron collision frequencies in the vicinity of the equilibrium temperature value may be written as

$$\nu_{\mathrm{col}}(T) = \nu_{\mathrm{col}}(T_0) + \delta\nu_{\mathrm{col}},$$
$$\delta\nu_{\mathrm{col}} = \delta T\left.\frac{\partial\nu_{\mathrm{col}}}{\partial T}\right|_{T_0} \ll \nu_{\mathrm{col}}(T_0),$$
$$\nu_*(T) = \nu_*(T_0) + \delta\nu_*,$$
$$\delta\nu_* = \delta T\left.\frac{\partial\nu_*}{\partial T}\right|_{T_0} \ll \nu_*(T_0), \tag{8}$$
$$\nu_i(T) = \nu_i(T_0) + \delta\nu_i,$$
$$\delta\nu_i = \delta T\left.\frac{\partial\nu_i}{\partial T}\right|_{T_0} \ll \nu_i(T_0),$$

where the excitation and ionization frequencies are defined by the following expressions [10, 20, 21]:

$$\nu_* = \nu_*^0\exp(-U_*/T),$$
$$\nu_{i1} = \nu_{i1}^0\exp(-U_*/T), \quad T > 2/3(U_i - U_*), \tag{9}$$
$$\nu_{i2} = \nu_{i2}^0\exp(-U_i/T), \quad T < 2/3(U_i - U_*),$$

where U_* and U_i are the excitation energy of the first level and the ionization energy of the working gas atoms, respectively. If the wave frequency ω is much higher than the characteristic frequency of energy transfer $\tilde{\nu}$ in collisions of plasma electrons with scattering centers, the process of energy exchange may be considered to be quasi-steady [22]. In this case, the electron temperature perturbation depends on the coordinates and the squared SW amplitude modulus $\delta T = \delta T(x, y, |A_1|^2)$ and may be determined using the wave-period-averaged equation of energy balance [22],

$$1/3\mathrm{Re}(\mathbf{j}\cdot\mathbf{E}^*) = \nabla\cdot\mathbf{Q} - P(T), \tag{10}$$

where $\mathbf{Q}$ is the vector of the heat flux transferred by electrons, $\mathbf{j}$ is the density of the high-frequency electron current, and $\mathbf{E}^*$ is the complex conjugate electric field of the wave.

The term $P(T) = -n_0\tilde{\nu}(T_0)(T - T_0)$ is the specific energy transferred by the electrons to the scattering centers with the characteristic frequency

$$\tilde{\nu}(T_0) = \gamma\nu_{\mathrm{col}}(T_0) + U_*\left.\frac{\partial\nu_*}{\partial T}\right|_{T_0} + U_i\left.\frac{\partial\nu_i}{\partial T}\right|_{T_0}, \tag{11}$$

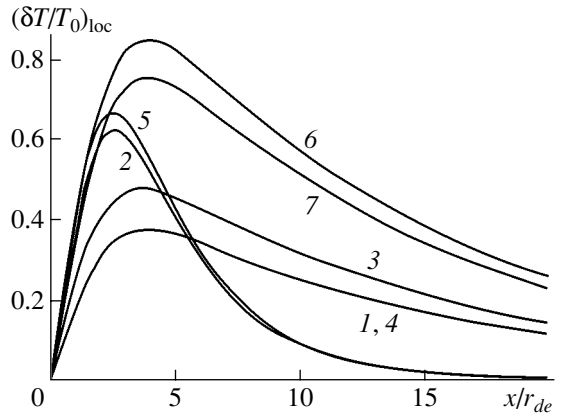

Fig. 2. Spatial distribution of the electron temperature in the local heating approximation. The values of the parameters $\sqrt{\varepsilon_0}\,\omega_{ce}/\omega_{pe}$, $\sqrt{\varepsilon_0}\,\omega/\omega_{pe}$, ν/ω, $\nu/\tilde{\nu}$, and μ are as follows: (*1*) 0.05, 0.2, 0.1, 10^3, 0.1; (*2*) 0.05, 0.4, 0.05, 10^3, 0.1; (*3*) 0.1, 0.2, 0.1, 10^3, 0.1; (*4*) 0.05, 0.2, 0.2, 10^3, 0.1; (*5*) 0.1, 0.4, 0.1, 10^3, 0.1; (*6*) 0.05, 0.2, 0.1, 10^3, 0.15; and (*7*) 0.05, 0.2, 0.2, 2×10^3, 0.1.

where n_0 is the unperturbed plasma density and the $\gamma = 2m_eM/(m_e + M)^2$ is the fraction of energy lost by the electrons in elastic collisions with scattering centers of mass M.

Note that, in the general case, the characteristic frequency $\tilde{\nu}$ is determined by both the elastic collision frequencies and the atom ionization and excitation frequencies.

The components of the heat flux vector $\mathbf{Q}$ in energy balance equation (10) are defined by the expression $\mathbf{Q}_i = -\chi_{ij}\partial T/\partial\xi_j$, where χ_{ij} is the tensor of the electron thermal conductivity of the plasma and $\boldsymbol{\xi} = (x, y)$. The left-hand side of Eq. (10) describes the heating of plasma electrons in the SW field. The terms on the right-hand side describe the electron energy losses per unit volume due to the finite thermal conductivity of the plasma and the energy transfer to the scattering centers.

The energy balance equation may be simplified if one assumes that the heat transfer largely occurs along the magnetic field, $\chi = \chi_{xx} \gg \chi_{xy}, \chi_{yx}, \chi_{yy}$. This condition is valid for collision frequencies much lower than the electron cyclotron frequency ($\nu \ll \omega_{ce}$) [22]. In view of this, Eq. (10) takes the form

$$-\frac{1}{\lambda_T^2}\frac{\partial^2}{\partial x^2}\frac{\delta T}{T_0} + \frac{\delta T}{T_0} = \left(\frac{\delta T}{T_0}\right)_{\rm loc}, \qquad (12)$$

where $\lambda_T^{-1} = 1/\sqrt{3m_e\nu\tilde{\nu}/(5T_0)}$ is the characteristic scale length of the electron thermal conductivity and the value

$$(\delta T/T_0)_{\rm loc} = -e{\rm Re}(\mathbf{V}_e\mathbf{E}^*)/(3\tilde{\nu}T_0) \qquad (13)$$

is the relative variation in the electron temperature in the local heating approximation.

Assuming that the wave propagates in the positive direction along the y axis and taking into account expressions (2), (6), and (7), obtained using the linear theory, one can write expression (13) in the form

$$\left(\frac{\delta T}{T_0}\right)_{\rm loc} = \frac{2}{3}\mu^2\frac{\omega}{\tilde{\nu}}\frac{\omega^2}{\omega^2 - \omega_{ce}^2}\exp(-2k_2''|y|)$$

$$\times\left\{2\frac{\nu}{\omega}\varepsilon_0^2\frac{\omega^2}{\omega_{pe}^2}\frac{\omega_{ce}^2}{\omega_{pe}^2}\exp(-2\lambda_1' x) + 2\frac{\nu}{\omega}\varepsilon_0\frac{\omega^2}{\omega_{pe}^2}\exp(-2\lambda_2' x)\right.$$
$$+\left[\frac{\omega^2 - \omega_{ce}^2}{\omega^2}\sin(\lambda_1'' - \lambda_2'')x - 2\frac{\nu}{\omega}\varepsilon_0\frac{\omega^2}{\omega_{pe}^2}\left(1 + \varepsilon_0\frac{\omega_{ce}^2}{\omega_{pe}^2}\right)\right. \qquad (14)$$
$$\left.\left.\times\cos(\lambda_1'' - \lambda_2'')x\right]\exp[-(\lambda_1' + \lambda_2')x]\right\},$$

where the dimensionless parameter $\mu = e|A_1|/(m_eV_{Te}^2)$ is the ratio of the electron energy in the wave field to the thermal electron energy.

The quantity $(\delta T/T_0)_{\rm loc}$ (Fig. 2) determines the spatial distribution of the electron temperature perturbations in the local heating approximation $|\nabla\cdot\mathbf{Q}| \ll |P(T)|$. Note that this approximation was used, e.g., in [2, 11, 17, 18, 22–26]. In our case, however, this approximation is invalid in view of the smallness of the frequencies ν and $\tilde{\nu}$. It can be shown that, in the problem under study, the condition of local heating may be reduced to the form $\omega_{pe}^2/(\nu\tilde{\nu}) \ll 1$. This implies that we are dealing with a situation in which electron heating is highly nonlocal [23]. Therefore, expression (14) characterizes only the spatial distribution of the SW power received by the plasma electrons due to their collisions with the scattering centers and fails to describe the spatial distribution of the plasma electron temperature.

In order to determine the spatial distribution of the temperature under conditions of nonlocal electron heating in the SW field, one must use Eq. (12) simultaneously with expression (14). In view of the fact that the thermal conductivity of metal is high compared to that of plasma, the boundary condition consisting in the continuity of the heat flux on the plasma–metal interface leads to a negligibly low heating of metal. In order to determine the plasma temperature, one must use the integral form of energy conservation during the heating of plasma electrons,

$$\int_0^\infty \delta T/T_0\,dx = \int_0^\infty (\delta T/T_0)_{\rm loc}\,dx,$$

which gives the following expression for the relative

temperature variation:

$$\frac{\delta T}{T_0} = \frac{2}{3}\mu^2 \exp(-2k_2''|y|)\{P_1 \exp(-2\lambda_1' x) + P_2 \exp(-2\lambda_2' x) + P_T \exp(-\lambda_T x) + [P_3 \sin(\lambda_1'' - \lambda_2'')x + P_4 \cos(\lambda_1'' - \lambda_2'')x]\exp[-(\lambda_1' + \lambda_2')x]\}. \quad (15)$$

Here, the following notation is used:

$$P_1 = -\frac{3}{5}\frac{\nu^2}{\omega^2}\frac{\varepsilon_0^3\omega^4}{\omega_{pe}^2}\frac{\omega_{ce}^2}{\omega_{pe}^2}\frac{\omega^2}{\omega^2 - \omega_{ce}^2},$$

$$P_2 = -\frac{3}{5}\frac{\nu^2}{\omega^2}\frac{\omega^2}{\omega^2 - \omega_{ce}^2}, \quad P_3 = -\frac{6}{5}\frac{\nu}{\omega}\frac{\varepsilon_0\omega^2}{\omega_{pe}^2}, \quad (16)$$

$$P_4 = -\frac{24}{5}\frac{\nu^2}{\omega^2}\frac{\varepsilon_0^2\omega^4}{\omega_{pe}^4}\frac{\omega^2}{\omega^2 - \omega_{ce}^2}\frac{\omega^2}{\omega^2 - 3\omega_{ce}^2},$$

$$P_T = \sqrt{\frac{6}{5}}\frac{\nu}{\tilde{\nu}}\frac{\nu}{\omega}\frac{\sqrt{\varepsilon_0}\omega}{\omega_{pe}}\frac{\omega^2}{\omega^2 - \omega_{ce}^2}.$$

Note that the relative variation of the electron temperature reaches its maximum value

$$\left(\frac{\delta T}{T_0}\right)_{\max} \approx \mu^2 \sqrt{\frac{8}{15}\varepsilon_0}\frac{\nu}{\tilde{\nu}}\frac{\nu}{\omega_{pe}}\frac{\omega^2}{\omega^2 - \omega_{ce}^2} e^{-2k_2''|y|} \quad (17)$$

at a distance of $x_{\max} \cong r_{de}$ from the plasma–metal interface. Because of the heating of plasma electrons, there is a heat flux deep into the plasma. Along with this, there is also a heat flux toward the grounded metal surface. However, in the case under consideration, the latter flux is negligibly small compared to the main flux deep into the plasma (Fig. 2).

The condition of weak heating, $\delta T \ll T_0$, $|\delta\nu| \ll \nu(T_0)$, leads to the following restriction on the wave field amplitude:

$$\mu^2 \sqrt{\frac{8}{15}\varepsilon_0}\frac{\nu}{\tilde{\nu}}\frac{\nu}{\omega_{pe}}\frac{\omega^2}{\omega^2 - \omega_{ce}^2}\frac{U_{*,i}}{T_0} \ll 1. \quad (18a)$$

At the same time, when solving the problem on SW self-interaction, one must take into account the fact that the results of linear theory are valid in the case

$$\mu\left(\frac{\omega^2}{\omega^2 - \omega_{ce}^2}\right)^{1/2} \ll 1. \quad (18b)$$

Numerical calculations (Fig. 3) revealed that these conditions are valid at field amplitudes for which $\mu \le 0.1$.

As was expected, an increase in the wave amplitude and electron collision frequency leads to an increase in the SW Joule loss and has a significant effect on the heating of plasma electrons. Note also that, as the wave frequency approaches the electron cyclotron frequency,

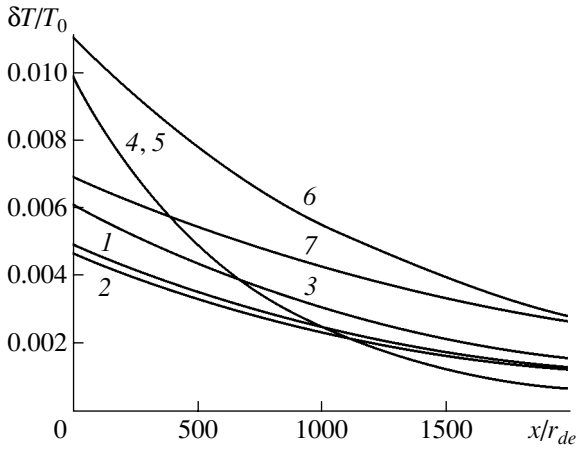

Fig. 3. Spatial distribution of the electron temperature in the vicinity of the perturbation source. Curves *1–7* correspond to the same plasma parameters as those in Fig. 2.

the efficiency of SW energy transfer to the plasma electrons increases (Fig. 3, curve *3*). An increase in the parameter $\nu/\tilde{\nu}$ causes an increase in both the temperature and the characteristic scale length of the electron thermal conductivity ($\lambda_T^{-1} \propto \sqrt{\nu/\tilde{\nu}}$), which leads to a more gradual decrease in the electron temperature deep into the plasma.

NONLINEAR DISPERSION RELATION

A variation in the plasma electron temperature causes a correction to the collision frequency $\delta\nu$ (see expressions (8)). Taking into account this correction and the correction to the plasma electron pressure $\delta p = n_0\delta T$ in the equation of motion for electrons,

$$\frac{\partial \mathbf{V}_e}{\partial t} = \frac{e}{m_e}\nabla\Psi - \frac{e}{m_e c}[\mathbf{V}_e, \mathbf{H}_0] - \frac{\nabla[p(T_0) + \delta p]}{n_0 m_e} - [\nu(T_0) + \delta\nu]\mathbf{V}_e, \quad (19)$$

and solving this equation simultaneously with the continuity and Poisson's equations, one can derive the following equation for the wave potential:

$$r_{de}^2\left\{\frac{\partial^4\Psi}{\partial x^4} + \frac{\partial^2\Psi}{\partial x^2}\left[\frac{\omega\omega'}{V_{Te}^2}(1-\alpha) - k_2^2(1+\beta)\right] + 2k_2^2\Psi\left[k_2^2\beta - \frac{\omega\omega'}{V_{Te}}(1-\alpha\beta)\right]\right\} = R_{\delta\nu} + R_{\delta p}. \quad (20)$$

Equation (20) with zero on the right-hand side is the equation for the SW potential in a linear approximation with respect to the wave field amplitude (see Eq. (1)). The right-hand side of Eq. (20) accounts for the nonlinear effects due to the variations in the collision fre-

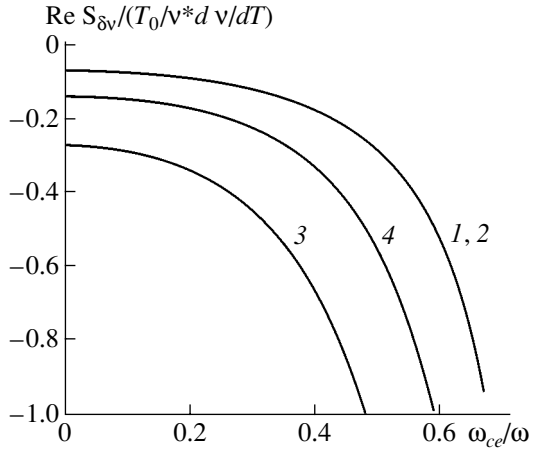

Fig. 4. Correction to the real part of the wavenumber caused by the variation in the electron collision frequency as a function of the magnetic field. The values of the parameters ω/ω_{pe}, ν/ω, and $\nu/\tilde{\nu}$ are as follows: (*1*) 0.1, 0.01, 10^3; (*2*) 0.2, 0.01, 10^3; (*3*) 0.1, 0.02, 10^3; and (*4*) 0.1, 0.01, 2×10^3.

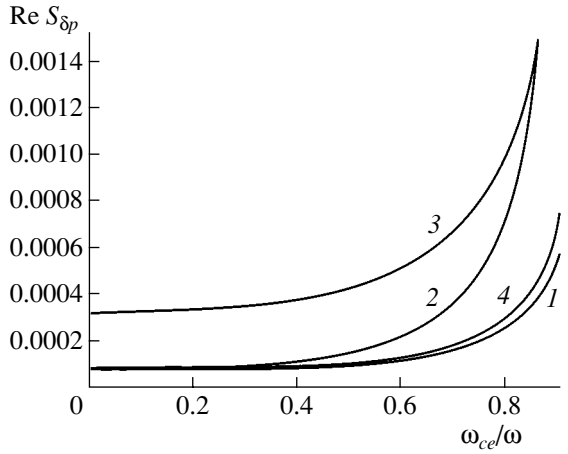

Fig. 5. Correction to the real part of the wavenumber caused by the variation in the electron pressure as a function of the magnetic field. Curves *1–4* correspond to the same plasma parameters as those in Fig. 4.

quency ($R_{\delta\nu}$) and electron pressure ($R_{\delta p}$). We do not present here the expressions for $R_{\delta\nu}$ and $R_{\delta p}$ because of their awkwardness. We will seek a solution to Eq. (20) for the SW potential in the form

$$\Psi = A_1\exp(-\lambda_1 x) + A_2\exp(-\lambda_2 x) + \Psi_{\delta\nu} + \Psi_{\delta p}, \quad (21)$$

where the nonlinear corrections $\Psi_{\delta\nu}$ and $\Psi_{\delta p}$ vary as μ^2.

Applying the boundary conditions for the potential and the normal component of the electron velocity on the interface, we derive the following nonlinear dispersion relation:

$$1 + k_2^2 r_{de}^2 - r_{de}^2(\lambda_1^2 + \lambda_1\lambda_2 + \lambda_2^2) = -\frac{\varepsilon_0\omega^2}{\omega_{pe}^2}\left(1 + i\frac{\nu}{\omega}\right)(S_{\delta\nu} + S_{\delta p}), \quad (22)$$

where

$$S_{\delta p} = \frac{4}{5}\mu^2\frac{\nu^2}{\omega^2}\frac{\omega^2}{\omega^2-\omega_{ce}^2}\left(1 - i\frac{\nu}{\omega}\right) \times\left(1 + i\frac{5}{6}\varepsilon_0^2\frac{\omega}{\tilde{\nu}}\frac{\omega^2}{\omega_{pe}^2}\frac{\omega_{ce}^2}{\omega_{pe}^2}\frac{\omega^2}{\omega^2-\omega_{ce}^2}\right)\exp(-2k_2''|y|),$$

$$S_{\delta\nu} = -i\frac{1}{5}\mu^2\frac{T_0}{\nu}\frac{\partial\nu}{\partial T}\bigg|_{T_0}\frac{\nu^3}{\omega^3}\frac{\omega^2}{\omega^2-\omega_{ce}^2}\frac{\omega_{pe}^2}{\varepsilon_0\omega^2} \times\left\{\frac{(\omega^2-\omega_{ce}^2)^2+\omega_{ce}^4}{2(\omega^2-\omega_{ce}^2)^2} - \frac{10}{3}\frac{\varepsilon_0\omega^2}{\omega_{pe}^2}\frac{\omega^2}{\nu\tilde{\nu}}\frac{\omega_{ce}^2}{\omega^2-\omega_{ce}^2} \times\left(1 + i\frac{\nu}{\omega}\frac{\omega^2}{\omega_{ce}^2}\frac{\omega^2+3\omega_{ce}^2}{\omega^2-\omega_{ce}^2}\right)\right\}\exp(-2k_2''|y|). \quad (23)$$

The complex wavenumber k_2 satisfying this dispersion relation has the form

$$k_2 = \frac{\omega}{V_{Te}}\sqrt{\varepsilon_0\frac{\omega^2-\omega_{ce}^2}{\omega_{pe}^2}\left(1 + i\frac{\nu}{\omega}\frac{\omega^2}{\omega^2-\omega_{ce}^2}\right)} \times(1 + S_{\delta\nu} + S_{\delta p}). \quad (24)$$

In the limiting case of $|A_1| \longrightarrow 0$, nonlinear dispersion relation (22) changes to linear dispersion relation (4) and its solution, given by expression (24), changes to expression (6).

RESULTS AND DISCUSSION

We will analyze the effect of the magnetic field on the phase characteristics of the SW. Numerical calculations reveal that an increase in the external magnetic field leads to an increase in the nonlinear corrections to the real part of the wavenumber, which are associated with the variations in the electron collision frequency (Fig. 4) and electron pressure (Fig. 5). Note that, in the entire range of variations of the magnetic field, the nonlinear correction due to the variations in the collision frequency significantly exceeds the correction due to the electron pressure perturbations.

The correction to the imaginary part of the wavenumber k_2 also increases with increasing external magnetic field and is largely determined by the perturbation of the electron collision frequency (Figs. 6 and 7).

One can see from expression (23) that the effect of the perturbation of the electron collision frequency ν on the SW dispersion is determined by the temperature dependence of the frequency ν. First, we will consider the case in which the electron collision frequency increases with temperature, $\partial\nu/\partial T|_{T_0} > 0$. This takes

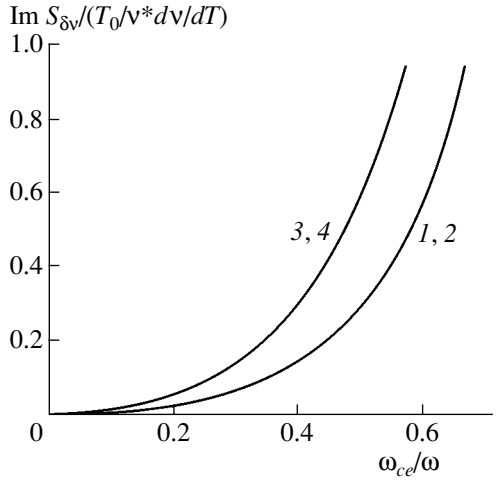

Fig. 6. Correction to the imaginary part of the wavenumber caused by the variation in the electron collision frequency as a function of the magnetic field. Curves *1–4* correspond to the same plasma parameters as those in Fig. 4.

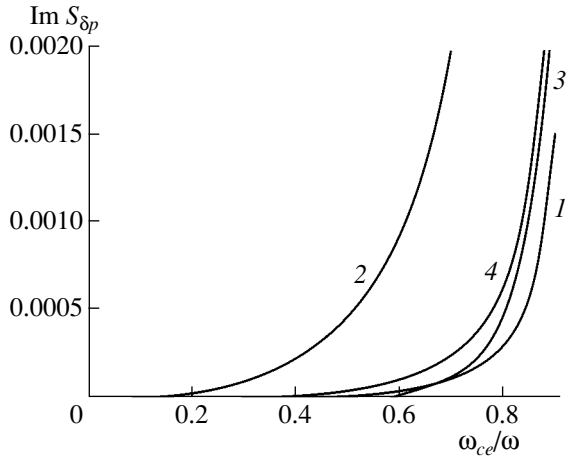

Fig. 7. Correction to the imaginary part of the wavenumber caused by the variation in the electron pressure as a function of the magnetic field. Curves *1–4* correspond to the same plasma parameters as those in Fig. 4.

place in the case of electron scattering by optical or acoustic phonons in semiconductor plasma ($\nu(T) \propto \sqrt{T}$, $T^{3/2}$ [17–19]) or in the case of a low-pressure gas discharge, when the electron collision frequency is determined by inelastic collisions leading to atom excitation [22]. In such a situation, the nonlinear shift of the real part of the wavenumber k_2 is negative ($\mathrm{Re}S_{\delta\nu} < 0$) and nonlinear damping rate (24) is larger than linear damping rate (6) ($\mathrm{Im}S_{\delta\nu} > 0$). In the opposite case, when the electron collision frequency decreases with increasing temperature, $\partial\nu/\partial T|_{T_0} < 0$ (e.g., in the case of elastic electron collisions with ions or impurities of gaseous plasma, when $\nu(T) \propto T^{-3/2}$, $T^{1/2}$ [22]), an opposite dependence takes place. In this case, the shift of the real part of the wavenumber is positive, and the damping rate decreases compared to its linear value.

As was mentioned above, the change in the electron temperature significantly depends on the parameter $\nu/\tilde{\nu}$. In view of this, nonlinear corrections to the complex wavenumber also depend on the mechanism of electron energy losses. The wave damping rate and the nonlinear shift of the wavenumber both increase with the parameter $\nu/\tilde{\nu}$: $\mathrm{Im}S_{\delta\nu}$, $\mathrm{Re}S_{\delta\nu} \propto \nu/\tilde{\nu}$. In the case of semiconductor plasma and high-pressure gas discharges, the main mechanism of energy losses is due to elastic collisions [22] and the ratio of the collision frequency ν to the characteristic frequency of energy transfer $\tilde{\nu}$ is described by

$$\nu/\tilde{\nu} = 0.5M/m_e \gg 1. \qquad (25)$$

As the pressure decreases and inelastic electron collisions leading to atom excitation become dominant [5, 22], this ratio decreases,

$$\nu/\tilde{\nu} = (1 + \nu_{\mathrm{col}}/\nu_*)T_0^2/U_*^2 \ll 0.5M/m_e. \qquad (26)$$

Expressions (25) and (26) allow one to conclude that nonlinear corrections to the wavenumber due to electron heating in the SW field are most significant at high pressures (Fig. 6, curves *1–4*).

CONCLUSIONS

We have theoretically studied the effect of plasma electron heating on the dispersion properties of high-frequency potential SWs propagating along the plasma–metal interface. We have considered the case of a finite-pressure dense plasma in an external steady magnetic field normal to the interface. We have derived and investigated a linear dispersion relation for SWs with allowance for electron thermal motion and electron collisions. It is demonstrated that the electron heating is highly nonlocal and is largely determined by the processes of heat transfer in the plasma. The spatial distribution of the plasma electron temperature has been found in a weak heating approximation. A nonlinear dispersion relation has been investigated. Analytical expressions have been derived for the nonlinear shift of the wavenumber and the spatial damping rate. Numerical analysis has been performed of the effect of the plasma parameters and the magnitude of the external magnetic field on the SW characteristics. The results of the investigation are applicable to both a semiconductor and gaseous plasma bounded by metal.

ACKNOWLEDGMENTS

This study was supported in part by the Science and Technology Center in Ukraine (project no. 1112).

REFERENCES

1. M. Moisan, J. Hurbert, J. Margot, and Z. Zakrzewski, *The Development and Use of Surface-Wave Sustained Discharges for Applications in Advanced Technologies Based on Wave and Beam Generated Plasmas* (Kluwer, Amsterdam, 1999), pp. 1–42.
2. N. A. Azarenkov and K. N. Ostrikov, Phys. Rep. **308**, 333 (1999).
3. N. A. Azarenkov, Zh. Tekh. Fiz. **57**, 1165 (1987) [Sov. Phys. Tech. Phys. **32**, 681 (1987)].
4. N. A. Azarenkov, A. N. Kondratenko, and Yu. O. Tyshetskii, Zh. Tekh. Fiz. **69** (11), 30 (1999) [Tech. Phys. **44**, 1286 (1999)].
5. A. V. Gurevich and A. B. Shvartsburg, *Nonlinear Theory of Radio Wave Propagation in Ionosphere* (Nauka, Moscow, 1973).
6. V. I. Karpman, *Non-Linear Waves in Dispersive Media* (Nauka, Moscow, 1973; Pergamon, Oxford, 1994).
7. J. Weiland and H. Wilhelmsson, *Coherent Nonlinear Interaction of Waves in Plasmas* (Pergamon, Oxford, 1976; Énergoizdat, Moscow, 1981).
8. A. N. Kondratenko, *Plasma Waveguides* (Atomizdat, Moscow, 1976).
9. N. A. Azarenkov, Yu. A. Akimov, and A. V. Gapon, Vestn. Khar'kovsk. Nats. Univ. Ser. Fiz. **496** (4), 29 (2000).
10. Yu. M. Aliev, A. G. Boev, and A. Shivarova, J. Phys. D **17**, 2233 (1984).
11. N. A. Azarenkov, K. N. Ostrikov, and M. Y. Yu, J. Appl. Phys. **84**, 4176 (1998).
12. A. G. Litvak and V. A. Mironov, *Nonlinear Thermal Phenomena in Plasmas* (Inst. Prikl. Fiz. Akad. Nauk SSSR, Gorki, 1979).
13. N. A. Azarenkov, Yu. A. Akimov, and V. P. Olefir, Vestn. Khar'kovsk. Nats. Univ. Ser. Fiz. **574** (4), 62 (2002).
14. V. I. Maslov, Fiz. Plazmy **16**, 394 (1990) [Sov. J. Plasma Phys. **16**, 225 (1990)].
15. N. A. Azarenkov, Yu. A. Akimov, and V. P. Olefir, Vopr. At. Nauki Tekh., Ser. 8: Fiz. Plazmy, No. 5, 92 (2002).
16. D. P. Schmidt, N. B. Meezan, W. A. Hargus, Jr., *et al.*, Plasma Sources Sci. Technol. **9**, 68 (2000).
17. F. G. Bass and Yu. G. Gurevich, *Hot Electrons and High-Power Electromagnetic Waves in Semiconductor and Gas Discharge Plasmas* (Nauka, Moscow, 1975).
18. F. G. Bass and Yu. G. Gurevich, Sov. Phys. J. **14**, 113 (1971).
19. N. N. Beletskiĭ, V. M. Svetlichnyĭ, D. D. Khalameĭda, and V. M. Yakovenko, *Microwave Phenomena in Heterogeneous Semiconductor Structures* (Naukova Dumka, Kiev, 1991).
20. Yu. M. Aliev, K. Ivanova, M. Moisan, *et al.*, Plasma Sources Sci. Technol. **2**, 145 (1993).
21. L. M. Biberman, V. S. Vorob'ev, and I. T. Yakubov, *Kinetics of Non-Equilibrium Low-Temperature Plasma* (Consultans Bureau, New York, 1987).
22. V. E. Golant, A. P. Zhilinskii, and S. A. Sakharov, *Fundamentals of Plasma Physics* (Atomizdat, Moscow, 1977; Wiley, New York, 1980).
23. Yu. M. Aliev, A. V. Maximov, and H. Schluter, Phys. Scr. **48**, 464 (1993).
24. Yu. M. Aliev, A. V. Maximov, I. Ghanashev, *et al.*, IEEE Trans. Plasma Sci. **23**, 409 (1995).
25. Yu. M. Aliev, V. Yu. Bychenkov, A. V. Maximov, *et al.*, Plasma Sources Sci. Technol., No. 1, 126 (1992).
26. Yu. M. Aliev, H. Schluter, and A. Shivarova, Plasma Sources Sci. Technol., No. 5, 514 (1996).

Translated by A. Bronshteĭn

Technical Physics, Vol. 49, No. 1, 2004, pp. 47–51. Translated from Zhurnal Tekhnicheskoĭ Fiziki, Vol. 74, No. 1, 2004, pp. 48–52.

GAS DISCHARGES, PLASMA

Polarization Correction and Effective Field in Plasma

K. P. Piskunov

State Research and Testing Institute of Problems of Technical Protection of Information, State Technical Commission of Russia, Voronezh, 394030 Russia

e-mail: mail@gniiptzi.vsi.ru

Received June 24, 2003

Abstract—The modified Lorentz–Mossotti method is used to determine the polarization correction to the mean macroscopic electric field in plasma as a function of the electron density, the density of the medium, and the electron and ion temperatures. It is demonstrated that, at high electron densities, the polarization correction may play a decisive role in estimating various electrodynamic characteristics of conducting media. © *2004 MAIK "Nauka/Interperiodica".*

INTRODUCTION

The formulation and solution of the problem of estimating the effective field $\mathbf{E}_{ef}$ in plasma have a long history. The beginning of investigations in this field is usually timed to the discovery of the Earth's ionosphere. The importance of solving this problem was determined by the practical need to perform correct calculations of the altitude at which the region with the critical electron density is located in the ionosphere. The first correct estimates of the field acting on an isolated electron in an ionospheric plasma were made in [1–5]. It turned out that, in this case, the acting field is approximately equal to the mean macroscopic field.

The assumption of the equality of the acting field to the mean macroscopic field for the ionospheric conditions is based on the relative smallness of the polarization correction.[1] This result, which was in fact obtained for a rarefied plasma (as was repeatedly stressed by Ginzburg [4–6]), is usually extended to other media in which the plasma parameters (the electron density, density of the medium, electron and ion temperatures, and others) significantly differ from the ionospheric ones.

Subsequent estimates of the effective field in plasma were associated with the use of the kinetic approach based on the solution of the Bogolyubov–Born–Green–Kirkwood–Yvon (BBGKY) hierarchy of kinetic equations.

Based on the solution of the BBGKY hierarchy of kinetic equations in the binary collision approximation, Kadomtsev [7] made an estimate of the analytical smallness of the polarization correction to the acting field in a plasma. It was demonstrated that, for conventional plasma media, this correction is always negligible. Kadomtsev [7] made corresponding estimates for three-particle distribution functions, assuming the probability of three-particle correlation to be zero. Since, in this approximation, the polarization correction can be directly found in terms of the second correlation functions [7], whose current values are a priori assumed to be negligible, one can refer to the negligibility of this correction without performing further kinetic analysis. In this sense, the result is predetermined and, once obtained, is valid only for highly rarefied plasma media in which the probability of three-particle correction is very low.

At the same time, the Coulomb interaction of particles in plasma is long-range and, since (by definition) the region bounded by the Debye radius contains many particles, the dipole interaction of screening correlation electron "clouds" shifted relative to ions is inevitably present in the external electric field. In addition, one must take into account the interaction between an isolated electron and a system of these "clouds." A rigorous inclusion of such a collective interaction is impossible in the approximation of two-particle correlation functions.[2] A further increase in the capabilities of the BBGKY method in application to dense plasma media involves significant and largely insurmountable mathematical difficulties.

Note that the use of a rigorous kinetic approach to determining different physical quantities in the thermodynamic limit leads, in quite a number of cases, to extremely simple results [9, 10]. A characteristic feature of these results is that they may be obtained using simple and physically descriptive models. In connection with this, we will demonstrate below the possibilities of the modified Lorentz–Mossotti approach as regards the determination of the polarization correction in various plasma media.

The objective of this study is to find the polarization correction to the acting field as a function of plasma parameters using the model notions of the Debye

[1] Detailed analysis of such an approach may be found in [6].

[2] This is treated in detail in [8].

1063-7842/04/4901-0047$26.00 © 2004 MAIK "Nauka/Interperiodica"

screening of a charge in an isolated small volume of a plasma medium.

BASIC POSTULATES AND RELATIONS

In the most general case, the expression for the effective electric field acting on an electron in an isotropic plasma may be written as

$$\mathbf{E}_{ef} = \mathbf{E} + a\mathbf{P}/\varepsilon_0, \quad (1)$$

where $\mathbf{E}$ is the mean macroscopic field; $\mathbf{P}$ is the polarization vector of the medium; ε_0 is the dielectric permittivity of a vacuum; and a is a coefficient that may depend on the electron density, the density of the medium, the electron and ion temperatures, and the charge screening radius associated with them.

As applied to dielectric media under certain model assumptions [11], the coefficient a may be set equal to 1/3. In this case, the quantity $\mathbf{P}/3\varepsilon_0$ in Eq. (1) is the so-called Lorentz polarization correction. It follows from experiment that, for plasma media, the coefficient $a \neq 0$; however, under certain (including ionospheric) conditions, one can assume that $a \approx 0$. Then, we can approximately assume that

$$\mathbf{E}_{ef} = \mathbf{E}. \quad (2)$$

In the case of $a \neq 0$ and for the harmonic dependence of the external field $\mathbf{E} = \mathbf{E}_0\exp(\pm i\omega t)$, the expression for the complex permittivity of the medium $\dot{\varepsilon}$ has the form [6]

$$\dot{\varepsilon} = 1 - \frac{e^2 N}{\varepsilon_0 \eta m_e \omega^2}\left[1 + \frac{e^2 a N}{\varepsilon_0 \eta m_e \omega^2}\right]^{-1}, \quad (3)$$

where N is the electron density of the medium, $\eta = 1 \pm i\Omega/\omega$ (or some other function of Ω and ω; examples of graphs of this function for the real $\eta_{Re\varepsilon}(\Omega/\omega)$ and imaginary $\eta_{Im\varepsilon}(\Omega/\omega)$ parts of $\dot{\varepsilon}$ are given, e.g., in [12]), and Ω is the effective collision frequency of electrons with heavy particles of the medium (in what follows, we will first assume that $\Omega = 0$; the obtained results may be readily generalized to the case of $\Omega \neq 0$). The rest of notation is conventional.

Note that relation (2), which was initially proved for relatively low values of N, is inapplicable to the case of high N. Indeed, at high electron densities, along with the polarization of an isolated small volume of the plasma medium, one must take into account the polarization associated with the electron scattering by the nearest ions and the effect of far charges in this volume on an individual electron. This effect was correctly ignored by Ginzburg [4–6], because the vectors of the corresponding partial polarization shifts of the electron component of an isolated volume under the conditions treated in [4–6] are equal in magnitude and opposite in direction. The restrictions on the electron temperature and density [4] determine the applicability range of Eq. (2).

In order to find a more general expression for the acting field in a plasma medium with allowance for charge screening, we will apply the Lorentz–Mossotti method [11] and isolate a spherical region in which charge separation of any significance is possible. The size of this region is assumed to be small compared to both the wavelength λ of the external field $\mathbf{E}$ and the characteristic scale on which the plasma parameters vary. The radius r_0 of this region may be related, e.g., to the Debye radius D of charge screening in a plasma, $r_0 \propto D$. In accordance with this method, we will seek the field acting on an isolated electron in the form of the sum of fields

$$\mathbf{E}_{ef} = \mathbf{E} + \mathbf{E}_1 + \mathbf{E}_2. \quad (4)$$

Here, $\mathbf{E}_1$ is the field produced by charges of the internal surface of the sphere that are formed under the effect of the external field $\mathbf{E}$ when all ions and electrons (except for the isolated electron) are removed from this spherical region, and $\mathbf{E}_2$ is the field characterizing the interaction of the isolated electron with all ions and electrons of the spherical plasma volume.

The use of these model concepts of the origin and character of the field $\mathbf{E}_1$ enables one to employ the analogous apparatus of the field calculations in a dielectric (see, e.g., [11]). In this case, the field $\mathbf{E}_1$ in both the plasma and the dielectric is expressed in terms of the polarization vector as

$$\mathbf{R}_1 = \frac{1}{3\varepsilon_0}\mathbf{P}. \quad (5)$$

In order to calculate the field $\mathbf{E}_2$, we will treat the problem of electron scattering in the ion field with allowance for the Debye screening of charge in a plasma, thereby taking into account the effect of far charges on the character of electron scattering. The Debye potential ϕ in the gas approximation ($e\phi \ll k_0 T_e$) is described by the expression [6]

$$\phi(r) = (e/4\pi\varepsilon_0 r)\exp(-r/D), \quad (6)$$

where

$$D = \{\varepsilon_0 k T_e T_i/[e^2(T_e + T_i)N]\}^{1/2}, \quad (7)$$

k is the Boltzmann constant, and T_e and T_i are the electron and ion temperatures, respectively.

The notion of the Debye radius of charge screening (along with Eq. (6)) may be generalized to the case of solid-state plasma (degenerate and nondegenerate). For example, Eqs. (6) and (7) for nondegenerate and degenerate semiconductors under conditions of thermodynamic equilibrium will be respectively rewritten as follows [13]:

$$\phi(r) = (e/4\pi\varepsilon_0\varepsilon_{st} r)\exp(-r/D), \quad (8)$$

$$D = \sqrt{\frac{\varepsilon_0\varepsilon_{st}kT}{e^2N}}, \quad (9)$$

$$D = \sqrt{\frac{\varepsilon_0\varepsilon_{st}kT}{e^2N_{c,v}}\Phi'_{1/2}(\zeta_{c,v}/kT)}, \quad (10)$$

where ε_{st} is the stationary permittivity of the semiconductor, $\Phi_{1/2}(\cdot)$ is the Fermi–Dirac integral [13], $\Phi'_{1/2}(z) = d\Phi(z)/dz$, $N_{c,v} = N/\Phi_{1/2}(\zeta_{c,v}/kT)$ is the effective density of states in the conduction band (c) or valence band (v), and $\zeta_{c,v}$ is the chemical potential for electrons (c) and holes (v).

Under conditions of complete degeneracy, expression (10) has the following simple form [13]:

$$D = \left(\frac{\pi}{3N}\right)^{1/6}\sqrt{\frac{\varepsilon_0\varepsilon_{st}h^2}{4\pi e^2m^*}}, \quad (11)$$

where h is the Planck's constant and m^* is the effective mass of the charge carrier.

In particular, we note that, in metals, in which the electron gas is highly degenerate and the impurity concentration is relatively low, formula (6) has a somewhat different form: concentric regions with an increased and a decreased correlation density of the screening charge are formed, and a charge halo is formed around the scattering Coulomb center [9]. Therefore, the applicability range of Eqs. (6)–(11) is restricted to semiconductors. For metals, these and subsequent formulas may be used with caution and only for making approximate estimates.

As in [6], we will solve the problem of finding the variation of the electron velocity $\delta\mathbf{v}$ after the electron scattering from a center of force characterized by the field potential given by Eq. (6). Here, $\delta\mathbf{v}$ is the variation of the electron velocity averaged over all impact parameters ρ and initial directions. In view of this averaging, one can write [6]

$$\delta\mathbf{v} = -(v/3\rho)\frac{d}{d\rho}(\rho\sin\theta)\mathbf{s}_0, \quad (12)$$

where $\mathbf{s}_0$ is the shift of an electron due to its scattering.

Over the time δt, an electron experiences, on the average, $Nv\delta t$ collisions. Then, on multiplying Eq. (12) by $Nv\delta t$ and performing integration over all impact parameters ρ, we will arrive at the following result:

$$\delta\mathbf{v}' = \int_0^{\rho_m}\delta\mathbf{v}Nv\delta t2\pi\rho d\rho = -(2\pi/3)v^2N\delta t\mathbf{s}_0(\rho\sin\theta)\Big|_0^{\rho_m}, \quad (13)$$

where ρ_m is the maximum impact parameter.

Based on Eq. (13), we can write

$$\delta\mathbf{v}'/\delta t = -(2\pi/3)v^2N\mathbf{s}_0(\rho\sin\theta)\Big|_0^{\rho_m}. \quad (14)$$

The quantity $\rho\sin\theta$ in Eq. (14) may be evaluated in a linear approximation based on the classical problem of electron scattering in the Coulomb field (see, e.g., [13]) if the approximation

$$\phi(r) \approx (e/4\pi\varepsilon_0 r)[1 - 2r/(3D)] \quad (15)$$

is used in Eq. (6) (this approximation produces a difference from the exponential of no more than 10% up to argument values of $(r/D) \sim 1$).

In order to find $\rho\sin\theta$, we will substitute $e\phi$ into the scattering integral [13],

$$\frac{\theta}{2} = \frac{\pi}{2} - \int_{r_{min}}^{\infty}\frac{\rho dr}{r^2\sqrt{1 - \frac{\rho^2}{r^2} - \frac{2e\phi(r)}{m_ev^2}}}, \quad (16)$$

where r_{min} is found from the condition that the radicand is zero.

Integration of Eq. (16) in view of (15) gives

$$\rho = \frac{e^2}{4\pi\varepsilon_0m_ev^2\sqrt{1 + e^2/3\pi\varepsilon_0m_ev^2D}}\cot(\theta/2). \quad (17)$$

Using Eq. (17), one can readily derive the limiting expression for $\rho\sin\theta$:

$$\rho\sin\theta = (e^2/[2\pi\varepsilon_0m_ev^2])\{1 + e^2/[3\pi\varepsilon_0m_ev^2D]\}^{1/2} \times \cos^2(\theta/2) \xrightarrow[\rho\to\infty]{} (e^2/[2\pi\varepsilon_0m_ev^2]) \times \{1 + e^2/[3\pi\varepsilon_0m_ev^2D]\}^{-1/2}. \quad (18)$$

Based on Eq. (18), expression (14) may be rewritten as

$$m_e\delta\mathbf{v}'/\delta t = -(e^2N/3\varepsilon_0) \times \{1 + e^2/[3\pi\varepsilon_0m_ev^2D]\}^{-1/2}\mathbf{s}_0. \quad (19)$$

After summing up Eq. (19) over all electrons $N^* = (4\pi/3)R^3N$ contained in the isolated spherical volume of radius R, we obtain the equation of motion for all electrons,

$$m_e\sum_{k=1}^{N^*}\delta\mathbf{v}'_k/\delta t = -(e^2N/3\varepsilon_0) \times \{1 + e^2/[3\pi\varepsilon_0m_ev^2D]\}^{-1/2}\sum_{k=1}^{N^*}\mathbf{s}_{0k}. \quad (20)$$

Using Eq. (20), one can readily determine the field $\mathbf{E}_2$ related to the polarization shift of electrons that is caused by the scattering in the Coulomb field with

allowance for the Debye screening of charges in the plasma,

$$\mathbf{E}_2 = -\frac{1}{3\varepsilon_0}\{1 + e^2/[3\pi\varepsilon_0 m_e v^2 D]\}^{-1/2}\mathbf{P}. \quad (21)$$

We substitute Eqs. (5) and (21) into Eq. (4) to derive the expression for the effective field as a function of **E** and **P**,

$$\mathbf{E}_{ef} = \mathbf{E} + \frac{1}{3\varepsilon_0}[1 - \{1 + e^2/[3\pi\varepsilon_0 m_e v^2 D]\}^{-1/2}]\mathbf{P}. \quad (22)$$

The second term in Eq. (23) is the sought polarization correction to the mean macroscopic field **E** in the plasma. Accordingly, the coefficient a in Eq. (3) is

$$a = \frac{1}{3}\left[1 - \left(1 + \frac{e^2}{3\pi\varepsilon_0 m_e v^2 D}\right)^{-1/2}\right]. \quad (23)$$

The quantity **P** is a function of **E**; therefore, in order to find the explicit expression for $\mathbf{E}_{ef}(\mathbf{E})$, one must, generally speaking, solve anew the problem of electron motion in the effective field, i.e., solve the following equation (here, we will again introduce the effective electron collision frequency Ω with the heavy particles of the medium):

$$\ddot{\mathbf{s}} + \Omega\dot{\mathbf{s}} = \frac{e}{m_e}[\mathbf{E} + a\mathbf{P}/\varepsilon_0], \quad (24)$$

where **s** is the electron shift caused by the effective field.

For harmonic fields, the solution to Eq. (24) may be written as

$$\mathbf{P}(\omega) = -\frac{\varepsilon_0\omega_0^2}{\omega^2 + a\omega_0^2 - i\omega\Omega}\mathbf{E}(\omega), \quad (25)$$

where $\omega_0 = \sqrt{e^2 N/\varepsilon_0 m_e}$ is the plasma frequency and $\mathbf{P} = eN\mathbf{s}$.

We substitute Eq. (25) into Eq. (22) to eventually derive the explicit expression for the effective field in the plasma,

$$\begin{aligned}\mathbf{E}_{ef} &= \mathbf{E}\left(1 - \frac{a\omega_0^2}{\omega^2 + a\omega_0^2 - i\omega\Omega}\right) \\ &= \mathbf{E}\left(1 - \frac{a\omega_0^2}{\eta\omega^2(1 + a\omega_0^2/\eta\omega^2)}\right).\end{aligned} \quad (26)$$

Note that expression (26) was derived irrespective of the methods for determining the quantity a (it is only its presence that is important). Therefore, this expression remains valid in the widest range of variations in the parameters appearing in this expression (i.e., in the region where expression (3) for $\dot{\varepsilon}$ remains valid. In this case, the right-hand side of Eq. (26) may be further generalized to the case in which $\eta \neq 1 \pm i\Omega/\omega$, as was indicated above). For a relative variation of the field in a plasma, one can use Eqs. (3) and (26) and write the following expression:

$$\frac{\Delta\mathbf{E}}{\mathbf{E}} = \frac{\mathbf{E} - \mathbf{E}_{ef}}{\mathbf{E}} = [\dot{\varepsilon}(a, \omega) - 1]a. \quad (27)$$

The physical reason for the deviation of the acting field in a plasma from the mean field consists in the polarization shift of the electron component relative to the ions, which arises both as a result of polarization of an isolated small volume of the plasma medium under the action of the external field and due to electron scattering by screened ions; i.e., this deviation is associated with the resultant shift of the screening cloud of the charge relative to the Coulomb center.

Indeed, in the absence of an external field, the Debye screening of charge in a plasma may be obtained as a result of either a direct solution of Poisson's equation with the use, e.g., a Boltzmann distribution [6, 8] or a kinetic treatment in the model of two-particle interaction. Therefore, it is natural to consider the polarization of a "quasi-particle" (ion + screening electron cloud) in the external electric field. As a result of summation of this polarization effect of all quasi-particles on an isolated electron, one can determine the field $\mathbf{E}_1$ (formula (5)), which turns out to be directed oppositely to the external field $\mathbf{E}_0$. Then, the isolated electron is scattered by the screened ion under the action of the external field. In this case, the polarization shift of scattering together with the field $\mathbf{E}_2$ in the thermodynamic limit (formula (22)) turns out to be directed along the external field (the electrons are decelerated by the field of quasi-particles). As follows from Eq. (22), the greatest contribution to $\mathbf{E}_2$ is made by slow electrons (they are decelerated by the field of quasi-particles to a greater extent). However, the summation of $\mathbf{E}_1$ and $\mathbf{E}_2$ eventually brings about a polarization shift opposite to **E** (see Eqs. (22) and (23)). A similar result is obtained in the case of the kinetic treatment of the problem: the correlation electron cloud shifts relative to the screened ion along the external field **E** [7].

Expression (26) for the effective field, obtained using the elementary approach, may be transformed to the form given in [7]. For this, one must assume that $e^2/3\pi\varepsilon_0 m_e v^2 D \ll 1$ (this is equivalent to the requirement that the plasma medium be rarefied). Then, taking into account the first two terms of expansion of a in the small parameter, one can use the following approximation:

$$a = \frac{1}{3}\left[1 - \left(1 + \frac{e^2}{3\pi\varepsilon_0 m_e v^2 D}\right)^{-1/2}\right] \approx \frac{e^2}{18\pi\varepsilon_0 m_e D}\frac{1}{v^2}. \quad (28)$$

Substituting Eq. (28) into (26) and averaging over the velocity in the high-frequency approximation (for weak fields), one can readily derive the following

expression for relative variation of the mean effective field:

$$\frac{\mathbf{E}-\mathbf{E}_{\mathrm{ef}}^{s}}{\mathbf{E}} \cong \frac{1}{8\pi NDD_0^2}\frac{\omega_0^2}{\omega^2 - i\omega\Omega} = \frac{1}{8\pi ND^3}\frac{\omega_0^2}{\omega^2 - i\omega\Omega}\frac{T_i}{T_i + T_e}, \quad (29)$$

where the superscript *s* implies averaging.

As was indicated above, the results of using the above classical approach to estimating the polarization correction in the problem of electromagnetic wave propagation in plasma media may be generalized to the case of solids characterized by intrinsic conductivity, e.g., semiconductors. In such solids, the Debye radius is determined on the basis of Eqs. (7)–(11) using the notion of the effective mass m^* of charge carriers, which may be much less than the electron mass. In this case, it is obvious that the field $\mathbf{E}_1$ is equal to $\mathbf{E}_1 = \mathbf{P}/3\varepsilon_0\varepsilon_{\mathrm{st}}$ and that Eq. (23) for the coefficient a with allowance for Eqs. (8) and (15) can be rewritten as

$$a = \frac{1}{3\varepsilon_{\mathrm{st}}}\left[1 - \left(1 + \frac{e^2}{3\pi\varepsilon_0 m_e v^2 D\varepsilon_{\mathrm{st}}}\right)^{-1/2}\right]. \quad (30)$$

As an illustrative example of inclusion of the decisive role of the obtained polarization correction, we will refer to the estimation of the skin depth for a ZnS piezoelectric semiconductor exhibiting electron conductivity with $N \approx 10^{16}$ cm^{-3} at a temperature $T = 273$ K [14]. The mean effective electron mass in ZnS is $m_e^* \sim 0.25m_e$, $\varepsilon_{\mathrm{st}} = 8.32$, and the electron relaxation frequency on optical lattice vibrations is $\Omega = 1/\tau \sim 7 \times 10^{11}$ s^{-1} [15]. For a wavelength of $\lambda = 1$ cm and chosen conditions, the coefficient a is approximately equal to 0.01. Estimates show that the skin depth ($d = \lambda/[12\pi \mathrm{Im}\sqrt{\varepsilon}\,]$) is approximately 2.5 times that for the case of $a = 0$. This ratio increases with increasing wavelength and tends to unity with decreasing wavelength. For example, for $\lambda = 10$ cm, this ratio increases to 7; for $\lambda = 0.1$ cm, it decreases to 1.15; and, for $\lambda = 0.01$ cm, it is close to 1.

CONCLUSIONS

The application of a systematic approach based on the classical concepts of the polarization of an isolated small plasma volume and electron scattering by a screened Coulomb center in an external field enables one to perform an analytical estimation of the polarization correction in dense plasma media (including degenerate solid-state plasma). In such media, the polarization correction may play a decisive role, thus significantly changing various electrodynamic characteristics of the media.

REFERENCES

1. C. G. Darwin, Proc. R. Soc. London, Ser. A **146**, 17 (1934).
2. L. I. Mandel'shtam, J. Phys. **4**, 9 (1941).
3. C. G. Darwin, Proc. R. Soc. London, Ser. A **182**, 152 (1943).
4. V. L. Ginzburg, Izv. Akad. Nauk SSSR, Ser. Fiz. **8**, 76 (1944).
5. V. L. Ginzburg, Zh. Éksp. Teor. Fiz. **18**, 487 (1948).
6. V. L. Ginzburg, *Theory of Radio Wave Propagation in Ionosphere* (OGIZ, Moscow, 1949).
7. B. B. Kadomtsev, Zh. Éksp. Teor. Fiz. **33**, 151 (1957) [Sov. Phys. JETP **6**, 117 (1958)].
8. Yu. L. Klimantovich, *Statistical Physics* (Nauka, Moscow, 1982).
9. J. M. Ziman, *Principles of the Theory of Solids*, 2nd ed. (Cambridge University Press, London, 1972; Mir, Moscow, 1966).
10. E. M. Lifshitz and L. P. Pitaevskii, *Physical Kinetics* (Nauka, Moscow, 1979; Pergamon, Oxford, 1981).
11. A. R. von Hippel, *Dielectrics and Waves* (Wiley, New York, 1954; Inostrannaya Literatura, Moscow, 1960).
12. V. L. Ginzburg, *The Propagation of Electromagnetic Waves in Plasmas* (Nauka, Moscow, 1967; Pergamon, Oxford, 1970).
13. V. L. Bonch-Bruevich and S. G. Kalashnikov, *Physics of Semiconductors* (Nauka, Moscow, 1977).
14. L. D. Landau and E. M. Lifshitz, *Course of Theoretical Physics*, Vol. 1: *Mechanics* (Nauka, Moscow, 1982; Pergamon, New York, 1988).
15. *Tables of Physical Data: A Reference Book*, Ed. by I. K. Kikoin (Atomizdat, Moscow, 1976).
16. A. I. Ansel'm, *Introduction to the Theory of Semiconductors* (Nauka, Moscow, 1978).

Translated by A. Bronshteĭn

Technical Physics, Vol. 49, No. 1, 2004, pp. 52–56. Translated from Zhurnal Tekhnicheskoĭ Fiziki, Vol. 74, No. 1, 2004, pp. 53–57.

SOLIDS

Structure and Properties of TiC–TiNi Composites Alloyed with Iron

V. P. Sivokha, Yu. P. Mironov, V. V. Ruday, and S. N. Kulkov

Institute of Strength Physics and Materials Science, Siberian Division, Russian Academy of Sciences, Akademicheskiĭ pr. 2/1, Tomsk, 634021 Russia

e-mail: kulkov@ms.tsc.ru

Received May 8, 2003

Abstract—The structure, composition, and mechanical properties of iron-alloyed TiC–TiNi composite materials are studied. When the titanium carbide framework is sintered with iron and then impregnated with titanium nickelide, iron atoms are found to diffuse into the matrix and form the *B*2 structure that is inhomogeneous (gradient) in chemical composition and properties and exhibits various temperatures of martensitic transformation. The latter fact shows up in the broadening of the martensitic transformation hysteresis and its shift toward low temperatures with increasing iron content. At room temperature, the strength properties of gradient-matrix TiC–TiNi composites are shown to increase with iron concentration. © *2004 MAIK "Nauka/Interperiodica".*

INTRODUCTION

An increase in the ductility of the matrix of a composite material via martensitic transformation was studied in [1–3]. When plastic shear is confined within thin intergranular layers of the matrix, this mechanism efficiently releases peak stresses that appear upon loading composite materials, and the deformation of titanium nickelide as the crystal lattice loses shear stability modifies its structural state.

However, this mechanism of stress relaxation in composite materials works only in a narrow temperature range near phase transition temperatures and at strictly defined stresses. The temperature–stress range of structural transformations that occur in the matrix of a composite material under an applied load may be extended by providing a chemical composition gradient in the matrix. In this case, its microvolumes could undergo structural transformations at different temperatures and stresses. As is known [4, 5], the parameters of martensitic transformation (MT) in TiNi can be controlled by alloying. For example, the substitution of iron or cobalt for nickel decreases the temperature at which the initial *B*2 structure transforms into the *B*19' martensitic phase and changes the sequence of phase transitions resulting in the formation of the intermediate *R* phase, which also transforms into *B*19' martensite under the action of applied stresses or when the temperature decreases [4, 5]. On the contrary, the alloying of TiNi with palladium, platinum, gold, or zirconium leads to an increase in the MT temperature [5]. Thus, the introduction of the elements listed above into different areas of the TiNi matrix can create a composition gradient, thereby producing microvolumes with different temperature and force parameters of martensitic transformation.

However, data for structural instability in the chemically graded matrix of a composite material are currently lacking. The purpose of this work is to investigate the structure and properties of a TiNi-based composite with a structurally unstable gradient matrix.

EXPERIMENTAL

A composite material was prepared via the method of powder metallurgy, which involves the sintering of Ti–15 wt % C ($TiC_{0.7}$) powder to form a TiC framework, followed by the impregnation of the framework with TiNi of equiatomic composition. A composition gradient was created using iron, since, as compared to other elements, a small increase in the iron concentration in TiNi significantly decreases the MT temperature and changes the sequence of MTs [4, 5]. This allows one to effectively control the MT temperature and force parameters upon the *B*2–martensite transition even if the alloying element content is low.

PZhR iron powder in amounts of 1, 3, 5, and 6 wt % was mixed with the TiC powder in a ball mill, and then the mixture was compacted at a pressure of 150 MPa. The preforms were sintered at a pressure of no lower than 10^{-2} Pa at 1550°C for 3 h. Under such sintering conditions, the porosity of TiC frameworks was 31–32 vol % irrespective of the iron content. The frameworks were then impregnated with TiNi at a pressure of no lower than 10^{-2} Pa at 1350°C for 10 min. After impregnation, the TiC framework-to-TiNi matrix weight ratio was 60 : 40. Test specimens were spark-cut and then mechanically finished with a diamond paste.

Phase analysis was performed with a DRON-UM1 X-ray diffractometer (CuK_α radiation) at room temperature. The temperature intervals of martensitic transfor-

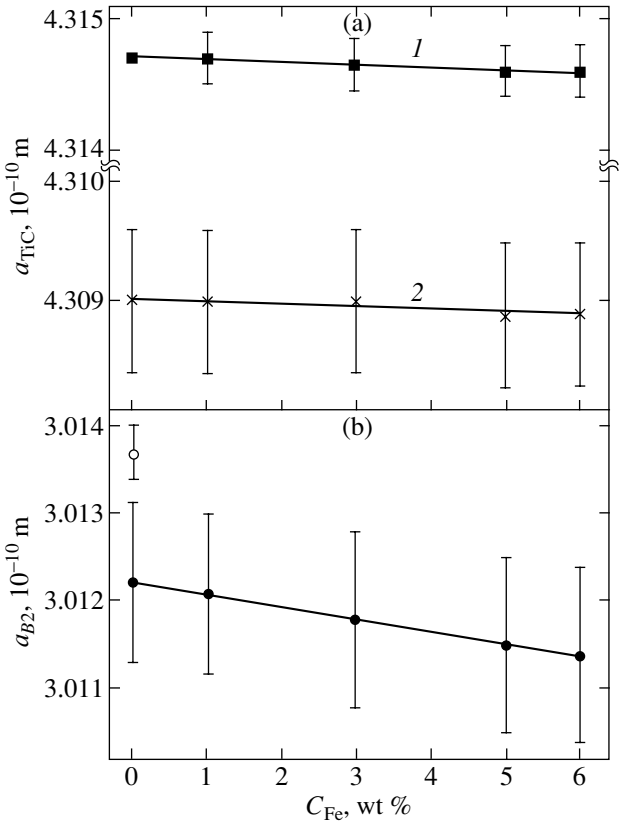

Fig. 1. Unit cell parameters of (a) titanium carbide and (b) titanium nickelide in (*1*) sintered TiC + Fe frameworks and (*2*) (TiC–TiNi) + Fe composite material vs. the iron concentration.

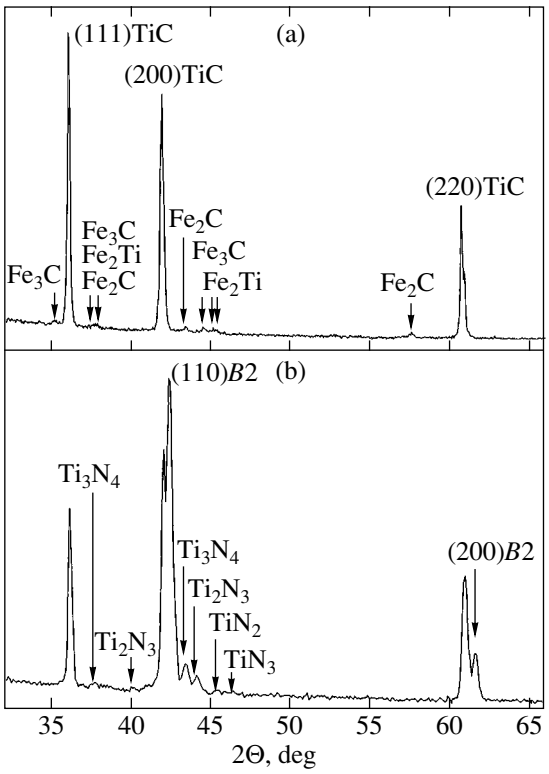

Fig. 2. X-ray diffraction patterns taken from the (a) sintered TiC + 6 wt % Fe framework and (b) TiC–TiNi + 6 wt % Fe composite material.

mations occurring in the matrix of the composite were determined from the temperature dependence of the electrical resistance. The strength properties of the composite were found from the critical breaking stress of the specimen subjected to three-point bending on an INSTRON-1185 testing machine.

RESULTS AND DISCUSSION

X-ray diffraction data show that the TiC frameworks sintered have the fcc lattice with a parameter of 0.43146 ± 0.00003 nm. The unit cell parameter (Fig. 1a) and the FWHM of diffraction reflections remain virtually unchanged with increasing iron content. The diffraction pattern taken from the TiC + 6 wt % Fe specimen (Fig. 2a) contains additional weak reflections, which may be assigned to iron-containing phases, such as Fe_2Ti, Fe_2C, and Fe_3C. These phases may form as a result of chemical interaction between the iron and titanium carbide upon sintering; their volume fraction does not exceed 5%.

After the impregnation of the frameworks, along with the reflections from the basic phases TiC and TiNi with the *B*2 structure, the reflections from nickel-enriched intermetallics (Ni_4Ti_3, Ni_3Ti_2, and Ni_3Ti) are also observed (Fig. 2b). The total volume fraction of these phases is about 8%. We failed to detect iron compounds present in the framework before impregnation probably because of their dissolution in the binder of the composite material.

Upon cooling, the matrix undergoes martensitic transformation, as follows from resistivity measurements (Fig. 3). The nonmonotonic run of the curve $\rho(T)$ during cooling and heating of the material and the formation of a hysteresis loop indicate martensitic transformation in the given temperature range [4, 5]. As the iron content rises, the hysteresis loop in the curve $\rho(T)$ broadens and shifts toward low temperatures (Fig. 3). Such behavior of the temperature dependence of the resistivity may be attributed to the formation of microvolumes where the MT temperatures are lower than in the rest of the matrix.

Such microvolumes with different MT temperatures may arise due to the dissolution of the iron in local areas of the matrix, which makes it graded in terms of composition and, hence, MT temperatures.

In all the hard alloys prepared by impregnating the TiC framework with TiNi, the FWHM of the reflections from the TiC increases (on average, by a factor of two as compared to the FWHM before impregnation)

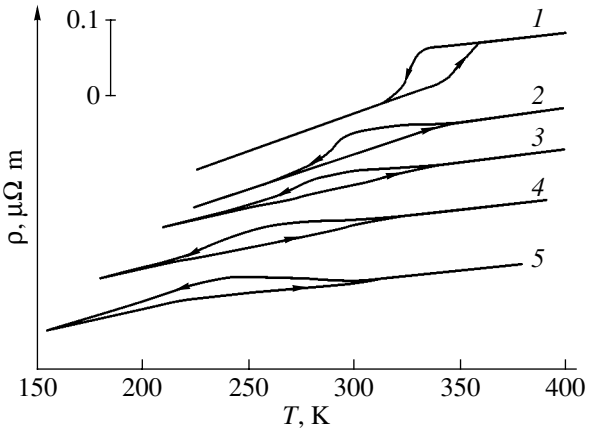

Fig. 3. Temperature dependences of the resistivity upon cooling and heating (*1*) Ti–50.0 at. % Ni, (*2*) TiC–TiNi, (*3*) TiC–TiNi + 1 wt % Fe, (*4*) TiC–TiNi + 3 wt % Fe, and (*5*) TiC–TiNi + 6 wt % Fe composites.

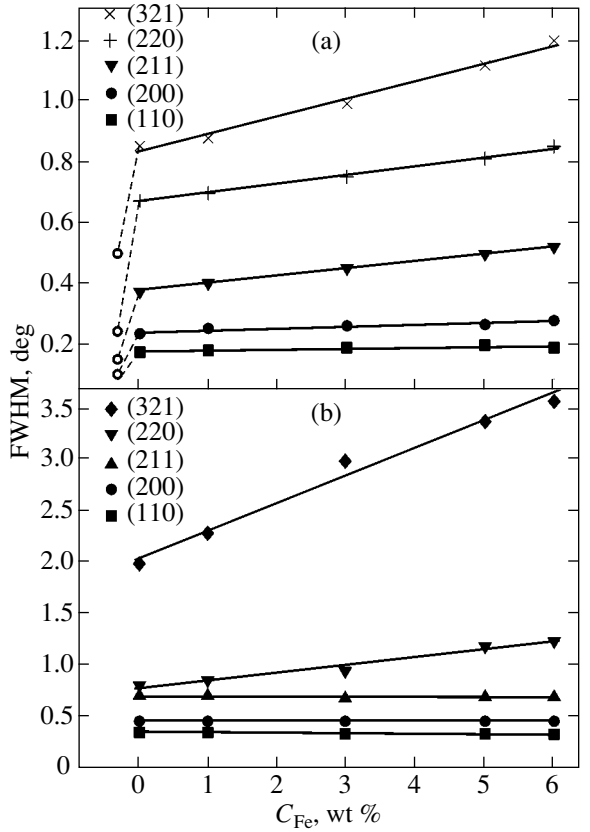

Fig. 4. FWHM of X-ray reflections from (a) the titanium carbide and (b) titanium nickelide in (TiC–TiNi) + Fe composites vs. the iron concentration (○, values for the sintered TiC framework).

(Fig. 4a) and shifts toward higher diffraction angles. The unit cell parameter of the TiC, a_{TiC}, decreases to 0.4309 nm and is virtually independent of the iron content in it (Fig. 1a). Since the unit cell parameter of titanium carbide depends substantially on the titanium-to-carbon ratio [6], the broadening of the reflections is most likely to be due to the variation of the TiC_x chemical composition in the microvolumes, while the decrease in the unit cell parameter indicates that x decreases from 0.7 to 0.6. Thus, during the impregnation of the TiC framework with TiNi, the carbon diffuses from the titanium carbide to the matrix and interacts with the titanium of the matrix, producing carbide particles with a composition other than the initial composition $TiC_{0.7}$. The matrix becomes nickel-enriched, as a result of which the MT temperature in it and the unit cell parameter of the *B*2 phase decrease (Figs. 3 and 1b, respectively). As the iron content in the composite grows, the unit cell parameter of the *B*2 phase somewhat diminishes (Fig. 1b). The reflections from high-index planes in the *B*2 phase become much broader, while the FWHM of the reflections from low-index planes remains unchanged (Fig. 4b). To separate out contributions to the reflection intrinsic broadening that are due to microstrains and small grain size, we used the technique [7] based on the construction of interpolation plots in the coordinates $[(\beta\cos\theta)/\lambda]^2$ – $[(\sin\theta)/\lambda]^2$, where β is the reflection FWHM in radians. It was found that, for both TiC and TiNi, the broadening of the reflections after impregnation is largely caused by microstrains in their crystal lattices. The choice of the quadratic coordinates for the interpolation plots is dictated by the fact that the shapes of the diffraction reflections from both the carbide phase and the *B*2 phase are closer to the Gaussian than to the Lorentzian form. Figure 5 shows the dependence of microstrains in the TiC and TiNi lattices on the content of iron introduced into the carbide framework. For the iron-free specimens, the value of $\langle\varepsilon^2\rangle^{1/2}$ is seen to be 1.1×10^{-3} for the *B*2 phase and 0.9×10^{-3} for the titanium carbide. As the iron content in the composite material rises, the microstrains build up in both TiC and TiNi, the buildup in the latter being greater.

There are two basic reasons for the broadening of diffraction peaks and the generation of microstrains. The first is thermal stresses arising upon cooling the composite material from the annealing temperature; the other, the presence of a composition gradient in the TiC and TiNi phases.

To estimate the contribution associated with the first reason, we calculated thermal stresses due to a mismatch in the thermal expansion coefficients (TECs) of the basic phases of the composite (for the calculation procedure, see [8]). Relevant parameters were taken from [6, 9]. The thermal stresses that appear at the carbide–binder interface upon cooling from the annealing temperature (400°C) to room temperature were found to be 100–160 MPa. Then, the average microstrains are 10^{-3} in the binder and 0.1×10^{-3} in the titanium carbide (Fig. 5, the dot-and-dash line). Since iron introduced in small amounts may change the TEC of the matrix insignificantly, this factor cannot give rise to high thermal stresses in either of the phases.

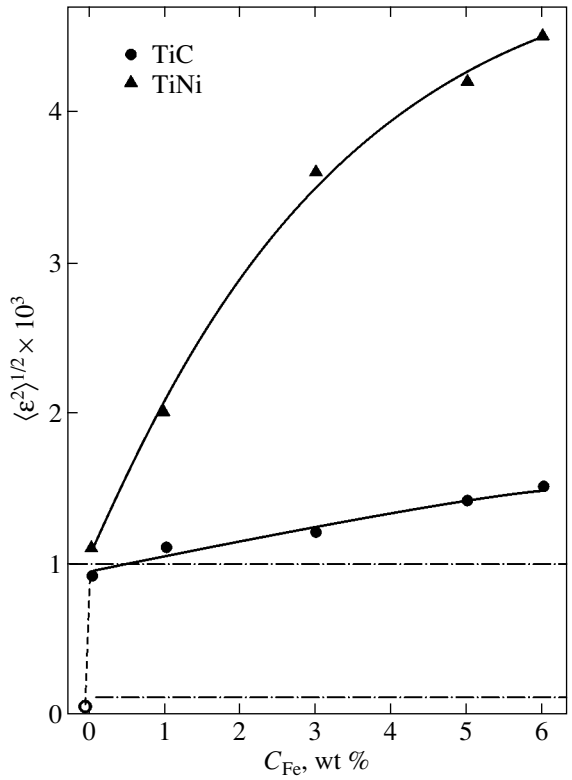

Fig. 5. Microstrains in the TiC and TiNi phases of the (TiC–TiNi) + Fe composite vs. the iron concentration (○, the value for the sintered TiC framework).

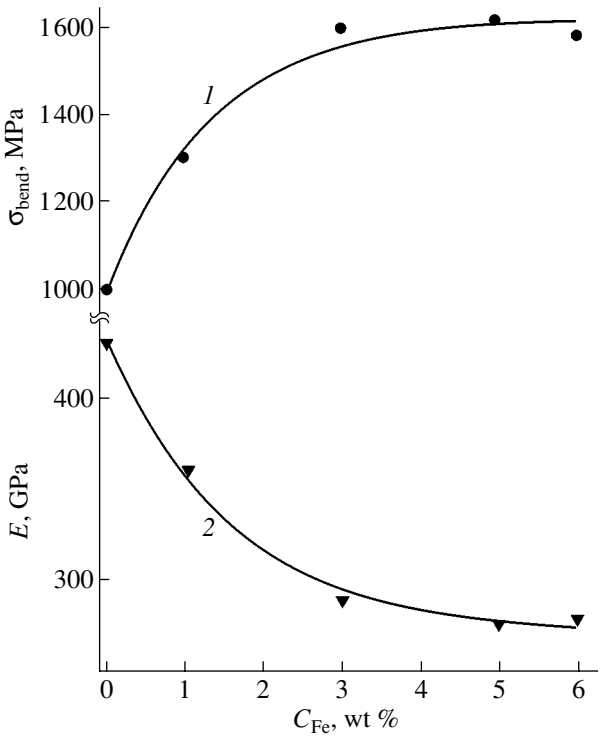

Fig. 6. (*1*) Bending strength and (*2*) elastic modulus of (TiC–TiNi) + Fe composites vs. the iron concentration.

Comparing the microstrains found experimentally, ε_1, with those obtained by calculation (see above), ε_2 (Fig. 5, the dot-and-dash line), shows that $\varepsilon_1 \approx \varepsilon_2$ for the *B*2 phase of the iron-free specimen and $\varepsilon_1 \gg \varepsilon_2$ for the titanium carbide. In other words, thermal stresses may be critical in distorting the lattice in TiNi with the *B*2 structure. For the carbide, the broadening of reflections is associated with the composition gradient (the formation of particles with different carbon content). In view of the intense diffusion of the carbon from the titanium carbide to the matrix during impregnation, we may assume that strains ε_1 in the titanium carbide result from the variation of the unit cell parameter a_{TiC}. Thus, the basic reason for microstrains in the TiC lattice is likely to be the formation of microscopic composition gradients and, to a lesser degree, static distortions of the lattice.

The increase in the iron concentration (Fig. 5) has a minor effect on the microstrains in the titanium carbide, whereas $\langle \varepsilon^2 \rangle^{1/2}$ in the matrix grows significantly (roughly by a factor of four). Therefore, one may assume that microstrains in the *B*2 phase are caused by thermal stresses and, even to a greater extent, by the formation of the composition gradient. The insignificant decrease in the unit cell parameter of the *B*2 phase (Fig. 1b) (the same is observed when TiNi is alloyed with Fe [4]) shows that the iron dissolves in the matrix. In addition, the related extension of the MT temperature range toward low temperatures (Fig. 3) indicates the nonuniform iron distribution and the formation of the graded *B*2 structure. As a result, microstrains in the matrix of the composite material grow considerably with increasing iron concentration.

The mechanical properties of such graded-composition materials are bound to reflect the structural modifications. In particular, the room-temperature strength of iron-alloyed composite materials is expected to be higher than in the initial TiC–TiNi material because of the formation of the graded-composition *B*2 matrix with a wide range of MT temperatures and martensitic shear stresses. Indeed, Fig. 6 shows that the bending strength of iron-alloyed composite materials increases with iron concentration, whereas the elastic modulus decreases. The increase in the bending strength can be explained by the formation of microvolumes with different compositions (a composition gradient) with an increase in the iron concentration in the matrix. Accordingly, peak stresses that appear upon loading the composite will relax in a wide range.

CONCLUSIONS

Iron introduced upon sintering the carbide framework does not change the stoichiometric composition

of titanium carbide; it only results in the formation of Fe_2Ti, Fe_2C, and Fe_3C compounds.

The impregnation of the carbide framework is accompanied by intense carbon diffusion to form additional TiC particles. The titanium carbide becomes nonuniformly depleted of carbon, and nickel-enriched Ni_mTi_n intermetallics form in the binder.

The alloying of the composite material with iron increases microstrains in the *B*2 structure fourfold, which may be related to the formation of the graded-composition matrix of the Ti–Ni–Fe ternary system.

A simultaneous increase in the ultimate strength and strain, along with a decrease in the elastic modulus, with content of the iron introduced into the framework, is due to the formation of graded-composition microdomains with the *B*2 structure. These microdomains provide the effective relaxation of peak stresses that arise upon loading the composite material.

ACKNOWLEDGMENTS

This work was supported by the Russian Foundation for Basic Research (project no. 02-02-17755), the Ministry of Atomic Energy of the Russian Federation (project no. 3.10-44), and CRDF (project PRHEGRANT REC-016).

REFERENCES

1. S. N. Kul'kov, T. M. Poletika, A. Yu. Chukhlomin, *et al.*, Poroshk. Metall., No. 8, 88 (1984).
2. S. N. Kul'kov, T. M. Poletika, and V. E. Panin, Poroshk. Metall., No. 7, 54 (1983).
3. D. Mari and D. C. Dunand, Metall. Mater. Trans. A **26**, 2833 (1995).
4. A. S. Savvinov, V. N. Khachin, and V. P. Sivokha, Izv. Vyssh. Uchebn. Zaved. Fiz., No. 7, 34 (1983).
5. V. N. Khachin, V. G. Pushkin, and V. V. Kondrat'ev, *Titanium Nickelide: Structure and Properties* (Nauka, Moscow, 1992).
6. S. S. Kiparisov, Yu. V. Umanskiĭ, and A. P. Petrov, *Titanium Carbide: Production, Properties, Application* (Metallurgiya, Moscow, 1987).
7. S. S. Gorelik, L. N. Rastorguev, and Yu. A. Skakov, *X-ray and Electron–Optical Analysis* (Metallurgiya, Moscow, 1970).
8. J. Gurland, Trans. ASME **50**, 1063 (1958).
9. I. I. Kornilov, O. K. Belousov, and E. V. Kachur, *Titanium Nickelide and Other Shape-Memory Alloys* (Nauka, Moscow, 1977).

Translated by K. Shakhlevich

Technical Physics, Vol. 49, No. 1, 2004, pp. 57–60. Translated from Zhurnal Tekhnicheskoĭ Fiziki, Vol. 74, No. 1, 2004, pp. 58–61.

SOLIDS

Dynamic Cracking Resistance of Structural Materials Predicted from Impact Fracture of an Aircraft Alloy

Yu. V. Petrov and E. V. Sitnikova

Research Institute of Mathematics and Mechanics, St. Petersburg State University, Universitetskaya nab. 7/9, St. Petersburg, 198504 Russia

e-mail: yp@YP1004.spb.edu

Received May 5, 2003

Abstract—A new approach to studying the dynamic strength properties of structural materials is demonstrated with fracture of 2024-T3 aircraft aluminum alloy. The central idea of this approach is the incubation time to failure. In [1], experimental data for dynamic fracture of this alloy were analyzed in terms of the classical fracture criterion, which is based on the principle of maximum critical stress intensity factor [2]. In [1], the dependence of the stress intensity factor limiting value (the dynamic fracture toughness K_{Id}, which was assumed to be a functional characteristic of the material) on the loading rate was also measured. The same experimental data were analyzed in terms of an alternative structure–time approach [3]. In this approach, the dynamic fracture toughness K_{Id} is considered as an estimable characteristic of the problem, so that determination of limiting loads does not require *a priori* knowledge of the loading-rate dependence of the dynamic fracture toughness. The incubation time to failure of the aircraft aluminum alloy is calculated. The difference in the loading-rate dependences of the dynamic fracture toughness, which is observed for various structural materials, is explained. The dynamic fracture toughness of the alloy under pulsed threshold loads is calculated. © *2004 MAIK "Nauka/Interperiodica".*

Quasi-brittle fracture of solids is known to be the result of cracking, the rate of which is governed by both the stressed state and structure of a solid. In the linear mechanics of fracture, it is assumed that the stress level in the vicinity of a crack depends on a single factor, namely, on the stress intensity factor K_I. For cracked specimens loaded statically, the limiting loads are determined in terms of the critical stress intensity factor (fracture toughness), which is considered to be a material constant. In static problems, the fracture criterion for regions with symmetrically loaded cracks has the form [2]

$$K_I \le K_{Ic}, \tag{1}$$

where K_I is the stress intensity factor and K_{Ic} is the static fracture toughness.

According to this criterion, a crack starts propagating when the stress intensity factor K_I reaches a certain critical value K_{Ic}, which is found experimentally for each material. This approach, fitting well the results of static fracture tests, is now universally accepted in engineering practice.

However, under conditions of dynamic loading, the situation changes. The strength properties of materials and constructions under static or dynamic loading are different. In the latter case, knowledge of only the critical stress intensity factor does not suffice to study fracture processes in cracked specimens. A series of experimental investigations into dynamic fracture has shown that the critical stress intensity factor is no longer a material constant. It varies in a complex manner with loading conditions, and the traditional approach becomes invalid. Therefore, other criteria are needed.

Owen *et al.* [1] experimentally studied the dynamic fracture toughness of 2024-T3 aircraft aluminum alloy. The authors used several schemes where notched-bar specimens of fixed length and width but various thickness had a fatigue crack at the end of the notch and were subjected to impact loading at different rates.

The fracture criterion used in [1] had the same form as under static loading, but the critical value of the stress intensity factor was conventionally considered as a function of the local loading rate. The time to failure was determined from the condition

$$K_I(t, P(t), a) \le K_{Id}(K_I'(t)). \tag{2}$$

Here, $P(t)$ is the generalized dynamic load and a is the crack initial length. The right of (2) represents the dynamic fracture toughness, which is a material function depending on the local loading rate $K_I'(t) = dK_I/dt$ in the scheme used. Such an approach is, however, inconvenient in applications, since it is difficult to predict the loading rate in advance. It is also inefficient because the dynamic fracture toughness may depend on both the history and the way of loading in a complex manner [4, 5]. For example, in [1], fracture was observed at the growth stage of the stress intensity factor and the loading rate was taken to be equal to the slope of the linear segment in the stress–strain curve.

1063-7842/04/4901-0057$26.00 © 2004 MAIK "Nauka/Interperiodica"

However, if threshold load pulses (e.g., ultimate pulses of given duration and critical amplitude) are applied, fracture may also take place after the stress intensity factor has reached a maximum. In this case, the dynamic fracture toughness may decrease with decreasing ultimate-pulse duration [4] rather than increase, as was observed in [1]. Thus, the dynamic fracture toughness appearing in condition (2) is not a material parameter, which brings about the need for alternative methods for testing the dynamic strength properties of materials.

Let us show that experimental data in [1] can be analyzed by using a much more convenient and efficient structure–time criterion [3]. For specimens with symmetrically loaded cracks, it has the form

$$\int_{t-\tau}^{t} K_{\mathrm{I}}(s)ds \le K_{\mathrm{Ic}}\tau, \tag{3}$$

where $K_{\mathrm{I}}(t)$ is the current value of the stress intensity factor and τ is the incubation time to failure, whose physical meaning is the characteristic relaxation time upon microfracture of a material.

An expression for the stress intensity factor as applied to the experiment [1] is found by solving the following initial- and boundary-value problem: an infinite elastic plane containing a semi-infinite crack $\Gamma_{\pm} = \{(x, y) : y = \pm 0, x \le 0\}$ is subjected to an impact load. The stressed state at the apex of the crack is determined from a solution to the equation of elastodynamics for an isotropic medium

$$\rho\frac{\partial^2 \mathbf{U}}{\partial t^2} = (\lambda + \mu)\,\mathrm{grad}\,\mathrm{div}\mathbf{U} + \mu\Delta\mathbf{U} \tag{4}$$

(which holds true at all inner points of this plane) with the boundary conditions

$$\sigma_{xy}|_{\Gamma_{\pm}} = 0, \quad \sigma_y|_{\Gamma_{\pm}} = -p(t), \tag{5}$$

the initial condition

$$\mathbf{U}|_{t\le 0} = 0, \tag{6}$$

and the energy condition

$$\forall t > 0\colon \mathbf{U} = \mathrm{const} + \mathrm{O}(r^{\beta}), \quad r \longrightarrow 0, \quad \beta > 0, \tag{7}$$

which provides the uniqueness of a solution to this problem. In (4), λ and μ are the Lamé parameters, $\mathbf{U} = \mathbf{U}(t, x_1, x_2)$ is the displacement vector, and Δ is the Laplacian. This initial- and boundary-value problem meets the experimental conditions in [1]. According to Hooke's law, the stress tensor components are related to the displacement vector as

$$\sigma_{ij} = \lambda\mathrm{div}\mathbf{U}\delta_{ij} + \mu\left(\frac{\partial U_i}{\partial x_j} + \frac{\partial U_j}{\partial x_i}\right). \tag{8}$$

The maximal breaking stress along the extension of the crack has the asymptotic

$$\sigma_y = \frac{K_{\mathrm{I}}(t)}{\sqrt{2\pi r}} + \mathrm{O}(1), \quad r \longrightarrow 0. \tag{9}$$

Let the crack edges be subjected to a linearly increasing impact load

$$p(t) = PU(t), \quad U(t) = tH(t), \tag{10}$$

where $H(t)$ is the Heaviside function and P is the loading rate.

Solving problem (4)–(7) by the Wiener–Hopf factorization method, we obtain the stress intensity factor in the form

$$K_{\mathrm{I}}(t) = P\varphi(c_1, c_2)Q(t), \tag{11}$$

where

$$Q(t) = \int_0^t U(s)\,f_0(t-s)ds, \quad f_0(t) = \frac{1}{2}t^{-\frac{1}{2}}, \tag{12}$$

and

$$\varphi(c_1, c_2) = \frac{4c_2\sqrt{(c_1^2 - c_2^2)}}{c_1\sqrt{\pi c_1}}$$

is the function of the velocities c_1 and c_2 of longitudinal and shear waves, respectively.

In our case, for a pulse given by (10), the function $Q(t)$ takes the form

$$Q(t) = \frac{2}{3}t^{\frac{3}{2}}H(t). \tag{13}$$

Let t_* be the time to failure. Substituting Eqs. (11) and (13) into criterion (3), we find the loading rate in terms of t_*:

$$P_* = \frac{15}{4}\frac{K_{\mathrm{Ic}}\tau}{\varphi(c_1, c_2)\left(t_*^{\frac{5}{2}} - (t_* - \tau)^{\frac{5}{2}}\right)}. \tag{14}$$

The critical value of the stress intensity factor $K_{\mathrm{Id}}(t)$ is the stress intensity factor at the instant of fracture. According to (11), it has the form

$$K_{\mathrm{Id}} = K_{\mathrm{I}}(t_*) = P_*\varphi(c_1, c_2)Q(t_*)$$

or

$$K_{\mathrm{Id}} = \frac{5}{2}\frac{t_*^{\frac{3}{2}}K_{\mathrm{Ic}}\tau}{t_*^{\frac{5}{2}} - (t_* - \tau)^{\frac{5}{2}}}. \tag{15}$$

This result will be used to analyze the experimental data in [1]. To estimate the incubation time to failure τ, we compare the data calculated by Eq. (15) with the experimental data. The result obtained with the structure–time criterion for the 2024-T3 alloy specimen

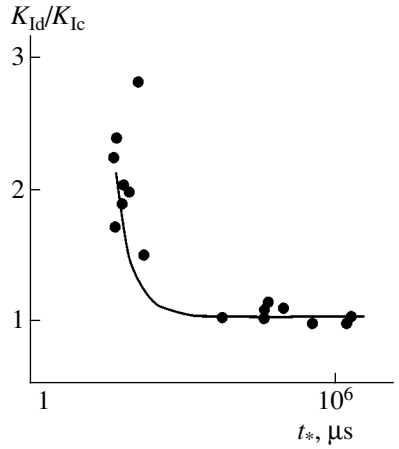

Fig. 1. Variation of the dynamic fracture toughness with time to failure t_*. Solid line, calculation with the structure–time criterion; circles, data points [1].

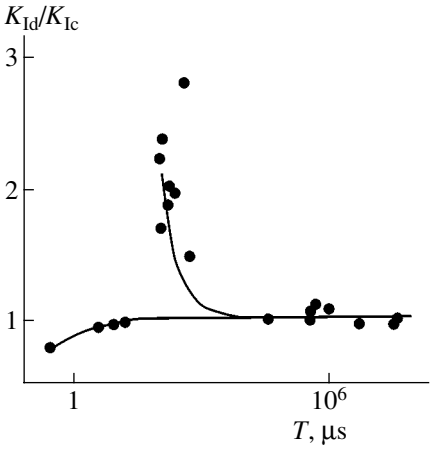

Fig. 2. Dependence of the dynamic fracture toughness of 2024-T3 alloy on the breaking pulse duration T. Upper curve, constant loading rates; lower curve, pulsed threshold loads.

(thickness $h = 2.54$ mm, $K_{Ic} \approx 30$ MPa m$^{1/2}$) is shown in Fig. 1. The incubation time τ found from this curve equals 40 μs.

In the problem under consideration, as well as in the experiments described in [1], the specimen failed at the growth stage of the stress intensity factor. However, the loading scheme where the dynamic fracture toughness is achieved at the stage of stress intensity factor decrease can be realized, as noted above [4]. Such a situation arises if we consider pulses of given duration and look for a threshold amplitude.

To study the dynamic fracture toughness at threshold loads, consider pulses $p(t) = PU(t)$ of various duration T, where

$$U(t) = \sin^2\left(\frac{\pi}{T}t\right)[H(t) - H(t-T)]. \qquad (16)$$

Using the solution to problem (4)–(7), we find the stress intensity factor in the form (11)–(12), where $U(t)$ is given by Eq. (16). Substituting (11) into criterion (3) yields an analytical expression for the threshold amplitude:

$$P_* = \frac{\tau K_{Ic}}{\varphi(c_1, c_2)\max\limits_t \int\limits_{t-\tau}^{t} Q(s)ds}. \qquad (17)$$

Here, P_* is the ultimate (threshold) pulse amplitude.

The time to failure t_* is the time over which

$$\max_t \int_{t-\tau}^{t} Q(s)ds$$

is reached.

Then, the expression for the dynamic fracture toughness $K_{Id} = K_I(t_*)$ takes the form

$$K_{Id} = \frac{\tau K_{Ic} Q(t_*)}{\max\limits_t \int\limits_{t-\tau}^{t} Q(s)ds}. \qquad (18)$$

By varying the pulse duration T, we find the time to failure t_* and the dynamic fracture toughness for each specific case.

Figure 2 compares the experimental behavior of the dynamic fracture toughness [1] with its behavior at threshold pulsed loads for this alloy. In the latter case, the dynamic fracture toughness decreases as the dura-

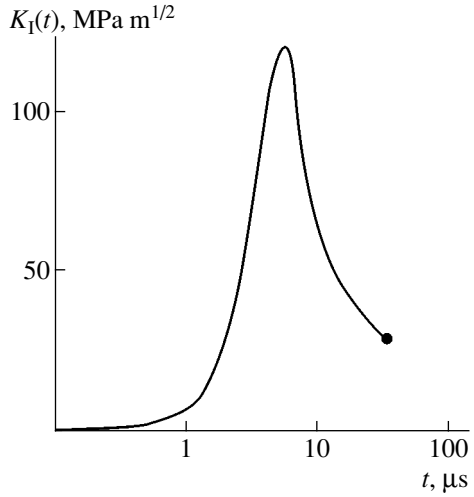

Fig. 3. Time dependence of the stress intensity factor for a threshold pulse of duration 8 μs. The dot indicates the time of fracture.

tion T of the threshold pulse approaches zero, whereas in [1] it rises without limit.

Figure 3 shows the stress intensity factor curve for a threshold pulse of duration 8 μs. The effect of "delayed fracture" is observed: the specimen fails when the stress intensity factor decreases after it has gone over a maximum. The same effect was predicted in [4] and observed experimentally in [5].

Thus, the structure–time criterion is an efficient means for studying problems of dynamic fracture. It allows one to predict the dynamic fracture toughness under various loading conditions. In the framework of the approach proposed, the critical value of the dynamic stress intensity factor is an estimable parameter. Its variation with loading rate depends on the loading history and experimental conditions.

Moreover, when describing the strength properties of a material under dynamic loading, one need not know the dependence of the dynamic fracture toughness on the loading rate. Two constants, K_{Ic} and τ, which characterize the strength properties of a material, will suffice to estimate ultimate loads.

REFERENCES

1. D. M. Owen, S. Z. Zhuang, A. J. Rosakis, *et al.*, Int. J. Fract. **90**, 153 (1998).
2. G. Irwin, J. Appl. Mech. **24**, 361 (1957).
3. Yu. V. Petrov, Preprint (Inst. for Problems of Mechanical Engineering, Russ. Acad. Sci., St. Petersburg, 1996), pp. 9–12.
4. D. A. Shockey, D. C. Erlich, J. F. Kalthoff, *et al.*, Eng. Fract. Mech. **23**, 311 (1986).
5. A. N. Berezkin, S. I. Krivosheev, Yu. V. Petrov, *et al.*, Dokl. Akad. Nauk **375**, 328 (2000) [Dokl. Phys. **45**, 617 (2000)].

Translated by K. Shakhlevich

Technical Physics, Vol. 49, No. 1, 2004, pp. 61–66. Translated from Zhurnal Tekhnicheskoĭ Fiziki, Vol. 74, No. 1, 2004, pp. 62–67.

SOLID-STATE ELECTRONICS

Prediction of the Properties of Two-Phase Composites with a Piezoactive Component

V. I. Aleshin, E. S. Tsikhotsky, and V. K. Yatsenko

Research Institute of Physics, Rostov State University, Rostov-on-Don, 344090 Russia

e-mail: aleshin@ip.rsu.ru

Received January 29, 2003

Abstract—The effective dielectric, piezoelectric, and elastic constants of two-phase macroscopically piezoactive 3–0 and 3–3 composites are calculated. It is assumed that one of the components is a polarized ferroelectric ceramic material and the other is an inactive material with variable elastic properties. The limiting case when the elastic compliances of the inactive material tend to infinity (porous ferroelectric ceramics) is considered. The adequacy of this model to production technologies of piezoelectric composites is discussed. Computational results are compared with experimental data. © *2004 MAIK "Nauka/Interperiodica".*

INTRODUCTION

The development of novel piezoactive composites needs reliable theoretical models allowing process engineers to predict the properties of the composites when the concentrations of the components vary in a wide range and their electromechanical parameters greatly differ. In a number of composite production technologies, initial powders are mixed in an appropriate proportion, pressed to a desired shape, and then sintered [1]. If the concentration of one component is sufficiently small, the final product will consist of a monolithic matrix, which is formed by the other (higher concentration) component, and isolated isometric inclusions. (Hereafter, we consider a two-phase composite assuming that the grain size in initial ceramic powders is much larger than the size of crystallites in the composite, so that each grain may be assigned the effective properties of associated ceramics. At the same time, the crystallite size must be much smaller than the characteristic size of the composite in order to provide the homogeneity of the entire system.) According to the classification currently adopted [2], such a composite is classified as a 3–0 composite. If the component concentration ratio in a composite prepared by the same technology is roughly 1/2, the components enter into the composite symmetrically, forming two infinite connected clusters. A composite thus prepared is referred to as a 2–3 composite.

Thus, the entire feasible range of concentrations may be covered and, hence, composites of various types can be produced by the same technology. Therefore, of great importance is the proper choice of a method for calculating the effective constants of a piezoelectric composite at different component concentrations. In other words, it is necessary to know how the effective constants of a composite depend on the component concentration for a given process of preparation. Unfortunately, exact solutions to three-dimensional problems of calculating the effective constants of inhomogeneous systems are unknown. As a result, a strict classification of composites by their structure is lacking.

The theory of heterogeneous systems subdivided two-phase composites into two large groups: matrix systems and binary mixtures. The concentration dependences of the effective constants in the two groups are much different.

In matrix systems, the variation of the concentration from 0 to 1 does not change the material structure qualitatively: at any concentration, one of the components forms a connected matrix that contains isolated inclusions of the second one. The system remains essentially asymmetric, and the formulas for effective constants yield their continuous concentration dependence throughout the 0–1 range. It should be noted that these formulas are, as a rule, used to calculate the effective constants of piezoactive composites, which is not universally true.

The case of binary systems has been considered above. These systems are known to exhibit critical concentration transitions: the metal–insulator transition (for properties described in terms of the second-rank tensor) or stiffness–compliance transition (for those described by the fourth-rank tensor). The transition of the first type means that, at some critical concentration m_1 ($0 < m_1 < 1$), the conducting component in the metal–insulator mixture forms an infinite cluster and the system as a whole becomes conductive. In the latter case, it is supposed that a composite is a mixture where the elastic compliances s_{ij} of one component tend to infinity (porous composite). The stiffness–compliance transition means that there exists some minimal (critical) concentration m_2($0 < m_2 < 1$) above which the rigid framework made up of the second component may lose

1063-7842/04/4901-0061$26.00 © 2004 MAIK "Nauka/Interperiodica"

stability. In other words, even if a composite with $m > m_2$ is prepared by an appropriate "mixing" technology, its properties are bound to depend significantly on many random factors related to the initial component distribution. The concentrations m_1 and m_2 do not necessarily coincide. In piezoactive materials, the transitions may be interrelated because of the piezoelectric effect.

In this work, we study the effective constants of binary piezoactive mixtures one component of which is a polarized ferroelectric ceramic material and the other is an inactive (passive) isotropic low-permittivity material with different elastic properties. In calculations, the effective medium method is used.

COMPUTATIONAL SCHEME

The effective medium method (or the self-consistency method) is based on the solution of the problem of interaction between a spherical inclusion and its environment. The properties of the environment are identified with the effective properties of the entire composite and are to be determined. Since of most practical interest are the properties of polarized piezoelectric composites, one should bear in mind that the medium is anisotropic in dielectric and elastic properties and shows the macroscopic piezoelectric effect. The problem of interaction between a spherical inclusion and an anisotropic piezoactive medium is reduced [3–5] to the solution of a set of linear equations that relate electric and elastic fields inside and outside (away from) the inclusion:

$$\begin{aligned} E_i - E_{0i} &= A_{ii}(D_{0i} - D_i) + H_{in}(u_{0n} - u_n), \\ \sigma_i - \sigma_{0i} &= H_{ji}(D_{0j} - D_j)B_{in}(u_{0n} - u_n). \end{aligned} \tag{1}$$

Here, E is the electric field, D is the induction, σ is the mechanical stress, and u is the strain. The variables with the subscript 0 refer to the medium, i.e., represent the volume-averaged components of the electric and elastic fields. The coefficients A, B, and H depend only on the properties of the medium and are independent of the properties of the inclusion. Expressions for these coefficients in the axisymmetric case are given in [3, 5]. Equations (1) do not impose any restrictions on the relationship between the electric and elastic fields inside the inclusion: they merely reflect the elastic (linear) character of inclusion–medium interaction.

To implement the self-consistency scheme, set (1) must be solved jointly with one of the sets of equations for the piezoelectric effect in the piezoactive component,

$$\begin{aligned} E_i &= \eta_{ij}^u D_j - h_{in}u_n, \\ \sigma_i &= -h_{ji}D_j + c_{in}^0 u_n, \end{aligned} \tag{2}$$

and with a similar set of equations for the passive (nonpiezoactive) component,

$$E_i = \eta D_i, \quad \sigma_i = c_{in}u_n. \tag{3}$$

In Eqs. (2) and (3), η_{ij} are the components of the reciprocal permittivity, c_{ij} are the stiffness coefficients, and h_{ij} are the piezoelectric constants.

By separately substituting (2) and (3) into (1) and averaging the resultant expressions for D_i and u_i over volume in the standard way, we arrive at a set of equations for the effective constants of the piezoactive composite:

$$\begin{aligned} \varepsilon_{11}^{*\sigma} &= \Delta_{11}X_1 + \Delta_{72}(H_{15}\varepsilon_{11}^{*\sigma} + B_{44}d_{15}^*), \\ \varepsilon_{33}^{*\sigma} &= \Delta_{33}X_3 + 2\Delta_{43}P_1 + \Delta_{63}P_3, \\ s_{11}^{*E} &= \Delta_{43}T_4 + \Delta_{44}(1 + T_1) + \Delta_{54}T_2 + \Delta_{64}T_3, \\ s_{12}^{*E} &= \Delta_{43}T_4 + \Delta_{44}T_2 + \Delta_{54}(1 + T_1) + \Delta_{64}T_3, \\ s_{13}^{*E} &= \Delta_{43}T_z + (\Delta_{44} + \Delta_{54})T_x + \Delta_{64}(1 + T_y), \\ s_{33}^{*E} &= \Delta_{63}T_z + 2\Delta_{64}T_x + \Delta_{66}(1 + T_y), \\ s_{44}^{*E} &= \Delta_{72}(A_{11}d_{15}^* + H_{15}s_{44}^{*E}) + \Delta_{77}(1 + H_{15}d_{15}^* + B_{44}s_{44}^{*E}), \\ d_{31}^* &= \Delta_{43}X_3 + (\Delta_{44} + \Delta_{54})P_1 + \Delta_{64}P_3, \\ d_{33}^* &= \Delta_{63}X_3 + 2\Delta_{64}P_1 + \Delta_{66}P_3, \\ d_{15}^* &= \Delta_{72}X_1 + \Delta_{77}(H_{15}\varepsilon_{11}^{*\sigma} + B_{44}d_{15}^*). \end{aligned} \tag{4}$$

Here,

$$\begin{aligned} X_1 &= 1 + A_{11}\varepsilon_{11}^{*\sigma} + H_{15}d_{15}^*, \\ X_3 &= 1 + A_{33}\varepsilon_{33}^{*\sigma} + 2H_{31}d_{31}^* + H_{33}d_{33}^*, \\ P_1 &= H_{31}\varepsilon_{33}^{*\sigma} + (B_{11} + B_{12})d_{31}^* + B_{13}d_{33}^*, \\ P_3 &= H_{33}\varepsilon_{33}^{*\sigma} + 2B_{14}d_{31}^* + B_{33}d_{33}^*, \\ T_1 &= H_{31}d_{31}^* + B_{11}s_{11}^{*E} + B_{12}s_{12}^{*E} + B_{13}s_{13}^{*E}, \\ T_2 &= H_{31}d_{31}^* + B_{12}s_{11}^{*E} + B_{11}s_{12}^{*E} + B_{13}s_{13}^{*E}, \\ T_3 &= H_{33}d_{31}^* + B_{13}(s_{11}^{*E} + s_{12}^{*E}) + B_{33}s_{13}^{*E}, \\ T_4 &= A_{33}d_{31}^* + H_{13}(s_{11}^{*E} + s_{12}^{*E}) + H_{33}s_{13}^{*E}, \\ T_x &= H_{31}d_{33}^* + (B_{11} + B_{12})s_{13}^{*E} + B_{13}s_{33}^{*E}, \\ T_y &= H_{33}d_{33}^* + 2B_{13}s_{13}^{*E} + B_{33}s_{33}^{*E}, \\ T_z &= A_{33}d_{33}^* + 2H_{31}s_{13}^{*E} + H_{33}s_{33}^{*E}. \end{aligned}$$

Table 1. Experimentally found constants $\varepsilon_{ij}^{\sigma}/\varepsilon_0$, d_{ij}(10^{-12} C/N), and s_{ij}^{E} (10^{-12} m^2/N) of PZT-5H [6]

$\varepsilon_{11}^{\sigma}/\varepsilon_0$	$\varepsilon_{33}^{\sigma}/\varepsilon_0$	d_{31}	d_{33}	d_{15}	s_{11}^{E}	s_{12}^{E}	s_{13}^{E}	s_{33}^{E}	s_{44}^{E}
3130	3400	–274	593	741	16.5	–4.78	–8.45	20.7	43.5

In (4), $s_{ij}^{*\sigma}$ are the components of the effective permittivity, s_{ij}^{*E} are the elastic compliances, and d_{ij}^{*} are the piezoelectric moduli of the composite. The matrix Δ (9×9), appearing in Eqs. (4), can be written in general form as follows:

$$\mathbf{\Delta} = (1-m)\begin{pmatrix} \boldsymbol{\eta}^{u} + \mathbf{A} & -\mathbf{h} + \mathbf{H} \\ -\mathbf{h}_t + \mathbf{H}_t & \mathbf{c}^{D} + \mathbf{B} \end{pmatrix}^{-1} + m\begin{pmatrix} \boldsymbol{\varepsilon} + \mathbf{A} & \mathbf{H} \\ \mathbf{H}_t & \mathbf{c} + \mathbf{B} \end{pmatrix}^{-1}. \quad (5)$$

The first and second terms in (5) correspond to the piezoactive and passive components, respectively; m is the concentration of the passive component; and the subscript t marks the transpose.

Self-consistent system of equations (4) is closed and allows one to calculate the complete set of effective constants for a piezoactive composite. The constituents of the composite enter into Eqs. (4) symmetrically, as follows, for example, from relationship (5). It is also assumed that the piezoactive component is homogeneous and polarized and also that its constants do not depend on its concentration (that is, the ferroelectric ceramic in the composite is equally polarized at any concentration).

RESULTS AND DISCUSSION

Calculations were performed for three cases where polarized PZT-5H ceramics was used as a piezoactive component and isotropic materials with greatly differing elastic properties were used as a passive component. These materials were (i) SiO_2, which has elastic properties of the same order as the ferroelectric ceramics, (ii) a formaldehyde-based polymer with elastic compliances one order of magnitude higher than those of the ferroelectric ceramics, and (iii) air bubbles as the limiting case (the material with an infinite compliance). The permittivity of the three passive components was small (~1). The constants of polarized PZT-5H were taken from [6]; the elastic and dielectric constants of SiO_2, from [7]. The constants of the polymer were measured with an original device described elsewhere [8]. All constants found experimentally and used in calculations are listed in Tables 1 and 2.

Figure 1 shows the calculation results for case (i). As follows from Fig. 1b, stiff SiO_2 inclusions in the piezoactive matrix suppress the piezoactive properties of the composite. The piezoelectric moduli of the material decrease nearly twofold at a SiO_2 concentration of 0.25 and vanish at a concentration of about 0.7. In the same concentration interval (from 0 to 0.7), the permittivity also drops significantly (Fig. 1a). As was noted above, such behavior is associated with the metal–insulator critical transition. It is known that the effective medium method gives the critical concentration $m_1 = 1/3$ for transitions of this type. In our case, this means that the piezoelectric composite loses piezoelectric properties if the concentration m of the passive component exceeds 2/3. The behavior of the elastic constants suggests that the material becomes almost isotropic in elastic properties at a concentration of 0.5 (Fig. 1c) with the piezoelectric modulus d_{33}^{*} remaining at a level of 100 pC/N, which is typical of several new piezoceramic compositions. Thus, by introducing appropriate admixtures into piezoceramic materials, one can vary their properties over a wide range and prepare composites of practical interest.

With less stiff polymeric inclusions present in the piezoactive matrix, the decrease in the piezoactive properties is not so drastic (Fig. 2b). They disappear, as before, at $m > 2/3$, but the sharp drop of the piezoelectric moduli d_{33}^{*} and d_{15}^{*} starts only with $m \approx 0.4$. It is noteworthy that, in the range $0 < m < 0.4$, where the d_{33}^{*} and d_{15}^{*} decline very smoothly, the elastic stiffness of the composite drops sharply (Fig. 2c). The components of the permittivity decrease almost linearly in the interval $0 < m < 2/3$ (Fig. 2a).

Having discussed this intermediate case, we pass to the limiting situation when the piezoactive matrix contains infinite-compliance (zero-stiffness) inclusions, i.e., air bubbles. In this case, we must put $\mathbf{c} = 0$ on the right of (5). The computational results for this case are shown in Fig. 3.

Note first of all that the stiffness coefficients of the composite (Fig. 3c) vanish at $m_2 = 0.5$. This point marks the critical stiffness–compliance transition. With $m > m_2$, A solution to the set of Eqs. (4) loses stability. Thus, the effective medium method predicts that a porous composite prepared by a mixing technology (the burn-

Table 2. Experimentally found constants $\varepsilon/\varepsilon_0$ and s_{ij}(10^{-12} m^2/N) of SiO_2 [7] and polymeric passive components

	$\varepsilon/\varepsilon_0$	s_{11}	s_{12}
SiO_2	4.5	10.5	–0.801
Polymer	4.3	255.1	–102.0

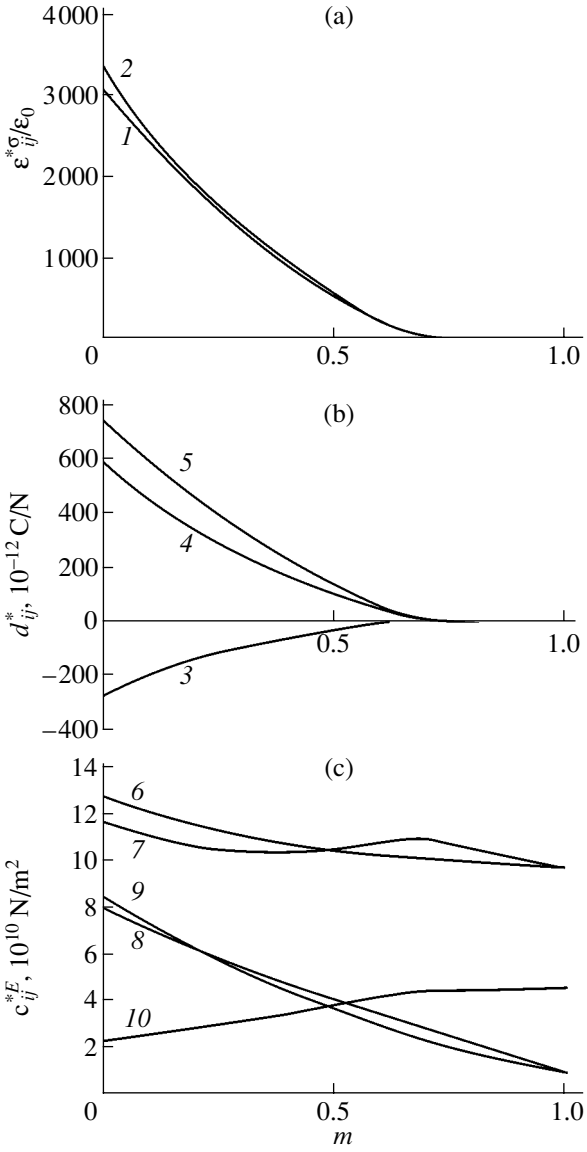

Fig. 1. Effective dielectric, piezoelectric, and elastic constants of the piezoactive composite vs. passive component (SiO_2) concentration. (*1*) $\varepsilon^{*\sigma}_{11}$, (*2*) $\varepsilon^{*\sigma}_{33}$, (*3*) d^*_{31}, (*4*) d^*_{33}, (*5*) d^*_{15}, (*6*) c^{*E}_{11}, (*7*) c^{*E}_{33}, (*8*) c^{*E}_{12}, (*9*) c^{*E}_{13}, and (*10*) c^{*E}_{44}.

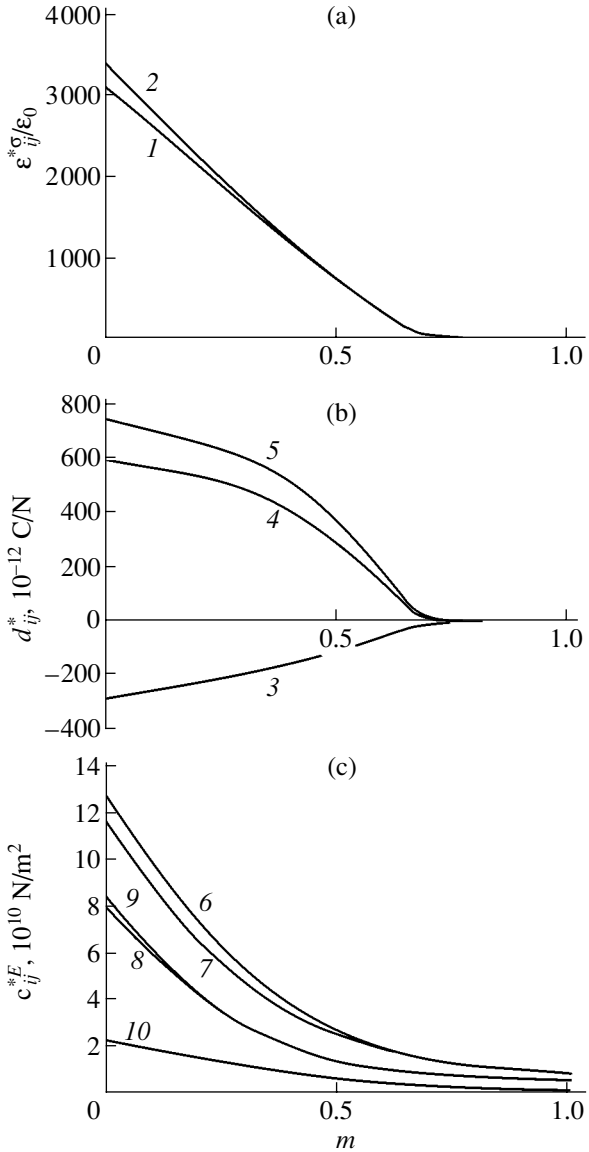

Fig. 2. Effective dielectric, piezoelectric, and elastic constants of the piezoactive composite vs. passive component (polymer) concentration. (*1*) $\varepsilon^{*\sigma}_{11}$, (*2*) $\varepsilon^{*\sigma}_{33}$, (*3*) d^*_{31}, (*4*) d^*_{33}, (*5*) d^*_{15}, (*6*) c^{*E}_{11}, (*7*) c^{*E}_{33}, (*8*) c^{*E}_{12}, (*9*) c^{*E}_{13}, and (*10*) c^{*E}_{44}.

ing-out of polymeric granules [1, 9], etc.) will have stable properties only if the bubble (pore) concentration is less than $m_2 = 0.5$. As regards this value of m_2, it should be noted that the effective medium method is approximate: at any rate, it does not give an exact geometrical distribution of the components in the system. Therefore, to refine the critical value of m_2 that manifests the stiffness–compliance transition, it would be reasonable to invoke numerical models similar to those used in the percolation theory to study metal–insulator transitions. Such analysis would also be of interest for asymmetric systems (such as a polymeric matrix with stiff inclusions), because it is impossible to arbitrarily increase the concentration of inclusions that have an arbitrary (isometric) shape to unity: at some $m < 1$, such inclusions necessarily form an infinite cluster. Thus, most matrix systems, as well as mixtures, are bound to exhibit critical concentrations.

Figure 3b shows the dependence of the piezoelectric moduli of the composite on the bubble concentration. The moduli d^*_{33} and d^*_{15} remain practically constant throughout the concentration range (the modulus d^*_{33}

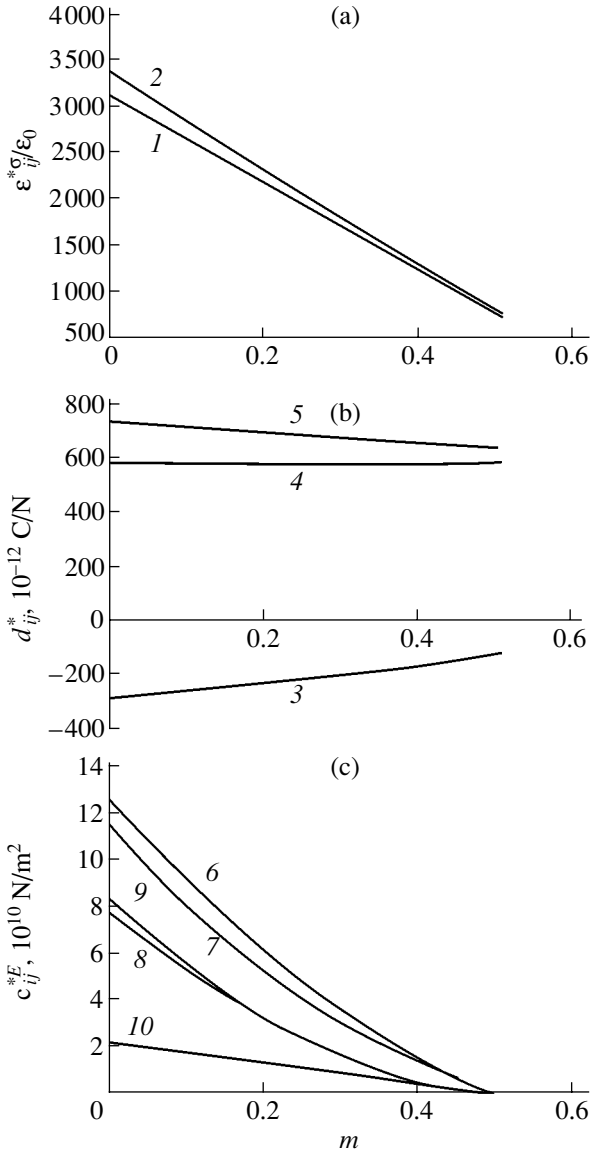

Fig. 3. Effective dielectric, piezoelectric, and elastic constants of the piezoactive composite vs. passive component (polymer) concentration. (*1*) $\varepsilon^{*\sigma}_{11}$, (*2*) $\varepsilon^{*\sigma}_{33}$, (*3*) d^*_{31}, (*4*) d^*_{33}, (*5*) d^*_{15}, (*6*) c^{*E}_{11}, (*7*) c^{*E}_{33}, (*8*) c^{*E}_{12}, (*9*) c^{*E}_{13}, and (*10*) c^{*E}_{44}.

even slightly increases). Such behavior may be explained by a fast growth of the compliances s^{*E}_{33} and s^{*E}_{44}. The monotonic decrease in the piezoelectric modulus d^*_{31} (at $m = 0.5$, it is less than half the starting value) may be associated with similar behavior of the Poisson's ratio $-s^{*E}_{13}/s^{*E}_{33}$, which also decreases more than twofold.

To conclude, our results are in good agreement with experimental data for porous ceramics [9–11]. Poor agreement of estimates [11] made by the method described in [12] with experiments is due to the fact that the anisotropy and piezoelectric activity of the medium were not taken into account adequately.

CONCLUSIONS

The analysis of the effective properties of two-phase piezoactive composites shows that the introduction of passive stiff low-permittivity admixtures even in small amounts (≈25%) suppresses markedly the piezoelectric activity of ferroelectric ceramics. Moreover, at a passive component concentration of ≈0.7, the composite totally loses piezoelectric properties. As was shown above, such behavior is associated with the metal–insulator critical transition.

If the passive component is of infinite elastic compliance (porous ferroelectric ceramics), the piezoelectric moduli d^*_{33} and d^*_{15} remain practically constant (the modulus d^*_{33} may even slightly increase) as its concentration grows. This may be related to the significant growth of the elastic compliance components s^{*E}_{33} and s^{*E}_{44}, which is likely to compensate for a decrease in the fraction of the piezoactive component in the system. The monotonic decrease in the piezoelectric modulus d^*_{31} may be associated with similar behavior of the Poisson's ratio $-s^{*E}_{13}/s^{*E}_{33}$.

At a certain pore concentration, porous ceramics exhibit the stiffness–compliance critical transition. The effective medium method, which was used in calculations, predicts this transition at $m_2 = 0.5$. This means that most present-day technologies can provide stable performance of porous ceramics only at pore concentrations $m < m_2$ (≈0.5).

The results obtained are consistent with available experimental data for porous ferroelectric ceramics.

ACKNOWLEDGMENTS

This work was supported by the Russian Foundation for Basic Research (grant no. 02-03-08057-inno).

REFERENCES

1. W. Wersing, K. Lubitz, and J. Mohaupt, Ferroelectrics **68** (1/4), 77 (1986).
2. R. E. Newnham, D. P. Skinner, and L. E. Cross, Mater. Res. Bull. **13**, 525 (1978).
3. V. I. Aleshin, *The Influence of the Piezoelectric Effect on the Properties of Anisotropic Crystals*, Available from VINITI, 1988, Moscow, No. 3902-B88.

4. V. I. Aleshin, Kristallografiya **36**, 1352 (1991) [Sov. Phys. Crystallogr. **36**, 766 (1991)].
5. V. I. Aleshin, J. Appl. Phys. **88**, 3587 (2000).
6. eFunda Inc., http://www.efunda.com/materials/piezo/material_data/matdata_index.cfm.
7. I. N. Frantsevich, F. F. Voronov, and S. A. Bakuta, *Elastic Constants and Elastic Moduli of Metals and Nonmetals: A Handbook* (Naukova Dumka, Kiev, 1982).
8. I. N. Moshchenko, A. N. Sadkov, V. K. Yatsenko, and M. I. Novgorodova, Izv. Vyssh. Uchebn. Zaved. Sev.-Kavk. Reg., Estestv. Nauki, No. 4, 56 (1997).
9. S. S. Lopatin and T. G. Lupeĭko, Izv. Akad. Nauk SSSR, Neorg. Mater. **27**, 1948 (1991).
10. I. Getman and S. Lopatin, Ferroelectrics **186**, 301 (1996).
11. A. N. Rybyanets, in *Proceedings of the International Conference on the Fundamental Problems of Piezoelectric Instrument Making, Tver', 2002*, pp. 48–72.
12. M. Marutake and T. Ikeda, J. Phys. Soc. Jpn. **11**, 814 (1956).

Translated by V. Isaakyan

Technical Physics, Vol. 49, No. 1, 2004, pp. 67–70. Translated from Zhurnal Tekhnicheskoĭ Fiziki, Vol. 74, No. 1, 2004, pp. 68–71.

OPTICS, QUANTUM ELECTRONICS

Investigation of Protein Mass Transfer by Real-Time Holographic Interferometry

N. M. Ganzherli, I. A. Maurer, and P. V. Granskiy

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

e-mail: nina@holo.ioffe.rssi.ru

Received July 3, 2003

Abstract—Based on holographic interferometry data for electrophoresis and diffusion, an algorithm for quantitatively analyzing protein mass transfer in gels is suggested and implemented. The diffusion coefficients of a number of proteins in different media are found. © *2004 MAIK "Nauka/Interperiodica".*

INTRODUCTION

Mass transfer causes phase inhomogeneities, which may be studied with optical methods, including holographic interferometry. The application of this method has greatly extended the potentialities of optical experiment for phase inhomogeneity visualization [1, 2].

Earlier [3], the efficiency of holographic interferometry in studying the electrophoresis kinetics of biological preparations in gels was demonstrated. The use of real-time holographic interferometry for electrophoresis visualization allows researchers to examine the process directly in the working zone of an electrophoresis column without adding dyes. Also, this method makes it possible to trace the separation and motion of protein fractions, control the shape of the protein fraction front, etc.

QUANTITATIVE ANALYSIS OF PROTEIN ELECTROPHORETIC SEPARATION IN GELS

In this work, we tried a technique for quantitatively analyzing interferograms that reflect the electrophoresis and diffusion of biological preparations in gels. Experiments were carried out with human donor albumin, which makes up almost 60% of the total mass of the blood plasma, under conditions of continuous electrophoresis in 7% polyacrylamide gel. This gel has proved efficient in studies of the protein electrophoretic separation kinetics. Since the system is homogeneous in this case, the heating of the gel is insignificant and the system is stable. A slight uniform heating and electric-field-induced ion redistribution in the buffer system show up as an inclination of finite-width fringes relative to their initial position; however, this inclination can be taken into account upon interferogram identification.

We used a 100-mm-high 35-mm-wide vertical electrophoresis column with polyacrylamide gel of thickness 7 mm. The 7×30-mm cross-sectional area of the column and gel allowed us to introduce a large volume of samples, which was significant for the analysis of low-concentration preparations. Under such conditions, proteins were uniformly distributed over the gel's cross section and enter the gel simultaneously.

During electrophoresis, the protein fractions concentrate and form narrow regions. When a protein fraction of mass m is localized in the working zone of the column, the interference pattern represents a family of finite-width fringes (in Fig. 1, the fringes are schematically shown as curves). The interferogram can be described through the number N of finite-width fringes (curves), the function $f(x)$ of distortion for each of the fringes, and the area S under the interference curve.

It is reasonable to assume that the area S, the mass m of the protein localized in a given zone, and the number N of finite-width fringes are interrelated. Finding this relationship is basic to the quantitative analysis of protein mixture electrophoresis.

It turns out that the distortion of fringes due to protein fraction localization and the area under the interference curve depend on the total amount m of protein rather than on the protein concentration in a sample under test (Fig. 2). For a fixed number N of fringes on

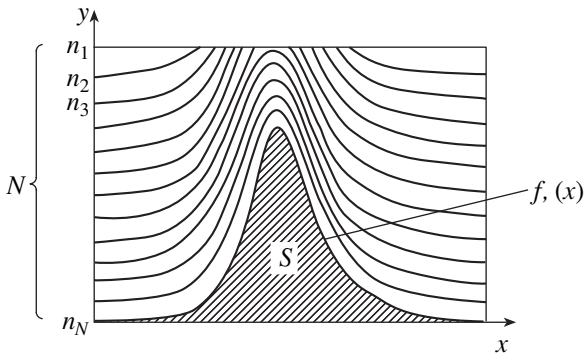

Fig. 1. Schematic representation of finite-width interferometric fringes taken from protein electrophoresis.

1063-7842/04/4901-0067$26.00 © 2004 MAIK "Nauka/Interperiodica"

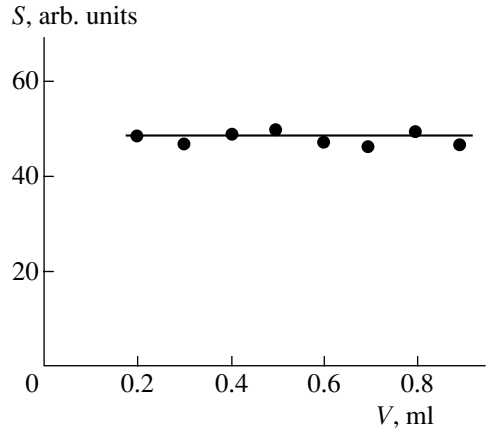

Fig. 2. Area *S* under the interferometric curve vs. volume *V* of protein with the protein mass *m* remaining unchanged.

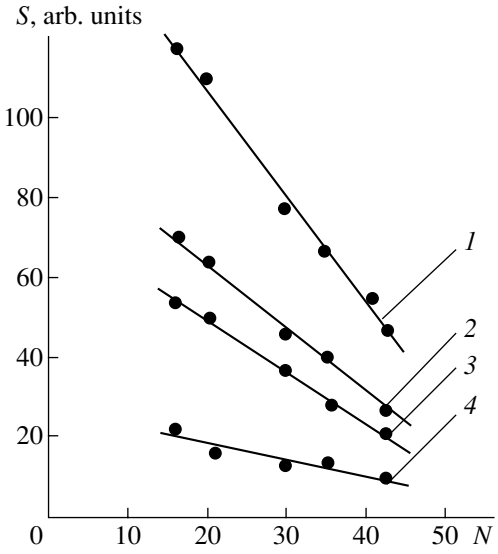

Fig. 3. Area *S* under the interferometric curve vs. the number *N* of finite-width fringes on the interferogram for the protein mass *m* = (*1*) 8, (*2*) 5, (*3*) 4, and (*4*) 0.5 mg.

the interferogram and for the same amount *m* of protein, the value of *S* varies with the protein volume *V* in the electrophoresis column only slightly.

We found the dependence of the area *S* under the curve on the number *N* of fringes for various amounts *m* of albumin introduced into the column (Fig. 3).

Furthermore, we experimentally substantiated the assumption that the shape of the interference curve $f(x)$ depends on an electrophoretic system, the electric field strength in an electrophoresis column, the gel porosity, etc., and remains unchanged if the basic parameters of electrophoresis are kept fixed. Also, it was shown that, when the amount *m* of protein is varied, the area *S* under the curve varies proportionally but the shape of the curve $f(x)$ remains constant. This fact allows one to construct an algorithm for separating out partial profiles from the overall complex profile $f(x)$ when proteins are insufficiently separated during electrophoresis.

QUANTITATIVE ANALYSIS OF PROTEIN DIFFUSION IN GELS

Diffusion, spontaneous transfer of a material between various parts of a system within one phase, depends on temperature, external field, the physical structure of biological preparations, and concentration gradient. Diffusion causes the smearing of protein zones and, consequently, degrades the resolving power of electrophoretic devices. The use of holographic interferometry for studying the diffusion of various materials and material mixtures is exemplified in [4–7].

Below, we give a technique for determining the diffusion coefficients of proteins in gels. Molecular diffusion, which occurs in the presence of a concentration gradient, is considered.

To determine the diffusion coefficient for albumin, a certain amount of protein was introduced in polyacrylamide gel. Electrophoresis was examined with real-time holographic interferometry. After the protein zone had been formed, the interferogram of the initial stage of diffusion was taken. Then, the electrophoresis current was switched off and the macroscopic directed diffusion of the protein in the gel, which was induced by the concentration gradient, started.

The technique for finding diffusion coefficients is elucidated in Fig. 4, which shows two distortions, $f_1(x)$ and $f_2(x)$, of the fringes in the zone where the protein of mass *m* is localized at the initial, t_1, and final, t_2, times.

The diffusion coefficient is found from Fick's first law

$$L = -D\frac{dc}{dx}, \tag{1}$$

where *L* is the protein flux through a unit cross section in a unit time in the direction of the concentration gradient, *D* is the diffusion coefficient, and dc/dx is the protein concentration gradient in the *x* direction.

Let us integrate both sides of (1) with respect to time:

$$\int_{t_1}^{t_2} L\,dt = -D\int_{t_1}^{t_2}\frac{\partial c}{\partial x}dt. \tag{2}$$

The integral on the right can be written as

$$\int_{t_1}^{t_2}\frac{\partial c}{\partial x}dt = \overline{\frac{\partial c}{\partial x}}\Delta t = \overline{\frac{\partial c}{\partial x}}(t_2 - t_1), \tag{3}$$

where $\overline{\partial c/\partial x}$ is the time-averaged derivative of $c(x)$ at a given point x_0 in the $t_1 - t_2$ time interval.

This point is selected so that the diffusate at this point is absent at the time t_1. To estimate $\overline{\partial c/\partial x}$, we use the approximation

$$\overline{\frac{\partial c}{\partial x}} \approx \frac{1}{2}\frac{\partial c(x_0, t_2)}{\partial x}. \quad (4)$$

Consider the left of equality (2):

$$\int_{t_1}^{t_2} L\,dt = \frac{1}{s}\int_{t_1}^{t_2}\frac{\partial m}{\partial t}dt = \frac{1}{s}(m_2 - m_1), \quad (5)$$

where m_1 and m_2 are the masses of the material passing through a section s at the point x_0 by the times t_1 and t_2, respectively.

For the point x_0 thus selected, $m_1 = 0$ and

$$m_2 = \int_{x_0}^{\infty} sc(x, t_2)dx. \quad (6)$$

Using these intermediate expressions, we recast Eq. (1) in the form

$$\int_{x_0}^{\infty} c(x, t_2)dx = -D\frac{1}{2}\frac{\partial c(x_0, t_2)}{\partial x}\Delta t. \quad (7)$$

Assuming that the distortion function $f(x)$ of the fringes is proportional to the concentration $c(x)$, we determine the area of the figure ABx_0 and the slope of the tangent to the curve $f_2(x)$ at the point x_0. Eventually, we find that the diffusion coefficient equals 1.2×10^{-4} cm^2/s.

As follows from the experiments, the concentration of the polyacrylamide gel as a supporting medium cannot be below 3%. Otherwise, the gel starts rapidly decomposing. We determined the diffusion coefficients for various concentrations of the gel. The diffusion coefficient of albumin vs. the gel concentration is plotted in Fig. 5.

The diffusion coefficient may also be found from Fick's second law

$$\frac{\partial c}{\partial t} = D\frac{\partial^2 c}{\partial x^2}, \quad (8)$$

which defines the rate of concentration variation due to diffusion. Here, D is the diffusion coefficient, t is the diffusion time, and x is the coordinate along which the protein diffuses.

The distortion of the fringe can be approximated with a high accuracy by a Gaussian function. During diffusion, the Gaussian curve expands. Assuming that the concentration curve $c(x)$ varies in proportion to the curve $f(x)$, we may represent $c(x)$ in Gaussian form:

$$c(x) = \frac{A}{b\sqrt{2\pi}}\exp\left[-\frac{1}{2}(x/b)^2\right], \quad (9)$$

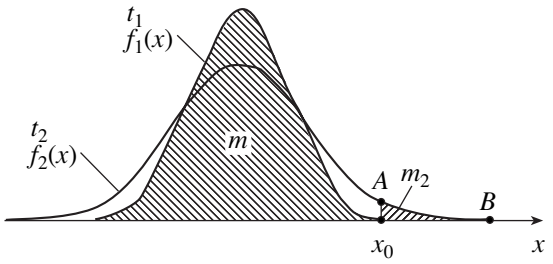

Fig. 4. Shape of the interferometric lines $f_1(x)$ and $f_2(x)$ for the initial, t_1, and final, t_2, times of diffusion.

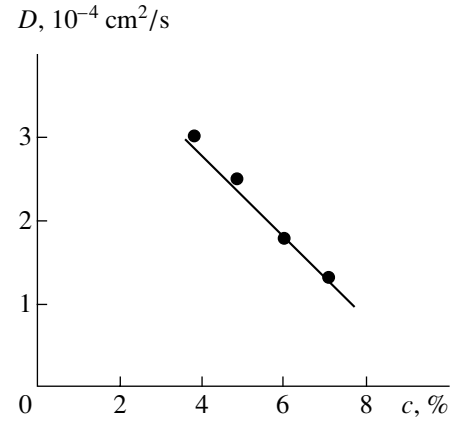

Fig. 5. Diffusion coefficient of albumin vs. gel concentration.

where A is a normalizing factor and b is a time-dependent parameter that characterizes the width of the distribution.

During diffusion, the parameter b grows; that is, the distribution expands. However, the mass of the protein introduced into the column and the area under the curve $c(x)$, which is proportional to the mass, remain the same; therefore, we may write

$$\int_{-\infty}^{+\infty} c(x)dx = A. \quad (10)$$

The second derivative of $c(x)$ with respect to x and the first time derivative of $c(x)$ are given by

$$\frac{\partial^2 c}{\partial x^2} = \frac{1}{b^4}(x^2 - b^2)c(x) = \frac{A}{b^5\sqrt{2\pi}}(x^2 - b^2)\exp\left[-\frac{1}{2}(x/b)^2\right], \quad (11)$$

$$\frac{\partial c}{\partial t} = \frac{1}{b^3}(x^2 - b^2)\frac{\partial b}{\partial t}c(x) = \frac{A}{b^4\sqrt{2\pi}}(x^2 - b^2)\exp\left[-\frac{1}{2}(x/b)^2\right]\frac{\partial b}{\partial t}. \quad (12)$$

Substituting formulas (11) and (12) into Eq. (8) yields an expression for the diffusion coefficient:

$$D = b\frac{\partial b}{\partial t}. \quad (13)$$

Let it seem reasonable to interpolate a set of data points y_k for the curve $c(x)$ by formula (9). The parameters b and A may be found by the least-squares method. At the same time, simple expression (13) for diffusion coefficient D, which was derived from Fick's second law, makes it possible to directly find D by measuring b and $\partial b/\partial t$. From (9), it follows that b is the half-width of the distribution $c(x)$ at a level of 0.6 of the peak. We may replace the quantity $\partial b/\partial t$ by $\Delta b/\Delta t$ and determine small variations Δb for short time intervals Δt. Such an approach to finding the diffusion coefficient (which is based on Fick's second law) yields $D = 1.6 \times 10^{-4}$ cm^2/s.

Thus, the two approaches give different values of the diffusion coefficient. This is because uncertainty in the position of the baseline for the curves $f(x)$ (Fig. 4) increases the calculation error and, accordingly, a scatter in the area of the figure ABx_0 and coefficient D. At the same time, this uncertainty influences b to a lesser extent.

CONCLUSIONS

We showed that real-time holographic interferometry may be useful for quantitative analysis of protein mass transfer in gels and for finding the diffusion coefficient. This method is illustrative and well compatible with modern data processing techniques. The technique for diffusion coefficient determination may be applied to study the diffusion of different protein components and makes it possible to find the diffusion coefficient under variable electrophoresis conditions.

ACKNOWLEDGMENTS

This work was financially supported by the Russian Foundation for Basic Research (grant no. 01-02-17854) and by the Scientific School (grant no. NSh-98.2003.2).

REFERENCES

1. Yu. I. Ostrovskiĭ, M. M. Butusov, and G. V. Ostrovskaya, *Holographic Interferometry* (Nauka, Moscow, 1977).
2. A. K. Beketova, A. F. Belozerov, A. N. Berezkin, *et al.*, *Holographic Interferometry of Phase Objects* (Nauka, Leningrad, 1979).
3. V. L. Barannikov, N. M. Ganzherli, S. B. Gurevich, *et al.*, Pis'ma Zh. Tekh. Fiz. **9**, 659 (1983) [Sov. Tech. Phys. Lett. **9**, 283 (1983)].
4. J. Pipman, S. Lipson, and J. Landam, Laser Elektro-Optik **8**, 24 (1976).
5. K. A. Stetson and G. Becsey, J. Opt. Soc. Am. **55**, 1694 (1965).
6. D. Paolleti and G. Schirrippa, Opt. Eng. **27**, 486 (1988).
7. A. V. Veniaminov, G. I. Lashkov, O. B. Ratner, *et al.*, Opt. Spektrosk. **60**, 142 (1986) [Opt. Spectrosc. **60**, 87 (1986)].

Translated by V. Isaakyan

Technical Physics, Vol. 49, No. 1, 2004, pp. 71–75. Translated from Zhurnal Tekhnicheskoĭ Fiziki, Vol. 74, No. 1, 2004, pp. 72–76.

**OPTICS,
QUANTUM ELECTRONICS**

Modulation of the Phase Difference between Polarization Modes in Single-Mode Fibers

O. I. Kotov, L. B. Liokumovich, S. I. Markov, A. V. Medvedev, and A. V. Khlybov

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

e-mail: kotov@radio.rphf.spbstu.ru

Received July 3, 2003

Abstract—Mechanisms of polarization modulation in a single-mode fiber that modulate the phase difference between polarization modes without affecting their amplitudes are considered. A coefficient that characterizes the efficiency of cylindrical piezoceramic modulators and is independent of their resonant properties is introduced. Analytical expressions for this coefficient for different modulation mechanisms are derived. The lateral pressure on the fiber is shown to provide the highest efficiency. For isotropic fibers, a modulator with a squeezing covering, which increases significantly its efficiency, is studied. For anisotropic fibers, the most appropriate way of phase difference modulation is longitudinal extension, in which case the birefringence axes do not have to be matched. In most cases, the measured and predicted efficiencies are in good correspondence. © *2004 MAIK "Nauka/Interperiodica".*

INTRODUCTION

In single-mode fibers with linear birefringence, polarization of light is analyzed in terms of a superposition of two orthogonal linearly polarized modes [1, 2]. In the general case, polarization at the exit from the fiber can be modulated by varying the amplitudes and phases of the polarization modes. In certain applications, it is, however, necessary to modulate the phase difference while keeping the amplitude ratio constant [3, 4]. In this paper, we accomplish such a modulation mechanically, extending and squeezing the fiber and also changing its bend radius. Figure 1 shows typical modulators subjected to these actions (the light guide is marked gray, the hatched areas are immovable parts, the transducer element that produces a varying mechanical effect is shown black, and vertical hatching indicates an adhesive). The analysis which follows relies on published data [1, 5–9] and additional calculations and experiments performed by the authors.

THEORETICAL ANALYSIS

Birefringence in single-mode fibers is described by the difference in propagation constants for orthogonal polarization modes E_x and E_y [1, 10]:

$$\beta = \beta_x - \beta_y = \frac{2\pi}{\lambda}(n_x - n_y) = \frac{2\pi}{\Lambda}, \qquad (1)$$

where λ is the optical wavelength in free space, n_x and n_y are the effective refractive indices of the core for the orthogonally polarized modes, and Λ is the polarization beat length.

The orthogonal directions of polarization of polarization eigenmodes are often referred to as the birefringence axes of a fiber. The modulation of the phase difference is determined by variations in the birefringence, $\delta\beta$, and length, δL, of the fiber:

$$\delta\varphi = L\delta\beta + \beta\delta L. \qquad (2)$$

The efficiency of polarization modulators is usually characterized through the ratio of the phase difference modulation index $\delta\varphi$ to the product of the applied voltage U and interaction length L [5]:

$$K' = \frac{\delta\varphi}{UL}\left[\frac{\text{rad}}{\text{V m}}\right]. \qquad (3)$$

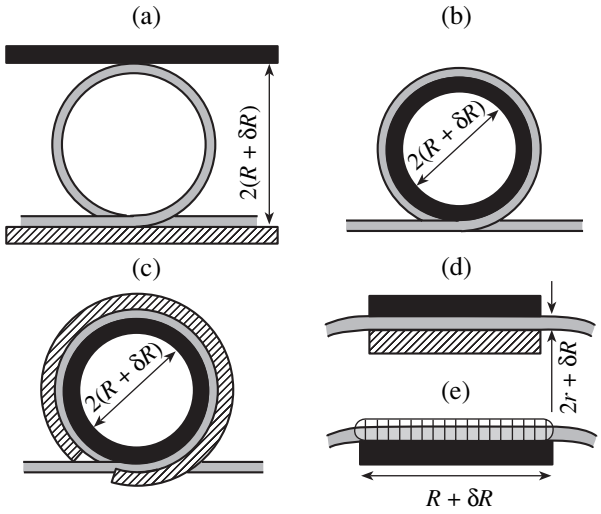

Fig. 1. Various designs of piezoceramic fiber-optic modulators of polarization mode phase difference.

However, this parameter depends strongly on frequency because of the resonance properties of electromechanical transducers (such as piezoceramic cylinders, etc.). It is therefore more appropriate to compare the efficiencies of fiber modulators with different kinds of perturbations applied to the fiber when the efficiency is defined as the ratio of the modulation index to a change δR in the piezoceramic cylinder radius with the interaction length L taken as a normalizing factor:

$$K = \frac{\delta\varphi}{\delta R L}\left[\frac{\text{rad}}{\mu\text{m m}}\right]. \quad (4)$$

Note that modulation efficiency may also be defined in terms of energy consumption, strain, etc. In this paper, we will hold definition (4) as the most convenient for practical use.

A mechanical perturbation of a single-mode fiber gives rise to several mechanisms of phase difference modulation. Their efficiency depends significantly on the type of perturbation and fiber. Below, we consider four modulation-inducing perturbations: bending (Fig. 1a), bending with tension (Figs. 1b, 1c), lateral pressure (Figs. 1c, 1d), and extension (Figs. 1b, 1c, 1e).

(1) Bending. To see how pure bending of an isotropic fiber affects its intrinsic birefringence, it is necessary to eliminate other types of deformation, in particular, longitudinal extension. To this end, several turns of the fiber are shaped into a circle, fixed together, and flattened in the diametral direction, so that the fiber experiences no extension when its radius R changes [11, Fig. 1]. As is known [1, 6], bending induces a static birefringence β_1 given by

$$\beta_1 = \frac{1}{2}C_S\frac{r^2}{R^2}, \quad (5)$$

where r is the fiber radius, $C_S = \pi n_1^3(p_{12} - p_{11})(1 + \mu)/\lambda$, μ is the Poisson's ratio of the fiber, n_1 is the refractive index of the core, and p_{11} and p_{12} are the elastooptic constants of quartz.

For $\lambda = 0.633$ μm, $C_S \approx 2.7 \times 10^6$ rad/m. The axes of induced birefringence are always uniquely related to the direction of the axis of the turns: the fast axis is perpendicular and the slow axis is parallel to it. Relationship (5) was verified experimentally.

The variation of the coil radius modulates the birefringence. Using formulas (2) and (5), we obtain an expression for the efficiency of this modulation mechanism:

$$K_1 = \frac{1}{2}C_S\frac{r^2}{R^3}. \quad (6)$$

This formula takes into account that the length of the deformed region of the fiber changes with the bend radius of the coil ($L = 2\pi RN$, where N is the number of turns). The bend-induced birefringence contains the factor of second order of smallness in r/R, which shows that the efficiency of modulation due to pure bending is low.

(2) Bending with tension. If the fiber is coiled around a cylinder, a change in the cylinder radius will change the longitudinal extension of the light guide. In this case, the mechanism of phase difference ($\delta\varphi$) modulation differs from that associated with pure bending. In the tension-coiled fiber, an additional linear birefringence β_2 appears [7],

$$\beta_2 = C_S\frac{2 - 3\mu}{1 - \mu}\frac{r}{R}\varepsilon \quad (7)$$

(where ε is the longitudinal strain in the fiber), which is added scalarly to the birefringence due to pure bending.

The directions of the optical axes in the coiled fiber are known: they are the same as in the case of pure bending and are related to the cylinder axis.

Formulas (5) and (7) give the total change in birefringence, $\beta_1 + \beta_2$, due to a change δR in the piezoceramic cylinder radius:

$$\begin{aligned}\delta\beta &= C_S\frac{r}{R^2}\delta R\left[\frac{r}{R} - \frac{2 - 3\mu}{1 - \mu}(1 - \varepsilon) + \frac{2 - 3\mu}{1 - \mu}\varepsilon\right] \\ &\approx C_S\frac{2 - 3\mu}{1 - \mu}\frac{r}{R^2}\delta R.\end{aligned} \quad (8)$$

Formula (8) assumes that $r \ll R$ and $\varepsilon \ll 1$ and therefore contains only the term for dynamic tension. Neglecting the geometrical elongation of the fiber yields the following expression for the modulation efficiency:

$$K_2 = C_S\frac{2 - 3\mu}{1 - \mu}\frac{r}{R^2}. \quad (9)$$

Within the above approximations, the modulation of birefringence is independent of the initial static tension of the fiber coiled around the cylinder, because (8) and (9) do not contain ε. This fact agrees with the results reported in [5] and is corroborated by our experiments.

(3) Lateral pressure. To simulate lateral pressure applied to cylindrical piezoceramic fiber polarization modulators, the fiber coiled around the cylinder was squeezed by tightened metal coverings (Fig. 1c). With an ac voltage applied to the ceramic cylinder, it changed the radius, exerting a lateral pressure on the fiber along its length. Polarization modulation arises from the modulation of birefringence (or phase difference $\delta\varphi = \delta\beta L$). The axes of the birefringence due to bending with tension are automatically aligned with the axes of birefringence induced by lateral pressure. In this case, the static component of birefringence caused by bending with tension adds up with the term β_3, which is approximately equal to [8]

$$\beta_3 = \frac{4C_S P}{\pi r E}, \quad (10)$$

where E is the Young's modulus of the fiber (77 GPa) and P is the squeezing force per unit length (N/m). Formula (10) was verified experimentally.

Note that formula (10) for the efficiency contains the hardly determinable parameter P, which must be expressed through δR.

Using methods of the elasticity theory [12], we found a relationship between the total displacement w of the extremities of the fiber that are on the axis of static force action and the force P:

$$w(P) \approx \frac{2}{\pi E} P \ln\left(\frac{\pi r E}{4P}\right). \quad (11)$$

In this expression, we neglect strains in the coverings and piezoceramic cylinder. Assuming that $\delta w = \delta R$, we find a relationship between changes in the radius of the cylinder, δR, and in the force P, δP:

$$\delta R \approx \delta P \frac{2}{\pi E} \ln\left(\frac{\pi r E}{4P}\right). \quad (12)$$

As follows from (10) and (12), the efficiency K of modulation due to lateral pressure on the fiber can be represented in the form

$$K_3 = \frac{2C_S}{r}\left[\ln\left(\frac{\pi r E}{4P}\right)\right]^{-1}. \quad (13)$$

It should be noted that the above formulas ignore the effect of protective polymeric coatings applied on the fiber. They, however, may reduce significantly the birefringence modulation due to lateral pressure, because their rigidity is one to two orders of magnitude lower than that of quartz.

Above, we considered birefringence modulation in isotropic single-mode fibers. However, polarization modulators based on anisotropic highly birefringent fibers (with a beat length $\Lambda \le 1$ cm) are of greater practical interest. Such fibers are known to exhibit a high intrinsic (fabrication-induced) birefringence. Anisotropic fibers are capable of reliably keeping the linear polarization of light at random external effects; therefore, they are widely used in polarization-sensitive pickups.

The analysis of the $\delta\varphi$ modulation mechanisms is also valid for anisotropic fibers. However, it should be kept in mind that a misalignment between the axes of intrinsic birefringence and axes of induced birefringence may cause unwanted coupling between polarization modes and change their amplitudes. In modulators with birefringent fibers, any of the above mechanisms may be accomplished only in the case of oriented fiber coiling (for example, in fibers with an elliptic cross section [13] that are matched to the axes of polarization eigenmodes).

Lateral squeezing of flat fiber regions is the easiest technique to induce phase difference modulation in an anisotropic fiber (Fig. 1d). However, it is necessary to align the birefringence axes with the direction of the force P in this case.

(4) Longitudinal extension. Unlike isotropic fibers, anisotropic fibers are appropriate for tension-related modulation of phase difference between polarization modes [9]. It has been shown [14] that such a modulation mechanism is efficient because of the nonuniform distribution of the Poisson's ratio μ and thermal expansion coefficient α over the cross section of an anisotropic fiber. In most cases, the cladding can be divided into two regions with different α_1, μ_1 and α_2, μ_2 where mechanical stresses govern intrinsic birefringence. If a fiber of length L elongates by a small amount δL, the intrinsic birefringence β changes by approximately [14]

$$\delta\beta = Q\beta\frac{\delta L}{L}. \quad (14)$$

Here, the dimensionless coefficient $Q = (\delta\beta L)/(\delta L\beta)$ is the ratio between the phase shifts that are associated with a change in the intrinsic birefringence and in the geometrical elongation of the fiber.

The quantity Q can be evaluated experimentally. It depends on the type of anisotropic fiber and is maximal (20–25) for bow-tie fibers.

The phase difference modulation index for this way of modulation is

$$\delta\varphi = (1 + Q)\beta\delta L. \quad (15)$$

The related efficiency is given by

$$K_4 = \frac{(1 + Q)\beta}{R}, \quad (16)$$

where R is the initial length of the stretcher (Fig. 1e) or the radius of the extending cylinder (Figs. 1b, 1c).

In modulators built around a piezoelectric cylinder, orientational fiber coiling is unnecessary, because here polarization modulation relies on tension alone. Birefringence induced by bending combined with tension is weak and randomly oriented with respect to the intrinsic birefringence axes; therefore, it does not affect the modulation efficiency. If an anisotropic fiber is coiled so that the intrinsic birefringence axes and the axes induced by bending with tension coincide, the efficiency may be improved.

EXPERIMENTAL

An experimental setup for studying the efficiency of fiber-optic modulators that modulate the phase difference between polarization modes is schematically shown in Fig. 2. Linearly polarized radiation from a He–Ne laser is applied to a single-mode fiber; passes through a fiber-optic modulator (M); and comes to a photodetector (D), the signal from which is displayed on an oscilloscope. An analyzer (A) is placed in front of the photodetector. The polarization of the radiation at the entrance to the fiber is adjusted with a half-wave plate ($\lambda/2$) so that the intensities of the polarization

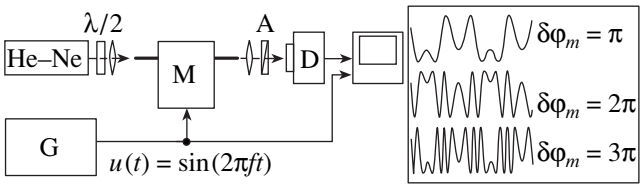

Fig. 2. Block diagram of the experimental setup and typical waveforms of phase-difference-modulated signals.

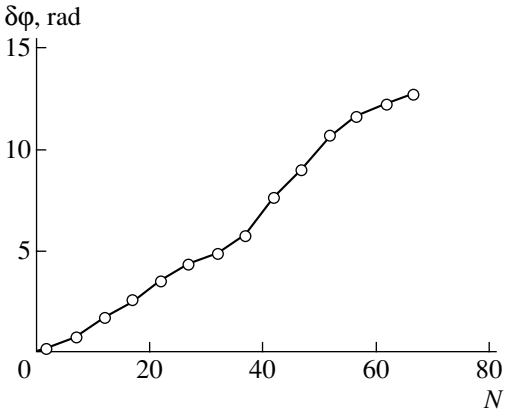

Fig. 3. Polarization modulation index versus the number N of turns of an anisotropic fiber coiled around a piezoceramic modulator (R = 1.6 cm, $\delta R = 7.3 \times 10^{-8}$ m, and Λ = 3 mm).

modes excited are the same. The modulator is controlled by a harmonic voltage generator (G). With the circuit tuned optimally and the phase difference modulated harmonically, $\delta\varphi(t) = \delta\varphi_m \sin(2\pi ft)$, the output signal of the photodetector is $U_D(t) = \cos(\varphi_0 + \delta\varphi_m \sin(2\pi ft))$. The modulation index $\delta\varphi_m$ can be readily determined from the waveform of such a signal (typical waveforms are shown in Fig. 2).

In experiments on measuring the efficiency of the modulators shown in Figs. 1a–1d, a standard isotropic single-mode fiber was used.

In the modulator with pure bending (Fig. 1a), the plates that squeeze the turns were actuated by an electromagnetic transducer (dynamic head).

The bending-and-tension mechanism of modulation was implemented with the fiber coiled around a piezoceramic cylinder (Fig. 1b).

Rigid tightened coverings (Fig. 1c) were used to exert lateral pressure. According to our theoretical estimates and measurements, such a way of pressure application provides the greatest modulation efficiency in this approach.

The design where a flat segment of the fiber without a protective cladding is squeezed (Fig. 1d) also used a piezoceramic transducer. Here, the displacement of the surfaces that squeeze the fiber was taken as δR in the formula for the modulation efficiency.

Experiments with anisotropic fibers used those with a stressing elliptic cladding with a beating length Λ = 3 and 1.7 mm ($Q \approx 11$).

In the experiments, we used several piezoceramic cylinders of different radii (9.5, 14, 16, and 37 mm) and $100 \times 17 \times 10$-mm piezoceramic bars. The piezoelectric transducers were excited by a harmonic voltage at a frequency of about 1 kHz (outside the range of natural resonances). Since a geometrical displacement provided by the transducer must be known to determine the modulator's efficiency, the oscillations of piezoceramic elements were calibrated in a series of special measurements. The measurements were carried out with a Michelson interferometer (one of the mirrors was displaced by the piezoceramic element) and fiber-optic Fabry–Perot interferometers (the fiber was fixed on the piezoceramic element). The efficiencies obtained are summarized in the table.

Table

Light guide	Modulator structure and basic modulation mechanism	Efficiency		
		theory	experiment	parameters used in calculation and experiment
Isotropic	Fig. 1a, pure bending	2.79×10^{-3}	2.51×10^{-3}	r = 50 μm; L = 0.35 m, R = 1.1 cm
	Fig. 1b, bending with tension	2.68	2.2	μ = 0.17; R = 0.95 cm, r = 50 μm, L = 2.21 m
	Fig. 1c, lateral pressure	7300	12.6*	P = 30 N/m; $E = 7.7 \times 10^{10}$ N/m^2; r = 62.5 μm; L = 2.21 m
	Fig. 1d, lateral pressure	8300	2000	P = 120 N/m; $E = 7.7 \times 10^{10}$ N/m^2; r = 62.5 μm; L = 0.08 m
Anisotropic	Fig. 1b, longitudinal extension	2.65	2.6	Q = 11; Λ = 3 cm; R = 0.95 cm; L = 2.45 m
	Fig. 1e, longitudinal extension	0.055	0.051	Q = 11; Λ = 1.7 cm; R = 8 cm
	Fig. 1d, lateral pressure	10^4	1700	P = 100 N/m; $E = 7.7 \times 10^{10}$ N/m^2; r = 50 μm; L = 0.08 m

* Fiber with protective polymeric cladding.

All the piezoceramic modulators had almost linear amplitude responses at voltages of up to several tens of volts.

The most important results of the experiments are as follows.

The experiments have corroborated the fact that the efficiency of lateral pressure application depends on the value of the static lateral pressure only slightly. It was found, in particular, that, instead of squeezing the cylinder with the fiber by the metal coverings, one may cover the fiber by wax (at frequencies above 1 kHz, wax acquires a sufficiently high rigidity and may be used as an outer covering).

In the design with an anisotropic fiber coiled around a piezoceramic cylinder, the mechanism associated with longitudinal extension dominates. In this case, the phase difference modulation index is expected to increase linearly with the length of the fiber in the modulator. In practice, however, a random orientation of the coiled fiber may induce a small additional birefringence associated with the bending-and-tension mechanism. The direction of the additional birefringence relative to the intrinsic birefringence is random. Therefore, experimental curves describing the modulation index vs. the number N of turns have a variable slope, whose average value is determined by pure tension (see Fig. 3). To reduce the parasitic irregular bending-induced modulation, the piezoelectric cylinder for a polarization modulator must have a radius as large as possible. It should also be noted that the protective cladding affects the efficiency of such a modulator insignificantly. When a composite material is under tension, internal stresses in its layers are proportional to the rigidity of the layer, so that the quartz, whose Young's modulus is 10 to 100 times higher than that of the protective cladding, takes over most of the load.

REFERENCES

1. S. C. Rachleigh, J. Lightwave Technol. **LT-1**, 312 (1983).
2. A. Yariv and P. Yeh, *Optical Waves in Crystals: Propagation and Control of Laser Radiation* (Wiley, New York, 1984; Mir, Moscow, 1987).
3. O. I. Kotov, L. B. Liokumovich, S. I. Markov, *et al.*, in *Proceedings of the 5th International Conference on Radar, Navigation, and Communication, Voronezh, 1999*, Vol. 3, pp. 1377–1388.
4. O. I. Kotov, L. B. Liokumovich, S. I. Markov, *et al.*, Pis'ma Zh. Tekh. Fiz. **26** (10), 28 (2000) [Tech. Phys. Lett. **26**, 415 (2000)].
5. A. R. Boyain, L. Martinez-Leon, J. L. Cruz, *et al.*, Appl. Opt. **38**, 6278 (1999).
6. R. Ulrich, S. C. Rashleigh, and W. Eickhof, Opt. Lett. **5**, 273 (1980).
7. S. C. Rashleigh and R. Ulrich, Opt. Lett. **5**, 354 (1980).
8. J.-I. Sakai and T. Kimura, IEEE J. Quantum Electron. **17**, 1041 (1981).
9. S. C. Rashleigh, in *Proceedings of the 1st International Conference on Optical Fiber Sensors, London, 1983*, pp. 210–213.
10. Yu. N. Kul'chin, *Distributed Fiber-Optics Measuring Systems* (Fizmatlit, Moscow, 2002).
11. O. I. Kotov, O. L. Marusov, and V. M. Nikolaev, Pis'ma Zh. Tekh. Fiz. **16** (10), 48 (1990) [Sov. Tech. Phys. Lett. **16**, 263 (1990)].
12. A. M. Kats, *Theory of Elasticity* (Lan', St. Petersburg, 2002).
13. Z. É. Arutyunyan, A. B. Grudinin, A. N. Gur'yanov, *et al.*, Kvantovaya Élektron. (Moscow) **17**, 1363 (1990).
14. M. P. Varnham, A. J. Barlow, D. N. Payne, and K. Okamoto, Electron. Lett. **19**, 699 (1983).

Translated by A. Khzmalyan

Technical Physics, Vol. 49, No. 1, 2004, pp. 76–82. Translated from Zhurnal Tekhnicheskoĭ Fiziki, Vol. 74, No. 1, 2004, pp. 77–83.

OPTICS, QUANTUM ELECTRONICS

Kinetics of Low-Temperature Initiation of $H_2/O_2/H_2O$ Mixture Combustion upon the Excitation of Molecular Vibrations in H_2O Molecules by Laser Radiation

A. M. Starik, N. S. Titova, and B. I. Loukhovitski

Baranov Central Institute of Aviation Motor Building, Moscow, 111116 Russia

e-mail: star@ciam.ru

Received April 17, 2003

Abstract—The initiation of $H_2/O_2/H_2O$ mixture combustion when asymmetric vibrations in H_2O molecules are excited by a resonant IR laser radiation is considered. It is shown that the vibrational excitation of the molecules gives rise to new efficient channels for the formation of chemically active O and H atoms and OH radicals. As a result, the chain mechanism of combustion in the mixtures is enhanced and, as a consequence, the induction time is cut and the ignition temperature is lowered. Even at a minor radiant energy flux delivered to the gas ($E_{in} \approx 2.5$ J/cm^2), the ignition temperature of the stoichiometric H_2/O_2 mixture containing only 5% of H_2O may become as low as 300 K. © *2004 MAIK "Nauka/Interperiodica".*

INTRODUCTION

The use of laser radiation to initiate combustion of various mixtures has been the subject of extensive discussion [1]. To date, only three methods of laser-initiated combustion have been tested experimentally. These are laser-induced thermal heating of a medium, photodissociation or photoionization of molecules under the action of laser radiation, and breakdown of a medium subjected to high-power radiation (laser spark) [2–10].

Among the methods that have been extensively investigated both theoretically and experimentally is local heating of a reacting mixture by laser radiation that is resonantly absorbed by vibrational–rotational transitions in molecules exhibiting high rates of vibrational–translational relaxation (for example, SF_6 or NH_3) [2–4] (these molecules are deliberately introduced into the mixture). In this case, the temperature of the gas mixture exposed to the radiation increases, which breaks molecular bonds and causes the formation of chemically active free radicals. These radicals initiate chain reactions. Such a method of combustion initiation (by CO_2 laser radiation with a wavelength $\lambda_I = 10.6$ μm) was applied to H_2/O_2, CH_4/O_2, C_2H_4/O_2, and C_3H_8/O_2 mixtures.

In the case of photochemical combustion initiation, the absorption of laser (usually UV) radiation by free–bound electron transitions in a molecule irradiated leads to photodissociation, i.e., to the formation of active atoms or radicals, which are responsible for the chain mechanism of combustion. With this method, the ignition of H_2/O_2, CH_4/O_2, and C_2H_2/O_2 mixtures exposed to radiations with $\lambda_I = 157$, 193, and 242 nm [5–7] was accomplished.

A laser spark arises when a high-power radiation pulse (of duration 10–100 ns and intensity 10^{10}–10^{11} W/cm^2) is incident on a very narrow region (of characteristic size 1 μm), causing multiphoton ionization of the gas in the beam channel and, as a result, electric breakdown of a medium. In this region, the temperature and pressure rise drastically (~10^6 K and ~10^8 Pa, respectively), which generates a shock wave starting from the beam axis and igniting the mixture. Laser-spark ignition has been accomplished in both gaseous and liquid media [8–10]. However, all these methods suffer from considerable disadvantages and their efficiency is low [10, 11].

Recently, a more effective method of combustion and detonation initiation has been proposed. It is based on the excitation of the electron states $O_2(a^1\Delta_g)$ and $O_2(b^1\Sigma_g^+)$ by laser radiation with wavelengths $\lambda_I = 1.268$ μm and 762 nm, respectively [12, 13]. In this case, combustion is initiated owing to high-rate chain reactions involving excited O_2 molecules. Here, the laser radiation energy required is much lower (~1 J/cm^2) than in the methods considered above. The basic cause for chain reaction intensification upon exciting O_2 molecules to the $a^1\Delta_g$ and $b^1\Sigma_g^+$ states is a decrease in the barrier of endoergic reactions involving excited molecules.

Vibrationally excited molecules also react 10^2–10^3 times faster than unexcited ones [14]. To excite molecular vibrations, IR laser radiation is necessary. In this case, a small amount of IR-active molecules must be introduced into a combustible gas mixture (for example, H_2/O_2). The most appropriate candidates are H_2O molecules, which absorb intensely the radiation of

1063-7842/04/4901-0076$26.00

HF, CO, and CO_2 lasers (of wavelengths 2.7, 5.6, and 10.4 μm, respectively) [15]. Importantly, the addition of these molecules does not affect the properties of combustibles [16].

The aim of this study is to analyze the kinetic mechanisms behind combustion initiation in $H_2/O_2/H_2O$ mixtures when the vibrational degrees of freedom of H_2O molecules are excited by laser radiation.

STATEMENT OF THE PROBLEM AND BASIC EQUATIONS

An H_2O molecule has three (symmetric, ν_1 = 3656 cm^{-1}; deformation, ν_2 = 1594 cm^{-1}; and asymmetric, ν_3 = 3755.8 cm^{-1}) vibrational modes. The rotational motion of an H_2O molecule and, correspondingly, a set of rotational energy levels can be described in terms of the asymmetric top model, where the energy of the top is characterized by three quantum numbers J, K_a, and K_c [17]. The first determines the total angular momentum; the second and third ones, its projection onto the axis of an oblate and prolate symmetric top. Therefore, a vibrational–rotational state of an H_2O molecule is defined by three vibrational, V_i (i = 1–3), and three rotational quantum numbers: $m \equiv V_1V_2V_3(JK_aK_c)$.

Radiations with $\lambda_I \approx 2.7$ and 5.6 μm are absorbed at the 000 ⟶ 001 (100) and 010 ⟶ 001 (100) vibrational–rotational transitions, respectively; radiation with $\lambda_I \approx 10.4$ μm, at the 000 ⟶ 010 transition [15]. Laser excitation of the vibrations will be effective if the condition $\tau_I \ll \tau_V$ is met, where τ_I is the time of induced transitions and τ_V is the relaxation time of an excited state. The asymmetric mode of an H_2O molecule has the longest relaxation time [18]. Therefore, the maximum excitation efficiency of molecular vibrations in H_2O is achieved when the radiation with $\lambda_I \approx 2.7$ μm is used.

We will analyze a stoichiometric H_2/O_2 mixture (containing 5% of water vapor) exposed to radiation with λ_I = 2.66 μm. In an H_2O molecule, this radiation is absorbed at the center of the spectral line for the $000(6_{34}) \longrightarrow 001(6_{33})$ transition. Consider the case when $\tau_I > \tau_t$, τ_r, and τ_{VV}, where τ_t and τ_r are the characteristic times of translational and rotational relaxations, respectively, and τ_{VV} is the time of vibrational–vibrational intramode exchange. One may assume that, at $t \geq \tau_I$, translational and rotational degrees of freedom in the molecule are in thermodynamic equilibrium and that the Boltzmann distribution of molecules over vibrational levels with a related vibrational temperature T_ξ is established very rapidly within each of the modes (ξ = 1, …, n, where n is the number of modes both in initial molecules of the reactive mixture and in resultant molecules).

Consider processes occurring in the central zone of the laser beam ($r < R_b$) with the Gaussian intensity distribution along its radius: $I(r, t) = I_0(t)\exp(-r^2/R_b^2)$, where R_b is the characteristic radius of the pulsed radiation beam. The duration of the pulse is τ_p, so that $I_0(t) = I_0$ at $0 < t \leq \tau_p$ and $I_0(t) = 0$ at $t > \tau_p$. The variation of the hydrodynamic parameters and reactant concentrations in the irradiation zone is specified by the hierarchy of characteristic times for different processes of macro- and microtransfer [19, 20]. For a vibrationally nonequilibrium reacting gas, these times are the time τ_a of propagation of acoustical vibrations across the beam; the times of multicomponent, τ_D^i, and thermal, τ_T^i, diffusions of an ith component; the time τ_{Di}^V of vibrational thermal diffusion; the time τ_λ of heat conduction; the time of change of the state of the medium under the action of the striction force τ_F, the time τ_I of induced transitions, the energy relaxation time τ_V for vibrationally excited states, the pulse duration τ_p, and the characteristic time τ_{iq}^{ch} of the chemical reaction that produces (destroys) a component responsible for the chain mechanism of the process (in our case, these components are O and H atoms and OH radicals).

For a stoichiometric H_2/O_2 mixture containing 5% water vapor ($H_2/O_2/H_2O$ = 0.633/0.317/0.05) exposed to radiation with λ_I = 2.66 μm, the estimation of these characteristic times under the conditions typical of numerical experiments (I_0 = 1–20 kW/cm^2, R_b = 10 cm, $P_0 = 10^3$–10^5 Pa, and T_0 = 300–700 K), we have $\tau_a \approx 2 \times 10^{-4}$ s, $\tau_I = 3 \times 10^{-6}$–$4.5 \times 10^{-4}$ s, $\tau_V = 3.4 \times 10^{-7}$–$2 \times 10^{-5}$ s, $\tau_D \sim \tau_T \sim \tau_\lambda \sim \tau_D^V$ = 0.3–10 s, and τ_F = 0.1–1 s. Consider the cases where $\tau_V < \tau_I \leq \tau_p < \tau_a \ll \tau_D, \tau_F$. Under these conditions, the coefficient k_ν of radiation absorption by H_2O molecules varies from 5×10^{-4} to 2×10^{-2} cm^{-1}. Therefore, the condition $k_\nu^{-1} \gg R_b$ is also valid. Hence, one may neglect the variation of the parameters along the direction of beam propagation and use the approximation of thin optical layer. The system of equations of state for the medium at the center of the zone exposed can then be written in the form

$$\frac{d\gamma_i}{dt} = G_i - \gamma_i \sum_{k=1}^{M_1} G_k,$$

$$\frac{d\varepsilon_\xi}{dt} = Q_{VV'}^\xi + Q_{VT}^\xi + Q_{ch}^\xi + Q_I^\xi,$$

$$\frac{dH}{dt} + \sum_{i=1}^{S} \frac{de_V^i}{dt} = \frac{k_\nu I_0(t)}{\rho},$$

$$Q_{VV'}^\xi = N \sum_{p=1}^{L_1} \frac{l_\xi L_{\xi,p}}{g_\xi^{l_\xi} g_p^{l_p}} W'_{\xi,p},$$

$$Q_{VT}^{\xi} = (\varepsilon_{\xi 0} - \varepsilon_{\xi})(1 - y_{\xi 0}) \sum_{i=1}^{M_1} W_{\xi,0}^{i} \gamma_i,$$

$$Q_{ch}^{\xi} = \sum_{r=1}^{L_2} \frac{\alpha_{ir}^{-} - \alpha_{ir}^{+}}{N_i} [(\chi_{r\xi}^{+} - \varepsilon_{\xi}) R_r^{+} - (\chi_{r\xi}^{-} - \varepsilon_{\xi}) R_r^{-}],$$

$$Q_I^{\xi} = l_{\xi I} \frac{k_{\nu} I_0(t)}{h \nu_1 N_{H_2O}},$$

$$L_{\xi,p} = \varepsilon_{\xi}^{l_{\xi}} (g_p + \varepsilon_p)^{l_p} - \varepsilon_p^{l_p} (g_{\xi} + \varepsilon_{\xi})^{l_{\xi}} \exp\left(\frac{l_{\xi}\theta_{\xi} - l_p \theta_p}{T}\right),$$

$$R_q^{+(-)} = k_{+(-)q} \prod_{j=1}^{n_q^{+(-)}} N_j^{\alpha_{jq}^{+(-)}}, \quad G_i = \sum_{q=1}^{M_2} S_{iq},$$

$$S_{iq} = \frac{(\alpha_{iq}^{-} - \alpha_{iq}^{+})}{N} (R_q^{+} - R_q^{-}),$$

$$k_{\nu} = \sigma_{mn} \left(\frac{g_n}{g_m} N_m - N_n\right),$$

$$\sigma_{mn} = \frac{\lambda_{mn}^2}{4\pi b_D} A_{mn} \sqrt{\frac{\ln 2}{\pi}} H(x, a), \quad \gamma_i = N_i / N,$$

$$\chi_{r\xi}^{+(-)} = (E_r / K\theta_{\xi}) \eta_{r\xi}^{+(-)}, \quad \eta_{r\xi}^{+(-)} = \beta_{r\xi}^2 / \sum_i \beta_{ri}^2,$$

$$\eta_{r\xi}^{-} = \beta_{r\xi}^2 T_{\xi}^2 \sum_i \beta_{ri}^2 \left(\sum_j \beta_{rj}^2 T_j\right)^{-2},$$

$$H = \frac{1}{\mu} \sum_{i=1}^{M_1} h_{0i} \gamma_i + \left[\frac{3}{2} + C_R\right] \frac{R}{\mu} T,$$

$$P = \frac{\rho R T}{\mu}, \quad N = P / KT,$$

$$e_V^i = \gamma_i \frac{R}{\mu} \sum_{j=1}^{Z} \theta_{ij} \varepsilon_{ij}, \quad C_R = \sum_{i=1}^{L} \gamma_i + \frac{3}{2} \sum_{i=L+1}^{S} \gamma_i,$$

$$\mu = \sum_{i=1}^{M_1} \mu_i \gamma_i, \quad E_r = \alpha_r E_{ar}^{+}, \quad \alpha_r = \frac{E_{ar}^{+}}{E_{ar}^{+} + E_{ar}^{-}},$$

$$\varepsilon_{\xi} = g_{\xi} y_{\xi} / (1 - y_{\xi}), \quad \varepsilon_{\xi 0} = \varepsilon_{\xi}(y_{\xi} = y_{\xi 0}),$$

$$y_{\xi} = \exp(-\theta_{\xi} / T_{\xi}), \quad y_{\xi 0} = y_{\xi}(T_{\xi} = T).$$

Here, ρ, T, and P are the density, temperature, and pressure of the gas, respectively; R is the universal gas constant; K is the Boltzmann constant; h_{0i} is the enthalpy of formation of an ith component in the mixture at T = 298 K; μ_i is the molecular mass of this component; S is the number of molecular components; L is the number of components consisting of linear molecules; θ_{ξ} is the characteristic vibrational temperature of a ξth mode; g_{ξ} is the degeneracy multiplicity of the ξth mode; Z is the number of modes in a molecule of an ith sort; $W'_{\xi,p} = \sum_{i=1}^{M_1} W_{\xi,p}^{i} \gamma_i$ and $W'_{\xi,p} = W_{\xi,p}\gamma_j$ ($\xi \longleftrightarrow i$, $p \longleftrightarrow j$) in the case of intra- and intermolecular V–V' exchange, respectively; $W_{\xi,0}^{i}$ is the rate constant of V–T relaxation due to collision with an ith partner; N_i is the concentration of molecules of an ith sort; M_1 is the number of atomic and molecular components in the mixture; l_{ξ} is the number of vibrational quanta that are lost or gained by a ξth mode in V–V' exchange (L_1 is the number of V–V' exchange channels); α_{iq}^{+} and α_{iq}^{-} are the stoichiometric coefficients of a qth reaction that produces an ith component; k_{+q} and k_{-q} are the rate constants of a forward (+) and back (–) qth reaction, respectively; L_2 is the number of reactions that produce (destroy) a molecule containing the mode ξ; M_2 is the number of reactions that form an ith component; β_{rj} are the coefficients of expansion of an rth reaction in coordinates of the normal modes; E_r is the part of the activation energy of an rth reaction that is accounted for by the vibrational degrees of freedom; $E_{ar}^{+(-)}$ is the activation energy of an rth chemical reaction that destroys (produces) a vibrationally excited molecule; $l_{\xi,I}$ is the number of vibrational quanta gained (lost) by the mode ξ due to induced transitions; N_m and N_n are the concentrations of H_2O molecules in the lower and upper states of the transition $m \longrightarrow n$, respectively; g_m and g_n are the degeneracy multiplicities of these states, respectively; h is the Planck constant; λ_{mn} is the wavelength at the center of the spectral line of the absorbing transition $m \longrightarrow n$; A_{mn} is the Einstein coefficient for this transition; b_D is the half-height Doppler width of a spectral line; and $H(x, a)$ is the Voigt function. The value of this function was evaluated with allowance for the joint action of the Doppler and collisional broadening of a spectral line. The values of the collisional broadening coefficient b_C^M (M = H_2, O_2, and H_2O) were taken the same as in [15]. The values of $\beta_{r\xi}$ were taken to be equal to unity (as in [21]). If the translational and vibrational degrees of freedom of molecules participating in a reaction are not in equilibrium, the rate constant of a chemical reaction is a function of T and T_{ξ} (in the framework of the mode approximation) and can be represented in the form

$$k_q = \varphi_q(T, T_{\xi}) k_q^0(T).$$

Here, $k_q^0(T)$ is the rate constant of a qth chemical reaction at $T_{\xi} = T$ and $\varphi(T, T_{\xi})$ is the nonequilibrium factor. The value of this factor at given T and T_{ξ} was calculated in the same way as in [21]. The kinetic model used in this paper to analyze combustion initiation takes into

consideration (i) 29 reversible chemical reactions with the participation of H, O, OH, H_2O, H_2, O_2, HO_2, H_2O_2, and O_3; (ii) vibrational–vibrational (V–V') exchange between the symmetric, deformation, and asymmetric (ν_1, ν_2, and ν_3) vibrational modes of H_2O; the modes of H_2 (ν_4), O_2 (ν_5), and OH (ν_6); the symmetric, deformation, and asymmetric vibrational modes of HO_2 (ν_7, ν_8, and ν_9) and O_3 (ν_{10}, ν_{11}, and ν_{12}) molecules; and between the modes of an H_2O_2 molecule (ν_{13}, ν_{14}, ν_{15}, ν_{16}, and ν_{17}); and (iii) vibrational–translational (V–T) energy relaxation for the modes ν_2, ν_4, ν_5, ν_6, ν_9, ν_{11}, and ν_{15}. The temperature dependences of the rate constants $k_q^0(T)$ for the chemical reactions,as well as of the rate constants for V–V' exchange ($W^i_{\xi,p}$ and $W_{\xi,p}$) and V–T relaxation ($W^i_{\xi,0}$) were taken the same as in [21].

COMBUSTION INITIATION IN THE $H_2/O_2/H_2O$ MIXTURE UNDER IRRADIATION WITH Λ_I = 2.66 μm

It is known that a gas mixture in a zone with a radius R_b is ignited when the condition $\tau_i^{ch} \le \tau_i^R$ is fulfilled. For an H_2/O_2 mixture, τ_i^{ch} is the formation time of active H and O atoms and OH radicals and τ_i^R is the time it takes for them to leave the reaction zone. In the problem under consideration, the latter is the diffusion time of the lightest carriers of the chain mechanism (H atoms); i.e., $\tau_i^R = \tau_D^H$. Thus, at $\tau_i^{ch} > \tau_D^H$, the value of the induction period τ_{in} is bound from above by the time τ_D^H. Figure 1 shows the dependences of τ_D^H (for R_b = 10 cm) and τ_{in} for different laser radiation energies delivered to the gas ($E_{in} = I_0\tau_p$) on the initial temperature of the $H_2/O_2/H_2O$ = 0.633/0.317/0.05 mixture at $P_0 = 10^3$ Pa. It is seen that the radiation with λ_I = 2.66 μm cuts considerably τ_{in} and decreases the self-ignition temperature T_{ign}. In a first approximation, the value of T_{ign} can be found from the relationship $\tau_{in}(T_{ign}, P_0, E_{in}) = \tau_D^H(T_{ign}, P_0)$. Even at E_{in} = 1 J/cm^2, T_{ign} equals 540 K instead of 750 K when E_{in} = 0. At $E_{in} \ge 2.7$ J/cm^2, ignition becomes possible at temperatures as low as 300 K. At E_{in} = 10 J/cm^2, a new tendency in the τ_{in} vs. T_0 dependence is observed (as T_0 decreases starting with 600 K, so does τ_{in}).

The decrease in the delay times of ignition, which are observed when asymmetric vibrations of an H_2O molecule are excited by laser radiation, are explained by a change in the formation kinetics of chemically active O and H atoms and OH radicals and also by an increase in the temperature of the mixture due to V–T-relaxation of the vibrational energy of the H_2O molecules excited. Figure 2 shows the mole fractions of the

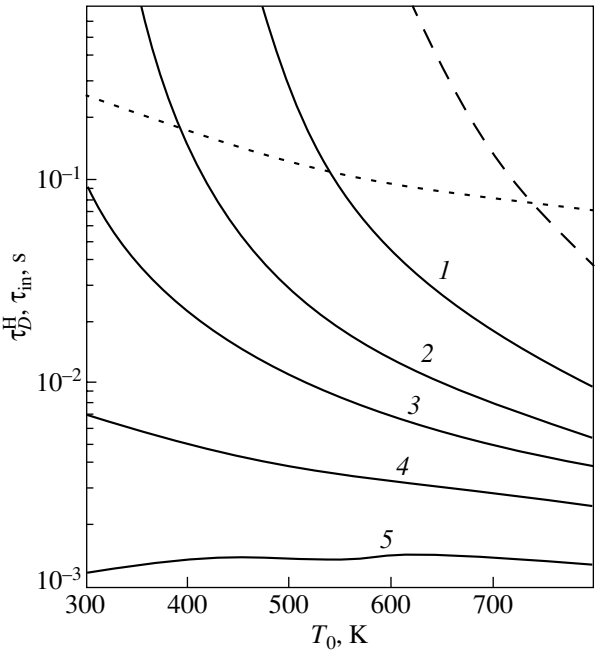

Fig. 1. Dependences $\tau_D^H(T_0)$ (dotted curve); $\tau_{in}(T_0)$ for the $H_2/O_2/H_2O$ = 0.633/0.317/0.05 mixture at $P_0 = 10^3$ Pa under the action of a radiation pulse with a duration $\tau_p = 10^{-4}$ s and E_{in} = (*1*) 1, (*2*) 2, (*3*) 3, (*4*) 5, and (*5*) 10 J/cm^2; and $\tau_{in}(T_0)$ in the absence of radiation (dashed curve).

initial and resultant components in the $H_2/O_2/H_2O$ mixture (at T_0 = 600 K and $P_0 = 10^3$ Pa) vs. time for E_{in} = 0 and 2.5 J/cm^2. It is seen that, when the gas is exposed to radiation with λ_I = 2.66 μm, both the dependences $\gamma_i(t)$ and the temperature dynamics change. While at E_{in} = 0 the maximum concentrations in the interval [0, τ_{in}] are typical of O_3, H, and HO_2, at E_{in} = 2.5 J/cm^2, the maximal concentrations are reached for H, O, and OH. Moreover, the latter components form even at $t < \tau_p$. After the relaxation of the laser energy absorbed by the H_2O molecules at $t > \tau_V$ (at the given T_0 and P_0, $\tau_V \approx 7 \times 10^{-5}$ s), the gas temperature begins to rise. At $t = \tau_p$, it reaches 840 K and subsequently (in the interval [τ_p, τ_{in}]) remains virtually unchanged.

Nevertheless, the basic reason for the decrease in τ_{in} is the appearance of new effective channels for the formation of H and O atoms and OH radicals rather than the heating of the medium. For example, if only the thermal mechanism worked, the value of τ_{in} would be 2.24×10^{-2} s. If, however, we also take into account the formation of H, O, and OH upon the excitation of H_2O molecules, the value of τ_{in} declines to 9.2×10^{-3} s. With decreasing T_0, the difference in the values of τ_{in} grows and reaches 5.7 times at T_0 = 300 K and E_{in} = 5 J/cm^2.

It should be noted that the excitation of asymmetric vibrations in an H_2O molecule by radiation with λ_I = 2.66 μm also causes symmetric vibrations in H_2O (because of the high-rate intramolecular V–V'

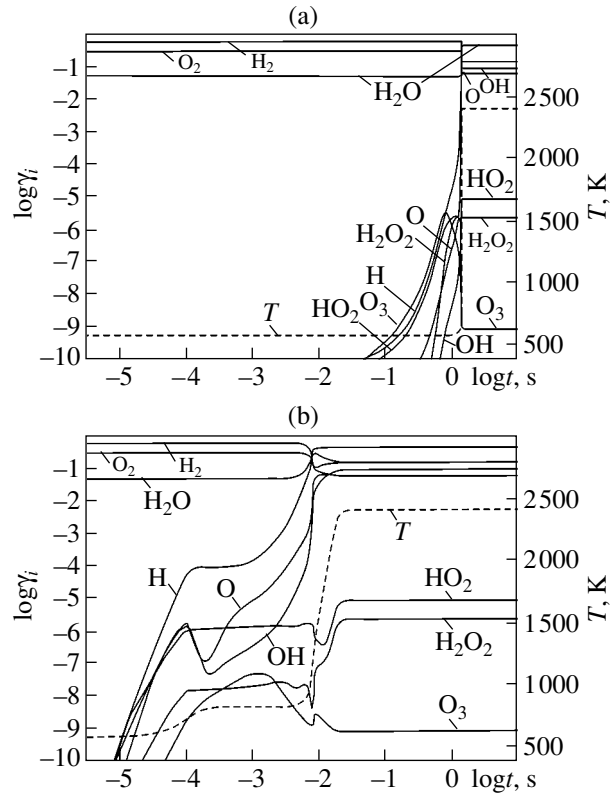

Fig. 2. Time dependences of the concentrations (mole fractions) of the components (solid curves) and temperature (dashed curves) for the $H_2/O_2/H_2O$ = 0.633/0.317/0.05 mixture at $P_0 = 10^3$ Pa and T_0 = 600 K (a) in the absence of radiation and (b) under the action of radiation with λ_I = 2.66 μm, E_{in} = 2.5 J/cm^2, and $\tau_p = 10^{-4}$ s.

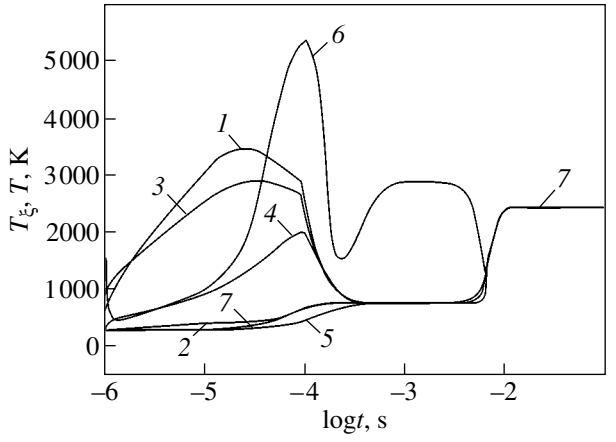

Fig. 3. Time dependences for the vibrational temperatures T_ξ of the modes (*1*) ν_1, (*2*) ν_2, (*3*) ν_3, (*4*) ν_4, (*5*) ν_5, and (*6*) ν_6 and (*7*) translational temperature T in the $H_2/O_2/H_2O$ = 0.633/0.317/0.05 mixture (kept at $P_0 = 10^3$ Pa and T_0 = 300 K) exposed to radiation with λ_I = 2.66 μm, E_{in} = 5 J/cm^2, and $\tau_p = 10^{-4}$ s.

exchange $H_2O(001) + M = H_2O(100) + M$) and vibrations in an H_2 molecule (due to the intermolecular *V–V'* exchange $H_2(V = 0) + H_2O(001) = H_2(V = 1) + H_2O(000)$). This is illustrated in Fig. 3, which shows the time variation of the vibrational, T_ξ (ξ = 1, 2, 3, 4, 5, 6), and translational, *T*, temperatures after the ignition of the $H_2/O_2/H_2O$ mixture at T_0 = 300 K and $P_0 = 10^3$ Pa for E_{in} = 5 J/cm^2. Note that initially ($t < 10^{-5}$ s) the vibrational temperature of OH grows due to the *V–V'* exchange $OH(V = 0) + H_2O(100) = OH(V = 1) + H_2O(000)$, while at $t > 10^{-5}$ s, it grows because the energy being released goes into OH vibrations when the chemical reaction $H_2 + O_2 = 2OH$ proceeds. Thus, the radiation with λ_I = 2.66 μm creates the situation where not only vibrationally excited H_2O molecules but also excited H_2 molecules and OH radicals enter into the chemical reactions within the interval $[0, \tau_p]$.

The excitation of H_2O and H_2 molecules provides new effective channels for the production of O and H atoms and OH radicals. Figure 4 demonstrates these channels for the cases when excited H_2O molecules in the $H_2/O_2/H_2O$ mixture are absent and when H_2O molecules are excited by the radiation with λ_I = 2.66 μm. The basic chain-initiating reaction at $T_0 < 800$ K (excitation is absent) is the reaction $H_2 + O_2 = 2OH$, which produces vibrationally excited OH radicals. Next, the OH radicals react with the H_2 molecules to form H atoms: $OH + H_2 = H_2O + H$. The H atoms react with the O_2 molecules to produce O atoms and OH radicals: $H + O_2 = OH + O$. This is one of the reactions responsible for chain propagation. The O atoms, when reacting with the H_2 molecules, produce OH and H. This is the second basic reaction in the chain mechanism of igniting the H_2/O_2 mixture. Another reaction in which the H atoms participate is the formation of HO_2: $H + O_2 + M = HO_2 + M$. At low temperatures ($T_0 < 800$ K) and sufficiently high pressures ($P_0 > 10^4$ Pa) of the mixture, the HO_2 molecules recombine intensely to form inactive hydrogen peroxide.

The radiation-induced (λ_I = 2.66 μm) excitation of asymmetric vibrations in H_2O molecules alters considerably the combustion initiation scheme. Indeed, excited H_2O molecules dissociate even at low temperatures. In addition, they react with O_2 molecules ~10^2 times more rapidly than unexcited ones. Therefore, the basic chain-initiating reactions in this case are those involving excited H_2O molecules: $H_2O(100, 001) + M = H + OH + M$, $H_2O(100, 001) + O_2 = OH + HO_2$, and $H_2O(100, 001) + O_2 = H_2O_2 + O$, which immediately produce active O and H atoms and OH radicals. Another important feature of this chain process is the presence of vibrationally excited H_2 molecules in the mixture. The rate of the branching chain reaction $H_2(V = 1) + O = H + OH$ far exceeds that of the reaction with the participation of unexcited H_2 molecules. That is why the delay time of ignition shortens and T_{ign} declines when molecular vibrations in H_2O molecules are excited by laser radiation.

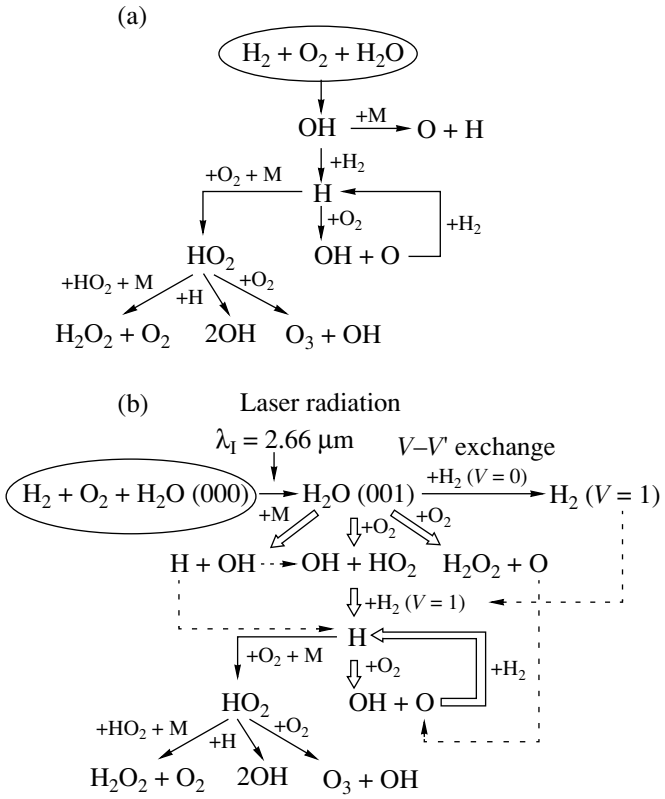

Fig. 4. Formation of the chain mechanism upon igniting the $H_2/O_2/H_2O$ mixture (a) without the excitation of the H_2O molecules and (b) with the excitation of molecular vibrations in the H_2O molecules by radiation with $\lambda_I = 2.66$ μm.

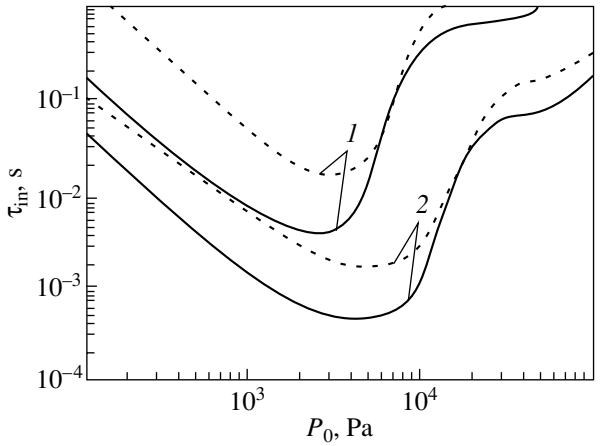

Fig. 5. Dependences $\tau_{in}(T_0)$ for the $H_2/O_2/H_2O$ = 0.633/0.317/0.05 mixture exposed to radiation with λ_I = 2.66 μm and E_{in} = (*1*) 5 and (*2*) 10 J/cm^2 for the cases when asymmetric vibrations in the H_2O molecules are excited (solid curves) and when the absorbed energy is totally spent on heating the gas (dashed curves).

Now let us compare the efficiency of the combustion initiation method proposed (which relies on the excitation of asymmetric vibrations in H_2O molecules) and the method of direct heating of a reactive mixture by resonant laser radiation (the total energy absorbed by the gas is spent on heating a medium), which is currently the subject of wide speculation. Figure 5 shows the associated dependences of τ_{in} on the initial pressure of the $H_2/O_2/H_2O$ = 0.633/0.317/0.05 mixture at T_0 = 300 K, the laser energy density E_{in} = 5 and 10 J/cm^2, and $\tau_p = 10^{-4}$ s. It is seen that, when H_2O molecules are excited by radiation with λ_I = 2.66 μm and E_{in} = 5 J/cm^2, τ_{in} may be five to ten times shorter. At E_{in} = 10 J/cm^2, this difference decreases slightly but remains significant (2.5–5 times) in the range $P_0 = 10^2$–10^4 Pa. Also, for each E_{in}, there exists the boundary value of the initial pressure, P_{0b}, above which the value of τ_{in} starts increasing. The higher E_{in}, the greater P_{0b}. It should be noted that the value of P_{0b} also depends on the initial temperature of the mixture (P_{0b} increases with T_0). Such behavior of the dependences $\tau_{in}(P_0)$ for the $H_2/O_2/H_2O$ mixture is explained by the fact that, at low T_0 ($P_0 > P_{0b}$), HO_2 molecules are formed by the reaction H + O_2 + M = HO_2 + M, the rate of which, at $T < 800$ K, is higher than the rate of the chain propagation reaction H + O_2 = OH + O. Under these conditions, the dominating process with the participation of HO_2 becomes the recombination reaction $2HO_2 + M = H_2O_2 + O_2 + M$ (Fig. 4). At low temperatures, H_2O_2 molecules are inactive and serve as a sink for the chain mechanism carriers (H atoms). An increase in E_{in} accelerates the H atom production by the reaction $H_2O(001) + M = H + OH + M$ and, accordingly, increases P_{0b}.

CONCLUSIONS

The excitation of asymmetric vibrations in an H_2O molecule by laser radiation leads to the formation of new effective channels for the production of chemically active O and H atoms and OH radicals. This, as well as the presence of vibrationally excited H_2 molecules, which form via the *V*–*V*' exchange $H_2O(001) + H_2(V = 0) = H_2O(000) + H_2(V = 1)$, enhances the chain mechanism of combustion in the $H_2/O_2/H_2O$ mixture. Accordingly, the induction time and the ignition temperature decrease. Even at a low radiation energy density delivered to the $H_2/O_2/H_2O$ mixture, E_{in} = 2.7 J/cm^2, the ignition temperature of the mixture kept at a low pressure ($P_0 = 10^3$ Pa) may be lowered down to 300 K. For combustion initiation, the excitation of asymmetric or symmetric vibrations in H_2O molecules by resonant laser radiation is much more (five to ten times) effective than the thermal heating of the medium by IR laser radiation. Since H and O atoms, as well as OH radicals, are also responsible for the chain mechanism in igniting hydrocarbon fuels, our method of chain reaction intensification, which is based on the laser-induced excitation of molecular vibrations in reactant molecules (or in IR-active molecules deliberately introduced into the mixture), may be efficient in various applications.

ACKNOWLEDGMENTS

This work was financially supported by the Russian Foundation for Basic Research (grant nos. 02-01-00703 and 02-02-16915).

REFERENCES

1. P. D. Ronney, Opt. Eng. **33**, 510 (1994).
2. W. M. Trott, J. Appl. Phys. **54**, 118 (1983).
3. B. Raffel, J. Warnatz, and J. Wolfrum, Appl. Phys. B **37**, 189 (1985).
4. M. A. Ranoff, M. D. Smooke, R. E. Teets, and J. A. Sell, Combust. Flame **103**, 253 (1995).
5. D. Lucas, D. Dunn-Rankin, K. Hom, and N. J. Brown, Combust. Flame **69**, 171 (1987).
6. B. E. Forch, F. E. Fendell, and H. W. Behrens, Proc. SPIE **1862**, 45 (1993).
7. M.-S. Chou and T. J. Zukowski, Combust. Flame **87**, 191 (1991).
8. J. A. Syage, E. W. Fournier, R. Rianda, and R. B. Cohen, J. Appl. Phys. **64**, 1499 (1988).
9. J. X. Ma, D. R. Alexander, and D. E. Poulain, Combust. Flame **112**, 492 (1998).
10. T. X. Phuoc and F. P. White, Combust. Flame **119**, 203 (1999).
11. M. H. Morsy, Y. S. Ko, and S. H. Chung, Combust. Flame **119**, 473 (1999).
12. A. M. Starik and N. S. Titova, Dokl. Akad. Nauk **380**, 332 (2001) [Dokl. Phys. **46**, 627 (2001)].
13. A. M. Starik and N. S. Titova, Kinet. Katal. **44**, 35 (2003).
14. G. G. Light, J. Chem. Phys. **68**, 2831 (1978).
15. V. A. Levin, A. A. Sorokin, and A. M. Starik, *Nonequilibrium Gas Flows with Physicochemical Conversions* (Mosk. Gos. Univ., Moscow, 1980), pp. 4–25.
16. E. V. Gurentsov, O. G. Divakov, and A. V. Eremin, Teplofiz. Vys. Temp. **40**, 416 (2002).
17. W. S. Benedict, M. A. Pollack, and W. J. Tomplinson, IEEE J. Quantum Electron. **5**, 108 (1969).
18. A. M. Starik and A. N. Khmelevskii, Khim. Fiz. **12**, 456 (1993).
19. V. I. Grabovskii and A. M. Starik, Kvantovaya Élektron. (Moscow) **21**, 365 (1994).
20. A. M. Starik and N. S. Titova, Kinet. Katal. **41**, 650 (2000).
21. A. M. Starik and N. S. Titova, Khim. Fiz. **19** (9), 61 (2000).

Translated by Yu. Vishnyakov

Technical Physics, Vol. 49, No. 1, 2004, pp. 8–18. Translated from Zhurnal Tekhnicheskoĭ Fiziki, Vol. 74, No. 1, 2004, pp. 10–20.

GASES AND LIQUIDS

On Nonlinear Resonant Four-Mode Interaction between Capillary Vibrations of a Charged Drop

S. O. Shiryaeva, A. N. Zharov, and A. I. Grigor'ev

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

e-mail: shir@uniyar.ac.ru

Received June 5, 2003

Abstract—Energy transfer from higher modes of capillary vibrations of an incompressible liquid charged drop to the lowest fundamental mode under four-mode resonance is studied. The resonance appears when the problem of nonlinear axisymmetric capillary vibration of a drop is solved in the third-order approximation in amplitude of the multimode initial deformation of the equilibrium shape of the drop. Although the resonant interaction mentioned above builds up the fundamental mode even in the first order of smallness, its amplitude turns out to be comparable to a quadratic (in small parameter) correction arising from nonresonant nonlinear interaction, since the associated numerical coefficients are small. *© 2004 MAIK "Nauka/Interperiodica".*

Nonlinear vibration of charged drops is of considerable interest in a variety of areas of science and technology. This problem has been repeatedly considered in statements of different degree of complexity and rigor (see, for example, [1–8] and references therein). Nevertheless, several issues of considerable interest remain to be investigated. It is still unclear, in particular, whether the fundamental mode amplitude may by resonantly enhanced by gaining energy from higher modes. This problem is of crucial importance in the theory of thunderstorm electricity in the context of lightning initiation by a corona near a large charged drop or water-covered hailstone in a storm cloud [9, 10]. Although such a mechanism seems plausible, considerable evidence for this mechanism is still lacking. According to full-scale experimental data [11], the self-charge of large drops and hailstones in clouds is too low for a corona discharge to be initiated near them or the surface of a drop to become unstable. At the same time, it is obvious that, when a drop elongates into a body close to a spheroid, the field strength near its vertices increases appreciably. A drop may extend into a spheroid through the excitation of the fundamental vibration mode upon resonant energy transfer from higher modes to the fundamental one [12–14]. However, calculations [12, 13] show that, under three-mode nonlinear resonant interaction between vibration modes, the lowest mode that can acquire energy from higher modes is the third one. The fundamental (second) mode is involved in resonant interaction with higher modes only in third-order calculations of the amplitude of the initial deformation of a drop, where four-mode resonances appear [8, 15]. Note that three-mode resonances arising in second-order calculations cause a first-order effect; i.e., the amplitude of a mode excited via energy exchange with higher modes is of the first order of smallness [13] and may exceed the amplitudes of initially excited higher modes. Therefore, a question of pure scientific interest that is pertinent to the theory of nonlinear interaction arises: Of which order of smallness will a mode be excited via resonant energy exchange upon four-mode interaction, which appears only in the third order of smallness? To tackle this question, we will solve the problem considered below.

(1) Consider a liquid drop with a radius R and charge Q in the absence of the environment and gravitational field. We consider an ideal incompressible perfectly conducting liquid with a density ρ and surface tension coefficient γ. Let the equilibrium spherical shape of the drop be subjected to a small-amplitude perturbation at zero time. Our aim is to trace the temporal evolution of the shape of the drop and to analyze its vibration under the action of capillary and electric forces. Considering that the flow inside the drop is caused by weak vibrations of its surface, one may assume that the flow is potential; i.e., the velocity field is characterized by a potential ψ. The electric field potential around the drop is denoted by ϕ. The shape of the drop is assumed to be axisymmetric for all time. In the dimensionless variables such that $\rho = 1$, $R = 1$, and $\gamma = 1$, the equation for the surface of the drop in the spherical coordinate system related to the drop's center of mass has the form

$$F(r, \vartheta, t) \equiv r - 1 - \xi(\vartheta, t) = 0, \qquad (1)$$

where r and ϑ are the spherical coordinates and $\xi(\vartheta, t)$ is the function describing the deviation of the shape from the spherical one ($|\xi(\vartheta, t)| \ll 1$).

The mathematical statement of the problem contains the Laplace equations for the liquid-velocity and electric-field potentials

$$\Delta\psi = 0, \quad \Delta\phi = 0; \qquad (2)$$

the boundedness conditions

$$r \longrightarrow 0\colon \nabla\psi \longrightarrow 0, \tag{3}$$

$$r \longrightarrow +\infty\colon \nabla\phi \longrightarrow 0; \tag{4}$$

the kinematic and dynamic boundary conditions

$$r = 1 + \xi(\vartheta, t)\colon -\frac{\partial\xi}{\partial t} + \nabla\psi\cdot\nabla F = 0, \tag{5}$$

$$\frac{\partial\psi}{\partial t} + \frac{1}{2}(\nabla\psi)^2 = p + p_q - p_{\text{at}} - p_\sigma; \tag{6}$$

the constancy condition for the volume of the drop

$$\int_V r^2\sin\vartheta\,dr\,d\vartheta\,d\varphi = \frac{4\pi}{3}, \tag{7}$$

$$V = \{r, \vartheta, \varphi | 0 \le r \le 1 + \xi;\ 0 \le \varphi \le \pi;\ 0 \le \varphi \le 2\pi\};$$

the stationarity condition for the center of mass

$$\int_V \mathbf{r} r^2\sin\vartheta\,dr\,d\vartheta\,d\varphi = 0; \tag{8}$$

the constancy condition for the total charge

$$\int_S \mathbf{n}\cdot\nabla\varphi\,dS = -4\pi Q, \tag{9}$$

$$S = \{r, \vartheta, \varphi | r = 1 + \xi;\ 0 \le \vartheta \le \pi;\ 0 \le \varphi \le 2\pi\};$$

the constancy condition for the electric potential over the surface of the drop

$$r = 1 + \xi(\vartheta, t)\colon \phi = \phi_S(t); \tag{10}$$

and the initial conditions

$$t = 0\colon \xi = \xi_0 P_0(\cos\vartheta) + \xi_1 P_1(\cos\vartheta) + \varepsilon\sum_{k\in\Omega} h_k P_k(\cos\vartheta), \quad \frac{\partial\xi}{\partial t} = 0. \tag{11}$$

In expressions (2)–(11), p_{at}, p, p_q, and p_σ are the atmospheric pressure, hydrodynamic pressure in equilibrium, electric field pressure, and capillary pressure, respectively; $\mathbf{n}$ is the unit vector that is normal to the surface of the drop; ϕ_S is the electric potential of the drop; ε is the initial deformation amplitude, which is a small parameter of the problem; Ω is the spectrum of modes specifying the initial deformation; h_k is the partial contribution of a kth mode to the initial deformation ($\sum_{k\in\Omega} h_k \sim O(1)$); $P_k(\cos\vartheta)$ is the kth-order Legendre polynomial; ξ_0 and ξ_1 are quantities defined in such a way that integral conditions (7) and (8) are valid at the initial time instant; and Δ is the Laplacian.

For convenience, we supplement the set of constants h_k so that $h_k \equiv 0$ for any $k \notin \Omega$.

(2) We will solve boundary problem (2)–(11) by the method of many scales up to the third order of smallness in initial perturbation amplitude ε. To this end, we represent all desired values in the form of expansions in powers of ε and assume that they depend not merely on time but on various time scales $T_j = \varepsilon^j t$ ($j = 0, 1, 2$). In this case, the time derivative is expressed via derivatives with respect to the time scales T_j as follows:

$$\frac{\partial}{\partial t} = \frac{\partial}{\partial T_0} + \varepsilon\frac{\partial}{\partial T_1} + \varepsilon^2\frac{\partial}{\partial T_2}.$$

Substituting the expansions

$$\xi = \varepsilon\xi^{(1)} + \varepsilon^2\xi^{(2)} + \varepsilon^3\xi^{(3)} + O(\varepsilon^4), \tag{12}$$

$$\psi = \varepsilon\psi^{(1)} + \varepsilon^2\psi^{(2)} + \varepsilon^3\psi^{(3)} + O(\varepsilon^4), \tag{13}$$

$$\phi = \phi^{(0)} + \varepsilon\phi^{(1)} + \varepsilon^2\phi^{(2)} + \varepsilon^3\phi^{(3)} + O(\varepsilon^4), \tag{14}$$

$$\phi_S = \phi_S^{(0)} + \varepsilon\phi_S^{(1)} + \varepsilon^2\phi_S^{(2)} + \varepsilon^3\phi_S^{(3)} + O(\varepsilon^4) \tag{15}$$

into boundary-value problem (2)–(11) and collecting terms with the same powers of ε, we come to the problems of different orders of smallness (see Appendix A). In expansions (14) and (15), $\phi^{(0)} = Q/r$ and $\phi_S^{(0)} = Q$ are zeroth-order solutions, which correspond to the equilibrium (spherical) surface of the drop.

Since the Laplace equations are linear, the functions $\psi^{(k)}$ and $\phi^{(k)}$ are obviously solutions to equations that are similar to (2). Subject to boundedness condition (3), (4), one may write

$$\psi^{(k)} = \sum_{n=1}^{\infty} r^n D_n^{(k)}(t) P_n(\cos\vartheta) \quad (k = 1, 2, 3), \tag{16}$$

$$\phi^{(k)} = \sum_{n=0}^{\infty} \frac{F_n^{(k)}(t)}{r^{n+1}} P_n(\cos\vartheta) \quad (k = 1, 2, 3). \tag{17}$$

The function describing the deviation of the surface from the spherical shape is represented by a similar expansion in Legendre polynomials:

$$\xi^{(k)} = \sum_{n=0}^{\infty} M_n^{(k)}(t) P_n(\cos\vartheta) \quad (k = 1, 2, 3). \tag{18}$$

Note that, when considering the problem in the third-order approximation, we can find the dependences of the first-order temporal coefficients in (16)–(18) on three time scales: $M_n^{(1)}(T_0, T_1, T_2)$, $F_n^{(1)}(T_0, T_1, T_2)$, and $D_n^{(1)}(T_0, T_1, T_2)$; the dependences of the second-order coefficients on two scales: $M_n^{(2)}(T_0, T_1)$, $F_n^{(2)}(T_0, T_1)$, and $D_n^{(2)}(T_0, T_1)$; and the dependences of the third-order coefficients only on T_0: $M_n^{(3)}(T_0)$, $F_n^{(3)}(T_0)$, and $D_n^{(3)}(T_0)$.

Sequentially using solutions (16)–(18) for $k = 1, 2, 3$ and the sets of first-, second-, and third-order bound-

ary conditions, we obtain differential equations for the coefficients $M_n^{(k)}(t)$, which characterize the temporal variation of the surface of the drop.

(3) Solving the first-order problem for the coefficients $M_n^{(1)}(t)$ (see Appendix A), we arrive at an equation harmonic in T_0:

$$\frac{\partial^2 M_n^{(1)}(t)}{\partial T_0^2} + \omega_n^2 M_n^{(1)}(t) = 0, \tag{19}$$

where $\omega_n^2 = n(n-1)(n+2-W)$ is the natural frequency of an nth mode of the surface vibrations and $W = Q^2/(4\pi)$ is the Rayleigh parameter characterizing the stability of the drop against the self-charge.

A general solution to Eq. (19) includes arbitrary functions: either one complex function or two real ones, which depend on the time scales T_1 and T_2:

$$\begin{aligned} M_n^{(1)}(t) &= A_n^{(1)}(T_1, T_2)\exp[i\omega_n T_0] + \text{c.c.} \\ &= 2a_n^{(1)}(T_1, T_2)\cos(\omega_n T_0 + b_n^{(1)}(T_1, T_2)) \end{aligned} \tag{20}$$

(hereafter c.c. stands for the complex conjugates to the preceding terms). Here, $A_n^{(1)}(T_1, T_2) = a_n^{(1)}(T_1, T_2)\exp[ib_n^{(1)}(T_1, T_2)]$ are the complex amplitudes and $a_n^{(1)}(T_1, T_2)$ and $b_n^{(1)}(T_1, T_2)$ are the real functions characterizing the amplitude and phase of vibrations. The form of the functions $A_n^{(1)}(T_1, T_2)$, $a_n^{(1)}(T_1, T_2)$, and $b_n^{(1)}(T_1, T_2)$ is found by solving the higher order problems.

(4) Considering the second-order problem (see Appendix A), we obtain an inhomogeneous differential equation for the evolutionary coefficients $M_n^{(2)}(t)$:

$$\begin{aligned} &\frac{\partial^2 M_n^{(2)}(t)}{\partial T_0^2} + \omega_n^2 M_n^{(2)}(t) \\ &= -2i\omega_n \frac{\partial A_n^{(1)}(T_1, T_2)}{\partial T_1} + \exp[i\omega_n T_0] \\ &+ \sum_{k=2}^{\infty}\sum_{m=2}^{\infty}\{[\gamma_{kmn} + \omega_k\omega_m\eta_{kmn}]A_k^{(1)}(T_1, T_2)A_m^{(1)}(T_1, T_2) \\ &\times \exp[i(\omega_k + \omega_m)T_0] + [\gamma_{kmn} - \omega_k\omega_m\eta_{kmn}]A_k^{(1)}(T_1, T_2) \\ &\times A_m^{(1)}(T_1, T_2)\exp[i(\omega_k - \omega_m)T_0] + \text{c.c.}\}. \end{aligned} \tag{21}$$

The constants γ_{kmn} and η_{kmn} are given in Appendix B.

In order that a solution to Eq. (21) be free of secular terms, the terms that depend on the time T_0 in the form $\exp[i\omega_n T_0]$ must be eliminated from its right-hand side. This will allow us to find the dependences of the functions $A_n^{(1)}(T_1, T_2)$ (or $a_n^{(1)}(T_1, T_2)$ and $b_n^{(1)}(T_1, T_2)$) on the time scale T_1. In the simplest case, we have

$$\frac{\partial A_n^{(1)}(T_1, T_2)}{\partial T_1} = 0, \tag{22}$$

which means that $A_n^{(1)}$, $a_n^{(1)}$, and $b_n^{(1)}$ are independent of T_1.

The careful examination of the inhomogeneity function for Eq. (21) shows that, if either of the two relationships $\omega_n = \omega_p \pm \omega_q$ is fulfilled for three modes of capillary vibrations with the numbers n, p, and q, the requirements of eliminating secular terms from solutions to similar equations (written for the modes n, p, and q) are reduced to a set of three coupled differential equations that define the dependence of the interrelated functions $A_n^{(1)}(T_1, T_2)$, $A_p^{(1)}(T_1, T_2)$, and $A_q^{(1)}(T_1, T_2)$ on the time scale T_1. In this case, it is customary to speak of internal three-mode resonant interaction between capillary vibrations of the drop, which was considered in [12, 13].

A general solution to Eq. (21) also includes either one complex, $A_n^{(2)}$, or two real, $a_n^{(2)}$ and $b_n^{(2)}$, arbitrary functions; however, these functions are dependent only on the time scale T_1. If three-mode resonant interactions are absent, the solution to Eq. (21) for vibration modes ($n > 2$) has the form

$$\begin{aligned} M_n^{(2)}(t) &= A_n^{(2)}(T_1)\exp[i\omega_n T_0] \\ &+ \sum_{k=2}^{\infty}\sum_{m=2}^{\infty}\{\lambda_{kmn}^{(+)}A_k^{(1)}A_m^{(1)}\exp[i(\omega_k + \omega_m)T_0] \\ &+ \lambda_{kmn}^{(-)}A_k^{(1)}\overline{A_m^{(1)}}\exp[i(\omega_k - \omega_m)T_0] + \text{c.c.}\}. \end{aligned} \tag{23}$$

Expressions for the constants $\lambda_{kmn}^{(+)}$ and $\lambda_{kmn}^{(-)}$ are given in Appendix B. The form of the functions $A_n^{(2)}(T_1)$, $a_n^{(2)}(T_1)$, and $b_n^{(2)}(T_1)$, where $A_n^{(2)}(T_1) = a_n^{(2)}(T_1)\exp[ib_n^{(2)}(T_1)]$, can be determined only by solving the third-order problem.

(5) Let us analyze in greater detail an inhomogeneous differential equation for the evolutionary coefficients $M_n^{(3)}(t)$, which follows from the set of third-order boundary conditions (see Appendix A):

$$\begin{aligned} &\frac{\partial^2 M_n^{(3)}(t)}{\partial T_0^2} + \omega_n^2 M_n^{(3)}t = -\left\{2i\omega_n\left[\frac{\partial A_n^{(2)}}{\partial T_1} + \frac{\partial A_n^{(1)}}{\partial T_2}\right]\right. \\ &\left. + G_n A_n^{(1)}\right\}\exp[i\omega_n T_0] + \sum_{k,g=2}^{\infty}\{H_{kgn}^{0(+)}\exp[i(\omega_k + \omega_g)T_0] \end{aligned}$$

$$\times (A_k^{(1)} A_g^{(2)}) + H_{kgn}^{0(-)} \exp[i(\omega_k - \omega_g)T_0](A_k^{(1)} \overline{A_g^{(2)}})\}$$

$$+ \sum_{k=2}^{\infty} \frac{1}{(2k+1)} \{[2(n-1)\omega_n\omega_k - \Xi_n](A_k^{(1)})^2$$

$$\times \exp[i(\omega_n + 2\omega_k)T_0] - (1-\delta_{n,k})[2(n-1)\omega_n\omega_k + \Xi_n]$$

$$\times (\overline{A_k^{(1)}})^2 \exp[i(\omega_n - 2\omega_k)T_0]\} A_n^{(1)}$$

$$+ D_{k,m}^{l,n}[\delta_{m,l+1}(\delta_{k,n-1} + \delta_{k,n+1})\chi_l \beta_{k,m,1,l,n}^{1(-)} + D_{k,n}^{l,m} H_{k,m,l,n}^{1(-)(+)}]$$

$$\times \exp[i\Psi_{k,l,m}^{(+)(-)} T_0](A_l^{(1)} \overline{A_m^{(1)}}) \quad (24)$$

$$+ D_{k,m}^{l,n} D_{k,l}^{m,n}[\delta_{m,l+1}(\delta_{k,n-1} + \delta_{k,n+1})\chi_l \beta_{k,m,1,l,n}^{2(+)} + H_{k,m,l,n}^{2(+)(+)}]$$

$$\times \exp[i\Psi_{k,l,m}^{(-)(-)} T_0](\overline{A_l^{(1)} \cdot A_m^{(1)}})$$

$$+ D_{k,l}^{m,n}[\delta_{m,l+1}(\delta_{k,n-1} + \delta_{k,n+1})\chi_l \beta_{k,m,1,l,n}^{2(-)} + D_{k,n}^{l,m} H_{k,m,l,n}^{2(-)(-)}]$$

$$\times \exp[i\Psi_{k,l,m}^{(-)(+)} T_0](\overline{A_l^{(1)} A_m^{(1)}}) A_k^{(1)},$$

where $\Psi_{k,l,m}^{(\pm)(\pm)} \equiv \omega_k \pm \omega_l \pm \omega_m$ and $\delta_{i,j}$ is the Kronecker delta.

Expressions for the coefficients used in (24) are given in Appendix B. For brevity, complex conjugates on the right-hand side of (24) are omitted.

By analogy, the requirement of eliminating secular terms from a solution to Eq. (24) allows one to determine the form of the functions $A_n^{(1)}(T_2)$ and $A_n^{(2)}(T_1)$. In the simplest case when resonant interactions between vibration modes are absent, this requirement has the form

$$2i\omega_n\left[\frac{\partial A_n^{(2)}(T_1)}{\partial T_1} + \frac{\partial A_n^{(1)}(T_2)}{\partial T_2}\right] + G_n A_n^{(1)}(T_2) \equiv 0;$$

hence, it readily follows that

$$b_n^{(1)}(T_2) = \frac{G_n}{2\omega_n} T_2, \quad (25)$$

whereas $a_n^{(1)}$ is independent of the time T_2 and $a_n^{(2)}$ and $b_n^{(2)}$ are independent of the time T_1. Expression (25) defines second-order corrections to the natural frequencies ω_n of the capillary vibrations of the drop (see (20)).

With initial conditions (11) met, a solution to Eq. (24) can be written in the form (provided that the components giving rise to secular terms are eliminated from its right-hand side)

$$M_n^{(3)}(t) = -\sum_{ki \in \Xi}^{\infty} \frac{h_k^2 h_n}{16} \left\{ \frac{[2(n-1)\omega_n\omega_k - \Xi_n]}{(2k+1)\omega_k(\omega_n + \omega_k)} \right.$$

$$\times [\cos((\omega_n + 2\omega_k)t) - \cos(\omega_n t)] + (1-\delta_{n,k})$$

$$\left. \times \frac{[2(n-1)\omega_n\omega_k + \Xi_n]}{(2k+1)\omega_k(\omega_n - \omega_k)} [\cos((\omega_n - 2\omega_k)t) - \cos(\omega_n t)] \right\}$$

$$+ \sum_{\substack{g=2 \\ k,l,m \in \Xi}}^{\infty} \frac{h_l h_m h_k}{4} (\lambda_{l,m,g}^{(+)} + \lambda_{l,m,g}^{(-)}) \left\{ \frac{H_{kgn}^{0(+)}}{[\omega_n^2 - (\omega_k + \omega_g)^2]} \right.$$

$$\times [\cos((\omega_k + \omega_g)t) - \cos(\omega_n t)] + \frac{H_{kgn}^{0(-)}}{[\omega_n^2 - (\omega_k - \omega_g)^2]}$$

$$\times [\cos((\omega_k - \omega_g)t) - \cos(\omega_n t)] + \sum_{k,l,m \in \Xi}^{\infty} \frac{h_l h_m h_k}{4}$$

$$\times \left\{ \frac{[\delta_{m,l+1}(\delta_{k,n-1} + \delta_{k,n+1})\chi_l \beta_{k,m,1,l,n}^{1(+)} + H_{k,m,l,n}^{1(+)(-)}]}{[\omega_n^2 - (\Psi_{k,l,m}^{(+)(+)})^2]} \right.$$

$$\times [\cos(\Psi_{k,l,m}^{(+)(+)} t) - \cos(\omega_n t)]$$

$$+ \frac{D_{k,m}^{l,n}[\delta_{m,l+1}(\delta_{k,n-1} + \delta_{k,n+1})\chi_l \beta_{k,m,1,l,n}^{1(-)} + D_{k,n}^{l,m} H_{k,m,l,n}^{1(-)(+)}]}{[\omega_n^2 - (\Psi_{k,l,m}^{(+)(-)})^2]}$$

$$\times [\cos(\Psi_{k,l,m}^{(+)(-)} t) - \cos(\omega_n t)]$$

$$+ \frac{D_{k,m}^{l,n} D_{k,l}^{m,n}[\delta_{m,l+1}(\delta_{k,n-1} + \delta_{k,n+1})\chi_l \beta_{k,m,1,l,n}^{2(+)} + H_{k,m,l,n}^{2(+)(+)}]}{[\omega_n^2 - (\Psi_{k,l,m}^{(+)(-)})^2]}$$

$$\times [\cos(\Psi_{k,l,m}^{(-)(-)} t) - \cos(\omega_n t)]$$

$$+ \frac{D_{k,l}^{m,n}[\delta_{m,l+1}(\delta_{k,n-1} + \delta_{k,n+1})\chi_l \beta_{k,m,1,l,n}^{2(-)} + D_{k,n}^{l,m} H_{k,m,l,n}^{2(-)(-)}]}{[\omega_n^2 - (\Psi_{k,l,m}^{(-)(+)})^2]}$$

$$\left. \times [\cos(\Psi_{k,l,m}^{(-)(+)} t) - \cos(\omega_n t)] \right\}.$$

Analyzing the form of the inhomogeneity function for Eq. (24), one easily checks that, along with three-mode resonant interaction (which appears in the second-order problem (see (21)), four-mode resonant interaction is also possible. It arises when the natural frequencies of modes with numbers n, p, q, and s satisfy either of the two relationships $\omega_p \pm \omega_q - \omega_s = \omega_n$ (see the triple sum in the inhomogeneity function for Eq. (24)). It may also happen that one of the modes takes part in resonant interaction twice (degenerate resonance). Moreover, in the third-order approximation considered, three-mode resonant interaction may also arise such that first-order modes, which are responsible for the initial deformation spectrum, exchange energy with second-order modes (see the double sum in the inhomogeneity function for Eq. (24)). Interactions of such types have not been found in the previous third-order calculations [2].

(6) Let us consider four-mode interaction more closely. To reflect the fact that the frequency combinations $\omega_p \pm \omega_q - \omega_s$ is close to the frequency ω_n, we introduce a mismatch parameter $\sigma \sim O(1)$ given by

$$\omega_p \pm \omega_q - \omega_s = \omega_n(1 + \varepsilon^2\sigma). \tag{26}$$

Supplementing (24) by similar equations for modes with the numbers p, q, and s and eliminating components responsible for secular terms in solutions from their right-hand sides, we derive a set of coupled differential equations for $A_j^{(i)}$ (where $i = 1, 2$; $j = n, p, q, s$). By way of example, we show such a set for the first of resonance situations (26) $\omega_p + \omega_q - \omega_s = \omega_n(1 + \varepsilon^2\sigma)$:

$$\begin{aligned}
-2i\omega_n\frac{\partial A_n^{(2)}(T_1)}{\partial T_1} &= 2i\omega_n\frac{\partial A_n^{(1)}(T_2)}{\partial T_2} + G_n(T_2)A_n^{(1)}(T_2)\\
&\quad - Y_n^{(+)}A_p^{(1)}(T_2)A_q^{(1)}(T_2)\overline{A_s^{(1)}(T_2)}\exp[i\omega_n\sigma T_2],\\
-2i\omega_p\frac{\partial A_p^{(2)}(T_1)}{\partial T_1} &= 2i\omega_p\frac{\partial A_p^{(1)}(T_2)}{\partial T_2} + G_p(T_2)A_p^{(1)}(T_2)\\
&\quad - Y_p^{(+)}A_n^{(1)}(T_2)\overline{A_q^{(1)}(T_2)}A_s^{(1)}(T_2)\exp[-i\omega_n\sigma T_2],\\
-2i\omega_q\frac{\partial A_q^{(2)}(T_1)}{\partial T_1} &= 2i\omega_q\frac{\partial A_q^{(1)}(T_2)}{\partial T_2} + G_q(T_2)A_q^{(1)}(T_2)\\
&\quad - Y_q^{(+)}A_n^{(1)}(T_2)\overline{A_p^{(1)}(T_2)}A_s^{(1)}(T_2)\exp[-i\omega_n\sigma T_2],\\
-2i\omega_s\frac{\partial A_s^{(2)}(T_1)}{\partial T_1} &= 2i\omega_s\frac{\partial A_s^{(1)}(T_2)}{\partial T_2} + G_s(T_2)A_s^{(1)}(T_2)\\
&\quad - Y_s^{(+)}\overline{A_n^{(1)}(T_2)}A_p^{(1)}(T_2)A_q^{(1)}(T_2)\exp[i\omega_n\sigma T_2]
\end{aligned} \tag{27}$$

(for the notation used, see Appendix B). For the second resonance situation ($\omega_p - \omega_q - \omega_s = \omega_n(1 + \varepsilon^2\sigma)$), the set of equations is similar to (27).

Set (27) must be supplemented by the requirements of eliminating secular terms from solutions to differential equations for the second-order amplitudes of modes with the numbers n, p, q, and s (see (21)). Assume that these modes are involved in none of the resonances other than those given by (26). This means that, in the second-order approximation, expressions like (22) are valid for the functions $A_n^{(1)}$, $A_p^{(1)}$, $A_q^{(1)}$, and $A_s^{(1)}$, so that these functions are independent of the time T_1. As a result, we find that the left-hand and right-hand sides of Eqs. (27) are functions only of T_1 and only of T_2, respectively. Since T_1 and T_2 in the method of many scales are independent variables, the left-hand and right-hand sides of Eqs. (27) should be set equal to a constant, for example, to zero.

For the functions $A_j^{(2)}$ (where $j = n, p, q, s$), we obtain

$$\frac{\partial A_j^{(2)}(T_1)}{\partial T_1} = 0;$$

hence, $A_j^{(2)}$, $a_j^{(2)}$, and $b_j^{(2)}$ are constants and equal their initial values, which are readily found from (11) and (12) in view of (18), (20), and (23):

$$a_j^{(2)} = -\frac{1}{4}\sum_{k\in\Omega}\sum_{m\in\Omega}(\lambda_{kmj}^{(+)} + \lambda_{kmj}^{(-)})h_k h_m; \quad b_j^{(2)} = 0.$$

Eventually, expression (23) for the second-order amplitudes takes the form

$$\begin{aligned}
M_n^{(2)}(t) = \sum_{k=2}^{\infty}\sum_{m=2}^{\infty}\{&\lambda_{kmn}^{(+)}[\cos((\omega_k + \omega_m)t) - \cos(\omega_n t)]\\
&+ \lambda_{kmn}^{(-)}[\cos((\omega_k - \omega_m)t) - \cos(\omega_n t)]\}\frac{h_k h_m}{2}.
\end{aligned} \tag{28}$$

For the functions $A_j^{(1)}$ (where $j = n, p, q, s$), we obtain complex equations by equating the real and imaginary parts to zero. For the determination of the functions $a_j^{(1)}$ and $b_j^{(1)}$ ($j = n, p, q, s$), we may write

$$\begin{aligned}
&a_n^{(1)}(T_2)\left(2\omega_n\left(\frac{\partial\beta_n^{(1)}(T_2)}{\partial T_2} - \omega_n\sigma\right) + G_n(T_2)\right)\\
&\quad - Y_n^{(\pm)}a_p^{(1)}(T_2)a_q^{(1)}(T_2)a_s^{(1)}(T_2)\cos[\varphi_{n,s,p,q}^{(-)(+)(\pm)}] = 0,\\
&2\omega_n\frac{\partial a_n^{(1)}(T_2)}{\partial T_2} - Y_n^{(\pm)}a_p^{(1)}(T_2)a_q^{(1)}(T_2)a_s^{(1)}(T_2)\\
&\quad\times\sin[\varphi_{n,s,p,q}^{(-)(+)(\pm)}] = 0,\\
&a_p^{(1)}(T_2)\left(2\omega_p\frac{\partial b_p^{(1)}(T_2)}{\partial T_2} - G_p(T_2)\right) + Y_p^{(\pm)}a_n^{(1)}(T_2)\\
&\quad\times a_q^{(1)}(T_2)a_s^{(1)}(T_2)\cos[\varphi_{n,s,p,q}^{(-)(+)(\pm)}] = 0,
\end{aligned}$$

$$2\omega_p\frac{\partial a_p^{(1)}(T_2)}{\partial T_2} + Y_p^{(\pm)}a_n^{(1)}(T_2)a_q^{(1)}(T_2)a_s^{(1)}(T_2) \times \sin[\varphi_{n,s,p,q}^{(-)(+)(\pm)}] = 0,$$

$$a_q^{(1)}(T_2)\left(2\omega_q\frac{\partial b_q^{(1)}(T_2)}{\partial T_2} - G_q(T_2)\right) + Y_q^{(\pm)}a_n^{(1)}(T_2) \times a_s^{(1)}(T_2)a_p^{(1)}(T_2)\cos[\varphi_{n,s,p,q}^{(-)(+)(\pm)}] = 0, \quad (29)$$

$$2\omega_q\frac{\partial a_q^{(1)}(T_2)}{\partial T_2} \pm Y_q^{(\pm)}a_n^{(1)}(T_2)a_s^{(1)}(T_2)a_p^{(1)}(T_2) \times \sin[\varphi_{n,s,p,q}^{(-)(+)(\pm)}] = 0,$$

$$a_s^{(1)}(T_2)\left(2\omega_s\frac{\partial b_s^{(1)}(T_2)}{\partial T_2} - G_s(T_2)\right) + Y_s^{(\pm)}a_n^{(1)}(T_2) \times a_q^{(1)}(T_2)a_p^{(1)}(T_2)\cos[\varphi_{n,s,p,q}^{(-)(+)(\pm)}] = 0,$$

$$2\omega_s\frac{\partial a_s^{(1)}(T_2)}{\partial T_2} - Y_s^{(\pm)}a_n^{(1)}(T_2)a_q^{(1)}(T_2)a_p^{(1)}(T_2) \times \sin[\varphi_{n,s,p,q}^{(-)(+)(\pm)}] = 0,$$

$$\beta_n^{(1)}(T_2) = \omega_n\sigma T_2 - b_n^{(1)}(T_2);$$

$$\varphi_{n,s,p,q}^{(-)(+)(\pm)}(T_2) \equiv \beta_n^{(1)}(T_2) - b_s^{(1)}(T_2) + b_p^{(1)}(T_2) \pm b_q^{(1)}(T_2).$$

The initial conditions for set (29) are also easily obtained from initial conditions (11) subject to (12), (18), and (20). The form of Eqs. (29) suggests that four-mode resonance may show up only if the amplitudes of at least three of interacting modes are other than zero at zero time. As an example, consider the situation where modes with the numbers p, q, and s are present in the spectrum specifying the initial deformation of the drop, while the nth mode is excited via mode interaction (that is, $p, q, s \in \Omega$; $n \notin \Omega$). In this case, the set of Eqs. (29) must be supplemented by the initial conditions

$$a_n^{(1)}(0) = 0; \quad \beta_n^{(1)}(0) = \pm\frac{\pi}{2}; \quad a_j^{(1)}(0) = \frac{h_j}{2}; \quad b_j^{(1)}(0) = 0 \quad (j = p, q, s). \quad (30)$$

Subject to initial conditions (30), solutions to set (29) yield the first-order amplitudes $M_j^{(1)}(t)$ (see (20)) as functions of the slow time scale $T_2 = \varepsilon^2 t$ for modes taking part in resonant interaction ($j = p, q, s, n$).

(7) Figure 1 shows the results of numerical calculations for the resonance situation $\omega_{17} + \omega_{21} - \omega_{30} = \omega_2$ with $W = 0.460245$ ($W = Q^2/4\pi\gamma R^3$ is the dimensionless parameter that characterizes the self-charge of the drop). It was assumed that the initial perturbation is specified by the 17th, 21st, and 30th modes, whose partial contributions to the amplitude of this perturbation ($\varepsilon = 0.1$) equal each other ($h_{17} = h_{21} = h_{30} = 1/3$). Since

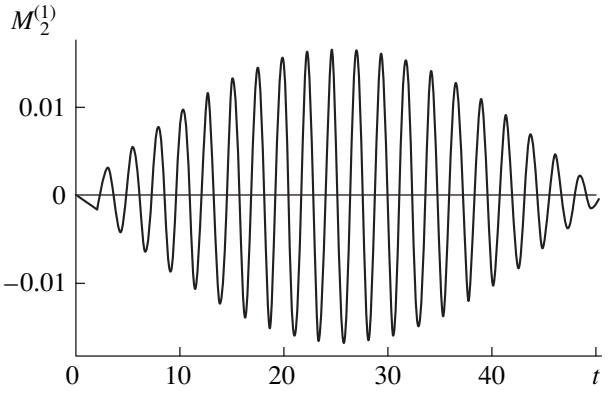

Fig. 1. Temporal dependence of the first-order evolutionary coefficient in the expansion in initial perturbation amplitude for the excited fundamental (second) mode of capillary vibrations of the drop. The parameter W exactly equals the resonant value $W = 0.46$. $\varepsilon = 0.1$ and $h_{17} = h_{21} = h_{30} = 1/3$.

the growth of the mode missing in the initial perturbation spectrum is of greatest interest, Fig. 1 and all other figures show the results obtained only for the second (fundamental) mode. From the plots presented, it follows that the first-order evolutionary coefficient $M_2^{(1)}(t)$ (see expansions (12) and (18)) of the mode excited via four-mode resonant interaction is of minor value (it is one order of magnitude smaller than the corresponding amplitudes of the 17th, 21st, and 30th modes) and does not exceed the second-order quantities. An increase in the relative amplitude ε of the initial perturbation merely decreases the period of resonant interaction without affecting the amplitude $M_2^{(1)}(t)$ (the calculation results for $\varepsilon = 0.3$ are shown in Fig. 2).

It is natural to suppose that the initial deformation of the surface of the drop is actually governed by a wider spectrum of modes (not only by the 17th, 21st, and 30th modes); then, the partial contribution from the modes of interest will decrease. Figure 3 shows the results obtained for $h_{17} = h_{21} = h_{30} = 1/12$ and $\varepsilon = 0.1$. Expectedly, the decrease in the partial contribution of resonantly interacting modes leads to a proportional decrease in the amplitude of the fundamental mode excited. Simultaneously, the period of resonant interaction extends considerably.

A change in the charge of the drop (in the parameter W) increases the frequency mismatch parameter in relationship (26), that is, deteriorates the conditions for resonant energy transfer from higher modes to the lowest (fundamental) mode. Figures 4 and 5 show the curves calculated for $W = 0$ and 0.87, respectively, i.e., for W lower and higher than the resonant value. The mismatch parameters in both cases are practically the same in magnitude but opposite in sign. It is easy to see that the change in sign of the charge causes the amplitude of the resonantly excited mode to grow and the resonant interaction period to extend. Note that, when the charge increases, the amplitude of the fundamental mode

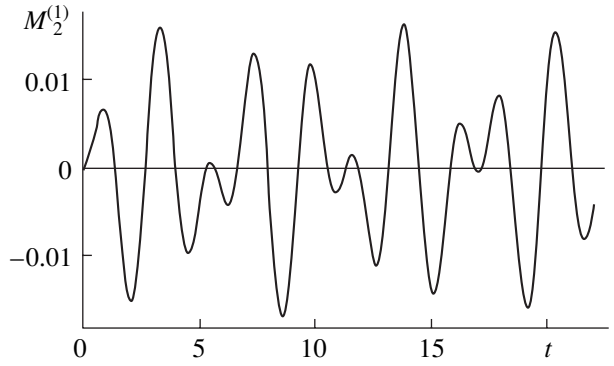

Fig. 2. Same as in Fig. 1 for ε = 0.3.

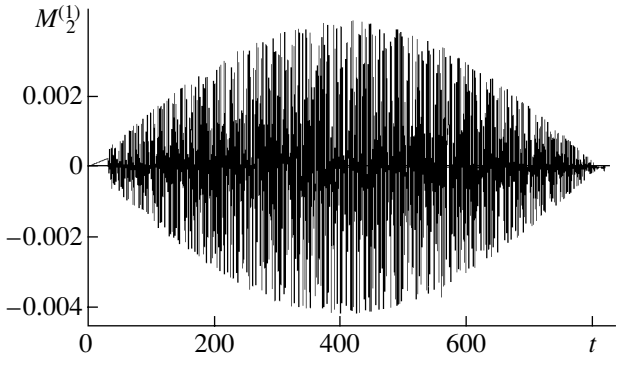

Fig. 3. Same as in Fig. 1 for $h_{17} = h_{21} = h_{30} = 1/12$.

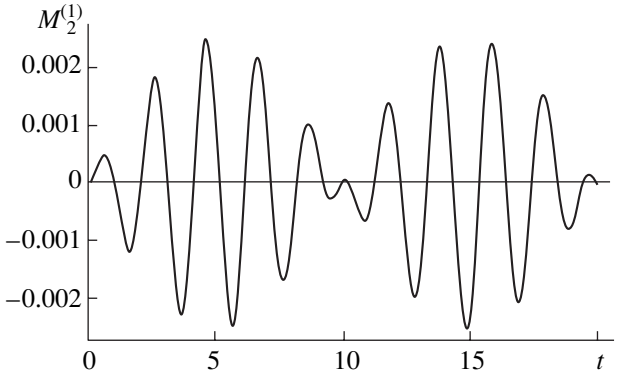

Fig. 4. Same as in Fig. 1 for $W = 0$.

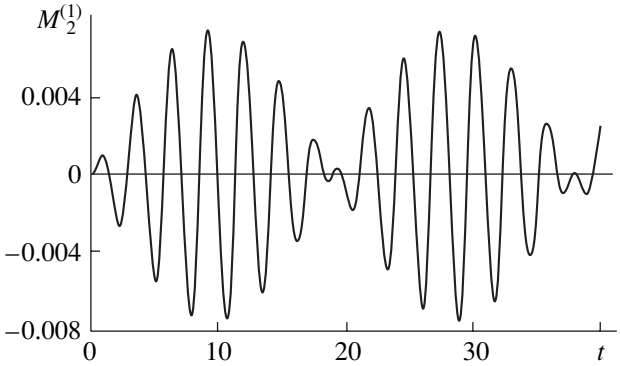

Fig. 5. Same as in Fig. 1 for $W = 0.87$.

diminishes to a considerably lesser extent, because an increase in the charge under normal (off-resonance) conditions builds up vibration modes.

Numerical calculations were also carried out for the second four-mode resonant situation $\omega_p - \omega_q - \omega_s = \omega_n(1 + \varepsilon^2\sigma)$ (see (26)), which takes place, for example, for the 34th, 30th, 10th, and 2nd modes with $W = 0.983454$. The results obtained for this situation coincide completely with those presented in Fig. 1 and are therefore omitted.

Calculation of the second-order correction $M_2^{(2)}(t)$ to the fundamental mode amplitude (see (12), (18), and (28)) by conventional methods from the theory of nonlinear vibration [1–6] shows that this correction (arising from nonresonant mode interaction) is comparable to $M_2^{(1)}(t)$. This is because the expression for the second-order correction $M_n^{(2)}(t)$ to the amplitude of an nth mode involves the coefficients

$$\lambda_{kmn}^{(-)} \sim \frac{1}{(\omega_n - \omega_k + \omega_m)(\omega_n + \omega_k - \omega_m)},$$

where the subscripts k and m run over the numbers of modes from the initial perturbation spectrum. It is obvious that, when k and m take the same values, $\lambda_{kkn}^{(-)} \sim 1/\omega_n^2$. Since the frequency of the second mode is much lower than possible frequencies of vibration modes, the coefficients $\lambda_{kk2}^{(-)}$ and, consequently, the corrections $M_2^{(2)}(t)$ turn out to be considerably greater than corrections $M_n^{(2)}(t)$ to higher modes. As a consequence, the contribution of the nonresonant second-order correction $\varepsilon^2 M_2^{(2)}(t)$ to the amplitude of the fundamental mode is comparable to the contribution from the resonance-induced first-order evolutionary coefficient ($\varepsilon M_2^{(1)}(t)$). This circumstance, together with the requirement that the asymptotic expansion of the amplitude of the fundamental mode excited be uniform, bounds the small parameter ε from above.

CONCLUSIONS

The asymptotic calculation of nonlinear capillary vibrations of a charged drop of an ideal incompressible liquid that is performed in the third-order approximation in amplitude of the multimode initial deformation reveals four-mode internal resonant interaction between modes, causing the fundamental mode to build up even if it is absent in the spectrum of initially excited modes. However, the amplitude of the fundamental

mode excited via resonant energy exchange with initially present higher modes, while formally being of the first order of smallness, does not exceed a second-order correction due to nonresonant nonlinear interaction. Therefore, the results obtained in this work could hardly shed light on the lightning initiation mechanism.

Also, in the third-order approximation, three-mode resonant interaction between the first-order amplitudes of initially excited modes with second-order corrections to the amplitudes arises.

APPENDIX A

BOUNDARY-VALUE PROBLEMS OF VARIOUS ORDERS OF SMALLNESS

Substituting expansions (12)–(15) into boundary-value problem (2)–(11) and collecting terms with the same powers of ε, we come to problems of various order of smallness. In the mathematics which follows, partial derivatives, for example, with respect to x are designated as ∂_x.

Separating out the terms with ε^1, we obtain the first-order problem stated as

$$\Delta\psi^{(1)} = 0; \quad \Delta\phi^{(1)} = 0;$$

$$r \longrightarrow 0: \psi^{(1)} \longrightarrow 0;$$

$$r \longrightarrow +\infty: \nabla\phi^{(1)} \longrightarrow 0;$$

$$r = 1: \partial_{T_0}\xi^{(1)} = \partial_r\psi^{(1)};$$

$$r = 1: \partial_{T_0}\psi^{(1)} = \frac{1}{4\pi}\partial_r\phi^{(0)}(\partial_r\phi^{(1)} + \xi^{(1)}\partial_{rr}\phi^{(0)}) + 2\xi^{(1)} + \Delta_\Omega\xi^{(1)};$$

$$\int_{-1}^{1}\xi^{(1)}d(\cos\vartheta) = 0; \quad \int_{-1}^{1}\xi^{(1)}P_1 d(\cos\vartheta) = 0;$$

$$r = 1: \int_{-1}^{1}\{\partial_r\phi^{(1)} + \xi^{(1)}(\partial_{rr}\phi^{(0)} + 2\partial_r\phi^{(0)})\}d(\cos\vartheta) = 0;$$

$$r = 1: \phi^{(1)} + \xi^{(1)}\partial_r\phi^{(0)} = \phi_S^{(1)}(t);$$

$$t = 0: \xi^{(1)} = \xi\sum_{k\in\Omega}h_k P_k(\cos\vartheta); \quad \partial_{T_0}\xi^{(1)} = 0.$$

The terms with ε^2 state the second-order problem:

$$\Delta\psi^{(2)} = 0; \quad \Delta\phi^{(2)} = 0;$$

$$r \longrightarrow 0: \psi^{(2)} \longrightarrow 0;$$

$$r \longrightarrow +\infty: \nabla\phi^{(2)} \longrightarrow 0;$$

$$r = 1: \partial_{T_0}\xi^{(2)} + \partial_{T_1}\xi^{(1)} = \partial_r\psi^{(2)} + \xi^{(1)}\partial_{rr}\psi^{(1)} - \partial_\vartheta\xi^{(1)}\partial_\vartheta\psi^{(1)};$$

$$r = 1: \partial_{T_0}\psi^{(2)} + \partial_{T_1}\psi^{(1)} + \xi^{(1)}\partial_{rT_0}\psi^{(1)} + \frac{1}{2}(\partial_r\psi^{(1)})^2 + \frac{1}{2}(\partial_\vartheta\psi^{(1)})^2 = \frac{1}{8\pi}\{2\xi^{(2)}\partial_r\phi^{(0)}\partial_{rr}\phi^{(0)} + (\xi^{(1)})^2 \times ((\partial_{rr}\phi^{(0)})^2 + \partial_{rrr}\phi^{(0)}\partial_r\phi^{(0)}) + (\partial_\vartheta\phi^{(1)})^2 + (\partial_r\phi^{(1)})^2 + 2\partial_r\phi^{(2)}\partial_r\phi^{(0)} + 2\xi^{(1)}(\partial_{rr}\phi^{(0)}\partial_r\phi^{(1)} + \partial_{rr}\phi^{(1)}\partial_r\phi^{(0)})\} + 2\xi^{(2)} + \Delta_\Omega\xi^{(2)} - 2(\xi^{(1)})^2 - 2\xi^{(1)}\Delta_\Omega\xi^{(1)};$$

$$\int_{-1}^{1}(\xi^{(2)} + (\xi^{(1)})^2)d(\cos\vartheta) = 0;$$

$$\int_{-1}^{1}(2\xi^{(2)} + 3(\xi^{(1)})^2)P_1 d(\cos\vartheta) = 0;$$

$$r = 1: \int_{-1}^{1}\left\{\partial_r\phi^{(2)} + \xi^{(1)}(\partial_{rr}\phi^{(1)} + 2\partial_r\phi^{(1)}) + \xi^{(2)}(\partial_{rr}\phi^{(0)} + 2\partial_r\phi^{(0)}) + (\xi^{(1)})^2\left(\frac{1}{2}\partial_{rrr}\phi^{(0)} + 2\partial_{rr}\phi^{(0)} + \partial_r\phi^{(0)}\right) - \partial_\vartheta\xi^{(1)}\partial_\vartheta\phi^{(1)}\right\}d(\cos\vartheta) = 0;$$

$$r = 1: \phi^{(2)} + \xi^{(1)}\partial_r\phi^{(1)} + \xi^{(2)}\partial_r\phi^{(0)} + \frac{1}{2}(\xi^{(1)})^2\partial_{rr}\phi^{(0)} = \phi_S^{(2)}(t);$$

$$t = 0: \xi^{(2)} = -\sum_{k\in\Omega}\frac{h_k P_0(\cos\vartheta)}{2k+1} - \frac{3}{2}\sum_{k,m\in\Omega}h_k h_m K_{km1}P_1(\cos\vartheta);$$

$$\partial_{T_0}\xi^{(2)} + \partial_{T_1}\xi^{(1)} = 0.$$

The third-order problem is defined by the terms with ε^3:

$$\Delta\psi^{(3)} = 0; \quad \Delta\phi^{(3)} = 0;$$

$$r \longrightarrow 0: \psi^{(3)} \longrightarrow 0;$$

$$r \longrightarrow +\infty: \nabla\phi^{(3)} \longrightarrow 0;$$

$$r = 1: \partial_{T_0}\xi^{(3)} + \partial_{T_1}\xi^{(2)} + \partial_{T_2}\xi^{(1)} = \partial_r\psi^{(3)} - \partial_\vartheta\xi^{(2)}\partial_\vartheta\psi^{(1)} - \partial_\vartheta\xi^{(1)}\partial_\vartheta\psi^{(2)} + \xi^{(2)}\partial_{rr}\psi^{(1)} + \xi^{(1)}(\partial_\vartheta\xi^{(1)}(2\partial_\vartheta\psi^{(1)} - \partial_{r\vartheta}\psi^{(1)}) + \partial_{rr}\psi^{(2)}) + \frac{1}{2}(\xi^{(1)})^2\partial_{rrr}\psi^{(1)};$$

$$r = 1: \partial_{T_0}\psi^{(3)} + \partial_{T_2}\psi^{(1)} + \partial_{T_1}\psi^{(2)} + \xi^{(1)}\partial_{rT_1}\psi^{(1)}$$
$$+ \partial_\vartheta\psi^{(1)}\partial_\vartheta\psi^{(2)} + \partial_r\psi^{(1)}\partial_r\psi^{(2)} + \xi^{(2)}\partial_{rT_0}\psi^{(1)} + \xi^{(1)}$$
$$\times(\partial_{rT_0}\psi^{(2)} + \partial_\vartheta\psi^{(1)}(\partial_{r\vartheta}\psi^{(1)} - \partial_\vartheta\psi^{(1)}) + \partial_r\psi^{(1)}\partial_{rr}\psi^{(1)})$$
$$+ \frac{1}{2}(\xi^{(1)})^2\partial_{rrT_0}\psi^{(1)} = \frac{1}{8\pi}\{2\xi^{(3)}\partial_r\phi^{(0)}\partial_{rr}\phi^{(0)} + (\xi^{(1)})^3$$
$$\times\left(\partial_{rr}\phi^{(0)}\partial_{rrr}\phi^{(0)} + \frac{1}{3}\partial_r\phi^{(0)}\partial_{rrrr}\phi^0\right) + 2(\partial_\vartheta\phi^{(1)}\partial_\vartheta\phi^{(2)}$$
$$+ \partial_r\phi^{(1)}(\xi^{(2)}\partial_{rr}\phi^{(0)} + \partial_r\phi^{(2)}) + \partial_r\phi^{(0)}\partial_r\phi^{(3)}$$
$$+ \xi^{(2)}\partial_r\phi^{(0)}\partial_{rr}\phi^{(1)}) + 2\xi^{(1)}(\xi^{(2)}((\partial_{rr}\phi^{(0)})^2$$
$$+ \partial_r\phi^{(0)}\partial_{rrr}\phi^{(0)}) + \partial_{rr}\phi^{(0)}\partial_r\phi^{(2)} + \partial_\vartheta\phi^{(1)}$$
$$\times(\partial_{r\vartheta}\phi^{(1)} - \partial_\vartheta\phi^{(1)}) + \partial_r\phi^{(1)}\partial_{rr}\phi^{(1)} + \partial_r\phi^{(0)}\partial_{rr}\phi^{(2)})$$
$$+ (\xi^{(1)})^2(\partial_{rrr}\phi^{(0)}\partial_r\phi^{(1)} + 2\partial_{rr}\phi^{(0)}\partial_{rr}\phi^{(1)}$$
$$+ \partial_r\phi^{(0)}\partial_{rrr}\phi^{(1)})\} + (2 + \Delta_\Omega)\xi^{(3)} + 2\xi^{(1)}((\xi^{(1)})^2$$
$$- (2 + \Delta_\Omega)\xi^{(2)}) - 2\xi^{(2)}\Delta_\Omega\xi^{(1)} + 3(\xi^{(1)})^2\Delta_\Omega\xi^{(1)}$$
$$- (\partial_\vartheta\xi^{(1)})^2\partial_{\vartheta\vartheta}\xi^{(1)} - \frac{1}{2}(\partial_\vartheta\xi^{(1)})^2\Delta_\Omega\xi^{(1)};$$

$$\int_{-1}^{1}(3\xi^{(3)} + 6\xi^{(1)}\xi^{(2)} + (\xi^{(1)})^3)d(\cos\vartheta) = 0;$$

$$\int_{-1}^{1}(\xi^{(3)} + 3\xi^{(1)}\xi^{(2)} + (\xi^{(1)})^3)P_1(\cos\vartheta)d(\cos\vartheta) = 0;$$

$$r = 1: \int_{-1}^{1}\left\{\partial_r\phi^{(3)} + \xi^{(3)}(\partial_{rr}\phi^{(0)} + 2\partial_r\phi^{(0)})\right.$$
$$+ \xi^{(2)}(\partial_{rr}\phi^{(1)} + 2\partial_r\phi^{(1)}) + (\xi^{(1)})^3\left(\frac{1}{6}\partial_{rrrr}\phi^{(0)}\right.$$
$$\left.+ \partial_{rrr}\phi^{(0)} + \partial_{rr}\phi^{(0)}\right) + (\xi^{(1)})^2\left(\frac{1}{2}\partial_{rrr}\phi^{(1)} + 2\partial_{rr}\phi^{(1)}\right.$$
$$\left.+ \partial_r\phi^{(1)}\right) + \xi^{(1)}(\xi^{(2)}(\partial_{rrr}\phi^{(0)} + 4\partial_{rr}\phi^{(0)} + 2\partial_r\phi^{(0)})$$
$$+ 2\partial_r\phi^{(2)} + \partial_{rr}\phi^{(2)} - \partial_\vartheta\xi^{(1)}\partial_{r\vartheta}\phi^{(1)}) - \partial_\vartheta\xi^{(2)}\partial_\vartheta\phi^{(1)}$$
$$\left.- \partial_\vartheta\xi^{(1)}\partial_\vartheta\phi^{(2)}\right\}d(\cos\vartheta) = 0;$$

$$r = 1: \phi^{(3)} + \xi^{(1)}\partial_r\phi^{(2)} + \xi^{(2)}\partial_r\phi^{(1)} + \xi^{(3)}\partial_r\phi^{(0)}$$
$$+ \frac{1}{2}(\xi^{(1)})^2\partial_{rr}\phi^{(1)} + \xi^{(1)}\xi^{(2)}\partial_{rr}\phi^{(0)}$$
$$+ \frac{1}{6}(\xi^{(1)})^3\partial_{rrr}\phi^{(0)} = \phi_S^{(3)}(t);$$

$$t = 0: \xi^{(3)} = -\sum_{k,m,l\in\Omega}\frac{h_k h_m h_l}{3(2l+1)}K_{kml}P_0(\cos\vartheta)$$
$$-\left(\frac{9}{5}h_2\sum_{k,m\in\Omega}h_k h_m K_{km1} + \sum_{g=0}^{\infty}\sum_{k,m,l\in\Omega}h_k h_m h_l K_{kmg}K_{gl1}\right)$$
$$\times P_1(\cos\vartheta);$$

$$t = 0: \partial_{T_0}\xi^{(3)} + \partial_{T_1}\xi^{(2)} + \partial_{T_2}\xi^{(1)} = 0.$$

Here, $K_{kmn} = (C^{n0}_{k0m0})^2$, where C^{n0}_{k0m0} are the Clebsch–Gordan coefficients.

APPENDIX B

NOTATION

$$\gamma_{kmn} = K_{kmn}[\omega_k^2(n-k+1) + 2n(m(m+1)-1)$$
$$+ (m(k+1) - k(2k-2n+7) + 3)nW/2]$$
$$+ \alpha_{kmn}[\omega_k^2/k + nW/2];$$

$$\eta_{kmn} = K_{kmn}(n/2 - k + 1) + \alpha_{kmn}(1 + n/(2m))/k;$$

$$K_{kmn} = (C^{n0}_{k0m0})^2;$$

$$\alpha_{kmn} = -C^{n0}_{k0m0}C^{n0}_{k(-1)m1}\sqrt{k(k+1)m(m+1)};$$

$$\lambda^{(\pm)}_{kmn} = (\gamma_{kmn} \pm \omega_k\omega_m\eta_{kmn})/(\omega_n^2 - (\omega_k \pm \omega_m)^2);$$

$$H^{0(+)}_{kgn} = (\Pi^0_{kgn} - \Pi^1_{kgn}\omega_k\omega_g - \Pi^2_{kgn}\omega_g^2)(\lambda^{+}_{kkg} + \lambda^{(-)}_{kkg});$$

$$H^{0(-)}_{kgn} = (\Pi^0_{kgn} + \Pi^1_{kgn}\omega_k\omega_g - \Pi^2_{kgn}\omega_g^2)(\lambda^{+}_{kkg} + \lambda^{(-)}_{kkg});$$

$$\Pi^0_{kgn} = (\omega_k^2(n-k+1) + 2n((k-1)(k+2) + g(g+1))$$
$$+ nW(3 - k(3-n+k) - g(3-n-k+g)))K_{kgn}$$
$$+ (\omega_k^2/k + nW)\alpha_{kgn};$$

$$\Pi^1_{kgn} = (g + k - n - 2)K_{kgn} - (n + k + g)\alpha_{kgn}/(gk);$$

$$\Pi^2_{kgn} = (g - n - 1)K_{kgn} - \alpha_{kgn}/g;$$

$$\Xi_n = 3(\omega_n^2 - n(n-1)W); \quad \chi_l = -\frac{9(l+1)}{(2l+1)(2l+3)};$$

$$\beta^{1(+)}_{kmgln} = \Pi^0_{kgn} - \Pi^1_{kgn}\omega_k(\omega_l + \omega_m) - \Pi^2_{kgn}(\omega_l + \omega_m)^2;$$

$$\beta^{1(-)}_{kmgln} = \Pi^0_{kgn} - \Pi^1_{kgn}\omega_k(\omega_l - \omega_m) - \Pi^2_{kgn}(\omega_l - \omega_m)^2;$$

$$\beta^{2(+)}_{kmgln} = \Pi^0_{kgn} + \Pi^1_{kgn}\omega_k(\omega_l + \omega_m) - \Pi^2_{kgn}(\omega_l + \omega_m)^2;$$

$$\beta^{2(-)}_{kmgln} = \Pi^0_{kgn} + \Pi^1_{kgn}\omega_k(\omega_l - \omega_m) - \Pi^2_{kgn}(\omega_l - \omega_m)^2;$$

$$H^{1(+)(-)}_{kmln} = \sum_{g=2}^{\infty}\beta^{1(+)}_{kmgln}\lambda^{(+)}_{lmg} + \sum_{g=1}^{\infty}\mu^{1(-)}_{kmgln} + \sum_{g=0}^{\infty}\mu^{0(-)}_{kmgln};$$

$$H^{1(-)(+)}_{kmln} = \sum_{g=2}^{\infty}\beta^{1(-)}_{kmgln}\lambda^{(-)}_{lmg} + \sum_{g=1}^{\infty}\mu^{1(+)}_{kmgln} + \sum_{g=0}^{\infty}\mu^{0(+)}_{kmgln};$$

$$H^{2(+)(+)}_{kmln} = \sum_{g=2}^{\infty}\beta^{2(+)}_{kmgln}\lambda^{(+)}_{lmg} + \sum_{g=1}^{\infty}\mu^{1(+)}_{kmgln} + \sum_{g=0}^{\infty}\mu^{0(+)}_{kmgln};$$

$$H^{2(-)(-)}_{kmln} = \sum_{g=2}^{\infty}\beta^{2(-)}_{kmgln}\lambda^{(-)}_{lmg} + \sum_{g=1}^{\infty}\mu^{1(-)}_{kmgln} + \sum_{g=0}^{\infty}\mu^{0(-)}_{kmgln};$$

$$\mu^{1(-)}_{kmgln} = \Lambda^{1}_{kmgln} - \Gamma^{1}_{kmgln}\omega_m\omega_k;$$

$$\mu^{1(+)}_{kmgln} = \Lambda^{1}_{kmgln} + \Gamma^{1}_{kmgln}\omega_m\omega_k;$$

$$\mu^{0(-)}_{kmgln} = \Lambda^{0}_{kmgln} - \Gamma^{0}_{kmgln}\omega_m\omega_k;$$

$$\mu^{0(+)}_{kmgln} = \Lambda^{1}_{kmgln} + \Gamma^{0}_{kmgln}\omega_m\omega_k;$$

$$\Lambda^{0}_{kmgln} = \frac{1}{2k}\Bigg\{K_{gln}[\alpha_{kmg}(2(k-2)\omega_k^2 - kn(2(k+2)W - l(3l+1))) + K_{kmg}(kn(4-6k(k+1) + (k^3 - 2(m+1)(m+2) - k^2(n-9) - k(3n+2m(m+3)-22))W) - (k-1)\times k(k-n-2)\omega_k^2)] - 2kn\alpha_{kmg}\times\sum_{\nu=1}^{[l/2]}(2l-4\nu+1)K_{g,l-2\nu,n}\Bigg\};$$

$$\Lambda^{1}_{kmgln} = ((g-n-1)K_{gln} - \alpha_{gln}/g)((m-1)K_{kmg} - \alpha_{kmg}/m)\omega_m^2 + Wnk((g+1)(l+n-g-2)K_{gln} + \alpha_{gln})K_{kmg};$$

$$\Gamma^{0}_{kmgln} = ((k-1)(k-2(n+1))K_{kmg}/2 - ((k-1)\times(m+n)-m)\alpha_{kmg}/(km))K_{gln} + ((k-1)\times(k-2)K_{kln}/2 - (k-2)\alpha_{kln}/k)K_{gmn};$$

$$\Gamma^{1}_{kmgln} = -((g-n-1)K_{gkn} - (n+k)\alpha_{gkn}/(kg))\times((m-1)K_{lmg} - \alpha_{lmg}/m) - ((g-n-1)K_{gln} - \alpha_{gln}/g)((m-1)K_{kmg} - \alpha_{kmg}/m);$$

$$G_n \equiv \sum_{k=2}^{\infty}\Bigg\{\frac{1}{(2k+1)}[2\Xi_n + \delta_{n,k}(2(k-1)\omega_k^2 + \Xi_k)] - (\delta_{k,n-1} + \delta_{k,n+1})\chi_{k-\delta_{k,n+1}}[\beta^{2(+)}_{k,k,1,n,n} + \beta^{1(-)}_{k,k,1,n,n}] - \sum_{g=0}^{\infty}\{(1-\delta_{g,0})(1-\delta_{g,1})[\beta^{2(+)}_{k,k,g,n,n}(\lambda^{(+)}_{kng} + (1-\delta_{k,n})\times\lambda^{(+)}_{nkg}) + \beta^{1(-)}_{k,k,g,n,n}(\lambda^{(-)}_{kng} + \lambda^{(-)}_{nkg}) + 2(1-\delta_{k,n})\Pi^{0}_{ngn}\lambda^{(-)}_{kkg}] + (2-\delta_{k,n})(1-\delta_{g,0})Z^{1}_{k,g,n} + Z^{0}_{k,g,n}]\}\Bigg\}(A^{(1)}_k\overline{A^{(1)}_k});$$

$$Z^{i}_{k,g,n} \equiv \mu^{i(+)}_{k,k,g,n,n} + \Lambda^{i}_{n,k,g,k,n} + \Lambda^{i}_{k,n,g,k,n} \quad (i = 0, 1);$$

$$D^{l,m}_{k,n} \equiv 1 - \delta_{k,n}\delta_{l,m};$$

$$Y^{(+)}_n = D^{p,s}_{q,n}(\delta_{p,n-1} + \delta_{p,n+1})(\delta_{s,q+1}\chi_q + \delta_{q,s+1}\chi_s)\beta^{1(-)}_{p,s,1,q,n} + D^{q,s}_{p,n}(\delta_{q,n-1} + \delta_{q,n+1})(\delta_{s,p+1}\chi_p + \delta_{p,s+1}\chi_s)\beta^{1(-)}_{q,s,1,p,n} + D^{p,s}_{q,n}D^{q,s}_{p,n}[(\delta_{s,n-1} + \delta_{s,n+1})(\delta_{q,p+1}\chi_p + \delta_{p,q+1}\chi_q)\beta^{2(+)}_{s,p,1,q,n} + H^{1(-)(+)}_{p,s,q,n} + H^{1(-)(+)}_{q,s,p,n} + H^{2(-)(-)}_{p,q,s,n} + H^{2(-)(-)}_{q,p,s,n} + H^{2(+)(+)}_{s,q,p,n} + H^{2(+)(+)}_{s,p,q,n}];$$

$$Y^{(-)}_n = D^{p,s}_{q,n}(\delta_{s,n-1} + \delta_{s,n+1})(\delta_{p,q+1}\chi_q + \delta_{q,p+1}\chi_p)\beta^{1(-)}_{s,p,1,q,n} + D^{q,p}_{s,n}(\delta_{q,n-1} + \delta_{q,n+1})(\delta_{p,s+1}\chi_s + \delta_{s,p+1}\chi_p)\beta^{1(-)}_{q,p,1,s,n} + D^{p,s}_{q,n}D^{q,p}_{s,n}[(\delta_{p,n-1} + \delta_{p,n+1})(\delta_{q,s+1}\chi_s + \delta_{s,q+1}\chi_q)\beta^{2(+)}_{p,s,1,q,n} + H^{1(-)(+)}_{s,p,q,n} + H^{1(-)(+)}_{q,p,s,n} + H^{2(-)(-)}_{s,q,p,n} + H^{2(-)(-)}_{q,s,p,n} + H^{2(+)(+)}_{p,q,s,n} + H^{2(+)(+)}_{p,s,q,n}];$$

$$Y^{(+)}_p = D^{p,s}_{q,n}(\delta_{n,p-1} + \delta_{n,p+1})(\delta_{q,s+1}\chi_s + \delta_{s,q+1}\chi_q)\beta^{1(-)}_{n,q,1,s,p} + D^{q,s}_{p,n}(\delta_{s,p-1} + \delta_{s,p+1})(\delta_{q,n+1}\chi_n + \delta_{n,q+1}\chi_q)\beta^{1(-)}_{s,q,1,n,p} + D^{q,s}_{p,n}D^{p,s}_{q,n}[(\delta_{q,p-1} + \delta_{q,p+1})(\delta_{s,n+1}\chi_n + \delta_{n,s+1}\chi_s)\beta^{2(+)}_{q,s,1,n,p} + H^{1(-)(+)}_{n,q,s,p} + H^{1(-)(+)}_{s,q,n,p} + H^{2(-)(-)}_{n,s,q,p} + H^{2(-)(-)}_{s,n,q,p} + H^{2(+)(+)}_{q,s,n,p} + H^{2(+)(+)}_{q,n,s,p}];$$

$$Y^{(-)}_p = (\delta_{n,p-1} + \delta_{n,p+1})(\delta_{q,s+1}\chi_s + \delta_{s,q+1}\chi_q)\beta^{1(+)}_{n,q,1,s,p} + (\delta_{s,p-1} + \delta_{s,p+1})(\delta_{q,n+1}\chi_n + \delta_{n,q+1}\chi_q)\beta^{1(+)}_{s,q,1,n,p} + (\delta_{q,p-1} + \delta_{q,p+1})(\delta_{s,n+1}\chi_n + \delta_{n,s+1}\chi_s)\beta^{1(+)}_{q,s,1,n,p} + H^{1(+)(-)}_{n,q,s,p} + H^{1(+)(-)}_{s,q,n,p} + H^{1(+)(-)}_{n,s,q,p} + H^{1(+)(-)}_{s,n,q,p} + H^{1(+)(-)}_{q,s,n,p} + H^{1(+)(-)}_{q,n,s,p};$$

$$Y^{(+)}_q = Y^{(-)}_q = D^{q,s}_{p,n}(\delta_{n,q-1} + \delta_{n,q+1})(\delta_{p,s+1}\chi_s + \delta_{s,p+1}\chi_p)\times\beta^{1(-)}_{n,p,1,s,q} + D^{p,s}_{q,n}(\delta_{s,q-1} + \delta_{s,q+1})(\delta_{p,n+1}\chi_n + \delta_{n,p+1}\chi_p)\times\beta^{1(-)}_{s,p,1,n,q} + D^{q,s}_{p,n}D^{p,s}_{q,n}[(\delta_{p,q-1} + \delta_{p,q+1})(\delta_{s,n+1}\chi_n + \delta_{n,s+1}\chi_s)\beta^{2(+)}_{p,s,1,n,q} + H^{1(-)(+)}_{n,p,s,q} + H^{1(-)(+)}_{s,p,n,q} + H^{2(-)(-)}_{n,s,p,q} + H^{2(-)(-)}_{s,n,p,q} + H^{2(+)(+)}_{p,s,n,p} + H^{2(+)(+)}_{p,n,s,q}];$$

$$Y_s^{(+)} = D_{p,n}^{q,s}(\delta_{p,s-1} + \delta_{p,s+1})(\delta_{n,q+1}\chi_q + \delta_{q,n+1}\chi_n)$$
$$\times \beta_{p,n,1,q,s}^{1(-)} + D_{q,n}^{p,s}(\delta_{q,s-1} + \delta_{q,s+1})(\delta_{n,p+1}\chi_p + \delta_{p,n+1}\chi_n)$$
$$\times \beta_{q,n,1,p,s}^{1(-)} + D_{p,n}^{q,s}D_{q,n}^{p,s}(\delta_{n,s-1} + \delta_{n,s+1})(\delta_{q,p+1}\chi_p$$
$$+ \delta_{p,q+1}\chi_q)\beta_{n,p,1,q,s}^{2(+)} + H_{p,n,q,s}^{1(-)(+)} + H_{q,n,p,s}^{1(-)(+)} + H_{p,q,n,s}^{2(-)(-)}$$
$$+ H_{q,p,n,s}^{2(-)(-)} + H_{n,q,p,s}^{2(+)(+)} + H_{n,p,q,s}^{2(+)(+)}];$$

$$Y_s^{(-)} = D_{p,n}^{q,s}(\delta_{n,s-1} + \delta_{n,s+1})(\delta_{p,q+1}\chi_q + \delta_{q,p+1}\chi_p)$$
$$\times \beta_{n,p,1,q,s}^{1(-)} + D_{s,n}^{q,p}(\delta_{q,s-1} + \delta_{q,s+1})(\delta_{p,n+1}\chi_n + \delta_{n,p+1}\chi_p)$$
$$\times \beta_{q,p,1,n,p}^{1(-)} + D_{p,n}^{q,s}D_{s,n}^{q,p}[(\delta_{p,s-1} + \delta_{p,s+1})(\delta_{q,n+1}\chi_n$$
$$+ \delta_{n,q+1}\chi_q)\beta_{p,n,1,q,s}^{2(+)} + H_{n,p,q,s}^{1(-)(+)} + H_{q,p,n,s}^{1(-)(+)} + H_{n,q,p,s}^{2(-)(-)}$$
$$+ H_{q,n,p,s}^{2(-)(-)} + H_{p,q,n,s}^{2(+)(+)} + H_{p,n,q,s}^{2(+)(+)}].$$

ACKNOWLEDGMENTS

This work was supported by the Russian Foundation for Basic Research (grant no. 03-01-00760).

REFERENCES

1. J. A. Tsamopoulos and R. A. Brown, J. Fluid Mech. **127**, 519 (1983).
2. J. A. Tsamopoulos and R. A. Brown, J. Fluid Mech. **147**, 373 (1984).
3. D. F. Belonozhko and A. I. Grigor'ev, Zh. Tekh. Fiz. **70** (8), 45 (2000) [Tech. Phys. **45**, 1001 (2000)].
4. S. O. Shiryaeva, Pis'ma Zh. Tekh. Fiz. **26** (22), 76 (2000) [Tech. Phys. Lett. **26**, 1016 (2000)].
5. S. O. Shiryaeva, Zh. Tekh. Fiz. **71** (2), 27 (2001) [Tech. Phys. **46**, 158 (2001)].
6. S. O. Shiryaeva, Izv. Ross. Akad. Nauk, Mekh. Zhidk. Gaza, No. 3, 163 (2001).
7. S. O. Shiryaeva, Zh. Tekh. Fiz. **72** (4), 15 (2002) [Tech. Phys. **47**, 389 (2002)].
8. A. N. Zharov, A. I. Grigor'ev, and S. O. Shiryaeva, Zh. Tekh. Fiz. **73** (6), 36 (2003) [Tech. Phys. **48**, 697 (2003)].
9. V. A. Dyachuk and V. M. Muchnik, Dokl. Akad. Nauk SSSR **248**, 60 (1979).
10. A. I. Grigor'ev and S. O. Shiryaeva, Phys. Scr. **54**, 660 (1996).
11. *Clouds and Cloudy Atmosphere: A Handbook*, Ed. by I. P. Mazin, A. Kh. Khrgian, and I. M. Imyanitov (Gidrometeoizdat, Leningrad, 1989).
12. S. O. Shiryaeva, D. F. Belonozhko, and A. I. Grigor'ev, Pis'ma Zh. Tekh. Fiz. **28** (22), 45 (2002) [Tech. Phys. Lett. **28**, 945 (2002)].
13. S. O. Shiryaeva, Zh. Tekh. Fiz. **73** (2), 19 (2003) [Tech. Phys. **48**, 152 (2003)].
14. S. O. Shiryaeva, A. I. Grigor'ev, and D. F. Belonozhko, Pis'ma Zh. Tekh. Fiz. **29** (6), 69 (2003) [Tech. Phys. Lett. **29**, 249 (2003)].
15. A. N. Zharov, A. I. Grigor'ev, and S. O. Shiryaeva, Pis'ma Zh. Tekh. Fiz. **29** (9), 75 (2003) [Tech. Phys. Lett. **29**, 388 (2003)].

Translated by N. Mende

Technical Physics, Vol. 49, No. 1, 2004, pp. 83–86. Translated from Zhurnal Tekhnicheskoĭ Fiziki, Vol. 74, No. 1, 2004, pp. 84–86.

OPTICS, QUANTUM ELECTRONICS

Anisotropic Acoustooptic Modulator of Nonpolarized Light Based on Diffraction by a Slow Acoustic Wave in a Paratellurite Crystal

S. N. Antonov and V. I. Mirgorodsky

Institute of Radio Engineering and Electronics (Fryazino Branch), Russian Academy of Sciences, pl. Vvedenskogo 1, Fryazino, Moscow Oblast, 141196 Russia

e-mail: olga-ant@yandex.ru

Received April 22, 2003

Abstract—The problem of anisotropic Bragg diffraction of nonpolarized light by a slow acoustic wave in a TeO_2 crystal is solved. Two independent acoustic waves are excited in the crystal. Nonpolarized light splits in the crystal into two orthogonally polarized eigenmodes, either diffracting by its associated acoustic beam. Conditions under which the angles of incidence and diffraction are the same for both diffraction processes are found. Depending on the acoustic frequency, the diffracted light at the exit from the crystal may be represented either by a single nonpolarized beam or by two orthogonally polarized beams with different directions and orthogonal polarizations. This may provide a high diffraction efficiency (up to 100%) for nonpolarized light in a TeO_2 crystal. Theoretical calculations are supported by experiments. Modulators capable of controlling a high-power laser operating at a wavelength of 1.06 μm are fabricated. © *2004 MAIK "Nauka/Interperiodica".*

Single-crystalline paratellurite (TeO_2) is the most promising acoustooptic material for the implementation of diffraction by a slow acoustic wave. This is because it offers not only an extraordinarily high acoustooptic quality M_2 but also a unique combination of optical, acoustic, and engineering properties. Today, a high (up to 100%) efficiency of slow-acoustic-wave diffraction is achievable only in the case of anisotropic acoustooptic diffraction of polarized light. However, the most powerful industrial solid-state or fiber lasers emit, as a rule, either nonpolarized or randomly polarized light (in particular, at a wavelength of 1.06 μm).

Thus, our aim was to find an acoustooptic means of controlling (modulating) nonpolarized optical radiation under the condition of high-efficiency diffraction by a slow acoustic wave. The idea of the method is illustrated by vector diagrams in Fig. 1. Let nonpolarized light be incident on the crystal at an angle α to its optical axis. In the crystal, the light splits into two orthogonally polarized eigenmodes, which are characterized by wavevectors $\mathbf{K}_1$ and $\mathbf{K}_2$ and associated with the indicatrices n_o and n_e of the crystal. Also, let there exist a single direction β for diffracted beams with vectors $\mathbf{K}_1'$ and $\mathbf{K}_2'$. The incident and diffracted optical modes are coupled through two acoustic beams with vectors $\mathbf{q}_1$ and $\mathbf{q}_2$. The wavevectors of the acoustic waves make angles γ_1 and γ_2, respectively, with the [110] crystallographic direction. In the general statement, a solution to the problem must simultaneously meet the momentum conservation law for two independent diffraction processes

$$\mathbf{K}_1 + \mathbf{q}_1 = \mathbf{K}_1', \quad (1)$$

$$\mathbf{K}_2 + \mathbf{q}_2 = \mathbf{K}_2'. \quad (2)$$

We calculated the angles of incidence and diffraction versus the direction and frequency of the acoustic

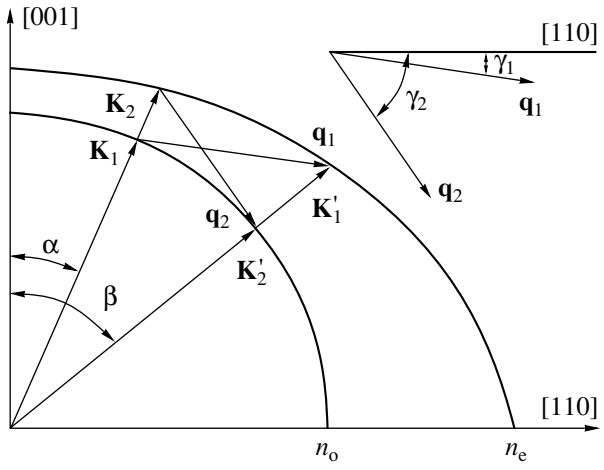

Fig. 1. Vector diagram of interaction between two optical eigenmodes diffracted by two acoustic waves and the angles these waves make with the [110] axis.

1063-7842/04/4901-0083$26.00 © 2004 MAIK "Nauka/Interperiodica"

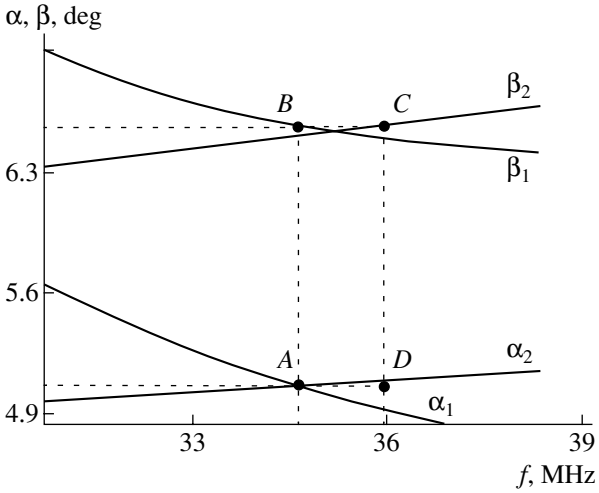

Fig. 2. Angles of incidence of orthogonally polarized optical eigenmodes in the crystal and their diffraction angles versus acoustic frequency.

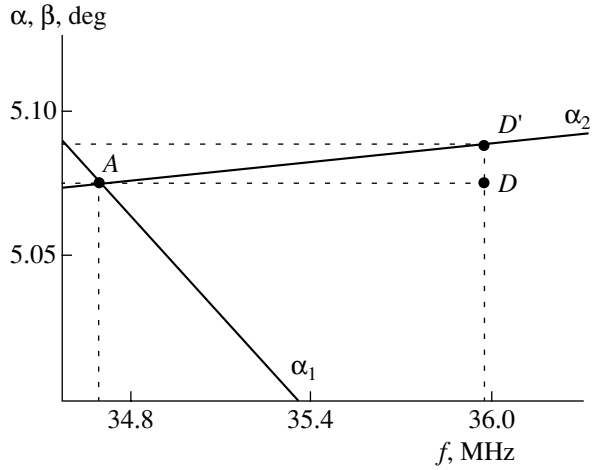

Fig. 3. Closer view of the vicinity of the points *A* and *D* that are shown in Fig. 2.

wave. The indicatrices of the refractive indices were approximated by ellipsoids of revolution [1]

$$\frac{n_1^2(\theta)\cos^2\theta}{n_o^2(1+\delta)^2} + \frac{n_1^2(\theta)\sin^2\theta}{n_e^2} = 1, \quad (3)$$

$$\frac{n_2^2(\theta)\cos^2\theta}{n_o^2(1-\delta)^2} + \frac{n_2^2(\theta)\sin^2\theta}{n_0^2} = 1, \quad (4)$$

where $n_1(\theta)$ is the refractive index of the extraordinary ray, $n_2(\theta)$ is the refractive index of the ordinary ray, θ is the angle relative to the optical axis, $\delta = \lambda_0\rho/2\pi n_0$ is the splitting factor, ρ is the specific rotation of polarization of light (rad/mm), and λ_0 is the wavelength of light in free space.

The calculations were performed for $\lambda_0 = 1.06$ μm, $\rho = 254.8$, $n_o = 2.208$, and $n_e = 2.352$. The velocity of an acoustic wave propagating in the (110) plane at a small angle to the normal to the optical axis was approximated [2] as $v = v_0(1 + b\varphi^2)$, where $b = 4.895$ is the elastic wave anisotropy parameter, φ is the angle between the phase velocity of sound and the normal to the optical axis, and $v_0 = 0.617 \times 10^6$ mm/s is the velocity of sound propagating perpendicularly to the crystal's optical axis.

The problem was solved for several practically important combinations of the angles γ_1 and γ_2. The choice of these angles and limitations imposed on them will be detailed below. A solution will be sought for the situation with $\gamma_1 = 4°$ and $\gamma_2 = 7°$, which was realized in practice. Figure 2 plots the angles of incidence α_1 and α_2 of the orthogonally polarized eigenmodes of the crystal and the respective diffraction angles β_1 and β_2 against the frequency of the acoustic wave. Note that the curves α_1 and β_1 refer to diffraction described by Eq. (1); the curves α_2 and β_2, to diffraction described by Eq. (2). The angles are counted from the optical axis, and Fig. 2 shows their values in the crystal. The points of interest in the plot are as follows. At the point *A*, the frequencies of both acoustic waves (the vectors $\mathbf{q}_1$ and $\mathbf{q}_2$) are roughly and the angles of incidence of the respective orthogonally polarized modes are exactly equal to each other, but the propagation directions of the diffracted beams do not coincide. At the points *B* and *C*, the diffraction angles of both optical modes are exactly the same. However, these angles become the same if the frequency of the acoustic wave with the wavevector $\mathbf{q}_2$ is shifted (by 1.326 MHz in this particular case). Figure 3 shows the closer view of the vicinity of the points *A* and *D*. It is seen that the frequency shift of the acoustic wave with the wave vector $\mathbf{q}_2$ changes the angle of incidence that corresponds to exact Bragg synchronism (*D* shifts to *D*'). Consequently, at a constant angle of incidence, a change in the frequency of this optical mode violates Bragg synchronism by the difference between the ordinates of the points *D*' and *D* (in our case, it is about 0.2 mrad). Such a small mismatch is however smaller than the divergence of the acoustic wave and does not degrade the diffraction efficiency. In particular, for this frequency and a piezoelectric transducer width of 10 mm, the acoustic wave divergence is 2 mrad. Note that such a weak dependence of the angle of incidence on the acoustic frequency makes it possible to use this acoustooptic diffraction mode in skew-angular anisotropic deflectors based on TeO_2 crystals [1, 2]. These deflectors rely on the possibility of significantly changing the acoustic frequency (and, hence, the diffraction angle) with the diffraction efficiency remaining almost unchanged.

We thus see that two acoustic beams in a TeO_2 crystal may be efficiently controlled (modulated) by nonpolarized light. However, one feature of such a light modulation technique should be noted. When light to be modulated is completely nonpolarized and a controlling optical beam does not have preferential polarization, the eigenmodes of the crystal are incoherent, so that their superposition at the exit of the crystal due to diffraction will also be completely nonpolarized. If, however, arbitrarily polarized light is applied to the crystal, the eigenmodes will be coherent. Since the light is diffracted by two acoustic waves with different frequencies, the polarization of the outgoing light, which is a superposition of two coherent but frequency-separated optical waves, is modulated in time. For most applications, the feature described above is of no significance. However, it should be taken into account in a number of cases.

To complete the discussion of the modulator's design, let us concentrate on limitations that are imposed on the acoustic wave propagation angles γ_1 and γ_2 in practice. The first limitation is associated with the fact that a slow acoustic wave traveling near the [110] axis is characterized by a high acoustic anisotropy. Therefore, a deflection of the acoustic wave phase front (the plane of the transducer) from this axis by a small angle γ causes a significant deflection of the energy vector by an angle φ. For the diffraction plane, in which the transducers are inclined in the case studied, the ratio φ/γ is approximately equal to 10 (for γ less than 7°). Consequently, a large value of γ is impracticable, because it requires the crystal to have large dimensions. The second limitation is imposed on the lower bound of the angle γ for the following reason. A TeO_2 crystal is optically active, with the activity being maximal along the optical axis and sharply decreasing away from it. For the acoustooptic interaction conditions considered, the eigenmodes of the crystal were shown [3] to become elliptically polarized (with the ellipticity being a function of the angle γ) when γ is smaller than 4°. Clearly, the polarizations of two optical eigenmodes due to diffraction by acoustic beams with different angles γ are not strictly orthogonal, while the basic idea of the method proposed relies on this orthogonality. Also, there exist limitations imposed on the frequency range, which is determined by the angles γ_1 and γ_2. In applications, the optimal values of the angles γ_1 and γ_2 apparently lie between 4° and 7°.

The table summarizes the frequency positions of the nodes *A*, *B*, *C*, *D*, and *D*' for several combinations of γ_1 and γ_2. The upper and lower figures are the frequency and angle of a node, respectively. It is seen that, in this range of γ_1 and γ_2, a modulator with an acoustic wave frequency of 24 to 50 MHz can be designed.

Figure 4 shows schematically the TeO_2 crystal. The "acoustic" end faces of crystal *1* were optically finished so that the planes of piezoelectric transducers *2* and *3* make angles of 7° and 4° to the [110] axis. Because of the acoustic anisotropy effect, the acoustic beams are deflected from the [110] axis by 60° and 40°, respectively. It is significant that, in such a geometry, there appears a region where the acoustic beams from the two piezoelectric transducers intersect. In this region, a parallax between the two beams at the exit of the crystal is absent upon acoustooptic interaction. In other words, the orthogonally polarized exit beams will be coincident in both the angular and coordinate spaces, which is undoubtedly important for applications.

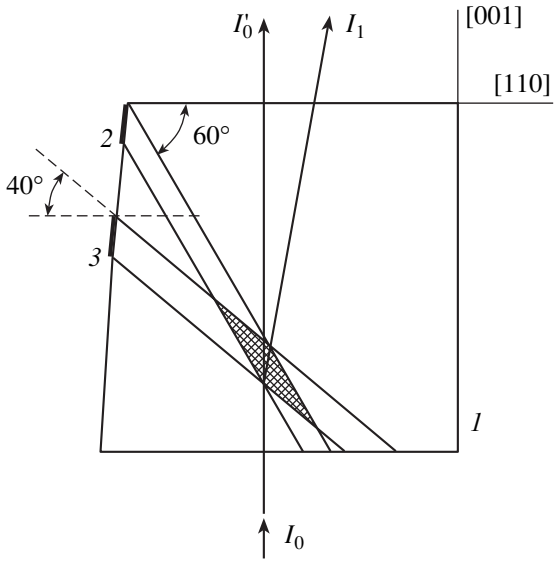

Fig. 4. Geometry of the TeO_2 crystal: I_0, incident optical beam; I_0', the zeroth diffraction order; and I_1, diffracted beam.

Table

γ_2, γ_1, deg	*A*, MHz; deg	*B*, MHz; deg	*C*, MHz; deg	*D*, MHz; deg	*D*', MHz; deg
4, 3	49.973, 2.616	49.973, 4.836	50.702, 4.836	50.702, 2.616	50.702, 2.606
5, 3	29.755, 3.591	29.755, 4.913	30.404, 4.913	30.404, 3.591	30.404, 3.593
6, 3	24.039, 4.236	24.039, 5.304	24.679, 5.304	24.679, 4.236	24.679, 4.249
6, 4	44.235, 4.348	44.235, 6.292	45.666, 6.292	45.666, 4.348	45.666, 4.341
7, 4	34.655, 5.075	34.655, 6.597	35.981, 6.597	35.981, 5.075	35.981, 5.0869

Our experiments confirmed the theoretical calculations. Several prototypes of acoustooptic modulators following the design shown in Fig. 4 were fabricated. The acoustic wave frequencies calculated differed from those providing equal diffraction angles in a particular crystal by no more than ±10%, which is apparently due to an inaccuracy in the orientation of the acoustic faces upon optical finishing. The modulators were used to control nonpolarized radiation from a high-power (25 W) 1.06-μm cw fiber laser. The overall diffraction efficiencies achieved in the experiments were no less than 85–90%.

REFERENCES

1. A. W. Warner, D. L. White, and W. A. Bonner, J. Appl. Phys. **43**, 4489 (1972).
2. V. P. Semenkov, Zh. Tekh. Fiz. **51**, 2090 (1981) [Sov. Phys. Tech. Phys. **26**, 1219 (1981)].
3. M. F. Bryzhina and S. Kh. Esayan, Zh. Tekh. Fiz. **47**, 1937 (1977) [Sov. Phys. Tech. Phys. **22**, 1123 (1977)].

Translated by A. Khzmalyan

Technical Physics, Vol. 49, No. 1, 2004, pp. 87–92. Translated from Zhurnal Tekhnicheskoĭ Fiziki, Vol. 74, No. 1, 2004, pp. 87–92.

RADIOPHYSICS

Influence of the Skin Effect on the Absorption of Electromagnetic Radiation by a Fine Metal Particle

I. O. Moiseev, A. A. Yushkanov, and Yu. I. Yalamov

Moscow Oblast State University, 107005 Moscow, Russia

e-mail: miosw@hotbox.ru

Received May 16, 2003

Abstract—The cross section of absorption of electromagnetic radiation by a fine spherical metal particle is calculated. The influence of the skin effect on the absorption cross-section is estimated for an arbitrary ratio between the free path and size of the particle. The results of this work are compared with those obtained earlier in the framework of classical electrodynamics. It is shown that taking into account the kinetic effects modifies essentially the known data for the skin effect in a spherical particle. © *2004 MAIK "Nauka/Interperiodica".*

INTRODUCTION

Fine metal particles (whose radius R is comparable to the free path Λ of an electron in a metal) are different from metal particulates in optical properties. For example, the absorption cross section of fine particles, which shows the nontrivial dependence on the ratio R/Λ, cannot be described by the equations of macroscopic electrodynamics [1, 2].

To date, the kinetic approach has been applied to fine particles whose radius R is much smaller than the skin depth δ. Accordingly, the skin effect was neglected [3–6].

To calculate the absorption cross-section of a fine spherical metal particle, we use here the method of moments to solve the kinetic equation for an electron in a metal. Variational moment boundary conditions for conduction electrons are formulated. The system of Maxwell equations for an electromagnetic field is solved jointly with the kinetic equation for electrons in a metal. This provides a consistent consideration of the influence of the skin effect on the absorption cross-section.

PROBLEM DEFINITION

Let a fine metal particle (FMP) be placed in the field of a plane electromagnetic wave $\mathbf{H} = \mathbf{H}_0\exp(-i\omega t)$. The radius of the particle is assumed to be much smaller than the wavelength; therefore, the nonuniformity of the external magnetic field H_0 is disregarded. The wave induces a vortex electric field inside the particle. If screening (skin effect) is absent, this vortex electric field has the form

$$\mathbf{E}_0 = \omega[\mathbf{H}_0 \times \mathbf{r}]/(2ic)\exp(-i\omega t),$$

where c is the velocity of light, H_0 is the magnetic field, and ω is the electromagnetic wave frequency.

The vortex electric field generates eddy currents inside the particle, causing it to absorb the electromagnetic energy. The electric field results in a deviation f_1 of the electron distribution function f from the Fermi equilibrium function [3, 4, 6]

$$f_0 = \begin{cases} 1, & 0 \le \varepsilon \le \varepsilon_F \\ 0, & \varepsilon_F < \varepsilon. \end{cases}$$

Here, $\varepsilon = mV^2/2$ is the energy of electrons with an effective mass m and velocity V and $\varepsilon_F = mV_F^2/2$ is the Fermi energy of electrons.

Let us introduce the spherical coordinate system such that its center coincides with the particle center and the polar axis is aligned with the magnetic field $\mathbf{H}_0$. It is assumed that the temperature is much lower than the degeneracy temperature of the electron gas and that the Fermi surface is of the spherical shape. The cross section σ of absorption of electromagnetic energy by a particle can be determined by the formula [7]

$$\sigma = \frac{1}{2}\left(\frac{8\pi}{cH_0^2}\right)\mathrm{Re}\int \mathbf{j}_\varphi \cdot \mathbf{E}_\varphi^* d^3r, \qquad (1)$$

where j_φ is the φth component of the eddy current inside the particle, E_φ^* is the complex conjugate to the φth component of the electric field inside the particle, and r is the running value of the radius.

The functions j_φ and E_φ^* are expressed through the function f_1. Consider the case of diffuse scattering of electrons by the particle surface [7]

$$f_1(\mathbf{r}, \mathbf{V}) = 0 \quad \text{at} \quad |r| = R \quad \text{and} \quad \mathbf{r} \cdot \mathbf{V} < 0. \qquad (2)$$

In an approximation linear in an external field, the function f_1 satisfies the Boltzmann kinetic equation for electrons where the collision integral is taken in the

relaxation time approximation [7, 8]:

$$-i\omega f_1 + \mathbf{V}\frac{\partial f_1}{\partial \mathbf{r}} + e\mathbf{V}\cdot\mathbf{E}\frac{\partial f_0}{\partial \varepsilon} = -\frac{f_1}{\tau}, \qquad (3)$$

where e is the electron charge and τ is the relaxation time.

MODEL AND CALCULATION

Owing to the symmetry of the problem, Eq. (3) is convenient to write in the spherical coordinates [9]:

$$\frac{\partial f_1}{\partial t} + V_r\frac{\partial f_1}{\partial r} + \frac{V_\varphi}{r\sin\Theta}\frac{\partial f_1}{\partial \varphi} + \frac{V_\Theta}{r}\frac{\partial f_1}{\partial \Theta}$$
$$+ \frac{V_\varphi^2 + V_\Theta^2}{r}\frac{\partial f_1}{\partial V_r} - \left(\frac{V_\varphi V_\Theta}{r}\cot\Theta + \frac{V_r V_\varphi}{r}\right)\frac{\partial f_1}{\partial V_\varphi} \qquad (4)$$
$$+ \left(\frac{V_\varphi^2}{r}\cot\Theta - \frac{V_r V_\Theta}{r}\right)\frac{\partial f_1}{\partial V_\Theta} + eV_\varphi E_\varphi\frac{\partial f_0}{\partial \varepsilon} = -\frac{f_1}{\tau},$$

where V_r, V_φ, and V_Θ are the velocity components along the respective spherical coordinates r, φ, and Θ.

Equation (4) is solved by the method of moments [9]. The function f_1 is represented as a combination of the moments C_r and C_rC_φ:

$$f_1 = \exp(-i\omega t)\delta(\varepsilon_f - \varepsilon)(a_1(r)C_\varphi + a_2(r)C_\varphi C_r)\sin\Theta,$$

where $\delta(\varepsilon_F - \varepsilon) = \delta(V - V_F)/(mV_F) = \delta(C-1)/mV_F^2$ is the Dirac delta function, $a_1(r)$ and $a_2(r)$ are the coefficients multiplying the moments, $C_r = V\cos\alpha/V_F$ and $C_\varphi = V\sin\alpha\cos\beta/V_F$ are the dimensionless velocity components, and α and β are angles in the velocity space.

The function f_1 can then be expressed as

$$f_1 = \frac{\exp(-i\omega t)}{mV_F^2}\delta(C-1)(a_1(r)C_\varphi + a_2(r)C_\varphi C_r)\sin\Theta.$$

We multiply Eq. (4) by C_φ and C_rC_φ in succession and integrate the result over the velocity space. After transformations, we obtain

$$10\nu a_1 + 6\frac{V_F}{r}a_2 + 2V_F\frac{\partial a_2}{\partial r} = 10eV_F\Psi_1,$$
$$\frac{V_F}{r}a_1 - V_F\frac{\partial a_1}{\partial r} - \nu a_2 = 0. \qquad (4')$$

Here, $\nu = 1/\tau - i\omega$, Ψ_1 is the electric field inside the particle (in the absence of screening, $\Psi_1^0 = i\omega H_0 r/(2c)$), and Ψ_2 is the electric field outside the particle. Let us introduce the dimensionless quantities

$$\Phi_1 = \frac{\Psi_1}{H_0};\quad \alpha_1 = \frac{a_1}{RH_0e};\quad \alpha_2 = \frac{a_2}{RH_0e};$$
$$\xi = \frac{r}{R}(0\le\xi\le 1);\quad x = \frac{R}{\tau V_F};\quad y = \frac{R\omega}{V_F}.$$

Substituting this notation into (4') and performing necessary transformations yields the system of equations

$$\begin{cases} 10z\alpha_1 + 6\dfrac{\alpha_1}{\xi} + 2\dfrac{\partial\alpha_2}{\partial\xi} = 10\Phi_1 \\ \dfrac{\alpha_1}{\xi} - \dfrac{\partial\alpha_1}{\partial\xi} - z\alpha_2 = 0. \end{cases} \qquad (5)$$

From this system, one can derive an equation for the function $\alpha_1(\xi)$:

$$\frac{\partial}{\partial\xi}\left(\xi^2\frac{\partial\alpha_1}{\partial\xi}\right) - 2\alpha_1 = 5z^2\xi^2\alpha_1 - 5\Phi_1 z\xi^2. \qquad (6)$$

Let the field outside the particle be designated as Ψ_2. Then, the Maxwell equations for the field Ψ_1 inside the particle and the field Ψ_2 outside the particle have the form

$$\begin{cases} \dfrac{\partial}{\partial r}\left(r^2\dfrac{\partial\Psi_1}{\partial r}\right) - 2\Psi_1 = i\omega\dfrac{4\pi}{c^2}j_\Phi r^2 \\ \dfrac{\partial}{\partial r}\left(r^2\dfrac{\partial\Psi_2}{\partial r}\right) - 2\Psi_2 = 0. \end{cases} \qquad (7)$$

Consider moment boundary conditions for the distribution function. The average power $\bar{Q}$ dissipated in the particle is given by

$$\frac{1}{2}\mathrm{Re}\int\mathbf{j}\cdot\mathbf{E}^* d^3r = 2\frac{m^3}{h^3}\left(\frac{1}{2\tau}\mathrm{Re}\int\delta(V-V_F)\frac{\varphi_1\varphi_1^*}{mV_F}d^3Vd^3r\right.$$
$$\left. + \frac{1}{4}\mathrm{Re}\int\frac{\delta(V-V_F)}{mV_F}V_r\varphi_1\varphi_1^* d^3VdS\right),$$

where the function φ_1 is defined as

$$f_1 = \frac{\partial f_0}{\partial\varepsilon}\varphi_1 = \frac{d(V-V_F)}{mV_F}\varphi_1.$$

The second term on the right of the equation for dissipated power can be represented as a sum of two terms that describe the energy fluxes transferred by electrons incident on the surface and reflected from it:

$$\frac{1}{4}\mathrm{Re}\int\frac{\delta(V-V_F)}{mV_F}V_F f_1 f_1^* VdS$$
$$= \frac{1}{4}\mathrm{Re}\int\frac{\delta(V-V_F)}{mV_F}V_F f_1 f_1^* d^3VdS$$
$$+ \frac{1}{4}\mathrm{Re}\int\frac{\delta(V-V_F)}{mV_F}V_F f_1 f_1^* d^3VdS.$$

For diffuse reflection of electrons from the surface (condition (2)), the second term on the right of this relationship vanishes. Accordingly, this term must be minimized. The quantities Ψ_1 and Ψ_2 on the particle surface

must satisfy the following boundary conditions:

$$\Psi_1\big|_{r=R} = \Psi_2\big|_{r=R};\ \left.\frac{\partial\Psi_1}{\partial r}\right|_{r=R} = \left.\frac{\partial\Psi_2}{\partial r}\right|_{r=R};\ a_1 = \frac{1}{\sqrt{3}}a_2;$$
$$\Phi_1\big|_{\xi=1} = \Phi_2\big|_{\xi=1};\ \left.\frac{\partial\Phi_1}{\partial\xi}\right|_{\xi=1} = \left.\frac{\partial\Phi_2}{\partial\xi}\right|_{\xi=1};\ \alpha_1 = \frac{1}{\sqrt{3}}\alpha_2. \quad (8)$$

Note that, with such boundary conditions, the fluxes are calculated accurate to within 4%. The current density j_φ can be represented as [9]

$$j_\varphi = 2\left(\frac{m}{h}\right)^3 e\int V_\varphi f_1 d^3V$$
$$= \frac{8em^2\pi V_F^2}{3h^3}a_1(r) = \frac{8e^2H_0Rm^2\pi V_F^2}{3h^3}\alpha_1(\xi). \quad (9)$$

Let us introduce the dimensionless quantity $w^2 = 32\pi^2e^2R^2m^2V_F^2/(3h^3c^2)$. In view of (9), system (7) of equations for the fields can be recast in the form

$$\begin{cases}\dfrac{\partial}{\partial\xi}\left(\xi^2\dfrac{\partial\Phi_1}{\partial\xi}\right) - 2\Phi_1 = -iyw^2\xi^2\alpha_1\\ \dfrac{\partial}{\partial\xi}\left(\xi^2\dfrac{\partial\Phi_2}{\partial\xi}\right) - 2\Phi_2 = 0.\end{cases} \quad (10)$$

As follows from the first equation of system (10), the influence of the conduction currents on the behavior of the field inside the particle at a given frequency is proportional to w^2. As w increases, interaction between the conduction currents and the variable electromagnetic field inside the particle is enhanced. Thus, this parameter characterizes the degree of influence of the skin effect on the absorption of the electromagnetic field by the particle.

Consider the behavior of the field and electrons inside the particle. It is described by the system of equations

$$\begin{cases}\dfrac{\partial}{\partial\xi}\left(\xi^2\dfrac{\partial\Phi_1}{\partial\xi}\right) - 2\Phi_1 = -iyw^2\xi^2\alpha_1\\ \dfrac{\partial}{\partial\xi}\left(\xi^2\dfrac{\partial\alpha_1}{\partial\xi}\right) - 2\alpha_2 = 5z^2\xi^2\alpha_1 - 5\Phi_1z\xi^2.\end{cases}$$

Let us make the substitutions $-iyw^2 = l_{21}$, $-5z = l_{11}$, and $5z^2 = l_{12}$. Then,

$$\begin{cases}\dfrac{\partial}{\partial\xi}\left(\xi^2\dfrac{\partial\Phi_1}{\partial\xi}\right) - 2\Phi_1 = l_{21}\xi^2\alpha_1\\ \dfrac{\partial}{\partial\xi}\left(\xi^2\dfrac{\partial\alpha_1}{\partial\xi}\right) - 2\alpha_2 = l_{12}\xi^2\alpha_1 + l_{11}\Phi_1\xi^2.\end{cases} \quad (11)$$

A solution to system Eq. (11) will be sought in the form $\Phi_1 = K\alpha_1$, where K is a coefficient. Let the operator

$$\xi^2\frac{\partial}{\partial\xi}\left(\xi^2\frac{\partial}{\partial\xi}\right)$$

be designated as Ω. Then, system (11) can be written as

$$\begin{cases}\Omega\Phi_1 = l_{21}\dfrac{\Phi_1}{K}\\ \Omega\Phi_1 = l_{12}K\Phi_1 + l_{11}\Phi_1.\end{cases}$$

Comparing the two above expressions, we find that $l_{21}/K = Kl_{11} + l_{12}$.

From this quadratic equation, we find the values of the coefficient K:

$$K_1 = \frac{-l_{12} + \sqrt{l_{12}^2 + 4l_{11}l_{21}}}{2l_{11}},$$

$$K_2 = \frac{-l_{12} - \sqrt{l_{12}^2 + 4l_{11}l_{21}}}{2l_{11}}.$$

Substituting the values of l_{11}, l_{12}, and l_{21} yields

$$K_1 = \frac{z}{2}\left(1 - \sqrt{1 + \frac{4iyw^2}{5z^3}}\right),$$

$$K_2 = \frac{z}{2}\left(1 + \sqrt{1 + \frac{4iyw^2}{5z^3}}\right).$$

Solving the first differential equation of system (11) and taking into account that K has two solutions and also that $\Phi_1 = K_{1,2}\alpha_1$ and $l_{21}/K_{1,2} = -\gamma_{1,2}^2$, we obtain

$$\frac{\partial}{\partial\xi}\left(\xi^2\frac{\partial\Phi_1}{\partial\xi}\right) - 2\Phi_1 = \frac{l_{21}}{K_{1,2}}\xi^2\Phi_1 = -\gamma_{1,2}^2\Phi_1\xi^2.$$

The general solution for Φ_1 has the form $\Phi_1 = \chi_1C_1 + \chi_2C_2$, where C_1 and C_2 are coefficients that can be found from boundary conditions (8) and

$$\chi_{1,2} = \frac{(\gamma_{1,2}\xi + i)\exp(i\gamma_{1,2}\xi) + (\gamma_{1,2}\xi - 1)\exp(-i\gamma_{1,2}\xi)}{\xi^2}.$$

Below, we will deal with the derivatives of χ_1 and χ_2. They are expressed as

$$\chi'_{1,2} = \frac{(i\gamma_{1,2}^2\xi^2 - 2\gamma_{1,2}\xi - 2i)\exp(i\gamma_{1,2}\xi)}{\xi^3}$$
$$+ \frac{(-i\gamma_{1,2}^2\xi^2 - 2\gamma_{1,2}\xi + 2i)\exp(-i\gamma_{1,2}\xi)}{\xi^3}.$$

Let us find the form of the coefficients C_1 and C_2 from boundary conditions (8). From the solution to the second differential equation in system (10) (the field outside the particle), we have $\Phi_2 = iyV_F\xi/(2c) + C_3/\xi^2$, where C_3 is a coefficient.

Since $\alpha_1 = \Phi_1/K_{1,2}$, one obtains

$$\alpha_1 = \frac{C_1}{K_1}\chi_1 + \frac{C_2}{K_2}\chi_2,$$

$$\alpha_2 = \frac{C_1}{zK_1}\left(\frac{\chi_1}{\xi} - \chi_1'\right) + \frac{C_2}{zK_2}\left(\frac{\chi_2}{\xi} - \chi_2'\right).$$

Substituting the expressions for Φ_1, Φ_2, α_1, and α_2 into boundary conditions (8) with $\xi = 1$ yields the set of equations

$$\begin{cases} C_1\chi_1 + C_2\chi_2 = \dfrac{iyV_F}{2c} + C_3 \\ C_1\chi_1' + C_2\chi_2' = \dfrac{iyV_F}{2c} - 2C_3 \\ \dfrac{C_1}{K_1}\chi_1 + \dfrac{C_2}{K_2}\chi_2 = \dfrac{C_1}{\sqrt{3}zK_1}(\chi_1 - \chi_1') + \dfrac{C_2}{\sqrt{3}zK_2}(\chi_2 - \chi_2'). \end{cases}$$

Solving it, we find expressions for C_1, C_2, and C_3. In terms of S_1 and S_2, they have the form

$$C_1 = \frac{3yV_F}{2c}S_1, \quad C_2 = \frac{3yV_F}{2c}S_2,$$

$$C_3 = C_1\chi_1 + C_2\chi_2 - \frac{iyV_F}{2c} = \frac{3yV_F}{2c}\left(S_1\chi_1 + S_2\chi_2 - \frac{i}{3}\right),$$

where

$$S_1 = \frac{i}{\dfrac{K_2}{K_1}\left(\dfrac{\chi_1 - \dfrac{\chi_1}{\sqrt{3}z} + \dfrac{\chi_1'}{\sqrt{3}z}}{\dfrac{\chi_2}{\sqrt{3}z} - \chi_2 - \dfrac{\chi_2'}{\sqrt{3}z}}\right)(2\chi_2 + \chi_2') + (2\chi_1 + \chi_1')},$$

$$S_2 = \frac{i}{\dfrac{K_1}{K_2}\left(\dfrac{\dfrac{\chi_2}{\sqrt{3}z} - \chi_2 - \dfrac{\chi_2'}{\sqrt{3}z}}{\chi_1 - \dfrac{\chi_1}{\sqrt{3}z} + \dfrac{\chi_1'}{\sqrt{3}z}}\right)(2\chi_1 + \chi_1') + (2\chi_2 + \chi_2')}.$$

Knowing the coefficients S_1 and S_2, we can find the absorption cross section σ in analytical form. According to (1),

$$\sigma = \frac{1}{2}\left(\frac{8\pi}{cH_0^2}\right)\mathrm{Re}\int \mathbf{j}_\varphi \cdot \mathbf{E}_\varphi^* d^3r$$

$$= 48\sigma_0 y^2 \int_0^1 \left(S_1\frac{\chi_1}{K_1} + S_2\frac{\chi_2}{K_2}\right)(S_1^*\chi_1^* + S_2^*\chi_2^*)\xi^2 d\xi, \quad (12)$$

where $\sigma_0 = \pi^2 ne^2 V_F R^4/(2mc^3)$; n is the charge carrier concentration; and S_1^*, S_2^*, χ_1^*, and χ_2^* are the respective complex conjugates.

The value of integral (12) can also be determined from the absorption equation [10]

$$\sigma = 4\pi\omega\alpha_m' V/c,$$

where α_m' is the magnetic moment calculated by the formula $\alpha_m' = -icC_3/(wH_0V)$ and V is the volume of the particle.

Carrying out relevant substitutions and changes of variables, we eventually obtain the formula for energy absorption:

$$\sigma = -\sigma_0\frac{48y}{w^2}\mathrm{Re}(C_3). \quad (13)$$

Thus,

$$F_\sigma(x, y, w) = -\frac{48y}{w^2}\mathrm{Re}(C_3).$$

Equations (12) and (13) are equivalent to each other. Therefore, the absorption cross section can be calculated by the expression that is free of integration, which considerably facilitates calculations.

RESULTS AND DISCUSSIONS

The calculated value of the absorption function allows one to describe the process of electromagnetic energy absorption by particles of different size.

The equations of macroscopic electrodynamics become invalid for particles whose linear sizes are comparable to the electron free path. For particles whose radius equals the electron free path ($x = 1$), the dependence of the absorption function $F(x, y, \omega)$ on the dimensionless frequency y of incident radiation is shown in Fig. 1 with $w = 3$. It is seen that the run of the function F_σ differs from that of F_{cl}, which was calculated in the framework of classical electrodynamics [10, 11]:

$$F_{cl}(x, y) = -\frac{24y}{w^2}\left(\frac{1}{k^2} - \frac{\sinh(2k) + \sin(2k)}{k(\cosh(2k) - \cos(2k))}\right),$$

where

$$k = w\sqrt{\frac{y}{2x}}.$$

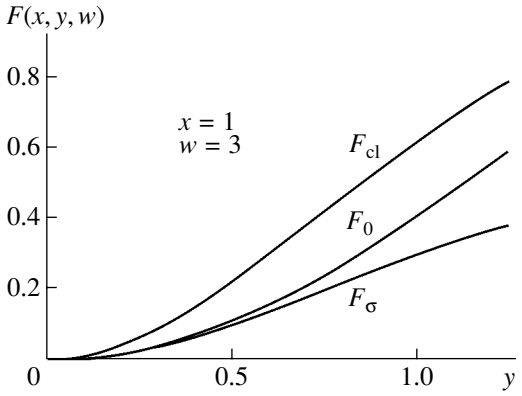

Fig. 1.

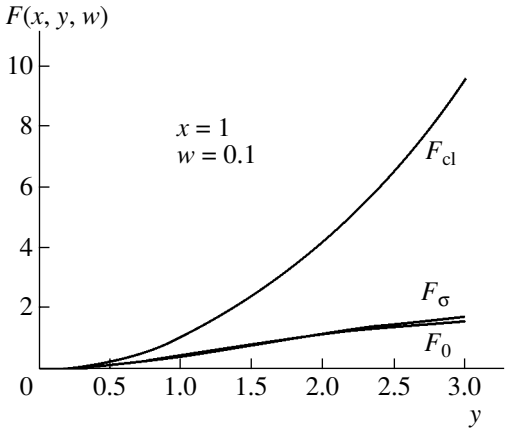

Fig. 2.

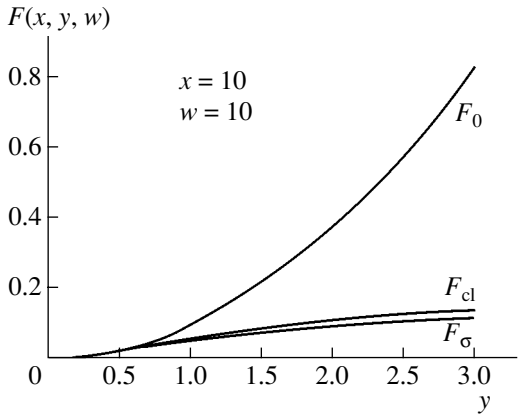

Fig. 3.

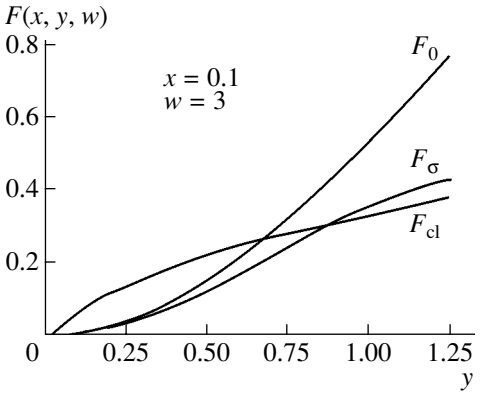

Fig. 4.

The kinetic increase of the function $F = F_0$, which does not consider the skin effect, also cannot cover all features of the electron behavior on the surface of the particle. The inclusion of the skin effect weakens the absorption (F_σ) in comparison with F_0, since field screening inside the particle is taken into account in this case.

Figure 2 shows that, for $w = 0.1$, the function F_σ, which was obtained in this paper, runs nearly in the same way as the function F_0, which was found from exact kinetic calculations without considering the skin effect. Note, however, that the skin effect is really insignificant at such w. Thus, the coincidence mentioned above indicates that the method suggested is fairly accurate. At the same time, with $x = 1$, macroscopic electrodynamics describes the absorption process inadequately, as demonstrated by the behavior of the curve F_{cl}.

The reverse situation is shown in Fig. 3, where $x = 10$ and the skin effect is pronounced ($w = 10$). Here, the classical macroscopic result F_{cl} describes the absorption process correctly, since the size of the particle is large (in comparison with the electron free path). This follows from the fact that the classical macroscopic result and F_σ are close to each other. At the same time, neglect of the skin effect leads to the overestimated value of the absorption cross section.

The dependence of the absorption function for $x = 0.1$ and $w = 3$ is shown in Fig. 4. When the dimension-

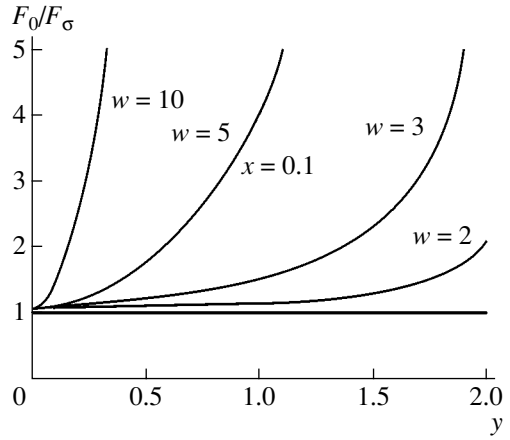

Fig. 5.

less frequency y is low, the functions F_σ and F_0 behave in a similar manner (i.e., the skin effect does not influence the absorption) but deviate markedly from the classical result F_{cl}. This is because x is small and electron scattering by the surface of the particle is essential for the process considered. As y increases, F_σ and F_0 diverge, since the skin effect shows up to a greater extent, while F_σ and F_{cl} approach each other.

Figure 5 demonstrates the y dependences of the ratio F_0/F_{cl} (the values of the absorption function $F(x, y, \omega)$ that were calculated without considering the skin effect). The rise in y and ω is seen to enhance the influence of the skin effect.

REFERENCES

1. Yu. I. Petrov, *Physics of Small Particles* (Nauka, Moscow, 1984), Chap. 7.
2. I. D. Morokhov, V. I. Petinov, L. I. Trusov, and V. F. Petrunin, Usp. Fiz. Nauk **133**, 653 (1981) [Sov. Phys. Usp. **24**, 295 (1981)].
3. A. G. Lesskis, V. E. Pasternak, and A. A. Yushkanov, Zh. Éksp. Teor. Fiz. **83**, 310 (1982) [Sov. Phys. JETP **56**, 170 (1982)].
4. A. G. Lesskis, A. A. Yushkanov, and Yu. I. Yalamov, Poverkhnost, No. 11, 115 (1987).
5. P. M. Tomchuk and B. P. Tomchuk, Zh. Éksp. Teor. Fiz. **112**, 661 (1997) [JETP **85**, 360 (1997)].
6. E. V. Zavitaev, A. A. Yushkanov, and Yu. I. Yalamov, *Absorption of Electromagnetic Radiation by a Finite-Length Cylindrical Particle*, Available from VINITI, Moscow, 2001, No. 2140-B.
7. J. M. Ziman, *Electrons and Phonons* (Clarendon Press, Oxford, 1960; Inostrannaya Literatura, Moscow, 1973).
8. W. A. Harrison, *Solid State Theory* (McGraw-Hill, New York, 1970; Mir, Moscow, 1972).
9. L. D. Landau and E. M. Lifshitz, *Course of Theoretical Physics*, Vol. 10: *Physical Kinetics* (Nauka, Moscow, 1979; Pergamon, New York, 1981).
10. L. D. Landau and E. M. Lifshitz, *Course of Theoretical Physics*, Vol. 8: *Electrodynamics of Continuous Media* (Nauka, Moscow, 1982; Pergamon, New York, 1984).
11. P. M. Morse and H. Feshbach, *Methods of Theoretical Physics* (McGraw-Hill, New York, 1952; Inostrannaya Literatura, Moscow, 1958), Vol. 1, Chap. 3.

Translated by M. Astrov

Technical Physics, Vol. 49, No. 1, 2004, pp. 93–103. Translated from Zhurnal Tekhnicheskoĭ Fiziki, Vol. 74, No. 1, 2004, pp. 93–103.

ELECTRON AND ION BEAMS, ACCELERATORS

Dynamics of Magnetic Insulation Failure and Self-Organization of an Electron Flow in a Magnetron Diode

A. V. Agafonov*, V. P. Tarakanov, and V. M. Fedorov****

** Lebedev Physical Institute, Russian Academy of Sciences, Leninskiĭ pr. 53, Moscow, 119991 Russia*

e-mail: agafonov@sci.lebedev.ru

*** Institute of High-Energy Densities, Joint Institute of High Temperatures, Russian Academy of Sciences, Izhorskaya ul. 13/19, Moscow, 125412 Russia*

Received May 7, 2003

Abstract—The dynamics of back cathode bombardment (BCB) instability in a magnetron diode (a coaxial diode in a magnetic field, $\mathbf{B} \equiv B_{0z} \equiv B_0$) is numerically simulated. The quasi-stationary regime of electron leakage across the high magnetic field ($B_0/B_{cr} > 1.1$, where B_{cr} is the insulation critical field) is realized. An electron beam in the electrode gap is split into a series of bunches in the azimuthal direction and generates the electric field component $E_\theta(r, \theta, t)$, which accelerates some of the electrons. Having gained an extra energy, these electrons bombard the cathode, causing secondary electron emission. The rest of the electrons lose kinetic energy and move toward the anode. Instability is sustained if the primary emission from the cathode is low and the secondary emission coefficient $k_{se} = I_{se}/I_{e,\,BCB}$ is greater than unity. The results of numerical simulation are shown to agree well with experimental data. A physical model of back-bombardment instability is suggested. Collective oscillations of charged flows take place in the gap with crossed electric and magnetic fields ($\mathbf{E} \times \mathbf{B}$ field) when the electrons and $\mathbf{E} \times \mathbf{B}$ field exchange momentum and energy. The self-generation and self-organization of flows are due to secondary electron emission from the cathode. © *2004 MAIK "Nauka/Interperiodica".*

INTRODUCTION

This work is devoted to the physics of magnetic insulation failure as applied to a dense electron flow in a vacuum coaxial diode subjected to crossed magnetic and electric fields $\mathbf{E} \times \mathbf{B}$ (hereafter, a magnetron diode (MD)). Interest in the physics of crossed-field devices (microwave oscillators, high-current-beam injectors, etc.) stems from the fact that they are widely used in high-power vacuum electronics. However, the theory of these devices still remains on the phenomenological level. A magnetic field $\mathbf{B}_0$ is known to suppress the mobility of electrons (ions) across the lines of magnetic force at distances exceeding the Larmor radius of the particles ($d_{ac} > r_{eL} = mv_{e,\,ac}/eB_0$). This effect, called magnetic insulation, is nearly absolute in axisymmetric systems with a sparse beam ($en_e d_{ac} \ll V_{ac}/d_{ac} = E_0$, where d_{ac} is the anode–cathode spacing and V_{ac} is the voltage across this spacing). In high-power electron devices, the beam is dense ($en_e d_{ac} \approx E_0$). In this case, a leakage current across the magnetic field is observed as a rule because of fluctuations in the charge distribution (flow turbulence).

The physics of crossed-field devices is difficult for theoretical analysis because the associated processes are highly nonlinear, and numerical simulation in terms of self-consistent models requires a great body of computation and dedicated programs [1–6]. The use of simplified models (such as the 1.5-D code) gives faulty results. For example, it was erroneously argued [7] that a beam in an MD becomes unstable when the emission current from the virtual cathode is high. It is, however, known that the virtual cathode instability is related to a high mobility of charged particles along the applied electric field **E** [8], while in MDs the charge mobility along the **E** field is suppressed by the transverse **B** field.

A beam in MDs becomes highly unstable when the primary electron current from the cathode is low. This result was repeatedly obtained in the midtwentieth century in experiments with diodes with cold secondary emission cathodes [1, 9, 10] and a low primary electron emission current ($I_{e0} < 1$ A).[1] With a voltage pulse (V_{ac} = 5–30 kV) applied to the diode and $k_{se}(w_{e,\,BCB}) > 1$, large-amplitude collective oscillations were observed. Individually, the oscillations of the charge and electric field inside the diode were measured neither in [9, 10] nor in other experiments because of technical difficulties. The high-amplitude beam oscillations inside the diode showed up as the emission enhancement more than a hundred times (the current through the diode dramatically increased, $\varepsilon_{eB} = I_{ac}/I_{e0} \gg 1$, when

[1] The effect of secondary emission from metal surfaces was discovered in 1902 upon studying cathode rays. The secondary emission coefficient ($k_{se} = I_{se}/I_{e,\,BCB}$) for all-metal cathodes used in [9] exceeds unity for bombarding electron energies $w_{e,\,BCB} >$ 200 eV. Composite (e.g., oxide) cathodes used in [10] have a lower energy threshold [11].

the magnetic field was applied) and as an additional heating of the cathode with a high efficiency ($\eta_{BCB} = P_{e,BCB}/P_{ac} \approx 30\%$) of energy extraction from the power supply. This effect was treated by many authors [1, 9, 10] as the generation of a beam due to the back bombardment of the cathode by fast electrons, which also heats up the cathode. However, attempts to construct a speculative theoretical model of this process have failed.

Electrons with an excess energy $\delta w_{e,c} \approx 300$ eV in the MD gap were discovered even in early experiments [12]; however, no adequate explanation to this phenomenon was given at the time. Note that the diodes operated in a high vacuum at voltages $V_{ac} > 1$ kV. Under such conditions, pair electron collisions and residual gas ionization may be neglected. Only the collective interaction of electrons with the electromagnetic field may affect noticeably the energy distribution. Oscillations that cannot be related to resonance structures are usually called turbulent oscillations. The study of electron flow turbulence in MDs is a challenge, because associated experimental data are hard to obtain. The oscillation (noise) level in the external circuit was relatively low and correlated weakly with the effect of magnetic insulation failure, thereby indicating that the oscillations are radiationless. Studies of crossed-field microwave devices over a period from the 1940s to the 1960s were summarized in [1]. The authors of those articles acknowledged that the problem of electron flow turbulence in crossed-field devices had remained unresolved. Later, the problem of stability of high-power beams in crossed fields has again attracted considerable attention because of research on controlled thermonuclear fusion. In particular, this problem is topical for plasma accelerators [13], magnetic isolation of hot plasma from cold walls [14], and the generation and transport of high-power beams [2, 3].

1. PHYSICS OF PROCESSES RESPONSIBLE FOR ELECTRON LEAKAGE TO THE ANODE IN MAGNETRON DIODES

"Strange" instability observed in experiments with MDs was first simulated in 1996 and then studied in detail in [15–18]. Simulation was performed with the electromagnetic PIC code KARAT [19, 20]. The results of simulation were verified by the fulfillment of the energy balance: $P_{ac} \cong P_{e,a} + P_{e,BCB}$; i.e., the consumed power and the power accumulated at the electrodes were equal to each other within several percent. Let us consider the steady-state conditions for electron leakage to the anode: $I_{e,a}(t) \approx I_{ac} = \text{const}$, $V_{ac}(t) \approx V_{a0}$, $eV_{a0} = m\nu_{e,a0}^2/2$, $\bar{\nu}_{e,a} < \nu_{e,a0}$. The law of conservation of energy in an MD is fulfilled if

$$P_{ac} = P_{e,a} + P_{e,BCB}; \quad P_{ac} = I_{ac}V_{ac};$$
$$P_{e,a} = \bar{I}_{e,a} m \bar{\nu}_{e,a}^2/2e; \tag{1}$$
$$P_{e,BCB} = \bar{I}_{e,BCB} m(\bar{\nu}_{e,BCB}^2 - \bar{\nu}_{e,c0}^2)/2e,$$

where $\bar{w}_{e,a} = m\bar{\nu}_{e,a}^2/2$, $\bar{w}_{e,BCB} = m\bar{\nu}_{e,BCB}^2/2$, and $\bar{w}_{e,c0} = m\bar{\nu}_{e,c0}^2/2$ are the mean energies of electrons that fall on the anode ($\bar{w}_{e,a} < eV_{ac}$), strike the cathode, and leave the cathode.

The computer model, unlike full-scale experiments [9, 10], allowed us to "measure" fields, currents, energy fluxes, and momentum fluxes in the diode gap. The azimuthal momentum changes because of the magnetic force moment, which is similar to the torque of a usual electric motor:

$$\dot{M}_{0B} = 2\pi \int_{r_c}^{r_a} \bar{J}_{er} B_0 r^2 dr = \int_{r_c}^{r_a} \bar{I}_{e,a} B_0 r dr = \frac{\bar{I}_{e,a}}{e} m \nu_{e,a0} r_a \frac{B_0}{B_{cr}}; \tag{2}$$
$$B_{cr} = \frac{2 r_a m \nu_{e,a0}}{e(r_a^2 - r_c^2)},$$

where B_0 is the applied axial magnetic field and B_{cr} is the critical magnetic field (or the Hull cutoff field).

Angular momentum (2) is transferred to the MD electrodes when electrons strike the anode (at an angle $\bar{\alpha}_{e,a}$) and when the electrons with an excess energy strike the cathode (at an angle $\bar{\alpha}_{e,BCB}$):

$$\dot{M}_{e,a} = (\bar{I}_{e,a}/e) m r_a \bar{\nu}_{e,a} \sin\bar{\alpha}_{e,a};$$
$$\dot{M}_{e,c} = (\bar{I}_{e,BCB}/e) m r_c (\bar{\nu}_{e,BCB} - \bar{\nu}_{e,c0}) \sin\bar{\alpha}_{e,BCB}; \tag{3}$$
$$\dot{M}_{e,a}/\dot{M}_{0B} + \dot{M}_{e,c}/\dot{M}_{0B} = 1.$$

The law of conservation of azimuthal momentum is fulfilled if condition (3) is met. Now let us see whether electrons may leak to the MD anode ($I_{e,a} \neq 0$) if the back bombardment of the cathode is neglected ($\dot{M}_{e,c} \approx 0$). Such an assumption comes into conflict with the law of conservation of energy ($\bar{\nu}_{e,a} < \nu_{e,a0}$) and azimuthal momentum (3), since the inequality $\dot{M}_{0B}/\dot{M}_{e,a} > B_0/B_{cr} > 1$ is valid. Thus, the effect of back cathode bombardment is a direct consequence of magnetic insulation failure in an MD. In [3, 6], such a regime was given the name back-bombardment instability.

2. CHARGE OSCILLATION IN A MAGNETRON DIODE UNDER THE CONDITION OF BACK-BOMBARDMENT INSTABILITY

It was shown [15–18] that an electron flow in a smooth-bore MD generates a self-organizing self-

renewable structure in the (r, θ) plane in the form of dense electron clouds rotating about the cathode. The azimuthal size of the clouds is close to the electrode spacing d_{ac} in the diode, which corresponds to the shielding length for electric field disturbances between the electrodes. An electric field produced by charge bunches is somewhat similar to the rotating rf electric field in the diode of a generator magnetron [21, 22].[2] It is known that magnetrons use a slow-wave structure (a slot anode with resonators), which generates a traveling azimuthal rf field $E_\theta(r_a, \theta - \omega t)$. According to the quantitative engineering model of electronic processes in a magnetron, which was developed in the 1940s–1950s [see, e.g., [1, 21, 22]), most electrons in the diode group to form needles. Electron needles rotate about the cathode with a velocity roughly equal to the drift velocity $v_{EB} = E_r/B_0$ synchronously with the traveling rf field $E_\theta(r_a, \theta - \omega t)$. Such conditions provide an efficient energy extraction from the beam. The magnetic field B_0 cannot confine the retarded electrons ($w_{e,a} < 0.3eV_{ac}$, $eE_a > ev_{e,a}B_0$), and they are attracted to the anode, which absorbs them to generate a current $I_{e,a}$. This current picks up the power $P_{ac} = \bar{I}_{e,a}\bar{V}_{ac}$ from the power supply. The electrons are emitted into needles from the surface of an "electronic sleeve" (its thickness is estimated as $\delta_{e,c} = (0.1–0.2)d_{ac}$), which is formed by a dense flow of electrons that both left the cathode and returned back to it. In the engineering model, the emission current from the cathode is assumed to be space-charge-limited (at the cathode the field $\mathbf{E}_c = 0$). This assumption is justified by satisfactory agreement between calculations and practice.

Such an approach to analyzing oscillations in a smooth-bore magnetron diode turned out to be inadequate, since both external ($P_\sim = 0$) and internal ($E_{a\theta} = 0$, $\mathbf{E}_c = 0$) sources of an rf field are absent in its gap. In an MD, the cathode (at the cathode, $\mathbf{E}_c = E_r(r_c, \theta - \Omega t, t) \neq 0$) with nonuniform nonstationary emission serves as a driver of a traveling rf electric field. This secondary emission occurs under the action of a backward electron flow, which implies a high level of beam oscillation. When studying a turbulent beam in an MD, one runs into obstacles both in solving nonlinear equations and in describing intermediate objects that make possible energy and momentum exchange between the electrons moving toward the anode and those returning to the cathode. The electrons taking part in collective oscillations exchange energy when interacting with the $\mathbf{E} \times \mathbf{B}$ field. In this process, the magnetic field remains virtually unchanged, while the electric field varies noticeably because of the space charge of the dense beam. In collective oscillations, the energy of the $\mathbf{E}$ field serves as an energy buffer ($\varepsilon_0\mathbf{E}^2/2 \sim \langle n_e m v^2\rangle/2$), which has been accepted by most of the researchers.

The process of momentum exchange between beam electrons in an MD (this process is a prerequisite for the oscillations to occur and the azimuthal momentum to be transferred to the electrodes; see (2), (3)) has been poorly understood because of the complexity of the problem. It is known that crossed-field devices have a nonzero density of the electromagnetic momentum ($\mathbf{g}_{EB} = \varepsilon_0\mathbf{E} \times \mathbf{B}_0$) for both rapidly and slowly time-varying fields (i.e., in the absence of electromagnetic waves; see, e.g. [25, §104 and Chap. 27]). Straightforward analysis shows that, in diodes with magnetic insulation of electrons, the electromagnetic momentum and the electron flow momentum per unit volume are comparable in order of magnitude ($|\mathbf{g}_{EB}| \sim |n_e m v_e|$). As applied to devices like MDs ($B_{0z} = \text{const}$, $\mathbf{E} = \nabla V(r, \theta, t)$), one can state the following. In the diode gap, almost radiationless high-amplitude ($\delta n_e \sim n_e$) rf oscillations of electric charges may take place even if the variation of the applied voltage and diode current is small. These oscillations are sustained when the charges interact with the $\mathbf{E} \times \mathbf{B}$ field and exchange momentum and energy with it. Nonstationary secondary electron emission from the cathode is also an important factor.

3. STATEMENT OF THE PROBLEM AND MATHEMATICAL MODEL

The dynamics of particles in a smooth-bore magnetron was simulated with the code KARAT in a combined way. An external voltage pulser $V_p(t)$ was connected to the diode through an *RL* circuit, and the diode was represented as a circuit element (Fig. 1) with the parameters $V_{ac}(t)$ (voltage across the diode) and $I_{ac}(t)$ (current through the diode). The overall circuit was described by the Kirchhoff equation

$$V_{ac}(t) = V_p(t) - R_1 I_{ac}(t) - L_1 dI_{ac}(t)/dt, \quad R_1 = 0.2k\Omega, \quad L_1 = 0.05\mu H. \tag{4}$$

The voltage pulse $V_p(t)$ has a smooth profile with a leading edge time of 2 ns. Typical voltage and current waveforms are given in Fig. 1. The current $I_{ac}(t)$, the sum of the electron leakage current to the anode and the capacitive (displacement) current, was calculated with

[2] Recall that a high-power pulsed microwave magnetron was first created early in the 1940s jointly by teams of researchers in the United States and Europe. Magnetron-equipped radars played an essential role in World War II. High peak microwave powers ($P_\sim = 0.1–1.0$ MW) generated by magnetrons in the centimeter range (with an efficiency $\eta_\sim = P_\sim/P_{ac} = 50–70\%$) were obtained due to an unexpectedly strong electron emission from the cathode ($J_e \sim 100\text{A/cm}^2$, $J_e/J_{e0} \sim 10^2$ [21, Chaps. 1, 12]), which far exceeded its thermionic emission. The reason was found to be secondary electron emission from the cathode, which is associated with the back bombardment of the cathode. In the 1930s and earlier, thin filaments were usually used as cathodes. These filaments readily burned out. The burnout of the cathode was observed, for example, upon testing (Leningrad, 1936–1937) the world's first prototypes of high-power ($\bar{P}_\sim \geq 0.3$ kW) multicavity magnetrons ($\lambda = 9$ cm) [23], which formed the basis for industrial magnetrons [21, 22]. In today's compact "nonincandescent" magnetrons [24], secondary emission is generated by weak cold emission.

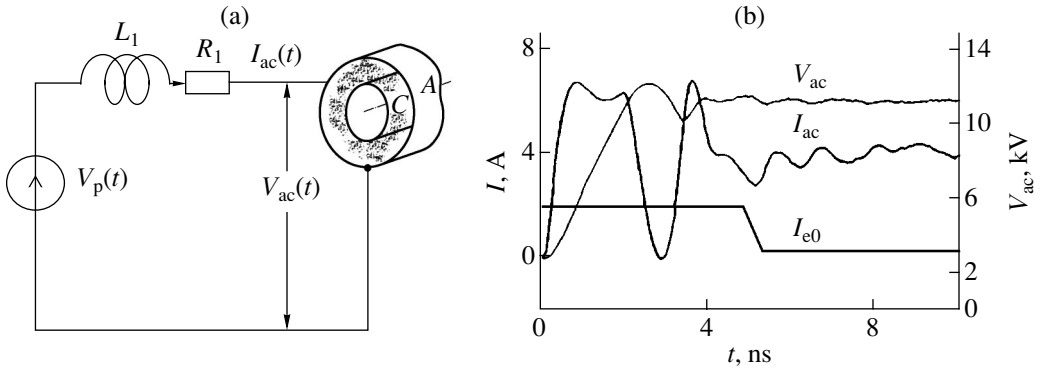

Fig. 1. (a) Connection of the magnetron diode to the external circuit and (b) waveforms of the voltage $V_{ac}(t)$ across the gap, primary emission current $I_{e0}(t)$, and current $I_{ac}(t)$ across the diode.

the Maxwell equations and a discrete model of charged plasma (the PIC method of coarse particles [27]). The dynamics simulation was carried out by using meshes with the number of nodes 10^5 and the number of coarse (primary and secondary) particles 10^5.

The code KARAT makes it possible to simulate both the 2D (r, θ) and 3D (r, θ, z) configurations of the diode. To test the numerical method and find the applicability domain of the 2D model, simulation was performed for an MD with parameters close to those obtained experimentally [9]: molybdenum cathode, r_c = 3.3 mm; anode, r_a < 10 mm; length, l_2 = 20 mm; $B_{0z} \equiv B_0$ = 0.2–0.3 T; and V_{ac} = 10– 30 kV. In [9], the cathode had annular extensions at the end faces and the magnetic field at the end faces was slightly increased (as in the diode of an AKh9 magnetron [21] in order to suppress electron leakage out of the diode. A negative voltage pulse was applied to the cathode, and the anode and evacuating system were under the zero potential, which is a standard practice in the technology of generator magnetrons and in experiments with MDs. In spite of shielding, the electrons in such a system may escape along the magnetic field toward the chamber walls. The results of simulation and experimental data from [9] coincided within 10% when (i) the gap was small compared with the longitudinal dimension of the diode, $d_{ac} \le 0.1 l_z$, and (ii) the field $B_0 = (1.1–1.2)B_{cr}$. At larger gaps ($d_{ac} > 0.1 l_z$) or $B_0 > 1.2 B_{cr}$, the results of simulation and experimental data differed markedly probably because of the influence of the cathode's end faces or electron escape toward the walls of the vacuum chamber.

High magnetic fields, B/B_{cr} = 1.5–3.0, are usually used to improve the efficiency of magnetrons [21, 22] and beam injectors [28, 29] based on diodes with magnetic insulation and secondary-emission cathodes. In the latter, the beam is injected into the gap from the near-cathode sheath (electronic sleeve). As was noted in Section 2, a theoretical model for this sheath is absent. In the last decade, a number of authors [7, 29] tried to fill this gap and suggested a simple model based on beam one-dimensional oscillations. In other words, they invoked the Langmuir's early idea of oscillation potentials [1, 8, 12]. Langmuir oscillation is known to be widely used in the plasma theory; however, this concept fails when applied in the theory of charged beams in crossed $\mathbf{E} \times \mathbf{B}$ fields. The theory of one-dimensional oscillation may explain electron leakage across the magnetic field provided that the restrictions related to the energy and momentum conservation laws are eliminated (see (1)–(3)).

Usually, smooth-bore MDs ($E_\theta(r_c) = E_\theta(r_a) = 0$) subjected to voltages V_{ac} < 100 kV may be analyzed under the assumption that the electromagnetic field is irrotational (curl$\mathbf{E}$ = 0, curl$\mathbf{B}$ = 0). At the same time, when analyzing the radial equilibrium of the device, one should take into account that the azimuthal current of the electrons displaces slightly ($\delta B/B_0 \approx 1\%$) the magnetic field toward the anode. Simulations based on the potential model and on equations that include relativistic effects and the rotational components of the electromagnetic fields gave nearly the same results for V_{ac} = 10–30 kV and r_a < 10 mm. A problem arising in numerical simulation is avoiding instabilities due to the discreteness of calculations.[3]

In the case of the 2D model, the integral quantities (current, charge, etc.) were calculated for an axial length l_{0z} = 10 mm. The basic design with the parameters r_c = 3.3 mm, r_a = 5.3 mm, d_{ac} = 2 mm, and B_0 = 0.25 T was used. The voltages (see Fig. 1) were V_p = 12 kV and

[3] High-power beams in crossed $\mathbf{E} \times \mathbf{B}$ fields are known to be subject to many instabilities especially when secondary emission takes place. In [30, 31], experimental and theoretical investigations into a single-wall secondary-emission microwave discharge in $\mathbf{E} \times \mathbf{B}$ fields were carried out. In our calculations, such a discharge occurred as a stray instability ("computer physics") when the mesh was insufficiently fine (N_r < 200, N_θ < 250). Fortunately, the fundamental microwave resonance of the discharge is near the double cyclotron frequency and away from the fundamental harmonic of BCB instability, which lies in the vicinity of the Larmor frequency ($\nu_{0,BCB} \approx \nu_{Le} = \nu_{ce}/2$). This fact allowed us to separate out, if any, the stray instability. However, stringent requirements for the fineness of the mesh prevented us from performing full-scale 3D calculations.

V_{ac} ($t > 6$ ns) $\approx V_{a0} = 11.2$ kV (quasi-stationary regime). In practice, the parameters l, t, ν, I, Q, V, W, E, B, R, ν, L, and C are usually measured in mm, ns, mm/ns, A, nC, kV, μJ, kV/mm, T, kΩ, GHz, μH, and pF. Also, $\varepsilon_0 = 1/36\pi$ pF/mm, $c = (\varepsilon_0\mu_0)^{-0.5} = 300$ mm/ns, $e = 1.6 \times 10^{-10}$ nC, and $m = m_e = 9.1 \times 10^{-13}$ μJ (ns/mm)^2. The equations which follow are given for the (r, θ) geometry in the nonrelativistic approximation and for a potential electric field:

$$v_{nr}(r, \theta, t) - v_{nr}(r_c, \theta_1, t_1) = \int_{t_1}^{t}\left(-\frac{E_r}{E_0}v_{EB} - v_{n\theta} + \frac{v_{n\theta}^2}{r\omega_{ce}}\right)\omega_{ce}dt_n, \quad (5)$$

$$\omega_{ce} = \frac{eB_0}{m}, \quad v_{EB} = \frac{E_0}{B_0},$$

$$mrv_{n\theta}(r, \theta, t) - mr_K v_{n\theta}(r_c, \theta_1, t_1) = mr_a v_{e,a0}\left(\frac{B_0}{B_{cr}}\frac{r^2 - r_c^2}{r_a^2 - r_c^2} - \frac{v_{EB}}{v_{e,a0}}\int_{t_1}^{t}\frac{rE_\theta}{r_a E_0}\omega_{ce}dt_n\right), \quad (6)$$

$$E_0 = V_{a0}/d_{ac}, \quad v_{e,a0} = (2eV_{a0}/m)^{0.5},$$

$$V_{cr} = V_{a0}(B_0/B_{cr})^2, \quad B_{cr} = \frac{2r_a m v_{e,a0}}{e(r_a^2 - r_c^2)}.$$

Here, v_{nr} and $v_{n\theta}$ are the velocity components of an nth electron, (r, θ, t) are path coordinates, $E_0 = 5.6$ kV/mm, $v_{EB} = 22.4$ mm/ns, $\omega_{ce} = 2\pi\nu_{ce}$, $\nu_{ce} = 28B_0 = 7$ GHz is the electron cyclotron frequency, $v_{e,a0} = 62.7$ mm/ns is the velocity of an electron with an energy $w_{e,a0} = eV_{a0} = 11.2$ keV, $V_{cr} = 14.5$ kV is the cutoff voltage, $B_{cr} = 0.22$ T is the critical cutoff field at $V_{ac} = 11.2$ kV, $B_0/B_{cr} = 1.13$; and $dt_n = 10^{-5}$–10^{-4} ns is the time step of calculation. Integration is carried out over the path of the nth electron.

Equations (5) and (6) describe the history of an arbitrary electron in a permanent magnetic field $\mathbf{B}_0$ and in an electric field with the components $E_r(r, \theta, t)$ and $E_\theta(r, \theta, t)$, which are found by numerically solving Poisson equation (7) with boundary conditions (8). Charge distribution (9) in the diode gap depends on the position (see (5), (6)) and amount of electrons (macroelectrons) that have escaped from the cathode over a time $\Delta t = (0 < t_1 < t)$ and stay in the gap at a current time instant t (that is, of those electrons that were not absorbed by the anode or cathode). The leakage current toward the anode $I_{e,a}(t)$ and the diode current $I_{ac}(t)$ are calculated by (10) and (11):

$$r\frac{\partial}{\partial r}\left(r\frac{\partial V}{\partial r}\right) + \frac{\partial^2 V}{\partial\theta^2} = \frac{-r^2 q_e}{\varepsilon_0}, \quad (7)$$

$$E_r(r, \theta, t) = \frac{-\partial V(r, \theta, t)}{\partial r}, \quad E_\theta(r, \theta, t) = \frac{-\partial V(r, \theta, t)}{r\partial\theta};$$

$$V(r_c, \theta, t) = 0, \quad V(r_a, \theta, t) = V_{ac}(t), \quad E_{\theta(r_c, \theta, t)} = E_\theta(r_a, \theta, t) = 0; \quad (8)$$

$$q_e(r, \theta, t) = \frac{1}{\delta r\delta\theta}\int_{r-\delta r/2}^{r+\delta r/2}dr\int_{\theta-\delta\theta/2}^{\theta+\delta\theta/2}e_n(r, \theta, t)d\theta, \quad \frac{\partial(rJ_r)}{\partial r} + \frac{\partial J_\theta}{\partial\theta} = -r\frac{\partial q_e}{\partial t}; \quad (9)$$

$$J_x(r, \theta, t) = \frac{1}{\delta r\delta\theta}\int_{r-\delta r/2}^{r+\delta r/2}dr\int_{\theta-\delta\theta/2}^{\theta+\delta\theta/2}e_n v_{nx}d\theta, \quad J_{x=r} \equiv J_r, \quad J_{x=\theta} \equiv J_\theta; \quad (10)$$

$$I_{e,a}(t) = -r_a I_{0z}\int_0^{2\pi}J_r(r_a, \theta, t)d\theta,$$

$$I_{ac}(t) = I_{e,a}(t) - \varepsilon_0 r_a l_{0z}\int_0^{2\pi}\frac{\partial}{\partial t}E_r(r_a, \theta, t)d\theta. \quad (11)$$

The cathode surface emits primary electrons ($I_{0e}(\theta, t)$ is given by initial conditions) and secondary electrons, which are knocked out of the cathode by fast electrons. The yield factor of secondary electrons ($k_{se} = I_{se}/I_{e,BCB}$) is calculated in view of empirical data [11] by the formula

$$k_{se} = \left((1 - \cos\alpha_{e,BCB}) + \frac{\pi}{2}k_{sm}w^{0.55}\exp(-0.45w)\right)h(w), \quad w = (w_{e,BCB} - w_{s1})/w_{s2}, \quad (12)$$

where k_{sm} is the maximal secondary emission coefficient ($k_{sm} = k_{sm}(\theta) < 3$), $\alpha_{e,BCB}$ is the angle between the electron velocity and the normal to the surface, $w_{e,BCB}$ is the kinetic energy of the electrons bombarding the cathode, $w_{s1} \approx 0.1$ keV is the threshold energy, $w_{s2} \approx 0.6$ keV is an energy parameter, and $h(w)$ is the Heaviside function.

The electrons escaping from the cathode were uniformly distributed in the energy interval $w_{e,c0} = (10$–$50)$ eV and had the same escape direction in terms of the polar and azimuthal angles. Note that the initial electron energy adopted in the simulation ($w_{e,c0} = 30$ eV) exceeds the value obtained in experiments (see, e.g., [11]) but is much smaller than the characteristic electron energy ($w_e \approx 1$ keV) in the near-cathode sheath. In the basic design of the diode, there exists a rather high energy threshold ($\delta w_e/w_{e,a} \approx 20\%$) for the electrons moving toward the anode. Note for comparison that, in [4], the electron escape toward the anode was simulated for $(B_0 - B_{cr})/B_0 \approx 1\%$ and, in [5], the computational scheme worked only for the case when electrons were emitted along the magnetic field. Under the steady-state conditions ($V_{ac}(t) \approx V_{a0}$), the electrons escaping from the cathode may reach the anode only if a nonstationary

azimuthal electric field $E_\theta(r, \theta, t)$ exists in the diode gap.

4. DYNAMICS OF ELECTRON FLOW FORMATION IN MAGNETRON DIODES

The dynamics of formation of electron clouds, secondary emission current, and electric field in the MD gap (basic design) is illustrated in Figs. 2–4. At the zero time, azimuth-uniform primary electron emission is accomplished (I_{e0} = 2 A). These electrons form an azimuth-uniform sheath near the cathode. The sheath thickens with increasing voltage $V_{ac}(t)$ (Fig. 2, the configuration at t = 3 ns). The electrons starting at earlier times cannot return to the cathode, because the electric field strength grows. After $\Delta t \approx 3$ ns, the voltage across the diode reaches a maximum value (12 kV) and the current through the diode drops to zero (the waveforms in Fig. 1). Over a time interval of 3–4 ns, the electric field strength $E_r(r, \theta, t)$ in the near-cathode sheath

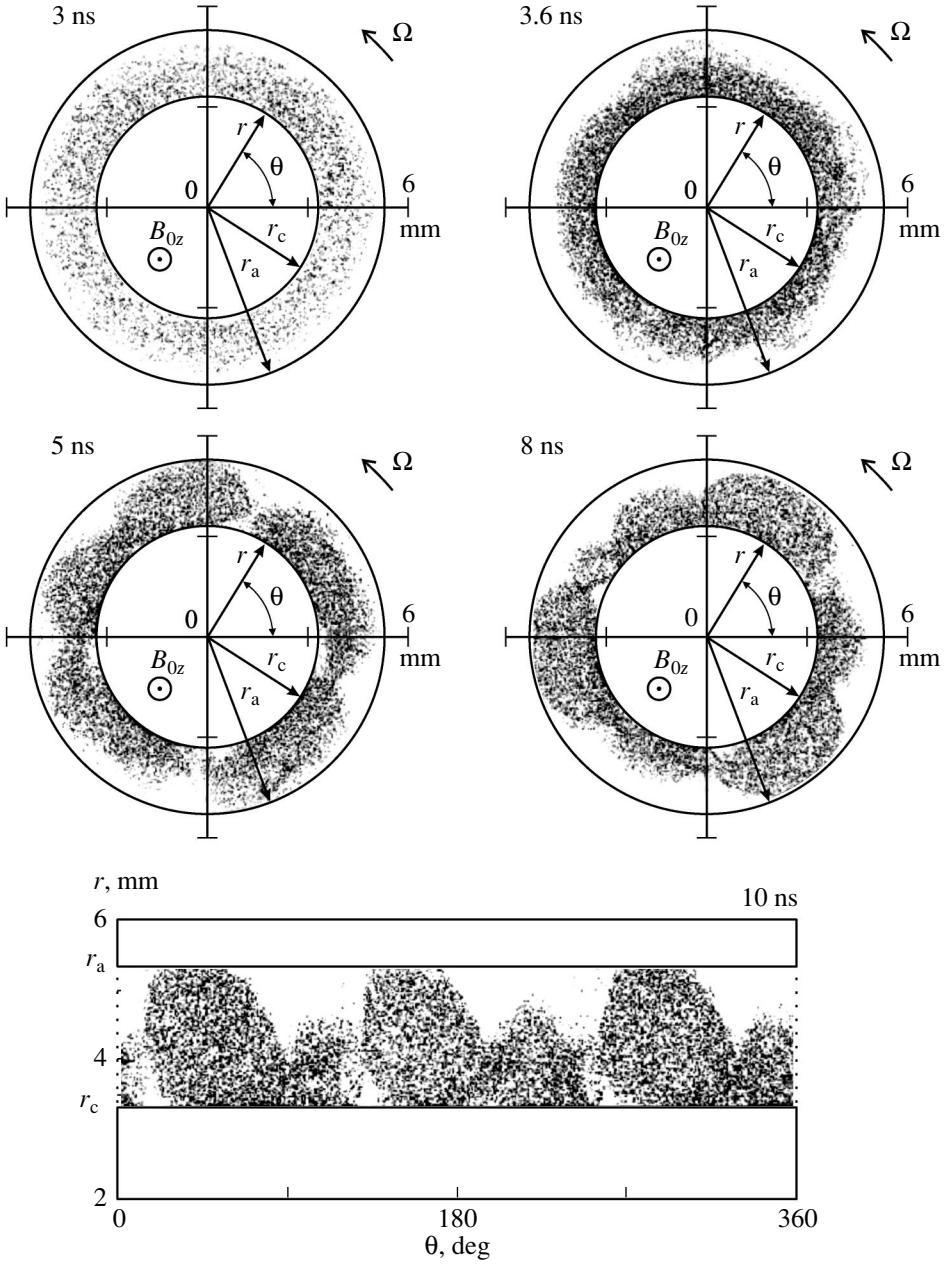

Fig. 2. Formation of electron clouds in the diode gap in the case of uniform initial emission ($I_{e0} \approx 2$ A) from the cathode. At the time 3.6 ns, secondary electrons are generated in an avalanche-like manner. By the time t = 8 ns, the quasi-stationary regime with a leakage current $I_{e,a} \approx 4$ A toward the anode and weak diode current oscillations ($\delta I_{ac}/I_{ac} \approx 1\%$) is established. At the bottom, rectangular sweep at the time t = 10 ns is shown.

decreases, which causes the intense electron bombardment of the cathode and a rapid growth of the secondary emission current (Figs. 1, 3, 4). The cathode bombardment at the stage where the field $\mathbf{E}(t)$ decreases has an effect opposite to that observed when the electrons are captured in the gap during the increase in $\mathbf{E}(t)$. In real experiments, the leading edge time is usually longer and secondary emission is stimulated by a special circuit that generates a high voltage spike $\delta V(t) \sim V_{ac}$ across the diode [28, 29].

The electrical parameters of the diode (the voltage $V_{ac}(t)$ and the current $I_{ac}(t)$) are stabilized with time. The steady regime is established by the time t = 8 ns (Fig. 1). However, the secondary emission current $I_{se}(t)$ continues to markedly oscillate even at $t > 8$ ns (Fig. 4). The current $I_{se}(t)$ exceeds the anode current amplitude $I_{e,a}(t) \approx I_{se}(t) - I_{e,BCB}(t)$ roughly by a factor of 10. The backward electron flow, which is represented by the bombardment current $I_{e,BCB}(t)$, also oscillates significantly. The charge distribution $q_e(r, \theta, t)$ in the gap turns out to be severely nonuniform (Fig. 2). As a result, an azimuthal field $E_\theta(r, \theta, t)$ with an amplitude of about $0.15E_0$ arises and the field $E_r(r, \theta, t)$ noticeably oscillates (Figs. 3, 5). The pattern of electron clouds (bunches) (Fig. 2, t = 8 ns) remains nearly the same with time and has a period of revolution about the axis of $t_a \approx 1$ ns. Assuming that the radial velocity in (5) equals zero and $E_r = -E_0$, we can estimate the azimuthal velocity at the midradius: $\langle v_{e\theta} \rangle = v_a \approx v_{EB}(1 + v_{EB}/r_a\omega_{ce}) = 25$ mm/ns. The period of revolution for these electrons is $t_{aa} = 2\pi r_a/v_a \approx 1$ ns, which coincides with the period of revolution of the clouds t_a. It can be assumed that most of the electrons in the gap are captured particles that have a kinetic energy lower than the potential energy ($w_e < eV(r, \theta, t)$). This assumption is substantiated by the shape of the electron energy distribution function (Fig. 6).

The mean energies of the electrons reaching the anode and bombarding the cathode are, respectively, $\bar{w}_{e,a}$ = 7.4 keV and $\bar{w}_{e,BCB}$ = 0.44 keV (Fig. 7). Using these values and those under the steady-state conditions (V_{ac} = 11.2 kV, $\bar{I}_{e,a}$ = 4 A, $\bar{I}_{e,BCB}$ = 36 A), we may numerically calculate the back bombardment efficiency, the power fluxes toward the electrodes, and the precision of energy balance (1): $\eta_{BCB} = (eV_{ac} - \bar{w}_{e,a})/eV_{ac} = 34\%$, P_{ac} = 44.8 kW, $P_{e,a}$ = 29.6 kW, $P_{e,BCB}$ = 14.8 kW, and $(P_{ac} - P_{e,a} - P_{e,BCB})/P_{ac} < 1\%$. Disagreement with experimental data in [9] is within 10%. In the computer model of the diode, the emission enhancement coefficient $\varepsilon_{eB} = I_{e,a}/I_{0e}$ may be taken high ($\varepsilon_{eB} > 100$), since the primary emission current under the steady-state conditions may be infinitesimal: $I_{0e}(t > 8 \text{ ns}) \approx 0$ (Fig. 1).

Consider the angular characteristics of the electrons at the MD electrodes and the conservation of the azimuthal momentum. The electrons approach the anode along the tangent ($\bar{\alpha}_{e,a} = 82°$), and their angular distribution is narrow. The mean angle of incidence to the cathode is $\bar{\alpha}_{e,BCB} = 24°$; that is, here the path is closer to the normal and the distribution function is wider (Fig. 7). Using the above values for the steady-state conditions and formulas (2) and (3), we check the balance in azimuthal momentum:

$$1 = \frac{\dot{M}_{e,a}}{\dot{M}_{0B}} + \frac{\dot{M}_{e,c}}{\dot{M}_{0B}} \approx \frac{B_{cr}\bar{w}_{e,a}^{0.5}\sin\bar{\alpha}_{e,a}}{B_0 w_{e,a0}^{0.5}} + \frac{B_{cr}r_c(\bar{w}_{e,BCB}^2 - \bar{w}_{e,c0}^{0.5})\bar{I}_{e,BCB}\sin\bar{\alpha}_{e,BCB}}{B_0 r_a w_{e,a0}^{0.5} I_{a,c}} = 0.71 + 0.29, \quad (13)$$

that is, the torque balance is fulfilled within 1%.

In the charge azimuth distribution (Fig. 2, t = 10 ns), three basic bunches and three bunches of lower intensity can be distinguished. In the frequency spectrum for the electric field, current density, etc., the frequencies corresponding to the period of revolution t_a and higher harmonics ($t_a/3$, $t_a/6$, …) are resolved. The spectrum

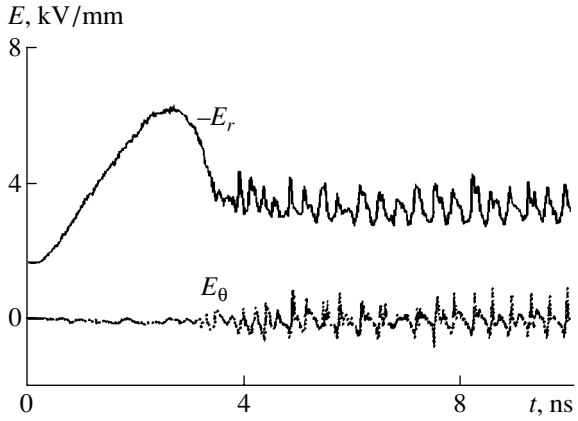

Fig. 3. Waveforms of the fields $E_r(t)$ and $E_\theta(t)$ at the point with the coordinates r = 0.4 cm and $\theta = 215°$.

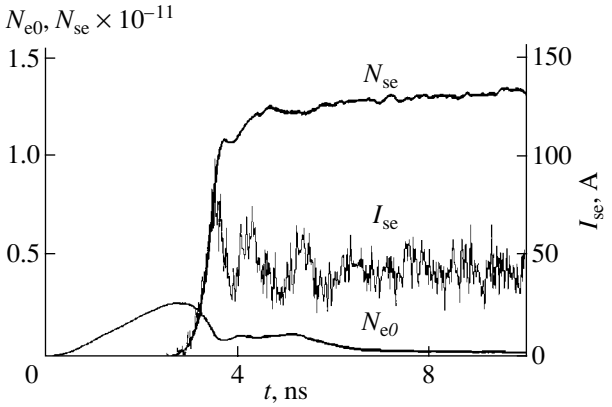

Fig. 4. Accumulation of primary, $N_{e0}(t)$, and secondary, $N_{se}(t)$, electrons in the electrode gap with time. The mean secondary emission current is specified by oscillations at a level $\delta I_{se}/I_{se} \approx 20\%$.

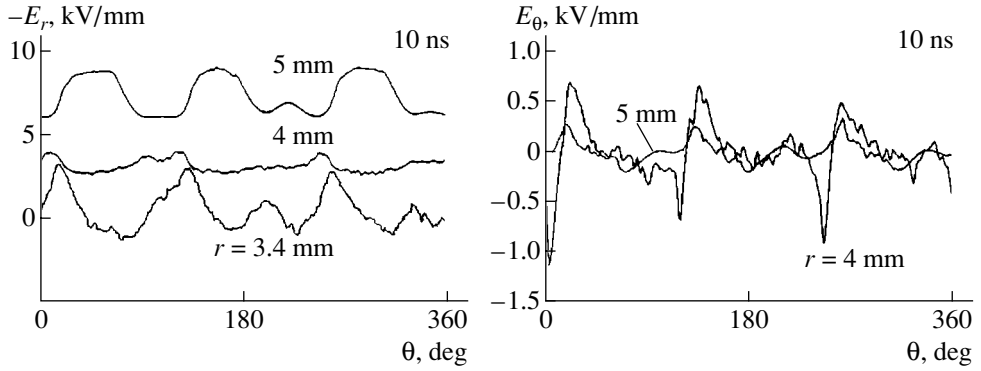

Fig. 5. Azimuthal distribution of the fields $E_r(r, \theta)$ and $E_\theta(r, \theta)$ at the time t = 10 ns.

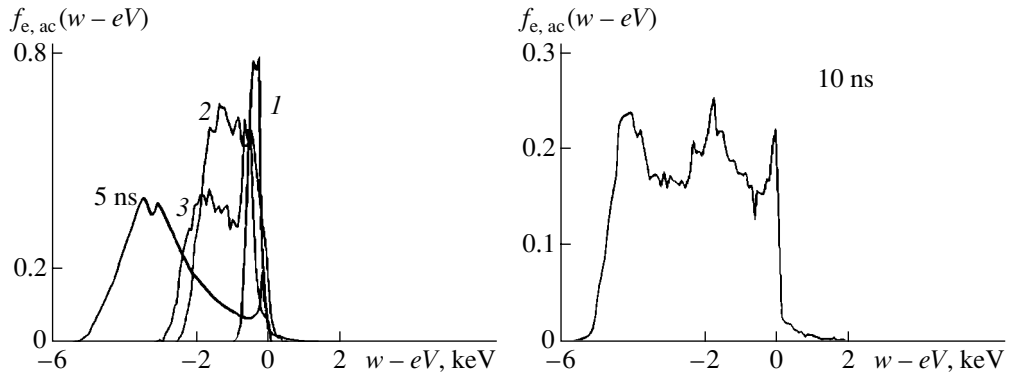

Fig. 6. Electron distribution function in the MD gap over total energy at different times (*w,* kinetic energy; *eV,* potential energy).

also contains the frequencies $\nu_{Le} = \nu_{ce}/2 = 3.5$ GHz, ν_{ce}, and its higher harmonics. In an ideal electrode system without azimuthal inhomogeneities (like the MD in Fig. 2), the spectrum is relatively simple with a peak near the Larmor frequency $\nu_{BCB} \approx \nu_{Le}$ (the fundamental frequency of BCB instability [3, 6]). In the presence of azimuthal inhomogeneities (the cathode axis is displaced, electron emission is nonuniform, etc.), the frequency spectrum becomes irregular.

5. RESULTS AND DISCUSSION

Macroscopic equations (2) and (3) can be obtained by summing the left- and right-hand sides of (6) over all electrons in the diode gap and averaging the sum over some time interval. Averaging is needed because the secondary electron flow $I_{se}(t)$, the leakage current toward the anode $I_{e,a}(t)$, the field $\mathbf{E}(r, \theta, t)$, etc. noticeably oscillate in time (Figs. 3–5). The equation thus obtained differs from (2) and (3) by the presence of the sum of the integrals that reflect the effect of the field E_θ on the electrons (see (6)). Under the steady-state conditions, the effects of the field $E_\theta(r, \theta, t)$ on the electrons moving toward the anode and on those bombarding the cathode compensate each other. In other words, the work done by the field $E_\theta(r, \theta, t)$ on the electrons in the gap over the time interval $\Delta t \gg t_{ce}$ equals zero:

$$\delta W_\theta = \int_{r_c}^{r_a} \int_0^{2\pi} r dr d\theta \int_t^{t+\Delta t} J_{e\theta}(r, \theta, t) E_\theta(r, \theta, t) dt \approx 0. \quad (14)$$

The electrons moving toward the anode are in phase with the decelerating field E_θ (r, θ, t). In this process, they lose energy and transfer the excess azimuthal momentum, which is gained via interaction with the field B_0, to the $\mathbf{E} \times \mathbf{B}$ field. Here, the $\mathbf{E} \times \mathbf{B}$ field serves as a time buffer for momentum and energy. The electrons that return to the cathode are in phase with the accelerating field $E_\theta(r, \theta, t)$ and acquire an excess energy and momentum. Such a mechanism is akin to electron–$\mathbf{E} \times \mathbf{B}$ field interaction in a magnetron (see Section 2). The difference is that integral (14) in the diode of a generator magnetron is less than zero, since the electric field removes energy (and azimuthal momentum) from the electrons and transfers it to anode resonators. Equations (3) for a magnetron also change, since the torque due to the electric field pressure exerted on the surface of anode slots is added. Therefore, the torque balance in a magnetron is usually fulfilled when the energy and momentum fluxes toward the cathode are minor.

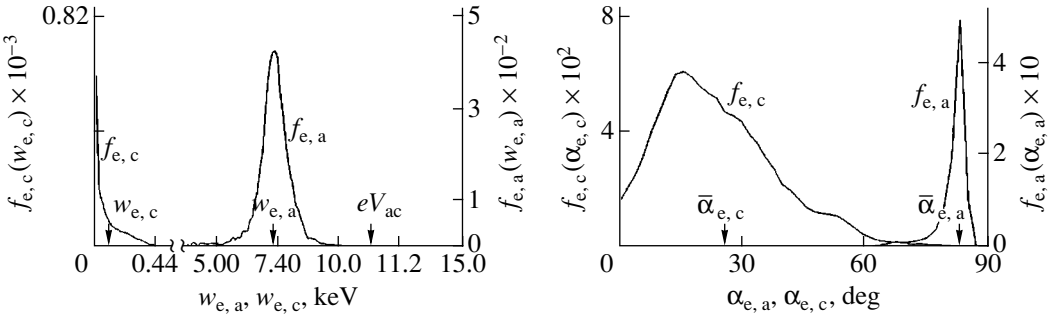

Fig. 7. Electron distribution functions over energy and angle of incidence (measured from the normal to the surface) for electrons bombarding the anode and cathode.

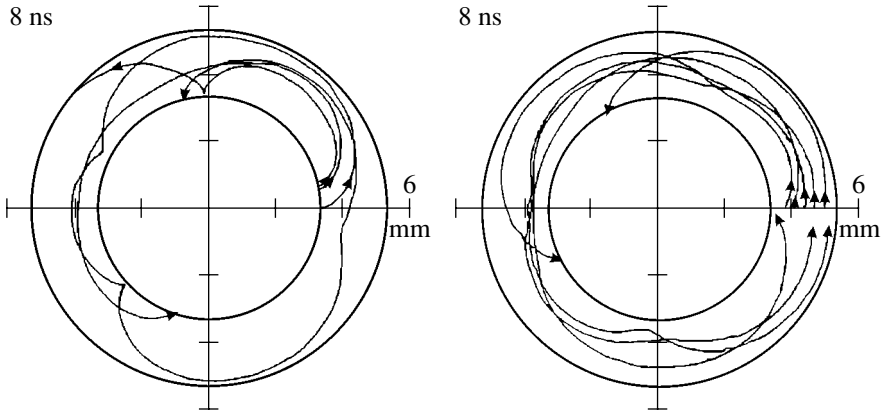

Fig. 8. Paths of several electrons within the time interval Δt = 7–8 ns. On the left: electrons start from the cathode (θ = 0°–20°); on the right: electrons go on traveling ($t \geq 7$ ns) with the azimuth angle θ = 0°.

The MD parameters are convenient to analyze in comparison to an electron diode with a high-emissivity cathode ($E_c = 0$) in the absence of a magnetic field ($B_0 = 0$). For the basic MD design (Fig. 2), the numerical simulation gives the limiting current $I_{cap,lim}$ = 216 A at a voltage V_{ac} = 11.2 kV. The electric charge in the electron diode for this case (**B** = 0) is $Q_{cap,lim}$ = 21.5 nC. This value is roughly equal to the charge in the MD gap, $Q_{se} = eN_{se}(t > 8\ \text{ns})$ = 21 nC, and exceeds the charge in a vacuum capacitor (no electrons in the gap) by a factor of 1.6: $Q_{cap} = C_0V_{ac} = Q_{cap,lim}/1.6$. Comparing these values with the MD parameters, we find the extent of magnetic insulation degradation $\chi_{e,a} = I_{e,a}/I_{cap,lim} \approx 1/50$ (the maximal value of $\chi_{e,amax} \approx 1/15$, was obtained in [9] for a platinum cathode and $B_0/B_{cr} \approx 1.1$), the relative secondary emission current $I_{se}/I_{cap,lim} \approx 1.5$, and $Q_{se}/Q_{cap,lim} \approx 1$. For estimation, we take the maximal charge in the MD to be equal to $Q_{em} = 1.5Q_{cap}$.

As was noted above, high instability in an MD diode occurs when the beam is dense. As a dense beam criterion, we choose the amount of the electron charge in the gap at the time t = 3 ns (the onset of instability, Fig. 4): $\xi_{e,ac}(t) = eN_{e0}(t = 3\ \text{ns})/Q_{em} \approx 1/4$. In practice, the amount of this charge can be estimated from only the MD external parameters (the waveforms in Fig. 1 and the diode dimensions). For example, $Q_{e0} = Q_a - Q_{cap} \approx$ 4.5 nC, where $Q_a(t = 3\ \text{ns}) = \int I_{a,c}(t)dt \approx 18.5$ nC is the charge accumulated at the MD anode and $Q_{cap} = C_0V_{ac}$ = 14 nC is the capacitive charge at V_{ac} = 12 kV. Thus, as a beam stability criterion, one may use the smallness condition for the exchange parameter ($\xi_{e,ac} \ll 1$):

$$\xi_{e,ac}(t_p) = Q_{e,a}/Q_e; \quad Q_{e,a}(t) = \int_0^{t_p} I_{e,a}(t)dt;$$
$$Q_{em} = e\int n_e dV \approx 1.5C_0/V_{a,c}, \tag{15}$$

where $\xi_{e,ac}$ is the exchange parameter, $Q_{e,a}(t_p)$ is the charge transferred by electrons for the process time t_p, and Q_{em} is the maximal charge in the diode.

A similar criterion was suggested long ago [13] upon estimating the stable acceleration domain for a collisionless plasma in a magnetic field. Experiments and numerical simulations of plasma acceleration between electrodes [32, 33] showed that the current sheath breaks down when the exchange parameter approaches unity for a homogeneous plasma or earlier if the plasma is inhomogeneous.

In crossed-field devices, multiple electron exchange in the gap ($\xi_{e,ac} \gg 1$) usually takes place during a voltage pulse V_{ac} ($t < t_p$) (see, e.g., [9, 10, 28, 29]). Our investigations show that strong BCB instability in the beam develops even at $\xi_{e,ac} > 0.3$. In many articles, magnetic-insulation devices were analyzed in terms of the hydrodynamic model of electron flow [1, 2] (which is also called the Brillouin, or drift, model). As follows from Fig. 8, the simulated paths of electrons in the MD gap, while extended in the azimuthal direction ($|v_{e,r}| \ll \bar{v}_{e,\theta}$), are not localized in a narrow layer. Instead, the electrons move across the gap (from the cathode to the anode and vice versa). Therefore, the hydrodynamic model does not work for "leaky" (at $\xi_{e,ac} > 0.3$) MDs.

The dynamics of BCB instability for electron flows in crossed-field devices was studied for voltages $V_{ac} \ll 0.5$ MV. However, for $V_{ac} \sim 1$ MV, the results will be qualitatively the same, since basic processes responsible for BCB instability occur in the near-cathode sheath, where electrons remain nonrelativistic. Such a conclusion relies, e.g., on data for kinetic losses of electrons arriving at the anode of a 2-MV magnetic-insulation beam-transport line [34].

In gaps with magnetic insulation of ions ($d_{ac} > r_{iL}$), the mechanism behind intense oscillation in electron beams exposed to crossed fields is apparently the same. In this case, the anode (or anodic plasma) serves as an emitter with nonstationary secondary emission. Such an assumption [35] is based on available experimental data for the generation of high-temperature proton beams by a gas-discharge magnetron-type injector and on the predictions of the BCB instability model.

CONCLUSIONS

The process of quasi-stationary electron leakage across a high magnetic field ($B_0/B_{cr} > 1.1$) is studied by numerical simulation. The electron flow in the gap splits into bunches in the azimuthal (drift) direction. The collective (turbulent) oscillation of charges takes place when the electrons interact with the $\mathbf{E} \times \mathbf{B}$ field and exchange momentum and energy with it. The self-organization of the turbulent electron flows is provided by the properties of the cathode ($J_{e0} \ll J_{cap,lim}$; the field at the cathode $E_c(t) \neq 0$) and also by secondary electron emission with a coefficient $k_{se} > 1$. The results of simulation are verified by the fulfillment of the energy balance, $P_{ac} = P_{e,a} + P_{e,BCB}$ (the spent power equals the absorbed power), and the torque balance (the torque arising when electrons leak across the magnetic field equals that transferred to the cathode and anode by incident electrons). The calculated results are shown to agree well with experimental data. The physical model of the oscillatory system (a dense charged particle flux in crossed $\mathbf{E} \times \mathbf{B}$ fields) applies to electron and ion beams. The numerical model of BCB instability may help to improve the efficiency of high-power microwave oscillators and high-current charged-beam injectors.

ACKNOWLEDGMENTS

This work was supported by the Russian Foundation for Basic Research (grant no. 03-02-17301).

REFERENCES

1. E. Okress, *Crossed-Field Microwave Devices* (Academic, New York, 1961; Inostrannaya Literatura, Moscow, 1961), Vol. 2.
2. E. P. Baranchikov, A. V. Gordeev, V. V. Korolev, and V. P. Smirnov, Zh. Éksp. Teor. Fiz. **75**, 2102 (1978) [Sov. Phys. JETP **48**, 1058 (1978)].
3. V. M. Fedorov, W. Schmidt, and Th. Westermann, in *Proceedings of the 9th International Conference on High Power Particle Beams (Beams'92), Washington, 1992*, Vol. 2, pp. 747–755.
4. A. V. Galagyan, I. V. Ruzhentsov, and A. A. Shadrin, Radiotekh. Élektron. (Moscow) **37**, 494 (1992).
5. S. D. Korovin and I. V. Pegel', Zh. Tekh. Fiz. **64** (3), 140 (1994) [Tech. Phys. **39**, 299 (1994)].
6. V. M. Fedorov, in *Proceedings of the 16th Symposium on Discharges and Electrical Insulation in Vacuum, Moscow, 1994*.
7. V. P. Gopinath, J. P. Verboncoer, and C. K. Birsdall, Phys. Plasmas **3**, 2766 (1996).
8. M. V. Nezlin, *Beam Dynamics in Plasmas* (Énergoatomizdat, Moscow, 1982).
9. R. L. Jepsen and M. W. Muller, J. Appl. Phys. **22**, 1196 (1951).
10. V. P. Tychinskiĭ and Yu. T. Derkach, Radiotekh. Élektron. (Moscow) **1**, 223 (1956).
11. I. M. Bronshteĭn and B. S. Fraĭman, *Secondary Electron Emission* (Nauka, Moscow, 1969).
12. I. M. Vigdorchik, Zh. Tekh. Fiz. **6**, 1657 (1936).
13. A. I. Morozov and L. S. Solov'ev, Dokl. Akad. Nauk SSSR **164**, 80 (1965) [Sov. Phys. Dokl. **10**, 834 (1965)].
14. B. B. Kadomtsev, in *Proceedings of the 6th European Conference on Controlled Fusion and Plasma Physics, Moscow, 1973*, Vol. 2, pp. 1–19.
15. A. V. Agafonov, V. M. Fedorov, and V. P. Tarakanov, in *Proceedings of the International Particle Accelerator Conference (PAC'97), Vancouver, 1997*, pp. 1299–1301.
16. A. V. Agafonov, V. P. Tarakanov, and V. M. Fedorov, Vopr. At. Nauki Tekh., Ser.: Yad. Fiz. Issled., Nos. 2–3, 137 (1997).
17. A. V. Agafonov, V. M. Fedorov, and V. P. Tarakanov, in *Proceedings of the 12th International Conference on High-Power Particle Beams (BEAMS'98), Haifa, 1998*, pp. 885–888.
18. A. V. Agafonov, V. P. Tarakanov, and V. M. Fedorov, in *Proceedings of the 17th Workshop on Charged Particle Accelerators, Protvino, 2000*, Vol. 2, pp. 67–70.
19. P. V. Kotetashvili, P. V. Rybak, and V. P. Tarakanov, Preprint No. 44 (Institute of General Physics, Russian Academy of Sciences, Moscow, 1991).

20. V. P. Tarakanov, *User's Manual for Code KARAT* (Berkley, Springfield, 1992).
21. *Centimeter-Range Magnetrons: A Collection of Articles* (Sov. Radio, Moscow, 1951).
22. S. I. Bychkov, *Theory and Application of Magnetron-Type Devices* (Sov. Radio, Moscow, 1967).
23. A. F. Alekseev and D. E. Malyarov, Zh. Tekh. Fiz. **10**, 1297 (1940).
24. M. F. Kopylov, Élektronika, Nos. 5–6, 28 (1996).
25. I. E. Tamm, *The Principles of Electricity Theory* (Nauka, Moscow, 1989).
26. R. P. Feynman, R. B. Leighton, and M. Sands, *The Feynman Lectures on Physics* (Addison-Wesley, Reading, 1965; Mir, Moscow, 1978).
27. Ch. K. Birdsall and A. B. Langdon, *Plasma Physics and Computer Simulation* (McGraw-Hill, New York, 1985; Énergoizdat, Moscow, 1989).
28. Y. M. Saveliev, W. Sibbnett, and D. M. Parkes, Phys. Plasmas **4**, 2319 (1997).
29. Yu. Ya. Volkolupov, A. N. Dovbnya, V. V. Zakutin, *et al.*, Zh. Tekh. Fiz. **71** (2), 98 (2001) [Tech. Phys. **46**, 227 (2001)].
30. D. A. Ganichev, V. F. Filatov, and S. A. Fridrikhov, Radiotekh. Élektron. (Moscow) **17**, 1639 (1972).
31. L. G. Blyakhman and V. E. Nechaev, Zh. Tekh. Fiz. **54**, 2163 (1984) [Sov. Phys. Tech. Phys. **29**, 1269 (1984)].
32. A. N. Babenko, E. P. Kruglyakov, and V. M. Fedorov, Prikl. Mekh. Tekh. Fiz., No. 3, 140 (1971).
33. K. B. Brushlinskiĭ and T. A. Ratnikova, Fiz. Plazmy **21**, 975 (1995) [Plasma Phys. Rep. **21**, 743 (1995)].
34. R. W. Stinnett and H. N. Woodal, in *Proceedings of the 5th International Conference on Pulsed Power, Arlington, 1985*, pp. 499–506.
35. V. M. Fedorov and V. P. Tarakanov, Trans. Fusion Technol. **35** (1T), 320 (1999).

Translated by V. Isaakyan