Technical Physics, Vol. 47, No. 12, 2002, pp. 1487–1494. Translated from Zhurnal Tekhnicheskoĭ Fiziki, Vol. 72, No. 12, 2002, pp. 1–8.

THEORETICAL AND MATHEMATICAL PHYSICS

Method of Integral Equation in the Theory of Weak-Guiding Inhomogeneous Optical Waveguides

L. I. Sotskaya and A. B. Sotskii

Institute of Applied Optics, Belarussian Academy of Sciences, Mogilev, 212793 Belarus

e-mail: ipo@physics.belpak.mogilev.by

Received May 15, 2002

Abstract—A method of integral equation is stated and applied in the analysis of the characteristic and noncharacteristic modes in 3D weak-guiding inhomogeneous optical waveguides. The dispersion curves for the modes of diffused channel waveguides are studied near the critical conditions in absorbing media. The accuracy of recovering the permittivity profile in a channel waveguide from the far-field radiation of the fundamental mode is estimated. *© 2002 MAIK "Nauka/Interperiodica".*

INTRODUCTION

Most optical waveguides are weak-guiding; that is, the permittivity of their interior is close to that of the environment [1, 2]. The key equation in the theory of such waveguides is the scalar equation [3]

$$\frac{\partial^2\psi}{\partial\xi^2} + \frac{\partial^2\psi}{\partial\eta^2} - V^2 b\psi = -V^2 f(\xi,\eta)\psi. \qquad (1)$$

Here, ψ is the transverse component of the electric field of the waveguide mode; the complex permittivity $\varepsilon(x, y)$ is represented as $\varepsilon(x, y) = \varepsilon_s + \Delta\varepsilon_0 f(\xi, \eta)$, where ε_s and $\Delta\varepsilon_0$ are, respectively, complex and real constants, $f(\xi, \eta)$ is a complex function satisfying the condition $f(\xi, \eta) \longrightarrow 0$ at $\rho = \sqrt{\xi^2 + \eta^2} \longrightarrow \infty$, $\xi = xM^{-1}$, $\eta = yM^{-1}$, and M is the scale factor characterizing the transverse dimension of a waveguide; $V = k_0 M\sqrt{\Delta\varepsilon_0}$, where $k_0 = 2\pi\lambda_0^{-1}$ is the free-space wave number; and $b = \gamma\Delta\varepsilon_0^{-1}$, where $\gamma = h^2 k_0^{-2} - \varepsilon_s$ and h is the mode propagation constant.

With appropriately chosen boundary conditions at infinity, Eq. (1) is the eigenfunction and eigenvalue problem. It can be solved with a number of numerical and analytical techniques [3, 4]. To date, the modes of circularly symmetric waveguides have been studied most extensively [3]. Waveguides of other types are usually investigated by variational and net methods [4]. However, these methods are totally inapplicable to noncharacteristic modes, while in the case of characteristic modes (eigenmodes), they fail under near-critical conditions and when the mode fields are calculated away from the waveguide.[1] Basically, problems such as finding conditions for the single-mode regime, the optimization of optical sensors, and the solution of inverse problems in integrated optics and fiber optics can be solved by the method of integral equation over area [3, 6, 7]. However, to date, this method was applied only to the eigenmodes of homogeneous waveguides with step real functions $f(\xi, \eta)$ [3, 6, 7]. In this work, we develop a version of the method of integral equation that can be applied to the study of various modes in inhomogeneous waveguides described by complex functions $f(\xi, \eta)$.

THE METHOD OF INTEGRAL EQUATION

Let a function $f(\xi, \eta)$ be analytical in a rectangular region G with the boundaries $\xi = \pm\alpha$ and $\eta = \pm\beta$ and take negligibly small values outside this region. Therefore, we can put $f(\xi, \eta) \equiv 0$ outside G and deal with a nonuniformly filled rectangular waveguide. Consider first the eigenmodes of such a waveguide. The fields of these modes are quadratically integrable functions [2, 5] that allow for the representation in the form of the Fourier integral

$$\psi(\xi,\eta) = \int_{\Gamma_1} \exp(-ik\xi)\hat{\psi}(k,\eta)dk, \qquad (2)$$

where the contour of integration Γ_1 coincides with the real axis of the complex plane k.

Substituting (2) into (1) yields an inhomogeneous ordinary differential equation for the function $\hat{\psi}(k, \eta)$. Its solution in view of the requirement $\lim_{|\eta|\to\infty} \hat{\psi}(k,\eta) = 0$ results in the relationship

$$\psi(\xi,\eta) = V^2(4\pi)^{-1}\int_{\Gamma_l} dk(i\lambda)^{-1}\iint_G f(\xi',\eta')\psi(\xi',\eta') \times \exp[ik(\xi'-\xi) - i\lambda|\eta'-\eta|]d\xi' d\eta'. \qquad (3)$$

[1] Modes are classified in terms of work [5].

1063-7842/02/4712-1487$22.00 © 2002 MAIK "Nauka/Interperiodica"

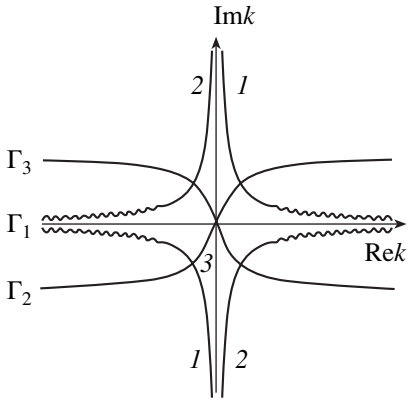

Fig. 1. Contours of integration Γ_l on the complex plane.

Here, $l = 1$, $\lambda = \sqrt{-V^2 b - k^2}$,

$$\mathrm{Re}\lambda < 0 \quad \text{for} \quad \mathrm{Im}b < 0, \tag{4}$$

$$\mathrm{Re}\lambda > 0 \quad \text{for} \quad \mathrm{Im}b > 0, \tag{5}$$

$$\mathrm{Im}\lambda < 0 \quad \text{for} \quad \mathrm{Im}b = 0. \tag{6}$$

The cuts of the functions $\lambda(k)$ are shown in Fig. 1 by wavy lines. They start at the branch points $k = \pm k_b$ ($k_b = \lambda(0)$) and run along the lines $\mathrm{Im}k = (\mathrm{Re}k)^{-1}V^2\mathrm{Im}b$, which are labeled *1*, *2*, and *3* according to conditions (4)–(6) (line *3* coincides with the imaginary axis). According to (4)–(6), $\mathrm{Im}\lambda < 0$ in the real axis; hence, the outer integral in (3) converges.

Expression (3) is an integral equation for the function $\psi(\xi, \eta)$ inside G and at the same time a computational formula for the mode field outside the waveguide. Changing the order of integration in (3) and taking advantage of the Macdonald function in the form

$$K_0(V\sqrt{b}R) = -0.5i\int_{-\infty}^{\infty}\lambda^{-1} \times \exp[ik(\xi' - \xi) - i\lambda|\eta' - \eta|]dk \tag{7}$$

(where $R = \sqrt{(\xi' - \xi)^2 + (\eta' - \eta)^2}$), we come to the well-known integral equation given, for example, in [6]. As applied to inhomogeneous waveguides, statement (3), however, seems to be preferable, since it does not involve the integration of singular function (7). One more advantage of such a statement is that there is no need to select an appropriate branch of the Macdonald function in describing the noncharacteristic modes of the waveguide. Yet, in this case, Eq. (3) needs some modification. In fact, for the noncharacteristic modes, conditions (4) and (5) should be replaced by [5]

$$\mathrm{Re}\lambda < 0 \quad \text{for} \quad \mathrm{Im}b > 0, \tag{8}$$

$$\mathrm{Re}\lambda > 0 \quad \text{for} \quad \mathrm{Im}b < 0. \tag{9}$$

Then, condition $\mathrm{Im}\lambda > 0$ will be fulfilled in the region between curves *2* (Fig. 1) in the case of (8) or between curves *1* in the case of (9). As a result, with the real axis retained as the contour of integration in (2) and (3), we come to the divergent integral over the variable k. This is quite natural because the noncharacteristic mode fields (nonrepresentable in terms of the standard Fourier integral [5]) grow with distance from the waveguide. To obviate this difficulty, (2) and (3) must involve integrals along the contours Γ_2 and Γ_3 that do not coincide with the real axis.

The contour Γ_2, which is parallel to the real axis at $|k| \longrightarrow \infty$ and envelops the branch points as shown in Fig. 1, meets condition (9). On this contour, $\mathrm{Im}\lambda < 0$ at $|k| \longrightarrow \infty$, which provides the existence of the right-hand part of Eq. (3). If the field varies as $\exp(i\omega t)$, this equation thus stated describes outgoing modes, since its solutions satisfy Eq. (1) and have the form of waves leaving the waveguide. To make sure of it, let us suppose that Eq. (3) is valid. Differentiating it, we obtain

$$\frac{\partial^2\psi}{\partial\xi^2} + \frac{\partial^2\psi}{\partial\eta^2} - V^2 b\psi = \int_{\Gamma_2} F(k, \eta)\exp(-ik\xi)dk,$$

$$F(k, \eta) = -\frac{V^2}{2\pi}\int_{-\infty}^{\infty} f(x, \eta)\psi(x, \eta)\exp(ikx)dx.$$

Here, $F(k, \eta)$ is an integer function of k such that $F(k, \eta) = O(k^{-1})$ at $|\mathrm{Re}k| \longrightarrow \infty$ [8]. Thus, the contour Γ_2 in the above expression allows for a deformation toward the real axis, which, according to the Fourier theorem [8], means that expression (1) is fulfilled identically. To estimate the behavior of the function $\psi(\xi, \eta)$ outside the waveguide, we will take into account that expression (3) can be represented as [9]

$$\psi(\xi, \eta) = \frac{V^2\exp(-ik_b\rho)}{\sqrt{8\pi ik_b\rho}}\iint_G f(\xi', \eta')\psi(\xi', \eta') \times \exp\left[i\left(\frac{\xi'\xi}{\rho} + \frac{\eta'\eta}{\rho}\right)\right]d\xi' d\eta' \tag{10}$$

for $|k_b\rho| \longrightarrow \infty$ and $l = 1$, 2, and 3.

The mode behavior mentioned above is a consequence of (10) and also of the inequalities $\mathrm{Re}k_b > 0$ and $\mathrm{Im}k_b > 0$, which result from (9). Note that $\mathrm{Im}k_b < 0$ when conditions (4)–(6) are met. According to (10), this means that the characteristic mode fields decay away from the waveguide.

The contour Γ_3 in Fig. 1 meets inequalities (8). At $|k| \longrightarrow \infty$, this contour runs parallel to the real axis and provides the condition $\mathrm{Im}\lambda < 0$. The adequacy of the associated integral equation to Eq. (1) is proved as above. In our case, from (10) and inequalities $\mathrm{Re}k_b < 0$ and $\mathrm{Im}k_b > 0$, which result from (8), it follows that the noncharacteristic mode fields under consideration

approach the waveguide from infinity. Such modes do not have a clear physical meaning [5]. At the same time, their consideration is helpful in elucidating the critical conditions for characteristic modes (see below).

Numerical solutions to the integral equations stated above will be sought by the method of quadratures; i.e., the integral over space variables in (3) is replaced by the finite sum [10]. The waveguide cross section is covered by a rectangular net of equidistant lines with the coordinates

$$\xi_i = i\Delta\xi \quad (i = 0, \ldots, \pm m;\ \Delta\xi = \alpha m^{-1}),$$

$$\eta_j = j\Delta\eta \quad (j = 0, \ldots, \pm n;\ \Delta\eta = \beta n^{-1}).$$

As a result, the region G will be partitioned into $2m \times 2n$ rectangles of sides $\Delta\xi$ and $\Delta\eta$. Now we make a linear interpolation of the product $\Phi(\xi, \eta) = f(\xi, \eta)\psi(\xi, \eta)$ within each of the rectangles by the rule

$$\Phi(\xi, \eta) = \Phi_0 + \Phi'_\xi(\xi - \xi_0) + \Phi'_\eta(\eta - \eta_0),$$

$$\Phi_0 = 0.25[\Phi(\xi_i, \eta_j) + \Phi(\xi_{i+1}, \eta_j) + \Phi(\xi_i, \eta_{j+1}) + \Phi(\xi_{i+1}, \eta_{j+1})],$$

$$\Phi'_\xi = (2\Delta\xi)^{-1}[\Phi(\xi_{i+1}, \eta_j) + \Phi(\xi_{i+1}, \eta_{j+1}) - \Phi(\xi_i, \eta_j) - \Phi(\xi_i, \eta_{j+1})], \quad (11)$$

$$\Phi'_\eta = 0.5(\Delta\eta)^{-1}[\Phi(\xi_i, \eta_{j+1}) + \Phi(\xi_{i+1}, \eta_{j+1}) - \Phi(\xi_i, \eta_j) - \Phi(\xi_{i+1}, \eta_j)],$$

where $\xi_0 = 0.5(\xi_{i+1} + \xi_i)$ and $\eta_0 = 0.5(\eta_{j+1} + \eta_j)$.

With (11), space integrals in (3) will take the form of a linear combination of the functions $\psi(\xi, \eta)$ at nodes of interpolation. Thus, assuming that ξ and η in (3) equal ξ_μ ($\mu = 0, \ldots, \pm m$) and η_ν ($\nu = 0, \ldots, \pm n$) and introducing the combined indices $\kappa = (m - \mu + 1)(2n + 1) + \nu - n$ and $\sigma = (m - i + 1)(2n + 1) + j - n$, we arrive at the set of homogeneous algebraic equations

$$\sum_{\sigma=1}^{N} M_{\kappa\sigma} X_\sigma = 0 \quad (\kappa = 1, 2, \ldots, N). \quad (12)$$

Here, $N = (2m + 1)(2n + 1)$, $X_\sigma = \psi(\xi_i, \eta_j)$, $i = m - \mathrm{Int}(\sigma(2n + 1)^{-1})$ (the operator Int means the separation of the integer part of a number), and $j = \sigma + n - (2n + 1)(m + 1 - i)$. The values $M_{\kappa\sigma}$ are represented by the sum of three integrals:

$$\int_{\Gamma_l} \lambda^{-1} D_1 D_2 dk, \quad \int_{\Gamma_l} \lambda^{-1} S_1 D_2 dk, \quad \int_{\Gamma_l} \lambda^{-1} D_1 S_2 dk, \quad (13)$$

where

$$D_1 = (ik)^{-1}(E_1^{(1)} - E_1^{(0)}),$$

$$S_1 = -U(i - \mu)k^{-2}[(A_1 - 1)E_1^{(1)} - (A_1 + 1)E_1^{(0)}],$$

$$D_2 = -(i\lambda)^{-1}(E_2^{(1)} - E_2^{(0)}),$$

$$S_2 = U(\nu - j - 1)\lambda^{-2}[(A_2 - 1)E_2^{(1)} + (A_2 + 1)E_2^{(0)}],$$

$$A_1 = 0.5ik\Delta\xi, \quad A_2 = -0.5i\lambda\Delta\eta,$$

$$E_1^{(l)} = \exp[2A_1(I + l)], \quad E_2^{(l)} = \exp[2A_2(J + l)],$$

$$I = (i - \mu)U(i - \mu) + (\mu - i - 1)[1 - U(i - \mu)],$$

$$J = (\nu - j - 1)U(\nu - j - 1) + (j - \nu)[1 - U(\nu - j - 1)],$$

$U(x) = 1$ at $x \geq 0$ and $U(x) = 0$ at $x < 0$.

It is easy to check that the integrands in (13) are continuous functions and their absolute values (at $|k| \longrightarrow \infty$) are on the order of $|k|^{-2}$. Therefore, integrals (13) converge uniformly and can be readily calculated numerically.

The key point in solving set (12) is finding the complex values of the spectral parameter b at which the determinant of the matrix M ($\det M$) vanishes. Note in this respect that the transition from conditions (4) and (5) to (8) and (9), respectively, causes the deformation of the contours of integration in (3) without crossing the branch points of the function $\lambda(k)$ (Fig. 1). It follows that at $V \neq 0$, $\det M$ is an analytical function of the complex variable b everywhere except the singular point $b = 0$ (at $V^2 b = 0$, the two branch points of the function $\lambda(k)$ merge together with the result that integrals (13) become divergent). Thus, the method of contour integration is efficient for solving the transcendental equation $\det M = 0$ [11].

As follows from calculations, the rank of the matrix M equals $N - 1$ if $\det M = 0$. This allows us to express all the components of the vector X through one of them, i.e., to calculate the spatial distribution of the mode field. Note also that the accuracy of solving Eq. (3) by the method of quadratures has the same order of magnitude as the accuracy of replacing the space integral in (3) by the finite sum [10]. Upon the interpolation of (11), the latter accuracy is a small of the order of $[\max(\Delta\xi, \Delta\eta)]^2$. This fact can be used to cut the size of calculation (see below).

ANALYSIS OF CHANNEL WAVEGUIDES NEAR THE CRITICAL CONDITIONS

Since Eq. (3) refers to the mode field in the interior of the waveguide, our approach ignores the field values away from the waveguide, which are usually difficult to take into account in variational and net methods. Because of this, the method of integral equation is efficient in studying the mode behavior near the critical conditions. Studies of this kind for circular and elliptic transparent homogeneous waveguides were performed in [6, 7]. However, the consideration of inhomogeneous guiding media with absorption is a topical problem. The conversion of characteristic modes to noncharacteristic ones has also been insufficiently explored. In

Table 1. Calculated values of b vs. discretization m of the guiding region

V	m	Re $b \times 10^3$	Re $b_0 \times 10^3$	Im $b \times 10^3$	Im $b_0 \times 10^3$
2.550	8	−3.337	−2.047	3.278	1.576
	16	−2.370	−2.019	2.001	1.591
	24	−2.175	−2.016	1.773	1.592
	32	−2.105	−2.016	1.694	1.593
	40	−2.073	−2.016	1.658	1.593
	48	−2.055	–	1.638	–
2.650	8	1.565	4.457	−0.1003	−0.1363
	16	3.734	4.583	−0.1273	−0.1358
	24	4.206	4.595	−0.1320	−0.1358
	32	4.376	4.597	−0.1337	−0.1358
	40	4.456	4.597	−0.1344	−0.1358
	48	4.499	–	−0.1348	–
3.486	8	91.96	103.9	−0.3018	−0.3327
	16	100.9	104.2	−0.3250	−0.3334
	24	102.8	104.2	−0.3297	−0.3336
	32	103.4	104.2	−0.3314	−0.3336
	40	103.7	104.2	−0.3322	−0.3336
	48	103.9	–	−0.3326	–

this section, these points are highlighted by an example of inhomogeneous optical channel waveguides.

Channel waveguides incorporated into the surface layer of insulators are used in many integrated optical devices [1]. They feature a large (in comparison with $\Delta\varepsilon_0$) jump of the permittivity at their surface [1]. Because of this, the modes of these waveguides are described (within the scalar approximation) by the odd (relative to the axis $\eta = 0$) eigenfunctions of Eq. (1) subject to $f(\xi, -\eta) = f(\xi, \eta)$ [12, 13].

Let us consider a diffused channel waveguide with

$$f(\xi, \eta) = 0.5(1 + i\delta)\exp(-\eta^2) \times [\mathrm{erf}(\xi + 0.5\phi) - \mathrm{erf}(\xi - 0.5\phi)], \quad (14)$$

where δ and ϕ are real parameters. The former is responsible for absorption, and the latter is the relative width of a diffusate strip [1].

The bounds of the region G will be $\beta = 2.5$ and α satisfying the condition

$$[\mathrm{erf}(\alpha + 0.5\phi) - \mathrm{erf}(\alpha - 0.5\phi)][2\mathrm{erf}(0.5\phi)]^{-1} = \exp(-\beta^2).$$

With such a choice, the properties of the nonuniformly filled rectangular waveguide are practically indistinguishable from those of a waveguide described by function (14) throughout the space. Since function (14) is even, set (12) is applicable only to the values of $\psi(\xi_i, \eta_j)$ in the fourth quadrant. In this case, the dimension of set (12) equals $N = (m + 1)(n + 1)$. Results that follow were found by solving this set for $m = n$.

In analytically estimating the ϕ dependence of the critical values of $V(V_c)$, it is of interest to consider the limiting case $\phi \longrightarrow 0$ in function (14) [14]. By virtue of the asymptotic expansion

$$0.5[\mathrm{erf}(\xi + 0.5\phi) - \mathrm{erf}(\xi - 0.5\phi)] = \Phi(\sqrt{\pi})^{-1}\exp(-\xi^2) + O(\phi^3), \quad (15)$$

we arrive (for $\phi \longrightarrow 0$) at Eq. (1) where

$$f(\xi, \eta) = (1 + i\delta)\exp(-\rho^2). \quad (16)$$

Table 1 and Fig. 2 show the solutions to the equation $\det M = 0$ subject to (16) and $\alpha = \beta = 2.5$. Table 1 illustrates the dependence of the roots of the equation $\det M = 0$ on the parameter m of discretization of the guiding region for $\delta = -0.001$ and values of V that are near-critical for the fundamental waveguide mode. The value $V = 2.55$ corresponds to the noncharacteristic mode (case (8)); $V = 2.65$, to the characteristic one.

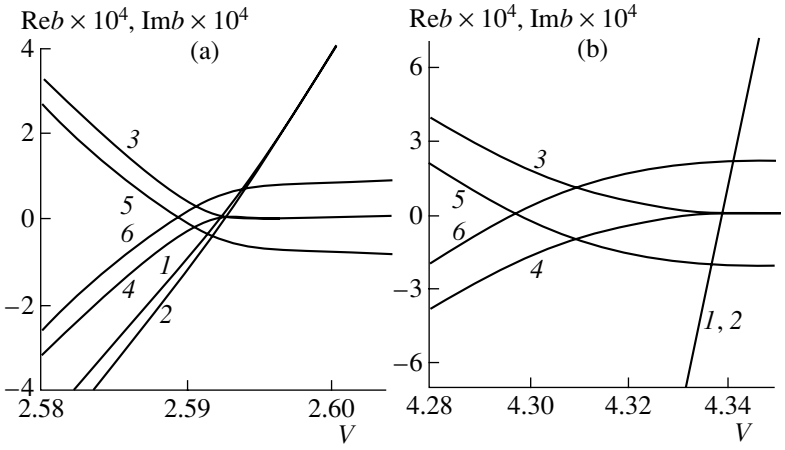

Fig. 2. Dependences $b(V)$ near the critical conditions for (a) the fundamental and (b) first higher modes of the channel waveguide. For Re$b(V)$: $\delta = 0$ (*1*) and $\pm 10^{-3}$ (*2*); for Im$b(V)$: $\delta = 0$ (*3*, *4*), -10^{-3} (*5*), and 10^{-3} (*6*).

According to Table 1, the convergence of the results near the critical conditions is relatively slow. In this situation, it is reasonable to take advantage of the above-mentioned quadratic dependence of the calculation accuracy on the step of discretization of the guiding region:

$$b(m) = b_0 + Cm^{-2} + O(m^{-3}), \quad (17)$$

where $b(m)$ is the zero of $\det M$ corresponding to the parameter of discretization m, $b_0 = b(\infty)$, and C is a constant.

Ignoring the small $O(m^{-3})$ in (17), we have

$$b_0 = b(m_1) + m_2^2[b(m_2) - b(m_1)](m_2^2 - m_1^2)^{-1}. \quad (18)$$

The values of b_0 calculated by (18) for $m_1 = m$ and $m_2 = m + 8$ are listed in Table 1. It is seen that expression (18) yields a reliable estimate of b_0 if m is sufficiently small. This expression becomes still more efficient if V deviates significantly from the critical value. In Table 1, this is indicated by the data corresponding to function (14) when $\phi = 2$, $\delta = -0.001$, and $V = 3.486$ (this waveguide will be studied in greater detail in the next section).

To verify the data obtained, we note that the modes of a radially symmetric waveguide that is described by functions (16) at $\rho < \beta$ and by $f(\xi, \eta) \equiv 0$ at $\rho > \beta$ allow for an independent calculation by the Runge–Kutta method [3]. It yields $b = b_1 = -2.044 \times 10^{-3} + i1.622 \times 10^{-3}$ at $V = 2.55$ and $b = b_2 = 4.561 \times 10^{-3} - i1.357 \times 10^{-4}$ at $V = 2.65$. These values are close to those given in the table. The small discrepancy between them are due to the use of different domains where $f(\xi, \eta) \equiv 0$, since the solution of the equation $\det M = 0$ subject to $f(\xi, \eta) \equiv 0$ and $\rho > \beta$ yields values of b_0 coincident with b_1 and b_2 to the decimal place mentioned above (for appropriate V and $m_1 \geq 32$).

Figure 2 shows the behavior of the $b(V)$ curves near the critical conditions for the fundamental, ψ_{11}, and the first higher, ψ_{21}, modes of a channel waveguide (the subscripts *1* and *2* refer to the number of maxima of the function $|\psi(\xi, \eta)|$ along the axes 0ξ and 0η for $\eta > 0$). The curves depicted were calculated with relationship (18) for $m_1 \geq 24$ and $m_2 = m_1 + 8$.

First, consider the dependences $b(V)$ for $\delta \neq 0$. Their shape implies that the real and imaginary parts of b vanish at different V. Therefore, since $\det M$ is an analytical function of b and V at $b \neq 0$ and $V \neq 0$, the solution to the equation $\det M = 0$ represents a continuous dependence $b(V)$. This means, in particular, that characteristic-to-noncharacteristic mode conversion takes place continuously. As follows from (4), (5), (8), and (9), the transition occurs at $V = V_c$, where V_c is a root of the equation $\mathrm{Im}\, b(V_c) = 0$. From Fig. 2, it follows that the characteristic mode ($V > V_c$) and noncharacteristic mode ($V < V_c$) are fast ($\mathrm{Re}\, b < 0$). This statement is valid for both $\delta > 0$ and $\delta < 0$, since the change of the sign of δ changes only the sign of $\mathrm{Im}\, b$, leaving the value of $\mathrm{Re}\, b$ the same (Fig. 2). Formally, this is explained by the fact that the solutions to Eqs. (1) and (3) found for the opposite signs of δ in (14) and (16) are complex conjugate functions that have different physical meanings. With $\delta < 0$, a characteristic mode becomes physically impossible when passing from the range $V > V_c$ to the range $V < V_c$ (case (8)), while with $\delta > 0$, it transforms into a leaky mode that can be observed experimentally. The latter situation may take place both in absorbing media (if environmental losses exceed those in the guiding channel) and in an amplifying waveguide.

Table 2. Critical values of V for the fundamental and first higher modes of the channel waveguide

ϕ	δ	$V_c^{(11)}$	$\bar{V}_c^{(11)}$	$V_c^{(21)}$	$\bar{V}_c^{(21)}$
0.5	0	4.905	4.881	8.169	8.169
0.5	$\pm 10^{-3}$	4.901	4.875	8.092	8.094
1	0	3.520	3.451	5.776	5.777
1	$\pm 10^{-3}$	3.516	3.447	5.722	5.723
0.5	0	2.941	2.819	4.714	7.717
0.5	$\pm 10^{-3}$	2.937	2.815	4.671	4.673

The case $\delta = 0$ is singular, since $b = 0$ for $V = V_c$ [6]. As was already noted, the matrix elements of set (12) become singular for $b = 0$; therefore, a rigorous calculation of appropriate V_c calls for the development of a special algorithm [6]. A simpler way, however, is to estimate V_c by extrapolating the curve $\mathrm{Re}\, b(V)$, which can be calculated in the ranges $V < V_c$ and $V > V_c$ corresponding to $|\mathrm{Re}\, b| \geq d > 0$. The validity of such an extrapolation follows from Fig. 2, where $d = 10^{-5}$. Here, the dependence $\mathrm{Re}\, b(V)$ for $V < V_c$ refers to two noncharacteristic modes described by complex conjugate functions $\psi(\xi, \eta)$.

From the above consideration, we can conclude that the functions $V_c(\delta)$ are even and reach maxima at $\delta = 0$. In particular, according to Fig. 2, $V_c(\pm 0.001) = 2.589$ and $V_c(0) = 2.592$ for the mode ψ_{11} and $V_c(\pm 0.001) = 4.299$ and $V_c(0) = 4.339$ for the mode ψ_{21}. This is also valid for model (14), as demonstrated by Table 2, where the critical values of V for various ϕ and δ are listed. Here, $V_c^{(\mu\nu)}$ stands for V_c for the mode $\psi_{\mu\nu}$ of a waveguide with the function $f(\xi, \eta)$ like (14). The value of $V_c^{(\mu\nu)}$ is found from the dependence $b(V)$ following the above scheme:

$$\bar{V}_c^{(\mu\nu)} = V_c^{(0)} \sqrt{\phi^{-1}\sqrt{\pi}}, \quad (19)$$

where $V_c^{(0)}$ is V_c for the mode $\psi_{(\mu\nu)}$ of a waveguide with the function $f(\xi, \eta)$ like (16).

Expression (19) represents an asymptotic approximation to $V_c^{(\mu\nu)}$ that follows from relationships (1) and (14)–(16). As follows from Table 2, it is quite an appropriate estimator of the single-mode conditions for $\phi \le 1.5$.

RECONSTRUCTION OF THE CHANNEL WAVEGUIDE PERMITTIVITY PROFILE

The solution of the inverse problem, i.e., the reconstruction of the permittivity profile $\Delta\varepsilon(x, y) = \varepsilon(x, y) - \varepsilon_s$ in inhomogeneous optical waveguides, is vital for designing waveguides with desired properties [1, 15, 16]. Among the $\Delta\varepsilon(x, y)$ reconstruction methods currently available, those where the spatial intensity distribution of the fundamental mode is measured are the most efficient. In these methods, the transverse component of the electric field of the mode $\psi(x, y)$ is recovered and then substituted into (1) to give a formula for directly computing $\Delta\varepsilon(x, y)$ [15, 16]. A technique for implementing such a procedure as applied to planar waveguides with $\varepsilon(x, y) = \varepsilon(y)$ was suggested in [17]. It is based on the measurement of the mode intensity angular distribution in the far-field zone with the subsequent determination of $\varepsilon(y)$ by the analytic continuation technique. In this section, we apply the method of integral equation to estimate the efficiency of the analytic approach for 3D waveguides.

Let the fundamental mode of an optical waveguide placed in the half-space $z < 0$ be emitted from the end face (the plane $z = 0$) of the waveguide into the half-space $z > 0$ filled with a homogeneous medium with a wave number k_a. The emission intensity is recorded by a photodetector array placed in the plane $z = L$. Assume that the cross section of the waveguide has two axes of symmetry: $x = 0$ and $y = 0$ (in the case of a channel waveguide placed in the domain $y > 0$, the axis of symmetry $y = 0$ is provided by a mirror surface brought into contact with the waveguide surface [17]). Then, if the mode is polarized along the $0x$ axis, we have (within the scalar approximation of the electric field components) [15, 16]

$$E_x\big|_{z=+0} = D\psi(x, y), \quad E_y \equiv 0, \tag{20}$$

where D is a constant.

For transparent media subject to (20) and $L \longrightarrow \infty$, the measured intensity S of the mode is given by [15, 16]

$$S(k, \varphi) = \left(\sqrt{1-k^2}\right)^3 (1 - k^2\sin^2\varphi)[\hat{\psi}(k, \varphi)]^2, \tag{21}$$

$$\hat{\psi}(k, \varphi) = \sqrt{s}\int_{-\infty}^{+\infty}\int \psi(x, y) \times \exp[ikk_a(x\cos\varphi + y\sin\varphi)]dxdy. \tag{22}$$

Here, $\hat{\psi}(k, \varphi)$ and $\psi(x, y)$ are real functions normalized so that $\max[S(k, \varphi)] = 1$, $k = \sin\theta$, θ and φ are the angles of the spherical coordinate system that are related to the Cartesian coordinates of the point of observation as $x_0 = r_0\cos\varphi$ and $y_0 = r_0\sin\varphi$, $r_0 = kL(\sqrt{1-k^2})^{-1}$, and the factor s is defined by the mode evenness condition $\psi(x, y) = s\psi(x, -y)$.

Some modification of the results of [17] in view of relationships (1), (21), and (22) yields the reconstruction formulas

$$\Delta\varepsilon(x, y) = \gamma + (k_a k_0^{-1})^2 R_2 R_1^{-1},$$

$$R_1 = \sum_{l=0}^{p} u_l \sum_{i=0}^{r} C_i(\varphi_l) I_{il}, \quad R_2 = \sum_{l=0}^{p} u_l \sum_{i=0}^{r} C_i(\varphi_l) I_{i+1\,l},$$

$$I_{il} = \int_0^{\infty} \exp[-\alpha(\varphi_l)k^2]k^{2i-0.5(s+3)} H(k, \varphi_l) \times [k^2 + (k_0 k_a^{-1})^2\gamma]^{-1} dk, \tag{23}$$

$$H(k, \varphi) = \cos(kk_a x\cos\varphi)[(1-s)\sin(kk_a y\sin\varphi) + (1+s)\cos(kk_a y\sin\varphi)],$$

$u_0 = u_p = 0.5$; $u_l = 1$ ($l = 1, 2, \ldots, p-1$); and the parameters $C_i(\varphi_l)$, $\alpha(\varphi_l)$, and γ are found by minimizing the sum

$$U = \sum_{l=0}^{p}\sum_{j=0}^{q} W_{lj}[\hat{\psi}(k_j, \varphi_l) - \hat{\psi}_r(k_j, \varphi_l)]^2$$

by the gradient method. Here, $W_{lj} = 1$ at $S(k_j, \varphi_l) \ge \bar{S}$ or $W_{lj} = 0$ at $S(k_j, \varphi_l) < \bar{S}$, $\varphi_l = 0.5\pi lp^{-1}$, and $k_j = jaq^{-1}$. The constants $\bar{S}$, a, p, and q depend on the dynamic range of intensity measurement, photodetector array dimension, and experimental sample size. The functions $\hat{\psi}(k_j, \varphi_l)$ are expressed through experimentally found values of $S(k_j, \varphi_l)$ in view of (21):

$$\hat{\psi}_r(k, \varphi) = \exp[-\alpha(\varphi)k^2][k^2 + (k_0 k_a^{-1})^2\gamma]^{-1} \times \sum_{i=1}^{r} C_i(\varphi)k^{2i-0.5(s+3)},$$

where the parameter r characterizes the order of reduction of the infinite product when the Fourier transform of the mode field is represented through its complex zeros [17].

It should be noted that expression (23) is strict (within the scalar approximation) only if $q > r + 1 + 0.5(1 - s)$, $r \longrightarrow \infty$, $p \longrightarrow \infty$, and $L \longrightarrow \infty$ [17]. Actually, however, r, p, q, and L are finite. This introduces systematic errors in the reconstruction. The errors can be estimated with the method of integral equation. In

this case, the function $S(k, \varphi)$ is first calculated for finite L for a waveguide with a known $\Delta\varepsilon(x, y)$ distribution, and then reconstructed function (23) is compared with the initial (known) distribution.

With regard for (1) and (20), as well as the fact that the function $E_x(x, y, z)$ in the domain $z > 0$ obeys the Helmholtz equation and satisfies Sommerfeld radiation condition, we find

$$S(k, \varphi) = (k_0 M)^4 (\Delta\varepsilon_0 L)^2 k_a^{-3}$$

$$\times \mathrm{Im}\left\{\left[\iint_G f(\xi, \eta)\psi(\xi, \eta)A(M\xi - x_0, M\eta - y_0, L)d\xi d\eta\right]^*\right.$$

$$\times \iint_G f(\xi, \eta)\psi(\xi, \eta)[\exp(-ik_a r_1)\zeta^{-2} r_1^{-3} \qquad (24)$$

$$\times (1 - ik_a r_1 + (M\xi - x_0)^2 r_1^{-2}(3 + 3ik_a r_1 - k_a^2 r_1^2))$$

$$\left. - (2\zeta)^{-1}\int_0^\infty \exp(-\zeta\tau)(B(\tau) + B(-\tau))d\tau]d\xi d\eta\right\},$$

where the asterisk means complex conjugation,

$$A(x, y, z) = (2\zeta)^{-1}\int_0^\infty \exp(-\tau\zeta)$$

$$\times [K(x, y, z + i\tau) + K(x, y, z - i\tau)]d\tau,$$

$$\zeta = \sqrt{k_a^2 + k_0^2\gamma}, \quad r_1 = \sqrt{(M\xi - x_0)^2 + (M\eta - y_0)^2 + L^2},$$

$$K(x, y, z) = z r_0^{-3}\exp(-ik_a r_0)(1 + ik_a r_0),$$

$$r_0 = \sqrt{x^2 + y^2 + z^2},$$

$$B(\tau) = \exp(-ik_a r_2) r_2^{-3}\{1 + ik_a r_2$$

$$- (L + i\tau)^2 r_2^{-2}(3 + 3ik_a r_2 - k_a^2 r_2^2)$$

$$+ (\zeta r_2)^{-2}[3 + 3ik_a r_2 - k_a^2 r_2^2 - ((M\xi - x_0)^2$$

$$+ (L + i\tau)^2) r_2^{-2}(15 + 15ik_a r_2 - 6k_a^2 r_2^2 - ik_a^3 r_2^3)$$

$$+ (M\xi - x_0)^2 (L + i\tau)^2 r_2^{-4}(105 + 105ik_a r_2$$

$$- 45k_a^2 r_2^2 - 10ik_a^3 r_2^3 + k_a^4 r_2^4]\},$$

$$r_2 = \sqrt{(M\xi - x_0)^2 + (M\eta - y_0)^2 + (L + i\tau)^2}.$$

Expression (24) is rigorous and involves the integration of the function $\psi(\xi, \eta)$ over the region G alone, for which Eq. (3) was stated. At $L \longrightarrow \infty$, this expression passes to (21) where

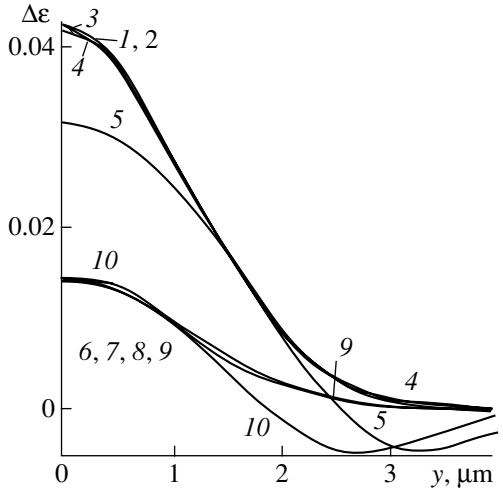

Fig. 3. Reconstruction of the function $\Delta\varepsilon(x, y)$ for the channel waveguide. $x = 0$ (*1*–*5*) and 2.2 μm (*6*–*10*). (*1*, *6*) Exact functions. Reconstructed functions are shown for $L = \infty$ (*2*, *7*), 5 (*3*, *8*), 1 (*4*, *9*), and 0.5 cm (*5*, *10*).

$$\hat{\psi}(k, \varphi) = \Delta\varepsilon_0 M^2\sqrt{s}[\gamma + (k_a k_0^{-1})^2 k^2]^{-1} \qquad (25)$$

$$\times \iint_G f(\xi, \eta)\psi(\xi, \eta)\exp[ikk_a M(\xi\cos\varphi + \eta\sin\varphi)]d\xi d\eta.$$

The equivalency of (22) and (25) follows from Eq. (1).

The reconstruction of the function $\Delta\varepsilon(x, y)$ for a single-mode ($\lambda_0 = 0.6328$ μm) channel waveguide with $f(\xi, \eta)$ of type (14) and $\phi = 2$, $M = 1.57$ μm, $\Delta\varepsilon_0 = 0.05$, and $\delta = 0$ is presented in Fig. 3. The curves were obtained by calculating $S(k_j, \varphi_l)$ from (24) at $L \neq \infty$ and from (21) and (25) at $L = \infty$. The integrals over the region G were calculated with the values of the function $\psi(\xi, \eta)$ found at the nodes of interpolation by solving set (12) for $m = n = 48$ (in this case, $\mathrm{Im}b = 0$ and the dependence $\mathrm{Re}b(m)$ coincides with that given in Table 1). In the calculations, $r = 3$, $a = 0.5$, $p = 20$, $q = 25$, and $\bar{S} = 10^{-6}$. The correct reconstruction of $\Delta\varepsilon(x, y)$ is observed at $L \geq 1$ cm. In the limit $L = \infty$, the function recovered is virtually indistinguishable from the exact function, indicating that the description of the mode field by the method of integral equation is accurate.

CONCLUSION

We showed that the method of integral equation is adequate in analyzing the properties of the characteristic and noncharacteristic modes of inhomogeneous optical waveguides. With this method, one can not only find the critical conditions and estimate errors involved in the solution to the inverse problem but also study polarization effects and effects of leakage observed in

anisotropic media, which are asymptotically described with Eq. (1) [2, 12, 13, 18, 19].

ACKNOWLEDGMENTS

This work was supported by the Belarussian Foundation for Basic Research (grant no. F99-125).

REFERENCES

1. L. D. Hutcheson, *Integrated Optical Circuits and Components* (Marcel Dekker, New York, 1987).
2. A. W. Snyder and J. D. Love, *Optical Waveguide Theory* (Chapman and Hall, London, 1983; Radio i Svyaz', Moscow, 1987).
3. A. I. Kleev, A. B. Manenkov, and A. G. Rozhnev, Radiotekh. Élektron. (Moscow) **38**, 769 (1993).
4. K. S. Chiang, Opt. Quantum Electron. **26** (3), S113 (1994).
5. V. V. Shevchenko, Izv. Vyssh. Uchebn. Zaved., Radiofiz. **15**, 257 (1972).
6. A. G. Rozhnev and A. B. Manenkov, Radiotekh. Élektron. (Moscow) **42**, 785 (1997).
7. A. B. Manenkov and A. G. Rozhnev, Opt. Quantum Electron. **30** (2), 61 (1998).
8. E. C. Titchmarsh, *Introduction to the Theory of Fourier Integrals* (Clarendon, Oxford, 1948; Gostekhizdat, Moscow, 1948).
9. A. B. Sotskii and L. I. Sorskaya, Opt. Quantum Electron. **31**, 733 (1999).
10. V. I. Krylov, V. V. Bobkov, and P. I. Monastyrnyĭ, *Computational Methods* (Nauka, Moscow, 1977), Vol. 2.
11. A. A. Romanenko and A. B. Sotskii, Zh. Tekh. Fiz. **68** (4), 88 (1998) [Tech. Phys. **43**, 427 (1998)].
12. N. N. Voĭtovich, B. Z. Katsenelenbaum, A. N. Sivov, and A. D. Shatrov, Radiotekh. Élektron. (Moscow) **24**, 1245 (1979).
13. A. B. Sotskii, I. U. Primak, and L. I. Sotskaya, Zh. Prikl. Spektrosk. **67**, 382 (2000).
14. L. I. Sotskaya and A. B. Sotskii, Opt. Spektrosk. **72**, 513 (1992) [Opt. Spectrosc. **72**, 276 (1992)].
15. W. Freude and A. Sharma, J. Lightwave Technol. **3**, 628 (1985).
16. S. J. Hosain, J. P. Meunier, E. Bourillot, *et al.*, Fiber Integr. Opt. **14**, 89 (1995).
17. A. B. Sotskii and L. I. Sotskaya, Opt. Spektrosk. **92**, 141 (2002) [Opt. Spectrosc. **92**, 112 (2002)].
18. L. I. Sotskaya, Zh. Tekh. Fiz. **70** (9), 89 (2000) [Tech. Phys. **45**, 1182 (2000)].
19. A. B. Sotskii and L. I. Sotskaya, Proc. SPIE **4358**, 212 (2001).

Translated by V. Isaakyan

Technical Physics, Vol. 47, No. 12, 2002, pp. 1495–1500. Translated from Zhurnal Tekhnicheskoĭ Fiziki, Vol. 72, No. 12, 2002, pp. 9–14.

ATOMS, SPECTRA, RADIATION

Pressure-Shock-Controlled Pulsed Molecular Beams

G. N. Makarov

Institute of Spectroscopy, Russian Academy of Sciences, Troitsk, Moscow oblast, 142190 Russia
e-mail: g.makarov@isan.troitsk.ru

Received May 6, 2002

Abstract—A method of controlling the duration of pulses of intense molecular beams is suggested. The idea of the method is the shortening of an initial molecular beam pulse by producing a pressure shock in front of a solid surface through which the beam passes. Experiments on shortening H_2, He, SF_6, $SF_6/H_2(1/10)$, and SF_6/He(1/10) molecular beam pulses are reported. The parameters of the beams incident on, and transmitted through, the surface are studied. The gas density in the initial beam and in the pressure shock before the surface is estimated. The intensity and duration of shortened molecular pulses are found as a function of the initial intensity, angle of incidence, and the diameter of a hole on the surface through which the beam passes. It is established that the duration of the shortened beam decreases greatly with increasing incident intensity and decreasing hole diameter. It is shown that intense pulsed H_2, He, SF_6, $SF_6/H_2(1/10)$, and SF_6/He(1/10) molecular beams with a pulse duration of ≤10–15 μs and an extent of ≤1–2 cm can be generated with the method suggested. © *2002 MAIK "Nauka/Interperiodica".*

INTRODUCTION

Intense ($\geq 10^{20}$ molecules/(sr s)) pulsed molecular beams are used in theoretical and applied investigations, such as the study of chemical reactions, elastic and inelastic collisions, molecule–surface interaction, etc. [1]. In recent years, they have also been widely employed in laser spectroscopy and photochemistry [2, 3], as well as in the laser separation of isotopes [4, 5]. The most commonly used way of generating intense molecular beams is their extraction by means of skimmers from dynamically cooled jets produced by pulsed nozzles [6]. The basic parameters of pulsed beams are beam intensity and velocity, pulse duration, and the spread of molecule velocities in a beam (the extent of gas cooling).

There are several types of nozzles producing pulsed molecular beams [6]. Among those employed most extensively are (1) solenoid-controlled nozzles (an automobile fuel injector is an example), (2) piezoelectrically controlled nozzles, and (3) nozzles of the current loop type [6, 7]. Nozzles of the first type generate pulses with a duration from several hundreds of microseconds to several milliseconds; those of the second type generate pulses of duration from 100 μs to 10 ms. Nozzles of the current loop type usually generate pulses of duration from 30–40 to 150 μs. It was reported, however, that molecular beams with pulses as short as ≅7 μs can be produced by these nozzles [8].

The generation of short (≤20 μs) pulses requires a sophisticated nozzle design and a high energy (≥20 J) to provide rapid nozzle opening [6]. At the high energies necessary to actuate the nozzle, it cannot operate with a high pulse repetition rate without cooling. Moreover, the service life of such nozzles is greatly cut compared with those operating under normal conditions. Yet short (in space and time) molecular beams are necessary for many experiments, especially for those where beam molecules are excited and then dissociate under the action of intense laser beams [9, 10]. In [11], we outlined a simple method of shortening intense pulsed molecular beams down to ≤10 μs. In this way, we give a detailed description of this method and discuss at length the shortening of intense pulsed molecular beams by producing a pressure shock before the solid surface.

EXPERIMENTAL

In the method suggested, a pressure shock [12–14] is produced by the interaction of an initial molecular beam with a solid surface. The pressure shock serves as a gas-dynamic shutter. The essence of the method is clarified in Fig. 1. An intense ($\geq 10^{20}$ molecules/(sr s)) wide-aperture (divergence $\omega \cong 0.05$ sr) pulsed molecular beam is incident on a solid surface (a thin (≅100 μm) metallic plate) with a hole of diameter 2–3 mm at its center. Upon the interaction of the intense supersonic molecular beam with the plate, a pressure shock with a characteristic size on the order of the mean molecule free path [12, 13] forms before the plate [15, 16].

When the beam intensity is low (the molecule initial concentration in the beam $N_{in} \leq 3 \times 10^{14}$ cm^{-3}), the pressure shock is absent [17, 18]. In this case, the incident beam, propagating inside a solid angle that depends on the area of the hole and the nozzle-to-surface spacing, passes completely through the hole. Molecules reflected from the surface decrease its intensity only slightly. If, however, the incident beam (pulse) intensity

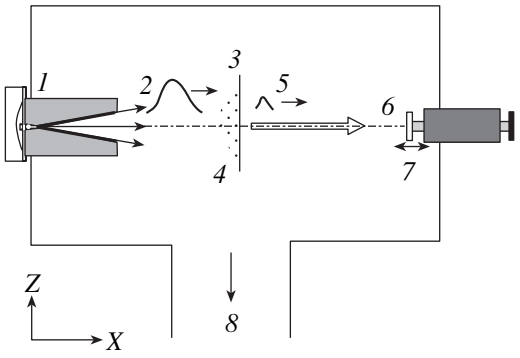

Fig. 1. Experimental scheme: (*1*) pulsed nozzle; (*2*) incident beam; (*3*) surface; (*4*) pressure shock; (*5*) transmitted beam; (*6*) pyroelectric detector; (*7*) travel; and (*8*) to turbomolecular pump.

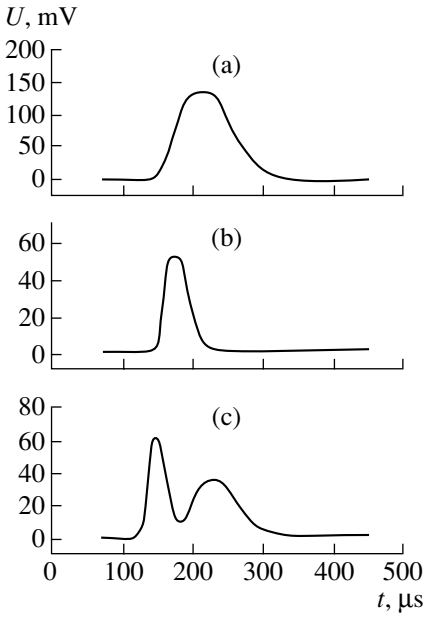

Fig. 2. Time evolution of the molecular beam pulse (SF_6/He = 1/10) passing through a hole of diameter h_0 = 2 mm in a plate for various pressures over the nozzle. The nozzle-to-surface and nozzle-to-detector spacings are 68 and 143 mm, respectively. The gas pressure over the nozzle is (a) 0.15, (b) 0.6, and (c) 1.5 atm. *U* is the signal from the pyroelectric detector.

is sufficiently high ($I_{in} \geq 10^{20}$ molecules/(sr s) or $I_{in} \geq 10^{15}$ cm^{-3}), only the front low-intensity part of the beam passes through the hole, while the remaining high-intensity part of the pulse is completely scattered ("absorbed") by a pressure shock forming before the surface. As a result, the initial molecular pulse shortens significantly. Here, one can draw the analogy to shortening an optical, for example, laser pulse by initiating optical breakdown at the focus of a lens or a telescope, where the tail of the pulse is completely absorbed by the resulting plasma.

In our experiments, we used a nozzle of the current loop type [7]. The hole diameter was 0.75 mm. The time it took for the nozzle to be half open varied within 50–100 μs depending on the gas pressure and composition over the nozzle. The gas pressure over the nozzle was varied from ≅0.1 to 7 atm. The cross section of the nozzle was a cone with an apex angle of 15° and a height of 35 mm. The vacuum chamber where the molecular beam was generated was evacuated to a pressure of $\cong 1 \times 10^{-6}$ torr with a turbomolecular pump. The number of molecules effusing from the nozzle per pulse was a function of the gas pressure and composition over the nozzle. In our experiments with SF_6, this value varied from $\cong 3 \times 10^{15}$ to $\cong 1 \times 10^{17}$ molecules/pulse. The method for measuring the number of molecules in a pulse is described in detail elsewhere [17, 18].

In the experiments, we studied the duration and intensity of shortened molecular pulses as functions of the incident beam intensity and hole diameter, as well as measured the molecule mean velocity and the molecule velocity spread in a shortened pulse. The measurements were made with the time-of-flight method that uses a pyroelectric detector with a time resolution of 3–5 μs as a molecular beam detector [19, 20]. The pyroelectric detector design is described in [21]. The time-of-flight method spectra of the molecules were taken at various distances from the nozzle. From these spectra, we determined the beam velocity and molecule velocity spread in the beams.

RESULTS AND DISCUSSION

1. Time evolution of shortened beam pulses. The experiments were carried out with SF_6, H_2, He, SF_6/H_2(1/10), SF_6/He(1/10), and SF_6/CH_4(1/10) molecular beams. In all the cases, a substantial shortening (2 to 7 times) of the initial molecular pulse was observed. Figure 2 shows the time evolution of the SF_6/He = 1/10 molecular beam (the time-of-flight method molecular spectrum) passing through the hole in the plate as the pressure over the nozzle is increased. The nozzle-to-surface spacing is 68 mm; the nozzle-to-detector spacing, 143 mm. When the pressure over the nozzle is low ($p \leq 0.15$ atm), i.e., when the incident beam intensity is low and a pressure shock is absent [17, 18], the molecular beam passes through the hole without shortening (Fig. 2a). As the pressure grows to the point where the pressure shock before the surface arises, the incident beam pulse shortens (Fig. 2b). As the incident beam intensity (pressure over the nozzle) grows further, the pulse becomes still shorter (Fig. 2c) and a secondary pulsed beam due to the pressure shock [22–24] appears.

As the nozzle–detector spacing increases, the signal induced by the secondary beam was found to decay much faster than the signal induced by the primary beam (Fig. 3). This is mainly attributable to two reasons. The first is the quadratic (for heavily cooled beams)

dependence of the signal on the detector–beam source spacing: the source of the secondary beam is closer to the detector; hence, its signal decays faster as the spacing increases. Second, the gas in the secondary beam is cooled to a much lesser extent than in the shortened primary beam. Therefore, the signal of the secondary beam decays faster also because this beam has a larger spread in the molecule velocities. At distances $x \geq 173$ mm from the nozzle (≥105 mm from the plate), the signal from the secondary beam is more than one order of magnitude weaker than that due to the primary beam.

To suppress the formation of the secondary beam, the plate was rotated so that it made an angle $\alpha = 60°–70°$ with the incident beam. In this case, the gas pressure and density in the oblique pressure shock generated before the plate were much smaller (the density was roughly $1/\cos^2\alpha$ times smaller) than in the normal pressure shock [12, 14]. As a result, the intensity of the secondary beam turned out to be negligibly small compared with that of the shortened primary beam. Figure 4 shows the dependence of the intensities of the shortened primary and secondary molecular beams (H_2) on the angle the incident beam makes with the surface. The secondary beam intensity is seen to drop sharply with increasing angle of incidence. This is to a great extent associated with the decrease in the gas pressure and density in the pressure shock. Note that at large angles of incidence ($\alpha \geq 75°$), when the condition $1/\cos^2\alpha \cong M_{in}$ (M_{in} is the Mach number for the initial beam) is met, a pressure shock at the surface does not form [12–14] and the incident beam does not shorten.

2. Duration of the shortened pulse vs. the gas pressure over nozzle and hole diameter. It was found that the higher the incident beam intensity and the smaller the diameter of the hole, the shorter the beam. For an incident beam intensity $I_{in} \geq 10^{21}$–10^{22} molecules/(sr s) ($N_{in} \geq 10^{16}$ cm^{-3}), the duration of a molecular beam pulse passing through a hole of diameter 2 mm was found to be ≤10–15 μs.

Figure 5 demonstrates the dependence of the duration of the H_2 molecular beam on the gas pressure over the nozzle. The incident beam of duration ≅78 μs strikes the plate at an angle $\alpha \cong 65°$. As follows from Fig. 5, at pressures $p \leq 0.15$ atm, the pulse does not shorten, while at pressures between 4 and 5 atm, the duration of the transmitted pulse is ≤10–12 μs.

The dependence of the duration of the H_2 molecular pulse passing through a hole of diameter 2 mm on the angle of incidence is presented in Fig. 6 (curve *1*). As the angle increases, the effective diameter of the hole is seen to decrease (in proportion to cosα, curve *2*). Therefore, this dependence can also be viewed as the dependence of the pulse duration on the hole diameter. The pulse passing through the hole shrinks with decreasing diameter. This is because the time taken to form a pressure shock decreases with decreasing area of the hole through which the molecules pass. At $\alpha \geq$ 60°–70°, the pulse duration depends on the angle of incidence only slightly, presumably because of the rapid decay of the pressure shock at the surface.

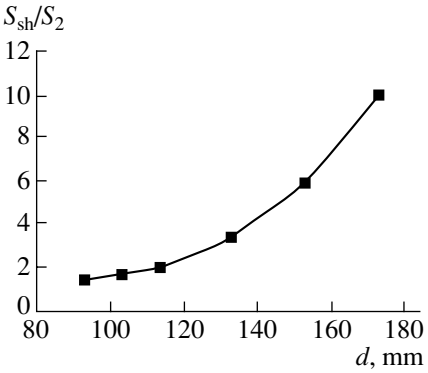

Fig. 3. Amplitude ratio of the signals induced in the detector by the shortened initial and secondary H_2 molecular beams vs. nozzle-to-detector spacing. The conditions are the same as in Fig. 2.

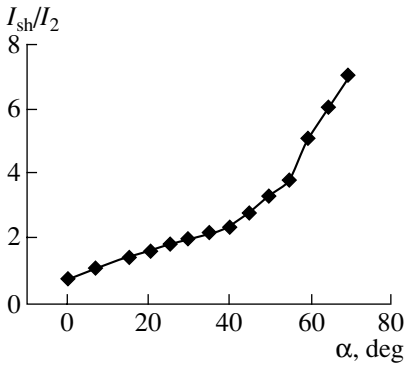

Fig. 4. Intensity ratio of the shortened initial and secondary H_2 molecular beams vs. angle on incidence. $P_{H_2} = 2.4$ atm. The conditions are the same as in Fig. 2.

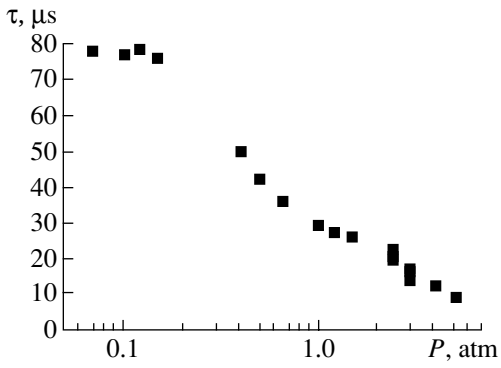

Fig. 5. Duration of the H_2 molecular beam pulse vs. gas pressure over the nozzle. The angle of incidence on the plate is $\alpha \cong 65°$. The conditions are the same as in Fig. 2.

In [11], measured data for the molecular beam intensity, pulse duration, and molecule mean velocity, as well as data for velocity spreads in the beams, were

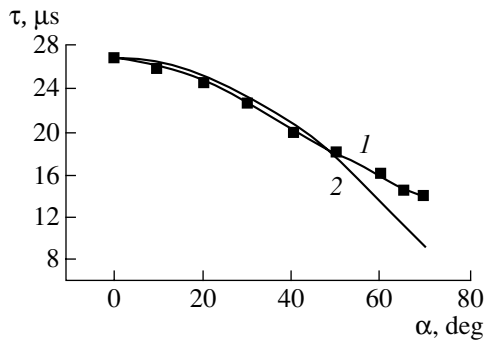

Fig. 6. Duration of the H_2 molecular beam pulse vs. angle of incidence on the plate. The conditions are the same as in Fig. 2.

reported for the case of an SF_6/He = 1/10 molecular beam passing through holes of different diameters. The nozzle–surface spacing was 68 mm. It was established that the pulse duration decreases substantially with decreasing hole diameter. Simultaneously, the beam mean velocity and the molecule velocity spread in the beam increase. Our results, including those for the other molecular beams (see table), are in full agreement with the results of early experiments. The mean velocity grows because more and more energetic molecules pass through the hole as the pulse shortens. The increase in the velocity spread is associated with the enhanced scattering of the beam by molecules reflected from the surface and hole walls. In the shortened pulse, the velocity spread was 1.5–2.0 times greater than in the incident one. At the same time, it also follows from the table that the gas in the shortened pulse was rather cold (the Mach number was $M \cong v/\Delta v = 5$–8). The durations of the SF_6 and SF_6/He(1/10) molecular beams passing through holes of diameter 1.7 and 2.0 mm, respectively, were ≤8 μs immediately behind the surface.

3. Gas density in the incident beam and pressure shock. Note first of all that the gas density in a pressure shock depends on the ratio $\gamma = c_p/c_v$, where c_p and c_v are the specific heats at constant pressure and volume, respectively. For a gas with a constant specific heat, the ultimate increase in the density in a normal pressure shock is defined by the ratio $\rho_{sh}/\rho_{in} = (\gamma + 1)/(\gamma - 1)$, where ρ_{in} and ρ_{sh} are the gas densities in the incident beam and in the pressure shock, respectively. Therefore, in the case of molecules with a low γ, a rather dense pressure shock arises at the surface [16, 17] (for example, in the case of SF_6, $\gamma \cong 1.1$ at $T \cong 300$ K [25] and $\rho_{sh}/\rho_{in} \approx 21$), which significantly shortens the initial pulse but at the same time generates an intense secondary beam. For a higher value of γ (for He and H_2, $\gamma \cong 1.66$ and 1.4, respectively [25]), the ultimate increase in the gas density in a pressure shock is much smaller ($\rho_{sh}/\rho_{in} \approx 4$ and 6, respectively). A pressure shock at the surface is of moderate density in this case. Hence, the secondary beam intensity is much lower than in the case of SF_6.

The ratio N_{sh}/N_{in}, where N_{sh} and N_{in} are the molecule concentration in the pressure shock and in the initial (primary) beam, respectively, can roughly be taken equal to the ratio $\Delta x_b/\Delta x_{sh}$, where Δx_b is the spatial extent of the incident beam and Δx_{sh} is the shock front width; that is, $N_{sh}/N_{in} \cong \Delta x_b/\Delta x_{sh}$. For SF_6, $\Delta x_b \cong 4.6$ cm and $\Delta x_{sh} \cong 3$ mm [17, 18]. Therefore, $N_{sh} \cong 15N_{in}$. For a SF_6 pressure over the nozzle of 1 atm, the number of effusing molecules per pulse is $N_b \cong 3.5 \times 10^{16}$ molecules/pulse

Measured parameters of the initial and shortened molecular beams

Gas composition	Pressure over nozzle, atm	Nozzle-to-detector spacing, mm	Initial beam				Hole diameter, mm	Shortened beam				
			pulse duration, μs	v, m/s	Δv, m/s	$v/\Delta v$		pulse duration, μs	v, m/s	Δv, m/s	$v/\Delta v$	relative intensity I_{sh}/I_{in}
H_2	2.0	143	91	2140	345	6.2	2	27	2400	480	5	0.06
	2.7	173	95	2330	333	7	1	30	2500	580	4.3	0.10
SF_6/CH_4 (1/10)	3.1	173	82	855	111	7.5	2.2	47	917	170	5.4	0.07
SF_6	5.8	113	112	480	51	9.5	2	34	545	78	7	0.11
							1.5	25	585	90	6.5	0.08
SF_6/He (1/10)	0.8	113	92	860	98	8.8	1.7	43	1100	144	7.7	0.11
	2.8	143	78	980	105	9.3	2	23	1095	240	4.6	0.09

[24]. The primary beam volume was estimated at $V_b \cong 14$ cm^3 [17, 24]. Hence, $N_{in} \cong 2.5 \times 10^{15}$ cm^{-3} and $N_{sh} \cong 3.8 \times 10^{16}$ cm^{-3}.

4. Beam scattering in the pressure shock. Let us estimate the beam attenuation due to scattering in the pressure shock by an example of the SF_6 molecular beam, whose parameters have been studied in detail [17, 18]. We will use the following expression for the molecular beam intensity passing through the pressure shock:

$$I_{sh}(t) = I_{in}(t)\exp[-\sigma N_{sh}(t)\Delta x_{sh}(t)]. \quad (1)$$

Here, $I_{in}(t)$ is the incident beam intensity, σ is the interaction cross section for SF_6 molecules, $N_{sh}(t)$ is the molecule concentration in the pressure shock, and $\Delta x_{sh}(t)$ is the extent (width) of the pressure shock at the surface. Note at once that relationship (1) is valid when the gas pressure is low and the molecules experience single collisions. In our case, this is true only at the initial stage of shock formation. Nevertheless, this relationship can apparently be applied as a rough estimator of the beam intensity at which the beam is nearly completely scattered in the shock. As was shown above, at an SF_6 pressure over the nozzle of 1 atm, the molecule concentration in the shock is $N_{sh} \cong 3.8 \times 10^{16}$ cm^{-3} and $\Delta x_{sh} \cong 3$ mm [17, 18]. If it is assumed that the interaction cross section for SF_6 equals the gas-kinetic cross section of this molecule ($\sigma \cong 2.4 \times 10^{-15}$ cm^2 [26]), the exponent in relationship (1) will be $(N_{sh}\sigma\Delta x_{sh}) \cong 28$. Consequently, an SF_6 molecular beam of such an intensity will be completely scattered in the pressure shock. Only the frontmost part of the pulse, in which the intensity is more than one order of magnitude lower than that mentioned above and the exponent is still within several unities, can pass through the beam-generated pressure pulse. The passage of precisely this portion of the pulse was observed in our experiments. Note that the intensity of the initial SF_6 molecular beam in our experiments was $I_{in} \geq 7 \times 10^{21}$ molecules/(sr s). The intensities of the H_2 and He beams were roughly one order of magnitude higher for the same pressures over the nozzle as for SF_6. Therefore, although the shortened beams obtained in the experiment were one or even two orders of magnitude less intensive than the initial ones, their intensities still remained relatively high.

CONCLUSION

We showed that the duration of a molecular beam passing through a solid surface can be controlled by producing a pressure shock before the surface. With this approach, molecular beams with a pulse duration of ≤10–15 μs and a spatial extent of ≤1–2 cm were obtained. If the pulses are sufficiently short, one succeeds in exciting (dissociating) almost all the beam molecules by a laser beam, thereby significantly improving the efficiency of the laser action on the molecular beam [9, 10].

For molecules with a low parameter γ (e.g., for SF_6), the detrimental side effect of the shortening method is the generation of a rather intense secondary beam from the forming shock pulse. To suppress the secondary beam, it is reasonable either to use low-intensity incident beams (when the pressure over the nozzle is low) or to expand the nozzle–surface spacing (when the gas density in the pressure shock is low [16, 17]). If the molecules have a high value of γ (H_2, He, as well as SF_6 + H_2 and SF_6 + He mixtures), the secondary beam is absent, as was shown experimentally. In this case, molecular beams with a pulse duration of ≤10 μs are easy to generate with our approach.

Note that in [27] the pulses of nitrogen molecular beams were shortened from 1.5 ms to 150–300 μs by scattering the beams in a background gas produced by the front portion of the pulse in the vacuum chamber. The molecular beams were generated with a solenoid-driven pulsed nozzle (automobile fuel injector).

ACKNOWLEDGMENTS

The author thanks V.V. Nesterov and A.N. Petin for their technical assistance.

This work was partially supported by the Russian Foundation for Basic Research (grant no. 00-03-33003) and by the CRDF (grant no. RC1-2206).

REFERENCES

1. *Atomic and Molecular Beam Methods*, Ed. by G. Scoles (Oxford Univ. Press, New York, 1988).
2. V. N. Bagratashvili, S. I. Ionov, and G. N. Makarov, in *Laser Spectroscopy of Vibrationally Excited Molecules*, Ed. by V. S. Letokhov (Nauka, Moscow, 1990).
3. V. N. Bagratashvili, V. S. Letokhov, A. A. Makarov, and E. A. Ryabov, *Multiple Photon Infrared Laser Photophysics and Photochemistry* (Academic, New York, 1985).
4. G. N. Makarov, V. N. Lokhman, D. E. Malinovskiĭ, and D. D. Ogurok, Kvantovaya Élektron. (Moscow) **25**, 545 (1998).
5. V. M. Apatin, V. N. Lokhman, G. N. Makarov, *et al.*, Opt. Spektrosk. **91**, 910 (2001) [Opt. Spectrosc. **91**, 852 (2001)].
6. W. R. Gentry, in *Atomic and Molecular Beam Methods*, Ed. by G. Scoles (Oxford Univ. Press, New York, 1988).
7. W. R. Gentry and C. F. Giese, Rev. Sci. Instrum. **49**, 595 (1978).
8. W. R. Gentry, Comments At. Mol. Phys. **9**, 113 (1980).
9. G. N. Makarov, D. E. Malinovsky, and D. D. Ogurok, Laser Chem. **17**, 205 (1998).
10. G. N. Makarov, D. E. Malinovskiĭ, and D. D. Ogurok, Zh. Tekh. Fiz. **69** (1), 35 (1999) [Tech. Phys. **44**, 31 (1999)].
11. G. N. Makarov, Pis'ma Zh. Éksp. Teor. Fiz. **75**, 159 (2002) [JETP Lett. **75**, 131 (2002)].

12. Ya. B. Zel'dovich and Yu. P. Raizer, *Physics of Shock Waves and High-Temperature Hydrodynamic Phenomena* (Nauka, Moscow, 1966; Academic, New York, 1967).
13. L. D. Landau and E. M. Lifshitz, *Course of Theoretical Physics*, Vol. 6: *Fluid Mechanics* (Nauka, Moscow, 1986; Pergamon, New York, 1987).
14. G. N. Abramovich, *Applied Gas Dynamics* (Nauka, Moscow, 1991), Part 1.
15. G. N. Makarov and A. N. Petin, Pis'ma Zh. Éksp. Teor. Fiz. **71**, 583 (2000) [JETP Lett. **71**, 399 (2000)].
16. G. N. Makarov and A. N. Petin, Chem. Phys. Lett. **323**, 345 (2000).
17. G. N. Makarov and A. N. Petin, Zh. Éksp. Teor. Fiz. **119**, 5 (2001) [JETP **92**, 1 (2001)].
18. G. N. Makarov and A. N. Petin, Chem. Phys. **266**, 125 (2001).
19. V. M. Apatin, L. M. Dorozhkin, G. N. Makarov, and L. M. Pleshkov, Appl. Phys. B **29**, 273 (1982).
20. V. M. Apatin and G. N. Makarov, Zh. Éksp. Teor. Fiz. **84**, 15 (1983) [Sov. Phys. JETP **57**, 8 (1983)].
21. R. V. Ambartzumian, L. M. Dorozhkin, G. N. Makarov, *et al*., Appl. Phys. **22**, 409 (1980).
22. V. M. Apatin, G. N. Makarov, and V. V. Nesterov, Pis'ma Zh. Éksp. Teor. Fiz. **73**, 735 (2001) [JETP Lett. **73**, 651 (2001)].
23. V. M. Apatin, G. N. Makarov, and V. V. Nesterov, Chem. Phys. Lett. **347**, 101 (2001).
24. G. N. Makarov, Zh. Éksp. Teor. Fiz. **120**, 1411 (2001) [JETP **93**, 1222 (2001)].
25. *CRC Handbook of Chemistry and Physics*, Ed. by D. R. Lide, 74th Edition (CRC Press, Boca Raton, 1993–1994).
26. I. Burak, J. I. Steinfeld, and D. G. Sutton, J. Quant. Spectrosc. Radiat. Trans. **9**, 959 (1969).
27. D. Bassi, S. Iannotta, and S. Niccolini, Rev. Sci. Instrum. **52**, 8 (1981).

Translated by V. Isaakyan

Technical Physics, Vol. 47, No. 12, 2002, pp. 1501–1508. Translated from Zhurnal Tekhnicheskoĭ Fiziki, Vol. 72, No. 12, 2002, pp. 15–22.

GASES AND LIQUIDS

The Shape of a Drop in a Constant Electric Field

A. A. Shutov

Karpov Institute of Physical Chemistry (Russian State Scientific Center), Obninsk Branch, Obninsk, Kaluga oblast, 249030 Russia

e-mail: fci@meteo.ru

Received November 30, 2001; in final form, May 6, 2002

Abstract—The variation of the shape of a drop immersed in an immiscible liquid under the action of an electric field is calculated. The charge is transferred both by ohmic current through the interface and by the convective component over the interface. A solution quadratic in the parameter that is the ratio of the electric pressure to the capillary pressure is analyzed. Conditions where the drop transforms into a spheroid that is prolate or oblate along the electric field vector are found. An experimental study of the drop deformation by electric forces is carried out. © *2002 MAIK "Nauka/Interperiodica".*

INTRODUCTION

The basic result of theoretical and experimental investigations into the deformation of a drop exposed to an electric field is that the drop takes the shape of a prolate or oblate spheroid or does not change its shape at all, depending on the relationship between the permittivity, conductivity, and viscosity of the medium inside and outside it. The theoretical description of this phenomenon is carried out by constructing electrohydrodynamic (EHD) models. In models [1–4], the statement of the problem of drop deformation does not use the boundary condition relating the surface charge density to the discontinuity of the electric induction at the interface. This is a serious disadvantage of these models from the electrohydrodynamics standpoint. Instead of this condition, the equality of the ohmic currents at the interface is used, while the effect of the permittivity shows up only through interfacial electric tensions. Nevertheless, such models describe adequately the drop deformation in a number of drop–environment liquid systems. For example, they predict correctly the deformation sign in highly viscous dielectric liquids [1, 2]. As a rule, the drop takes the shape of an oblate spheroid when the environment is of higher conductivity. Otherwise, the drop is extended along the field direction [2]. The objective of this paper is to construct a model that would allow us to elucidate the limits of applicability of the approximations used in [1–4] and to describe the behavior of systems that cannot be analyzed within this scheme. Our model takes into consideration the surface charge kinetics in an ordinary layer, which, in the general case, involves the variation of the charge with time, surface convection, and ohmic currents toward the moving interface. Physically, a charge relationship can be derived from the balance of charge flows with allowance for interface deformation [5]. A more general approach to the derivation of an equation describing the evolution of the surface charge is averaging the equation of conservation of space charge near the interface. In [6], a charge equation for the problem of a fluctuating double electric layer was derived. Here, we consider a steady-state problem for an ordinary layer with regard for the surface convection of charges [7].

STATEMENT OF THE PROBLEM

An equation of charge transfer must meet the definition of surface charge density σ in electrodynamics. The density σ is introduced by averaging the equation $\mathrm{div}\mathbf{D} = \rho_e$ along the direction normal to the interface in such a way [8] that

$$D_{ne} - D_{ni} = \sigma, \quad \sigma = \int_{f-h}^{f+h} \rho_e dN, \tag{1}$$

where $\mathbf{D}$ is the electric induction, ρ_e is the space charge density, dN is a length element in the normal direction, and $N = f$ is the equation of interface. The unit vector of the normal $\mathbf{n}$ to the interface is directed from medium i to medium e, and the lower and upper limits of integration refer to liquids i and e, respectively.

Let us integrate the equation of conservation of space charge in the orthogonal coordinate system q_1, q_2, q_3 where the coordinates q_1 and q_2 form an orthogonal grid on the interface $q_3 = f(q_1, q_2, t)$. Then, taking into account (1), we obtain [7]

$$\frac{\partial}{\partial t}(\sigma\sqrt{g_0}) + \frac{\partial}{\partial q_1}(\sigma V_1\sqrt{g_{11}}) + \frac{\partial}{\partial q_2}(\sigma V_2\sqrt{g_{22}}) + \sqrt{g_0}(\lambda_e E_{ne} - \lambda_i E_{ni}) = 0, \tag{2}$$

where g_{ii} are the metric tensor elements for the surface $q_3 = f(q_1, q_2, t)$; $i = 1, 2$; and $g_0 = g_{11}g_{22}$.

The elements g_{ii} are defined as

$$g_{ii} = \left(\frac{\partial x}{\partial q_i}\right)^2 + \left(\frac{\partial y}{\partial q_i}\right)^2 + \left(\frac{\partial z}{\partial q_i}\right)^2; \quad i = 1, 2,$$

where

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} x \\ y \\ z \end{pmatrix}(q_1, q_2, q_3 = f(q_1, q_2, t)).$$

In our case, the metric coefficients coincide with the definition of the surface metric induced by the space metric [9]. In particular, in the spherical coordinate system r, Θ, ϕ, the coordinates Θ, ϕ form an orthogonal grid on the body of revolution $r = f(\Theta, t)$. The length element on such a surface has the form

$$\begin{aligned} ds^2 &= g_{11}d\Theta^2 + g_{22}d\phi^2 \\ &= (f^2 + f'^2)d\Theta^2 + f^2\sin^2\Theta d\phi^2. \end{aligned} \tag{3}$$

Here, $f' = \partial f/\partial\Theta$.

Consider the steady-state axisymmetric problem of the shape of a drop suspended in a liquid that is immiscible with the drop in a uniform electric field. The origin of the spherical coordinate system is located at the drop center. In the absence of the electric field, the drop is a sphere with a radius r_0. The angle Θ is measured from the field vector E_0. The drop radius depends only on the angle Θ. We assume that the translational motion of the drop is absent and the electric current in the liquids flows in the absence of the space charge.

Let us introduce the dimensionless variables marked by bars for the radius, pressure, field strength, electric potential, and surface charge as follows:

$$\bar r = \frac{r}{r_0}, \quad \bar p = \frac{pr_0}{T}, \quad \bar V = \frac{\mu_e + \mu_i}{\varepsilon_e\varepsilon_0 E_0^2 r_0}V,$$

$$\bar E = \frac{E}{E_0}, \quad \bar\varphi = \frac{\varphi}{E_0 r_0}, \quad \bar\sigma = \frac{\sigma}{\varepsilon_e\varepsilon_0 E_0},$$

where $E = -\nabla\varphi$; $\mu_{e,i}$, $\varepsilon_{e,i}$, and $\lambda_{e,i}$ are the viscosity, permittivity, and conductivity of the liquids; T is the surface tension; and ε_0 is the dielectric constant. The subscripts i and e refer to the liquids inside and outside the drop, respectively.

The dimensionless set of equations has the form

$$\nabla p_{e,i} = \alpha M_{e,i}\Delta\mathbf{V}_{e,i}, \quad \nabla\cdot\mathbf{V}_{e,i} = 0, \quad \Delta\varphi_{e,i} = 0, \tag{4}$$

$$r = f: V_r = V_\Theta\frac{df}{fd\Theta}, \tag{5}$$

$$\begin{aligned} &\alpha\beta\frac{d}{d\Theta}(\sigma V_\Theta f\sin\Theta) \\ &+ (\lambda E_{ne} - E_{ni})f^2\sqrt{1 + \frac{f'^2}{f^2}}\sin\Theta = 0, \end{aligned} \tag{6}$$

where

$$\lambda = \frac{\lambda_e}{\lambda_i}, \quad \alpha = \frac{\varepsilon_e\varepsilon_0 E_0^2 r_0}{T}, \quad \beta = \frac{\varepsilon_e\varepsilon_0 T}{\lambda_i(\mu_i + \mu_e)r_0},$$

and Δ is Laplacian.

Here, the impenetrability equation for a steady-state axisymmetric flow degenerates into condition (5), which defines the interface as a current surface. Relationship (6) is charge equation (2) at the interface $r = f(\Theta)$ with metric coefficients (3). Hereafter, bars over the dimensionless values are omitted. Equations (4)–(6) are supplemented by the following boundary conditions at the interface $r = f(\Theta)$:

$$[V_\Theta] = 0, \quad [V_r] = 0, \quad [\varphi] = 0, \tag{7}$$

$$E_{ne} - \varepsilon E_{ni} = \sigma, \tag{8}$$

$$\begin{aligned} &[p]\mathbf{n} - \alpha[M\boldsymbol{\Pi}] \\ &= -K\mathbf{n} + \alpha(E_{ne}\mathbf{E}_e - \varepsilon E_{ni}\mathbf{E}_i - (E_e^2 - \varepsilon E_i^2)\mathbf{n}/2), \end{aligned} \tag{9}$$

$$r \longrightarrow \infty: \varphi \longrightarrow -r\cos\Theta.$$

Here, $[A] = A_e - A_i$ denotes a discontinuity in the value of A.

The curvature K and the normal and tangential components Π_n and Π_τ of the viscous stress vector $\boldsymbol{\Pi}$ are respectively given by

$$\begin{aligned} K &= \left(\frac{1}{f} + \frac{d^2}{d\Theta^2}\frac{1}{f}\right)\Big/\left(1 + \frac{f'^2}{f^2}\right)^{3/2} \\ &- \frac{1}{\sin\Theta}\frac{d}{d\Theta}\left(\frac{\cos\Theta}{f}\right)\Big/\left(1 + \frac{f'^2}{f^2}\right)^{1/2}, \end{aligned}$$

$$\Pi_n = \left(\tau_{rr} - 2\frac{f'}{f}\tau_{r\Theta} + \frac{f'^2}{f^2}\tau_{\Theta\Theta}\right)\Big/\left(1 + \frac{f'^2}{f^2}\right),$$

$$\Pi_\tau = \left(\tau_{r\Theta}\left(1 - \frac{f'^2}{f^2}\right) + (\tau_{rr} - \tau_{\Theta\Theta})\frac{f'}{f}\right)\Big/\left(1 + \frac{f'^2}{f^2}\right),$$

where τ_{ij} are the components of the viscous stress tensor in the spherical coordinate system. The dimensionless parameters in (4)–(9) are $\varepsilon = \varepsilon_i/\varepsilon_e$, $\mu = \mu_i/\mu_e$, $M_i = \mu/(\mu + 1)$, and $M_e = 1/(\mu + 1)$.

The interaction between the electric and hydrodynamic fields in our self-consistent problem can be represented as follows. The discontinuity of the electric induction is expressed, according to (8), in terms of the surface charge density. The surface charge density σ is found during the solution of the problem, this parameter, according to (6), relating the liquid tangential velocity at the interface to the electric fields. Dynamic boundary condition (9), in its turn, also considers the pressure and surface tension. Equation (5) defines the drop boundary (unknown in advance). In the general case, vortex motion arises inside the drop and in the environment. In some cases, the liquid may be stagnant. At the uncharged interface between perfect dielectrics

or at the free surface of a perfect conductor, tangential electric forces do not arise and the problem allows for a static solution in which the drop shape is determined from the balance between the forces normal to the interface [10]. We seek a solution to set (4)–(9) in the form of the expansion in powers of the parameter α:

$$f(\Theta) = 1 + \alpha f_1(\Theta) + \alpha^2 f_2(\Theta) + \ldots,$$

$$\sigma(\Theta) = \sigma_0(\Theta) + \alpha\sigma_1(\Theta) + \alpha^2\sigma_2(\Theta) + \ldots,$$

$$\Phi(r, \Theta) = \Phi^{(0)}(r, \Theta) + \alpha\Phi^{(1)}(r, \Theta) + \alpha^2\Phi^{(2)}(r, \Theta) + \ldots,$$

where Φ stands for p, V, or φ.

Expanding the boundary conditions in the parameter α, which also enters into f, and collecting terms with the same powers of α, we get problems formulated on the surface $f = 1$ in each of the approximations. General problem (4)–(9) will be solved up to the second approximation. The zeroth approximation formally corresponds to the absence of the electric field ($E_0 = 0$). In this approximation, the drop has the constant radius $f = 1$ and the difference between the pressure inside it and the ambient pressure is equal to the capillary pressure $p_i^{(0)} = 2$. Since the next approximations are formulated on the surface $f = 1$, we can use general solutions for a sphere. A solution to the hydrodynamic problem in the Stokes approximation has the form [11]

$$p_i = \text{const} + M_i\sum_n (2n+2)(2n+3)a_n\omega_n,$$

$$p_e = \text{cosnt} + M_e\sum_n 2n(n-1)b_n\pi_n,$$

$$\begin{aligned}\mathbf{V}_i = &\sum_n a_n((n+3)r^2\nabla\omega_n - 2nr\omega_n)\\ &+ \sum_n (n+1)c_n\nabla\omega_n,\end{aligned} \tag{10}$$

$$\mathbf{V}_e = \sum_n b_n((2-n)r^2\nabla\pi_n + 2(n+1)r\pi_n) - \sum_n nd_n\nabla\pi_n,$$

$$\pi_n = r^{-n-1}P_n(\eta), \quad \omega_n = r^nP_n(\eta),$$

where $P_n(\eta)$ is the Legendre polynomial and $\eta = \cos\Theta$.

Similarly, the general solution to the Laplace equation for the potential has the form

$$\varphi_e = -r\cos\Theta + \sum_n B_nP_n/r^{n+1}, \quad \varphi_i = \sum_n C_nP_nr^n. \tag{11}$$

Here, a_n, b_n, c_n, d_n, B_n, and C_n are constants.

RESULTS OF CALCULATION

Set (4)–(9) in the first approximation has the form

$$\nabla p_{i,e}^{(1)} = M_{i,e}\Delta\mathbf{V}_{i,e}^{(0)}, \quad \text{div}\mathbf{V}_{i,e}^{(0)} = 0, \quad \Delta\varphi_{i,e}^{(0)} = 0$$

with the boundary conditions on the surface $f = 1$

$$V_{\Theta i}^{(0)} = V_{\Theta e}^{(0)}, \quad V_{ri}^{(0)} = V_{re}^{(0)} = 0, \tag{12}$$

$$\lambda E_{re}^{(0)} - E_{ri}^{(0)} = 0, \quad E_{re} - \varepsilon E_{ri} = \sigma_0, \tag{13}$$

$$\begin{aligned}&\varphi_i^{(0)} = \varphi_e^{(0)}, \quad r \longrightarrow \infty\colon \varphi_e^{(0)} \longrightarrow -r\cos\Theta,\\ &\left[M\left(\frac{1}{r}\frac{\partial V_r^{(0)}}{\partial\Theta} + \frac{\partial V_\Theta^{(0)}}{\partial r} - \frac{V_\Theta^{(0)}}{r}\right)\right] = \sigma_0E_\Theta^{(0)},\\ &\left[p^{(1)} - 2M\frac{\partial V_r^{(0)}}{\partial r}\right] = \frac{d}{d\eta}(1-\eta^2)\frac{df_1}{d\eta} + 2f_1\\ &\quad + \frac{E_{re}^{(0)2} - E_{\Theta e}^{(0)2} - \varepsilon(E_{ri}^{(0)2} - E_{\Theta i}^{(0)2})}{2}.\end{aligned} \tag{14}$$

Using representations (10) and (11), we find the following solution to this problem:

$$\begin{aligned}p_i^{(1)} &= 14M_iUr^2P_2(\eta) + \text{const},\\ p_e^{(1)} &= 4M_eUP_2(\eta)/r^3 + \text{const},\\ V_{ri}^{(0)} &= 2U(r^3 - r)P_2(\eta),\\ V_{\Theta i}^{(0)} &= U(3r - 5r^3)\sin\Theta\cos\Theta,\\ V_{re}^{(0)} &= 2U\left(\frac{1}{r^2} - \frac{1}{r^4}\right)P_2(\eta),\\ V_{\Theta e}^{(0)} &= -2U\frac{1}{r^4}\sin\Theta\cos\Theta,\\ \varphi_e^{(0)} &= -\left(r + \frac{\lambda-1}{2\lambda+1}\frac{1}{r^2}\right)\cos\Theta,\\ \varphi_i^{(0)} &= -\frac{3\lambda}{2\lambda+1}r\cos\Theta,\\ U &= \frac{9\lambda(1-\lambda\varepsilon)}{10(2\lambda+1)^2}, \quad \sigma_0 = 3\frac{1-\lambda\varepsilon}{2\lambda+1}\cos\Theta.\end{aligned} \tag{15}$$

It should be noted that boundary condition (13) is not used in solving the problem in the first approximation and serves only for the determination of the density σ_0. Two conditions (13) and (14), which state the fields in classical electrodynamics, are replaced in this approximation by conditions (12) and (14). As is seen from charge equation (6), such a situation takes place at $\alpha\beta \ll 1$. Therefore, the linear approximation is valid only in the limiting cases of highly viscous media or of a high-conductivity drop. The first-order contribution to the drop radius is given by

$$f_1 = FP_2(\eta) = \frac{3}{4(2\lambda+1)^2}\left(1+\lambda^2-2\lambda^2\varepsilon + \frac{3}{5}\lambda(1-\lambda\varepsilon)\frac{3\mu+2}{\mu+1}\right)P_2(\eta). \quad (16)$$

The condition $F > 0$ corresponds to the transformation of the drop into a spheroid extended along the field direction. When $F < 0$, the drop takes the shape of an oblate spheroid [2, 3].

In the next approximation, the separation of terms of the same order with respect to α relates quadratic terms in p and f with linear terms of the other functions. In the second-order approximation, the problem is stated in the same manner; however, now the charge equation includes the convective term unlike the first approximation. A solution to the electric problem has the form

$$\varphi_i^{(1)} = A_1 r\cos\Theta + A_2 r^3 P_3,$$

$$\varphi_e^{(1)} = A_3\frac{\cos\Theta}{r^2} + A_4\frac{P_3}{r^4},$$

$$A_1 = W + \frac{18}{5}\frac{\lambda(\lambda-1)}{(2\lambda+1)^2}F, \quad A_2 = 4W\frac{2\lambda+1}{4\lambda+3},$$

$$W = \frac{12}{5}\beta U\frac{1-\lambda\varepsilon}{(2\lambda+1)^2},$$

$$A_3 = A_1 - \frac{6}{5}\frac{(\lambda-1)(2\lambda+1)}{(2\lambda+1)^2}F,$$

$$A_4 = A_2 - \frac{9}{5}\frac{\lambda-1}{(2\lambda+1)^2}F.$$

Using expansion (10), we determine the pressures and velocities:

$$p_e^{(2)} = \frac{M_e}{r^3}\left(4B_1P_2 + \frac{56}{5}B_2\frac{P_4}{r^2}\right),$$

$$p_i^{(2)} = M_i r^2(14B_5 + 22B_6 r^2 P_4),$$

$$V_{re}^{(1)} = \frac{2P_2}{r^2}\left(B_1 + \frac{B_3}{r^2}\right) + \frac{2P_4}{r^4}\left(2B_2 + \frac{B_4}{r^2}\right),$$

$$V_{\Theta e}^{(1)} = \frac{2\sin\Theta\cos\Theta}{r^4}\left(B_3 + \frac{1}{3}\left(B_2 + \frac{B_4}{r^2}\right)(7P_2-1)\right),$$

$$V_{ri}^{(1)} = 2P_2(B_5 r^3 + B_7 r) + 2P_4(2B_6 r^5 + B_8 r),$$

$$V_{\Theta i}^{(1)} = -\sin\Theta\cos\Theta\left(5B_5 r^3 + 3B_7 r + \frac{7}{3}\left(B_6 r^5 + \frac{5}{4}B_8 r^3\right)(7P_2-1)\right).$$

The coefficients B_n are related to each other by the following relationships:

$$B_1 = 3FU - \frac{6}{35}FUG - \frac{3}{5}FU\frac{14\lambda+1}{2\lambda+1} - \frac{3}{70}W\frac{40\lambda^2(2\varepsilon-1) + 54\varepsilon\lambda - 130\lambda - 69}{(2\lambda+1)(4\lambda+3)},$$

$$B_2 = \frac{23}{21}FU - \frac{10}{21}W\frac{6\varepsilon\lambda + 4\lambda - 3}{4\lambda+3},$$

$$\frac{B_1+B_3}{2} = \frac{B_1+B_5}{5} = -\frac{B_1+B_7}{3} = \frac{2B_2+B_4}{17} = \frac{B_2+B_6}{3} = \frac{2B_2+B_8}{6} = -\frac{1}{7}FU.$$

Taking into account the first and second approximations, we come to the following expression for the latitude dependence of the drop radius:

$$f = 1 + \alpha A_{12}P_2 + \alpha^2(A_{20} + A_{22}P_2 + A_{24}P_4), \quad (17)$$

where

$$A_{22} = \frac{79-94\lambda}{35(2\lambda+1)}F^2 + \frac{1}{4}\left(5 + 3G - \frac{6}{35}G^2\right)FU + \frac{W}{14}\left(\frac{20\varepsilon\lambda^2 + 116\lambda^2 + 24\varepsilon\lambda + 251\lambda + 114}{(2\lambda+1)(4\lambda+3)} + \frac{3(5+G)}{20}\frac{80\varepsilon\lambda^2 - 40\lambda^2 + 54\varepsilon\lambda - 130\lambda - 69}{(2\lambda+1)(4\lambda+3)}\right),$$

$$A_{24} = \frac{26}{35}F^2 + \frac{1}{15}\left(1 - \frac{11}{21}G\right)FU + \frac{W}{21}\left(2\frac{3\varepsilon\lambda - 12\lambda + 16}{4\lambda+3} + \frac{9+G}{3}\frac{6\varepsilon\lambda + 4\lambda - 3}{4\lambda+3}\right),$$

$$A_{20} = -\frac{F^2}{5}, \quad G = M_i - M_e = \frac{\mu-1}{\mu+1}.$$

Here, $A_{12} = F$ and A_{20} is a constant found from the condition of drop volume constancy. The terms in (17) that are proportional to F^2 and FU coincide with data in [4]. The contribution of charge surface convection is allowed for by the parameter W. Figures 1–3 show the dependences of the coefficients A_{kj} in relationship (17) on the relative conductivity λ for various μ, β, and ε. As is seen in Fig. 1, the μ dependence of A_{12} is weak; for A_{22} and A_{24}, the character of this dependence is the same. Comparing the coefficients A_{kj} shows that the contribution of the second approximation may exceed the term linear in α by several orders. Such a behavior of the coefficients is due to the fact that series (17) is virtually a two-parametric expansion in α and β. Here, the procedure of determining the coefficients is performed up to the term linear in β, which is the basic

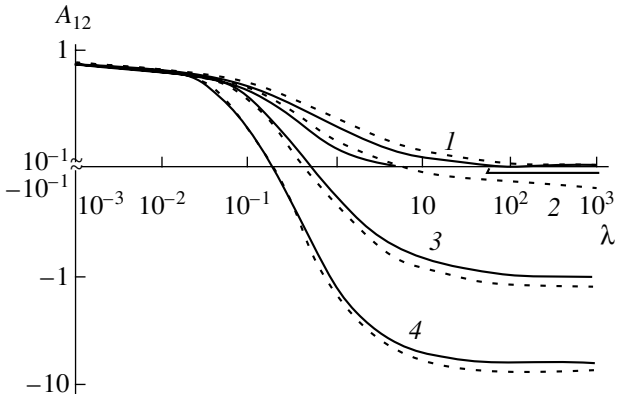

Fig. 1. Coefficient A_{12} as a function of the relative conductivity λ. Solid lines, μ = 0.01; dashed lines, μ = 100. ε = (*1*) 0.1, (*2*) 0.5, (*3*) 2, and (*4*) 10.

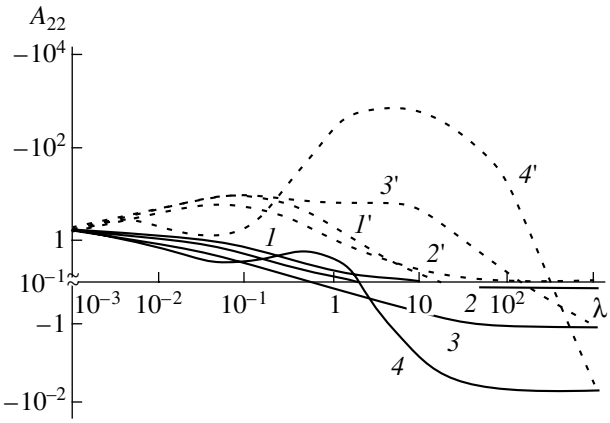

Fig. 2. Coefficient A_{22} as a function of λ at μ = 1. Solid lines, β = 0.01; dashed lines, β = 100. (*1–4*) The same as in Fig. 1.

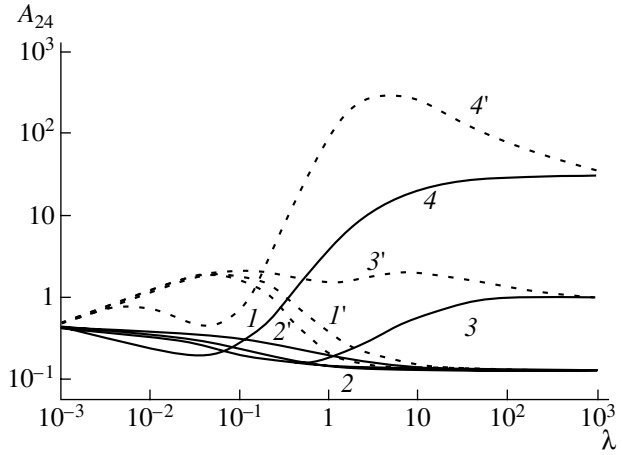

Fig. 3. Coefficient A_{24} as a function of λ at μ = 1. The notation is the same as in Fig. 2.

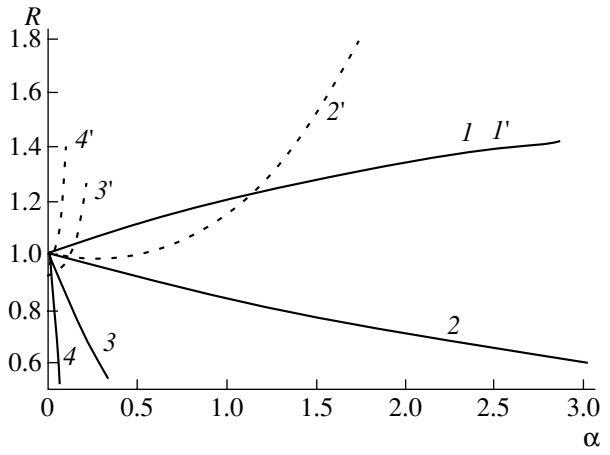

Fig. 4. Variation of the drop deformation with electric field strength. λ = 10 and μ = 10. The notation is the same as in Fig. 2.

term of the expansion at certain values. The drop remains spherical if λε = 1 and $F = 0$; that is, if the conditions ε = 1 and λ = 1 are fulfilled simultaneously.

Let us define the degree of drop deformation as

$$R = \frac{r_1}{r_2}, \tag{18}$$

where $r_1 = f(\Theta = 0)$ and $r_2 = f(\Theta = \pi/2)$ are the lengths of the drop semiaxes in the longitudinal and transverse directions with respect to the field vector, respectively.

The case $R > 1$ corresponds to a prolate spheroid; $R < 1$, to an oblate one.

Figure 4 shows the dependences of the drop deformation R on the field strength for various ε, λ, μ, and β. These curves were constructed in the range $0 \le \alpha \le \alpha_k$, where α_k is the maximum value of α at which the curve $f(\Theta)$ remains convex. By a convex curve, we mean a closed curve that lies in full on one side of a tangent to it at any point. Such a choice of the α range follows from observations that the drop remains stable as long as its shape is ellipsoidal and takes a dumbbell-like shape at the instant of disintegration [2, 12]. The parameter β prevents drop compression. From Fig. 4, as β grows, the drop compression decreases and under certain conditions the drop becomes extended.

Qualitatively, such an effect can be explained by convective charge transfer over the drop surface largely from the equator to the poles. When surface charge convection is absent (β = 0), the deformation is related to the flow inside the drop as follows [2, 3]. At $\lambda\varepsilon \ll 1$, the liquid near the interface, according to (15), flows from the equator to the poles. At the same time, the function F in relationship (16) is positive; that is, the spheroid is extended along the field. When $\lambda\varepsilon \gg 1$, the flow is directed from the poles to the equator, the function $F < 0$, and the spheroid is oblate. If β > 0, then, according to (7), the gradient of the convective flux of the charges coincides with the direction of fastest increase of the ohmic current, which reaches a maximum near the

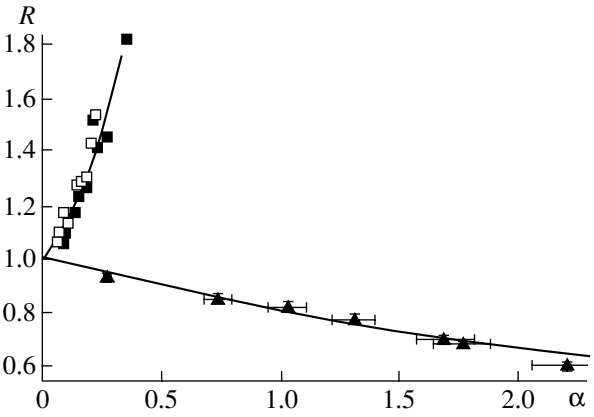

Fig. 5. Same as in Fig. 4. Solid curves are the results of calculation. (■) CO in PMPS, $r_0 = 0.15$ mm, $\varepsilon = 1.9$, $\mu = 1.18$, $\lambda = 6.7 \times 10^{-3}$, and $\beta = 0.36$; (□) water in PMPS, $r_0 = 0.5$–0.8 mm, $\varepsilon = 30$, $\mu = 8.4 \times 10^{-4}$, $\lambda = 10^{-6}$, and $\beta = 10^{-3}$; (▲) PMPS in CO, $r_0 = 0.25$ mm, $\varepsilon = 0.527$, $\mu = 0.85$, $\lambda = 150$, and $\beta = 56.5$.

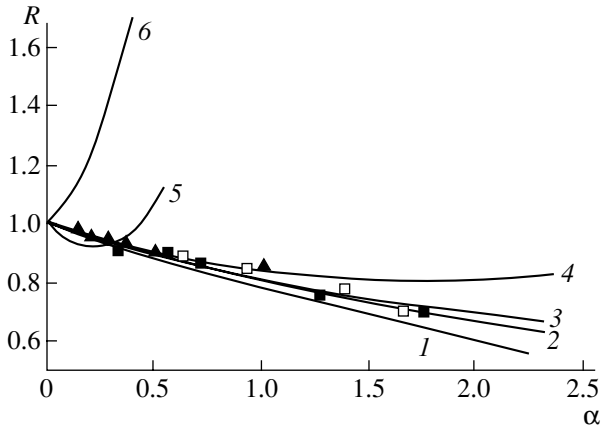

Fig. 6. Same as in Figs. 4 and 5. (*1–4*) PMPS in CO, $\varepsilon = 0.527$, $\mu = 0.85$, and $\lambda = 150$. (*1*) $\beta = 0$; (*2*) ■, $r_0 = 0.3$ mm and $\beta = 47.8$; (*3*) □, $r_0 = 0.2$ mm and $\beta = 71.7$; (*4*) ▲, $r_0 = 0.9$ mm and $\beta = 159$; (*5*) PMPS in CO, $r_0 = 0.05$ mm, $\varepsilon = 1.3$, $\mu = 16.3$, $\lambda = 4.3$, and $\beta = 28$; (*6*) BO in PMPS, $r_0 = 0.05$ mm, $\varepsilon = 0.74$, $\mu = 0.061$, $\lambda = 0.23$, and $\beta = 8.8$.

poles. Thus, one can conclude that surface charge convection causes an additional liquid motion from the equator toward the poles regardless of the drop deformation.

For $0.01 \le \mu \le 100$, $0.02 \le \varepsilon \le 50$, $10^{-3} \le \lambda \le 10^3$, $\beta \le 1000$, and $\alpha \le \alpha_{jk}$, the drop deformations calculated are within $0.5 \le R \le 2$.

EXPERIMENTAL

The experimental study of drop deformation was carried out in a rectangular cell with the base area 4 × 4 cm and an adjustable height of up to 2 cm. The side walls were made of glass, and a direct voltage from 0 to 10 kV was applied to the metallic base. The liquid drop under study was injected into a liquid inside the cell through a glass capillary 0.010–0.005 cm in diameter. A pair of immiscible liquids was chosen from organic and mineral oils, organosilicon polymer polymethylphenylsiloxane (PMPS), and water. In the experiment, we measured the density, viscosity, permittivity, conductivity, and interfacial tension of the liquid pair. The drop image was photographed with the help of a long-focus optical system and a pulsed light source (pulse duration 1 ms). The drop size was measured using photographs magnified 50 to 100 times. The measurement error for $\rho_{i,e}$, $\mu_{i,e}$, $\varepsilon_{i,e}$, and potential (field strength) did not exceed 1%; for the conductivity, 15%. The interfacial tension was measured by the hanging drop method or by the method of capillary rise with an accuracy of ≈5%. The properties of the liquids are presented in Table 1, and the interfacial tensions for several pairs of the liquids are listed in Table 2. Except for the bobbin oil–castor oil pair, the other liquids are practically mutually insoluble. The surface tension in the water–PMPS and water–castor oil systems equaled 38 and 18 dyn/cm, respectively.

Table 1. Properties of the liquids

Liquid	$\rho_{e,i}$, g/cm^3	$\mu_{e,i}$, p	$\varepsilon_{e,i}$	$\lambda_{e,i}$, Ω^{-1} m^{-1}
PMPS	1.1	11.9	2.69	1.43×10^{-12}
CO	0.97	14	5.1	2.1×10^{-10}
BO	0.87	0.73	1.9	6.0×10^{-12}
TO	0.88	0.67	2.36	9.46×10^{-12}
Water	1.0	0.01	81	10^{-6}

Table 2. Interfacial tension, dyn/cm

Liquid	Liquid			
	1	2	3	4
PMPS	×	1.2	0.14	0.16
CO		×	0.6	0.72
BO			×	–
TO				×

Figures 5 and 6 show the dependences of the drop deformation on the field strength for various pairs of liquids (the drop is named first): castor oil (CO) in PMPS, water in PMPS, and PMPS in CO. Symbols in the figures are data points; the curves, results of calculation obtained using relationship (18). According to approximation (16) linear with respect to α, the reversal of the drop–environment combination reverses the extension/flattering deformation when the product $\lambda\varepsilon$ considerably differs from unity, as, for example, for the PMPS–CO combination (Fig. 5, curves *1*, *3*). This turns out to be valid when the parameter β is far from its critical values. Figure 6 demonstrates the dependences of

R on α for PMPS drops in CO with drop initial radii of 0.03, 0.02, and 0.009 cm (curves *2–4*, respectively). It is seen that the compression of the drop decreases as its radius decreases. This effect is more pronounced for the PMPS–bobbin oil (BO) combination with a drop radius of 0.005 cm. Within the experimental error, the PMPS–BO and BO–PMPS combinations demonstrate only extension (Fig. 6, curves *5, 6*). Similar behavior is observed for the PMPS–turbine oil (TO) and BO–CO combinations.

The calculations of the drop shape were carried out under the assumption that the drop deformation is small (formally for $\alpha < 1$). At the same time, a number of theoretical curves plotted to α_k and several data points extend into the range $\alpha > 1$ (Figs. 4–6). Such a situation takes place when $F < 1$, $U < 1$, and $\beta \ll 1$. In this case, the expansion of the drop radius contains the parameter α only in the combinations αF and αU. Therefore, the radius of convergence of the power series for f in α may be greater than unity.

In Fig. 5, the errors of data points for one of the curves are shown. The measurement error for R does not exceed 1%, and the parameter α calculated from the measured values of E_e, r_0, and T has an error of ≈7%. In the moderate deformation range, the agreement between the experimental and theoretical data is satisfactory. However, at $R \approx 1$, as in [2], the experimental deformations systematically exceed the calculated data.

The calculated curves in the figures are plotted up to the critical value of α_k corresponding to the ultimate convexity of the drop shape. The experimental values of α_k coincide with the calculated ones within the experimental error for the pairs of liquids considered. In general, the determination of α_k should be carried out within the problem of stability of the steady-state solution with regard for various mechanisms of drop stability loss [12–17]. The hypothesis for the stability of convex shapes used by us implies the development of a neck on the critical spheroid. In particular, a flattened drop in the critical field contracts most intensely along the line passing through the poles, transforming into a toroidal body. Observations, which are in complete agreement with data reported in [2], indicate that the final shape is not a torus but a bell-shaped figure, which disintegrates because of complicated motions.

If the drop is extended, one can expect its partition into two parts at the neck formed at the spheroid equator. Such a disintegration, accompanied by the emergence of satellite droplets, was observed in our experiments.

Under certain conditions, the disintegration is preceded by the ejection of jets (or individuals droplets) from the poles for an extended drop or of a fan-like jet from the equatorial region for an oblate spheroid. For example, in our calculations, the water drop disintegrates into two parts in air ($\lambda = 0$, $\varepsilon = 81$, $\mu \gg 1$, and $\beta \gg 1$) at $\alpha_k = 0.329$ and $R = 1.74$. However, in a weaker field at $\alpha_k = 0.21$, the ejection of droplets from the poles is observed. This is due to the excitation of higher harmonics of capillary perturbations in regions of the highest field strength [10, 18]. At the same time, in the water–PMPS and water–CO systems, the ejection of droplets is absent: the drop splits into two roughly equal parts. The same was observed in [2] for a water drop in a medium with a viscosity that was one order of magnitude higher.

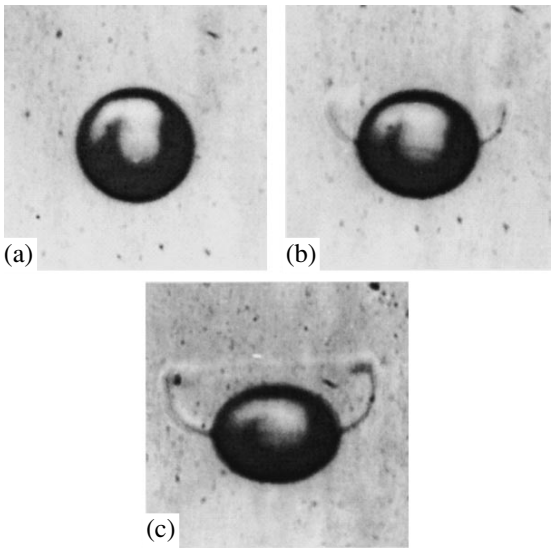

Fig. 7. Shape of a PMPS drop in castor oil. The field vector is parallel to the plane of the photograph and directed vertically. E_0 = (a) 0, (b) 2.5×10^5, and (c) 3.5×10^5 V/m.

Figure 7 displays the photos of drops in the PMPS–CO system. An initially spherical drop falls in lighter castor oil (Fig. 7a). Under the action of the electric field, it transforms into an oblate spheroid. At a certain field strength, a sharp edge appears at the equator from which the drop liquid flows out (Fig. 7b). The stronger the field, the more intense the outflow (Fig. 7c). In the latter case, a structure similar to a ring wake vortex forms. When carrying out the experiments, we did not select pairs of liquids with neutral buoyancy, because convection in the cell intensifies with growing field strength. From the deflection of the fan-like jet from the horizontal in Figs. 7b and 7c, one can conclude that convection has a minor effect in this case in comparison with the effect of the gravitational field. A similar transformation of a drop was observed in [2] in a much more viscous system. However, in [2], in a constant field the liquid outflow did not occur, while in a variable field, the ejection of individual droplets from the equatorial region of the spheroid, rather than a continuous fan-like jet, was observed.

CONCLUSION

The models that do not allow for the motion of the charge along the interface between liquids derive the

dependences of the drop shape only on the conductivity, permittivity, and viscosity of the liquids. Experimental data for highly viscous liquids with a total viscosity of hundreds of poises confirm these results [1, 2]. At the same time, in low-viscous emulsions of weakly conducting liquids, the extension or compression of the drop along the field direction depends considerably on its size. This discrepancy is explained by the fact that these models use a truncated equation for surface charge transfer, which takes into account only the ohmic current through the interface. Convective charge transfer causes additional liquid flows, which act on the drop unilaterally, preventing its compression. The finer the drop, the higher the probability of finding the extension of an emulsion drop in a strong electric field virtually irrespective of other parameters of weakly conducting liquids. The assumption that the spheroid retains the convex shape up to the critical field strength at which the drop disintegrates into equal parts is in good agreement with observations.

ACKNOWLEDGMENTS

The work was supported by the Russian Foundation for Basic Research and by the administration of Kaluga oblast (projects no. 01-01-96002 and 02-01-96035)

REFERENCES

1. G. I. Taylor, Proc. R. Soc. London, Ser. A **291**, 159 (1966).
2. S. Torza, R. G. Cox, and S. G. Mason, Philos. Trans. R. Soc. London, Ser. A **269**, 295 (1971).
3. J. R. Melcher and G. I. Taylor, Annu. Rev. Fluid Mech. **1**, 111 (1969).
4. O. O. Ajaji, Proc. R. Soc. London, Ser. A **344**, 499 (1978).
5. A. J. Mestel, J. Fluid Mech. **274**, 93 (1994).
6. V. V. Gogosov and G. A. Shaposhnikova, Int. J. Appl. Electromagn. Mater. **1**, 45 (1990).
7. A. A. Shutov, N. M. Alontseva, and V. Ya. Shkadov, in *Proceedings of the Regional Competition of Projects in Natural Sciences* (Éĭdos, Kaluga, 2000), Vol. 1, p. 67 (ISBN 5-938810-006-2).
8. V. G. Levich, Yu. A. Vdovin, and V. A. Myamlin, *Course of Theoretical Physics* (Fizmatgiz, Moscow, 1962), Vol. II.
9. B. A. Dubrovin, S. P. Novikov, and A. T. Fomenko, *Modern Geometry* (Nauka, Moscow, 1979).
10. G. I. Taylor, Proc. R. Soc. London, Ser. A **280**, 383 (1964).
11. H. Lamb, *Hydrodynamics* (Cambridge Univ. Press, Cambridge, 1932), Art. 336.
12. S. B. Sample, B. Raghupathy, and C. D. Hendricks, Int. J. Eng. Sci. **8**, 97 (1970).
13. P. R. Brazier-Smith, Phys. Fluids **14**, 1 (1971).
14. A. A. Sabry, Physica A (Amsterdam) **101**, 223 (1980).
15. C. E. Rosenkilde, Proc. R. Soc. London, Ser. A **312**, 473 (1969).
16. Z. C. Feng and L. G. Leal, Int. J. Multiphase Flow **22**, 93 (1996).
17. K. J. Cheng and J. B. Cheddock, Phys. Lett. A **106A**, 51 (1984).
18. A. I. Grigor'ev and S. O. Shiryaeva, Zh. Tekh. Fiz. **57**, 1706 (1987) [Sov. Phys. Tech. Phys. **32**, 1020 (1987)].

Translated by N. Mende

Technical Physics, Vol. 47, No. 12, 2002, pp. 1509–1517. Translated from Zhurnal Tekhnicheskoĭ Fiziki, Vol. 72, No. 12, 2002, pp. 23–31.

GAS DISCHARGES, PLASMA

Study of the Reverse Current Range in a Knudsen Diode with Surface Ionization under Overneutralized Conditions in the Presence of Electron Emission from the Collector: Part I. Large Electrode Spacing

V. I. Babanin*, V. I. Sitnov*, A. V. Solov'ev, and A. Ya. Énder***

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

*** Institute of Chemistry and Applied Engineering Science, Aalborg University of Esbjerg, 6700 Esbjerg, Denmark*

Received April 24, 2002

Abstract—The reverse current range of a Knudsen diode with surface ionization is studied under overneutralized conditions in the presence of electron emission from the collector. The difficulty of the problem is associated with the need to consider three particle flows: ions and electrons from the emitter and electrons from the collector. This work consists of two parts. In the first part, the potential distributions and *I–V* characteristics for plasma regimes, where the potential distribution has a quasi-neutrality portion, are calculated. In the second, the diode with a small electrode spacing, where the near-electrode potential jumps overlap, is investigated. The asymptotic behavior of the current under high negative voltages is examined. Data obtained are compared with the classical Langmuir results for the diode with two opposing electron and ion flows. © *2002 MAIK "Nauka/Interperiodica".*

INTRODUCTION

The study of a Knudsen diode with surface ionization (SIKD) in the presence of electron emission from the collector is of great importance. The electron emission may greatly affect the properties of thermionic converters (TCs) filled with a Cs–Ba compound, where a hot barium-covered collector offers a high emissivity. Furthemore, finding the self-consistent potential distributions in a three-flow SIKD may reveal intriguing physical effects. In [1–4], SIKDs were studied without considering the emission from the collector. The first attempt to take the emission into account was made in [5]. However, the calculations were performed only for some arbitrarily selected conditions.

A basic parameter of an SIKD is the degree of neutralization

$$\gamma = \frac{n_i^+(0)}{n_e^+(0)}, \qquad (1)$$

where $n_i^+(0)$ and $n_e^+(0)$ are the concentrations of ions and electrons emitted from the collector immediately on its surface. Recently [6, 7], the method for computing potential distributions (PDs) in an SIKD has been suggested, and the *I–V* characteristics of the device in the presence of electron emission from the collector in underneutralized conditions ($\gamma < 1$) have been analyzed. In the calculations, it was assumed that the emission from the collector is unrestricted; that is, the electron flow into the plasma is defined by the potential barrier near the collector rather than by its work function. It has been shown [6] that, under negative voltages in the underneutralized regime, the PD has a minimum near the collector (i.e., a virtual anode arises), while the PD near the emitter is monotonic. Also, in the underneutralized regime, the saturation current in the reverse branch of the *I–V* characteristic has been found to be proportional to the saturation current in the forward branch, with the proportionality coefficient being small (about 0.2) [7]. In this work, we study the PDs and *I–V* characteristics in the reverse current range under overneutralized conditions ($\gamma > 1$). In this case, unlike the underneutralized regime, the PDs are more complicated: they are no longer monotonic near the emitter.

This work consists of two parts. In the first part, as in [6, 7], the analysis is performed for plasma conditions, where the potential distribution has a quasi-neutrality portion (potential plateau). The results are compared with experimental data for Cs–Ba-filled TCs. In the second part, an SIKD with a small electrode spacing, where the near-electrode potential jumps overlap, is studied. In this case, electron emission from the emitter can be neglected if γ is high and one can consider only two opposing flows of unlikely charged particles: ions from the emitter and electrons from the collector. The asymptotic behavior of the current at small electrode spacings and high negative voltages is investi-

1063-7842/02/4712-1509$22.00

gated. These data are compared with the classical Langmuir results for the problem of two opposing flows of electrons and ions [8].

CALCULATION METHOD

We assume that the diode operates under overneutralized conditions and that the emission from the collector is unrestricted. Then, at moderate negative voltages, the PD has a minimum in the electron-enriched near-collector region; that is, a virtual anode (VA) forms [6]. However, unlike underneutralized conditions, the PD between the emitter and the top of the electron VA is nonmonotonic: because of the excess of the ion space charge near the emitter, a potential maximum arises, which may be called a virtual cathode (VC) by analogy with the VA.

The potential diagram for this case is depicted in Fig. 1. The distance from the emitter is measured in terms of the Debye length $\lambda_D = [kT_E/(4\pi e^2 n_e^+(0))]^{1/2}$, and the potentials, emitter and collector work functions, and applied voltage are expressed through kT_E/e (e is the charge of an electron, k is the Boltzmann constant, and T_E is the emitter temperature). The associated dimensionless parameters are as follows: ζ is the distance to the emitter; η, potential; χ_E, emitter work function; χ_C, collector work function; and u, applied voltage. The potential diagram exhibits three characteristic potentials: plasma potential η_p, potential at the top of the electron VA η_m, and that at the top of the ionic VC η_M.

As in [6], the concentrations of the particles at the point with the potential η can be represented as

$$n_i = n_i^+(0)F_1(\eta, \eta_M), \quad n_{eE} = n_e^+(0)F_{eE}(\eta, \eta_m),$$
$$n_{eC} = n_e^-(\eta_m)F_{eC}(\eta, \eta_m, \Theta). \tag{2}$$

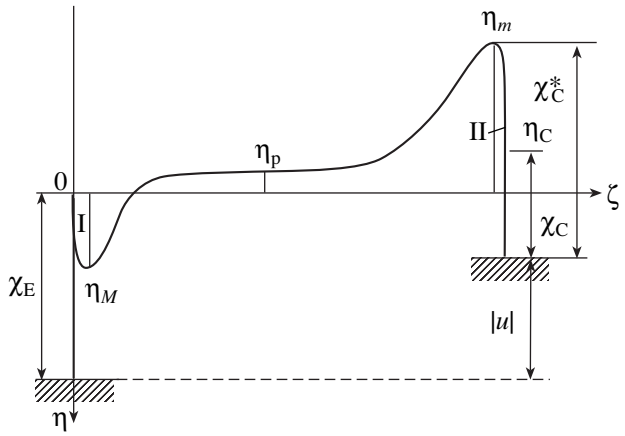

Fig. 1. Potential diagram for an SIKD under overneutralized conditions and negative voltages in the presence of electron emission from the collector. η_m, potential at the top of the electron VA; η_M, potential at the top of the ionic VC; η_p, plasma potential; χ_E, emitter work function; χ_C, collector work function; and u, applied voltage.

Here, the subscripts E and C indicate the place where the particles originate, $\Theta = T_E/T_C$ (T_C is the collector temperature), and $n_e^-(\eta_m)$ is the concentration of electrons emitted from the collector at the point with the potential η_m (the top of the VA). Under the assumption that the particles are emitted with the semi-Maxwellian velocity distribution functions, the functions F in the interval between the points with the potentials η_M and η_m have the form

$$F_i(\eta, \eta_M) = \exp(-\eta_M)\mathrm{exers}(\eta_M - \eta),$$
$$F_{eE}(\eta, \eta_m)$$
$$= \begin{cases} 2\mathrm{exers}(\eta) - \exp(\eta_m)\mathrm{exers}(\eta - \eta_m), & \eta > 0 \\ 2\exp(\eta) - \exp(\eta_m)\mathrm{exers}(\eta - \eta_m), & \eta < 0, \end{cases} \tag{3}$$
$$F_{eC}(\eta, \eta_m, \Theta) = \mathrm{exers}[\Theta(\eta - \eta_m)],$$
$$\mathrm{exers}(x) = \exp(x)(1 - \mathrm{erf}(\sqrt{x})),$$
$$\mathrm{erf}(x) = \frac{2}{\sqrt{\pi}}\int_0^x \exp(-t^2)dt. \tag{4}$$

Formulas (3) can also be used for finding the concentrations in regions I and II in the potential diagram (Fig. 1), which are adjacent to the electrodes. Specifically, in region I, the electron concentration is found directly by these formulas, while in region II, the ion concentration is found by formula (3) for F_{eE} at $\eta < 0$ if η is replaced by $-\eta$ and η_m, by $-\eta_m$. The expression for the concentration of collector electrons in region II is given by

$$F_{eC}^{II}(\eta, \eta_m, \Theta) = 2\exp[\Theta(\eta - \eta_m)] - \mathrm{exers}[\Theta(\eta - \eta_m)].$$

In [6], two conditions for finding the potentials η_p and η_m were used at $\gamma < 1$: the plasma quasi-neutrality condition and the condition of vanishing total charge on the outer part of the ionic VA (between the points with the potentials η_p and η_m). In the region where the PD is monotonic, the total charge vanishes along with the integral of the concentration over the potential. The result of integration of the function $F_{i(e)}$ over the interval (η_1, η) will be denoted as

$$G_{i(e)}(\eta, \eta_1; \ldots) = \int_{\eta_1}^{\eta} F_{i(e)}d\eta', \tag{5}$$

where the ellipsis stands for other arguments of the function $G_{i(e)}$ (sometimes they will be omitted). For the integral of the total concentration over this interval, we have

$$G(\eta, \eta_1) = \gamma G_i(\eta, \eta_1) - \beta G_{eC}(\eta, \eta_1) - G_{eE}(\eta, \eta_1); \tag{6}$$

therefore, a set of equations for the characteristic potentials can be written as

$$\begin{aligned}\gamma A_1 + \beta B_1 + C_1 &= 0,\\ \gamma A_2 + \beta B_2 + C_2 &= 0,\\ \gamma A_3 + \beta B_3 + C_3 &= 0.\end{aligned} \tag{7}$$

Here $\beta = n_e^-(\eta_m)/n_e^+(0)$,

$$\begin{aligned}A_1 &= F_i(\eta_p, \eta_M) = \exp(-\eta_M)\mathrm{exers}(\eta_M - \eta_p),\\ B_1 &= -F_{eC}(\eta_p, \eta_m, \Theta) = -\mathrm{exers}[\Theta(\eta_p - \eta_m)],\\ C_1 &= -F_{eE}(\eta_p, \eta_m)\\ &= \exp(\eta_m)\mathrm{exers}(\eta_p - \eta_m) - 2\exp(\eta_p),\\ A_2 &= -G_i(\eta_p, \eta_M) = \exp(-\eta_M)[\mathrm{Ir}(\eta_M - \eta_p) - 1],\\ B_2 &= G_{eC}(\eta_p, \eta_M; \eta_m, \Theta)\\ &= \frac{1}{\Theta}\{\mathrm{Ir}[\Theta(\eta_p - \eta_m)] - \mathrm{Ir}[\Theta(\eta_M - \eta_m)]\},\\ C_2 &= G_{eE}(\eta_p, \eta_M; \eta_m) = 2[\exp(\eta_p) - \mathrm{Ir}(\eta_M)]\\ &\quad - \exp(\eta_m)[\mathrm{Ir}(\eta_p - \eta_m) - \mathrm{Ir}(\eta_M - \eta_m)],\\ A_3 &= G_i(\eta_p, \eta_m; \eta_M)\\ &= \exp(-\eta_M)[\mathrm{Ir}(\eta_M - \eta_m) - \mathrm{Ir}(\eta_M - \eta_p)],\\ B_3 &= -G_{eC}(\eta_p, \eta_m; \Theta) = \frac{1}{\Theta}\{1 - \mathrm{Ir}[\Theta(\eta_p - \eta_m)]\},\\ C_3 &= -G_{eE}(\eta_p, \eta_m)\\ &= \exp(\eta_m)[\mathrm{Ir}(\eta_p - \eta_m) + 1] - 2\exp(\eta_p).\end{aligned} \tag{8}$$

To calculate $G_{i(e)}$, we use the formula

$$\begin{aligned}\int_{\eta_1}^{\eta}\mathrm{exers}(\eta')d\eta' &= \mathrm{Ir}(\eta) - \mathrm{Ir}(\eta_1),\\ \mathrm{Ir}(x) &= \mathrm{exers}(x) + 2\sqrt{x/\pi}.\end{aligned} \tag{9}$$

Unlike γ, the parameter β in (7) is not independent. It is related to the external parameters χ_E, Θ, and u through the relationship

$$\beta = \Theta^{-3/2}\exp[(\eta_m - u)\Theta - \chi_E(\Theta - 1)], \tag{10}$$

which is readily found if we assume that the VA emits electrons like an electrode with the effective work function $\chi_C^* = \chi_E - \eta_m + u$ (Fig. 1) and the temperature T_C. In this case, it is sufficient to express the concentrations $n_e^-(\eta_m)$ and $n_e^+(0)$ entering into the definition of β through the corresponding emission current densities $j_e^-(\eta_m)$ and $j_e^+(0)$ and take advantage of the Richardson equation. Simultaneously, it is easy to derive a formula for the density of the electron current through the diode:

$$\frac{j}{j_e^+(0)} = \exp(\eta_m) - \frac{\beta}{\sqrt{\Theta}}. \tag{11}$$

Here, the first and second terms are the contributions of emitter and collector electrons, respectively, to the current. Formulas (10) and (11) coincide with the corresponding formulas in [6].

In calculating the *I–V* characteristic for a given degree of neutralization γ as the parameter defining a point in the *I–V* curve, it is suitable to specify the potential η_m at the top of the electron VA. Set (7) can be viewed as a set of linear equations for γ and β. With this in mind, to find η_M and η_p for given γ_0 and η_m, one can proceed as follows. Let us prescribe, along with η_m, a tentative value of η_M. Then, from the compatibility condition for the set of linear equations

$$\det\begin{vmatrix}A_1 & B_1 & C_1\\ A_2 & B_2 & C_2\\ A_3 & B_3 & C_3\end{vmatrix} = 0 \tag{12}$$

one can find the plasma potential η_p, and the solution of the set yields the values of γ and β. The value of γ found differs, as a rule, from γ_0. Therefore, to solve the problem for a given γ_0, it is necessary to repeat the above procedure by varying η_M until γ coincides with γ_0. In other words, this procedure defines γ as a function of η_M, and η_M is found by numerically solving the equation $\gamma(\eta_M) - \gamma_0 = 0$. The solution to this equation, as well as the solution to Eq. (12), can be found, for example, by the chord method. The voltage at the desired point in the *I–V* characteristic is found by the formula

$$u = \eta_m - \frac{1}{\Theta}\left[\ln(\beta) + \frac{3}{2}\ln(\Theta) + \chi_E(\Theta - 1)\right], \tag{13}$$

which is readily derived from (10), and the electron current is calculated directly from formula (11).

The parameters corresponding to the point α in the *I–V* characteristic where $\eta_p = \eta_m$, i.e., where the electron current vanishes, will be marked by the subscript α. In [6], a method allowing one to analytically relate the parameters β and η_p at the point α was described. It uses the fact that, in the layer (η_p, η_m), there always (in particular, when $\eta_p \longrightarrow \eta_m$) exists a point η^* where the derivative of the total charge density with respect to the potential vanishes. Applying this expedient to overneutralized conditions, one can readily show that the relation between the parameters β_α and η_α remains the same as in the underneutralized regime:

$$\beta_\alpha = \frac{\exp(\eta_\alpha)}{\sqrt{\Theta}}. \tag{14}$$

Obviously, the formulas for u_α and j_α also remain valid:

$$u_\alpha = (\eta_\alpha - \chi_E)\left(1 - \frac{1}{\Theta}\right) - \frac{1}{\Theta}\ln\Theta,$$
$$\frac{j_\alpha}{j_e^+(0)} = \exp(\eta_\alpha)\left(1 - \frac{1}{\Theta}\right). \quad (15)$$

These formulas follow from (11), (13), and (14). However, unlike the underneutralized regime, where the parameter η_α is found only from the quasi-neutrality condition, in our case it is found simultaneously from the first two equations of set (7), which at the point α take the form

$$\gamma\exp(-\eta_{M\alpha})\mathrm{exers}(\eta_{M\alpha} - \eta_\alpha) - \exp(\eta_\alpha)\left(1 + \frac{1}{\sqrt{\Theta}}\right) = 0,$$
$$\gamma\exp(-\eta_{M\alpha})[\mathrm{Ir}(\eta_{M\alpha} - \eta_\alpha) - 1]$$
$$+ \Theta^{-3/2}\exp(\eta_\alpha)\{1 - \mathrm{Ir}[\Theta(\eta_{M\alpha} - \eta_\alpha)]\} \quad (16)$$
$$+ 2[\exp(\eta_\alpha) - \mathrm{Ir}(\eta_{M\alpha})]$$
$$- \exp(\eta_\alpha)[1 - \mathrm{Ir}(\eta_{M\alpha} - \eta_\alpha)] = 0.$$

An algorithm for the numerical solution of set (16) is similar to that for the solution of set (7).

It should be noted that for the calculation of the *I–V* characteristic in the plasma regime, the characteristic potentials in the PD will suffice. However, if necessary, the entire PD can be constructed for any point in the *I–V* curve, i.e., as was described in [6]. The absolute value of the electric field strength at a point η is known: $|\varepsilon| = [-2G(\eta, \eta_1)]^{1/2}$, where $G(\eta, \eta_1)$ is the integral of concentration (6) over the potential from the potential η_1 at the point with $\varepsilon = 0$ (the top of the VC or VA) to the potential η. It should be emphasized that the function $G(\eta, \eta_1)$ thus defined is always negative. From the known field strength and using the relationship $\varepsilon = -d\eta/d\zeta$, one finds the dependences $\zeta(\eta)$, i.e., the PD in the monotonicity ranges. For example, for the region between the top of the VA and the potential plateau, we have

$$\zeta_m - \zeta = \int_{\eta_m}^{\eta} [-2G(x, \eta_m)]^{-1/2} dx. \quad (17)$$

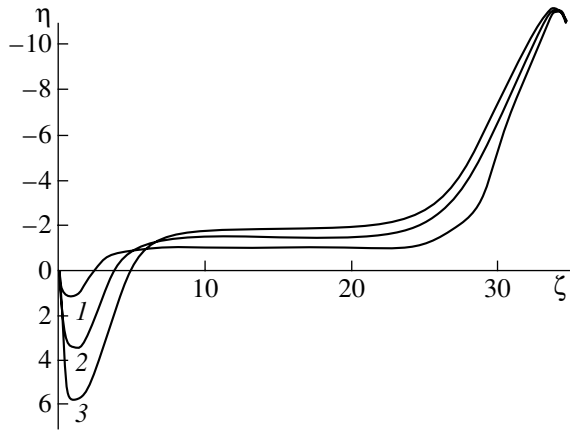

Fig. 2. Potential distribution for $\chi_E = 15$, $\Theta = 2$, and $u = -20$. $\gamma = 10$ (*1*), 100 (*2*), and 1000 (*3*). Corresponding points I–III are shown in Fig. 4.

RESULTS OF CALCULATIONS

With the approach described above, we performed the calculations for three values of the parameter γ in the range of negative voltages ($u < u_\alpha$) (Figs. 2–6). Figure 2 exemplifies the potential distributions; Figs. 3 and 4, the dependences of the potentials at the characteristic points and the voltage dependence of the electron current. At the point α, the current is always positive but goes to the range of reverse currents as the voltage deviates slightly from this point.

As follows from Fig. 4, the reverse current in the vicinity of the point α grows sharply with voltage magnitude. Subsequently, the growth of the current slows down markedly and the curves tend to saturation. At high $|u|$, the reverse current is more than twice as high as the emission current of the emitter $j_e^+(0)$ and is virtually independent of γ. Since at $\gamma > 1$ the forward saturation current coincides with $j_e^+(0)$, the basic result of the calculation can be stated in the following way: The reverse saturation current under overneutralized conditions in the presence of electron emission from the collector is proportional to the forward saturation current, with the proportionality factor being about 2. Note that the proportionality between the reverse and forward saturation currents was also observed in the underneutralized regime [7], but the proportionality factor was found to be about 0.2. The fact that the ratio of the reverse and forward saturation currents at $\gamma > 1$ exceeds that at $\gamma < 1$ by one order of magnitude can be used to identify the operating conditions of the SIKD.

Let us compare our results with those obtained earlier for a nonemitting collector (for details, see [3]). If the electron emission from the collector is absent, the collector potential η_C is negative, and $\gamma > 1$, the PD has an ionic VC (as in Fig. 1) but the electron VA is absent. In [3], detailed calculations for the underneutralized regime were performed and the principle of symmetry was stated. According to this principle, if the potential and charged particle concentration distributions are known for given γ and η_C, the corresponding distributions for the degree of neutralization $1/\gamma$ and collector potential $-\eta_C$ are found by substituting $-\eta$ for η and commuting the subscripts referring to the particles. In particular, the distribution of the ion concentration and the ion current density under overneutralized conditions coincide with those of the electron concentration

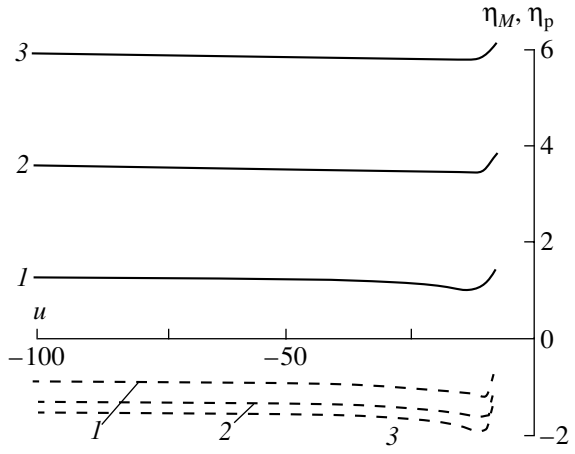

Fig. 3. Voltage dependences of the characteristic potentials η_M (solid lines) and η_p (dashed lines) for $\gamma = 10$ (*1*), 100 (*2*), and 1000 (*3*).

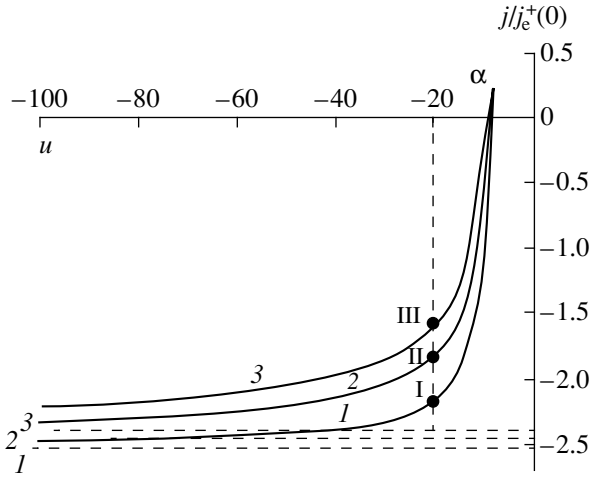

Fig. 4. Voltage dependences of the electron current for $\gamma = 10$ (*1*), 100 (*2*), and 1000 (*3*). $\chi_E = 15$ and $\Theta = 2$; dashed lines depict associated asymptotic limits.

and electron current density in the underneutralized regime.

The voltage dependences of the ion current j_i^0 [in units of $j_e^+(0)(m_e/m_i)^{1/2}$] that were obtained by using data in [3] (see also [9]) and the principle of symmetry are shown in Fig. 5a by dashed curves. The ionic current grows appreciably at voltages from 0 to several units of kT_E/e because of the increase in emitter electron concentration (and, consequently, the additional neutralization of the ion charge near the VC) due to electron reflection from the potential jump at the collector. The ion current is saturated after all the electrons have been reflected. This effect also shows up in the voltage dependences of the ionic current in the absence of the emission from the collector (solid lines in Fig. 5a).

Roughly, the reverse current vs. voltage dependence in the presence of the emission from the collector can be found within a simple model assuming that the collector electrons have no effect on the ionic current from the emitter and the electron current from the collector j_{eC} and the ionic current are related as

$$\frac{j_{eC}}{j_i}\left(\frac{m_e}{m_i}\right)^{1/2} = 1. \qquad (18)$$

This relationship approximates the equality condition for the total charges in the spacing between the virtual electrodes. Substituting j_i^0 into (18) yields j_{eC}, which equals the reverse current at high negative voltages: $j = j_i^0 (m_i/m_e)^{1/2}$. For $\gamma = 10$, the dependence of $j/j_e^+(0)$ on u is depicted in Fig. 5b by the dashed line. For comparison, the result of the self-consistent calculation is also shown. In Fig. 6, the dependence of $(j_e/j_i)(m_e/m_i)^{1/2}$ on u obtained by the self-consistent calculation (solid curve) is presented. At intermediate $|u|$, the curve runs much below unity. At the same time, as follows from Fig. 5a, j_i exceeds j_i^0 at these $|u|$. In finding the electron current, both effects compensate each other; as a result, the model suggested adequately describes the behavior of the reverse current (Fig. 5b).

We stress that the behavior of the curves in Figs. 4 and 5, which were obtained in the presence of electron emission from the collector, reflects (1) the penetration of collector electrons into the emitter region and their

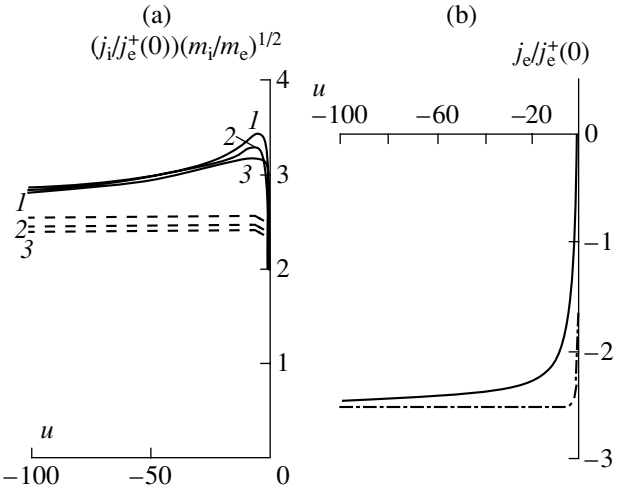

Fig. 5. (a) Voltage dependences of the ion current with (solid curves) and without (dashed curves) emission from the collector (in both cases, the zero voltage is placed at the point α). $\chi_E = 15$, $\Theta = 2$. $\gamma = 10$ (*1*), 100 (*2*), and 1000 (*3*). (b) Voltage dependence of the electron current for $\gamma = 10$. Solid curve, exact calculation; dash-and-dot line, approximate calculation with using the ion current and relationship (18) without considering emission from the collector.

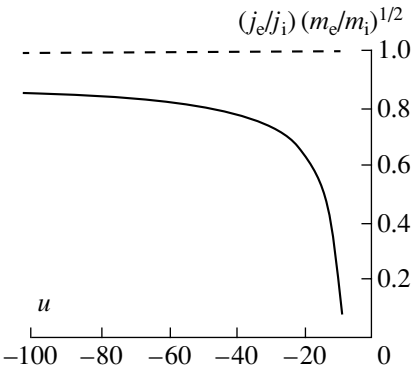

Fig. 6. Voltage dependence of the ratio between the electron current from the collector and the ion current from the emitter. Solid curve, exact calculation for $\gamma = 10$; dashed line, model representation.

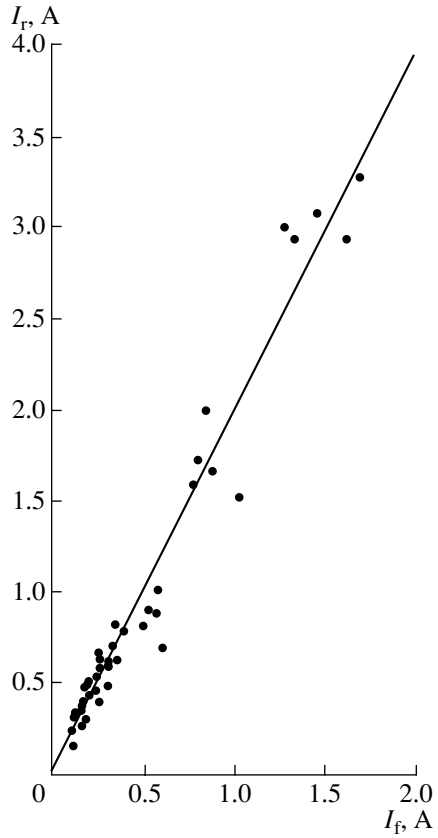

Fig. 7. Experimental dependence of the reverse saturation current I_r on the forward saturation current I_f under overneutralized operating conditions of the SIKD. Dots, experimental data; solid line, processing of the experimental data by the least squares method.

effect on the neutralization of the ionic space charge near the emitter and (2) the effect of emitter electrons on the passage of ions into the near-collector region and the neutralization of the electron charge near the collector.

It should also be noted that the approximate equations for the current coincide, at high voltages, with the zeroth-order equations used in finding asymptotic solutions at $u \longrightarrow \infty$ (see Appendix, Eqs. (A2), (A3)). Therefore, the approximate calculation yields an accurate value of the reverse saturation current.

As follows from Fig. 4, the electron current declines slowly with increasing γ when $\gamma > 10$. With data in [9] used to find the ultimate (at $u \longrightarrow \infty$) value of the ion current, we obtain the saturation electron current values 2.54, 2.46, and 2.41 for $\gamma = 10$, 100, and 1000, respectively. These asymptotic values of the electron current are shown by dashed lines in Fig. 4. It can be shown (see Appendix) that the electron current asymptotically tends to two if $\gamma \longrightarrow \infty$.

On the other hand, Fig. 3 implies that the VC must disappear, i.e., the ion current must become saturated, when γ decreases to some critical value γ^*. Hence, at $\gamma < \gamma^*$, the electron current must decrease. It is shown in the Appendix (Fig. 8) that $\gamma^* = 2.47$ for $|u| = \infty$. In the transition region ($\gamma < 2.47$), the PD changes and calls for special investigation.

Note that χ_C does not appear in the formulas for calculating the *I–V* characteristics and finding the characteristic points in the PD and, thereby, does not affect the PD. This is, however, true until the value of χ_C becomes too small. Indeed, a decrease in χ_C for a given u causes the collector potential η_C to rise. If η_C exceeds η_M, ions near the collector partially reflect and formulas (3) for F_i become incorrect. The work function χ_C at which $\eta_C = \eta_M$ is found from the equality $\chi_C = \chi_E + u - \eta_M$ (Fig. 1). Substituting the values of u and η_m at the point α into this equality yields the minimal collector work function for which ions *a fortiori* will not reflect from the collector region for any $u < u_\alpha$. For example, for $\gamma = 10$, $\chi_E = 15$, and $\Theta = 2$, the minimal value of χ_C turns out to be 5.3.

COMPARISON WITH EXPERIMENT

The asymptotic behavior of the analytical *I–V* characteristics was confirmed experimentally. The experiments were performed in Cs–Ba-filled TCs with the setup described in [10]. The results of the experimental study under overneutralized conditions at moderate (to 2000 K) emitter temperatures were published in [11].

In the experiments, the collector temperature was kept at a relatively high level (1000–1300 K). The presence of Ba allowed us to vary the emitter work function (and, hence, the parameter γ) over a wide range. The high temperature and the low work function of the collector provided unrestricted electron emission from the collector. At the same time, ion emission from the col-

lector was negligibly small. Thus, the experiments imitated conditions used in the calculations.

In the experiments, we varied the emitter temperature, as well as the cesium and barium vapor pressure, and recorded the *I–V* characteristics. From the *I–V* curves, we determined the forward, I_f, and reverse, I_r, currents. It was found that if the electron emission from the emitter under overneutralized conditions is increased severalfold (e.g., by increasing the emitter temperature), the reverse current increases roughly by the same factor.

To determine the proportionality coefficient between the forward and reverse saturation currents, the experimental values of these currents were plotted as data points in the plane (I_f, I_r) (Fig. 7). The proportionality coefficient found by linearly approximating the experimental I_r vs. I_f dependence with the least squares method was equal to 1.9 ± 0.3. This value is in reasonable agreement with the calculated asymptotic limit j^0 of the reverse current: for the range of γ considered, $j^0/j_e^+(0) \approx 2.5$ (see Fig. 4 and Appendix). The value of I_r/I_f found experimentally is somewhat less than the calculated value, possibly because the data points were taken at voltages lower than 10 V, where the asymptotic behavior of the current is not quire clear. High reverse voltages could not be used, since the TC might pass into the arc regime. Moreover, data points corresponding to small γ (close to γ^*) could decrease the current ratio.

Thus, in this work we elaborated a method of self-consistent calculation of the potential distribution and *I–V* characteristics under overneutralized conditions of SIKD operation in a range of reverse currents. It was shown that the potential distribution has an ionic virtual cathode near the emitter and an electron virtual anode near the collector. The calculations of the potential distribution and *I–V* branch in the range of reverse currents (at $u < u_\alpha$) for several values of the degree of neutralization were performed. For overneutralized conditions, the calculated reverse saturation current was found to exceed the forward saturation current by a factor of 2.5 in a wide range of γ. This is in good agreement with the experimental data. The fact that the ratio of the reverse and forward saturation currents under overneutralized conditions exceeds that in the underneutralized regime by one order of magnitude can be used for experimentally identifying SIKD operating conditions. In the case of a large electrode spacing, the *I–V* characteristic is constructed by solving a set of transcendental equations (7) and finding the potentials at points with a zero electric field. In the regions of monotonicity, the potential distribution was calculated only for greater clarity. The situation is quite different when the spacing is small. In this case, the plasma region disappears, the electrode regions overlap, and the *I–V* characteristic can be constructed only if the complete potential distribution is known. An SIKD with three particle flows and a small electrode spacing will be considered in part II of this work.

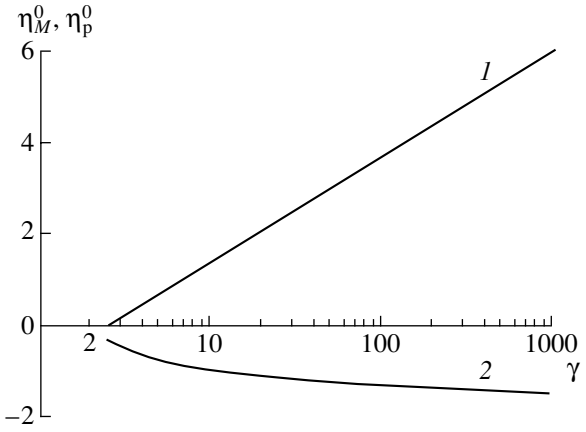

Fig. 8. γ dependence of the zeroth-order approximation parameters (*1*) η_M^0 and (*2*) η_p^0.

ACKNOWLEDGMENTS

This work was supported by the Russian Foundation for Basic Research (project no. 00-02-16882).

APPENDIX

Asymptotic Behavior of Electron Current under High Negative Voltages

Consider the potential evolution at $u \longrightarrow -\infty$. In this case, clearly, $\eta_m \longrightarrow -\infty$. An asymptote will be found under the assumption that a quasi-neutrality plateau in the potential distribution exists, i.e., that the electrode spacing is small. Using the asymptotic behavior of the functions exers(x) and Ir(x) in set (4) at $\eta_m \longrightarrow -\infty$ (namely, $\mathrm{exers}(x) \sim 1/\sqrt{\pi x}$, $\mathrm{Ir}(x) \sim 2\sqrt{x/\pi} + 1/\sqrt{\pi x}$ at $x \longrightarrow \infty$), one can show that η_M, η_p, β, and j are expanded in the small parameter $\omega = 1/\sqrt{-\eta_m}$:

$$\begin{aligned} \eta_M &= \eta_M^0 + \eta_M^1\omega + \ldots, \\ \eta_p &= \eta_p^0 + \eta_p^1\omega + \ldots, \\ \beta &= \beta^0 + \beta^1\omega + \ldots, \\ j &= j^0 + j^1\omega + \ldots. \end{aligned} \tag{A1}$$

Then, in the zeroth-order expansion in ω, we have the following set of equations for η_M, η_p, and β:

$$\begin{aligned} &\gamma\exp(-\eta_M^0)\mathrm{exers}(\Delta\eta^0) - 2\exp(-\eta_p^0) = 0, \\ &\gamma\exp(\eta_M^0)[\mathrm{Ir}(\Delta\eta^0) - 1] - 2[\mathrm{Ir}(\eta_M^0) - \exp(\eta_p^0)] = 0, \\ &\beta^0 = \gamma\sqrt{\Theta}\exp(-\eta_M^0), \end{aligned} \tag{A2}$$

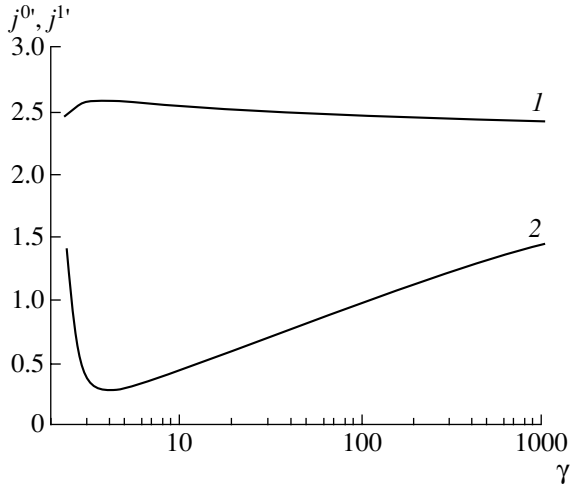

Fig. 9. γ dependence of the (*1*) zeroth-order current j^0 and (*2*) coefficient j^1 by ω in the first-order correction to the current.

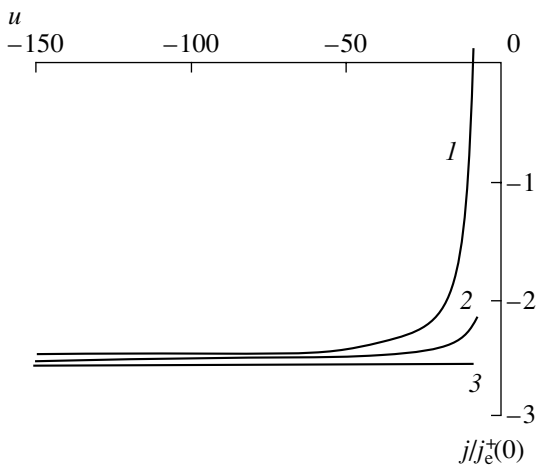

Fig. 10. Voltage dependences of the current in various approximations for γ = 10: (*1*) exact solution, (*2*) j_{as} (calculation by formula (A7)), and (*3*) asymptotic limit j^0.

where $\Delta\eta^0 = (\eta_M^0 - \eta_p^0)$. The zeroth-order approximations for η_M^0 and η_p^0 are found from the first two equations of set (A2). It is interesting that these equations coincide with equations for overneutralized conditions in the absence of electron emission from the collector [3] at high negative values of *u*, i.e., under ion current saturation conditions. This is because the concentration of electrons emitted from the collector is negligibly small in the plateau region and near the emitter when $|u| \longrightarrow \infty$. Thus, finding η_M^0 and η_p^0 is reduced to using the results of the problem solved previously.

Once η_M^0 and η_p^0 have been found, the parameter β^0 is determined from the third equation of set (A2). With formula (11), we obtain a zeroth-order expression for the electron current:

$$j^{0'} = \left|\frac{j^0}{j_e^+(0)}\right| = \frac{\beta^0}{\sqrt{\Theta}} = \gamma e^{-\eta_M^0}. \qquad \text{(A3)}$$

It is easy to check that (A3) is equivalent to (18), i.e., that asymptotic relationship (18) is exact. The γ dependences of η_M^0 and η_p^0 are depicted in Fig. 8. It is seen that η_M^0 is a nearly linear function of ln(γ) and vanishes at $\gamma = \gamma^* = 2.47$. In addition, the slope of this linear dependence is close to unity. Then, from (A3) it follows that the current $j^{0'}$ must be almost independent of γ and be close to γ*. This is illustrated in Fig. 9 (curve *1*). In a wide range of γ, $j^{0'} \approx 2.5$. Note also that, according to Fig. 8, γ* is the lower limit of the degree of neutralization below which the solution with an ionic VC is absent at high negative voltages.

From set (A2) at $\gamma \longrightarrow \infty$, one can find analytical asymptotic formulas for the parameters of the zeroth-order approximation:

$$\eta_M^0 \sim \ln\gamma, \quad \eta_p^0 \sim -\frac{1}{2}\ln(\ln\gamma),$$
$$j^0(\gamma) = \gamma\exp(-\eta_M^0) \longrightarrow 2. \qquad \text{(A4)}$$

Thus, if the degree of neutralization is infinitely large, the ratio of the reverse and forward saturation currents equals two. It should be noted, however, that this ratio tends to saturation (i.e., to two) very slowly (Fig. 8) and actually the saturation is never reached.

After tedious transformations, we find for the coefficients by ω in (A1)

$$\eta_M^1 = \frac{\beta^0}{\sqrt{\pi\Theta}}\frac{(p_2 - p_1)}{d}, \quad \eta_p^1 = \frac{\beta^0}{\sqrt{\pi\Theta}}\frac{(q_1 - q_2)}{d},$$
$$\beta^1 = \gamma\exp(-\eta_M^0) \qquad \text{(A5)}$$
$$\times\left[\frac{\sqrt{\Theta}}{2} - \sqrt{\Theta}\eta_M^1 - \frac{\sqrt{\Theta}}{2}\mathrm{Ir}(\Delta\eta^0) - \sqrt{\pi}\Theta\exp(\Delta\eta^0)\right],$$

where $d = q_1p_2 - q_2p_1$,

$$q_1 = -\gamma/\sqrt{\pi\Delta\eta^0}\exp(-\eta_M^0),$$
$$p_1 = -\gamma[\mathrm{exers}(\Delta\eta^0) - 1/\sqrt{\pi\Delta\eta^0}]$$
$$\times\exp(-\eta_M^0) - 2\exp(-\eta_M^0),$$
$$q_2 = \frac{1}{\Delta\eta^0}[\exp(-\eta_M^0)(1 - \mathrm{Ir}(\Delta\eta^0))$$
$$+ \mathrm{exers}(\Delta\eta^0) - 2\mathrm{exers}(\eta_M^0)],$$
$$p_2 = \frac{1}{\Delta\eta^0}[-\gamma\exp(-\eta_M^0)\mathrm{exers}(\Delta\eta^0) + 2\mathrm{exers}(\eta_M^0)].$$

From (11), using the value of β^1, we find for the coefficient by ω in current expansion (A1)

$$j^{1'} = \left|\frac{j^1}{j_e^+(0)}\right| = \frac{\beta^1}{\sqrt{\Theta}}. \quad \text{(A6)}$$

The function $j^{1'}(\gamma)$ is shown in Fig. 9 (curve *2*). It is seen that, unlike $j^{0'}$, $j^{1'}$ significantly depends on γ. Figure 10 demonstrates the exact solution j, asymptotic dependence (first-order approximation) $j_{as} = j^0 + j^1\omega$, and the asymptotic limit j^0. The first-order approximation gives a slightly overestimated value of the current. However, as an estimate, this value can be used even at relatively low reverse voltages.

REFERENCES

1. R. G. Mc. Intyre, J. Appl. Phys. **33**, 2485 (1962).
2. W. Ott, Z. Naturforsch. A **22**, 1057 (1967).
3. A. Ya. Énder, Candidate's Dissertation (Leningrad, 1972).
4. S. Kuhn, Plasma Phys. **23**, 881 (1981).
5. R. G. Mc. Intyre, Proc. IEEE **51**, 760 (1963).
6. V. I. Sitnov and A. Ya. Énder, Zh. Tekh. Fiz. **68** (4), 37 (1998) [Tech. Phys. **43**, 382 (1998)].
7. V. I. Sitnov, A. Ya. Énder, and E. V. Yakovlev, Zh. Tekh. Fiz. **69** (4), 26 (1999) [Tech. Phys. **44**, 373 (1999)].
8. I. Langmuir, Phys. Rev. **33**, 954 (1929).
9. V. I. Babanin, V. I. Kuznetsov, A. S. Mustafaev, *et al.*, Zh. Tekh. Fiz. **48**, 754 (1978) [Sov. Phys. Tech. Phys. **23**, 444 (1978)].
10. V. I. Babanin, Yu. A. Dunaev, A. S. Mustafaev, *et al.*, Zh. Tekh. Fiz. **43**, 1916 (1973) [Sov. Phys. Tech. Phys. **18**, 1211 (1973)].
11. V. I. Babanin, Yu. A. Dunaev, A. S. Mustafaev, *et al.*, Zh. Tekh. Fiz. **42**, 1662 (1972) [Sov. Phys. Tech. Phys. **17**, 1326 (1972)].

Translated by V. Isaakyan

Technical Physics, Vol. 47, No. 12, 2002, pp. 1518–1523. Translated from Zhurnal Tekhnicheskoĭ Fiziki, Vol. 72, No. 12, 2002, pp. 32–38.

GAS DISCHARGES, PLASMA

Study of the Reverse Current Range in a Knudsen Diode with Surface Ionization under Overneutralized Conditions in the Presence of Electron Emission from the Collector: Part II. Small Electrode Spacing

V. I. Babanin*, V. I. Sitnov*, A. V. Solov'ev, and A. Ya. Énder***

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

*** Institute of Chemistry and Applied Engineering Science, Aalborg University of Esbjerg, 6700 Esbjerg, Denmark*

Received April 24, 2002

Abstract—A Knudsen diode with surface ionization under overneutralized conditions is studied in the range of reverse current. The electrode spacing is assumed to be small, which causes the potential jumps near the electrodes to overlap. The asymptotic behavior of the current under high negative voltage is investigated. © *2002 MAIK "Nauka/Interperiodica".*

INTRODUCTION

This paper continues our study of a Knudsen diode with surface ionization (SIKD) in the case of unrestricted electron emission from the collector under overneutralized conditions, that is, at degrees of neutralization $\gamma = n_i^+(0)/n_e^+(0) > 1$ (where $n_{i(e)}^+(0)$ is the concentration of ions (electrons) leaving the emitter (collector) immediately on the electrode surface). The diode involves three flows of particles: electrons and ions from the emitter and electrons from the collector.

In the first part of this work [1], we considered an SIKD with a large electrode spacing, where the potential distribution exhibits a quasi-neutrality region (potential plateau) and the potential varies within the narrow near-electrode regions. The potential distribution (PD) and *I–V* characteristics in the range of reverse currents were examined. It was shown that, unlike underneutralized conditions [2, 3], both a virtual anode (VA) near the collector and an electron virtual cathode (VC) near the emitter arise in this range of the current and that the reverse saturation current in the *I–V* characteristic is more than twice as high as the forward saturation current.

This paper is the second part of the work, where we study the diode with a small electrode spacing, where the potential plateau disappears and the near electrode potential jumps overlap. Note that the use of a small electrode spacing is one way of improving the efficiency of thermionic converters (TCs), which are, in essence, SIKDs [4]. Therefore, the analysis of small-spacing SIKDs is of great applied interest. The method for calculating the current and constructing the potential distribution in such devices, as well as associated calculations, are given in Section 1 of this paper.

If γ is large and the applied voltage is high, the effect of electron emission from the collector in small-spacing SIKDs is insignificant. In this case, the problem of finding the reverse current becomes similar to the Langmuir problem of two opposing particle flows in a diode [5]. In [5], Langmuir studied the neutralization of the space charge produced in the electrode spacing by electrons emitted from the cathode and by ions originating at the anode. In our case, ions are generated on the hot emitter due to surface ionization, while electrons are emitted by the cooler collector, the electron source being under a negative potential relative to the ion source in both cases. Consequently, the SIKD collector and emitter can be viewed as the cathode and the anode, respectively, in terms of the Langmuir problem. Therefore, when comparing associated results, one should change signs of the current and voltage values. In Section 2 of this part, we study the asymptotic behavior of the SIKD reverse current under high negative voltages and compare our data with the classical Langmuir results.

1. ANALYSIS OF A SMALL-SPACING SIKD

In the case considered in [1], the PD has a quasi-neutral plateau; in other words, the sum of the widths of the near-electrode regions within which the potential varies is much less than the electrode spacing *d*. In this case, by varying *d*, we vary only the width of the plateau without changing the passing current and the PD in the near-electrode regions. However, if the electrode spacing *d* is sufficiently small, when the near-electrode regions overlap (see the potential diagram in Fig. 1), the

1063-7842/02/4712-1518$22.00

current becomes d dependent. In Fig. 1 and thereafter, the variables and units of measure are the same as in Part I [1].

At high voltages ($|U| \gg kT_E/e$, where k is the Boltzmann constant, e is the electron charge, and T_E is the emitter temperature), the critical spacing at which the plateau in the PD disappears depends largely on the width of the near-collector region across which the potential drops. The width of this region can be estimated if it is assumed that the space charge is produced mainly by collector electrons with negligibly small initial thermal velocities. In this case, the electron current, width of the region, and the potential difference across the region are related by the well-known Child's law for a vacuum diode [6]:

$$j_{3/2} = \frac{1}{9\pi}\left(\frac{2e}{m_e}\right)^{1/2}\frac{U^{3/2}}{d^2}, \quad (1)$$

where m_e is the mass of an electron, U is the applied voltage, and d is the electrode spacing in the diode.

Using (1), we obtain for the dimensionless critical spacing

$$\tilde{\delta} = \frac{2}{3}\pi^{1/4}|u|^{3/4}\left(\frac{|j|}{j_e^+(0)}\right)^{1/2} \approx 1.25|u|^{3/4}. \quad (2)$$

Here, we took into account the fact that the relative reverse current roughly equals two in a wide range of γ under overneutralized conditions and high dimensionless voltages $|u|$ [1]. The basic difference between the PD at the critical point in the *I–V* characteristic and the PD in a vacuum diode [6] is that there is an inflection point within the spacing where the second-order derivative vanishes. Therefore, a more exact estimator of the critical spacing appears to be

$$\delta_c = 2\tilde{\delta}. \quad (3)$$

Further calculations supported the validity of such an estimate: at $\delta = \delta_c$, the calculated current exceeds the current in the plasma regime by no more than 10–20%. The critical spacings for a number of voltages are listed in Table 1. For given u, the electrode spacing will be called small if $\delta < \delta_c$. It should be noted that the above estimate becomes invalid if $|u|$ is sufficiently small.

In [1], three characteristic potentials in the *I–V* characteristic at $\delta > \delta_c$ (plasma potential η_p, potential at the top of the VA η_m, and potential at the top of the VC η_M) were calculated by three equations. The first one was derived from the condition of plasma neutrality; the second one, from the condition that the total charge between the top of the VA and the plateau (i.e., between the points with the potentials η_m and η_p) equals zero; and the third equation, from the condition that the total charge between the plateau and the top of the VC (i.e., between the points with the potentials η_p and η_M) equals zero. At small δ, the plasma potential η_p in the PD is absent and the number of the desired parameters

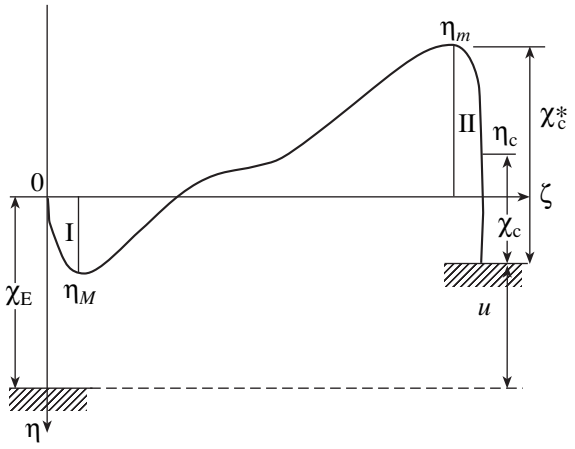

Fig. 1. Potential diagram for an SIKD under overneutralized conditions in the range of reverse currents for small electrode spacings.

diminishes by one. Of the three equations, only the one following from the zero total charge between the VA and VC tops (the points with the potentials η_M and η_m) remains. It can be written in the form

$$\gamma A + \beta B + C = 0. \quad (4)$$

Here, $\beta = n_{eC}(\eta_m)/n_e^+(0)$; $n_{eC}(\eta_m)$ is the concentration of collector electrons at the VA top; and the coefficients A, B, and C are found by summing the corresponding coefficients in the equations reflecting the absence of the charge between the plateau and the VA and VC tops [1]:

$$\begin{aligned} A &= G_i(\eta_M, \eta_m) = \exp(-\eta_M)[\mathrm{Ir}(\eta_M - \eta_m) - 1], \\ B &= -G_{eC}(\eta_M, \eta_m) = \frac{1}{\Theta}\{1 - \mathrm{Ir}[\Theta(\eta_M - \eta_m)]\}, \\ C &= -G_{eE}(\eta_M, \eta_m) \\ &= \exp(\eta_m)[\mathrm{Ir}(\eta_M - \eta_m) + 1] - 2\mathrm{Ir}(\eta_M), \end{aligned} \quad (5)$$

where $G_{i(e)}(\eta, \eta_1)$ are the integrals of the ion (electron) concentrations over the potential between η_1 and η, $\mathrm{Ir}(x) = \mathrm{exers}(x) + 2\sqrt{x/\pi}$, $\mathrm{exers}(x) = \exp(x)(1 - \mathrm{erf}(\sqrt{x}))$, and erf(x) is the error function integral.

At the same time, when δ is small, there exists an expression that equates the electrode spacing as a func-

Table 1

$\|u\|$	δ_c
40	20
70	30
100	40
136	50

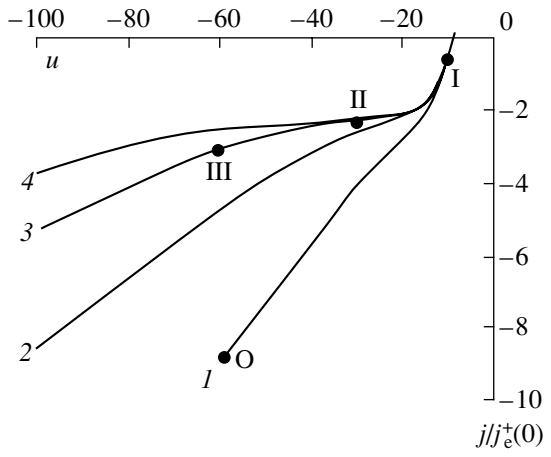

Fig. 2. Voltage dependence of the electron current for $\delta = 10$ (*1*), 15 (*2*), 20 (*3*), and 25 (*4*). $\chi_E = 15$, $\Theta = 2$.

tion of the characteristic potentials to a given value. In essence, the calculation of this spacing is reduced to the solution of Poisson's equation for the PD in the entire electrode spacing. In this case, it is necessary to specify the collector work function χ_C, since it, in turn, specifies the width of the near-collector region in the PD. Recall that for large electrode spacings, the PDs in the near-electrode regions were calculated only for greater clarity. In this case, the characteristic potentials and the current do not depend on χ_C and can be found without specifying this parameter as long as the electron emission from the collector is unrestricted.

A computing algorithm for a small electrode spacing is the following. One prescribes η_m and a tentative value of η_M. From (4), the parameter β is found and the voltage u is determined by the formula

$$u = \eta_m - \frac{1}{\Theta}\left[\ln\beta + \frac{3}{2}\ln\Theta + \chi_E(\Theta - 1)\right], \qquad (6)$$

which was derived in [1]. Then, from the relationship $\chi_E - \eta_C = \chi_C - u$ (Fig. 1), the collector potential η_C is found. Once all the characteristic potentials have been determined, one can calculate the PDs in the central region (between the VA and VC tops) and also in the regions adjacent to the electrodes (regions I and II in Fig. 1) as was described in [1]. Eventually, the widths of all three regions are found and the entire electrode spacing as a function of the characteristic potentials takes the form

$$\delta(\eta_M, \eta_m, \eta_C) = \int_{\eta_m}^{\eta_M} \frac{d\eta}{\sqrt{-2G(\eta, \eta_m)}} + \int_0^{\eta_M} \frac{d\eta}{\sqrt{-2G^I(\eta, \eta_M)}} + \int_{\eta_m}^{\eta_C} \frac{d\eta}{\sqrt{-2G^{II}(\eta, \eta_m)}}. \qquad (7)$$

Here, $G(\eta, \eta_a) = \gamma G_i(\eta, \eta_a) - \beta G_{eC}(\eta, \eta_a) - G_{eE}(\eta, \eta_a) = -\varepsilon^2(\eta)$, η_a is the potential at an extreme point (η_m or η_M), $\varepsilon(\eta)$ is the potential-dependent dimensionless electric field strength, and the superscripts I and II refer to the related near-electrode regions. By varying η_M with η_m fixed, we find η_M such that the electrode spacing equals some given value δ. Along with the electrode spacing, one also calculates the dependences $\zeta(\eta)$ in all three regions to obtain the desired PD. Finally, the current density j is found by the formula

$$\frac{j}{j_e^+(0)} = \exp(\eta_m) - \frac{\beta}{\sqrt{\Theta}}, \qquad (8)$$

which was also derived in [1]. Formulas (6) and (8) define the parametric dependence of j on u, i.e., the I–V characteristic in a voltage range for which given δ is small. It should be emphasized, however, that this algorithm is also applicable to a $\delta - u$ domain that admits a narrow plateau in the PD. On the other hand, this approach has a natural restriction when δ is sufficiently small ($|u|$ is high). Indeed, as δ shrinks ($|u|$ grows), the VC height η_M also decreases and vanishes at a certain $\delta = \delta_0$ ($u = u_0$) when the ion emission becomes restricted. This approach was applied until the VC disappeared.

The voltage dependences of the current thus obtained for $\gamma = 10$ and various electrode spacings are depicted in Fig. 2. For $\delta = 10$, the curve is drawn up to the point where $\eta_M = 0$ (the point *O*); the others are "cut" at $u = -100$. If δ is sufficiently large, the current curves flatten. The horizontal portions of the curves correspond to the reverse saturation current in the presence of the plateau in the PD (for example, the segment near point II in curve *4* for $\delta = 25$ and the PD at point II (Fig. 3, curve *2*)). As $|u|$ grows within the horizontal portions, the near-collector region in the PD widens and the plateau shrinks. Subsequently, the near-electrode regions overlap and the current grows noticeably. The larger the δ, the higher the $|u|$ at which the curve starts rising. The rise voltage at various δ agrees with estimates given in Table 1. Figure 3, which shows the PD at $\delta = 20$ and various voltages, illustrates the disappearance of the quasi-neutral plateau as $|u|$ grows.

2. COMPARISON WITH THE LANGMUIR PROBLEM OF TWO OPPOSING ELECTRON AND ION FLOWS IN A DIODE

If the electrode spacing is small, the relative contribution of emitter electrons to the total concentration decreases with increasing γ. Therefore, it is of interest to compare our results with the solution to the problem of electron and ion flows propagating from the opposite electrodes. The expression for the electron current in such a diode was first found by Langmuir [5], who solved the problem assuming that the particles had zero initial velocities. For the case of unrestricted electron and ion emission, the neutralization of the electron

space charge by ions was shown to increase the electron current j_e by a factor of 1.87 compared with the ion-free current $j_{3/2}$ that obeys Child's law (1). The ion and electron currents are related by the formula $(j_i/j_e)\sqrt{m_i/m_e} = 1$. Note that we found a mistake in Langmuir's work [5]. In the case of unrestricted ion emission, the relative electron current equals 1.8655, rather than 1.8605 as in [5] (we calculated the same integral as Langmuir to the 12th decimal place).

If the initial velocities are zero and the particle emission is unrestricted, the condition of zero electron field strength is imposed directly on the electrodes. If the initial velocities are thermal, the zero-field points are the VA and VC tops. Therefore, for comparison with the Langmuir results, we will relate the electron current to the current $j^*_{3/2}$ that follows from Child's law at the effective spacing d^* (spacing between the VA and VC tops) and voltage U^* (the potential difference between the VA and VC tops). These effective values are found during calculations.

Since the transition from the plasma condition to conditions where the near-electrode regions heavily overlap can be accomplished by narrowing the spacing, let us consider the results of calculation when δ is varied and γ is relatively high. By way of example, we present the results for $\gamma = 1000$, $\chi_E = 15$, $\chi_C = 3$, $\Theta = 2$, and $u = -25$. For these values, $\delta_c \simeq 15$, as follows from (1) and (2). During the calculations, δ was varied from 25 ($\simeq 1.7\delta_c$) to 0.415 (the distance δ_0 at which the VC disappears). Figure 4 (curve *1*) shows the δ dependence of the current ratio $I = |j|/j^*_{3/2}$. When δ is large and decreases, I drops roughly following the quadratic law. This means that the current j in the plasma regime is constant and the current $j^*_{3/2}$ grows in proportion to $1/\delta^{*2}$. At $\delta \simeq \delta_c$, the decrease of I slows down and I tends to saturation. However, at $\delta \simeq 3.3$, the curve $I(\delta)$ passes a minimum and then the current increases sharply. The minimal value of I is nearly twice as high as the Langmuir value.

Table 2 lists the ratio of the electron current to the emitter current, $I' = |j|/j^+_e(0)$, for different δ. It is seen that, for $\delta \lesssim 3$, the collector current exceeds $j^+_e(0)$ by two or three orders of magnitude. Hence, in this range of δ, the effect of emitter electrons is negligibly small and the situation is close to that considered by Langmuir. In this situation, one might expect the smooth saturation of the $I(\delta)$ curve. Note that the collector work function was taken sufficiently small, $\chi_C = 3$, so as to provide unrestricted emission from the collector for any δ. Actually, however, the collector work function is greater and the electron current may become saturated prior to reaching the ultimate point with $\eta_M = 0$.

The reason for the anomalous behavior of curve *1* in the range of small δ in Fig. 4 can be clarified by analyzing the δ dependences of δ^*/δ and $u^*/|u|$ (Fig. 5). For any δ, the effective spacing differs from the actual value insignificantly. However, u^* decreases substantially. If the particles leaving the virtual electrodes had zero velocity, a decrease in u^* with δ would not affect the value of I. However, together with corrections for thermal spread, which shift the zero-field points from the electrode tops into the electrode spacing, there is another correction due to the variation of the charged particle concentration distribution and the PD between the electrodes. A decrease in u^* must cause the PD and current ratio I to deviate further from the Langmuir values.

In [7] current passage in a vacuum diode was considered, assuming that electrons are emitted with thermal velocities. In the presence of the ion current, the asymptotic expression for the electron current under

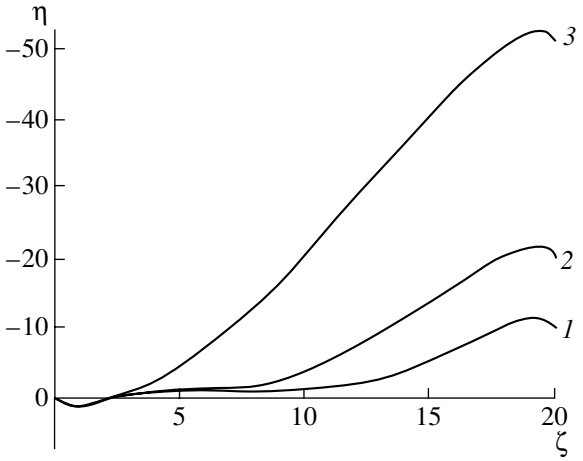

Fig. 3. Potential distribution for $\delta = 20$ and $u = 10$ (*1*), 30 (*2*), and 60 (*3*). Points I–III are shown in the *I–V* characteristics in Fig. 2.

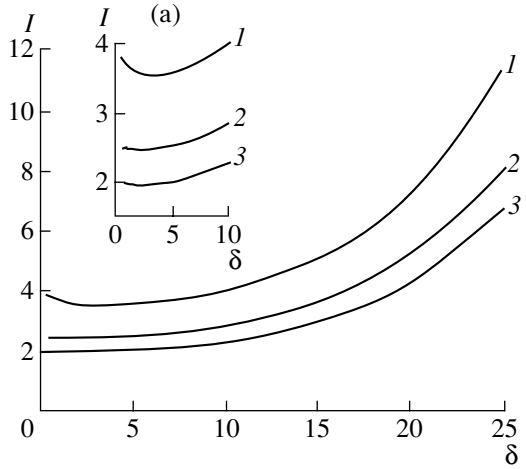

Fig. 4. Current ratio I vs. electrode spacing δ. $I = |j|/j^*_{3/2}$ (*1*) and j/j_T (*2*) for $B = 1.9$ and (*3*) $I = j/j_T$ for $B = 3.3$. Panel (a) shows the same on another scale for a range of small δ.

Table 2

δ	*I'*
24.9	1.8
14.6	2.4
7.7	6.3
5.9	10.1
2.9	35.5
1.49	117
1.1	209
0.57	597
0.415	937

Table 3

γ	$u = -25$	$u = -100$	$u = -1000$	$u = -10\,000$
5	21	4.2	0.30	0.005
10	10	2.1	0.13	0
10^2	6	1.0	0.07	0
10^3	6	1.0	0.06	0

high negative voltages takes the form (in our designations) [7]

$$j_T = -j^*_{3/2}K_T, \quad K_T = 1 + B(u^*)^{-1/2}. \qquad (9)$$

Here, $B = 2.7\Theta^{-1/2}$ (the factor $\Theta^{-1/2}$ arises because u^* is measured in terms of kT_E/e, and the spread in electron velocities depends on the collector temperature T_C). The coefficient K_T is responsible for the second correction mentioned above. It is natural to relate the current to j_T, not to $j^*_{3/2}$, i.e., to divide the dependence $I(\delta)$ (Fig. 4, curve *1*) by K_T (at $\Theta = 2$, $B = 1.9$ in (9)). The resulting dependence is depicted by curve *2* in Fig. 4. It is seen to have a smoother peak than curve *1*.

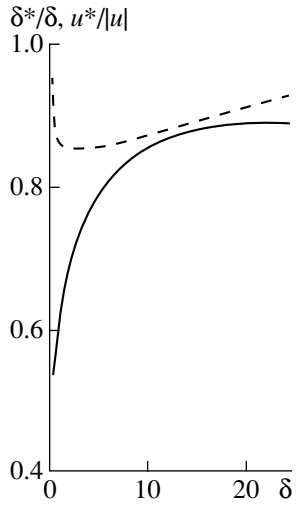

Fig. 5. δ dependences of δ*/δ (dashed line) and $u^*/|u|$ (solid line).

Thus, the nonmonotonic run of the dependence $I(\delta)$ can be explained by the second-type correction for thermal spread. However, in our problem, this spread must include velocity spreads for both electrons and ions. In SIKDs, the latter have an even higher temperature than the collector electrons. We will assume that, in the presence of ions, the K_T vs. u dependence retains its form and only the coefficient B changes. The curves were constructed for different values of B in (9). Curve *3* in Fig. 4 implies that the selection $B = 3.3$ eliminates completely the minimum in the j/j_T vs. δ curve and allows the curve to reach saturation smoothly and approach the Langmuir value for j/j_T at δ/δ_0.

The comparison with the Langmuir problem is complicated by the fact that the problem of opposing electron and ion flows in a diode has not yet been solved for the case when the particles emitted from the electrodes have semi-Maxwellian velocity distributions. This problem is of independent importance. When compared with our calculations, its solution will help in elucidating effects associated with the presence of the third flow (electrons from the emitter).

The above analysis was carried out for actual operating voltages of SIKDs (at $T_E = 2000$ K, $|u| = 25$ corresponds to a voltage of about 4 V). Formally, one can take much higher voltages and pass to the limit $|u| \longrightarrow \infty$. In this case, the thermal velocities of the particles can be ignored. Moreover, in this case, emitter electrons become localized in the infinitesimal neighborhood of the emitter and must not influence the PD and passing current. Consequently, one can expect that the solution to our problem at $|u| \longrightarrow \infty$ will approach the Langmuir solution over a wide range of γ. However, the asymptotic behavior of the current at high $|u|$ must depend on γ.

To verify these statements, we calculated the electron current at various γ over a wide voltage range. For each voltage value, we took the maximum possible δ value for the conditions of unrestricted ion emission (δ_0). The dependences obtained indicate that, at $u > 10\,000$, the diode current is well approximated by the formula

$$\frac{|j|}{j^*_{3/2}} \approx A[1 + B(u^*)^{-1/2}]. \qquad (10)$$

The coefficients A and B were found by the least squares method. It was established that A equals 1.87 for any γ, that is, coincides with the Langmuir value, while B depends on γ: $B = 4.34$ for $\gamma = 5$, 3.76 for $\gamma = 10$, 3.34 for $\gamma = 100$, and 3.305 for $\gamma = 1000$. The γ dependence of B is associated with the increased effect of emitter electrons on the current.

Note that these results were obtained numerically. To reliably estimate the coefficients A and B (to three significant figures in our case), it was necessary to carry out calculations in low-field regions near the virtual

electrodes especially carefully, since these regions introduce the greatest errors into the calculated values of δ and, eventually, limit the accuracy of finding the constants A and B. Our method makes it possible to accurately calculate the PD in an SIKD.

Table 3 gives the deviations (in percent) of the values of I calculated by asymptotic formula (10) for $\delta = \delta_0$ over a wide range of γ and u. The asymptotic formula obtained by processing the calculations for $u > 10\,000$ also adequately describes the behavior of I at much lower voltages. Such an asymptotic voltage dependence of the relative current was also found by calculating the curve $I(\delta)$ at a relatively low voltage, $|u| = 25$ (Fig. 4). Moreover, the coefficient $B = 3.3$ in (9) found in this way is nearly coincident with the coefficient B in (10) for $\gamma = 1000$. The deviations from the asymptotic behavior at low $|u|$ and small γ are explained by the joint effect of emitter electrons and thermal velocities of the particles on the current. A deviation of about 6% at $u = -25$ and large γ is associated only with the thermal spread and can be reduced by including higher order terms in the asymptotic formula.

Thus, our analysis showed that the solution to our problem approaches the Langmuir solution [5] in a wide range of γ at sufficiently high voltages.

CONCLUSION

The basic results of the investigations performed in Parts I [1] and II of this work are as follows.

A method of self-consistently calculating the PDs and I–V characteristics in an SIKD under overneutralized conditions in a range of reverse current is elaborated. It is shown that the PD has an ionic virtual cathode near the emitter and an electron virtual anode near the collector. Calculations were carried out in a range of reverse currents for a number of degrees of neutralization γ.

It is established that, under overneutralized conditions, the reverse saturation current is approximately 2.5 times higher than the forward saturation current over a wide range of γ. This result agrees well with experimental data. The fact that the reverse-to-forward saturation current ratio under overneutralized conditions is roughly one order of magnitude higher than that in the underneutralized regime can be used as an express experimental criterion for identifying SIKD operation conditions.

A computing technique in the case of small electrode spacings, where the potential distribution lacks a quasi-neutral plateau, is developed. It should be noted that the use of small electrode spacings is one way to improve the TC efficiency. The technique developed can be useful in designing small-spacing converters.

Our results are compared with those obtained by Langmuir for the problem of two opposing electron and ion flows in a diode. In the limit of very high negative voltages, the current equals the Langmuir value. It is demonstrated that the addition of the third, even minor, particle flow (electrons from the emitter) at low voltages or large electrode spacings drastically changes the potential distribution and the electron current value compared with the Langmuir solution.

The computing method and the results obtained can be used for finding the potential distributions and I–V characteristics in the forward current range on the right of the point α and also in the intermediate (between over- and underneutralized conditions) regime. This is of particular importance for the optimization of TC parameters. It is also shown that the electron virtual anode in the potential distribution near the collector disappears at the point α. In addition, the asymptotic behavior of the current at high voltage [1] suggests that a decrease in γ causes the ionic virtual cathode to disappear. Hence, one can expect radically new potential distributions in the range on the right of the point α and at intermediate values of γ.

ACKNOWLEDGMENTS

This work was supported by the Russian Foundation for Basic Research (project no. 00-02-16882).

REFERENCES

1. V. I. Babanin, V. I. Sitnov, A. V. Solov'ev, and A. Ya. Énder, Zh. Tekh. Fiz. **72** (12), 23 (2002) [Tech. Phys. **47**, 1509 (2002)].
2. V. I. Sitnov and A. Ya. Énder, Zh. Tekh. Fiz. **68** (4), 37 (1998) [Tech. Phys. **43**, 382 (1998)].
3. V. I. Sitnov, A. Ya. Énder, and E. V. Yakovlev, Zh. Tekh. Fiz. **69** (4), 26 (1999) [Tech. Phys. **44**, 373 (1999)].
4. Yu. V. Nikolaev, S. A. Eryomin, S. S. Kalmykov, *et al.*, AIP Conf. Proc. **361**, 1299 (1996).
5. I. Langmuir, Phys. Rev. **33**, 954 (1929).
6. L. N. Dobretsov and M. V. Gomoyunova, *Emission Electronics* (Nauka, Moscow, 1966); I. Langmuir, Phys. Rev. **2**, 450 (1913).
7. I. Langmuir, Phys. Rev. **21**, 419 (1923); V. R. Bursian, Zh. Russ. Fiz.-Khim. O-va. **54**, 139 (1922).

Translated by V. Isaakyan

Technical Physics, Vol. 47, No. 12, 2002, pp. 1524–1528. Translated from Zhurnal Tekhnicheskoĭ Fiziki, Vol. 72, No. 12, 2002, pp. 39–43.

GAS DISCHARGES, PLASMA

Parameters of Electronic Detonation in Solid Dielectrics

Yu. N. Vershinin

Institute of Electrophysics, Ural Division, Russian Academy of Sciences, ul. Amundsena 106, Yekaterinburg, 620016 Russia

e-mail: lfd@iep.uran.ru

Received March 29, 2002

Abstract—Using KCl as an example, the electronic detonation parameters are studied in solid dielectrics at supersonic velocities of pulsed discharge channel propagation from the anode. The technique used to estimate the detonation parameters is based on theoretical approaches and experimental methods involved in the physics of detonation, explosion, and dense nonideal plasma. © *2002 MAIK "Nauka/Interperiodica"*.

INTRODUCTION

The investigation of processes in a discharge gap with a solid dielectric that precede the occurrence of a breakdown channel bridging the gap showed that the propagation of breakdown channels is accompanied by primary phase transitions of solid dielectrics to the melt or plasma. The first case corresponds to a cathode discharge solely with a subsonic velocity v_{dis}; the second, solely to a supersonic discharge from the anode [1]. The properties of the latter process have led researchers to describe it in terms of shock waves and methods of the detonation theory. The mechanism of the supersonic propagation of a discharge channel has been named electronic detonation [1, 2].

In [3], a detonation process was defined as "the occurrence and existence of a steady-state shock wave + energy release zone combination." To date, several variations of this phenomenon have been established and studied. Along with the well-known chemical detonation of explosives, there also exist light and nuclear detonations, as well as hybrid processes, such as light-chemical detonation, electrotechnical detonation, etc. [4]. The general rules for the formation of a detonation wave are similar for all the variations. An initial material is compressed by a shock wave with a velocity D equal to the detonation velocity. In a detonation wave, the velocity u of all states is the same but the values of pressure p and volume V differ. The difference between these parameters depends on how the energy sustaining the propagation of the shock wave is released. Depending on the conditions for energy release, detonation may be overcompressed, undercompressed, overheated, or overcooled [5, 6]. The most general and most frequently used approach to the quantitative description of the detonation wave parameters is based on the classical concepts, which require the following relation to be fulfilled in the wave:

$$D = u + C, \tag{1}$$

where u and C are the mass velocity and velocity of sound in the energy release zone, respectively.

Equation (1), known as the Chapman–Jouguet relationship, describes the steady-state detonation process.

ESTIMATION OF ENERGY RELEASE ZONE PARAMETERS IN AN ELECTRONIC DETONATION WAVE

In order to estimate the parameters of an electronic detonation wave within the Chapman–Jouguet detonation theory, it is appropriate to begin by finding the possible ranges of these parameters. To this end, one can use experimental Hugoniot adiabats; i.e., the dependences of the density ρ_{sh} of the dielectric or of its compressibility $\delta = \rho_{sh}/\rho_0 = V_0/V_{sh}$ on the shock wave pressure p_{sh}. Hugoniot adiabats for many solid and liquid dielectrics are well known and presented elsewhere [7, 8]. Below, all the parameters of electronic detonation waves will be calculated for potassium chloride as an example. Thus, using Hugoniot adiabats, one can find the effective polytropic index in the equation pV^n = const for condensed media:

$$n = \frac{\delta_{sh} + 1}{\delta_{sh} - 1}. \tag{2}$$

If detonation is in any way initiated in a dielectric, the values of p and $n(p)$ found from Hugoniot adiabats allow one to calculate (in the energy release zone of the detonation wave)

the compressibility

$$\delta_D = \frac{n + 1}{n}, \tag{3}$$

the detonation velocity

$$D = \left[\frac{p(n + 1)}{\rho_0}\right]^{1/2}, \tag{4}$$

1063-7842/02/4712-1524$22.00

the specific heat of explosive transformation

$$Q = \frac{D^2}{2(n^2-1)}, \quad (5)$$

the pressure in the detonation wave

$$p_D = 2\rho_0(n-1)Q, \quad (6)$$

the internal energy

$$U_D = \frac{p_D}{\rho_0\delta_D}\frac{1}{n-1}, \quad (7)$$

and the temperature [9]

$$T = T_0\exp\left[-\int_{V_0}^{V}\left(\frac{dU_D}{dp_D}\right)^{-1}dV\right]. \quad (8)$$

As was mentioned in [10], Eq. (7) can be applied in the range where the temperature $T_0(p, U)$ can be either found experimentally or calculated reliably by methods of statistical physics. According to [9], the temperature $T = 3800$ K at a pressure $p \approx 3.3 \times 10^{10}$ Pa can be taken as the reference point T_0.

Hugoniot adiabats for potassium chloride have been studied experimentally at a pressure $p_{sh} \cong 400$ GPa. As the pressure p_{sh} increases in the range 20–400 GPa, the parameters of the detonation process in the Chapman–Jouguet approximation (Eqs. (2)–(7)) vary in the ranges $1.92 \le n \le 4.38$, $5.2 \le D \le 17.1$ km/s, $1.23 \le \delta_D \le 1.53$, $6.8 \le Q \le 54.4$ MJ/kg, $10 \le p_D \le 200$ GPa, $4.8 \le U_D \le 71.0$ MJ/kg, and $1500 \le T \le 9300$ K.[1]

For small velocities of electronic detonation and, hence, relatively low pressures in the energy-release zone, the specific heat Q sustaining the shock wave is of Joulean nature [2]. However, this mechanism can be applied within certain limits. For example, as the detonation velocity increases to 12×10^3 m/s and, hence, the detonation wave pressure increases to 80 GPa, the conductivity of alkali-halide crystals increases by six orders of magnitude [11]. Under these conditions, Joule losses can no longer sustain the shock wave. For such electronic detonation velocities, the mechanism of Q formation should be changed.

Estimates of the degree of ionization in KCl at $p_D \ge 100$ GPa, $T \ge 10^4$ K, and high densities ($\delta > 1.5$) have given rise to the model of dense nonideal plasma, where the general laws of detonation are retained [12, 13]. In this model, detonation in the energy release zone becomes a two-step process. It starts with the occurrence of a plasma in the state of a physical cluster behind the shock wave front. A physical cluster is known to be stable if a medium where it has originated has certain values of p and T. If the medium disappears or its parameters change, the cluster spontaneously breaks down with the release of the stored energy U_0 [14]. During electronic detonation, the high pressure of the electron gas, which initiates the breakdown of solid dielectrics in the vicinity of the discharge channel, may be a factor responsible for the breakdown of the cluster.

For a dense plasma with a degree of ionization $x_e \ge 0.5$, the equations of state in the averaged sphere model can be approximated by formulas [12, 15][2]

$$\frac{U}{NkT} \cong -0.8\Gamma, \quad (9)$$

$$\frac{U}{NkT} = \alpha\Gamma + b\Gamma^{2/3} + c\Gamma^2. \quad (10)$$

Here, U is the internal energy,

$$a = 0.4497\exp(-0.5471/\theta) - 0.4434,$$

$$b = -1.502\exp(-0.5471/\theta) - 1.0478,$$

$$c = 0.4816\exp(-0.16/\theta) + 0.2362.$$

In formulas (9) and (10), Γ and θ are the parameters of nonideality for the ion and electron subsystem, respectively. They are given by

$$\Gamma = \frac{z_1 z_2 e^2}{kTd} > 1, \quad \theta_e = \frac{kT}{E_f} < 1, \quad (11)$$

where $d = (0.75\pi n_e)^{1/3}$ and E_f is the Fermi energy.

From (9)–(11), one comes to

$$1.25a(\theta) + 1.25\Gamma^{1/2}b(\theta) + 1.25\Gamma c(\theta) = 1. \quad (12)$$

If the functions $n_i(D_{eD})$ and $n_e(D_{eD})$ are known, a solution to Eq. (12) allows one to estimate the temperature $T(D_{eD})$ of the nonideal plasma and then calculate the parameters of nonideality Γ and θ. Then from the values of Γ, T, and n_e, the energy U stored in the plasma can be determined with Eq. (9). The conductivity of the plasma is calculated by the well-known Spitzer–Harm formula where the Coulomb logarithm at $\Gamma > 1$ equals [12]

$$\ln\Lambda = \ln\left(\frac{3\sqrt{2}}{\Gamma}\right). \quad (13)$$

RESULTS AND DISCUSSION

As follows from the aforesaid, to determine the real parameters of a material in an electronic detonation wave, one should first find the experimental dependence of the discharge channel velocity v_{dis} and the detonation velocity D on the breakdown voltage U_{br}. The dependence $v_{dis}(U_{br})$ was found to have a threshold value of the voltage pulse ($U_{br} \ge 150$ kV) starting from which the velocity of the anode discharge increases by

[1] If the detonation of solid explosives proceeds with $D \cong 8$–10 km/s, n is usually in the range $2.6 < n < 3.2$, depending on the material density [11].

[2] The formulas are presented in the Gaussian system of units, which is commonly accepted in the plasma physics.

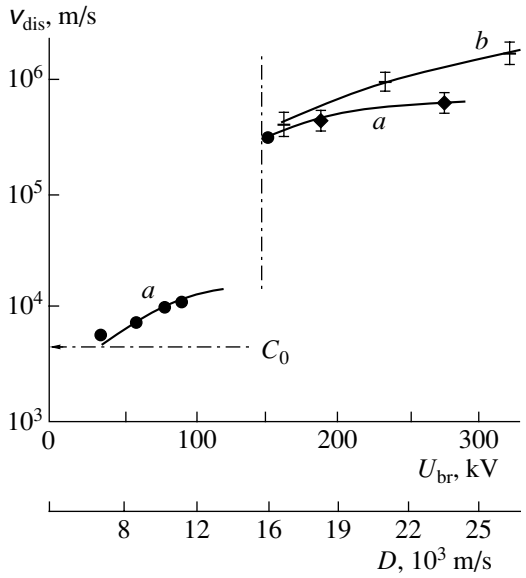

Fig. 1. Velocities of discharge channel propagation from the anode for the breakdown of (*a*) NaCl and (*b*) KCl due to rectangular voltage pulses [1]. C_0 is the velocity of sound.

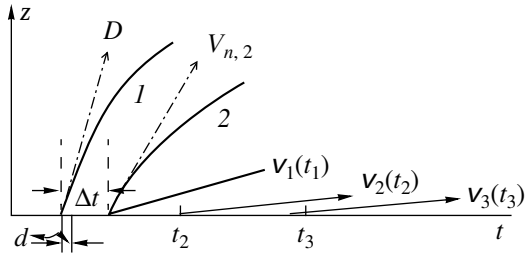

Fig. 2. Experimental data derived from chromatograms of the plasma pulsed discharge plasma flowing from the anode; (*1*) leading edge of the plasma glow and (*2*) trailing edge of the high-speed plasma glow; $v_i(t_i)$ are the time dependences of the low-speed plasma flow velocities (i = 1, 2, 3), Δt is the time of high-speed plasma flow and z = 0 is the coordinate of the dielectric surface.

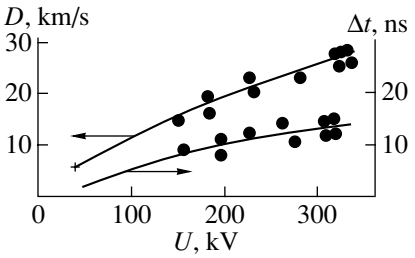

Fig. 3. Detonation velocity D and the flow duration Δt of high-speed plasma vs. amplitude of a high-voltage rectangular pulse U_{br}.

one to two orders of magnitude up to $v_{dis} \cong 1.5 \times 10^3$ km/s. This is observed most vividly when the breakdown is due to breakup by rectangular voltage pulses (Fig. 1).

The dependence $D(U_{br})$ can be determined with the well-known technique based on measuring the shock wave velocity v_{sh} in air. This shock wave is generated by detonation products when the detonation wave reaches the interface, where $v_{sh} \cong D$ [16]. It was found experimentally that the plasma flows out of the nanosecond discharge channel with different velocities (Fig. 2). This suggests a considerable difference in the plasma thermodynamic properties [17].

It is natural to assume that the high-speed plasma flowing out of the leading part of the discharge channel lies in the energy release zone. It has been found that both the velocity $v_{sh} \cong D$ and the outflow time Δt vary with voltage pulse amplitude (Fig. 3). A comparison of these dependences with the values of D calculated from the experimental Hugoniot adiabats show that the velocities of electronic detonation are higher than those expected from the shock wave experiments (see above) at $U_{br} > 150$ kV. Moreover, the anomalous increase in the velocities v_{dis} is observed in this very range of voltages and detonation velocities (Fig. 1). Obviously, this effect is related with a change in the energy release mechanism in the electronic detonation wave. This effect is also in agreement with the above-mentioned transition from the Joulean to cluster mechanism of energy release. In the detonation approximation, a difference between the velocities v_{dis} and D at U_{br} = const attests that an oblique shock appears.

From the experimental dependences $D(U_{br})$, one can compare the parameters of an electronic detonation wave in the detonation approximation and in the case of the cluster mechanism of energy release. In the first case, for velocities up to D = 26 km/s, the calculation is carried out by formulas (2)–(8). In the second case, for velocities $D < 14$ km/s, expressions (9)–(13) are used. Recall that if relationship (1) is fulfilled, the parameters of the detonation wave (p_D, U_D, T_D, Q_D) define only conditions for sustaining the shock wave with a velocity D and are not related to a specific mechanism of releasing the energy Q_D.

To apply formulas (9)–(13), the effective ion charge numbers z should be determined or specified. To this end, the specific heat Q in the detonation wave was calculated using the technique of determining a change in the enthalpy for a known composition of detonation products [16]. For the electronic detonation process, we applied sequential ionization schemes and, hence, the values of Q such that they sustain a detonation wave with velocities $D > 1.4 \times 10^4$ m/s [1]. To the values of Q desired, there corresponded ionization schemes with different charge numbers, which allowed us to choose appropriate values of z_{eff} (Fig. 4).

A comparison of these two approaches to estimating the parameters of an electronic detonation wave at velocities $D > 14$ km/s is shown in Figs. 5–7. The dependences obtained agree qualitatively; quantitatively, they differ 1.4–1.6 times. The parameter of non-

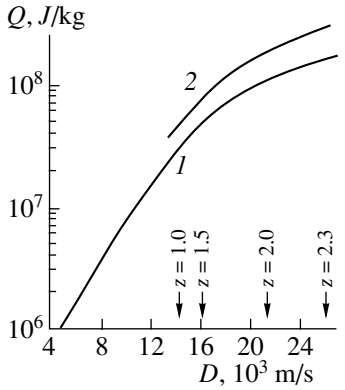

Fig. 4. Relationship between the specific heat Q, detonation velocity D, and effective charge numbers z in a physical cluster; (*1*) calculation by (7) and (*2*) calculation by (9).

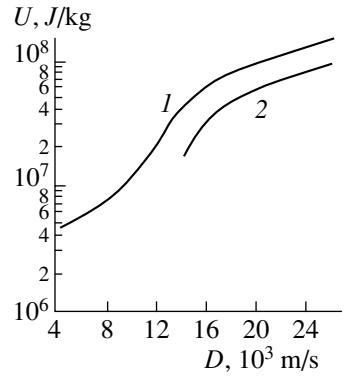

Fig. 5. Internal energy of the material in the electronic detonation wave; (*1*) calculation by (7) and (*2*) calculation by (9).

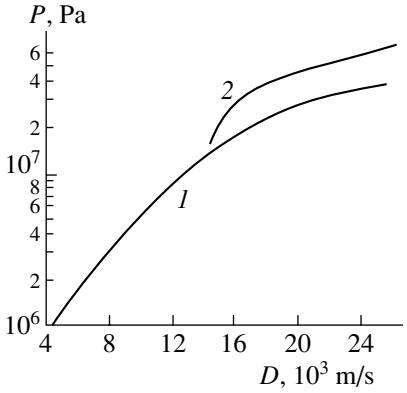

Fig. 6. Pressure in the electronic detonation wave; (*1*) calculation (6) and (*2*) pressure of the degenerate electron Fermi gas.

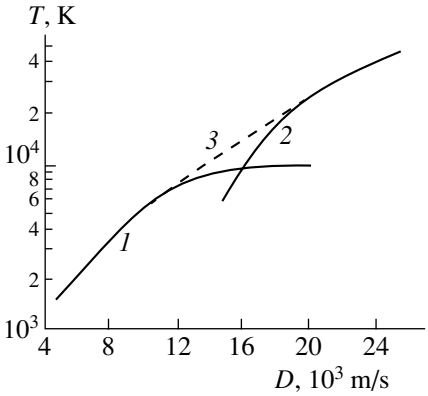

Fig. 7. Temperature of the material in the electronic detonation wave; (*1*) calculation by (8), (*2*) calculation by (10), and (*3*) envelope.

ideality for a dense plasma with $\delta = 1.56$ lie in the ranges $10 < \Gamma < 15$ and $0.1 < \theta < 0.5$ for $14 < D < 26$ km/s. According to the classification in [12], the electronic subsystem in the plasma is degenerate, while the ionic subsystem is classical. Both subsystems are characterized by strong interaction. The ultimate conclusion on this issue can be made when z_{eff} and, accordingly, the electron density n_e are determined experimentally.

Nevertheless, there is good reason to believe that the detonation rules following from the Chapman–Jouguet approximation combined with the aforementioned technique for quantitatively determining the dependence $Q(D)$ allow the estimation of the electronic detonation parameters in the entire range of detonation velocities and for power densities $10^{13} \le q \le 10^{16}$ W/m^2.

Next, it was found that a high-speed plasma flowing out of the discharge channel to air forms a spherical sector with an angle $\alpha \cong \pi/2$. In this case, the shock wave produces a spherical surface $S_1 = 1.81R^2(t)$. The surface propagation $R(t)$ follows the point explosion law [19][3]

$$R(t) = \left(\frac{Q_0}{\alpha\rho_0}\right)^{1/5} t^{2/5}. \quad (14)$$

Here, Q_0 is the energy of a point explosion generating a spherical shock wave with a surface $S_0 = 4\pi R^2(t)$, $\rho_0 = 1.29$ kg/m^3 is the air density, and $\alpha = 0.5$ at $\gamma = 5/3$.

On the other hand, the length of the energy release zone is $l = u\Delta t$ and the material mass in the zone is given by

$$m = \pi r_0^2 u \Delta t \rho_0 \delta_{eD}, \quad (15)$$

where r_0 is the U_{br}-dependent radius of the discharge channel [20], $u = D/(n + 1)$ is the mass velocity of the material in the zone, and Δt is the outflow time (Fig. 3).

[3] This law was found earlier and used in [18].

The values of Q calculated from

$$Q = \frac{S_1}{S_0}\frac{R^5(t)\alpha\rho_0}{mt^2} \quad (16)$$

(Fig. 4) differ from those calculated by (7) by a factor of 1.3–1.8. Taking into account the errors in determining Δt and r_0, the agreement between these values can be considered satisfactory.

According to the detonation theory, the material in the energy release zone is of higher density and pressure and releases more thermal energy in comparison with the material outside the zone. As was mentioned above, for a pulsed discharge from the anode, this shows up in the different velocities with which the plasma flows out of the leading and channel parts of the discharge. The properties of a channel plasma have been discussed most fully as applied to the electropulsed destruction of materials [21]. It has been found that upon the pulsed breakdown of solid dielectrics, the maximum density and specific energy of the channel plasma are observed at the instant t_0 the channel crosses the discharge gap. For KCl at $t = t_0 + 10^{-7}$ s, these parameters equal p = 1.0–1.5 GPa, $\rho \cong 0.1$ g/m^3, and $Q \le 1.3 \times 10^6$ J/kg (the particle concentration in the plasma is $n \cong 3 \times 10^{21}$ cm^{-3}).

At this instant, the plasma temperature $T \cong 5000$ K is minimum. At the maximum of the through current in the first period, the temperature increases up to (15–18) $\times$ 10^3 K, which is accompanied by a further decrease in the pressure ($p \cong 0.5$ GPa) and plasma density ($\rho \cong 10^{-4}$ g/cm^3). The parameter of nonideality Γ drops from 2.0–2.5 to 0.2–0.5.

If we compare these values with the associated values for electronic detonation in KCl (see above), we can easily check that the parameters in the energy release zone do exceed those in the channel part.

CONCLUSION

A technique for estimating expected and actual parameters of electronic detonation is offered. It is applied to the process taking place at supersonic velocities of a pulsed discharge channel propagating from the anode in solid dielectrics. The technique can be helpful in studying electronic detonation in solid dielectrics with different properties and in various applications.

REFERENCES

1. Yu. N. Vershinin, *Electron Heat and Detonation Processes upon Electrical Breakdown of Solid Dielectrics* (Ural. Otd. Ross. Akad. Nauk, Yekaterinburg, 2000).
2. Yu. N. Vershinini and D. S. Il'ichev, Dokl. Akad. Nauk **365**, 617 (1999) [Dokl. Phys. **44**, 220 (1999)].
3. A. M. Prokhorov, V. I. Konov, I. Ursu, *et al.*, *Interaction of Laser Radiation with Metals* (Nauka, Moscow, 1988).
4. V. J. Tarzhanov, in *Proceedings of the V Zababachin Scientific Talks, 1998.*
5. V. I. Tarzhanov, Fiz. Goreniya Vzryva, No. 5, 81 (1985).
6. Ya. B. Zel'dovich and A. S. Kompaneets, *Theory of Detonation* (Gostekhizdat, Moscow, 1953).
7. *Properties of Condensed Matter under High Pressures and Temperatures*, Ed. by R. F. Trunin (VNIITF, Arzamas, 1992).
8. *Solids under Pressure*, Ed. by W. Paul and D. M. Warshauer (McGraw-Hill, New York, 1963), p. 495.
9. S. B. Kormer, M. V. Sinitsyn, V. D. Urlin, *et al.*, Zh. Éksp. Teor. Fiz. **48**, 1033 (1965) [Sov. Phys. JETP **21**, 689 (1965)].
10. V. E. Fortov and Yu. G. Krasnikov, Zh. Éksp. Teor. Fiz. **59**, 1645 (1970) [Sov. Phys. JETP **32**, 897 (1971)].
11. L. V. Al'tshuller, Usp. Fiz. Nauk **85**, 179 (1965) [Sov. Phys. Usp. **8**, 140 (1965)].
12. V. E. Fortov and I. T. Yakubov, *Physics of Nonideal Plasma* (Inst. Khim. Fiz., Chernogolovka, 1984; Hemisphere, New York, 1990).
13. Yu. N. Vershinin and D. S. Il'ichev, Dokl. Akad. Nauk **374**, 187 (2000) [Dokl. Phys. **45**, 454 (2000)].
14. P. P. Kulik, G. É. Norman, and L. S. Polak, Khim. Vys. Énerg. **10**, 203 (1976).
15. S. Tanaka, S. Metake, X. Z. Yan, *et al.*, Phys. Rev. A **32**, 1779 (1985).
16. F. A. Baum, K. P. Stanyukovich, and B. I. Shekhter, *Physics of Explosion* (Fizmatgiz, Moscow, 1959).
17. Y. N. Vershinin, A. A. Podrezov, and S. V. Barahvostov, in *Proceedings of the International Conference on the Properties and Applications of Dielectric Materials, Tokyo, 1991*, p. 1189.
18. M. P. Tonkonogov, Yu. D. Il'yushenkov, and I. N. Pereletov, Zh. Tekh. Fiz. **52**, 260 (1982) [Sov. Phys. Tech. Phys. **27**, 170 (1982)].
19. V. P. Korobeĭnikov, *Problems in the Theory of Point Explosion* (Nauka, Moscow, 1985).
20. Yu. N. Vershinin, Dokl. Akad. Nauk **347**, 614 (1996) [Phys. Dokl. **41**, 157 (1996)].
21. B. V. Semkin, A. F. Usov, and V. I. Kurets, *Foundations of Electrical Pulsed Destruction of Materials* (Nauka, St. Petersburg, 1995).

Translated by M. Astrov

Technical Physics, Vol. 47, No. 12, 2002, pp. 1529–1532. Translated from Zhurnal Tekhnicheskoĭ Fiziki, Vol. 72, No. 12, 2002, pp. 44–48.

SOLIDS

Fractal Cluster Model of Spalling

G. G. Savenkov

Krasnoznamenets State Unitary Research and Production Enterprise, St. Petersburg, 195043 Russia

e-mail: sog@hotbox.ru

Received March 1, 2002; in final form, May 13, 2002

Abstract—A model of spalling that takes into account the shape and defect structure of the fracture surface is developed. Spalling strength and its relation to various characteristics of the spalling process are studied theoretically and experimentally. The model proposed eliminates a number of disadvantages encountered in the conventional determination of spalling strength. © *2002 MAIK "Nauka/Interperiodica".*

In spite of the recent progress in generally understanding the spalling process, adequate models of phenomena attendant on spalling are still lacking.

Moreover, there are a number of problems in the experimental determination of spalling strength. This parameter is measured by various methods [1] because of the complex wave pattern in shock-stressed specimens, and the relation between the spalling strength and the defect concentration in the specimens is as yet little understood. Moreover, the defect size distribution vs. applied load and the type of the fracture surface in relation to the spalling strength still remain unclear [2, 3].

In this work, by analogy with [3], we proceed from the concept that material particles in loading and unloading waves move as a set of mesofluxes with different velocities relative to each other. This concept was repeatedly confirmed experimentally by Meshcheryakov and colleagues.

Along with the mean particle velocity $\langle u \rangle$, stressing in this case is characterized by two more parameters, namely, the deviation of the velocity from the mean velocity (the "pulsation" velocity $\Delta u = u - \langle u \rangle$ [4]) in terms of turbulent hydrodynamics and the difference $\delta u = u_{\max} - u_{\min}$ between the maximum and minimum velocities of the mesofluxes. Methods of their determination are given in detail elsewhere [5].

MODEL OF THE PROCESS

Let us consider the stages of spalling. Aptukov [6] and several other researchers discuss two stages of spalling, while Naĭmark and Belyaev [7] divide the process into three. It was shown [3, 8] that the spalling crack contour and the spalling surface are fractal clusters consisting of micro- and mesodefects of the first and second levels (according to the terminology of Panin and coworkers [9]). Therefore, we assume that a spalling crack is formed in four stages: the first stage, the explosive appearance of multiple micro- and mesocracks; the second stage, coalescence of individual cracks into groups (finite-size clusters); the third stage, coalescence of finite clusters into an infinite percolation (fractal) cluster; and the fourth stage, disintegration of the specimen into parts (the formation of a spalling "plate"). The fourth stage is not obligatory; therefore, we may consider only the first three.

Based on the model proposed, we will determine the spalling strength of materials. In the general form, the spalling strength σ_s vs. uniaxial strain ε_s dependence is given by [10] $\sigma_s = K\varepsilon_s + (4/3)\tau$ or, in terms of strain rate,

$$\sigma_s = K\langle u \rangle / c_s + (4/3)(\tau_0 + \mu_\tau \dot{\gamma}), \qquad (1)$$

where K is the bulk modulus, τ is the shear strength, τ_0 is the static component of τ, $\dot{\gamma}$ is the shear strain rate, c_s is the longitudinal plastic wave velocity, and μ_τ is the shear dynamic viscosity.

Since the first stage is of explosive type and, hence, can hardly be analyzed numerically, it is obvious that basic processes occur at the second and third stages. Then, we can assume that the specimen is a cracked medium with a shear strain rate determined from the relationship [11]

$$\dot{\gamma} = (1/G)d\tau/dt + \Delta h \int_0^\infty V(l)n(l, t)dl. \qquad (2)$$

Here, G is the shear modulus, Δh is the displacement of the crack edges, $n(l, t)dl$ is the number of cracks per unit surface, $V(l)$ is the crack velocity, and t is time. Based on the theory of percolation clusters, we will determine the crack size distribution function $n(l, t)$. According to [12], the probability p of the existence of a finite cluster of n nodes (in our case, cracks or defects) that is normalized in the form

$$\sum_{n=1}^{\infty} p_n = 1$$

has the asymptotics

$$p_n = p(n) = An^{1-m}\exp(-an^{\lambda}) \quad (3)$$

for large n, where A is the normalizing factor, $a = a(p, d_1)$, $m = m(p, d_1)$, and d_1 is the dimension of the topological space (in our case, $d_1 = 2$).

The exponents m and λ at $d_1 = 2$ have the following values:

$$\begin{aligned} m(0 < p < p_*) &= 1, \quad m(p_*) = 2, \\ \lambda(0 < p < p_*) &= 1, \quad \lambda(p_*) = 0, \end{aligned} \quad (4)$$

where $p_* \geq 1/7$ is the critical probability of the formation of an infinite cluster [13]. The coefficients $a(p = 0, 15 > 1/7) \cong 0.36$ and $a(p_*) = 0$ at $d_1 = 2$ [12].

The size of a finite cluster consisting of n defects with the mean defect length l_0 is $L = nl_0$. By definition, the probability density can be written as $p_n dn = f(L)dL$, where $f(L)$ is the probability density of the finite cluster sizes falling into the range $(L, L + dL)$. Taking (3) into account, the probability density is given by

$$f(L) = (A/l_0)(L/l_0)^{1-m}\exp(-a(L/l_0)^{\lambda}). \quad (5)$$

By definition, the fraction of clusters with a size larger than L is

$$Z(L) = \int_L^{\infty} f(L)dL.$$

With (5) and (4), we have

$$Z(L) = (A/l_0)\exp(-a(L/l_0))$$

or $L = nx$ (x is the characteristic defect size). Then,

$$n(x) = (A/l_0)\exp[-an(x/l_0)].$$

This agrees with the Rozin–Ramler law [14]

$$n(x) = n_0\exp[-b(x/l_0)] \quad (6)$$

with the distribution function

$$F(x) = 1 - n_0\exp[-b(x/l_0)],$$

which is a particular case of the Weibull distribution function [14].

It should be noted that dependence (6) is a theoretical estimate of the distribution n, unlike the defect size distribution

$$N = N_0\exp(-R/R_1)$$

experimentally found in [15] upon shock stressing. Here, N is the number of defects with a size larger than R, N_0 is the total number of defects, and R_1 is the mathematical expectation of the crack length distribution function.

Since cracks responsible for the shear strain in Eq. (2) are shear cracks, the relationships [16]

$$x = \Delta u \Delta t \quad \text{and} \quad l_0 = \delta u \Delta t$$

are valid.

Substituting these dependences into (6) and taking into account that the number of spalling cracks are equal to that of shear cracks yields

$$n(x) = 0.5n_0\exp[-b(\Delta u/\delta u)], \quad (7)$$

where n_0 is the number of shear cracks with the mean length l_0 per unit volume.

It is reasonable to assume that the crack velocity upon shock-wave loading is limiting and constant; i.e., $V \cong 0.7c_R$ (c_R is the velocity of the Rayleigh surface waves) [16]. This assumption is also supported by experimental data [17]. For Δh, as in [11], we choose the critical crack opening δ_{cr} in the Leonov–Panasyuk model [18]. Substituting (7) and the value of V into (2), putting $c_R = \delta_{cr}/\Delta t$, and taking into account that $\mu_\tau = \rho\Delta u l_0 = \rho\Delta u \delta u \Delta t$, we obtain Eq. (1) in the form

$$\sigma_s \cong K\langle u\rangle/c_s + (4/3)\tau_0 - 0.5n_0\rho\Delta u\delta u\delta_{cr}. \quad (8)$$

Here, the first term in (2) is neglected because of its smallness and ρ is the target density. As compared to the conventional relationships for spalling strength, expression (8) has both advantages and disadvantages. First, it removes the question of which sound velocity (longitudinal c_1 or volume c_0) should be substituted into the standard formula [1]

$$\sigma_s = \varphi\rho u_s + \Delta\sigma, \quad (9)$$

where $\varphi = 0.5$ or $\varphi = 1/(1 + c_1/c_0)$ and $u_s = V_{max} - V_{min}$ is the difference between the maximum and minimum velocities of the free surface.

Second, it is unnecessary to determine $\Delta\sigma$, whose form is not quite clear, and, finally, Eq. (9) itself becomes unnecessary for the calculation of spalling strength.

The disadvantages of Eq. (8) are the complexity of determining c_s under nonsteady loading and difficulties in calculating n_0 and δ_{cr}. The difficulty in determining the latter characteristic (δ_{cr}) is related mainly to the fact that spalling is a highly dynamic process and also to insufficiently advanced techniques for the determination of the crack resistance parameters under high-rate stressing.

Now, consider the effect of clustering on spalling from other points of view. As was already mentioned, finite clusters join together at the third spalling stage to form a fractal cluster. In this case [19], the density of particles (defects) making up this cluster is a function of the type

$$n_1 \cong \beta H^{D-d_1}, \quad (10)$$

where β is the normalizing factor, H is the spalling zone width (for an axisymmetric shock, $H \cong B$, where B is the spalling zone diameter), and D is the fractal (Hausdorff–Bezikovich) dimension [19].

In our case, n_1 specifies n_0 in the relationships derived above. Then, with (10) taken into account, Eq. (8) for spalling strength takes the form

$$\sigma_s = K\langle u\rangle/c_s + (4/3)\tau_0 - 0.5\beta\rho\Delta u\delta u\delta_{cr}H^{D-d_1}, \quad (11)$$

which shows that spalling strength increases with fractal dimension (by definition, $D < d_1$). This conclusion totally coincides with the conclusion drawn in [8] and physically means that σ_s increases with the spalling surface.

Now let us discuss the effect of dynamic stressing parameters on the fractal dimension. To this end, we take advantage of the fact that the propagation of an elastoplastic front in a structurally nonuniform medium is a random process. It is described by the Fokker–Planck equation [3], which determines the nonequilibrium practicle velocity distribution function $f(x, V, t)$ in a medium:

$$\partial f/\partial t + V\partial f/\partial x + (F_1/m)\partial f/\partial t$$
$$= -\partial(Q_1V)/\partial V + 0.5\partial^2(Q_2 f)/\partial V^2,$$

where F_1 is the applied force, m is the particle mass, Q_1 is the drift coefficient, and Q_2 is the diffusion (percolation) coefficient.

The solution to this equation corresponds to self-organization in a system upon the interaction of two transfer phenomena (drift and diffusion). The Fokker–Planck equation is known to have two types of solution: steady-state solutions, where the argument is independent of time, and non-steady-state solutions, which can be found in the self-similar mode alone [20]. In the latter case, a function of two arguments x and t (in the unidimensional case) is replaced by a function of one variable, $y = x/a(t)$:

$$f(x, t) = l_0^\alpha\varphi(y),$$

where the functions $l_0(t)$ and $\varphi(y)$ and the exponent α are to be determined. Note that the function $\varphi(y)$ can be found only if the scaling properties of the drift force (Q_1mf) and the diffusion component Q_2f are specified [20].

It was also shown in [20] that, for a stochastic system to be self-similar, the following relationship should be fulfilled:

$$l_0 = [\mu(1-\xi)]^{1/(1-\xi)}t^{1/(1-\xi)},$$

where $\mu = \rho\Delta u\Delta h$ is the dynamic viscosity [4], Δh is the mesoflux width, $\xi = 1 - D$, and l_0 is the characteristic scale specifying the characteristic value of the variable x.

Substituting μ and ξ into this relationship and taking into account that $l_0 = (\delta u)t$, we have

$$(\delta u)t = [(\rho\Delta u\Delta h)D]^{1/D}t^{1/D}. \quad (12)$$

Table 1. Testing results for 12Kh18N10T steel

Shock speed, m/s	Δu, m/s	δu, m/s	U_s, m/s
108	0	33	–
181	13.0	48	97
216	24.0	81	66
248.5	28.0	64	81
285	30.0	12	102
308	38.0	14	100
344.5	43.0	30	100
447	–	171	84.5

Table 2. Testing results for KhN75VMYu alloy

Shock speed, m/s	Δu, m/s	δu, m/s	U_s, m/s
217.8	24.0	13.3	164.7
230.1	25.8	18.3	168.0
251.8	25.8	19.2	160.1
276.2	25.2	1.1	174.9
299.2	25.3	16.6	153.3
343.0	27.1	8.875	177.3
350.4	25.4	4.9	189.5
363.5	29.9	6.6	183
386.6	41	13.7	176.2
298.2	49.5	2.18	193.7

Note: Values are averaged over two or three tests.

Thus, the fractal dimension is determined by δu, t, and Δu. Hence, the spalling strength of materials depends on the velocity characteristics $\langle u\rangle$, Δu, δu, and the time to failure t (see (11)).

From Eq. (12), it follows that the fractal dimension increases with decreasing δu (thus favoring an increase in the spalling strength with a decrease in the fractal dimension) and increases with decreasing Δu. As a result, the function $\sigma_s = \Phi(\Delta u)$ is uncertain.

RESULTS AND DISCUSSION

Planar 7- to 35-mm-thick 12Kh18N10T steel specimens (Table 1) and 5.6-mm-thick KhN75VMYu nickel-based alloy specimens were shock-stressed under conditions of uniaxial deformation at striker speeds of 100–500 m/s. The striker thickness in all cases was 1.98–2.0 mm. The testing results are given in Figs. 1 (12Kh18N10T) and 2 (KhN75VMYu). Figure 3 shows the results of calculating the fractal dimension of the spalling crack contour (the technique for its determination is given in [3]).

From the results given in the tables and figures, it follows that the spalling model proposed and the relationships obtained do not contradict the experimental

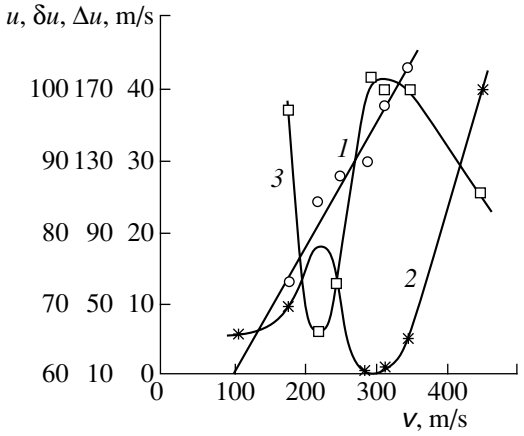

Fig. 1. Effect of the shock speed on the parameters (*1*) Δu, (*2*) δu, and (*3*) u_s for 12Kh18N10T steel.

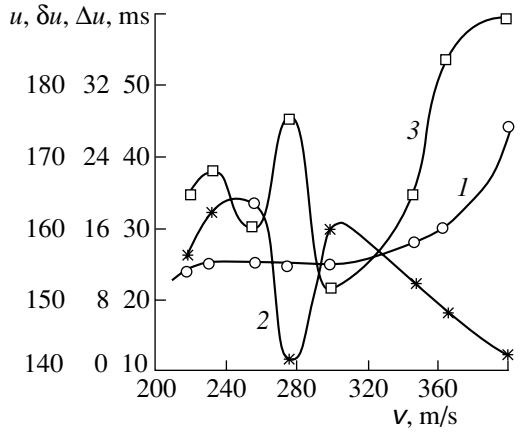

Fig. 2. The same as in Fig. 1 for KhN75VMYu alloy.

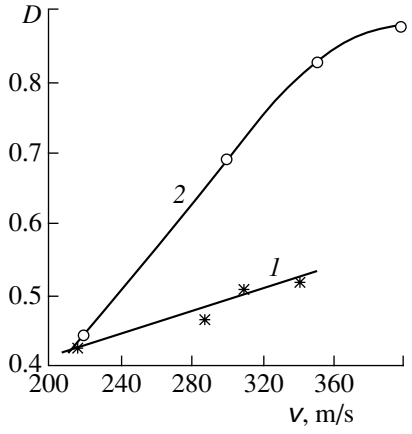

Fig. 3. Effect of the shock speed on the fractal dimension for (*1*) 12Kh18N10T and (*2*) KhN75VMYu.

data. For example, the spalling rate (which defines the spalling strength) is clearly seen to be in antiphase with u_s and δu, whereas there is no clear relationship between u_s and Δu. Figure 3 also illustrates the increase in the fractal dimension with shock speed, which in general supports the conclusion that spalling strength grows with fractal dimension.

Thus, we can conclude that our model adequately fits the actual process of spalling.

REFERENCES

1. G. I. Kanel', Prikl. Mekh. Tekh. Fiz., No. 2, 194 (2001).
2. S. A. Novikov, Fiz. Goreniya Vzryva, No. 6, 77 (1985).
3. B. K. Barakhtin, Yu. I. Meshcheryakov, and G. G. Savenkov, Zh. Tekh. Fiz. **68** (10), 43 (1998) [Tech. Phys. **43**, 1175 (1998)].
4. L. D. Landau and E. M. Lifshitz, *Course of Theoretical Physics*, Vol. 6: *Fluid Mechanics* (Pergamon, New York, 1987; Nauka, Moscow, 1988).
5. Yu. I. Meshcheryakov and A. K. Divakov, Preprint No. 25, LfiMash (Leningrad, 1989).
6. V. I. Aptukov, Fiz. Goreniya Vzryva, No. 2, 122 (1985).
7. O. B. Naĭmark and V. V. Belyaev, Fiz. Goreniya Vzryva, No. 4, 115 (1989).
8. G. G. Savenkov, B. K. Barakhtin, and Yu. I. Meshcheryakov, in *Proceedings of the V International Workshop "Modern Problems of Strength," Novgorod, 2001*, Vol. 2, p. 43.
9. V. E. Panin, Izv. Akad. Nauk, Mekh. Tverd. Tela, No. 5, 88 (1999).
10. G. V. Stepanov, *Elasto-Plastic Deformation and Destruction of Materials under Impulsive Loading* (Naukova Dumka, Kiev, 1991).
11. O. V. Kovalenko and V. K. Sirotkin, Prikl. Mekh. Tekh. Fiz., No. 4, 138 (1985).
12. D. Staufer, Phys. Rep. **54** (1), 23 (1979).
13. S. A. Molchanov, V. F. Pisarenko, and A. Ya. Reznikova, in *Computational Seismology*, Vol. 19: *Mathematical Methods in Seismology and Geodynamics* (Nauka, Moscow, 1986), pp. 3–8.
14. V. M. Kuznetsov, *Mathematical Models of Explosion* (Nauka, Novosibirsk, 1977).
15. *Shock Waves and High-Strain-Rate Phenomena in Metals*, Ed. by M. A. Meyers and L. E. Murr (Plenum, New York, 1981; Nauka, Moscow, 1984).
16. V. M. Finkel', *Physics of Destruction* (Metallurgiya, Moscow, 1970).
17. Yu. I. Meshcheryakov and G. G. Savenkov, Prikl. Mekh. Tekh. Fiz., No. 3, 138 (1993).
18. V. V. Panasyuk, *Limit Equilibrium of Brittle Solids with Cracks* (Naukova Dumka, Kiev, 1968).
19. A. I. Olemskoĭ and A. Ya. Flat, Usp. Fiz. Nauk **163** (12), 1 (1993) [Phys. Usp. **36**, 1087 (1993)].
20. A. I. Olemskoĭ, Usp. Fiz. Nauk **168** (3), 287 (1998) [Phys. Usp. **41**, 269 (1998)].

Translated by K. Shakhlevich

Technical Physics, Vol. 47, No. 12, 2002, pp. 1533–1537. Translated from Zhurnal Tekhnicheskoĭ Fiziki, Vol. 72, No. 12, 2002, pp. 49–53.

SOLIDS

Mechanical, Structural, and Spectroscopic Properties of C_{70} Fullerite Phases Produced under High Pressure and Shear

V. D. Blank[1,2], K. V. Gogolinsky[2], V. N. Denisov[1], V. A. Ivdenko[1], B. N. Mavrin[1], N. R. Serebryanaya[1], and S. N. Sulyanov[2]

[1] *Institute of Spectroscopy, Russian Academy of Sciences, Troitsk, Moscow oblast, 142190 Russia*
e-mail: mavrin@isan.troitsk.ru

[2] *Technological Institute of Superhard and Novel Carbon Materials, Troitsk, Moscow oblast, 142190 Russia*

Received March 18, 2002; in final form, June 11, 2002

Abstract—X-ray diffraction patterns, Raman spectra, and the hardness of C_{70} fullerite subjected to a high pressure with shear are investigated. It is shown that these conditions favor the phase transformation of molecular fullerite into the hard amorphous phase. The hardness of a specimen removed from a diamond anvil cell loaded up to 26 GPa under shear deformation applied is found to be equal to 30 GPa. © *2002 MAIK "Nauka/Interperiodica"*.

INTRODUCTION

The superhard high-pressure phases of C_{60} fullerite that are stable under normal conditions were obtained by us earlier [1–5]. The hardness [6] and elasticity [7] of some of them were higher than those of diamond. In searching for novel superhard materials, we investigated C_{70} fullerite at high pressures P and high temperatures T [8]. In [8], several polymorphic modifications were reported and their nonequilibrium P–T phase diagram was presented. Also, we succeeded in producing superhard phases. The superhard phases of both fullerites were obtained in both the crystalline and amorphous states [3, 8], the amorphous phases offering the higher hardness [4]. The existence of the amorphous C_{60} and C_{70} phases obtained under high pressures and temperatures was confirmed in several studies [1–10] and is beyond question. The most widely debated issue has been structural and phase transformations in C_{70} at room temperature. In early studies, authors reported the observation of phase transitions at ≈1 [11], ≈2 [12], and ≈5.5 GPa [12]; however, their existence was not confirmed [13, 14] and subsequently attention was mainly focused on the transition to the amorphous carbon phase. Using X-ray diffraction methods, the authors of [15] observed irreversible amorphization under compression stresses exceeding 18 GPa. At the same time, Raman and photoluminescence spectra [11, 16] indicate reversible amorphization arising upon releasing the pressure P = 31 GPa in high-pressure cells. These results were obtained under high pressures directly in diamond anvil cells. Recently, when investigating Raman spectra [13] and electrical and mechanical properties of C_{70} [14], it was found that the pressure-induced amorphous phase begins to form at a pressure of about 12 GPa and the specimen becomes completely amorphous in the range 18–20 GPa [11, 13, 14]. After relief, the amorphous phase reverts to the initial structure of fullerite if the pressure does not exceed 35 GPa [14]. At pressures higher than 35 GPa, the specimen remains amorphous even after the pressure has been removed. At pressures above 10 GPa, the electrical resistance of C_{70} decreases gradually by 4–5 orders of magnitude with a minimum at 20 GPa. As the pressure rises further, the resistance starts increasing and attains its initial value at $P > 35$ GPa [14]. The microhardness of the amorphous phase measured by nanoindentation after the pressure 48 GPa has been released was about 30 GPa [14].

In this paper, we studied X-ray diffraction patterns, Raman spectra, and mechanical properties of C_{70} fullerite at high pressures in a diamond cell where the specimen was under shear. The application of shear deformations narrows the hysteresis of structural phase transformations and homogenizes the phase [17]. We expected that, under shear, the transformation of molecular C_{70} fullerite into the hard phase will occur at lower pressures and that other phase transitions will appear. C_{70} fullerite of purity higher than 99% was produced in the Institute of Metallorganic Chemistry (Nizhni Novgorod).

X-RAY DIFFRACTION

Our investigations were performed in an X-ray diamond anvil cell (DAC) with shear under pressures of up to 50 GPa [18]. High shear strains in the specimen are produced by rotating one of the anvils with respect to the other about the axis of stressing. The rotation is performed using a thrust bearing.

In order to clarify the effect of shear strains on the phase transitions, a pressure-transfer medium was not applied. The nonhydrostatic character of the pressure due to the uniaxial stressing of the specimen in the DAC

1063-7842/02/4712-1533$22.00 © 2002 MAIK "Nauka/Interperiodica"

was compensated for in part by using a gasket that was plastic at high pressures. The surface of the immovable anvil was covered by a ruby powder to measure the pressure from the pressure-dependent shift of narrow luminescent R lines of ruby (R_1 = 694.2 nm and R_2 = 692.7 nm). The pressure distribution in the specimen was found from the luminescence spectra of ruby, whose particles (1–3 μm in size) were uniformly applied on the working surfaces of the anvils. The particle spacing was about 10 μm.

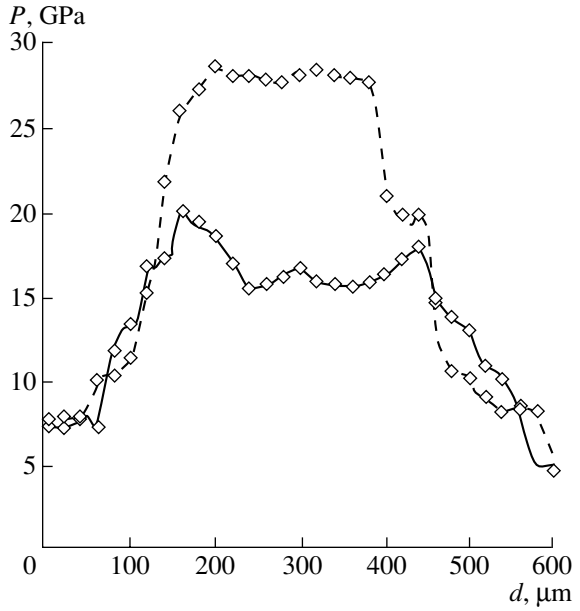

Fig. 1. Pressure distribution in the C_{70} specimen along the diameter (d) of the diamond anvil. Solid and dashed curves correspond to the situation before and after shear, respectively.

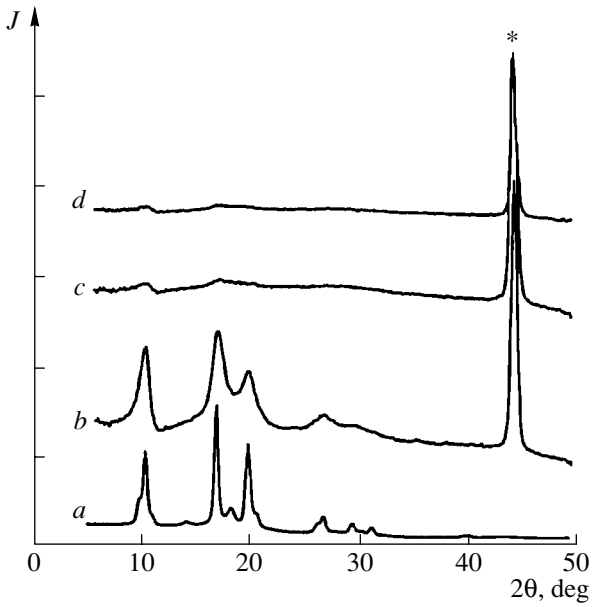

Fig. 2. Diffraction patterns from C_{70} removed from the diamond cell after the action of high pressure under shear: (*a*) initial C_{70}, (*b*) 12 (22 GPa after applying shear stress), (*c*) 16 (30), and (*d*) 26 (40) GPa. The asterisk indicates the diffraction peak from the steel gasket.

The pressure was evaluated from the more intense line R_1. The gasket with a hole where the specimen was mounted was attached to the movable anvil. The cell with the specimen was assembled, and stressing was accomplished by rotating the clumping nut of a spring-and-lever stressing system. The smoothness of stressing was provided by disk springs. Diffraction patterns from C_{70} pressed in the gasket were taken with a KARD-6 X-ray diffractometer (CuK_α) [19] before and after the action of pressure. The Raman spectra were recorded under pressure directly in the DAC.

Earlier, it was shown theoretically that plastic deformation is a universal mechanism for reducing internal strains [17] and favors polymorphic transformations. It turned out that when C_{60} is subjected to shear under pressure, phase transition to the superhard phase occurs. This transition is accompanied by the giant effect of self-multiplication, the pressure at the center increasing by more than twofold [20]. Figure 1 shows the pressure distribution in C_{70} along the diamond anvil diameter before and after shear. After the rotation of the anvil, the pressure at the center of the specimen grows up to 30 GPa. By analogy with C_{60}, this phenomenon testifies to the formation of the superhard phase in C_{70}.

The diffraction patterns from C_{70} taken out of the cell after it had been subjected to a high pressure and shear are shown in Fig. 2. Since the sizes of the specimen (≈250 μm in diameter and 20 μm in thickness) are very small, we did not remove it from the gasket; therefore, the diffraction patterns also contain reflections from the gasket made of stainless steel. For each of the diffraction patterns, the pressure value after shear is indicated. The diffraction pattern from initial C_{70} is presented at the bottom of Fig. 2. After being subjected to a pressure of 22 GPa, C_{70} recovers the crystalline state. However, the significant broadening of the diffraction lines and the enhanced background under the most intense diffraction peaks indicate the breaking of long-range order in the lattice and partial amorphization. Under the action of higher pressures, the amorphization increases; however, a small amount of the crystalline phase is present even after 40 GPa.

It should be noted that the diffraction patterns obtained from the C_{70} specimens after the action of a high pressure and shear at room temperature and those obtained after the action of a high pressure at a high temperature differ. High temperatures (T > 1000 K) lead to irreversible amorphization at a pressure of 5–13 GPa [8, 13]. In the range 2θ = 25°–30°, the diffraction patterns exhibit an intense halo. This fact suggests that this amorphous phase has a layered structure [8]. The diffraction patterns of the C_{70} amorphous phase that were obtained in the diamond cell at room temperature do not contain the halo.

RAMAN SCATTERING

Raman spectra were excited by laser radiation (514.5 nm) with a power density of about 50 W/cm^2. For each of the pressures, the Raman spectra of fullerite were first recorded in the absence of shear; next, at the same pressure after the application of shear deformation and, finally, after pressure relief and removing the specimen from the diamond cell.

The spectra of C_{70} fullerite in the diamond cell can be recorded only in the range above 1400 cm^{-1}, where, as is known [8], the Raman lines of fullerite are the most intense and the measurements are unaffected by the strong line 1333 cm^{-1} from the diamond anvils. In this range (at atmospheric pressure), the most intense line of fullerite is the high-frequency ("graphite-like") line at 1567 cm^{-1}. The doublet at 1444–1468 cm^{-1} is slightly weaker, and the faint line at 1511 cm^{-1} can also be seen.

The behavior of these lines with an increase in pressure was investigated previously [11, 13] (see Fig. 3). Wide bands at wave numbers above 1700 cm^{-1} belong to the Raman spectrum at combined frequencies of the diamond anvils. They can serve as an internal standard for estimating the changes in the fullerite spectrum intensity with increasing pressure. At a pressure of 7 GPa, only two bands at 1505 and 1592 cm^{-1} (under atmospheric pressure, at 1468 and 1567 cm^{-1}, respectively) can be seen. The low-frequency component of the doublet at 1444–1468 cm^{-1} and the line at 1511 cm^{-1} virtually disappear, thereby confirming the data in [11, 13]. At higher pressures (12 and 16 GPa), the intensities of the lines at 1502 and 1592 cm^{-1} decrease and the low-frequency component of the doublet (lying at 1460 and 1480 cm^{-1} for 12 and 16 GPa, respectively) becomes visible. In the range 7–16 GPa, the line at 1592 cm^{-1} almost does not shift but broadens with pressure. At 26 GPa, the spectrum intensity increases and only one very broad band that peaked at 1620 cm^{-1} is observed. The appearance of the single broad band at high pressures in the range above 1400 cm^{-1} instead of the line spectrum of initial fullerite was associated [11, 13, 14] with the crystalline-to-amorphous transition in fullerite.

The application of shear deformation considerably modify the spectra. The spectra shown in Fig. 4 were obtained at the same pressures and from the same specimens as in Fig. 3 but under shear. After the application of shear deformation, the pressure in the diamond cell increases (due to self-multiplication) from 7 (before shear) to 10 GPa (after shear), from 12 to 22 GPa, from 16 to 30 GPa, and from 26 to 40 GPa. At a pressure of 7 GPa with shear (Fig. 4), the form of the fullerite spectrum is close to that obtained at 16 GPa without shear. Indeed, in both cases, one can observe a broad line near 1590 cm^{-1} and a faint line at 1480 cm^{-1}. At a pressure of 12 GPa, the line at 1590 cm^{-1} splits into a doublet at 1538–1580 cm^{-1}. At 16 GPa, the spectrum intensity increases noticeably, and the high-frequency component of the doublet broadens and shifts toward 1610 cm^{-1}. At 26 GPa, this component becomes dominant (1640 cm^{-1}). Thus, shear deformation not only favors the phase transformations in fullerite but also uncover new features of the phase transition in the pressure range 12–16 GPa.

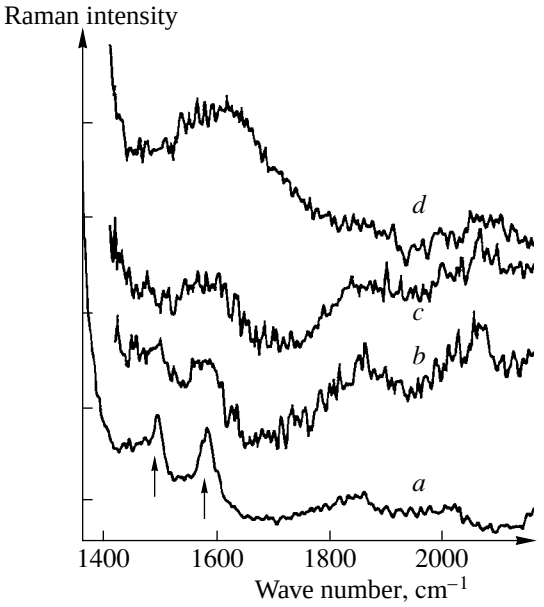

Fig. 3. Raman spectra from C_{70} fullerite in the diamond cell under high pressure before applying shear deformation: (*a*) 7, (*b*) 12, (*c*) 16, and (*d*) 26 GPa. Arrows (here and in Fig. 4) point to the bands discussed in the text.

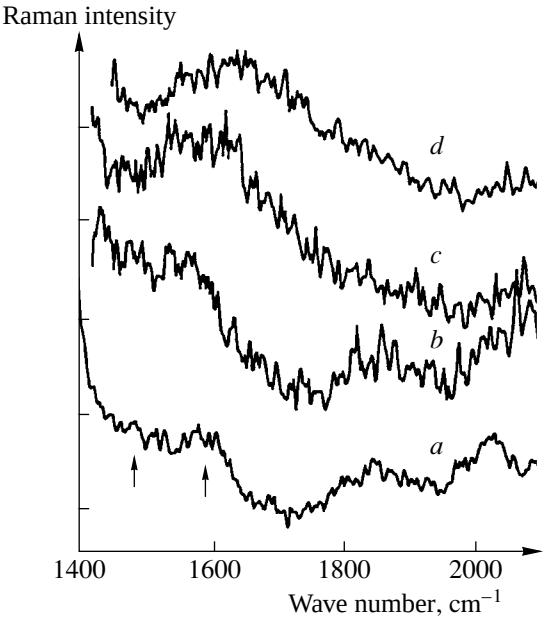

Fig. 4. Raman spectra from C_{70} fullerite in the diamond cell after applying shear deformation to the specimens the spectra from which are shown in Fig. 3. Due to self-multiplication, the pressure applied to the specimens increased from (*a*) 7 to 10, (*b*) 12 to 22, (*c*) 16 to 30, and (*d*) 26 to 40 GPa.

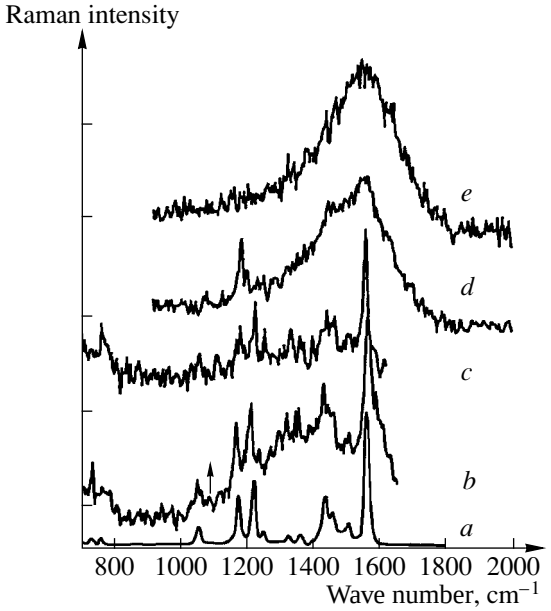

Fig. 5. Raman spectra from (*a*) initial C_{70} fullerite and (*b–e*) the specimens removed from the diamond cell the spectra from which are shown in Fig. 4: (*b*) after applying the pressure 7 GPa and shear deformation (10 GPa after shear), (*c*) 12 (22), (*d*) 16 (30), and (*e*) 26 (40) GPa.

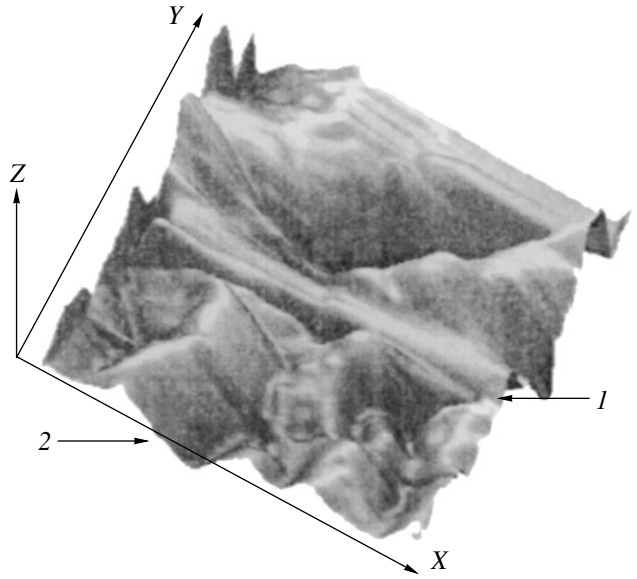

Fig. 6. Scratched region on the surface of the C_{70} specimen removed from the diamond cell and investigated by diffraction (Fig. 2d) and spectroscopic (Fig. 5e) methods. The sizes of the scratch are shown in the application. Arrow *1* points to the scratch, and arrow *2* shows a hexagonal crystallite.

After releasing the pressure and removing the specimen from the diamond cell, the Raman spectra could be recorded in the entire frequency range, since the contribution from the spectrum of the diamond anvils was absent. Figure 5 shows the spectra of fullerite removed from the cell after applying the pressure and shear. For pressures of 7 and 12 GPa (Figs. 5b, 5c), the spectrum is seen to consist of the same narrow bands as initial fullerite before applying the pressure (Fig. 5a). Therefore, one may assume that the molecular skeletons of C_{70} are still retained. However, after applying the pressure 7 GPa, the high-frequency line (1567 cm^{-1} in initial fullerite) shifts to 1569 cm^{-1}, while after applying 12 GPa, it shifts to 1557 cm^{-1}. At the same time, the other lines do not noticeably change their positions. The low-frequency shift of the "graphite-like" line by more than by 10 cm^{-1} may be due to the appearance of fourfold-coordinated carbon atoms (i.e., sp^3 bonds) in fullerite molecules. As is known, molecules of original fullerite have only threefold-coordinated carbon atoms. As a rule, sp^3 bonds appear in the process of polymerization (or dimerization) of molecules. Specifically, in the Raman spectra of fullerite containing mainly dimers $(C_{70})_2$, a new line near 1550 cm^{-1} was observed [21].

After applying 16 GPa, the line at 1557 cm^{-1} becomes very broad (with a width greater than 200 cm^{-1}), but a number of molecular fullerite lines still remain (Fig. 5d). However, after applying 26 GPa, only one broad intense line at 1558 cm^{-1} is seen (Fig. 5e). Therefore, it can be assumed that, after applying 16 GPa, the structure of the specimen is still two-phase, whereas after 26 GPa, the specimen is nearly monophase. The spectrum of this phase consists of one nearly symmetric line at 1558 cm^{-1}. In the case of carbon materials, such a spectrum was observed in diamond-like films [22], as well as in spatially polymerized C_{60} fullerite [3]. Both materials contain a significant amount of sp^3 bonds and offer good elastic properties. After applying the pressure 26 GPa, our specimen (unlike initial fullerite) was found to be stable against laser radiation and did not transform into amorphous carbon or graphite at a 20-fold increase in the laser radiation density.

HARDNESS MEASUREMENTS

The hardness was measured by the sclerometry method (scratching under constant indentation load), which was carried out with the specimens removed from the gasket after the application of high pressure and shear. The use of a nanosclerometer with a nanoindenter made of superhard C_{60} fullerite [4] (produced in the Technological Institute of Superhard and Novel Carbon Materials at a pressure of 13 GPa and a temperature of 1770 K) made it possible to quantitatively determine the hardness of the superhard materials. The hardness H is defined as $H = kP/b^2$, where k is the form factor of the indenter, P is the indentation load, and b is the scratch width. The shape and dimensions of the scratch were determined with a Nanoscan instrument based on the principles of atomic force microscopy [4]. The time of scratching was 2 s; the width and length of the scratch, 0.6 and 2.5 μm, respectively; and the maximum indentation load, 0.1 N. The obtaining of

scratches and subsequent scanning were performed by the same probe.

Figure 6 shows the scratch region on the C_{70} specimen surface. Initially, the pressure applied to the specimen was 26 GPa, which increased to 40 GPa under shear (the diffraction pattern in Fig. 2d). Against the background of rounded amorphous features, one can see the hexagonal prism of the C_{70} crystal (shown by arrow *2* at the bottom of Fig. 6). The scratch runs through the amorphous parts of the specimen. The hardness determined from the scratch equals 30 GPa, which completely coincides with that obtained in [14] for an amorphous specimen after the action of a pressure exceeding 35 GPa.

Thus, our diffraction, spectroscopic, and hardness measurements confirm the irreversible transformation of C_{70} molecular fullerite into the superhard amorphous phase at high pressures and room temperature. Under shear, this transformation occurs at lower pressures.

ACKNOWLEDGMENTS

This work was financially supported by the Russian Foundation for Basic Research (project no. 02-02-17470) and INTAS (grant no. 00237).

REFERENCES

1. V. D. Blank, S. G. Buga, M. Yu. Popov, *et al.*, Zh. Tekh. Fiz. **64** (8), 153 (1994) [Tech. Phys. **39**, 828 (1994)].
2. V. D. Blank, S. G. Buga, N. R. Serebryanaya, *et al.*, Phys. Lett. A **205**, 208 (1995).
3. V. D. Blank, S. G. Buga, N. R. Serebryanaya, *et al.*, Phys. Lett. A **220**, 149 (1996).
4. V. Blank, M. Popov, N. Lvova, *et al.*, J. Mater. Res. **12**, 3109 (1997).
5. V. D. Blank, S. G. Buga, G. A. Dubitsky, *et al.*, Carbon **36**, 319 (1998).
6. V. Blank, M. Popov, G. Pivovarov, *et al.*, Diamond Relat. Mater. **7**, 427 (1998).
7. V. D. Blank, V. M. Levin, V. M. Prokhorov, *et al.*, Zh. Éksp. Teor. Fiz. **114**, 1365 (1998) [JETP **87**, 741 (1998)].
8. V. D. Blank, N. R. Serebryanaya, G. A. Dubitsky, *et al.*, Phys. Lett. A **248**, 415 (1998).
9. Y. Iwasa, T. Furudate, T. Fukawa, *et al.*, Appl. Phys. A **64**, 251 (1997).
10. A. G. Lyapin, V. V. Brazhkin, E. L. Gromnitskaya, *et al.*, Appl. Phys. Lett. **76**, 712 (2000).
11. A. K. Sood, N. Chandrabhas, D. V. S. Muthu, *et al.*, Philos. Mag. B **70**, 347 (1994).
12. K. P. Meletov, A. A. Maksimov, and I. I. Tartakovskiĭ, Zh. Éksp. Teor. Fiz. **111**, 262 (1997) [JETP **84**, 144 (1997)].
13. S. Wasa, K. Suito, M. Kobayashi, and A. Onodera, Solid State Commun. **114**, 209 (2000).
14. J. R. Patterson, S. A. Catledge, Y. K. Vohra, *et al.*, Phys. Rev. Lett. **85**, 5364 (2000).
15. C. Cristides, I. M. Thomas, T. J. S. Dennis, and K. Prassides, Europhys. Lett. **22**, 611 (1993).
16. N. Chandrabhas, A. Sood, D. V. S. Muthu, *et al.*, Phys. Rev. Lett. **73**, 3411 (1994).
17. V. D. Blank, Yu. Ya. Boguslavskiĭ, M. I. Eremets, *et al.*, Zh. Éksp. Teor. Fiz. **87**, 922 (1984) [Sov. Phys. JETP **60**, 525 (1984)].
18. V. D. Blank, A. A. Vorontsov, N. R. Serebryanaya, and A. Yu. Tserr, Fiz. Tekh. Vys. Davleniĭ **2**, 82 (1992).
19. S. N. Sulyanov, A. N. Popov, and D. M. Kheiker, J. Appl. Crystallogr. **27**, 934 (1994).
20. N. R. Serebryanaya, V. D. Blank, V. A. Ivdenko, and L. A. Chernozatonskii, Solid State Commun. **118**, 183 (2001).
21. S. Lebedkin, W. E. Hull, A. Soldatov, *et al.*, J. Phys. Chem. B **104**, 4101 (2000).
22. V. N. Denisov, L. A. Kuzik, N. Lvova, *et al.*, Phys. Lett. A **239**, 328 (1998).

Translated by Yu. Vishnyakov

Technical Physics, Vol. 47, No. 12, 2002, pp. 1538–1542. Translated from Zhurnal Tekhnicheskoĭ Fiziki, Vol. 72, No. 12, 2002, pp. 54–58.

SOLIDS

Fracture of Spheroplastic under Static and Dynamic Stressing

S. A. Atroshenko, S. I. Krivosheev, Yu. V. Petrov, A. A. Utkin, and G. D. Fedorovskiy

Research Institute of Mathematics and Mechanics,
St. Petersburg State University, Universitetskaya nab. 7/9, St. Petersburg, 198904 Russia
e-mail: dyn@unicorn.ath.spbu.ru

Received March 25, 2002

Abstract—The fracture of a composite material, a spheroplastic consisting of a polyester resin matrix and glass microspheres as a filler, is studied experimentally and theoretically under static and dynamic stressing. A shock is generated by a pulsed magnetic field. The fracture type in relation to the shock parameters and material structure is analyzed. A method for testing the dynamic behavior of the material based on the incubation time accumulation is suggested. © *2002 MAIK "Nauka/Interperiodica".*

INTRODUCTION

The pulsed stressing of heterogeneous gas-filled media has not been adequately explored, although the shock compression of porous materials is widely used in the physics of shock waves. Interest in such materials is dictated by the need to study their behavior as targets used in inertial thermonuclear fusion and in buoyant facilities of deep-sea apparatuses, as well as by the need to synthesize novel materials capable of damping pulsed loads. Of great practical interest is a material where glass or carbon microspheres are embedded in a binder: spheroplastic [1]. Such materials have high structural and insulating properties, a relatively high impact strength, and a small weight. For example, the dynamic strength of a material consisting of glass microspheres (42 vol % or 27.7 wt %) embedded in an epoxy binder was measured to be 0.24 GPa [2]. The reinforcement of viscoelastic binders by glass microspheres decreases greatly the density of the resulting composite and improves its damping properties. The embedding of microspheres is a promising way to synthesize heterogeneous materials with desired properties.

MATERIALS AND EXPERIMENTAL METHODS

In this work, we studied the deformation and strength properties (in particular, static and dynamic cracking resistance) of a novel spheroplastic consisting of a polyester resin matrix and a reinforcing filler (glass microspheres, 41 vol %). The sphere size determined by microstructure analysis was found to vary from 6–60 to 12–60 μm (from 21 to 31 μm on average) in various specimens. The spread in the filler particle sizes is distinctly seen in Fig. 1. The parameters of the filler are listed in Table 1. The mean density of the spheroplastic found by weighing was $\rho = (0.79 \pm 0.01) \times 10^3$ kg/m^3.

The static mechanical characteristics of the material [3] were determined with a P-0.5 breaking machine at an extension rate of 1.67×10^{-4} m/s. The Young's modulus under extension was found to be $E = 2400 \pm 50$ MPa; the rupture strength, $\sigma_c = 12.4 \pm 0.9$ MPa; and

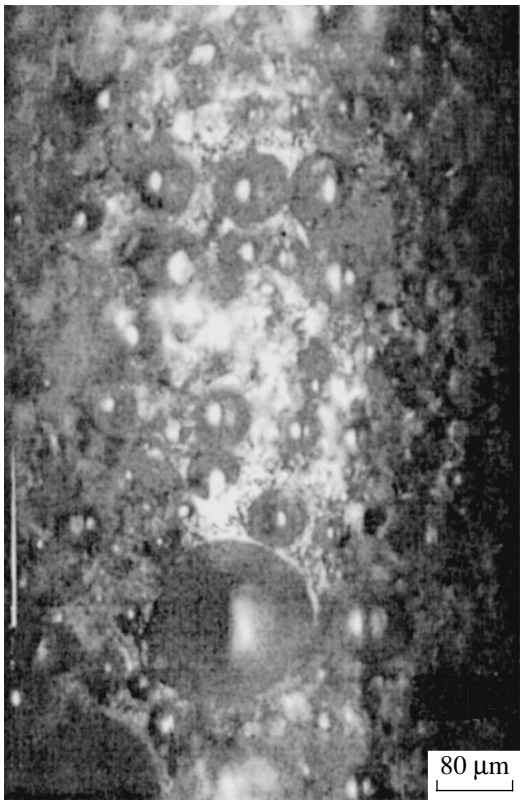

Fig. 1. Fracture surface of the spheroplastic.

1063-7842/02/4712-1538$22.00 © 2002 MAIK "Nauka/Interperiodica"

the fracture toughness (critical stress intensity factor), $K_{Ic} = 0.52 \pm 0.03$ MPa $\sqrt{m}$. The longitudinal wave velocity determined with the interferometric method was $c_1 = 2196 \pm 50$ m/s.

The dynamic cracking resistance of the material was estimated with a new testing method that is based on the concept of incubation time [4]. According to this approach, the basic parameter specifying the critical characteristics of dynamic fracture is the incubation time τ of a material. It is specific for each material and can be found by testing cracked specimens. In our experiments, the specimens were plates measuring $120 \times 120 \times$ (9–16) mm with a 60-mm-long middle through edge notch of width 2.2 mm passing into a narrower symmetrical notch of depth 1.5 mm and width 0.18 mm. The stressing scheme and the shape of the specimen are shown in Fig. 2. When the current passes through the lines of the stressing device, they experience a magnetic pressure, which is applied to the notch edges. This pressure is uniformly distributed over the specimen and is uniquely related to the current. A current pulse is generated with a magnetic device ($\xi = 2$) [5] and is measured with a Rogowski loop on a TDS-754C oscilloscope. The pulse shape and width are defined by the parameters of the pulse-generating device, and the amplitude also depends on the parameters of the stressing device. The pressure is found by the formula $p(t) = P\sin^2(\pi t/T)$, where T is the pulse width. For strip lines of width b, the pulse amplitude P is given by $P = (\mu_0/2)(I/b)^2$, where I is the current amplitude and $\mu_0 = 4\pi \times 10^{-7}$ H/m. The pulsed magnetic device can generate microsecond current pulses of peak power up to 1 MA and, thereby, produce controllable stress pulses of amplitude about 1 GPa and width of >1 μs.

The experiments were carried out for the stress pulse durations 2.76 and 4.40 μs. Figure 3 shows the length of a crack grown in the specimen vs. pulse amplitude. The threshold amplitude was determined by extrapolating the crack length to zero. They were found to be 1.68 ± 0.5 and 11.2 ± 0.4 MPa, respectively, for the above stress pulse widths.

RESULTS AND DISCUSSION

According to our approach, two material constants, K_{Ic} and τ, are responsible for the static and dynamic cracking resistances of a material on a given scale level. The critical value of the external load, as well as the rate dependence of the fracture toughness for a given structure stressed symmetrically about the crack line, can be found with the incubation time criterion [4]

$$\int_{t-\tau}^{t} K_I(t')dt \le \tau K_{Ic}, \qquad (1)$$

where $K_I(t)$ is the stress intensity factor as a function of time and τ is the incubation time.

Table 1. Filler parameters

Specimen no.	Size, μm	Wall thickness, μm
1	48.5	15.5
3	35.1	14.7
4	21.1	14.4
6	30.5	13.1

The stress intensity factor was found by calculation. To do this, we considered the problem of a semi-infinite crack whose edges are stressed by pressure pulses of a given shape. With a unit instantaneous pulse applied to the notch edges, the stress intensity factor is given by $K_I^{(\delta)}(t) = \alpha/2/\sqrt{t}$, where $\alpha = 4c_2\sqrt{c_1^2 - c_2^2}/(c_1\sqrt{\pi c_1})$ [6]. If a pulse of a given shape is applied, the stress intensity factor is found by the time convolution of the stress applied with $K_I^{(\delta)}(t)$. Thus, the stress intensity factor is experimentally determined from the relationship

$$K_I(t) = \alpha P\int_0^t \frac{\sin^2(\pi(t-s))/T}{2\sqrt{s}}ds. \qquad (2)$$

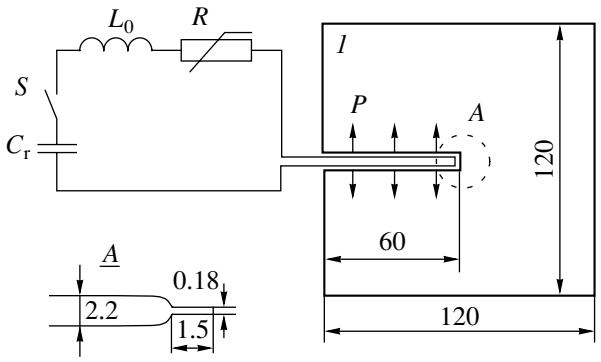

Fig. 2. Pulse-generating circuit and specimen stressing scheme. C_r, reservoir capacitor; L_0, self-inductance of current pulse generator; *S*, high-voltage switch; *R*, nonlinear resistor; and *1*, specimen. Sizes are given in millimeters.

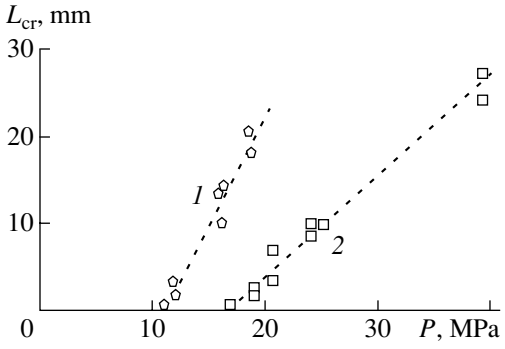

Fig. 3. Experimental data for spheroplastic fracture by individual pulses of duration (*1*) 4.40 and (*2*) 2.76 μs.

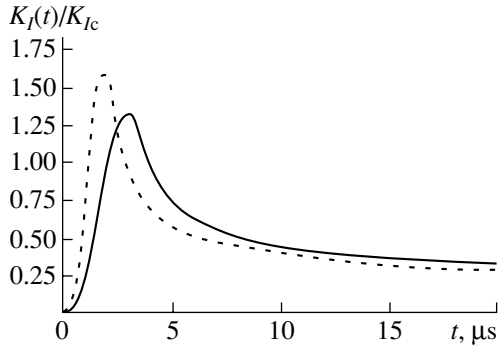

Fig. 4. Stress intensity factor vs. time. Continuous and dashed lines refer, respectively, to the durations 4.40 and 2.76 μs.

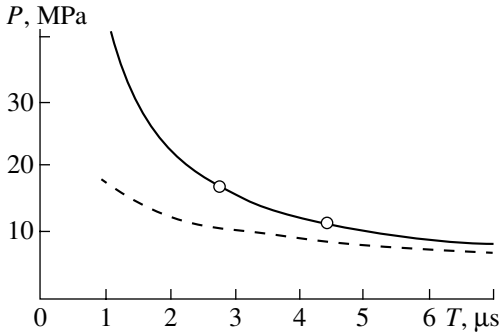

Fig. 5. Breaking amplitude P vs. pulse duration T.

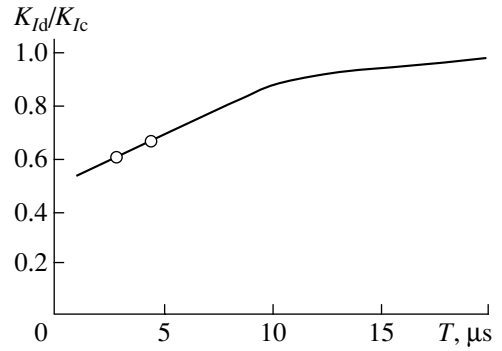

Fig. 6. K_{Id}/K_{Ic} vs. pulse duration. Circles show the dynamic fracture toughness for the durations used in the experiment.

This relationship for the stress intensity factor is valid as long as the problem of semiinfinite crack is applicable, i.e., until the waves reflected from the boundaries reach the notch tip. For pulse models used in the tests, one should consider times $t \le 27.3$ μs, i.e., time periods within which the longitudinal wave travels along the notch edge. For these times, the stress intensity factor vs. time is shown in Fig. 4. The value of $K_I(t)$ grows, reaches a maximum, and then decreases. The instant the maximum is attained slightly exceeds that of pulse application ($T/2$).

The incubation time is found from the parameters found experimentally: threshold pulse amplitude P and its duration T. From these parameters, one can determine the stress intensity factor for a threshold pulse. A threshold pulse is a breaking pulse such that any decrease in its amplitude no longer causes fracture. Hence, the maximal value of the integral in (1) transforms inequality (1) into an equality. From this equality, one finds the incubation time τ. The incubation time thus found for the spheroplastic considered is $\tau = 3.93 \pm 0.25$ μs for a pulse of duration $T = 2.76$ μs and 3.94 ± 0.30 μs for $T = 4.40$ μs. Thus, the incubation time equals 3.93 μs.

Figure 5 shows the threshold pulse amplitude vs. pulse duration. Circles are data points, and the continuous curve stands for the threshold amplitude calculated by criterion (1). The threshold amplitude determined with the conventional critical stress intensity factor criterion ($K_I(t) \le K_{Ic}$) is somewhat lower (dashed curve in Fig. 5). For the pulse durations $T = 2.76$ and 4.40 μs, the breaking amplitude found with the critical stress intensity factor amounts to 64 and 76%, respectively, of that observed in the experiment.

The dynamic fracture toughness, i.e., the value of the stress intensity coefficient at the instant of fracture onset ($K_{Id} = K_I(t_*)$, where t_* is the instant of fracture onset) can be found with criterion (1). In this case, first the instant of fracture onset t_* is found from (1), and then the stress intensity coefficient at this instant is determined from (2). In the experiments where the fracture was caused by threshold pulses, the dynamic fracture toughness was found to be $K_{Id} = 0.61K_{Ic}$ and $0.67K_{Ic}$ for $T = 2.76$ and 4.40 μs, respectively. The instant of fracture onset t_* was 5.4 and 6.1 μs, respectively. In both cases, the fracture has a delay, i.e., occurs some time after the stress intensity coefficients have reached their maximum (Fig. 6).

The specific features of the deformation behavior and fracture of heterogeneous materials are associated with elastic stress sources arising at the continuous interfaces between homogeneous structure elements during deformation. The sources are due to a mismatch between the properties of dissimilar constituents. In the stress field, the constituents undergo various deformations and the complex shape of the interfaces can be retained if the elastic strains exceed their mean value.

The micrographs of the fracture surface are shown in Figs. 1, 7, and 8, and Table 2 lists the crack length vs. filler size. The fracture of the spheroplastic takes place largely between the glass microspheres. The crack propagates in the binder, usually bypassing the spheres. The spheres are crossed in few cases, as follows from Fig. 7.

As a rule, two cracks originate at the notch, which usually propagate at a certain angle to each other (8°–12°). At distances of 2 or 4 mm from the fracture site, the crack may sharply change its propagation direction,

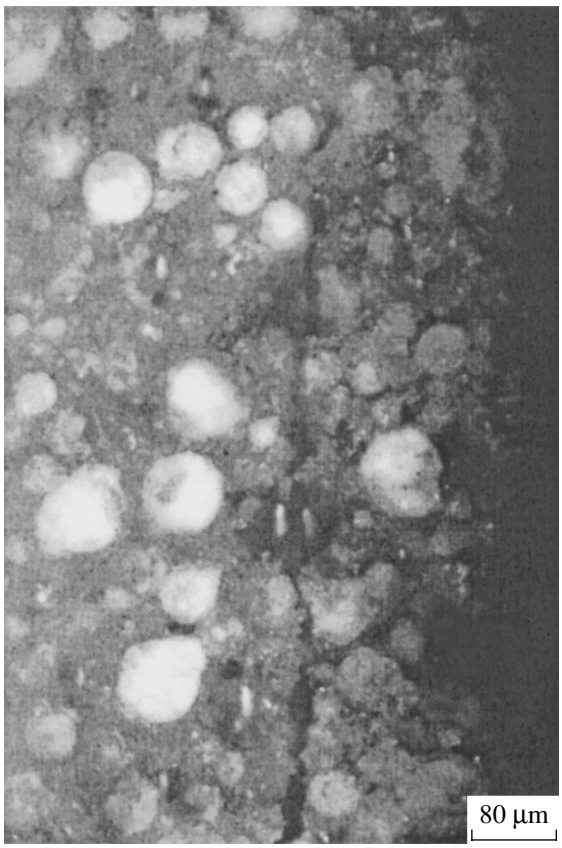

Fig. 7. Crack propagation in the spheroplastic.

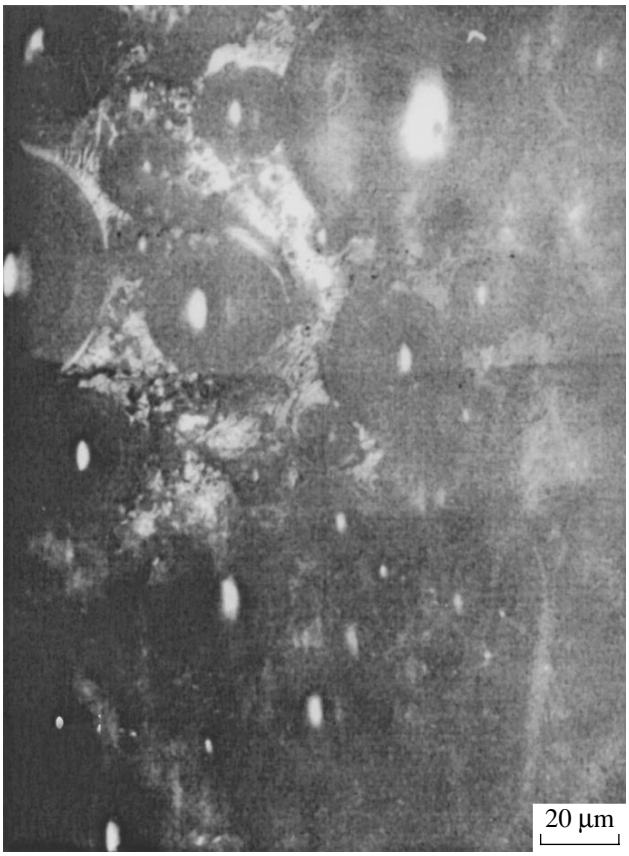

Fig. 8. Traces of plastic flow and microcracks on the fracture surface of the spheroplastic.

deviating from the initial one by 8°–10°. Then, passing roughly the same distance, the crack may start propagating in the initial direction again. The crack travel varies from 7 to 27 mm, depending on the shock intensity. As follows from Table 2, the smaller the sphere radius, the shorter the final length of the crack. This can be explained by the increase in the number of obstacles over the crack path.

The microstructure of the spheroplastic at a high magnification is depicted in Fig. 8. The traces of the irreversible deformation of the binder and microcracks on several areas surrounding the glass microspheres are seen. At an even higher magnification, one can observe the pattern of filler fracture typical of other plastics, e.g., of polymethylmethacrylate: specular region, rough region, and parabolic region.

Table 2. Final crack length vs. filler size

Specimen no.	Crack length, mm		Filler size, μm
	first	second	
1	24	27	48.5
3	9.6	11.9	35.1
4	7	11	21.1
6	9	11.2	30.5

Comparing the fracture of the statically and dynamically stressed material, one can conclude that the number of broken microspheres in the former case far exceeds that in the case of pulsed stressing.

CONCLUSION

The dynamic strength properties of the spheroplastic were tested with a new method based on the incubation time concept.

The threshold amplitude values obtained with the new method differ markedly from those predicted with the conventional stress intensity factor criterion, and the values of the dynamic fracture toughness turn out to be much less than K_{Ic}.

It is shown that the fracture due to threshold pulses originates at the crack tip when the stress intensity factor starts decreasing.

The fracture analysis of the spheroplastic shows that the smaller the spheres, the smaller the crack travel for a given stress level.

ACKNOWLEDGMENTS

This work was jointly supported by the Russian and Belarussian Foundations for Basic Research (grant no. 02-01-81036), the Russian Foundation for Basic Research (grant no. 99-01-00718), CJU project no. 97-01-05009, program "Integration," and the Grant Award Center at St. Petersburg State University.

REFERENCES

1. P. G. Krzhechkovskiĭ, Doctoral Dissertation (Nikolaev, 1989).
2. L. J. Weirick, *Shock Compression of Condensed Matter* (Elsevier, Amsterdam, 1992), pp. 99–102.
3. I. M. Kernshteĭn, V. D. Klyushnikov, E. V. Lomakin, and S. A. Shesterikov, *Foundations of Experimental Mechanics of Fracture* (Mosk. Gos. Univ., Moscow, 1989).
4. N. Morozov and Y. Petrov, *Dynamics of Fracture* (Springer-Verlag, Berlin, 2000).
5. S. I. Krivosheev and Yu. V. Petrov, Preprint No. 142, IPMASh RAN (Institute of Problems in Machine Science, Russian Academy of Sciences, St. Petersburg, 1997).
6. G. P. Cherepanov, *Mechanics of Brittle Fracture* (Nauka, Moscow, 1974).

Translated by V. Isaakyan

Technical Physics, Vol. 47, No. 12, 2002, pp. 1543–1546. Translated from Zhurnal Tekhnicheskoĭ Fiziki, Vol. 72, No. 12, 2002, pp. 59–63.

SOLIDS

On the Mechanism of Fibril Rotation at the Early Stages of Amorphous–Crystalline Polymer Reorientation

B. M. Ginzburg*, N. Sultonov, and A. A. Shepelevskii***

** Institute of Problems in Machine Sciences, Russian Academy of Sciences, Vasil'evskiĭ ostrov, Bol'shoĭ pr. 61, St. Petersburg, 199178 Russia*

e-mail: ginzburg@tribol.ipme.ru

*** Tajik State National University, ul. Rudaki 17, Dushanbe, 734000 Tajikistan*

Received April 24, 2002

Abstract—The microdeformation behavior of polyamide-6 (PA6) oriented films is studied by X-ray diffraction at the early stages of reorientation at an angle of 45° to the direction of primary orientation. In the elastic strain range (up to 22%), the shear of crystallites and the rotation of fibrils take place simultaneously. The rotation is provided by the mutual slip of neighboring lamellas inside the fibrils. © *2002 MAIK "Nauka/Interperiodica".*

INTRODUCTION

For flexible-chain crystalline polymers, the shear of crystallites seems to be a basic deformation process occurring at the supermolecular level. The study of polymer reorientation at different angles α relative to the primary orientation direction is a suitable technique for gaining information on the properties of polymer systems at the supermolecular level [1, 2]. In [1], X-ray diffraction was used to study the reorientation of polyethylene films. Upon analyzing the initial (elastic) deformation stages, it was assumed that the crystallites are subjected to shear and gradually turn toward the direction of secondary extension. However, the mechanism of fibril rotation and the reasons why fibrils remain intact have not been clarified. In this work, an attempt is made to elucidate a mechanism of fibril rotation upon PA6 stretching.

EXPERIMENTAL

The starting material was 70-μm-thick films of PK-4 commercial oriented polyamide. The films were additionally oriented by extending by 60% along the orientation direction at 150°C. Then they were annealed in the free state at T_1 = 200°C for 2 h at a pressure of 0.13 Pa. This caused the films to shrink by 7%. After the annealing, the films were cut into specimens at an angle $\alpha = 45°$ to the direction of primary stretching in the form of a double-sided blade with working parts measuring 20 × 10 mm. The specimens were then extended in steps at T_2 = 20°C. X-ray diffraction patterns were taken from the stressed specimens and from those in the free state (i.e., after relaxation for 24 h). Small-angle and large-angle X-ray scattering (SAXS and LAXS) patterns were taken from the same section of the specimen simultaneously. Quantitative SAXS and LAXS measurements were performed with a KRM-1 X-ray camera and a DRON-2.0 diffractometer, respectively. Ni-filtered CuK_α radiation was used. The crystallite size was calculated by the Scherrer formula.

RESULTS AND DISCUSSION

Figures 1a and 1b show the experimental scheme, the mutual arrangement of reflections in SAXS and LAXS patterns, the crystallite faces, and the directions of macromolecule axes in the initial state, and Fig. 1c demonstrates the hypothetical changes in the scattering patterns at the early stage of deformation (within the elastic range). The scheme uses the Tsvankin model of fibrils (with beveled crystallites) [3], where the direction of long periods is aligned with the macromolecule axes in the crystallites. It is also shown that shear in the crystallites is accompanied by fibril rotation [1].

Figure 2 shows X-ray patterns taken from PA6 specimens at different values of elongation ε_0. For $\varepsilon_0 \le 20$–25%, the changes in the scattering patterns and specimen sizes are reversible (Fig. 2b). In the SAXS patterns, the centers of streak reflections shift sideways from the initial direction **M** of the macromolecule axes (cf. Fig. 1c) according to the shear of the crystallites. In the LAXS patterns, (002) reflections rotate in the azimuth direction toward the direction of secondary extension. At the same elongations, (200) reflections shift insignificantly. These changes in the scattering pattern are explained by intracrystallite shear along the (002) planes, which is accompanied by the rotation of the macromolecule axes (Fig. 1c), and, at the same time, by the absence of tangible shear in the crystallites along the (200) planes. One may assume that the shear of the crystallites along the (200) planes also takes place but the *b* axes do not rotate. However, other experiments indicate that the shear along the (200) planes is absent. Indeed, the shear of crystallites usually widens reflec-

1063-7842/02/4712-1543$22.00 © 2002 MAIK "Nauka/Interperiodica"

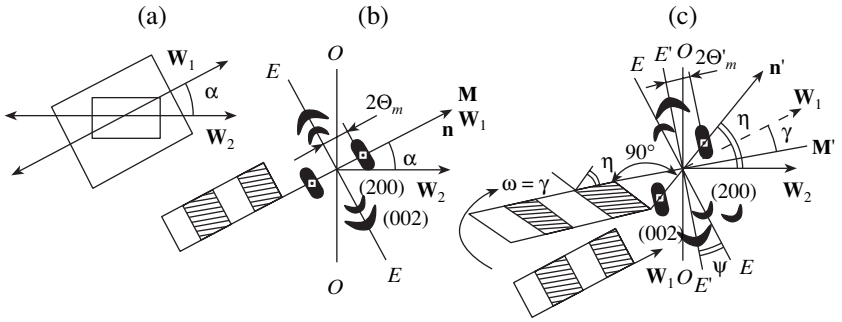

Fig. 1. Experimental scheme. (a) Reorientation scheme, (b) initial shape of crystallites and fibrils and associated X-ray patterns, and (c) the same as in (b) under reversible (elastic) "reorientation" extension. $\mathbf{W}_1$, direction of primary extension; $\mathbf{W}_2$, direction of secondary extension; α, reorientation angle; η, shear angle of crystallites; ψ, divergence angle of equatorial reflections; $\mathbf{n}$, direction of the normal to the end faces of the crystallite; $\mathbf{M}$ and $\mathbf{M}'$, initial and new directions of the macromolecule axes, respectively; γ, rotation of the macromolecule axes; ω, rotation of fibrils; $2\Theta_m$, angular position of the intensity maximum of the small-angle reflection; and d_0, long period.

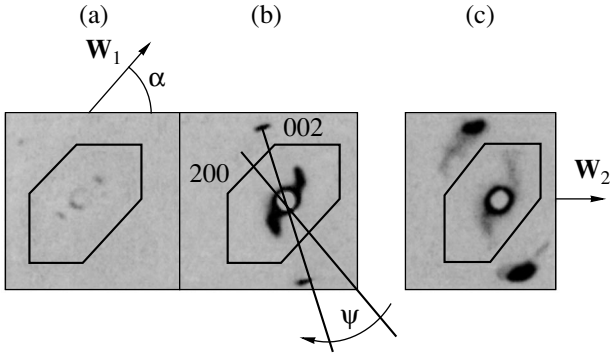

Fig. 2. SAXS and WAXS patterns taken from the PA6 films under reorientation at an angle $\alpha = 45°$ for $T_2 = 20°C$. ψ is the divergence angle of (002) and (200) equatorial reflections. Reorientation extension is horizontal. Strain ε_0 = (a) 0, (b) 22, and (c) 35%.

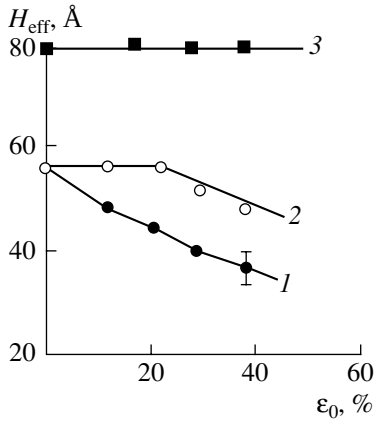

Fig. 3. Effective transverse crystallite size H_{eff} vs. strain of the PA6 specimen under reorientation as measured from the (002) reflection for the (*1*) stressed and (*2*) unstressed specimen and from (*3*) the (200) reflection for the stressed or unstressed specimen.

tions from shear planes in the radial direction. Figure 3 shows the effective transverse crystallite size H_{eff} vs. strain. The widening of the (002) reflection does take place; accordingly, the crystallite size decreases, the decrease being reversible for strains of up to ≈22%: after the stress is relieved, H_{eff} recovers to the initial value. At the same time, the width of the (200) reflection and, hence, the crystallite size remain unchanged. This argues against the supposition that shear deformation occurs along the (200) planes.

The change in the (002) reflection width goes in parallel with a decrease in the reflection intensity (Fig. 4). This may be due to a decrease in the size of coherent scattering areas under shear: in region *A* (Fig. 5), these areas retain their transverse size, whereas in regions *B* and *C*, they shrink. This fact can be used for interpreting a decrease in the intensity of reflections from paratropic planes, which is often observed when highly oriented systems are deformed along the axis of orientation [4].

When estimating the shear strain of the crystallites, we assumed that the mean direction of the normal *n* to the upper and lower faces of the crystallites coincides with a straight line connecting the center of the SAXS pattern with the maximal intensity of the reflection. In the initial state, the directions of *n*, macromolecule axes $\mathbf{M}$, and primary orientation $\mathbf{W}_1$ coincide (Fig. 1b). After the extension, they diverge (Fig. 1c). In the first approximation, however, we will assume that the direction $\mathbf{M}$ of the macromolecule axes in the crystallites and the direction $\mathbf{M}_1$ of the fibril axes (the direction of the long periods) coincide. Let us draw a new direction of the macromolecule axes $\mathbf{M}'$ perpendicularly to the line *E*'*E*', which passes through the centers of the (002) reflections, since precisely these (002) planes are the shear planes. Then, knowing the new direction of the normal *n*', one can find the angles of intracrystallite shear (see table). Since the position of the (200) reflection remains unchanged in the elastic strain range of the specimens, the angle γ through which the direction $\mathbf{M}'$

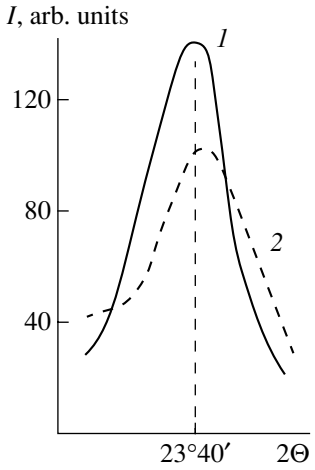

Fig. 4. Changes in the (002) and (200) reflection profiles. Strain ε_0 is (*1*) 0 and (*2*) 20%.

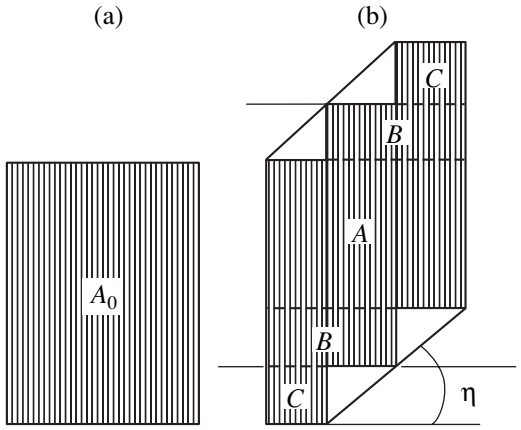

Fig. 5. Diagram explaining the decrease in the size of coherent scattering areas and in the intensity of reflections from the shear planes: (a) initial crystallite and (b) crystallite after shear deformation.

rotates equals the angle ψ, which is the difference between the azimuth rotations of the reflections (002) and (200) (Fig. 1c, table). Then, the angle γ through which the (002) reflections rotate under the elastic strain $\varepsilon_0 = 12$ and 22% equals, respectively, 8° and 15°; i.e., they are roughly half as large as the angles η. This means that shear in the crystallites involves not only intracrystallite slip with macromolecule rotation but also pure shear along the macromolecule axes without rotation. The long periods under deformation were estimated by the Bragg formula and from the angular position $2\Theta_m$ of the projection of the small-angle reflection maximum onto the new "meridian" (i.e., onto the new direction of the macromolecule axes; Fig. 1c). With strains of 10–12%, the long period increases by ≈7% and then remains unchanged until necking. Figure 6 shows the deformation scheme for a fibril fragment at $\varepsilon_0 \approx 22\%$ with all the geometrical relationships met. In the initial (unstrained) state of the PA6 specimens, the transverse and longitudinal crystallite sizes are $H_0 = 58 \pm 2$ Å (according to the (002) reflection profile) and $L_0 = 52 \pm 2$ Å (according to the meridional (0.14.0.) reflection profile). Since the transverse and longitudinal crystallite sizes are roughly the same and the specimens are strained at an angle of 45°, the square cross section *OBCA* (Fig. 6a) takes the rhombic shape *OB'C'A'*; the shear angles η_1 and η_2 (Fig. 6b) are roughly equal to each other, ≈15°, so that the total shear angle $\eta = 30°$. This is consistent with the fact that the rotations of macromolecules in the crystallites are half the shear angle η (see table).

If the position of the crystallite apexes *O* remains unchanged, the long period must decrease and macromolecules in the amorphous regions must rotate in opposition to the crystallite rotation. Actually, the long period obtained by projecting the maximum position of the small-angle reflection intensity on the direction **M'** (Fig. 1c) grows by ≈7%. Just such a value of the long period d_ε is shown in Fig. 6b. To match d_ε with the position of the apex *O*, the apex should be placed at the point O_1. In other words, interlamellar slip in the direction $\mathbf{W}_\lambda$ by $\lambda \approx 20$ Å (≈34% of the length of the crystallite edge *OA'*) must occur. As a result, the structure shown in Fig. 6c forms. The angle δ between the direction of the long period and that of the macromolecule axes in the crystallites is ≈4°, which nearly does not affect the determination of the long period and the crystallite shear angle. Therefore, even the first approximation used in selecting the direction of a "new meridian" and estimating the long period under strain proved to be adequate. Such a scheme explains the fact that fibrils remain intact (elastic deformation of the specimens) and rotate toward the direction of secondary extension, as well as validates the use of the Tsvankin model [3] for treating experimental data. It seems that this scheme is also valid for other polymers. It is worth noting, however, that if the directions $\mathbf{M}_f$ and **M** diverge significantly, it is necessary to employ the Seto–Hara fibril model [5] (where these directions diverge) and the long period and shear angles should be estimated by the method of successive approximations. In fact, this

Changes in the shear angle η of crystallites, divergence angle ψ of equatorial reflections, and long period under the reorientation of the PA6 films for $\alpha = 45°$

Specimen strain ε, %	Shear angle η of crystallites, deg	Divergences angle ψ, deg	Long period *d*, Å
0	0	0	100
12	15 (0)	8 (0)	107
22	30 (0)	15 (0)	107
35 (9)	55 (22)	27 (10)	107

Note: In parentheses, the residual values of these parameters after stress release are given.

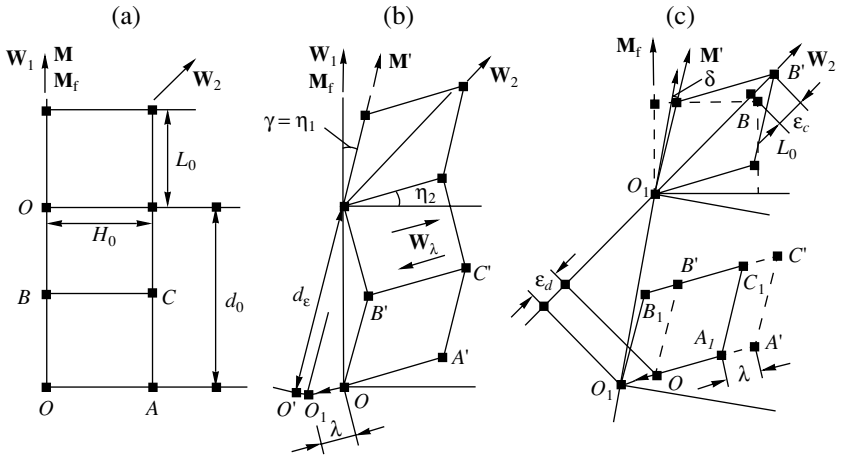

Fig. 6. Schematic representation of the fibrillar structure of the PA6 specimen under shear deformation at an angle of 45° to the direction of primary orientation. A fibril fragment (two crystallites and the amorphous region in between) is shown. (a) Initial state, (b) deformed state, and (c) state after interlamellar slip.

method was used in this work; however, a single iteration sufficed because of the smallness of δ.

Consider the contribution of the shear strain of the crystallites and long periods to the total strain of the specimen (Fig. 6c). Having projected the apexes *B* and *B*' of the crystallite onto the direction $\mathbf{W}_2$ of secondary extension, we see that the crystallite elongation in the direction of secondary extension is 17%, that is, roughly equals the strain of the whole specimen. Similarly, having projected the points *O* and O_1 onto the same direction, we see that the strain of the long period ε_d in the direction $\mathbf{W}_2$ is ≈26%, which far exceeds the specimen strain ε_0. The relationship $\varepsilon_d > \varepsilon_0$ is typical of the nonuniform deformation of long periods [6]. This agrees with the results of [7], where it was shown that the same PA6 films exhibit the sharply nonuniform deformation of long periods under elastic extension at room conditions. Interfibrillar slip under these conditions is virtually absent.

REFERENCES

1. B. M. Ginzburg, N. Sultonov, and D. Rashidov, J. Macromol. Sci., Phys. **B9**, 609 (1974).
2. B. M. Ginzburg and N. Sultonov, Zh. Tekh. Fiz. **71** (2), 129 (2001) [Tech. Phys. **46**, 258 (2001)].
3. V. I. Gerasimov, Ya. V. Genin, A. I. Kitaigorodsky, and D. Ya. Tsvankin, Kolloid Z. Z. Polym. **250**, 518 (1972).
4. B. M. Ginzburg, A. A. Shepelevskii, N. Sultanov, and Sh. Tuichiev, J. Macromol. Sci., Phys. **B41**, 357 (2002).
5. T. Seto and T. Hara, Rep. Prog. Polym. Phys. Jpn. **12**, 189 (1969).
6. B. M. Ginzburg and Sh. Tuichiev, J. Macromol. Sci., Phys. **B31**, 291 (1992).
7. B. M. Ginzburg and N. Sultonov, Vysokomol. Soedin., Ser. A **43**, 1140 (2001).

Translated by V. Isaakyan

Technical Physics, Vol. 47, No. 12, 2002, pp. 1547–1553. Translated from Zhurnal Tekhnicheskoĭ Fiziki, Vol. 72, No. 12, 2002, pp. 64–71.

OPTICS, QUANTUM ELECTRONICS

Optical-to-Acoustic Energy Conversion Efficiency upon Interaction of Pulsed Laser Radiation with a Liquid: II. Conversion Efficiency Measurement by Holographic Interferometry upon Acoustooptic Interaction

G. V. Ostrovskaya

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

Received May 29, 2002

Abstract—The efficiency of conversion of pulsed CO_2 laser energy to acoustic energy is measured with holographic interferometry. The method suggested is simple and does not require that fringe shifts in the interferogram be converted to local pressure values in the acoustic wave. The efficiency measured for the thermal mechanism of energy conversion is in good agreement with analytical calculations [1] where the temperature dependence of the volume thermal expansion coefficient of water is taken into account. © *2002 MAIK "Nauka/Interperiodica".*

INTRODUCTION

In our previous work [1], we derived relationships for the efficiency of laser-to-acoustic energy conversion for the acoustooptic (thermal) mechanism of interaction and calculated the efficiency for the specific case of the interaction between a pulsed transversely excited atmospheric-pressure (TEA) CO_2 laser with water taking into account the temperature dependence of the volume thermal expansion coefficient of water. This paper, which is a continuation of [1], is concerned with the experimental measurement of this efficiency by the method of holographic interferometry.

The starting formula for calculating the conversion efficiency has the form (see formula (3) in [1])

$$\eta = \frac{1}{\rho_0 c_0 E_L}\int_S\left\{\int_{-\infty}^{+\infty}[P_S(t)]^2 dt\right\}dS, \qquad (1)$$

where ρ_0 is the initial density of the medium, c_0 is the velocity of sound in the medium, E_L is the laser radiation energy, and P_S is the excess pressure in the acoustic pulse with respect to the initial pressure P_0 at an arbitrary point S on the wave front surface.

To calculate the efficiency by formula (1), one must know the pressure time profile throughout the wave front surface rather than at a given point of the space. The space–time pattern of the pressure field $P(r, \theta, t)$ is most conveniently obtained with the method of holographic interferometry, which was applied by us earlier in studying hydrodynamic phenomena accompanying the interaction of CO_2 laser radiation with water [2–4].

EXPERIMENT

The experimental setup is depicted in Fig. 1. Mirror *1* directs the radiation of a pulsed TEA CO_2 laser toward water-filled cuvette *2* normally to the water surface. Circular diaphragm *3* of diameter 2 cm cuts the most uniform part of the laser beam, and BaF_2 lens *4* (F = 30 cm) narrows the laser beam. Having passed through the diaphragm and the lens, the radiation had an energy of ≈1.5 J. The laser energy density ε_L on the surface was varied between 0.75 and 8 J/cm^2 by varying the focal spot size. The latter, in turn, was varied by varying the lens–water surface spacing. The spot size was determined from its print on heat-sensitive paper.

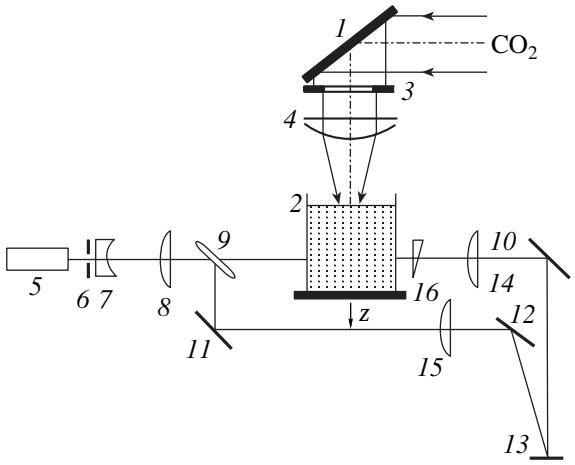

Fig. 1. Experimental setup.

1063-7842/02/4712-1547$22.00 © 2002 MAIK "Nauka/Interperiodica"

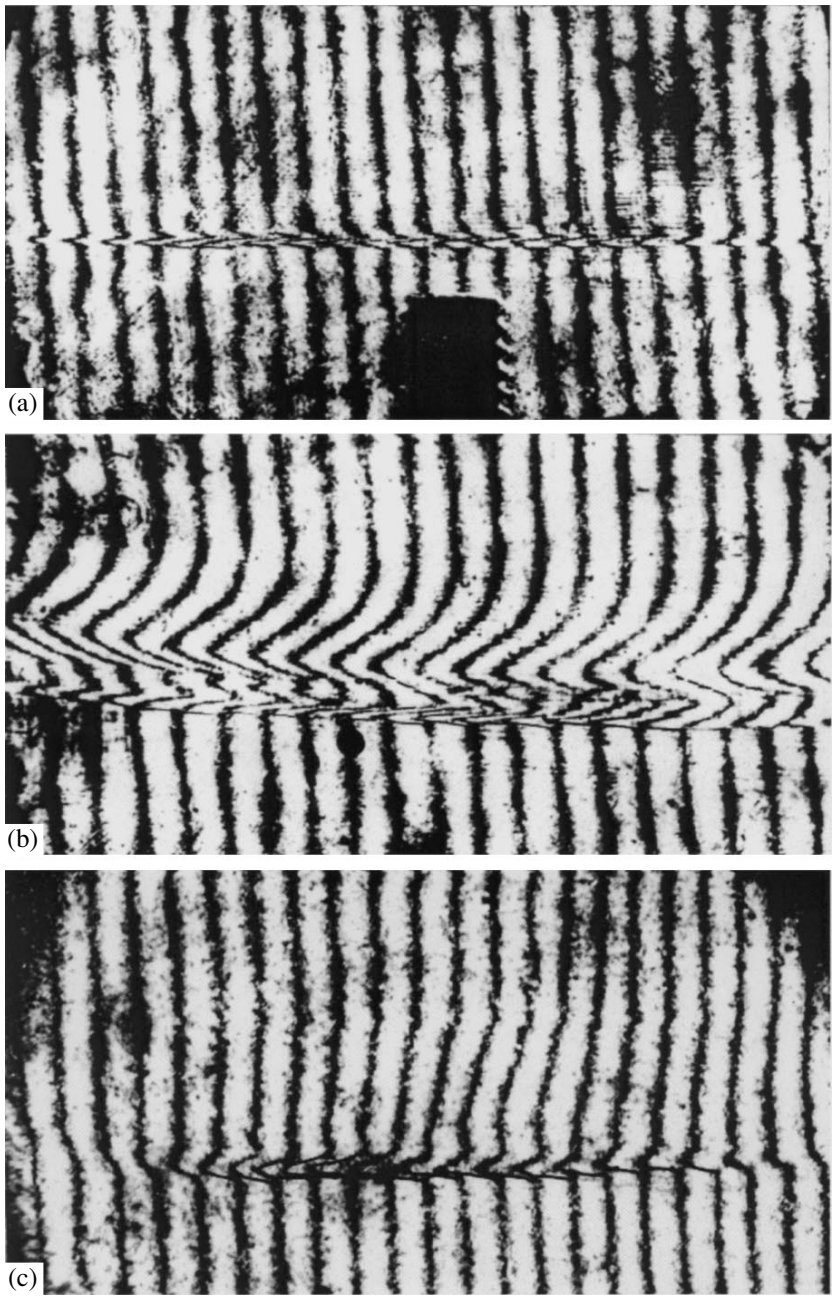

Fig. 2. Holographic interferograms of an acoustic wave generated when the radiation of a CO_2 laser interacts with the free surface of water at energy densities ε_L = (a) 1 and (b) 3 J/cm^2. The distance from the surface is d = 7 cm. (c) Interaction with the rigid water–germanium interface. ε_L = 0.5 J/cm^2, d = 0.6 cm.

Holograms were recorded with the radiation of an OGM-2 pulsed ruby laser *5*. A laser beam passed through diaphragm *6* of diameter ≈1.5 mm and then was expanded to a diameter of 20 cm by means of a telescopic system (lenses *7* and *8*). Semitransparent mirror *9* split the laser beam into object and reference beams, which were then made coincident in the plane of hologram *13* with mirrors *10–12*. The object beam passed through the cuvette parallel to the water surface at a distance $d = c_0\tau_d$ from it (τ_d is the time delay between the pulses from the CO_2 and ruby lasers). The holograms were recorded using the focused image technique: the cuvette plane (corresponding to the CO_2 laser beam axis) was projected onto the hologram plane with objective lens *14*. To make the mode structures of the object and reference beams coincident in the hologram plane, objective lens *15*, which is similar to lens *14*, was introduced into the reference beam. Glass wedge *16*, which was rotated through a certain angle between the first and second exposures, was used to obtain holo-

graphic interferograms in finite-width fringes. Typical holographic interferograms are shown in Figs. 2a and 2b. Figure 2c also shows the interferogram of an acoustic wave generated at the rigid (germanium–water) interface in our early experiments.

INTERFEROGRAM PROCESSING AND EFFICIENCY CALCULATION WITH THE USE OF ABELIAN INVERSION

Fringe shifts are known to be integral along the line of observation:

$$K(y, z) = \frac{1}{\lambda}\int_{x_1}^{x_2}\Delta n(x, y, z)dx, \qquad (2)$$

where $\Delta n(x, y, z) = n(x, y, z) - n_0$; $n(x, y, z)$ and n_0 are, respectively, the local value of the refractive index in an object and its initial value in the environment; and x_1 and x_2 are the boundaries of the object.

The value of Δn and the local pressure $P(x, y, z)$ are related as $\Delta n(x, y, z) = (n_0 - 1)\xi P(x, y, z)$, where ξ is the compressibility factor (4.5×10^{-5} atm^{-1} for water) and $n_0 = 1.33$. With this in mind, the relationship between the fringe shift and pressure can be written in the form

$$K(y, z) = A\int_{x_1}^{x_2}P(x, y, z)dx, \qquad (3)$$

where $A = (n_0 - 1)\xi/\lambda = 0.22$ atm^{-1} cm^{-1} if $\lambda = 6.94 \times 10^{-5}$ cm.

Prior to calculating the efficiency, it is necessary to construct the spatial pressure field from the fringe shifts in the interferogram. If the energy distribution within the laser spot on the liquid surface is assumed to be uniform, the experiment geometry can be considered to be symmetric about the vertical z axis (Fig. 1). Then, expression (3) transforms into the Abelian integral

$$K(y, z) = 2A\int_{|y|}^{R_0}\frac{P(R, z)PdR}{\sqrt{R^2 - y^2}}, \qquad (4)$$

where R is the distance from the z axis and R_0 is the radius corresponding to the object boundary.

The spatial pressure distribution $P(R, z)$ in an acoustic pulse can be found by numerically solving integral equation (4) (Abelian inversion) [5, 6]. The spatial pressure distribution $P(R, z)$ obtained by processing the interferogram in Fig. 2a is shown in Fig. 3a. To find the distribution along the z direction, which corresponds to the time profile of acoustic pulses on the scale $z = c_0t$, Abelian inversion was accomplished for several sections of the interferogram that ran normal to the z axis in the interval between A–A' and B–B', i.e., between the beginning and end of an acoustic perturbation. The true spacing between these sections was $\Delta l = 0.6$ mm, or

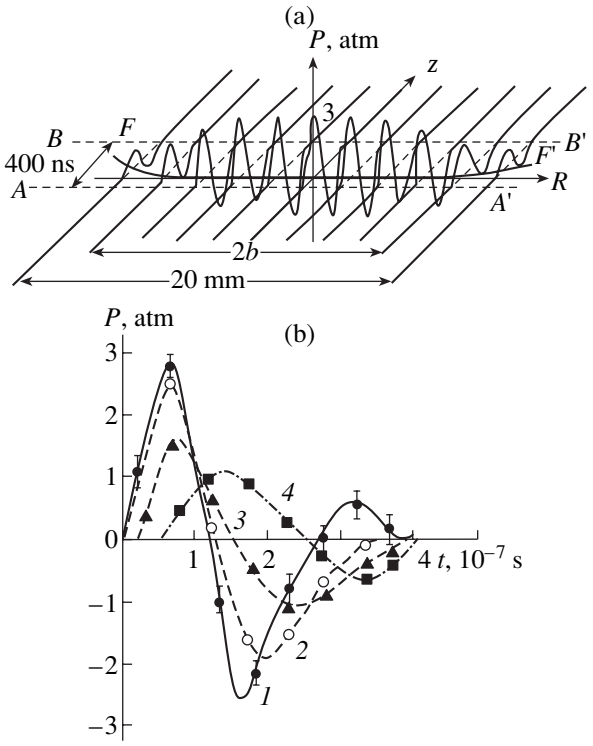

Fig. 3. (a) Two-dimensional pressure distribution in the acoustic wave obtained by processing the interferogram in Fig. 2a and (b) time profiles of the acoustic pulse at distances r from the z axis (*1*) 0–3, (*2*) 5, (*3*) 7, and (*4*) 9 mm.

$\Delta l/c_0 = 400$ ns on the time scale. It is seen that the wave front surface F–F', passing through the maxima of the first positive pressure pulse, is virtually flat. It is curved due to diffraction divergence only at the periphery of the front and shows up in the time shift of the pressure pulses.

Figure 3b demonstrates the time profiles of pressure pulses at different distances from the axis of symmetry. The time t is counted from the section A–A'. Curve *1* is the profile of the pulse averaged in the range R = 0–3 mm. Vertical bars show a spread in the measured pressure values at a given time instant for these intervals of R. Away from the axis of symmetry, the time shift due to the distortion of the wave front is observed, the negative pulse amplitude decreases markedly relative to the maximum of the positive pulse, and both pulses somewhat widen. Such changes in the acoustic pulse shape at the wave front periphery, as well as the small positive pressure peak appearing in curve *1* after the negative pulse, are typical diffraction distortions of the pulse at distances on the order of diffraction length $L_{DF} = a^2/2c_0\tau$ [7]. In our case (the radiation spot radius $a \approx 1$ cm, positive pulse duration τ = 120–150 ns, and the velocity of sound in water $c_0 = 1.5 \times 10^5$ cm/s), $L_{DF} \approx 25$ cm and the wave travel is $d \approx 7$ cm; i.e., the diffraction distortion is minor. Therefore, rejecting the time shift due to the wave front curvature, we can con-

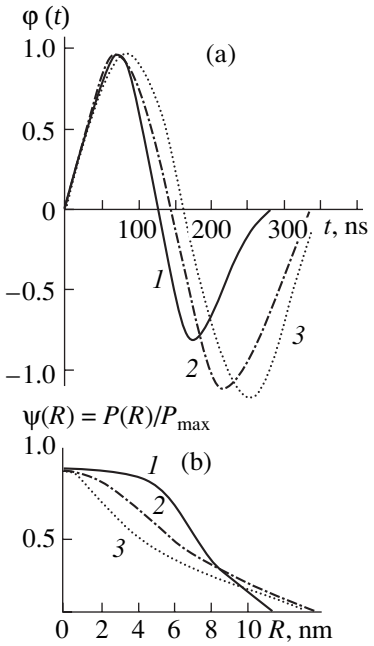

Fig. 4. Experimentally found functions (a) $\varphi(t)$ and (b) $\psi(R)$ describing, respectively, the time profile of the acoustic pulse and the spatial pressure distribution on the acoustic wave front at energy densities in the focal spot (*1*) 1, (*2*) 2, and (*3*) 3 J/cm^2.

struct the averaged pulse profile (Fig. 4a, curve *1*) $\varphi(t) \approx \bar{\varphi}(t, R) = P(R, t)/P_F(R)$, where averaging is over R between 0 and 5 mm. Here, $P_F(R) = P(R, t_{max})$ is the radial pressure distribution over the front and t_{max} is the time position of the positive pressure maximum at a given R. In turn, $P_F(R)$ can be represented as $P_F(R) = P_{max}\psi(R)$, where $P_{max} = P_F(R = 0)$. Thus, with a fairly good accuracy, the pressure field in the acoustic wave can be given by

$$P(R, t) = P_{max}\psi(R)\varphi(t). \quad (5)$$

Substituting (5) into (1) yields an expression for the conversion efficiency:

$$\eta = \frac{P_{max}^2}{\rho_0 c_0 E_L}\int_0^\infty \varphi^2(t)dt \int_S \psi^2(R)dS = \frac{\Phi_t \Psi_S P_{max}^2}{\rho_0 c_0 E_L}, \quad (6)$$

where

$$\Phi_t = \int_0^\infty \varphi^2(t)dt, \quad (7)$$

$$\Psi_S = \int_S \psi^2(R)dS = 2\pi\int_0^\infty \psi^2(R)RdR. \quad (8)$$

The functions $\varphi(t)$ and $\psi(R)$, which are found by processing the holographic interferograms corresponding to various energy densities in the focal spot, are depicted in Fig. 4.

The integral Φ_t can be found by either numerically integrating the functions $\varphi(t)$ found experimentally or approximating them by analytical functions, for example, segments of a sinusoid. In the latter case,

$$\Phi_t = (1/2)[\Delta t_+ + (P^-/P^+)^2 \Delta t_-], \quad (9)$$

where Δt_+ and Δt_- are the respective durations of the positive and negative pulses and P^+ and P^- are their amplitudes.

The quantity $\Psi_S P_{max}^2$ entering into (6) can be found by numerically integrating the experimentally found pressure distribution $P_F(R) = P_{max}\psi(R)$ (Fig. 3b) over the acoustic wave front:

$$\Psi_S P_{max}^2 = 2\pi P_{max}^2 \int_0^\infty \psi^2(R)RdR = 2\pi\int_0^\infty P_F^2(R)RdR. \quad (10)$$

The values of Φ_t and $\Psi_S P_{max}^2$ calculated by formulas (9) and (10) for various energy densities in the focal spot are listed in Table 1 (columns *a*). Here, the conversion efficiencies obtained by substituting these values of Φ_t and $\Psi_S P_{max}^2$ into (6) are also presented.

Thus, constructing the space–time pressure field from an interferogram, one can find the laser-to-acoustic energy conversion efficiency. However, the procedure associated with Abelian inversion in several sections of the interferogram is tedious and not very accurate. Below, we describe another much simpler way of calculating the optical-to-acoustic energy conversion efficiency that does not require Abelian inversion.

CALCULATION OF THE CONVERSION EFFICIENCY FROM INTERFEROMETRIC DATA INTEGRAL ALONG THE LINE OF OBSERVATION

Let us represent the pressure distribution over the acoustic wave front as $P_F(R) = P_{max}\psi(R)$, where $\psi(R)$ is a certain analytical function that has a maximum at $R = 0$ and decreases monotonically with increasing R. A number of approximating functions $\psi(R)$ are given in Table 2, and their shapes are depicted in Fig. 5. Figure 5 also contains data points found by Abelian inversion from the interferogram (Fig. 2a) for the section $z = 0$ and the wave front *F*–*F*'. Table 2 also lists the values of Ψ_S calculated by formula (8) for each of the approximating functions and divided by the area of a circle of radius *a*.

Table 1. Φ_t and $\Psi_S P^2_{max}$ calculated (*a*) with Abelian inversion and (*b*) from the integral data

ε_L, J/cm^2	Φ_t, ns		$\Psi_S P^2_{max}$, cm^2 atm^2		$n \times 10^7$	
	a	*b*	*a*	*b*	*a*	*b*
1.0	115 ± 10	130 ± 10	8.8 ± 2	10 ± 1	4.5 ± 1.5	5.7 ± 1.0
1.5	115 ± 10	160 ± 20	17.6 ± 2	17 ± 2	8.9 ± 2	12 ± 3
2.0	190 ± 20	160 ± 20	30 ± 3	31 ± 4	25 ± 5	22 ± 5
3.0	200 ± 20	190 ± 20	38 ± 5	43 ± 5	33 ± 8	36 ± 8

Table 2. $\Psi_S/\pi a^2$ and Γ_a calculated for various approximating functions

Curve no. in Fig. 5	$\psi(R)$	$\Psi_S/\pi a^2$	Γ_a
1	$\psi(R) = 1$ for $R \le a$, $\psi(R) = 0$ for $R \le a$	1.0	$\pi/4 = 0.79$
2	$\psi(R) = [1 - (R/a)^4]$ for $R \le a$	8/15	$5\pi/24 = 0.65$
3	$\psi(R) = [1 - (R/a)^2]$ for $R \le a$	1/3	$3\pi/16 = 0.59$
4	$\psi(R) = [1 - \lvert R/a \rvert]$ for $R \le a$	1/6	$\pi/6 = 0.52$
5	$\psi(R) = \exp[-(R/a)^2]$	0.5	0.50

As follows from Table 2, Ψ_S depends significantly on the approximating function. Specifically, its values for curves *1* and *4* differ sixfold.

Now, for each of the functions listed in Table 2, we calculate the maximal shift of the fringes at the axis of the interferogram, putting $z = 0$ and $y = 0$ in (3). Then,

$$K_0 = K(y = 0, z = 0) = 2AP_{\max}\int_0^{x'}\psi(x)dx, \qquad (11)$$

where $x' = +\infty$ for Gaussian function (5) and $x' = a$ for the other functions.

Having found $P_{\max}$ from (11), we come to an expression for $\Psi_S P^2_{\max}$ (this quantity enters into formula (6) for the efficiency):

$$\Psi_S P^2_{\max} = \frac{K_0^2}{A^2}\Gamma_a, \qquad (12)$$

where

$$\Gamma_a = \frac{\Psi_S}{4}\left[\int_0^a \psi(x)dx\right]^{-2}. \qquad (13)$$

The dimensionless coefficient Γ_a, whose values for various functions $\psi(R)$ are listed in the last column of Table 2, depends neither on $P_{\max}$ nor on *a* and varies with the approximating function only slightly. In particular, for functions (2)–(4), which approximate most closely the actual distributions found by Abelian inversion (Fig. 4b), the mean value of Γ_a is 0.59 (≈7%).

The calculation of $\Psi_S P^2_{\max}$ by formula (12) does not require Abelian inversion. The only parameter to be measured in the interferogram is the fringe shift K_0 on the acoustic wave front at the axis of symmetry. Strictly speaking, such a calculation is valid only if the wave front is flat, since formula (11) for K_0 ignores its curvature. Indeed, when the wave is illuminated along the line $z = 0$, integration is only over the flat part of the front from $-b$ to $+b$ (Fig. 3a and data points in Fig. 5); hence, its periphery does not contribute to $K_{0,b}$ measured from the interferogram.

The curvature of the wave front can be included if Γ_a in (12) is replaced by $\Gamma_{a,b} = (K_{0,a}/K_{0,b})^2\Gamma_a$, where $K_{0,a}$ and $K_{0,b}$ are integrals (11) calculated, respectively, for

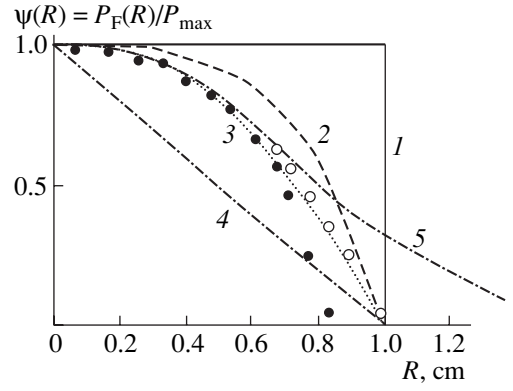

Fig. 5. (*1–5*) Analytical functions (see Table 2) approximating the experimental distribution $\psi(R)$ and experimental pressure distributions (●) over the section $z = 0$ and (○) over the wave front *F*–*F*' found from the interferogram in Fig. 2a.

Table 3. Comparison of experimental and theoretical values of η and P_{max}

ε_L, J/cm^2	$\eta \times 10^7$			P_{max}, atm			Conditions on liquid surface
	1	2	3	1	2	3	
1.0	5.1	1.5	4.7	3.0 ± 0.2	2.4 (1.4)	3.6 (2.2)	Free surface
1.5	10.5	2.25	10	4.2 ± 0.3	3.6 (1.8)	7.2 (3.6)	
2.0	24	3.0	21	6.0 ± 0.7	4.8 (1.9)	12 (4.8)	
3.0	35	4.5	57	7.5 ± 0.5	7.2 (2.2)	24 (7.2)	
0.5	260	100	300	8.4 ± 0.5	6.2	8.0	Rigid interface

Note: 1, experiment; 2, linear theory; 3, nonlinear theory.

$x' = a$ and $x' = b$. Then, formula (12) takes the form

$$\Psi_S P_{max}^2 \approx \frac{\Gamma_{a,b} K_{0,b}^2}{A^2}. \quad (14)$$

For parabolic function *3* in Table 2, the ratio of the integrals $K_{0,b}/K_{0,a} = (3.2)(b/a)[1 - (1/3)(b/a)^2]$. Putting $b/a = 0.8$ (which corresponds to the spatial distribution in Fig. 3a and data points in Fig. 5), we obtain $K_{0,a}/K_{0,b} \approx 1.06$ and $\Gamma_{a,b} \approx 0.66$.

The values of $\Psi_S P_{max}^2$ calculated by formula (14) for various energy densities are summarized in Table 1 (columns *b*). In our experiments, the energy density on the surface was increased by decreasing the laser spot size; as a result, as ε_L grows, the diffraction distortion of the wave front increases and the ratio b/a diminishes. Therefore, the value $\Gamma_{a,b} \approx 0.66$ was used only for ε_L = 1 and 1.5 J/cm^2. At higher energy densities, the ratio b/a was determined from the interferogram geometrically. For example, at ε_L = 2 and 3 J/cm^2, $b/a \approx 0.7$ and 0.6, respectively, which corresponds to $\Gamma_{a,b} \approx 0.76$ and 0.91. These values of $\Gamma_{a,b}$ were used in the calculations. As follows from Table 1, the values of $\Psi_S P_{max}^2$ calculated by formulas (10) and (14) always coincide within the experimental error (10–15% on average).

It should be noted that the function Φ_t, which is necessary to calculate the conversion efficiency by formula (6), can also be found with a reasonable accuracy directly from the shift of the fringe near the *z* axis of symmetry (associated data are given in Table 1). In this case, an extension of the first positive pulse due to the superposition of time-shifted peripheral pulses is compensated for by a decrease in the negative pulse amplitude, so that the values of Φ_t calculated by (9) remain within the experimental error with which the values of Φ_t were calculated using Abelian inversion in several sections of the interferogram (Table 1). In both cases, the error in determining Φ_t is 10–20 ns and is comparable to the time resolution of holographic interferometry, which is defined by the width of a ruby laser pulse (≈20 ns).

Obviously, our simplified method of calculating the conversion efficiency from the integral (along the line of observation) fringe shifts in an interferogram is the more accurate η, the smaller the wave front curvature due to diffraction divergence.

RESULTS AND DISCUSSION

The measured conversion efficiencies for the acoustooptic (thermal) interaction mechanism are summarized in Table 3, where the conversion efficiencies η are averaged over the data arrays picked up by the two measuring methods (Table 1). This table also lists the theoretical data for η calculated by the method described in [1] in the linear and nonlinear approximations, i.e., respectively, without and with considering the temperature dependence of the volume thermal expansion coefficient of water β.

As follows from Table 3, the measured values of η for the free boundary are in good agreement (within ≈10%) with the theoretical values of η calculated in view of the thermal nonlinearity. The only exception is the overestimated theoretical value $\eta = 57 \times 10^{-7}$ for ε_L = 3 J/cm^2. The reason for such a high value is likely to be the disregard of heat conduction in the theoretical model adopted, as well as the linear approximation of the temperature dependence $\beta(T)$. In fact, within our model, the water surface temperature had to rise to 100°C for ≈130 ns; that is, the evaporation had to start and the associated evaporation-induced pressure pulse had to appear immediately after the positive pulse. However, as follows from the interferogram in Fig. 2b, the evaporation-induced pressure pulse lags behind the beginning of the acoustooptic pulse by ≈300 ns, which indicates that the actual surface temperature at the instant the acoustooptic pulse forms was somewhat lower than that taken in the calculations. Note also that the evaporation threshold (≈1.5 J/cm^2) observed experimentally [2] during the CO_2 laser–water interaction is much higher than the energy density (0.4 J/cm^2) needed to heat the surface to 100°C. This is another demonstration that heat conduction should be taken into account.

Table 3 also lists the maximal pressure values P_{max} in the positive pulse that were calculated from interferograms using Abelian inversion, as well as the theoretical values of P_{max} calculated in the linear and nonlinear approximations (formulas (15) and (29) in [1]). At first glance, the experimental values of P_{max} are close to those found within the linear theory. However, this coincidence is accidental if we take into account that the pressure was measured at a large distance d from the surface, while the calculations were made for pressure pulses formed immediately on the surface. With regard for the fact that the distance $d = 7$ cm is comparable to the diffraction length $L_{DF} = a^2/2c_0\tau$ (which varies from 25 to 6 cm as the energy density varies from 1 to 3 J/cm^2), it becomes clear that the diffraction distortion of the pulse amplitude should be taken into account for the experimental and theoretical pressure values to be compared correctly. According to [7], with $d/L_{DF} \approx 1$ (or with the energy density $\approx$2.5 J/cm^2 for our case), the pulse amplitude at the wave axis decreases roughly threefold. The parenthesized figures in Table 3 are pulse amplitudes estimated at a distance of 7 cm from the surface with the diffraction effects included. In view of corrections for diffraction, the theoretical values obtained with allowance for the thermal nonlinearity are in fairly good agreement with P_{max} found experimentally.

The pattern is completed by the lowest row in Table 3, where the experimental and theoretical values of the conversion efficiency and maximal pressure in the acoustic pulse are listed for the case when the laser radiation acts on the rigid (water–germanium) interface. The experimental values were found by processing the interferogram in Fig. 2c, which was obtained in our earlier experiments. The laser radiation density at the water–germanium interface with regard for the radiation absorbed in the germanium was 0.5 J/cm^2, and the wave travel from the water surface was $d = 6$ mm. Since d is small, a correction for diffraction divergence was not introduced in this case. As follows from Table 3, in this case, too, the experimental data are in good agreement with the calculations allowing for the thermal nonlinearity.

In the experiments described, the interferograms of acoustic and shock waves were obtained in an energy density interval that was much wider than shown in Table. 3. However, at ε_L above 3 J/cm^2, the acoustooptic pressure pulse almost coincides with the evaporation-induced pulse. The evaporative mechanism of generating acoustic and shock waves in a liquid and the conversion efficiency for this mechanism will be considered in further publications.

ACKNOWLEDGMENTS

The author thanks I.I. Komissarova, V.N. Filippov, and E.N. Shvedova for their participation in the experiment.

This work was supported by the Russian Foundation for Basic Research (grant no. 00-15-96771, the school of thought of Academician Yu.N. Denisyuk).

REFERENCES

1. G. V. Ostrovskaya, Zh. Tekh. Fiz. **72** (10), 95 (2002) [Tech. Phys. **47**, 1299 (2002)].
2. I. I. Komissarova, G. V. Ostrovskaya, V. N. Filippov, and E. N. Shedova, Zh. Tekh. Fiz. **62** (2), 34 (1992) [Sov. Phys. Tech. Phys. **37**, 130 (1992)].
3. I. I. Komissarova, G. V. Ostrovskaya, V. N. Filippov, and E. N. Shedova, Zh. Tekh. Fiz. **67** (2), 138 (1997) [Tech. Phys. **42**, 247 (1997)].
4. G. V. Ostrovskaya and E. N. Shedova, Izv. Akad. Nauk, Ser. Fiz. **61**, 1342 (1997).
5. N. G. Preobrazhenskiĭ, in *Abel Inversion and Its Generalization* (Novosibirsk, 1978), pp. 6–24.
6. Yu. I. Ostrovskiĭ, G. V. Ostrovskaya, and M. M. Butusov, *Holographic Interferometry* (Nauka, Moscow, 1979).
7. V. É. Gusev and A. A. Karabutov, *Laser Optoacoustics* (Nauka, Moscow, 1991).

Translated by V. Isaakyan

Technical Physics, Vol. 47, No. 12, 2002, pp. 1554–1560. Translated from Zhurnal Tekhnicheskoĭ Fiziki, Vol. 72, No. 12, 2002, pp. 72–78.

RADIOPHYSICS

Transformation of Plane Wave Fields Due to a Parameter Jump in a Free Magnetodielectric Layer

O. N. Rybin, N. I. Slipchenko, and L. N. Shul'ga

National University of Radio Electronics, Kharkov, 61166 Ukraine

Received March 12, 2002

Abstract—The transformation of plane electromagnetic waves caused by a synchronous stepwise temporal change in the permittivity and permeability of a layer of a free anisotropic magnetodielectric medium is considered. No constraints are imposed on the layer thickness and the jump magnitude. The problem is solved analytically with the integral equation method. Exact expressions for the electric component of the field throughout the space after the jump are found. The expressions derived are analyzed from the physical viewpoint. *© 2002 MAIK "Nauka/Interperiodica".*

INTRODUCTION

The radiation and propagation of electromagnetic waves in media where the permittivity ε and permeability μ change stepwise in time have been analyzed in a variety of works (see, e.g., [1–4]). The importance of those papers for the development of modern radiophysics is difficult to overestimate. However, they considered parameter jumps in either infinite [1, 2] or semi-infinite [3, 4] media. Therefore, it would be logical to extend this study by introducing one more boundary, especially with regard for the fact that an insulating layer serves as the 1D model of a dielectric waveguide [5]. This makes the theoretical study of wave transformation in a magnetodielectric layer due to a step change in its parameters particularly important for further advances in present-day optoelectron transmission lines.

In this work, we study the transformation of plane electromagnetic waves (a plane monochromatic wave and a Gaussian pulse) in a magnetodielectric layer of a free isotropic nonconducting magnetodielectric medium. The layer is formed by a step change in the permittivity and permeability from ε_0 and μ_0, respectively, to ε_1 and μ_1 at the initial time instant $t = 0$ in the domain $x \in [0, d]$. It is assumed that the permittivity and permeability in the domains $x \in (-\infty, 0)$ and $x \in (d, +\infty)$ remain constant and equal to ε_0 and μ_0.

Let a *TM* wave with an electric component $E_0(t, x)$ be present in a homogeneous isotropic nonconducting free (unbounded) medium prior to the time instant $t = 0$. Such a statement allows one to readily reduce this (generally three-dimensional) problem to the one-dimensional case [6] and still makes it possible to reveal general laws governing the transformation of plane electromagnetic waves in isotropic nonconducting media with parameters that sharply vary in a layer of arbitrary size and location.

In our statement ($t > 0$), the electromagnetic field behavior in the domain $x \in [0, d]$ can be described by the Volterra integral equation of the second kind [7]. In this case, it takes the form

$$E(t, x) = E_0(t, x) + \int_0^{+\infty} dt' \int_0^d dx' K(t, t', x, x') E(t', x'), \tag{1}$$

where

$$K(t, t', x, x') = [1 - m^2]\delta(t - t')\delta(x - x') - \frac{1 - a^2 m^2}{2a^2} m^2 \frac{\partial}{\partial t} \delta(v_0(t - t') - |x - x'|) \tag{2}$$

is the difference kernel of integral equation (1). Here, $m = \sqrt{\mu_0/\mu_1}$, $a = \sqrt{\varepsilon_0/\varepsilon_1}$, $v_0 = c/\sqrt{\varepsilon_0\mu_0}$, c is the velocity of light in a vacuum, $\delta(t)$ is the delta function, and $\Theta(t)$ is the Heaviside function.

A solution to Eq. (1) in the domain $x \in [0, d]$ can be found by the resolvent method [8] using the integral

$$E(t, x) = E_0(t, x) + \int_0^{+\infty} dt' \int_0^d dx' R(t, t', x, x') F(t', x'), \quad t > 0, \tag{3}$$

where the resolvent $R(t, t', x, x')$ is the solution to the integral equation

$$R(t, t', x, x') = K(t, t', x, x') + \int_0^{+\infty} dt' \int_0^d dx'' K(t, t'', x, x'') R(t'', t', x'', x') F(t', x'). \tag{4}$$

The construction of the resolvent $R(t, t', x, x')$ satisfying integral equation (4) is a complicated problem

1063-7842/02/4712-1554$22.00

consisting of two steps. First, a resolvent neglecting the effect of boundaries is constructed; then, the part describing the contribution of the field due to the boundaries $x = 0$ and $x = d$ appearing at the time instant $t = 0$ is added. In other words,

$$R(t, t', x, x') = \lfloor R_1(t, t', x, x') + R_2^{M, N}(t, t', x, x') \rfloor \Theta(x')\Theta(d - x'). \quad (5)$$

The expression for the resolvent in the boundary-free problem has been derived previously [4]:

$$R(t, t', x, x') = \left[1 - \frac{1}{m^2}\right]\delta(t - t')\delta(x - x') - \frac{1 - a^2m^2}{2m^2}\frac{\partial}{\partial t}\delta(v_1(t - t') - |x - x'|), \quad (6)$$

where $v_1 = c/\sqrt{\varepsilon_1\mu_1}$.

An expression for the second part of the resolvent, $R_2^{M, N}(t, t', x, x')$, is constructed in a way similar to that used in [7], where only the permittivity of the layer experienced a jump (Fig. 1). Figure 1 depicts schematically the layer appearing at the zero time instant. It is split into zones indexed by two figures. The figures indicate the number of rereflections for the world lines of backward (first figure) and forward (second figure) waves existing prior to a corresponding time instant in the layer ($t > 0$).

Thus, according to the rereflection principle [7], an expression for the second part of the resolvent, $R_2^{M, N}(t, t', x, x')$, can be found in the form

$$R_2^{M, N}(t, t', x, x') = -\frac{1 - a^2m^2}{2m^2}\frac{\partial}{\partial t}\left\{\sum_{l=1}^{M}\left(\frac{1 - a^2m^2}{1 + a^2m^2}\right)^l \times \delta(v_1(t - t') - x + (-1)^l(x' - d/2) - (l - 1/2)d) + \sum_{s=1}^{N}\left(\frac{1 - a^2m^2}{1 + a^2m^2}\right)^s \delta(v_1(t - t') + x + (-1)^{s+1}(x' - d/2) - (s + 1/2)d)\right\}. \quad (7)$$

It is clear (see Fig. 1) that the integer variables M and N in (7) must meet the inequality

$$|M - N| \le 1. \quad (8)$$

Also, Fig. 1, in view of inequality (8), implies that any set of the integer variables M and N corresponds to a certain space–time domain of the layer. Mathematically, this can be represented by the set of inequalities

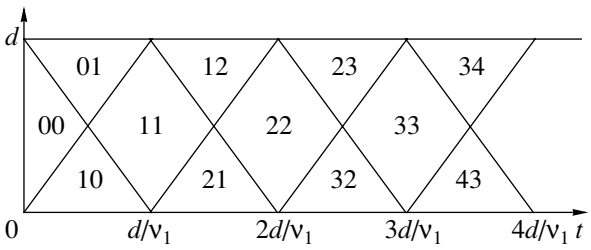

Fig. 1. Computational algorithm.

$$M, N = N - 1, N \Rightarrow \begin{cases} x > v_1(t - (N - 1)d/v_1) \\ x > -v_1(t - (N - 1)d/v_1) + d; \end{cases}$$

$$M, N = M, M - 1 \Rightarrow \begin{cases} x < v_1(t - (N - 1)d/v_1) \\ x < -v_1(t - (N - 1)d/v_1) + d; \end{cases} \quad (9)$$

$$M, N = M, M = N, N \Rightarrow \begin{cases} x < v_1(t - (N - 1)d/v_1) \\ x < -v_1(t - Nd/v_1) + d \\ x > -v_1(t - (N - 1)d/v_1) + d \\ x > v_1(t - Nd/v_1). \end{cases}$$

FIELD IN THE DISTURBED LAYER

Let the initial (primary) field be generated by a plane monochromatic wave of unit amplitude

$$E_0(t, x) = e^{i(\omega_0 t - kx)}, \quad (10)$$

where $k = \omega_0/v_0$ is the wave number and ω_0 is the frequency of the primary wave.

Substituting (5)–(7) into (10) and (3) yields an expression ($t > 0$) for the electrical component of the field in the layer $x \in [0, d]$ arising at the zero time instant. This expression has arbitrary values of the integer variables M and N, which specify space–time domains of the layer:

$$E^{M, N}(t, x) = E_1(t, x) + E_2^{M, N}(t, x). \quad (11)$$

Here,

$$E_1(t, x) = \Theta(-t + x/v_1)A_0^+ e^{i(am\omega t - kx)} + \Theta(-1 + (d - x)/v_1)A_0^- e^{-i(am\omega t + kx)} + \Theta(t - x/v_1)B_0^+ e^{i(\omega t - kx/am)} + \Theta(t + (x - d)/v_1)B_0^- e^{i(\omega t + kx/am)},$$

where

$$A_0^{\pm} = \pm\frac{1 \pm am}{2m^2}am, \quad B_0^{\pm} = \frac{1 \pm am}{2m^2},$$

and

$$E_2^{M,N}(t, x) = \sum_{l=1}^{M}\{\Theta(-t + (x + 2ld)/v_1) \times \Theta(t - (x + (2l-1)d)/v_1)A_l^+ e^{i(am\omega t - k(x-2ld))} + \Theta(t - (x + 2ld)/v_1)B_l^+ e^{i(\omega t - k(x-2ld)/am)}\} + \sum_{s=1}^{N}\{\Theta(-t - (x + (2s+1)d)/v_1) \times \Theta(t + (x - 2sd)/v_1)A_s^- e^{-i(am\omega t + k(x-2sd))} + \Theta(t + (x - (2s+1)d)/v_1)B_s^- e^{i(\omega t + k(x-(2s+1)d)/am)}\},$$

where

$$A_j^{\pm} = A_0^{\pm}\left(\frac{1-am}{1+am}\right)^j, \quad B_j^{\pm} = B_0^{\pm}\left(\frac{1-am}{1+am}\right)^j.$$

Hereafter, the waves appearing in the formulas will be designated according to their amplitudes $A_j^{\pm}$ and $B_j^{\pm}$.

The structure of Eq. (11) reflects the general picture of transforming the primary monochromatic wave $E_0(t, x)$ in the layer $x \in [0, d]$ where the permittivity and permeability change stepwise. At the instant of the jump, the primary wave is split into opposing forward, A_0^+, and backward, A_0^-, monochromatic waves that have new frequencies ($am\omega$) and amplitudes compared with those of the primary wave. Furthermore, two plane-parallel interfaces separating the three space domains appear at the time instant $t = 0$. Therefore, either of the waves A_0^+ and A_0^-, partially reflecting from the interfaces $x = 0$ and $x = d$, initiate respective backward, A_1^-, and forward, A_1^+, waves. The latter, rereflecting from the interfaces, generate respective forward, A_2^+, and backward, A_2^-, waves, etc. Also, at the instant the disturbed layer appears jumpwise, a forward monochromatic wave B_0^+ appears in it. This wave, which is the transmitted primary wave with a new value of the wave number (k/am), partially reflects from the interface $x = d$ and gives rise to a backward monochromatic wave B_0^-. When striking the interface $x = 0$, the wave B_0^- initiates the process of generating pairs of respective forward and backward waves B_s^+ and B_s^- ($s = 1, \ldots, N$).

Of interest is an expression for the electric component of the field under steady-state conditions. It can be derived from (11) if the superscript M for the zone MM (or N for the zone NN) approaches infinity. The corresponding limit on the right of (11) can easily be found if it is taken into account that the space–time domains where the waves $A_j^{\pm}$ ($j = 1, 2, \ldots$) exist are bounded by plane-parallel boundaries moving in the same direction with the same velocities. (This directly follows from the forms and structures of the arguments of the Heaviside functions standing as factors by the amplitudes $A_j^{\pm}$ in formula (11).) These moving planes asymptotically go to infinity, thereby leaving the finite domain of the space free of the field due to the forward and backward waves $A_j^{\pm}$. Moreover, the one-sided boundaries of the fields induced by the waves $B_j^{\pm}$ ($j = 0, 1, \ldots$), which also go to infinity, asymptotically "make up" the fields of the two waves in the layer. The amplitudes of these waves are easy to find by using the formula for the sum of an infinite geometric progression [9] (in this case, the inequality $(1 - am)/(1 + am) < 1$ is fulfilled). With the aforesaid in mind, we arrive at an expression for the transformed component of the electric field in the layer under steady-state conditions:

$$\lim_{M \to +\infty} E^{M,M}(t, x) = A^+ e^{i(\omega t - kx/am)} + A^- e^{i(\omega t + k(x-d)/am)}, \tag{12}$$

where

$$A^{\pm} = \frac{1 \pm am}{2m^2}\frac{1}{1 + am - (1 - am)e^{\pm 2ikd/am}}.$$

Thus, a jump of the permittivity and permeability in a layer of a magnetodielectric space transforms the field of a plane monochromatic wave into a field that is, under steady-state conditions, the superposition of the fields of forward and backward monochromatic waves with new values of the wave number and amplitude.

Now let us derive an expression for the transformed component of the electric field in the domain $x \in [0, d]$ (in the same layer created stepwise) when the primary field has the form of a Gaussian electromagnetic pulse

$$E_0(t, x) = e^{-(\omega_0 t - kx)^2}, \quad k = \omega_0/v_0. \tag{13}$$

Substituting (5)–(7) into (13) and (3) and integrating the Gaussian with the formula for differentiation of parameter-depending integral [10] yields

$$E^{M,N}(t, x) = E_1(t, x) + E_2^{M,N}(t, x), \tag{14}$$

where

$$E_1(t, x) = \Theta(-t + x/v_1)A_0^+ e^{-(am\omega t - kx)^2} + \Theta(-t + (d - x)/v_1)A_0^- e^{-(am\omega t + kx)^2} + \Theta(t - x/v_1)B_0^+ e^{-(\omega t - kx/am)^2} + Q(t + (x - d)/v_1)B_0^- e^{-(\omega t + kx/am)^2},$$

$$E_2^{M,N}(t, x) = \sum_{l=1}^{M} \{\Theta(-t + (x + 2ld)/v_1) \times \Theta(t - (x + (2l - 1)d)/v_1)A_l^+ e^{-(am\omega t - k(x - 2ld))^2} + \Theta(t - (x + 2ld)/v_1)B_l^+ e^{-(\omega t - k(x - 2ld)/am)^2}\} + \sum_{s=1}^{N} \{\Theta(-t - (x + (2s + 1)d)/v_1) \times \Theta(t + (x - 2sd)/v_1)A_s^+ e^{-(am\omega t + k(x - 2sd))^2} + \Theta(t + (x - (2s + 1)d)/v_1)B_s^+ e^{-(\omega t + k(x - (2s + 1)d)/am)^2}\},$$

and the expressions for the amplitude coefficients coincide completely with those in formula (11).

The structures of expressions (11) and (14) coincide completely. Thus, we can conclude that the transformation of a plane monochromatic electromagnetic wave is physically totally identical to that of a Gaussian electromagnetic pulse when the parameters of a layer of a free magnetodielectric space undergo a step change.

With the approach used in deriving Eq. (12), we obtain an expression for the field electric component in (14) under steady-state conditions:

$$E^{M,M}(t, x) \overset{M \gg 1}{\cong} \sum_{l=0}^{M} \{B_l^+ e^{-(\omega t - k(x - 2ld)/am)^2} + B_l^- e^{-(\omega t + k(x - (2l + 1)d)/am)^2}\}. \quad (15)$$

Even if the structure of field (12) for each of the subdomains *MM* is retained, the expression for field (15) cannot be reduced to the form of (12), because the sum in (15) cannot be found with the formula for the sum of an infinite geometric progression.

FIELD IN UNDISTURBED DOMAINS OF THE SPACE

Of interest also are expressions for the fields outside the disturbed layer $x \in [0, d]$. With regard for the structure of expressions (11) and (14) for the fields, one may conclude that each of the components of these fields interacts with one of the two interfaces dividing the space into three domains. Therefore, expressions for the fields in the domains $x \le 0$ and $x \ge \alpha$ ($t > 0$) can be derived if the interaction of each plane component of fields (11) and (14) with these interfaces is considered. First, we will find a solution to the problem of diffraction of a nonsteady plane wave given in the general case by

$$E_0(t, x) = E_0(t - (x - \beta)/v_1), \quad v_1 = c/\sqrt{\varepsilon_1\mu_1}, \quad (16)$$

which propagates in medium *1* with a velocity v_1 normally to the interface $x = \beta$ between media *1* and *2* (the velocity of the plane wave in medium *2* is $v_2 = c/\sqrt{\varepsilon_2\mu_2}$). It is assumed that both media are homogeneous, isotropic, and nonconducting and have the permittivity and permeability ε_j and μ_j ($j = 1, 2$). The mathematical statement of the diffraction problem has the form

$$\left(\frac{\partial^2}{\partial x^2} - \frac{1}{v_1^2}\frac{\partial^2}{\partial t^2}\right)E(t, x) = 0, \quad x > \beta;$$

$$\left(\frac{\partial^2}{\partial x^2} - \frac{1}{v_2^2}\frac{\partial^2}{\partial t^2}\right)E(t, x) = 0, \quad x < \beta; \quad (17)$$

$$E(t, x)|_{x = \beta + 0} = E(t, x)|_{x = \beta - 0},$$

$$\left.\frac{\partial E(t, x)}{\partial x}\right|_{x = \beta + 0} = \left.\frac{\partial E(t, x)}{\partial x}\right|_{x = \beta - 0}.$$

Using direct Fourier transformation, we write the function of the primary wave as

$$E_0(t - (x - \beta)/v_1) = \int_{-\infty}^{+\infty} d\omega E_0(\omega) e^{i\omega(t - (x - \beta)/v_1)} = \int_{-\infty}^{+\infty} d\omega E_0(\omega) e^{i(\omega t - k_1(x - \gamma_1))}, \quad (18)$$

where $k_1 = \omega/v_1$ and $\gamma_1 = \beta/v_1$.

Now, using the condition for radiation [11], we write the expressions for the transmitted, $E_{tr}(t, x)$, and reflected, $E_{ref}(t, x)$, waves in the form of (18):

$$E_{tr}(t, x) = \int_{-\infty}^{+\infty} d\omega E_{tr}(\omega) e^{i(\omega t - k_2(x - \gamma_2))}, \quad (19)$$

$$E_{ref}(t, x) = \int_{-\infty}^{+\infty} d\omega E_{ref}(\omega) e^{i(\omega t - k_1(x - \gamma_1))}, \quad (20)$$

where $k_2 = \omega/v_2$ and $\gamma_2 = \beta/v_2$.

According to the principle of superposition [11], the solution to problem (17) has the form

$$E(t, x) = \int_{-\infty}^{+\infty} d\omega e^{i\omega t} \times \begin{cases} E_0(\omega)e^{-ik_1(x-\gamma_1)} + E_{\text{ref}}(\omega)e^{ik_1(x-\gamma_1)}, & x > \beta \\ E_{\text{tr}}(\omega)e^{-ik_2(x-\gamma_2)}, & x < \beta. \end{cases} \quad (21)$$

To find the Fourier transforms $E_{\text{tr}}(\omega)$ and $E_{\text{ref}}(\omega)$, we substitute expression (21) into the boundary conditions for the set of differential equations (17) to obtain a set of linear algebraic equations of second order with respect to the Fourier transforms. The solution to this set is given by

$$E_{\text{tr}}(\omega) = \frac{2}{1 + \sqrt{\varepsilon_1\mu_1/\varepsilon_2\mu_2}} E_{\text{tr}}(\omega) = R,$$

$$E_{\text{ref}}(\omega) = \frac{1 - \sqrt{\varepsilon_1\mu_1/\varepsilon_2\mu_2}}{1 + \sqrt{\varepsilon_1\mu_1/\varepsilon_2\mu_2}} E_{\text{tr}}(\omega) = T,$$

where R is the reflection coefficient for the Fourier transform of the plane wave reflecting from the interface $x = \beta$ between media *1* and *2* and T is the transmission coefficient of the Fourier transform of the plane wave passing through the interface $x = \beta$.

To obtain the components $E_{\text{tr}}(t, x)$ and $E_{\text{ref}}(t, x)$ of the electric field, we now consider the electric field component as a function of three variables:

$$E(t - (x - \beta)/v_j) = E(t, x, v_j); \quad j = 1, 2. \quad (23)$$

Then, using direct Fourier transformation and taking into account that the coefficients R and T are independent of the variable of integration ω, we come to

$$E_{\text{tr}}(t, x) = \int_{-\infty}^{+\infty} d\omega T E_0(\omega) e^{i(\omega t - k_2(x-\gamma_2))} = T\int_{-\infty}^{+\infty} d\omega E_0(\omega) e^{i\omega(t-(x-\beta)/v_2)} = TE(t, x, v_2), \quad (24)$$

$$E_{\text{ref}}(t, x) = \int_{-\infty}^{+\infty} d\omega R E_0(\omega) e^{i(\omega t + k_1(x-\gamma_1))} = R\int_{-\infty}^{+\infty} d\omega E_0(\omega) e^{i\omega(t-(x-\beta)/(-v_1))} = RE(t, x, -v_1). \quad (25)$$

Thus, to find expressions for the electric field components in the undisturbed domains of the space, one must be sure that formulas (24) and (25) are applicable to individual terms of expressions (11) and (14). Such an approach is valid, since both a plane monochromatic wave and a Gaussian pulse can be represented in the form of (16). Therefore, we will apply formulas (24) and (25) to fields (11) and (14).

In order to find an expression for the primary wave reflected from the disturbed layer into the domain $x < 0$ (wave *1*), one should apply formula (25) to the primary wave functions in (10) and (13). Next, in order to find an expression for the primary wave transmitted into the domain $x > d$ (wave Q_0^+), one should apply formula (24) to the expressions for the waves B_0^+ and B_0^- in (11) and (14). Formula (24) should also be applied to the expressions for the waves A_s^- and B_s^- ($s = 0, \ldots, N$) in order to find expressions for these waves transmitted into the domain $x < 0$ (waves P_s^- and Q_s^-, respectively). Finally, formula (24) should be applied to the expressions for the waves A_l^+ and B_l^+ ($l = 1, \ldots, M$) in order to find expressions for these waves transmitted into the domain $x > d$ (waves P_l^+ and Q_l^+, respectively).

Before we write expressions for the fields in the undisturbed domains, it should be noted that in expression (25) for the primary wave reflected from the interface $x = 0$, parameters with the subscripts 1 and 2 refer, respectively, to the undisturbed domains and the disturbed layer. For the other waves, these subscripts in expressions (24) and (25) should be understood in the reverse sense.

With the aforesaid in the last two paragraphs, expressions for the electric field component in the undisturbed domains of the magnetodielectric space are as follows:

$$E^N(t, x) = \Theta(t + x/v_0)IF(\omega t + kx) + \Theta(-t + (d - x)/v_0)P_0^-F(-am(\omega t + kx)) + \Theta(t + (x - d)/v_0)Q_0^-F(\omega t + k(x - d)) + \sum_{s=1}^{N}\{\Theta(-t + ((2s + 1)d - x)/v_0) \times \Theta(t + (x - 2sd)/v_0)F(-am(\omega t + k(x - 2sd))) + \Theta(t + (x - (2s + 1)d)/v_0) \times F(\omega t + k(x - (2s + 1)d))\}, \quad x < 0, \quad (26.1)$$

$$E^N(t, x) = \Theta(-t + x/v_0)P_0^+F(iam(\omega t - kx)) + \Theta(t - x/v_0)Q_0^+F(\omega t - kx) + \sum_{l=1}^{M}\{\Theta(-t + (x + 2l)/v_0) \times \Theta(t - (x + (2l + 1))/v_0)P_l^+F(am(\omega t - k(x - 2ld)))$$

$$+\Theta(t-(x+2ld)/v_0)$$
$$\times Q_l^+ F(\omega t - k(x-2ld))\},\quad x<d,$$
$$F(z)=\begin{cases}e^{iz}, & E_0(t,x)=e^{i(\omega t-kx)}\\ e^{-z^2}, & E_0(t,x)=e^{-(\omega t-kx)^2},\end{cases} \tag{26.3}$$

where

$$I=\frac{1-am}{1+am},\quad P_j^+=a^2\left(\frac{1-am}{1+am}\right)^j,$$
$$P_j^-=-a^2\left(\frac{1-am}{1+am}\right)^{j+1},$$
$$Q_j^+=\frac{a}{m}\left(\frac{1-am}{1+am}\right)^j,\quad Q_j^-=\frac{a}{m}\left(\frac{1-am}{1+am}\right)^{j+1}.$$

It is also of obvious interest to find long-term approximations of expressions (26), i.e., to find fields (26) under steady-state conditions. For a plane monochromatic primary wave, the transition from expressions (26.1) and (26.2) to their long-term approximations is based on the same reasoning as in going from (11) to (12). Note only that in this case, the waves P_s^- ($s=1,\ldots,N$) and P_l^+ ($l=1,\ldots,M$) asymptotically go to infinity; hence, in the finite domain of the space, the electric component of the field under steady-state conditions is given by

$$E(t,x)=\begin{cases}\lim\limits_{M\to+\infty}E^N(t,x)=(I+W^-)e^{i(\omega t+kx)}, & x<0\\ \lim\limits_{M\to+\infty}E^N(t,x)=W^+e^{i(\omega t-kx)}, & x>d,\end{cases} \tag{27}$$

where

$$W^\pm=\frac{a}{m}\frac{1\pm am}{1+am-(1-am)e^{\pm 2ikd}}.$$

Comparing expressions (12) and (27) shows that the latter can be derived from the former directly by applying to (12) formulas (24) and (25).

If a primary wave is a Gaussian pulse, the long-term approximations of expressions (26.1) and (26.2) have the form

$$E(t,x)$$
$$=\begin{cases}E^N(t,x)\overset{N\gg 1}{\cong}Ie^{-(\omega t+kx)^2}+\sum\limits_{s=0}^{N}Q_s^-e^{-(\omega t+k(x-(2s+1)d))^2},\\ x<0\\ E^M(t,x)\overset{M\gg 1}{\cong}\sum\limits_{l=0}^{M}Q_l^+e^{-(\omega t-k(x-2ld))^2},\quad x>d.\end{cases} \tag{28}$$

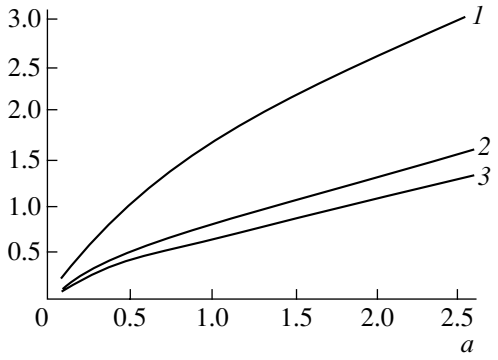

Fig. 2. Absolute value of the amplitude of the primary wave transmitted into the domain $x > d$ vs. relative change in the permittivity a for relative changes in the permeability m = 0.7 (*1*), 1.2 (*2*), and 1.4 (*3*) at $kd = 10^6$.

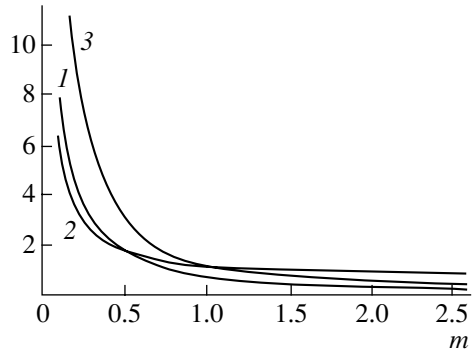

Fig. 3. Absolute value of the amplitude of the primary wave transmitted into the domain $x > d$ vs. relative change in the permeability m for relative changes in the permittivity a = 0.6 (*1*), 1.1 (*2*), and 1.3 (*3*) at $kd = 10^6$.

As in the case of a plane monochromatic primary wave, this result can be found directly by applying formulas (24) and (25) to expression (15).

Because of the simplicity of expressions (27), it is worth comparing the absolute value of the forward wave amplitude in the domain $x > d$ with the primary wave amplitude. The associated numerical results for a monochromatic primary wave are shown in Figs. 2 and 3. For certain values of relative changes in the permittivity and permeability, disturbed layer thickness, and primary wave frequency, the absolute value of the transmitted primary wave amplitude is seen to be enhanced.

CONCLUSION

In this work, we considered the transformation of a plane monochromatic electromagnetic wave and a Gaussian pulse for the case when the permittivity and permeability undergo a jump in a layer of a free (unbounded) magnetodielectric isotropic space. At the instant both parameters change stepwise, two layers form in the layer. One of them and the primary wave

copropagate, while the other moves in the opposite direction. The new waves and the primary wave interact with the two planar interfaces dividing the space into three domains and generate sets of forward and backward waves throughout the space. These waves are of two types: those induced by the parameter jump in the layer and those initiated when the jump-induced waves and the primary wave cross the boundaries of the layer. For each of the waves, the domains of existence are bounded by one or two plane-parallel boundaries moving in the opposite directions. The velocity of these boundaries depends on the material parameters in domains where they propagate. Once all the moving parallel planes have asymptotically gone to infinity, the finite domain of the space is free of the jump-induced waves. The numerical study of the amplitude of a plane monochromatic primary wave transmitted through the nonsteady layer suggests that a monochromatic primary wave can be amplified within our model of disturbance of a free magnetodielectric space.

REFERENCES

1. S. V. Afanas'ev, Izv. Vyssh. Uchebn. Zaved., Radiofiz. **34**, 605 (1991).
2. N. I. Slipchenko, L. N. Shul'ga, and O. N. Rybin, Radioélektron. Informatika **12** (3), 22 (2000).
3. F. R. Morgenthaler, IRE Trans. Microwave Theory Tech. **MTT-6**, 167 (1958).
4. N. I. Slipchenko, L. N. Shul'ga, and O. N. Rybin, Zh. Tekh. Fiz. **71** (10), 123 (2001) [Tech. Phys. **46**, 1326 (2001)].
5. B. V. Nikol'skiĭ and T. I. Nikol'skaya, *Electrodynamics and Radio-Wave Propagation* (Nauka, Moscow, 1989).
6. M. B. Vinogradova, O. V. Rudenko, and A. P. Sukhorukov, *Theory of Waves* (Nauka, Moscow, 1990).
7. A. G. Nerukh and N. A. Khizhnyak, *Modern Problems of Nonstationary Macroscopic Electrodynamics* (NPO Test-Radio, Kharkov, 1991).
8. V. S. Vladimirov, *Equations of Mathematical Physics* (Nauka, Moscow, 1976; Marcel Dekker, New York, 1971).
9. F. P. Yaremchuk and P. A. Rudchenko, *Algebra and Elementary Functions: A Handbook* (Naukova Dumka, Kiev, 1987).
10. G. M. Fikhtengol'ts, *Course of Integral and Differential Calculus* (Gostekhizdat, Moscow, 1948), Vol. II.
11. A. Sommerfeld, *Vorlesungen über theoretische Physik*, Bd. 4: *Optik* (Dieterich, Wiesbaden, 1950; Inostrannaya Literatura, Moscow, 1953).

Translated by V. Isaakyan

Technical Physics, Vol. 47, No. 12, 2002, pp. 1561–1569. Translated from Zhurnal Tekhnicheskoĭ Fiziki, Vol. 72, No. 12, 2002, pp. 79–87.

RADIOPHYSICS

Open Resonators for Low Dielectric Loss Measurements

S. N. Vlasov, E. V. Koposova, S. E. Myasnikova, and V. V. Parshin

Institute of Applied Physics, Russian Academy of Sciences, Nizhni Novgorod, 603950 Russia

e-mail: vlasov@hydro.appl.sci-nnov.ru

Received April 12, 2002

Abstract—Open resonators are studied based on the scattering theory. In particular, a theory of open resonators with a dielectric plate is constructed. This theory is used to develop a technique for measuring small dielectric loss. Measurements of the loss tangent in diamond plates show that the volume absorption in the best samples is $\approx 7 \times 10^{-6}$ in the high-frequency part of the millimeter-wave range. A significant absorption in thin surface layers that is caused by surface treatment techniques (polishing, vacuum soldering, etc.) is revealed. © *2002 MAIK "Nauka/Interperiodica".*

INTRODUCTION

Open resonators have long been used in various facilities that measure the complex permittivity in the millimeter- and sub-millimeter-wave ranges [1–3]. However, ultra-low-loss materials recently developed for power electronics, such as high-purity silicon [4] and particularly CVD diamond [5], have sent researchers to revise the existing concepts of open resonator operation.

When developing windows through which the energy of megawatt millimeter-wave generators (gyrotrons) is extracted, it was found that thin surface layers of a diamond disk (used as the window through which the microwave energy is extracted) feature an increased absorption.

These thin (of thickness on the order of the surface roughness) layers have been shown to appear as a result of surface treatment by various mechanical and/or chemical techniques, such as polishing, chemical etching, etc. The contamination is the most severe when the disk is soldered to a metal fixture. For instance, after soldering a diamond disk to a metal waveguide and then to a cooler to create energy-extraction windows, an excessive absorption arises in the thin surface layer, which is often higher than that in the disk volume. This excessive absorption breaks the thermal conditions of the window to such an extent that it fails. We will not discuss here the nature of the excessive absorption. Note only that the problem of how to eliminate it is yet to be resolved.

In this paper, we describe a method for measuring extremely low dielectric losses in high-quality CVD diamond plates with the help of Fabri–Perot open resonators [5–7]. The method relies on the refined theory of open resonators [8] that allows for dispersion introduced into a resonator by a plane-parallel dielectric plate several wavelengths thick with highly absorbing thin surface layers.

The eigenfrequencies of the resonator modes are found by analyzing the field structure at the exit of a resonator excited by a Gaussian beam. The frequencies and the Q factors of the eigenmodes are determined with allowance for the absorption in the plate and the coupling losses in the mirrors. The frequencies and the Q factors of the eigenmodes obtained by numerical analysis are used to justify simplified analytical formulas that relate the measured Q factors to the losses in the plate.

1. DESCRIPTION OF A RESONATOR WITH A PLANE-PARALLEL DIELECTRIC PLATE

Consider the resonator illustrated in Fig. 1. It is formed by nontransparent metallic mirrors with a radius of curvature R_0 and diameter $2a_m$ that are L apart. The ohmic reflection coefficient of both mirrors is $R_{\text{ohm}} \approx -1$.

A coupling element (dielectric film) with a reflection coefficient jR (where j is the imaginary unit) and transmission coefficient T makes an angle of 45° with the resonator axis. The coupling film divides the resonator into two arms with lengths L_1 and L_2 ($L_1 + L_2 = L$). With such a description, the film can be viewed as a four-port network. Ports *1* and *2* are directed toward the dielectric plate and mirror II, respectively. Radiation is fed to the resonator through port *3* and is extracted through ports *3* (reflection from the resonator) and *4* (transmission through the resonator).

We assume that the coupling film is lossless and dispersion-free; therefore, for real R and T,

$$R^2 + T^2 = 1.$$

The distance from the coupling film is measured from its center. The projection of the film onto the plane perpendicular to the resonator axis is a circle of diameter $2a_{\text{ex}}$. Arm *1* contains a plane-parallel dielectric plate of diameter $2a_{\text{pl}}$, thickness L_{pl}, and refractive index $n_{\text{pl}} = n\sqrt{1 - j\tan\delta_{\text{pl}}}$, where n is real and $\tan\delta_{\text{pl}}$ is the loss tangent of the dielectric material.

1063-7842/02/4712-1561$22.00

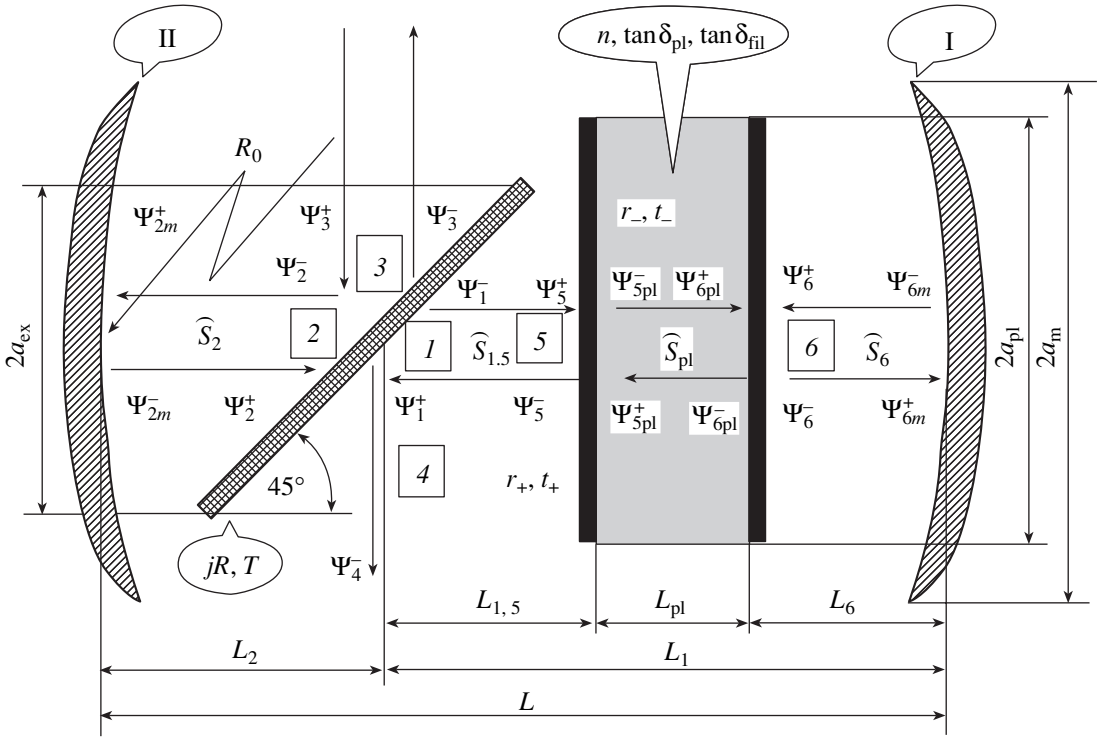

Fig. 1. Resonator with a dielectric plate: I and II are mirrors.

The planar surfaces of the plate are assumed to be covered by thin films (Fig. 2) of thickness L_{fil} and refractive index $n_{fil} = n\sqrt{1 - j\tan\delta_{fil}}$, where $\tan\delta_{fil}$ is the loss tangent of the films. The plate is placed at a distance L_6 from the first mirror perpendicular to the resonator axis and at a distance $L_{1,5}$ (between points *1* and *5*) from the coupling film ($L_{1,5} + L_{pl} + L_6 = L_1$). The distance between the coupling film and the left film on the dielectric plate is $\tilde{L}_{1,5} = L_{1,5} - L_{fil}$, and the distance between the right film on the dielectric plate and the first mirror is $\tilde{L}_6 = L_6 - L_{fil}$.

The fields in a resonator with a dielectric plate will be described in terms of the theory of open structures proposed in [9, 10]. In this theory, the fields varying with time as $\exp(j\omega t)$, where ω is the circular frequency, are assumed to have transverse (relative to the propagation direction) components. We use linearly polarized electric fields with amplitudes denoted by $\Psi_i^{\pm}$. In each cross section of the resonator, the fields are represented as two counterpropagating wave beams (Fig. 1). An input wave beam with an amplitude Ψ_3^+ impinges on the coupling film and excites the resonator. Output wave beams with amplitudes Ψ_3^- and Ψ_4^- leave the film and go out of the resonator.

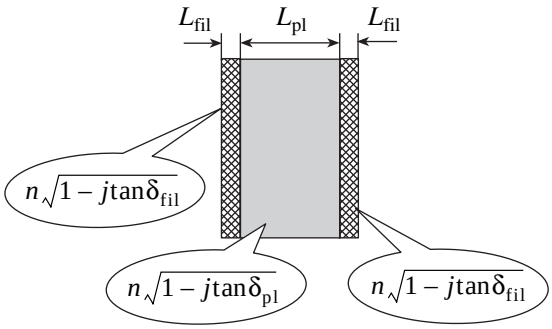

Fig. 2. Surface layers on the dielectric plate.

Let us denote the amplitudes of the beams that travel from the coupling film towards the first and second mirrors by $\Psi_{1,2}^-$, respectively; the amplitudes of the beams that come to the film from these directions, by $\Psi_{1,2}^+$. Equations that relate the wave beams on the coupling film have the form

$$\Psi_1^- = T\Psi_2^+,$$

$$\Psi_2^- = T\Psi_1^+ + jR\Psi_3^+,$$

$$\Psi_3^- = jR\Psi_2^+.$$

This formulation neglects beam diffraction near the coupling film.

The variation of the beam amplitudes $\Psi(x, y, z)$ with the transverse coordinates x and y and longitudinal

coordinate z as the beams propagate in the resonator is described by a relationship that follows from the representation of the field as the superposition of plane waves. This relationship will be written in detail for the beam propagating towards the second mirror. It has the form

$$\Psi_{2m}^{+} = \hat{S}_2\Psi_2^{-} = \hat{\Phi}^{\mathrm{i}}(k_x, k_y)S_{L_2}\hat{\Phi}^{\mathrm{f}}(x, y)\Psi_2^{-}, \quad (1)$$

where Ψ_{2m}^{+} is the amplitude of the beam incident on the second mirror, S_2 is the propagation operator, $\hat{\Phi}^{\mathrm{f}}(x, y)$ and $\hat{\Phi}^{\mathrm{i}}(k_x, k_y)$ are the operators of forward and inverse double Fourier transformations, $S_{L_2} = \exp(-jk_zL_2)$, k_zL_2 is the phase advance,

$$k_z = \sqrt{k^2 - k_x^2 - k_z^2} \quad (2)$$

is the longitudinal wave number of the plane wave, $k = \omega/c = 2\pi f/c$, c is the velocity of light, f is frequency, and k_x and k_y are the transverse wave numbers.

We will describe resonator mirrors as phase correctors, so that the amplitude Ψ_{2m}^{-} of the beam that leaves the second mirror is related to the amplitude of the beam incident on this mirror as

$$\Psi_{2m}^{-} = \hat{R}_0\Psi_{2m}^{+}$$
$$= R_{\mathrm{ohm}}\exp\left\{-2jk\left[R_0\sqrt{1 - \frac{x^2 + y^2}{R_0^2}} - R_0\right]\right\}\Psi_{2m}^{+}. \quad (3)$$

In formula (3), positive values of R_0 correspond to a concave mirror and $\hat{R}_0$ denotes the operator of beam transformation by the mirror.

Relationships similar to (1) can be written for the amplitudes of wave beams propagating in each section of the resonator. The variation of the beam structure as it travels from the coupling film to the left end face of the dielectric plate is described by the operator $S_{1,5}$,

$$\Psi_5^{+} = \hat{S}_{1,5}\Psi_1^{-} = \hat{\Phi}^{\mathrm{i}}(k_x, k_y)S_{L_{1,5}}\hat{\Phi}^{\mathrm{f}}(x, y)\Psi_1^{-}, \quad (4)$$

and from the left end face of the dielectric plate to its right end face, by the operator $\hat{S}_{\mathrm{pl}}$:

$$\Psi_{6\mathrm{pl}}^{+} = \hat{S}_{\mathrm{pl}}\Psi_5^{-} = \hat{\Phi}^{\mathrm{i}}(k_x, k_y)S_{L_{\mathrm{pl}}}\hat{\Phi}^{\mathrm{f}}(x, y)\Psi_{5\mathrm{pl}}^{-}. \quad (5)$$

Note that the refractive index n_{pl} of the plate is complex. With such a refractive index, relationship (2) inside the dielectric plate must be replaced by

$$k_z = \sqrt{n_{\mathrm{pl}}^2k^2 - k_x^2 - k_y^2}.$$

The variation of the beam structure from the right end face of the dielectric plate to the first mirror is described by the operator S_6:

$$\Psi_{1m}^{+} = \hat{S}_6\Psi_6^{-} = \hat{\Phi}^{\mathrm{i}}(k_x, k_y)S_{L_6}\hat{\Phi}^{\mathrm{f}}(x, y)\Psi_6^{-}. \quad (6)$$

The beam incident on the first mirror is related to the beam leaving it through an equation similar to (3).

On the left end face of the left film, the field is formed by the incident beam with the amplitude Ψ_5^{+} and the outgoing beam with the amplitude Ψ_5^{-}; on the right end face of the left film, by the incident beam with the amplitude $\Psi_{5\mathrm{pl}}^{+}$ and the outgoing beam with the amplitude $\Psi_{5\mathrm{pl}}^{-}$. These amplitudes are related as [11]

$$\Psi_5^{-} = r_+\Psi_5^{+} + t_+\Psi_{5\mathrm{pl}}^{+}, \quad (7)$$

$$\Psi_{5\mathrm{pl}}^{-} = r_-\Psi_{5\mathrm{pl}}^{+} + t_-\Psi_5^{+}, \quad (8)$$

where [11] $r_{\pm}$ are the reflection coefficients for the waves incident on the film from the second mirror and from the plate and $t_{\mp}$ are the transmission coefficients for the waves incident on the film from the second mirror and from the plate, respectively.[1] We will represent these coefficients through the thicknesses and refractive index of the films:[2]

$$r_+ = \frac{r_{12} + r_{23}\exp(-2j\varphi_0)}{1 + r_{12}r_{23}\exp(-2j\varphi_0)},$$

$$r_- = \frac{r_{32} + r_{21}\exp(-2j\varphi_0)}{1 + r_{32}r_{21}\exp(-2j\varphi_0)},$$

$$t_- = \frac{t_{12}t_{23}\exp(-j\varphi_0)}{1 + r_{12}r_{23}\exp(-2j\varphi_0)},$$

$$t_+ = \frac{t_{32}t_{21}\exp(-j\varphi_0)}{1 + r_{32}r_{21}\exp(-2j\varphi_0)},$$

$$\varphi_0 = kn_{\mathrm{fil}}L_{\mathrm{fil}}, \quad r_{12} = -r_{21} = \frac{1 - n_{\mathrm{fil}}}{1 + n_{\mathrm{fil}}},$$

$$r_{23} = -r_{32} = \frac{n_{\mathrm{fil}} - n_{\mathrm{pl}}}{n_{\mathrm{fil}} + n_{\mathrm{pl}}}, \quad t_{12} = \frac{2}{1 + n_{\mathrm{fil}}},$$

$$t_{21} = \frac{2n_{\mathrm{fil}}}{1 + n_{\mathrm{fil}}}, \quad t_{23} = \frac{2n_{\mathrm{fil}}}{n_{\mathrm{pl}} + n_{\mathrm{fil}}}, \quad t_{32} = \frac{2n_{\mathrm{pl}}}{n_{\mathrm{pl}} + n_{\mathrm{fil}}}.$$

On the left side of the right film, the field is formed by the incident and reflected beams with the amplitudes $\Psi_{6\mathrm{pl}}^{+}$ and $\Psi_{6\mathrm{pl}}^{-}$, respectively; on the right side of the

[1] We neglect diffraction inside the films assuming that they are thin.

[2] Here, we assume that the refractive index of the resonator medium equals unity and the velocity of electromagnetic waves in this medium equals the velocity of light in a vacuum.

right film, by the beams with the amplitudes Ψ_6^+ and Ψ_6^-. These beams are related as

$$\Psi_{6pl}^- = r_-\Psi_{6pl}^+ + t_-\Psi_6^+, \quad (9)$$

$$\Psi_6^- = r_+\Psi_6^+ + t_+\Psi_{6pl}^+. \quad (10)$$

A solution to the system of equations (1)–(10) yields the fields scattered by the resonator, i.e., the beams with the amplitudes Ψ_3^- and Ψ_4^-.

We will characterize these fields by two coefficients that are the functionals of the fields scattered by the resonator. These are the coefficient of reflection from the resonator, K_r^0, and the coefficient of transmission through it K_t^0.

In this so-called single-mode [8] case, the reflection coefficient (transfer coefficient upon reflection) and transmission coefficients (transfer coefficient upon transmission) are given by

$$K_{r,t}^0 = \frac{P_{r,t}^0}{P},$$

where $P = \iint |\Psi_3^+|^2\, dxdy$ is the power of the beam that excites the resonator, $P_{rt}^0 = P_{rt}N_{rt}$, $P_r = \iint |\Psi_3^-|^2\, dxdy$ is the power of the beam leaving the resonator toward the incident beam, and $P_t = \iint |\Psi_6^-|^2\, dxdy$ is the power of the beam transmitted through the resonator.

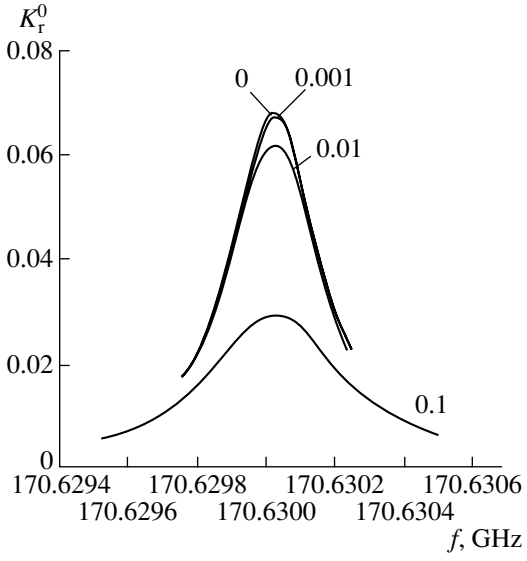

Fig. 3. Coefficient of reflection from the resonator as a function of frequency. The figures by the curves are the values of $\tan\delta_{fil}$.

The beam–detector coupling coefficients $N_{r,t}$ characterize the portion of the incident radiation that is picked up by the detector:

$$N_r = \frac{\int \Psi_3^- \Psi_{r,out}^* ds \int \Psi_3^{*-} \Psi_{r,out} ds}{P_r P_{r,n}},$$

where function $\Psi_{r,out}^*$ describes the field structure of the detector operating mode on the coupling film; $P_{r,n} = \int |\Psi_{r,out}|^2\, ds$ is the power of the beam with this structure; and

$$N_t = \frac{\int \Psi_4^- \Psi_{t,out}^* ds \int \Psi_4^{*-} \Psi_{t,out} ds}{P_t P_{t,n}},$$

where $\Psi_{t,out}$ is the field structure of the detector operating mode on the coupling film and $P_{t,n} = \int |\Psi_{t,out}|^2\, ds$ is the power of the beam with this structure.

Hereafter, the beam of the detector operating mode is assumed to be Gaussian:

$$\Psi_{r,out} \sim \Psi_{t,out} \sim \exp\left(-\frac{x^2+y^2}{2a_{r,t}^2}\right)$$

with a half-width a_r for the detector that measures the reflected wave and a_t for the detector that measures the transmitted wave.

A method for solving the system of equations (1)–(10) is reported in [8]. It was solved by iterations at a given frequency. As the initial value of the functions Ψ_1^+, Ψ_{5pl}^+, and Ψ_6^+, we usually took the function Ψ_3^+. The transverse distribution of the exciting wave beam Ψ_3^+ was taken to be Gaussian:

$$\sim \exp\left(-\frac{x^2+y^2}{2a_b^2}\right),$$

where a_b is the half-width of the exciting beam.

The amplitudes of the reflected and transmitted output beams were calculated from the relationships

$$\Psi_3^- = iR\hat{S}_2\hat{R}_2\hat{S}_2\{T\Psi_1^+ + jR\Psi_3^+\},$$

$$\Psi_4^- = T\Psi_3^+ + jR\Psi_1^+,$$

into which the function Ψ_1^+ found above was substituted.

Then, the frequency was varied and the response of the resonator was measured. At certain resonance frequencies f_{res}, the reflection coefficient increased and the transmission coefficient decreased.

Figure 3 plots the coefficient of reflection from the resonator vs. frequency at the following realistic

parameters: L = 400.31035 mm, R_0 = 240.583 mm, R_{ohm} = 0.9992, L_{pl} = 1.84285, L_{fil} = 0.001 mm, n = 2.3808, a_m = 60 mm, a_{ex} = 60 mm, and $\tan\delta_{pl} = 10^{-5}$. The speed of light is taken to be $c = 2.9969518 \times 10^{11}$ mm/s. We assume that the exciting beam width is equal to the width of the beam in the resonator near the coupling film and that the dielectric plate is at the center of the resonator.

The maxima of the reflection coefficient correspond to a particular resonator mode. This mode can be characterized by a longitudinal index m_s that is an integer equal to the number of field maxima for this mode on the resonator axis. The index m_s can roughly be found as an integer nearest to a number $\tilde{m}_s$:

$$\tilde{m}_s = 2\frac{L + (L_{pl} + 2L_{fil})(n-1)}{\lambda} - \frac{\arccos(g)}{\pi} \approx 458.003,$$

where

$$g = \left(1 - \frac{L}{R_0}\right),$$

$\lambda = c/f_{res}$, and f_{res} is found from the solution to system (1)–(10).

When the plate executes plane-parallel motion along the resonator axis, the frequency of the mode changes in general. However, for a given plate thickness, the resonator length can be chosen such that the eigenfrequency of one of the modes remains almost unchanged. This effect is illustrated in Fig. 4, which plots the resonance frequency for the mode with the longitudinal index m_s = 458 vs. plate displacement L at several resonator lengths (at $L = L_0$, which keeps the frequency almost unchanged; at $L = L_0 + 5$ μm; and at $L = L_0 + 10$ μm). Figure 5 shows the width of the resonance curve of this mode vs. plate displacement from the resonator center at $L = L_0$ for several values of the film loss tangent. The width varies because, as the plate shifts, the field distribution inside it changes, which causes the ohmic loss to change.

In the experiment, the dielectric loss tangent in the plate was determined from the maximum values of the mode Q factor. It was found that one can construct a much simpler theory for a resonator with such a plate inside that adequately describes the behavior illustrated in Fig. 5.

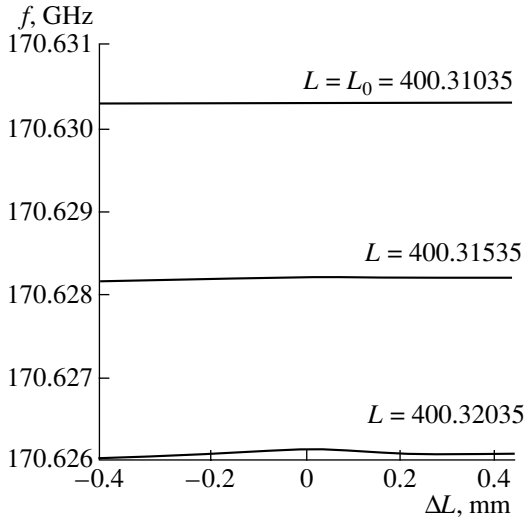

Fig. 4. Eigenfrequency versus plate displacement ΔL from the resonator center at various L (mm).

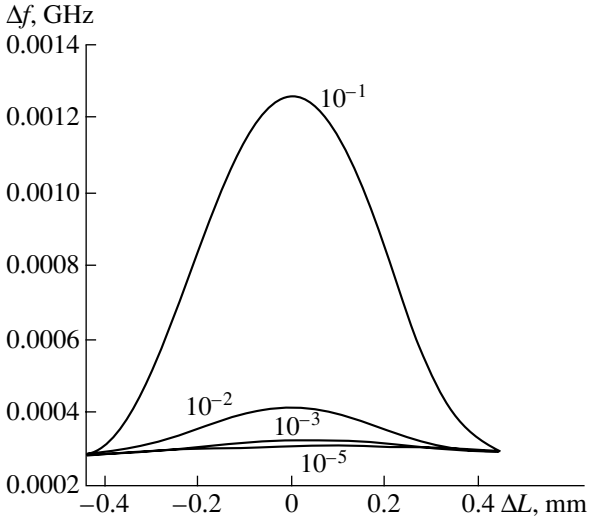

Fig. 5. Resonance curve width versus plate displacement ΔL from the resonator center at various loss tangents in the films. The figures by the curves are the values of $\tan\delta_{fil}$.

RESONATOR WITH A PLANE-PARALLEL PLATE AND INFINITE PLANE MIRRORS

Let us apply system (1)–(10) to a resonator with infinite plane mirrors (in this case, the integral operators $\hat{S}_i$ become numbers; in particular, $\hat{S}_2 = \exp[-jkL_2]$, etc.) and find the eigenfrequencies of the resonator by setting the exciting field equal to zero. The system of integral equations is then reduced to the transcendental characteristic equation

$$\begin{aligned}&[-1 + r_+T^2R_{ohm}\exp[-2jk(L_2 + L_{1,5})]]\\ &\times[-1 + r_+R_{ohm}\exp(-2jkl_6)] - \exp(-2jkn_{pl}L_{pl})\\ &\times[-r + T^2R_{ohm}D\exp[-2jk(L_2 + L_{1,5})]]\\ &\times[-r_- + R_{ohm}D\exp(-2jkL_6)] = 0,\end{aligned} \quad (11)$$

where $D = r_+r_- - t_+t_-$.

To find a solution to Eq. (11), we use the perturbation technique, assuming that small parameters are the

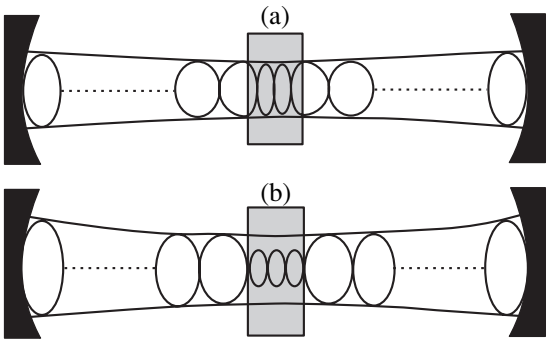

Fig. 6. Field structure in a resonator with a plate. The sides of the plate are located at (a) the maxima and (b) zeros of the electric field.

dielectric loss tangents ($\tan\delta_{pl}$, $\tan\delta_{fil} \ll 1$), the ohmic losses in the mirrors ($\delta_{ohm} = 1 - R_{ohm}^2 \ll 1$), the coefficient of reflection from the coupling film ($\delta_{coupl} = 1 - T^2 \ll 1$), and the thickness of the absorbing films ($\varphi_0 = knL_{fil} \ll 1$).

In the zeroth approximation (when the small parameters are zero), the equation for the eigenfrequencies takes the form

$$-2r_+\cos[k^0(L_2 + l_{1,5} - L_6)]\sin[k^0 n(L_{pl} + 2L_{fil})] + r_+^2\sin[k^0(L_2 + L_{1,5} + L_6 - nL_{pl} - 2L_{fil})] = \sin[k^0(L_2 + L_{1,5} + nL_{pl} + 2L_{fil})]. \quad (12)$$

It follows from Eq. (12) that, if

$$nk^0(L_{pl} + 2L_{fil}) = m_1\pi, \quad (13)$$

$$k^0(L_2 + L_{1,5} + L_6) = m_2\pi, \quad (14)$$

where m_1 and m_2 are the integer numbers of half-waves falling on the resonator axis inside and outside a dielectric plate with absorbing films, respectively, there exists a mode with a wave number k^0 and frequency that does not change when the plate executes plane-parallel motion along the resonator axis [2].

In the first-order approximation of the perturbation theory, the variation in the eigenwavenumber $\Delta k = j\Delta k''$ of this mode with the index $m_1 + m_2$ due to the coupling losses and thermal losses in the mirrors $\Delta k''_{mir}$, thermal losses in the dielectric plate $\Delta k''_{pl}$, and thermal losses in the films that cover the plate $\Delta k''_{fil}$ is given by

$$\Delta k''_R = \frac{(\delta_{ohm} + \delta_{coupl})(1 - r_+^2)}{2\{(1 - r_+^2)(L_2 + L_{1,5} + L_6) + n\langle L_{pl} + 2L_{fil}\rangle\langle r_+^2 + 1 + 2r_+\cos[2k^0(L_2 + L_{1,5})]\rangle\}},$$

$$\Delta k''_{pl} = \frac{n\tan\delta_{pl}k^0\left\{L_{pl}[r_+^2 + 1 + 2r_+\cos\langle 2k^0(L_2 + L_{1,5})\rangle] + \frac{4L_{fil}}{(n+1)^2}[1 - n^2 + (n^2+1)\cos\langle 2k^0(L_2 + L_{1,5})\rangle]\right\}}{2\{(1 - r_+^2)(L_2 + L_{1,5} + L_6) + n[L_{pl} + 2L_{fil}][r_+^2 + 1 + 2r_+\cos\langle 2k^0(L_2 + L_{1,5})\rangle]\}}, \quad (15)$$

$$\Delta k''_{fil} = \frac{4n\tan\delta_{fil}k^0 L_{fil}[1 - \cos\langle 2k^0(L_2 + L_{1,5})\rangle]\frac{n^2}{(n+1)^2}}{\{(1 - r_+^2)(L_2 + L_{1,5} + L_6) + n[L_{pl} + 2L_{fil}][r_+^2 + 1 + 2r_+\cos\langle 2k^0(L_2 + L_{1,5})\rangle]\}}.$$

First, consider the width of the resonance curve, $\Delta k'' = \Delta k''_R + \Delta k''_{pl}$, as a function of the plate position in the resonator (of length $L_2 + L_{1,5}$) without the covering absorbing films:

$$\Delta k''_R + \Delta''_{pl} \approx \frac{\tan\delta_{pl}}{2} \times \frac{\frac{\delta_{ohm} + \delta_{coupl}}{k^0 nL_{pl}\tan\delta_{pl}} + \frac{\{1 + 2r_+\cos\langle 2k^0(L_2 + L_{1,5})\rangle + r_+^2\}}{1 - r_+^2}}{\frac{L_1 + L_{1,5} + L_6}{nL_{pl}} + \frac{\{1 + 2r_+\cos\langle 2k^0(L_2 + L_{1,5})\rangle + r_+^2\}}{1 - r_+^2}}. \quad (16)$$

As follows from this expression, the Q factor $Q = k^0/2\Delta k''$ of the mode of interest oscillates as the plate moves along the resonator axis, reaching extrema when $2k^0(L_2 + L_{1,5}) = s\pi$, where s is an integer.

When s is even, the sides of the dielectric plate are located in the zeros of the electric field of the standing wave in the resonator. In this case, the electric field amplitude in the dielectric plate is minimal (Fig. 6a). The Q factor of this mode is

$$Q_+ = \frac{k^0(L_2 + L_{1,5} + L_6 + L_{pl})}{\delta_{ohm} + \delta_{coupl} + \tan\delta_{pl}k^0 L_{pl}}. \quad (17)$$

When s is odd, the sides of the dielectric plate are located in the maxima of the electric field of the standing wave and the electric field amplitude in the dielectric plate is maximal (Fig. 6b). The Q factor of this

mode is

$$Q_- = \frac{k^0(L_2 + L_{1,5} + L_6 + n^2 L_{pl})}{\delta_{ohm} + \delta_{coupl} + n^2 \tan\delta_{pl} k^0 L_{pl}}. \quad (18)$$

If the dielectric is lossless, the mode with Q_-, which has the maximal electric field amplitude, also has the higher Q factor due to the energy stored in the plate. If the dielectric losses prevail and the energy stored in the plate is small compared to that stored in the remaining space of the resonator, the mode with Q_+ has the higher Q factor. As a rule, the latter case is the one observed experimentally [2]. As follows from (16), when

$$\frac{\delta_{ohm} + \delta_{coupl}}{k^0 n L_{pl} \tan\delta_{pl}} = \frac{L_2 + L_{1,5} + L_6}{n L_{pl}}$$

or

$$\tan\delta_{pl} = \frac{\delta_{ohm} + \delta_{coupl}}{k^0(L_2 + L_{1,5} + L_6)}$$

the Q factor does not oscillate.

The dielectric losses are determined by measuring the extremum values of the Q factors, Q_+ and Q_-, as well as the Q factor of the empty resonator mode with the same index $m_1 + m_2$ at the same frequency:

$$Q_0 = \frac{k^0(L_2 + L_{1,5} + L_6 + nL_{pl})}{\delta_{ohm} + \delta_{coupl}}. \quad (19)$$

Next, by eliminating the losses in the empty resonator, $\delta_{ohm} + \delta_{coupl}$, from any pair of equations (17)–(19), we come to the following three expressions for $\tan\delta_{pl}$ in terms of Q_0, Q_+, and Q_-. The corresponding values are denoted as $\tan\delta_{pl-}$ when (17) and (19) are used; $\tan\delta_{pl+}$, (18) and (19); and $\tan\delta_{pl0}$, (17) and (18):

$$\tan\delta_{pl+} = \left(\frac{1}{Q_+} - \frac{1}{Q_0}\right)\left[\frac{L_2 + L_{1,5} + L_6}{L_{pl}} - (n-1)\right], \quad (20)$$

$$\tan\delta_{pl-} = \left(\frac{1}{Q_-} - \frac{1}{Q_0}\right)\left[\frac{L_2 + L_{1,5} + L_6}{n^2 L_{pl}} + \frac{(n-1)}{n}\right], \quad (21)$$

$$\tan\delta_{pl0} = \frac{1}{n^2 - 1}\left[\left(\frac{1}{Q_+} - \frac{1}{Q_-}\right)\frac{L_2 + L_{1,5} + L_6 + nL_{pl}}{L_{pl}} + (n+1)\left(\frac{n}{Q_+} + \frac{1}{Q_-}\right)\right]. \quad (22)$$

When using formulas (20)–(22), one should keep in mind that the expressions

$$\frac{1}{Q_-} - \frac{1}{Q_0} \quad \text{and} \quad \frac{1}{Q_+} - \frac{1}{Q_-}$$

can take both positive and negative values.

Table 1

Disk no.	Thickness, mm	Frequency, GHz	$\tan\delta_{pl-} \times 10^4$	$\tan\delta_{pl+} \times 10^5$
87	0.41	152	6.5	7.6
93	0.38	166	4.0	5.8
128	0.37	170	1.7	2.5
131	0.46	136	2.0	9.6

Thus, $\tan\delta_{pl}$ can be measured at two positions of the dielectric plate: when the dielectric loss in the resonator is minimal (this case is used to measure relatively large values of $\tan\delta_{pl}$, up to 10^{-2}) and maximal (this case is applied to ultra-low-absorption materials).

The technique that does not use the Q factor of an empty resonator is very convenient when the parameters of the plate vary in the course of the experiment, for example, when the temperature dependence of $\tan\delta_{pl}$ is taken.

Undoubtedly, results of calculating $\tan\delta_{pl}$ from measured values of Q_0, Q_+, and Q_- must be independent of the plate position inside the resonator. Therefore, when we use the three versions to calculate $\tan\delta_{pl}$, a self-check of the result becomes possible.

When examining diamond disks (and, earlier, silicon disks), we found out that the theoretical $\tan\delta_{pl}$ depends strongly on their position in the resonator, the value calculated from Q_- and Q_0 being significantly higher than that calculated from Q_+ and Q_0 (Table 1). This disagreement may arise, in particular, because of additional losses on the surface of the dielectric plate. Under the assumption that these additional losses are concentrated in a thin surface layer, it follows from formulas (15) that they are introduced only into the mode with Q_-:

$$Q_- \approx \frac{k^0[L_2 + L_{1,5} + L_6 + n^2(L_{pl} + 2L_{fil})]}{\delta_{ohm} + \delta_{coupl} + n^2 \tan\delta_{pl} k^0 L_{pl} + 4n^2 k^0 L_{fil} \tan\delta_{fil}}. \quad (23)$$

The extremal values of the resonance curve width calculated from formulas (17) and (23) agree with those predicted by the more rigorous theory (Eqs. (1)–(10)) with an accuracy of four significant figures for the same resonator length and at the resonance frequency calculated from solutions to (1)–(10).

From the assumption that the surface absorbing layer on the plate sides is thin, we can evaluate the absorption of a traveling wave separately in the volume and at the surface of the plate.

Moreover, with this assumption, the values of $\tan\delta_{pl}$ and $\tan\delta_{fil}$ found can be used to estimate the volume, P_{pl}, and surface, P_{fil}, losses in the plate for the

Table 2

Frequency, GHz	$\tan\delta_{pl} \times 10^5$	Absorption in the plate volume $\frac{P_{pl}}{\tilde{P}} \times 10^4$	Absorption in the surface layers $\frac{P_{pl}}{\tilde{P}} \times 10^4$
22DB1 (ITER-1) plate thickness L_{pl} = 2.22 mm, diameter $2a_{pl}$ = 119 mm			
56.585	5	3.5	4.0
84.867	3.5	4.4	3.0
113.153	3.0	4.9	4.0
141.432	2.45	5.2	5.0
169.710	2.45	6.2	5.0
197.990	2.5	7.1	5.0
22DB5 (ITER-3) plate thickness L_{pl} = 1.86 mm, diameter $2a_{pl}$ = 104 mm			
67.846	2.3	1.8	8
101.741	2.0	2.8	10
135.645	1.7	3.2	12
169.549	1.4	3.1	14
203.448	1.35	4.1	18
22DB6 plate thickness L_{pl} = 1.85 mm, diameter $2a_{pl}$ = 106 mm			
68.276	2.2	2.0	13
102.393	1.9	2.6	16
136.514	1.6	2.8	19
170.630	1.4	3.1	22
204.747	1.3	3.6	24
53DB1 plate thickness L_{pl} = 1.798 mm, diameter $2a_{pl}$ = 106 mm			
70.182	1.05	0.9	4.4
105.257	1.00	1.3	5.3
140.332	0.79	1.4	5.9
175.407	0.64	1.4	6.6

traveling wave at its resonance frequency. The losses can be calculated from the following expressions:

$$\frac{P_{tot}}{\tilde{P}}(\tan\delta_{pl}, \tan\delta_{fil}) \approx \frac{P_{pl}}{\tilde{P}}(\tan\delta_{pl}, 0)$$

$$+ \frac{P_{fil}}{\tilde{P}}(0, \tan\delta_{fil}) = -1 - |K|^2 - |R_{pl}|^2,$$

where

$$K(\tan\delta_{pl}, \tan\delta_{fil}) = \frac{t_+ t_- \exp(-ikn_{pl}L_{pl})}{1 + r_+ r_- \exp(-2ikn_{pl}L_{pl})}$$

is the transfer coefficient of the plate with the film,

$$R_{pl}(\tan\delta_{pl}, \tan\delta_{fil}) = \frac{r_+ + r_- \exp(-2ikn_{pl}L_{pl})}{1 + r_+ r_- \exp(-2ikn_{pl}L_{pl})}$$

is the coefficient of reflection from the plate, P_{tot} are the total losses in the films and plate, and $\tilde{P}$ is the power of the wave incident on the plate.

DIAMOND PLATES WITH LOW DIELECTRIC LOSSES

The absorption in the diamond plates was measured with a setup built around a Fabri–Perot open resonator [12]. The measurements were conducted at the resonance frequencies of the plates placed approximately at the center of the resonator perpendicular to its axis. The sensitivity of the setup in terms of $\tan\delta_{pl}$ is $\sim 10^{-7}$ for $\approx\lambda/2$ thick diamond plates. The plates measured in the experiment were several half-waves thick and the loss tangent was $\tan\delta_{pl}$ was $\sim 10^{-6}$. Therefore, the sensitivity margin was sufficient for precision measurements.

We tested several plates $2a$ = 60 mm in diameter fabricated by the DIGAZKRON Co. (Moscow) and four disks from De Beers. The parameters of the disks are summarized in Tables 1 and 2.

For all the plates, the refractive index was n = 2.38 and the surface roughness was 0.5–1.0 μm. Metal flanges were soldered to the De Beers disks through aluminum rings by thermocompression. To these flanges, cooling fittings and flanges needed to fasten the window to the gyrotron and overmoded waveguide were then welded.

Table 1 demonstrates the discrepancy between the values of $\tan\delta_{pl}$ calculated from Q_- and Q_0 and from Q_+ and Q_0, which indicates the presence of the additional surface absorbing layer on the plates. Leaving aside the nature of this layer, we only note that all our attempts to appreciably clean the surfaces by various mechanical and chemical techniques both separately and in combination failed. As yet, the only reliable technique for cleaning the surfaces is to heat them to 600°C in air. Apparently, the oxygen present in air completely burns out impurities from the plate surfaces.

After the heating, the values of $\tan\delta_{pl+}$ calculated from Q_- and Q_0 and from Q_+ and Q_0 coincide within the accuracy of the experiment. It is noteworthy that the value of $\tan\delta_{pl+}$ remains unchanged, while the value of $\tan\delta_{pl-}$ becomes equal to $\tan\delta_{pl+}$. This result corroborates the assumption that the additional surface absorption occurs in a very thin layer.

The thickness of this layer is impossible to measure mechanically. We suppose that it is comparable to that of the layer damaged by the mechanical treatment (about 0.5–1.0 μm). On the other hand, it was found

that the resonance frequency of the disk increases by ≈150 MHz after each heating cycle, which corresponds to the thinning of the plate by approximately 1 μm. These data have given impetus to the development of the above theory.

Table 2 lists the values of $\tan\delta_{pl}$ and absorption coefficients for four De Beers windows calculated under the additional assumption that the absorbing films are ≈1 μm thick. It can be seen that the surface absorption is much higher than the volume absorption. This caused the overheating of the windows and rf breakdowns over its surface. Direct calorimetric measurements of the power absorbed by window 3 (thermal conductivity 18 W/(cm K)) in the operating gyrotron agree with the theoretical results on the volume and surface losses. It should be noted that the other three windows failed during the operation, probably due to the overheating.

From the frequency behavior of the loss tangent, it follows that $\tan\delta_{pl}$ varies inversely with frequency in the low-frequency region, then stabilizes, and even slightly grows above 170 GHz. This indicates that the absorption mechanism in CVD diamond changes from predominantly free-carrier absorption to lattice absorption presumably by impurity centers and lattice imperfections. Unfortunately, this also means that the losses in the window increase with gyrotron operating frequency.

Note that the minimal loss tangent measured in the high-frequency part of the millimeter-wave range was $\approx 7.0 \times 10^{-6}$ at 170 GHz. This loss depends on the diamond quality and is as yet far from the theoretical limit of $\sim 10^{-8}$ [13].

CONCLUSION

Based on the resonator method, a technique for measuring low losses in dielectrics is developed. Our measurements showed that the best diamond disks have a loss tangent of $\approx 7.0 \times 10^{-6}$ at 170 GHz.

An excessive absorption in thin surface layers of diamond disks, which is much higher than the volume absorption of the disk, is found. This effect is due to various surface treatments (polishing, soldering, chemical treatment, etc.).

A technique for measuring and evaluating this excessive absorption is developed.

Loss tangents are measured for diamond disks and real diamond power-extracting windows of high-power gyrotrons.

ACKNOWLEDGMENTS

This work was supported by the Russian Foundation for Basic Research (project no. 00-02-16423), the Council in Support of Leading Scientific Schools (project no. 00-15-96772), the research programs "Scientific Instrument Making" and "Quantum and Nonlinear Processes," and INTAS (grant no. 2173).

We thank Yu.K. Verevkin for submitting the results of illuminating diamond plates with the radiation of a 3-μm high-power XeCl laser.

REFERENCES

1. R. A. Valitov, S. F. Dyubko, V. V. Kamyshan, *et al.*, *Technology of Submillimetric Waves* (Sov. Radio, Moscow, 1969).
2. Yu. A. Dryagin and V. V. Parshin, Int. J. Infrared Millim. Waves **13**, 1023 (1992).
3. A. L. Cullen, IEEE Trans. Microwave Theory Tech. **MTT-24**, 534 (1976).
4. V. V. Parshin, R. Heidinger, B. A. Andreev, *et al.*, Int. J. Infrared Millim. Waves **16**, 863 (1995).
5. V. V. Parshin, V. G. Ralchenko, and V. I. Konov, in *Conference Digest of the 23rd International Conference of Infrared and Millimetre Waves, Colchester, 1998*, p. 232.
6. J. R. Brandon, S. E. Coe, R. S. Sussmann, *et al.*, Fusion Eng. Des. **53**, 553 (2001).
7. M. Thumm, Int. J. Infrared Millim. Waves **19**, 3 (1998).
8. S. N. Vlasov and E. V. Koposova, Radiofizika **44**, 940 (2001).
9. N. G. Bondarenko and V. I. Talanov, Izv. Vyssh. Uchebn. Zaved., Radiofiz. **7**, 313 (1964).
10. V. I. Talanov, Izv. Vyssh. Uchebn. Zaved., Radiofiz. **8**, 260 (1965).
11. M. Born and E. Wolf, *Principles of Optics* (Pergamon, Oxford, 1969; Nauka, Moscow, 1970).
12. A. F. Krupnov, M. Yu. Tretyakov, V. V. Parshin, *et al.*, J. Mol. Spectrosc. **202**, 107 (2000).
13. B. M. Garin, Fiz. Tverd. Tela (Leningrad) **32**, 1917 (1990) [Sov. Phys. Solid State **32**, 1917 (1990)].

Translated by A. Khzmalyan

Technical Physics, Vol. 47, No. 12, 2002, pp. 1570–1575. Translated from Zhurnal Tekhnicheskoĭ Fiziki, Vol. 72, No. 12, 2002, pp. 88–93.

**SURFACES,
ELECTRON AND ION EMISSION**

New Approach to Surface Ionization and Drift-Tube Spectroscopy of Organic Molecules

O. A. Bannykh, K. B. Povarova, and V. I. Kapustin

Baĭkov Institute of Metallurgy and Materials Science, Russian Academy of Sciences, Leninskiĭ pr. 49, Moscow, 117911 Russia

e-mail: Povarova@ultra.imet.ac.ru

Received April 15, 2002

Abstract—A new physical model of ionization of organic molecules from the class of amines on the oxidized surface of transition metals is suggested. According to this model, the process involves the capture of protons or hydroxyl groups forming on the oxide surface upon water molecule adsorption. The adequacy of the model is demonstrated experimentally with test amines, such as Novocaine (procaine), bencaine, Dimedrol (diphenylhydramine), etc. A theory of drift motion of ion beams that includes the space charge effect is proposed. It is shown that the quantity $P_i = \mu j/\varepsilon_0 v_g^2$ (where μ is the ion mobility, j is the ion current density, ε_0 is the permittivity, and v_g is the longitudinal velocity of an ionized gas) plays the role of perveance for intense drift ion beams. A new type of drift-tube spectrometer that uses an ion source due to surface ionization is developed.

INTRODUCTION

The ionization of atoms and molecules on a heated body surface has been long known. The pioneering works on this subject studied the ionization of alkali metal atoms and alkali halides on the surface of refractory and noble metals (see, e.g., [1]). Experimental data obtained at that time were successfully treated in terms of the well-known Saha–Langmuir equation [2].

In the mid-1960s, the selective surface ionization of molecules of several organic materials (specifically, amines) on the heated and preoxidized surface of refractory metals (tungsten, molybdenum, and rhenium) was discovered. This effect was observed in both a medium vacuum [3] and air [4]. The surface ionization was assumed to proceed in two stages: (1) dissociation of an organic molecule M on the adsorbent surface into fragments (M–H) and (M–R), where H is a hydrogen atom and R is a radical of the molecule M, and (2) desorption of fragments $(M–H)^+$ and $(M–R)^+$ as ions with the transfer of one electron to the solid [5, 6]. Within such a consideration, the Saha–Langmuir equation yields the following expression for the ion current from the solid surface [5, 6]:

$$I_i(T) = e\nu S\gamma_i(T)\beta_i(T, E), \qquad (1)$$

where T is temperature, E is the electric field strength at the solid surface, $I_i(T)$ is the current of ions of i sort from the solid surface, e is the charge of an electron, ν is the organic molecule flux per unit surface area of the solid, S is the surface area of the solid, $\gamma_i(T)$ is a correction factor (in [5], it was called the coefficient of conversion of the organic molecule flux to particles of i sort), and $\beta_i(T, E)$ is the surface ionization coefficient for particles of i sort. The coefficient β_i is given by [5]

$$\beta_i(T) = \frac{1}{1 + \frac{1}{A_i(T)}\exp\left(\frac{V_i - \varphi - (eE)^{1/2}}{kT}\right)}, \qquad (2)$$

where V_i is the adiabatic ionization potential for particles of i sort, $A_i(T)$ is the ratio between the partition functions for the ionized and neutral states of particles of i sort, and φ is the so-called "ionic work function" of a solid.

In our opinion, the conventional approach to treating the surface ionization of organic molecules is inadequate for the following reasons. First, it makes researchers introduce the concept of ionic work function, which does not have a clear physical meaning; second, it fails in explaining the experimentally found electric field dependence of the ion current; third, it does not explain the appearance of ions like $(M + H)^+$, $(M – H – 2nH)^+$, and $(M – R – 2nH)^+$ in the ion spectrum; fourth, the origin of the "concentration dependence" of the ion current (i.e., the dependence of the ion current on the organic molecule flux toward the solid surface) remains unclear; and, fifth, it cannot explain the very effect of ionization selectivity as applied to amine molecules on the oxidized surface of refractory metals.

1063-7842/02/4712-1570$22.00

PHYSICAL MODEL OF SURFACE IONIZATION

When considering the surface ionization of amine organic molecules, one should take into account the specific electronic configuration of amine groups entering into associated compounds, as well as the specific structure of the metal oxide surface.

Figure 1 shows a model of the oxide surface containing Brensted acid-type centers *1* (BACs) and Brensted base-type centers *2* (BBCs) [7]. These centers, which form on the surface of W, Mo, Re, Al, Zr, Mg, etc. oxides, consist of hydrogen ions and a hydroxyl group chemisorbed, respectively, on oxygen ions and metal oxides as a result of the dissociative adsorption of water molecules on the oxide surface. An oxide surface containing active BACs and BBCs may initiate exchange reactions between protons and hydroxide ions.

Nitrogen atoms in a molecule of an organic compound from the class of amines are known to have a pair of free valence electrons [8]. They can take on a proton to produce a secondary ion. Simultaneously, the proton acquires a closed electron orbital:

$$R_1 : \overset{\cdot\cdot}{N} : \begin{matrix} R_2 \\ \cdot\cdot \\ R_3 \end{matrix} + H^{\oplus} \longrightarrow [R_1 : \overset{\cdot\cdot}{N} : \begin{matrix} R_2 \\ H \\ \cdot\cdot \\ R_3 \end{matrix}]^{\oplus}. \qquad (3)$$

Instead of a proton, reaction (3) may also involve an ion of an alkali metal, for example, sodium. In this case, however, the secondary ion will be less stable, because the electron affinity of sodium, $S_{Na} = +0.08$ eV, is less than that of a hydrogen atom, $S_H = +0.75$ eV [1].

Thus, amine ions on the oxide surface may form without electron exchange between the organic molecules and the oxide surface. In this case, the overall process of ionization can be represented by a sequence of ionization reactions:

$$BAC\} + M \rightleftarrows (M + H)^{+} M \rightleftarrows (M - H)^{+} + H_{2\,gas}, \quad (4a)$$

$$\{BBC\} + M \rightleftarrows (M + OH)^{-} M \rightleftarrows (M - H)^{-} + H_2O_{gas}, \quad (4b)$$

$$\{BAC\} + MM \rightleftarrows (M - R)^{+} + (R + H)_{gas}. \quad (4c)$$

Since the Saha–Langmuir equation is inapplicable to reactions (4a)–(4c), the rate of surface ionization (the ion current) for the organic molecules can be calculated from these reactions with methods used in the theory of absolute reaction rates [9]:

$$I_i(T) = A\frac{P^n}{T^{5/2}}\exp\left(-\frac{\Delta W - (eE)^{1/2}}{kT}\right), \qquad (5)$$

Here, A is a constant; P is the partial pressure of the organic molecule vapor or gas at the oxide surface; ΔW is the activation energy of ion desorption from the oxide surface; and n is the reaction order, which depends on the order of molecule–surface bonds.

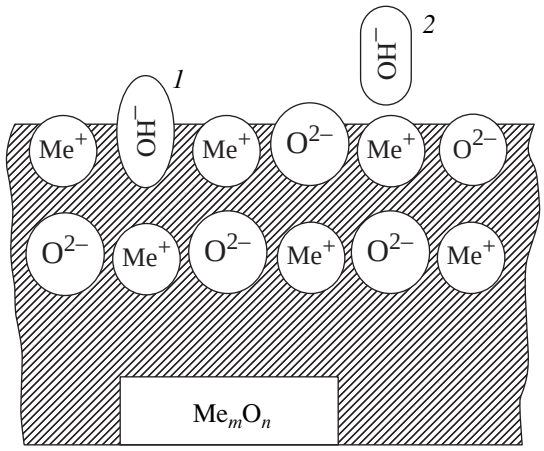

Fig. 1. Model of Me_mO_n oxide surface: (*1*) Brensted acid-type centers and (*2*) Brensted base-type center.

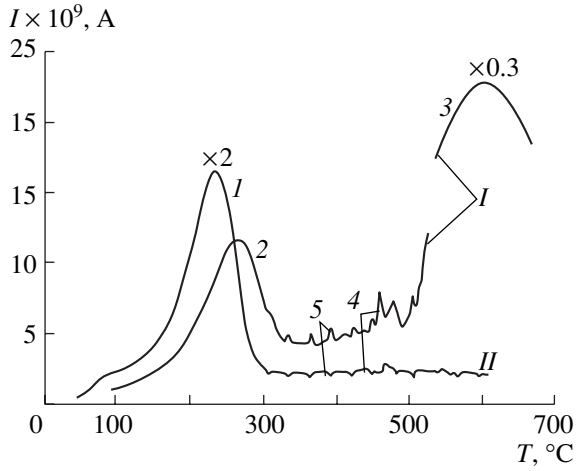

Fig. 2. Thermograms of the background current for (*I*) positive and (*II*) negative ions from the oxidized molybdenum surface in air.

Figure 2 demonstrates the temperature dependences of the background ion current in air. Curve *I* is the current of positive ions from the oxidized molybdenum surface, and curve *II* is that of negative ions. Peak *1* of the ion current corresponds to the hydroxide ion desorption from BBCs; peak *2*, to proton desorption from BBCs; peak *3*, to proton desorption from BACs; peaks *4*, to the phase transition in the molybdenum oxide at a temperature of 467°C; and peaks *5*, to thermal vibrations of hydrogen ions on the oxide surface. Table 1 lists the basic properties of Brensted centers on oxidized molybdenum calculated from Fig. 2. The fundamental frequency of hydrogen ion vibrations on the surface of oxidized molybdenum calculated from the position of peaks *5* was found to be 6.4×10^{11} Hz.

DESIGN OF A NEW DRIFT-TUBE SPECTROMETER

We developed and tested a new drift-tube spectrometer where the vapor of organic molecules delivered to

Table 1. Parameters of Brensted centers for oxidized molybdenum

Ion being desorbed	Type of Brensted center	Desorption maximum temperature, °C	Activation energy of desorption, eV
H^+	Base	268	1.36
OH^-	"	234	1.27
H^+	Acid	610	2.34

Table 2. Ionization parameters for Novocaine on oxidized molybdenum

Amount of Novocaine in air flow, ng	Energy of ion desorption activation, eV	Order of ionization reaction	Ionization efficiency, C/mol
3300	1.77	–1/6	20
400	1.29	1/2	100
50	1.26	1	200

the instrument with a carrier gas undergoes surface ionization (Fig. 3).

Ion emitter *1*, where the surface ionization of organic molecules is performed, is made of oxidized molybdenum. The working temperature of the emitter within 200–500°C is kept with heater *2*. An analyzed sample M containing amine organic molecules is delivered with the carrier gas (atmospheric-pressure air) flow Q_1. The transport of ions in the instrument is accomplished by means of another atmospheric-pressure air flow Q_0.

Organic molecule ions produced on the flat working surface of emitter *1* are extracted by ion-optical lens *3*, directed to analyzer *4* of the transverse drift mobility, and detected in the circuit of ion collector *6*. Ion suppressor *5* serves to improve the ion collection efficiency.

It is known [10, 11] that in air, the drift velocity of ions V_D depends on the electric field strength E in the drift space:

$$V_D = \mu_0(1 + \alpha E^2)E, \tag{6}$$

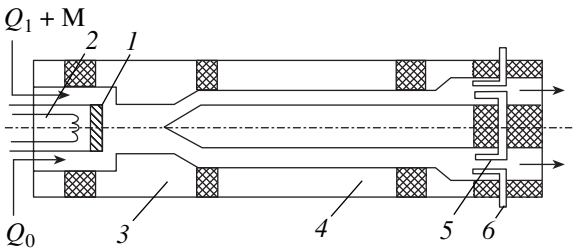

Fig. 3. Drift-tube spectrometer with the emitter using selective surface ionization of organic molecules.

where μ_0 is the low-field ion mobility and α is the nonlinearity parameter of the drift mobility (μ_0 and α may be both positive and negative according to the structure of an organic molecule and its electric charge).

To identify ions in transverse-mobility drift-tube spectrometers, an electric field of strength up to 30 kV/cm is usually applied between the outer and central electrodes of analyzer *4*. This field separates ions by the nonlinearity parameter of their drift mobility [12, 13]. The voltage across the analyzer has two components: a pulse asymmetric voltage with a frequency 500–800 kHz (a positive pulse of amplitude U and duration τ and a negative pulse of amplitude $U/2$ and duration 2τ) and a constant spectrum-sweeping voltage varying in the interval –10…+10 V.

SPACE CHARGE IN THE DRIFT MOTION OF IONS

The drift motion of ions in an atmospheric-pressure gas is usually studied without considering their space charge [10]. However, if the ion current exceeds 10^{-9}–10^{-10} A, the effect of the space charge may become essential.

In the absence of longitudinal gas flow, the solution of Poisson's equation for the drift motion of ions from the ion emitter surface in a flat-side gap of width d yields the following expression for the I–V characteristic of such a diode:

$$j = \frac{9}{8}\frac{\varepsilon_0\mu}{d^3}U^2, \tag{7}$$

where j is the density of the ion current to the collector and U is the emitter-to-collector voltage.

By jointly solving Poisson's equation and the equation of motion in the drift tube where a gas flows along the z axis, one obtains an expression for the radius r of an initially axisymmetric ion beam of radius r_0:

$$r = r_0[1 + P_i^* z/2]^{1/2}. \tag{8}$$

The quantity $P_i = \mu j/\varepsilon_0 v_g^2$ (where μ is the ion mobility, j is the ion current density, ε_0 is the permittivity, and v_g is the longitudinal velocity of the ionized gas) plays the role of perveance for intense ion beams.

By way of example, Fig. 4 depicts the trajectories of an initially axisymmetric ion beam of radius 0.5 mm when it moves in a 9-mm-long drift tube of radius 1 mm (the tube wall is shown by the dashed line). The curves were obtained with formula (8) for various ion currents and a transport gas flow rate of 3.5 l/min along the tube axis. It is seen that the tube "capacity" is no more than 7×10^{-9} A.

Similarly, the width a of a ribbon-shaped ion beam entering into a flat gap parallel to its walls will vary as

$$a = a_0[1 + P_i^* z], \tag{9}$$

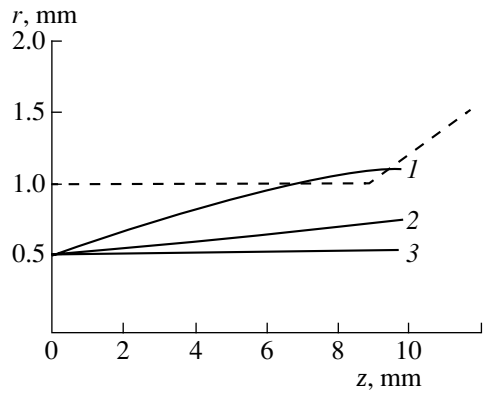

Fig. 4. Calculated boundaries of an axisymmetric ion beam in a drift tube of diameter 1 mm. The carrier gas flow rate is 3.5 l/min, and the beam current is (*1*) 1×10^{-8}, (*2*) 3×10^{-9}, and (*3*) 3×10^{-10} A.

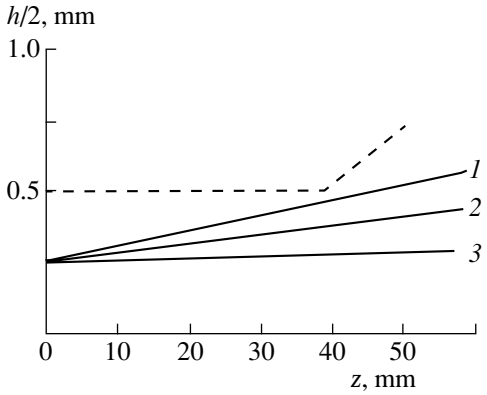

Fig. 5. Calculated boundaries of a ribbon-shaped ion beam in a flat channel with a plate spacing of 1 mm. The carrier gas flow rate is 3.5 l/min, and the beam current is (*1*) 5×10^{-11}, (*2*) 3×10^{-11}, and (*3*) 1×10^{-11} A.

where a_0 is the initial width of the ion beam.

Figure 5 shows the trajectories of a ribbon-shaped beam with an initial width of 0.5 mm passing in a flat channel 35 mm wide and 40 mm long. The plate spacing in the channel is 1 mm, and the transport gas flow rate is 3.5 l/min. The capacity of this channel is seen to be no higher than 5×10^{-11} A.

For axisymmetric and ribbon ion beams, the space charge induces a potential along their axes. It depends on the ion current and transport gas flow rate and may reach 200–300 V according to our calculations.

When ion beams move in an asymmetric flat gap, the transverse component of the ion current appears even if the potential difference between the flat-channel walls is absent.

Our approach, taking into account the effect of the space charge on the drift motion of ion beams, made it possible to optimize the design of the drift-tube spectrometer shown in Fig. 3.

SURFACE IONIZATION OF AMINES

Organic molecules of test amines underwent surface ionization in the drift-tube spectrometer shown in Fig. 3. Test microsamples of weight from 1ng to 10 μg were introduced into the carrier gas flow over a helical heater and evaporated by pulsed heating for 2–3 s. Current pulses from the first collector were recorded with an *x*–*y* recorder, the ion emitter temperature being fixed. Time scan was accomplished along the *x* axis, and the pulses were recorded along *y* axis. From five to seven measurements were made for each emitter temperature, each emitter–first collector voltage, and each sample weight. Varying the emitter–first collector voltage, we constructed the *I*–*V* characteristics of the diode gap.

All the *I*–*V* curves can be subdivided into two regions, the demarcation line between them lying in the interval 120–200 V depending on the emitter temperature, sample weight, and transport gas flow rate. The "low-voltage" region of the curves is well described by Eq. (7); the "high-voltage" one, by Eq. (5). With (7), we found the ion current for the zero electric field at the emitter surface. The current value was then used to determine the activation energy of ion desorption from the emitter surface and the order of the desorption reaction, which are involved in Eq. (5).

Figure 6 shows the experimental dependences of the positive ion current on the sample weight introduced into the surface ionization zone for Novocaine,

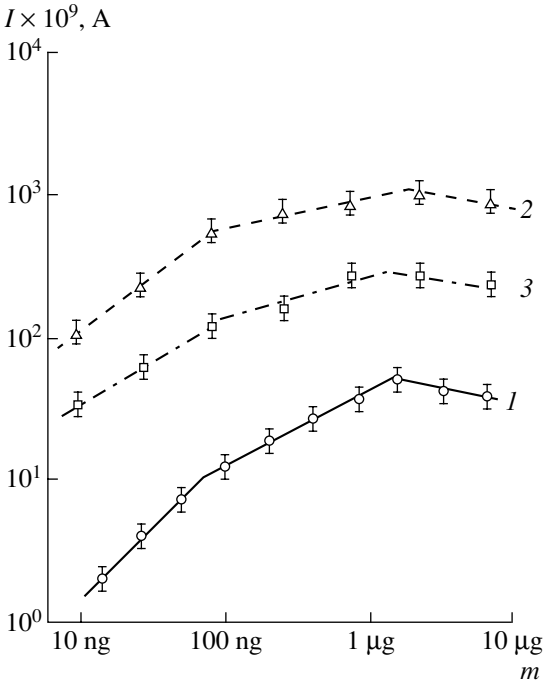

Fig. 6. Dose dependence of the positive ion current for (*1*) Novocaine, (*2*) bencaine, and (*3*) Dimedrol.

Table 3. Nonlinearity parameter of the drift mobility for several organic molecules from the class of amines

Peak in spectrum	$\alpha \times 10^{10}$, cm^2/V^2				
	novocaine	papaverine	dimedrol	bencaine	quinine
Basic	+0.018	−0.045	−0.034	+0.034	+0.023
Satellite	−0.65	−0.21	−0.53	−0.50	−0.94

bencaine, and Dimedrol. All the dependences have characteristic regions differing in slope and, hence, order of the ion desorption reaction. Table 2 lists the desorption orders and activation energies for Novocaine that were determined from the temperature dependences of the positive ion currents in these characteristic regions.

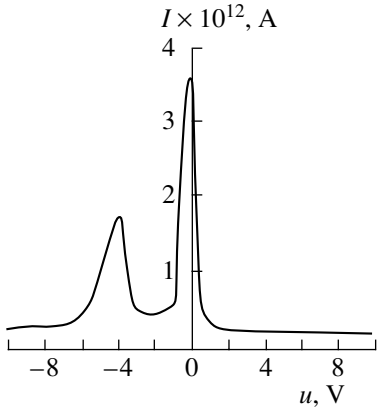

Fig. 7. Drift spectrum of quinine hydrochloride at a concentration of 1×10^{-10} g/cm^3 in an air flow.

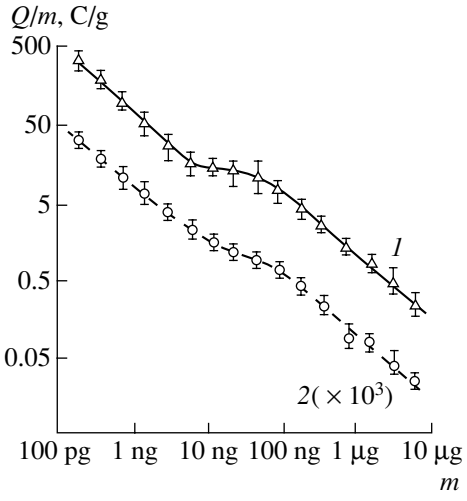

Fig. 8. Sensitivity of the drift-tube spectrometer vs. Novocaine amount in terms of (*1*) emitter and (*2*) collector current.

DRIFT SPECTRA FOR TEST AMINES

The drift-tube spectrometer with surface ionization that is described in this paper (Fig. 3) offers higher analytical and operating capabilities compared with conventional instruments.

Figure 7 demonstrates the drift spectrum of one of the test amines, quinine hydrochloride, for its concentration in the carrier gas flow 1×10^{-10} g/cm^3. The drift spectrum of this substance, as well as of the other amines studied, has two peaks: basic (at a sweep voltage close to zero) and satellite (at a negative bias voltage of 4.15 V). From the positions of these peaks, one can find the nonlinearity parameter of the drift mobility α of the ions. Note that the positions of the peaks are specific for each of the organic molecules. Table 3 summarizes the values of α for several amine compounds studied in this work.

Thus, our instrument allows for the identification of a class of organic molecules in a sample: with the emitter material selected appropriately, the very presence of the ion current from the emitter indicates the presence of amine compound molecules in the sample.

Organic molecules in a sample are analyzed quantitatively with the ion current from the emitter and the ion current from the collector. Figure 8 shows the concentration dependences of the instrument sensitivity in terms of the emitter current (curve *1*) and collector current (curve *2*) for Novocaine.

The type of organic molecules from the amine group is identified from the positions of the peaks in the drift spectrum, i.e., from the nonlinearity parameters of the drift mobility $\alpha_{1,2}$ of ions responsible for the peaks.

The instrument is insensitive to the humidity of carrier and transport gases. It makes it possible to analyze solid and liquid microsamples and samples of surface deposits, as well as to analyze air samples when air passes through the instrument or by trapping organic impurities of air with special sorbents. The sensitivity of the instrument is 1×10^{-6} g (by weight) or 3×10^{-17} g/cm^3 (by concentration) in terms of the emitter current and 1×10^{-12} g or 3×10^{-13} g/cm^3 in terms of the collector current. The drift-tube spectrometer measures $200 \times 50 \times 60$ mm, weighs 0.8 kg, and can be used in compact portable equipment for detecting narcotics and explosives [14].

CONCLUSIONS

(1) A new physical model of amine selective ionization on the oxidized transition metal surface in air is suggested. The model explains adequately experimental data for this phenomenon, namely, the dependence of the ion current on the amine concentration, temperature, and electric field strength at the metal oxide surface.

(2) A new approach to treating the drift motion of ions in an atmospheric-pressure gas is suggested. It takes into account the effect of the ion space charge on the ion path, including in the presence of a gas flow in the ion-drift tube. The concept of perveance for the drift motion of intense ion beams is introduced.

(3) A new type of a drift-tube spectrometer with surface ionization is developed. It is shown that the sensitivity of the new instrument is two or three orders of magnitude higher than that of similar spectrometers containing ion sources on radioactive isotopes.

(4) The nonlinearity parameters of the drift mobility are found experimentally for Novocaine, bencaine, papaverine, quinine hydrochloride, and Dimedrol.

REFERENCES

1. L. N. Dobretsov and M. V. Gomoyunova, *Emission Electronics* (Nauka, Moscow, 1966).
2. U. Kh. Rasulev and E. Ya. Zandberg, Prog. Surf. Sci. **28**, 181 (1988).
3. É. Ya. Zandberg and N. I. Ionov, Dokl. Akad. Nauk SSSR **141**, 139 (1962).
4. US Patent No. 5038544, G 01N33/00 (1991).
5. É. G. Nazarov and U. Kh. Rasulev, *Nonstationary Processes of Surface Ionization* (Fan, Tashkent, 1991).
6. U. Kh. Rasulev, Int. Ion Mob. Spec. **4** (2), 13 (2001).
7. S. R. Morrison, *The Chemical Physics of Surface* (Plenum, New York, 1977; Mir, Moscow, 1980).
8. V. M. Potapov, *Organic Chemistry* (Prosveshchenie, Moscow, 1970).
9. A. I. Lushpa, *Foundations of Chemical Thermodynamics and Kinetics of Chemical Reactions* (Mashinostroenie, Moscow, 1981).
10. E. McDaniel and E. Mason, *The Mobility and Diffusion of Ions in Gases* (Wiley, New York, 1973; Mir, Moscow, 1976).
11. E. G. Nazarov, R. A. Miller, G. A. Eiceman, *et al.*, Int. J. Ion Mob. Spec. **4** (2), 43 (2001).
12. W. McGann, Proc. SPIE **2937**, 78 (1996).
13. B. Carnahan, S. Day, V. Kouznetsov, *et al.*, Proc. SPIE **2937**, 106 (1996).
14. Review of the In-Vision Technologies Inc., ICAO Journal, No. 12, 11 (1995).

Translated by V. Isaakyan

Technical Physics, Vol. 47, No. 12, 2002, pp. 1576–1579. Translated from Zhurnal Tekhnicheskoĭ Fiziki, Vol. 72, No. 12, 2002, pp. 94–97.

EXPERIMENTAL INSTRUMENTS AND TECHNIQUES

Visualization of the Langmuir–Blodgett Film Structure in Liquid-Crystal Cells

B. N. Klimov and E. G. Glukhovskoy

Saratov State University, Bol'shaya Kazachya ul. 112a, Saratov, 410012 Russia

e-mail: GluhovskoyEG@info.sgu.ru

Received May 14, 2002

Abstract—The use of liquid-crystal (LC) cells for the visualization of the Langmuir–Blodgett (LB) film structure is described. Images appearing in LC cells upon LB film application are presented. The images are analyzed in relation to the LB film structure (application conditions) and to the relief formed by indium tin oxide (ITO) electrodes. © *2002 MAIK "Nauka/Interperiodica".*

INTRODUCTION

The behavior of liquid crystals are known to be highly sensitive to the state of the surface with which they are in contact.

The use of LCs for the visualization of the surface relief on substrates is also well known. However, this approach is today used only for research purposes.

The recent advances in Langmuir–Blodgett (LB) technique may regenerate interest in this domain of LC applications. In particular, if during the deposition of an LB monolayer one of the plates in an LC cell serves as a dummy sample, one can visualize the macrostructure of the resulting LB film.

The macrostructure of the LB film under test can be judged from the structure of the dummy sample with an express method proposed in this work

PREPARATION OF SAMPLES AND OBSERVATIONS OF ORIENTATION EFFECTS IN LC CELLS

The cross section of LC cells used for studying the LB film macrostructure is schematically shown in Fig. 1. The cell consists of two spaced glass plates (substrates).

As substrates, we used pure glass plates, as well as glass plates with strip transparent ITO electrodes, which are usually used in LC screens. The gap between the plates is provided by 6-μm-thick polyethylene stripes applied along the perimeter.

The dummy plate was prepared in the usual way: the orienting plane was obtained by rubbing with a felt-covered roller or special brushes on an industrial setup produced in the Volga Research Institute.

The other plate was covered by monolayers (MLs) of the octadecyldimethylamine salt of polyamide acid by the vertical LB technique [1]. After the deposition, the samples were subjected to thermal treatment with the resulting polymerization of polyamide acid to polyimide [2, 3].

The cells thus prepared were filled with ZhK-807 nematic liquid crystal in the isotropic state (at a temperature between 85 and 90°C). The material is provided by the Volga Research Institute.

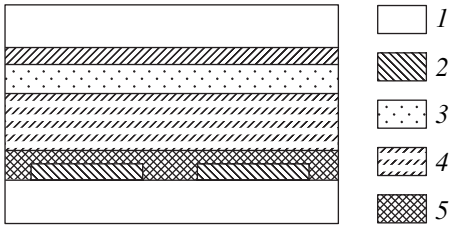

Fig. 1. (*1*) Glass, (*2*) ITO electrodes, (*3*) test coating, (*4*) LC, and (*5*) Langmuir–Blodgett film.

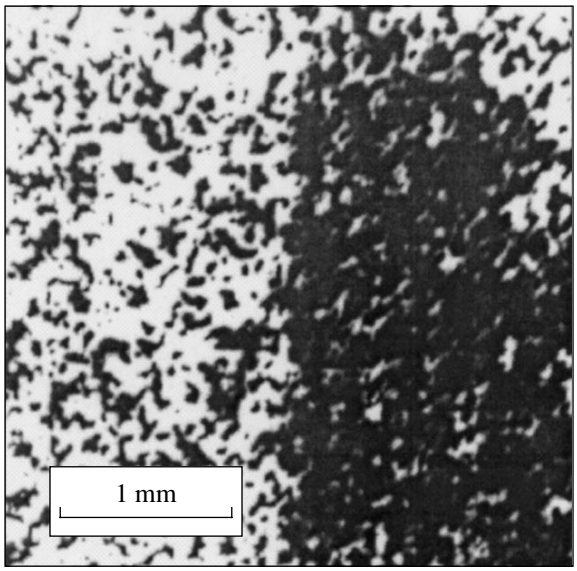

Fig. 2.

1063-7842/02/4712-1576$22.00 © 2002 MAIK "Nauka/Interperiodica"

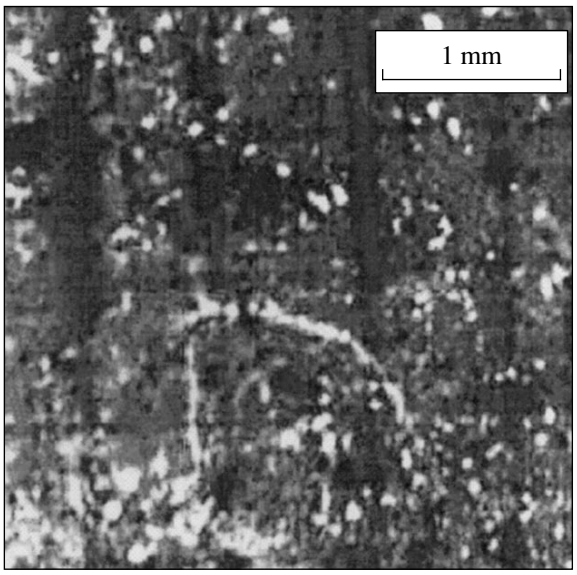

Fig. 3.

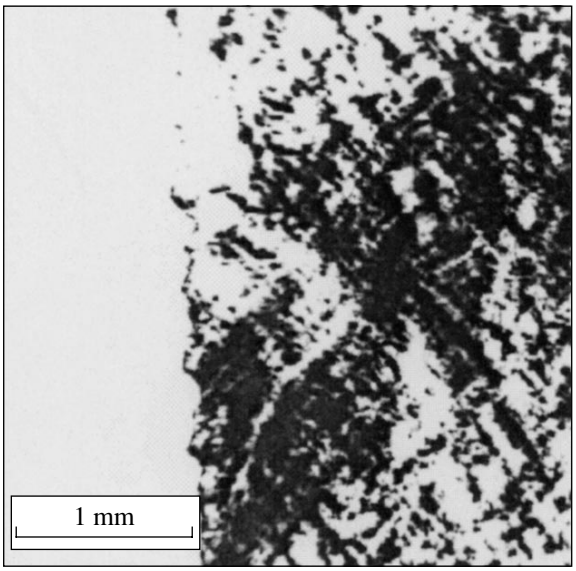

Fig. 4.

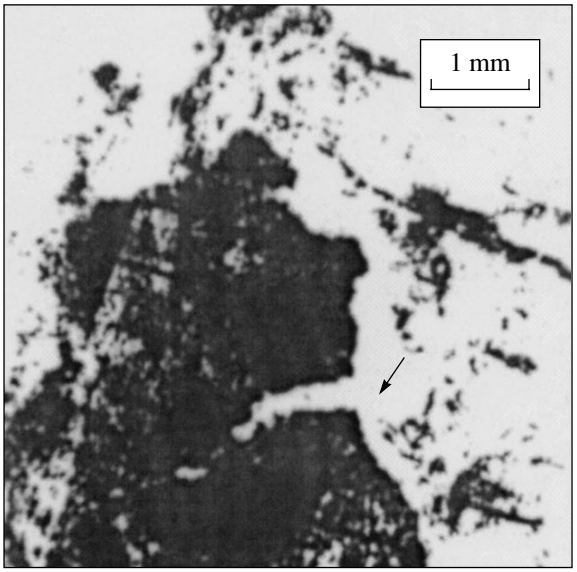

Fig. 5.

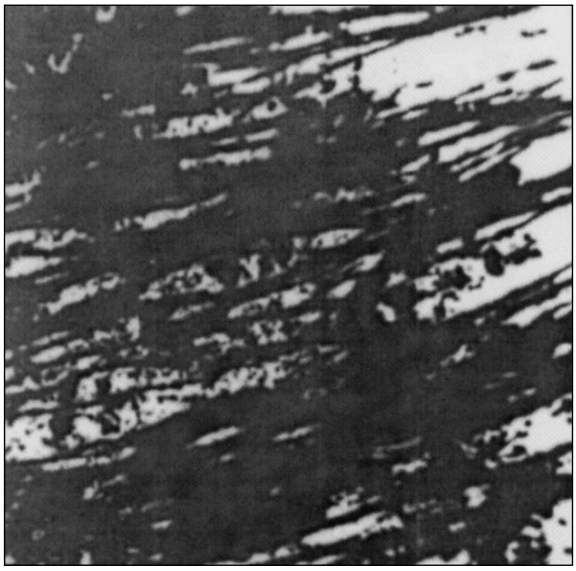

Fig. 6.

After the samples had been cooled to room temperature, the pattern containing light and dark regions was observed in crossed polaroids.

PARAMETERS OF THE SAMPLES AND CONDITIONS OF LB FILM PREPARATION

Orientation effects observed in LC cells are illustrated in Figs. 2–8. The conditions for film preparation are listed in the table.

According to previous studies [2–4], at low pressures (less than 20 dyn/cm), orientation effects in LC cells are virtually absent. In this case, the patterns obtained in crossed polaroids exhibits a grain structure, indicating the polycrystallinity of the LB films. The size of grains on the patterns corresponds to that of single-crystal domains in the film deposited. Estimated visually, this size varies from 20 to 300 μm.

ANALYSIS OF THE PATTERNS

Single monolayer. Though orientation effects are lacking in this case, the LC cell patterns may still contain useful information. For example, they can definitely testify to the fact that a single monolayer is deposited. Figure 2 shows the distinct boundary between the

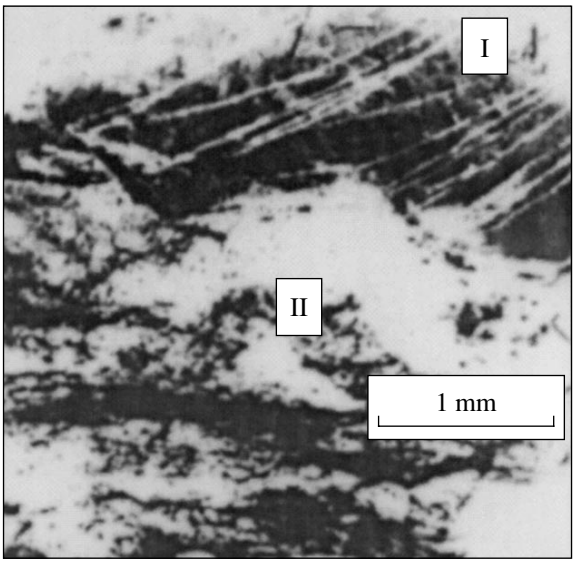

Fig. 7.

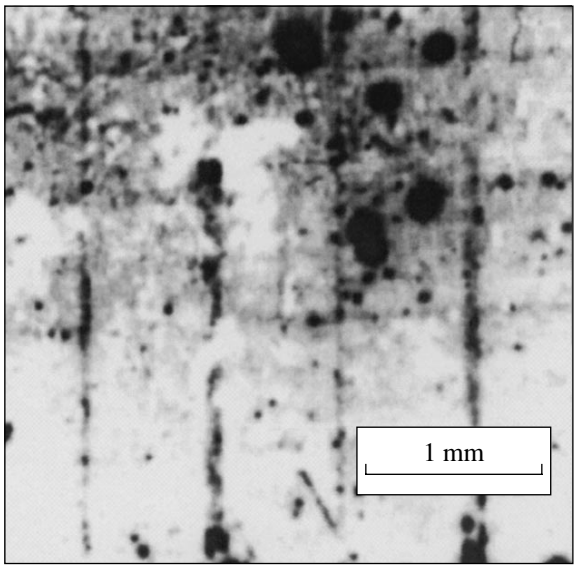

Fig. 8.

bare part of the substrate on the left and the part covered by one monolayer on the right. It should be noted that the grain size varies: the grain size between 100 and 130 μm becomes prevalent in the patterns of the LC cells with the LB film.

LB films deposited at different pressures. Collapse of LB film. Comparing the images in Figs. 2–4 indicates that an increase in the deposition pressure not only coarsens the grains up to 300 μm (Fig. 3) but also results in the formation of alignment texture (Fig. 4).

A further rise in the pressure in the monolayer leads to its collapse. Films that were transferred at the collapse pressure can easily be identified with the help of the LC cell. For example, Fig. 5 shows the clear-cut contour of an appreciable LB film fragment that broke down during deposition (marked by an arrow). Such cracks are typical of films compressed to the single-crystal state. In this state, the film loses elasticity and has no time to melt even at small variations of the working surface area of the cell.

Another type of collapse is illustrated in Fig. 6. Clearly distinguishable folds indicate the formation of a multilayer structure after the collapse of the monolayer. Such deformations are usually irreversible. When the excessive pressure is released, the folds in the film do not straighten out. Most often, a monolayer deposited after severe compression is fragmentary rather than continuous. Figure 7 shows two regions with different structures: region I, where the film becomes multilayer, and region II with the initial stage of collapse. Two interpretations of this pattern are possible: either the structure of the under layer may be seen through voids in the upper layer or the region that has become free of collapsed layer I is filled with layer II, which is formed by separated and unfolded pieces of layer I.

Influence of substrate relief (electrodes under the LB film). Usually, both the method of preparing the substrate and the initial relief on its surface are important for LB film deposition. The presence of electrodes on the glass substrate surface also influences the LC orientation in the cell. A cell with transparent ITO elec-

Table

Pattern no.	π, mN/m	ML transfer type	Number of MLs	Presence of electrodes	Notes
2	15	Z	1	No	ML boundary
3	20	Z	1	Yes	
4	25	Z	1	No	LC meniscus edge
5	35	Z	5	Yes	ML breaks
6	50(?)	Z	1	No	ML collapse
7	30(?)	Z	1	No	Expanded after collapse
8	–	–	–	Yes	

trodes applied in LC screens is shown in Fig. 8. Here, an LB film was not applied on the electrodes.

A comparison of Fig. 8 with Figs. 3 and 5 (see table for LB film deposition conditions) implies that the orientation effect due to the LB film is greater than that attributed to the electrodes. Note that the electrodes are about two orders of magnitude thicker than the LB film.

From the cell images, we can conclude that the orientation effects due to the LB films grow with both the number of layers and the deposition pressure. The images gain more contrast and sharpness. This may be illustrated by Fig. 5, where the electrodes are hardly seen under the film.

CONCLUSION

The aforesaid not only confirms the possibility of using LB films for orienting liquid crystals but also demonstrates the potentiality of the technique for observing the macrostructure of mono- and multilayer LB films. As it is rather illustrative, this method may also be very convenient for many applications, especially upon choosing a new material and refining the deposition technique.

ACKNOWLEDGMENTS

This work was supported by the Ministry of Education of the Russian Federation (grant no. T00-1.5-3225).

REFERENCES

1. L. M. Blinov, Usp. Fiz. Nauk **155** (3), 443 (1988) [Sov. Phys. Usp. **31**, 623 (1988)].
2. S. Baker, A. Seki, and J. Seto, Thin Solid Films **180**, 263 (1989).
3. S. N. Shtykov, B. N. Klimov, G. Yu. Naumenko, *et al.*, Zh. Fiz. Khim. **73**, 1711 (1999) [Russ. J. Phys. Chem. **73**, 1519 (1999)].
4. B. N. Klimov, E. G. Glukhovskoy, Yu. P. Volkov, *et al.*, in *Physics of Semiconductors and Semiconductor Electronics: Collection of Articles*, Ed. by B. N. Klimov and A. I. Mikhaĭlov (Gos. UNTs "Kolledzh," Saratov, 2001).

Translated by A. Sidorova

Technical Physics, Vol. 47, No. 12, 2002, pp. 1580–1583. Translated from Zhurnal Tekhnicheskoĭ Fiziki, Vol. 72, No. 12, 2002, pp. 98–101.

BRIEF COMMUNICATIONS

Generation of High-Power Electron Beams in Magnetron Guns with Cold Secondary-Emission Cathodes

A. N. Dovbnya, V. V. Zakutin, N. G. Reshetnyak, V. P. Romas'ko, Yu. Ya. Volkolupov, and M. A. Krasnogolovets

Uskoritel' Research Complex, Kharkov Institute of Physics and Technology, National Scientific Center, Akademicheskaya ul. 1, Kharkov, 61108 Ukraine

e-mail: dovbnya@nik.kharkov.ua

Received March 12, 2002

Abstract—The generation of powerful electron beams in a single magnetron gun and in a set of magnetron guns with cold secondary-emission cathodes is studied. The possibility of generating tubular electron beams of current 50–100 A, particle energy 30–100 kV, and pulse power 1–5 MW is shown. Such beams can be used as electron sources in the technology of accelerators, as well as in ordinary and multibeam high-power microwave devices. © *2002 MAIK "Nauka/Interperiodica".*

INTRODUCTION

In recent years, injection magnetron guns with normal [1–6] and reflex [6, 7] cold secondary-emission cathodes (SECs) have attracted particular interest. These sources offer a number of advantages (long service life, high current density, simple design, tubular beams, etc.), which allow for their use in accelerators, reliable powerful microwave sources [3], and high-speed high-voltage devices [8]. In injection magnetron guns with cold SECs, the cathode is subjected to backward bombardment by primary electrons (the secondary emission coefficient of the cathode is above unity), which gain energy when moving in a decaying electric field. Primary electrons can be produced by field (cold) emission, emission from insulating inclusions on the cathode surface, or emission from an additional thermionic cathode [6, 9]. Under these conditions, the electrons are multiplied by secondary emission, their density builds up in an avalanche-like manner, an electron sheath forms near the cathode, and an electron beam is formed and extracted from the gun. At the early stage of electron sheath formation, the secondary-emission multiplication process takes place because the electrons, when following a cycloidal path, gain energy in the decaying electric field. Once the electrons have been accumulated (steady-state stage), the process is governed by electric fields due to space-charge density oscillations. Of interest is the stable generation of electron beams with a high peak power in injection magnetron guns with cold SECs. In this work, we study the generation of high-power electron beams in a single gun and in a set of the guns. We also trace a correlation between the beam current and the electric and magnetic fields, as well as evaluate the transverse size of the beam.

EXPERIMENT

Experiments on generating powerful electron beams in magnetron guns with SECs were performed with the setup shown in Fig. 1. The magnetron gun is fed by modulator *1*, which forms a voltage pulse of amplitude 4–200 kV, duration ≈4 μs, and pulse repetition rate 10–50 Hz. A negative-polarity flat-top pulse with an overshoot is applied to copper cathode *5*. Anode *6* (made of stainless steel or copper) is grounded through the resistor R_3. The secondary-emission process is triggered in a decaying electric field produced by the falling edge (duration 0.6 μs, steepness 50–100 kV/μs) of an overshoot that is specially produced on the top of the cathode voltage pulse [4]. The magnetic field is generated by solenoid *4*. The beam current is measured with Faraday cup *7* and resistor R_4; the cathode voltage, with the divider R_1R_2; the anode current, with the resistor R_3; and the beam size, with prints on an X-ray film and molybdenum foil that are placed on the Faraday cup. The magnetron gun is positioned in stainless-steel vacuum chamber *3*, where the pressure is kept at a level of

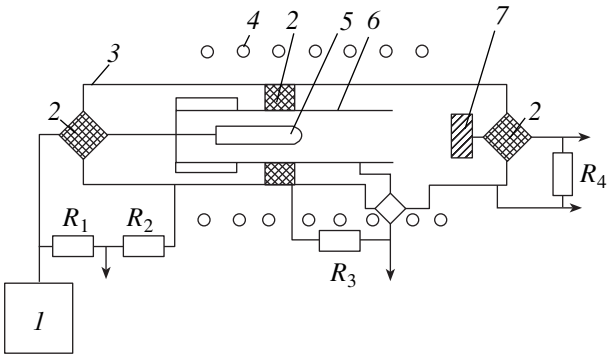

Fig. 1. Experimental setup.

1063-7842/02/4712-1580$22.00 © 2002 MAIK "Nauka/Interperiodica"

$\sim 10^{-6}$ torr and is insulated from the chamber with insulators *2*.

In the experiments, two approaches to generate magnetic fields necessary for beam generation and transportation were employed: a pulsed discharge through a solenoid [10] and dc feeding of a solenoid [11]. Figure 2 shows the distribution of the magnetic field along the solenoid axis for both cases, as well as the arrangement of the magnetron gun and the Faraday cup.

In the former case (Fig. 2, curve *1*), the axial field of the solenoid can be very uniform (±5%) and high (5000–6000 Oe). However, one should take into account the decay of the pulsed field when it diffuses through the walls of the cavity and vacuum chamber, which may change the longitudinal distribution of the pulsed magnetic field (see, e.g., [9, 12]). The repetition rate of magnetic field pulses depends on the value of the reservoir capacitor and the selection of the switch. If the pulse rate is low, the cathode surface is contaminated for the time between the pulses. Due to electron bombardment, the contaminants may fall into the anode–cathode gap, causing its vacuum breakdown [9].

In the latter case, a high density of the feed power and water cooling are necessary to provide a high permanent magnetic field by passing current through the solenoid. Here, the field strength is limited by the heat being released in the solenoid and the magnetic field distribution along the solenoid axis is less uniform (Fig. 2, curve *2*). Also, due to current ripple in the solenoid, the repetition rate of pulses driving the modulator should be related to the mains frequency.

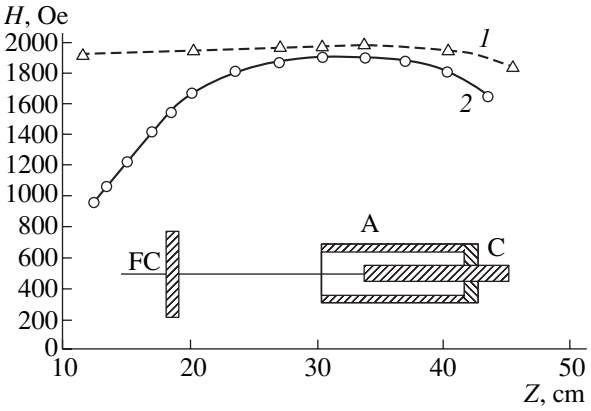

Fig. 2. Distribution of the magnetic field along the solenoid axis and the arrangement of the gun (A, anode; C, cathode) and the Faraday cup (FC). (*1*) Pulsed magnetic field; (*2*) static magnetic field.

RESULTS AND DISCUSSION

1. Generation of High-Power Beams in a Single Magnetron Gun

In this case, high-power beams can be produced if the cathode diameter is large. This is because the beam current (and consequently, beam power) can be increased by raising the voltage across the cathode–anode gap. However, the voltage cannot exceed the breakdown voltage of the gap. The larger the cathode diameter, the lower the electric field strength at the cathode and the lower the breakdown probability. Therefore, higher cathode–anode voltages can be used. In our experiments, the cathode diameter in the magnetron guns were varied from 40 to 80 mm; the anode diameter, from 50 to 140 mm; and the voltage amplitude, from 20 to 120 kV.

In a ≈100-mm-long single magnetron gun, stable beam generation was achieved at a cathode diameter of 40 mm, anode diameter of 78 mm, and voltage amplitude of 100 kV (hereafter, the flat-part amplitude is meant). The beam current was found to be about 50 A, which corresponds to a microperveance of ≈1.6 and a beam power of ≈5 MW. The magnetic field strength was ≈1800 Oe in this case. Figure 3 plots the beam current at the Faraday cup against the cathode voltage pulse amplitude. It is seen that the beam current obeys the Child–Langmuir law. During the measurements, to each voltage value there corresponded the optimum value of the magnetic field at which the beam current amplitude was the highest. It turned out that the gun readily withstands the voltage 120 kV. Hence, a peak power of up to 8 MW can be obtained. We also measured the electron current to the anode. It varied from 1 to 10% of the beam current, depending on the experimental conditions.

Next, we studied the beam parameters as functions of the electric and magnetic fields. When the cathode voltage amplitude is higher or lower than the optimal value (with the magnetic field fixed), the conditions for secondary-emission multiplication are violated and the process ceases. If the top of a voltage pulse is harmonically modulated, electron bunches arise at the exit of the gun at those instants of time at which the sinusoid falls (forced modulation [4]). The beam current vs. magnetic field dependence shows that, as the magnetic field grows (with the cathode voltage amplitude fixed), the beam current amplitude at the Faraday cup first grows sharply, then exhibits a plateau, and sharply falls. Such behavior reflects changes in the electron trajectories and in the conditions under which the electrons gain energy in the anode–cathode gap with increasing magnetic field. When the magnetic field amplitude varies, so does the shape of the beam current pulse. At a cathode voltage of 55 kV and a magnetic field of 700 Oe, the beam current pulse breaks down into spikes of amplitude ≈30 A and duration of 10–30 ns. When the magnetic field grows to ≈1200 Oe, the beam current pulse becomes flat.

In a ≈100-mm-long magnetron gun with a copper cathode of diameter 40 mm and a stainless-steel anode of diameter 50 mm, we obtained a beam current of 50 A at a voltage of 30 kV and a magnetic field of 2200 Oe.

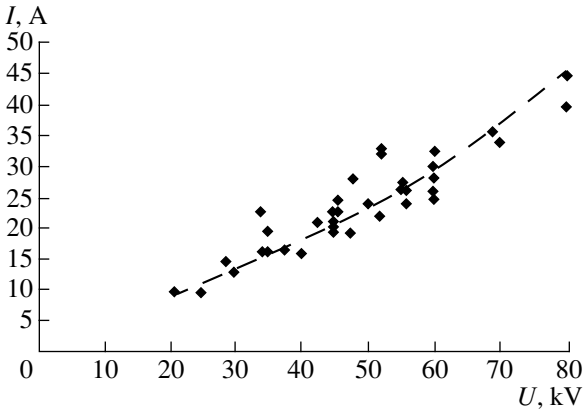

Fig. 3. Beam current at the Faraday cup vs. cathode voltage amplitude. Dashed line describes the Child–Langmuir law.

This corresponds to a microperveance of ≈10 and a pulse power of ≈1.5 MW.

In [6], single magnetron guns of both normal (cathode diameter 50 mm, anode diameter 60 mm) and reflex type (cathode diameter 54 mm, anode diameter 43 mm) generated a beam current of ≈100 A at a voltage of 40 kV in a magnetic field of ≈1800 Oe (the cathodes were made of the BeCu alloy and the anodes, of stainless steel). This corresponds to the beam power ≈4 MW and the microperveance 12, which agrees with our results.

2. Production of Multiple Electron Beams in a Set of Magnetron Guns

Another way of increasing the beam power is the use of a system of parallel-connected injection magnetron guns. The beam current, power, microperveance, and size for each of the guns are small, whereas the total beam current and power of the system are significant.

The generation of stable electron beams with a system of parallel-connected SEC magnetron guns may pose difficulties like those associated with parallel-connected switches. When a partial beam is generated or one of the guns breaks down, the voltage amplitude at the cathodes of other guns decreases. This may quench generation or cause beam instability. In actual systems, a voltage decrease at the cathodes of the other guns due to parasitic inductances and capacitances lasts several nanoseconds. It was shown [13] that the formation time of the electron beam in SEC guns (hence, the rise time of the beam current pulse) may be decreased to 2 ns (if secondary-emission multiplication is triggered by the nanosecond trailing edge of a voltage pulse with a steepness of more than 300 kV/μs). Thus, in a system of the guns, the formation of the electron sheath and the electron beam is possible even if one of the guns breaks down. If the feeding modulator has a low output resistance, the amplitude of the generation-driving pulse decreases insignificantly (in our experiments, by less than 20% at an output resistance of the modulator of 2000 Ω). It was demonstrated [14] that the beam generation continues when the cathode voltage changes by 30% (with the magnetic field fixed), which exceeds the above value. Thus, the difficulties associated with the beam generation in a system of the guns are quite surmountable.

In a set of magnetron guns, the electric strength can by reduced because of an increase in the total surface area of the cathodes (the breakdown voltage is inversely proportional to $S^{0.1}$, where S is the surface area of the cathodes). If the number of the guns is eight or more, the decrease in the breakdown voltage is 15–20%. In a set of eight guns with cathode and anode diameters of 5 and 22 mm, respectively, the breakdown voltage was shown to diminish roughly by 30% (from 70 to 50 kV) as compared with a single gun with the same electrode geometry. Note that this decrease is not only due to the above effect but also results from the asymmetric arrangement of the guns, poor vacuum conditions, and other adverse factors.

In our experiments, high-power beams were generated in a system of six and eight guns. The cathodes and anodes were made of copper, and the gun length was ≈100 mm.

In the former case, the system generated six beams. All the guns were arranged in a circle of diameter 60 mm. The anode and cathode diameters were 5 and 26 mm, respectively. For a cathode voltage of ≈40 kV and a magnetic field strength of ≈2000 Oe, each of the guns generated a beam current of 18 A with a microperveance of ≈2.5. The total beam current was ≈100 A with a peak power of ≈4 MW.

In the latter case, eight guns were arranged in a circle of diameter ≈70 mm. The anode and cathode diameters were 5 and 22 mm, respectively. For a cathode voltage of ≈30 kV and a magnetic field strength of ≈2000 Oe, the total beam current was ≈60 A with a peak power of ≈2 MW. The microperveance of each of the beams was ≈2.

3. Beam Size

In the SEC guns, the beams traveled a distance of 50–100 mm from the anode surface. It was found that they have a ring-shaped cross section with a uniform azimuth distribution of intensity. The inner diameter of the ring is roughly equal to the cathode diameter, and the "wall" thickness is 1–2 mm.

The outer and inner diameters of the beam generated in the single magnetron gun (with anode and cathode diameters of 40 and 70 mm, respectively) were measured to be 45 and 41 mm, respectively. For the magnetron diode (single gun) with cathode and anode diameters of 80 and 140 mm, respectively, it was found that, when the electric field is nonuniform within 5% in the transverse direction, the distinct beam trace in the azi-

muth direction is surrounded by a wide diffuse region. For a more nonuniform field, the beam is disrupted in the azimuth direction (over a length of about 20 mm). This indicates that the coaxiality of the electrodes must meet stringent requirements.

In the system of magnetron guns, the beams were found to be arranged in a circle of a given diameter (60 or 70 mm). The sizes of each of the beams equaled those of the beam in the single gun (the outer and inner diameters were, respectively, ≈9 and ≈5 mm). The intensity of the beams was uniformly distributed in the azimuth direction.

CONCLUSION

Our experiments show that magnetron guns with cold secondary-emission cathodes can provide high-power electron beams. We studied the correlation between the beam current and the electric and magnetic fields. Beam generation in a set of magnetron guns and its electrical strength were considered.

REFERENCES

1. J. F. Skowron, Proc. IEEE **61** (3), 69 (1973).
2. S. A. Cherenshchikov, Élektron. Tekh., Ser. 1, No. 6, 20 (1973).
3. A. N. Dovbnya, V. V. Zakutin, V. F. Zhiglo, *et al.*, in *Proceedings of the Fifth European Particle Accelerator Conference, EPAC*, Ed. by S. Myers, A. Pacheco, R. Pascual, *et al.* (Institute of Physics Publishing, Bristol, 1996), Vol. 2, p. 1508.
4. A. N. Dovbnya, V. V. Mitrochenko, V. V. Zakutin, *et al.*, in *Proceedings of the 1997 Particle Accelerator Conference, Vancouver, 1997*, Ed. by M. Comyn, M. K. Craddok, M. Reiser, and J. Thomson, 1997, Vol. 3, p. 2820.
5. A. V. Agafonov, V. P. Tarakanov, and V. M. Fedorov, Vopr. At. Nauki Tekh., Ser. Yad.-Fiz. Issled. **1** (2–3), 134 (1997).
6. Y. M. Saveliev, W. Sibbbett, and D. M. Parkes, Phys. Plasmas **4**, 2319 (1997).
7. G. I. Churyumov, Radioélektronika, No. 7, 77 (1997).
8. A. I. Vishnevskiĭ, A. I. Soldatenko, and A. I. Shendakov, Izv. Vyssh. Uchebn. Zaved., Ser. Radioélektron. **11**, 555 (1968).
9. A. N. Dovbnya, V. V. Zakutin, N. G. Reshetnyak, *et al.*, Vopr. At. Nauki Tekh., Ser. Yad.-Fiz. Issled., No. 1 (28), 53 (1997).
10. H. Knoepfel, *Pulsed High Magnetic Fields* (North-Holland, Amsterdam, 1970; Mir, Moscow, 1972).
11. I. I. Kifer, *Ferromagnetic Materials Tests* (Gosénergoizdat, Moscow, 1962).
12. V. V. Zakutin and A. M. Shenderovich, Prib. Tekh. Éksp., No. 5, 238 (1976).
13. N. I. Aĭzatskiĭ, A. N. Dovbnya, V. V. Mitrochenko, *et al.*, Vopr. At. Nauki Tekh., Ser. Yad.-Fiz. Issled., No. 3 (34), 38 (1999).

Translated by V. Isaakyan

Technical Physics, Vol. 47, No. 12, 2002, pp. 1584–1586. Translated from Zhurnal Tekhnicheskoĭ Fiziki, Vol. 72, No. 12, 2002, pp. 102–104.

BRIEF COMMUNICATIONS

Long-Lived Plasmoids Produced in Humid Air as Analogues of Ball Lightning

A. I. Egorov and S. I. Stepanov

Konstantinov Institute of Nuclear Physics, RAS, Gatchina, Leningrad oblast, 188300 Russia

Received June 11, 2002.

Abstract—The results of an experimental study of long-lived plasmoids, analogues of ball lightning, are presented. The study focuses mainly on the establishment of the conditions for the prolonged existence of plasmoids in the atmosphere. Experimental setup is described and the observation data displayed in the form of photographic pictures. © *2002 MAIK "Nauka/Interperiodica".*

The large-scale questioning of eye-witnesses became the basic method of study of this natural phenomenon [1–5]. Analyses of the subjective evidence, rich in conjectures and fantasies, gave birth to scores of hypotheses as to its nature. However, there is still no conclusive evidence that ball lightning, i.e., an airborne, long-lived, isolated, and luminous formation, has been reliably reproduced in a laboratory.

The most general conditions required for the origin of ball lightning are known: high electrical activity of the atmosphere and a high concentration of water vapor in the air [2]. Water vapor possesses a number of unique properties: it is lighter than air, does not support combustion, and inhibits the oxidation of inflammable substances, but at the same time, it is capable of interacting with heated carbon or iron with the formation of molecular hydrogen. An extremely inhomogeneous distribution of electron density across an H_2O molecule induces large effective charges on the hydrogen and oxygen atoms. Dipole water molecules attach to free ions and envelope aerosol particles. The binding energy of the first H_2O molecule to the H^+ ion is 7.18 eV, for the second, 1.6 eV, and for the fifth, 0.51 eV [6]. The binding energy of water dipoles to other ions, both positive and negative, is of about the same magnitude.

In studies [6, 7], computer simulation of the interaction of water molecules with H^+ and OH^- ions in air saturated with water vapor is reported. It is shown that, as hydrated ions of opposite sign come close together, additional water molecules are drawn into the gap between them and a stable cluster is formed. This cluster consists of two ions of opposite sign and a hydrate envelope. Water molecules prevent the ions from coming close enough for recombination; therefore, the lifetime of the ions in the cluster rises to tens of minutes, i.e., by 12–13 orders of magnitude. On the other hand, because of the dipole–dipole interaction of the clusters, first chains and then spatial structures arise; i.e., a cloud of cold plasma or plasmoid forms. The internal energy of a plasmoid is insufficient for the ionization of its molecules but can accumulate a huge amount of energy (up to 500 J/l) from an external source. This energy should be given to a plasmoid at its formation.

In Fig. 1, a diagram of the setup is shown in which a large population of ions can be introduced into a cloud of warm air saturated with water vapor. The main part of the setup comprises a bank of 0.6 mF capacitors; this bank could be charged to 5.5 kV. A polyethylene vessel 18 cm in diameter was filled to a height of 15 cm with weakly conducting tap water. A ring-shaped copper electrode was placed at the bottom of the vessel and connected by an insulated copper busbar to one terminal of the bank of capacitors. The second terminal was connected to a cylindrical electrode positioned at the center of the vessel close to the water surface and pointing into the half-space of air. As the central electrode, a rod of spectral-pure carbon 5–6 mm in diameter was used in most cases. The quartz tube surrounding the central electrode protruded above it by 2–4 mm and above the water by 3–8 mm. The electrical resistance of the water contained between the submerged electrodes was 1.0–1.2 kΩ.

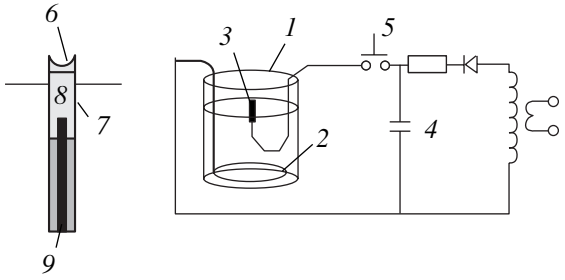

Fig. 1. Setup for generating long-lived plasmoids: (*1*) polyethylene vessel, (*2*) ring electrode, (*3*) central electrode, (*4*) bank of 0.6 mF capacitors, (*5*) discharger, (*6*) drop of water or water suspension, (*7*) quartz tube, (*8*) carbon or metal electrode, and (*9*) copper busbar.

1063-7842/02/4712-1584$22.00 © 2002 MAIK "Nauka/Interperiodica"

To produce an artificial ball lightning, two to three drops of water (about 0.1 ml) were applied to the central electrode of carbon. Fast closing-breaking of the discharger caused a plasma jet to issue from the central electrode with a light popping sound and an isolated luminous plasmoid separated from the jet. It slowly floated up in the air and in 0.2–0.5 s disintegrated. The detachment of the plasmoid from the plasma jet is shown in Fig. 2.

The ring made of a thin (0.1 mm) copper or nichrome wire weighed on an analytical balance placed in the path of the ball lightning was partially sputtered and in the plasma jet even melted, probably on account of intensive recombination on the metal surface. The ends of the retrieved wire fragments carried characteristic fused balls (Fig. 3).

In most cases, the central electrode was connected to the negative terminal of the bank of capacitors. At some optimum potential difference between the electrodes, the plasmoid assumed a rounded shape. For the setup with the given dimensions, this potential difference was 4.2–4.8 kV. The discharge current did not exceed 30–50 A.

In the horizontal electrical field produced by a plane condenser, the ball lightning stretched to the sides and its lifetime decreased. The approximate equality of volume charges of the positive and negative ion populations does not exclude the possession of a small excess charge by the plasmoid.

The size of plasmoids differed between the discharges but was predominantly in the range 10–18 cm. The coloring of the "water" ball lightning is close to that of a gas discharge generated in humid air at low pressure. The lilac-colored central part of the plasmoid is surrounded by a diffuse yellowish envelope. A small admixture of sodium or calcium salts gives the plasmoid core a yellow or orange color.

In changing to iron, copper, or aluminum central electrodes, the basic features of the phenomenon remained the same. The pulsed discharge from a metal electrode wetted with water without organic additives produced a bright plasma jet that dislodged a flying plasmoid. Its coloring depended on the emission spectrum of the excited atoms in the electrode: iron plasmoids were whitish, copper ones were greenish, and aluminum ones were white with a red glint.

The ball lightning temperature can be estimated from the velocity of its vertical ascent. If the plasmoid is considered as a cloud of warm humid air 14 cm in diameter ascending in the atmosphere at 293 K with a velocity of 1.0–1.2 m/s, then the calculated average temperature of the plasmoid does not exceed 330 K.

The lifetime of an artificial ball lightning depends on quite a number of factors: the size and geometry of the central electrode, the voltage across the electrodes, the amplitude and duration of the current pulse, and the temperature and electrical conductivity of water applied to the central electrode. It is also possible to change the lifetime of the plasmoid by introducing a dispersed phase into it. This is done by applying a suspension of some substance to the central electrode: it is sputtered and dispersed by the pulsed discharge and gets into the plasmoid together with the ion population.

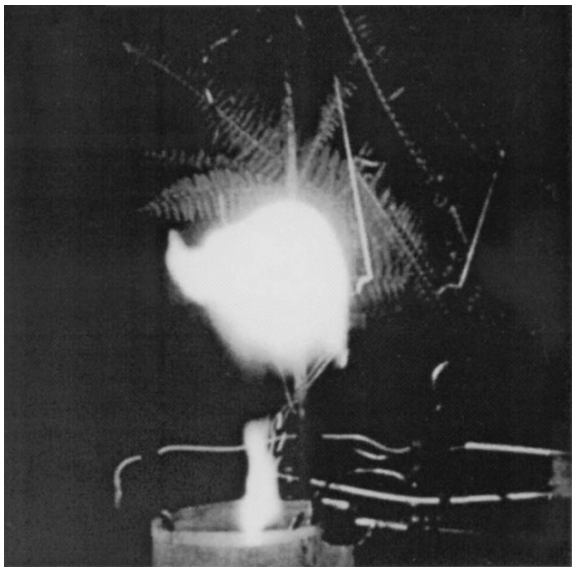

Fig. 2. Floating plasmoid at the moment of detachment from the plasma jet. The diameter of the vessel with water is 18 cm.

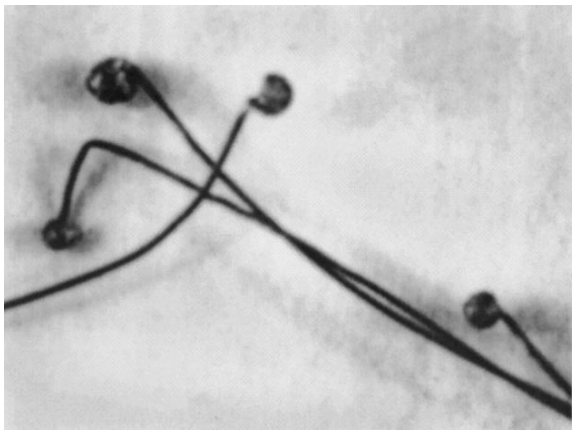

Fig. 3. Enlarged photograph of the wires after contact with the plasma jet.

Scores of substances were tried: carbon powders, hydrocarbons with a high heat of combustion, soot, a mixture of carbon and fine silica, carbonyl iron powder, a colloidal suspension of Fe_3O_4, clay, samples of soil, sawdust, colophony, and other natural substances. For the detailed study, colloidal graphite and fine iron oxides were chosen, although it is possible that other more suitable suspensions exist. It is only necessary that the suspension have high electrical resistivity.

In most of the experiments, a suspension of 3 g colloidal graphite, 8–10 ml acetone (wetting agent), and

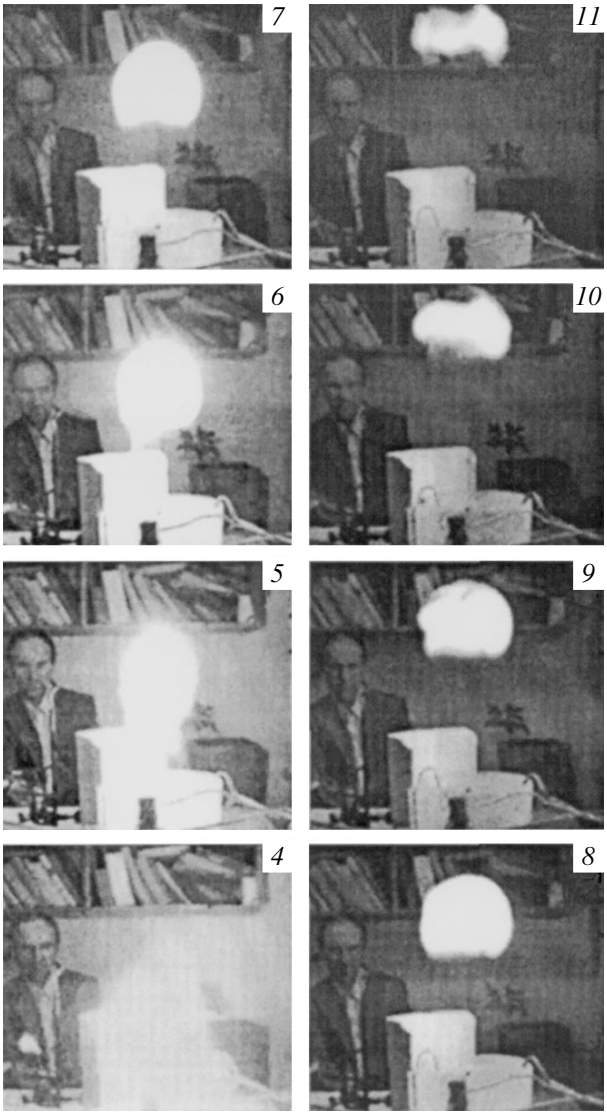

Fig. 4. Stages of origination, flight, and disintegration of the artificial ball lightning produced using a colloidal graphite, acetone, and water. Figures are the frame numbers in the videofilm. The time interval between the frames is 0.2 s. The complete film "fball02.avi" can be seen at ftp://biod.pnpi.spb.ru/pub/people/stepanov.

90 ml of water was applied to the central electrode. Under an electric discharge through the layer of this suspension there emerged a floating, rounded plasmoid (Fig. 4), which slowly ascended in the air and disappeared in 0.3–0.8 s; long-lived plasmoids rarely occurred. The core of the plasmoid had a flame coloring, i.e., the dispersion of carbon and destruction of the organic substances under the pulsed discharge produced glowing particles of soot and chemical radicals.

Thus, as a result of introducing a considerable population of ions of both signs into a cloud of warm air saturated with water vapor, a rounded luminous long-lived plasmoid emerges. This plamoid accumulates the energy spent by the electrical discharge upon formation of the ion pairs and conserves it for a long time. The plasmoid temperature slightly exceeds room temperature, and the plasmoid disintegrates with an explosion only on the rare occasion when it contains a lot of hydrogen. On metal surfaces, the recombination rate of the ions is much higher and, at the same time, aerosols of fine carbon and oxidized iron introduced into the plasmoid do not reduce its lifetime. Laboratory discharges are not as grandiose as natural discharges; however, artificial ball lightnings are readily reproducible and accessible for detailed study.

Artificial ball lightning is an exceedingly beautiful physical phenomenon. Luminous globes floating up in a darkened room are an unforgettable sight, especially for those who have happened to observe short-lived natural ball lightning. For such people, the similarity of the two phenomena is undoubted.

ACKNOWLEDGMENTS

The authors are deeply indebted to G. D. Shabanov, S. E. Emelin, and A. L. Pirozerskiĭ for invaluable help in experimental work and to V. M. Lobashev for his constant interest and support.

REFERENCES

1. J. D. Barry, *Ball Lightening and Bead Lightening: Extreme Forms of Atmospheric Electricity* (Plenum, New York, 1980; Mir, Moscow, 1983).
2. I. P. Stakhanov, *On Physical Nature of Ball Lightening* (Nauchnyĭ Mir, Moscow, 1996).
3. S. Singer, *Nature of Ball Lightning* (Plenum, New York, 1971; Mir, Moscow, 1973).
4. A. I. Grigor'ev, *Fire Killers* (Debyut, Yaroslavl', 1990).
5. O. A. Sinkevich, Teplofiz. Vys. Temp. **35** (4), 651 (1997); **35**, 968 (1997).
6. S. V. Shevkunov, Zh. Éksp. Teor. Fiz. **119**, 485 (2001) [JETP **92**, 420 (2001)].
7. S. V. Shevkunov, Dokl. Akad. Nauk **379**, 181 (2001) [Dokl. Phys. **46**, 467 (2001)].

Translated by B. Kalinin