Technical Physics, Vol. 48, No. 6, 2003, pp. 661–668. Translated from Zhurnal Tekhnicheskoĭ Fiziki, Vol. 73, No. 6, 2003, pp. 1–8.

THEORETICAL AND MATHEMATICAL PHYSICS

Transport in the Two-Terminal Aharonov–Bohm Ring

V. A. Geyler, V. V. Demidov, and V. A. Margulis

Ogarev Mordovian State University, Saransk, 430000 Russia

e-mail: geyler@mrsu.ru

Received February 8, 2002

Abstract—The transmission coefficient of a nanodevice—a quantum ring with two one-dimensional conductors attached—is found. The Hamiltonian of a nanodevice is constructed in terms of the theory of self-adjoint extensions of symmetric operators. It is shown that, in this case, the transmission coefficient coincides with that determined by the Feynman sum rule for the probability amplitudes. The transmission coefficient of the nanodevice is studied as a function of the electron energy, magnetic field, and the relative positions of the conductor contacts and the ring. © *2003 MAIK "Nauka/Interperiodica".*

INTRODUCTION

New physical phenomena observed in curvilinear nanostructures, such as a ring quantum interferometer [1–4], provoke interest in their theoretical study [5–13]. The simplest model of such a nanostructure is a one-dimensional ring with one-dimensional conductors attached symmetrically to it. This model, which was first proposed in [5], was then studied in a number of papers [6–12]. As applied to transport problems, such an ideal model is a crucial simplification of real conditions, because it allows one to describe the electron motion only by a longitudinal part of the wave function (single-mode regime). An advantage of this model is that it results in a simple explicit expression for the transmission coefficient of a nanodevice that is convenient for analysis. However, this model requires matching the wave functions of the channels and the ring at the contact points. The numerical results obtained for the multimode regime show that, in this case, an important role is played by the electron modes reflected back into the channels from the contact regions [13]. There are two fundamentally different approaches to this problem. The most widely used approach proposed in [6] involves introducing a priori the scattering matrix of a contact that satisfies both the current density conservation and time-reversal symmetry. This implies that the complex scattering matrix should be a unitary symmetric matrix of the third order. Such a matrix can be described by five independent real parameters. These parameters are empirical parameters that should be specified by using additional and, sometimes, nonphysical arguments, such as the requirement for the matrix elements to be real, the smallness of some of the matrix elements, etc. In greater detail, this approach was analyzed in [9].

Papers [14–23] are based on a fundamentally different approach. In those papers, it was proposed to describe electrons in a nanodevice by the perturbation of a Hamiltonian H_0 that is a direct sum of the Hamiltonians of the channels (H_1 and H_2) and the one-dimensional ring (H_r): $H_0 = H_r \oplus H_1 \oplus H_2$. This perturbation is produced in two steps. First, H_0 is restricted within the functions that vanish at the contact points, and then the symmetric operator obtained is extended to a self-adjoint one. All such extensions are described by the boundary conditions at the contact points; these conditions are similar to those in the theory of zero-range potentials [24, 25]. In this model, the parameters of the boundary conditions are empirical parameters. In contrast to the theory of zero-range potentials, the parameters form a non-diagonal Hermitian matrix. The parameters of the boundary conditions gain physical sense when they are related to the corresponding scattering lengths.

In this paper, we use the second approach. In some sense, this approach is more general as compared to the first one. In particular, the results of [6–11] can be obtained in the framework of the second approach if one assumes that the parameters of the boundary conditions depend on energy as const/$\sqrt{E}$. This corresponds to a self-adjoint extension into a wider space of states. Such models are constructed and discussed in detail in [24].

HAMILTONIAN AND KREIN FUNCTION OF A NANODEVICE

As was noted above, our model consists of a ring $\mathbb{S}_\rho = \{x \in \mathbb{R}^2: |x| = \rho\}$ of radius ρ with half-lines W_1 and W_2 attached to the ring at points A_1 and A_2 (Fig. 1). We will use polar coordinates on the ring $\mathbb{S}_\rho$ such that the points A_j correspond to angles ϕ_j and the half-lines W_j correspond to two instances of the positive semiaxis $\mathbb{R}_j^+ = \{x \in \mathbb{R}: x \geq 0\}$ ($j = 1, 2$) (thus, the point $0 \in \mathbb{R}_j^+$ will correspond to the point $\phi_j \in \mathbb{S}_\rho$). We assume that the ring is subject to a magnetic field perpendicular to

1063-7842/03/4806-0661$24.00

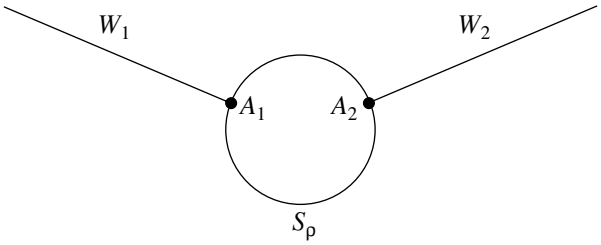

Fig. 1. Two-channel Aharonov–Bohm ring. The dots denote the points where channels are attached to the ring.

the plane of the ring; we also assume that the field magnitude depends only on the distance from the center of a ring. In this case, the Hamiltonian of an electron in the ring has the form

$$H_r = \frac{\hbar^2}{2m^*\rho^2}\left(-i\frac{d}{d\phi} + \eta\right)^2. \quad (1)$$

Here, m^* is the effective electron mass and η is the number of quanta of the magnetic field flux through the ring; i.e., $\eta = 2\pi\Phi_0^{-1}\int_0^\rho B(r)dr$, where $\Phi_0 = 2\pi\hbar c/|e|$ is the quantum of the magnetic flux. In an isolated channel $\mathbb{R}_j^+$, the operator

$$H_j = -\frac{\hbar^2}{2m^*}\frac{d^2}{dx^2}$$

with the Neumann condition $\psi_j'(0) = 0$ at the point $x = 0$ is used as an unperturbed Hamiltonian. Thus, the electron dynamics in the system consisting of an isolated ring $\mathbb{S}_\rho$ and channels $\mathbb{R}_j^+$ is described by the self-adjoint operator $H_0 = H_r \oplus H_1 \oplus H_2$.

The wave function ψ of any nanodevice consisting of a ring $\mathbb{S}_\rho$ and channels $\mathbb{R}_j^+$ is represented by a column matrix $\psi = (\psi_r, \psi_1, \psi_2)^t$, where ψ_r and ψ_j are the functions at the ring $\mathbb{S}_\rho$ and the channels $\mathbb{R}_j^+$, respectively. If ψ_r, ψ_1, and ψ_2 are equal to zero at the points A_j, the value of $\langle\psi|H|\psi\rangle$ for the Hamiltonian H of a nanodevice should coincide with $\langle\psi|H_0|\psi\rangle$. Hence, H is a self-adjoint extension of a symmetric operator V obtained by restricting H_0 within the functions ψ vanishing at the points A_1 and A_2. It is well known that all the self-adjoint extensions of the operator V are described by the boundary conditions at the points A_j [24]. The boundary values for ψ are represented by $\psi_i(0)$, $\psi_j'(0)$, and $\psi_r(0)$, as well as by the jumps of the derivatives $\psi_r'(\phi_j + 0) - \psi_r'(\phi_j - 0)$ (here, $\psi_r'(\phi)$ denotes the derivative with respect to the curve length $s = \rho\phi$). Hereafter, we restrict our consideration to the boundary conditions of a rather general form:

$$\begin{cases}\psi_r(\phi_j) = \beta_j[\psi_r'(\phi_j + 0) - \psi_r'(\phi_j - 0)] + \alpha_j\psi_j'(0)\\ \psi_j(0) = \bar{\alpha}_j[\psi_r'(\phi_j + 0) - \psi_r'(\phi_j - 0)] + \gamma_j\psi_j'(0)\end{cases} \quad (2)$$

($j = 1, 2$).

The parameters α_j, β_j, and γ_j make up the matrix

$$P = \begin{pmatrix}\beta_1 & 0 & \alpha_1 & 0\\ 0 & \beta_2 & 0 & \alpha_2\\ \bar{\alpha}_1 & 0 & \gamma_1 & 0\\ 0 & \bar{\alpha}_2 & 0 & \gamma_2\end{pmatrix}. \quad (3)$$

The hermicity of the matrix P provides the self-adjointness of H. Thus, the Hamiltonian in our model is described by eight independent real parameters. The substitution of the zero elements of the matrix P for nonzero ones without violating the hermicity of P will also yield a self-adjoint (but, in some sense, nonlocal) operator [26]. On the other hand, a nontrivial scattering matrix for H will take place only for nonzero parameters α_j, β_j, and γ_j. Further, we will use the so-called Krein Q matrix for a pair of operators (V, H_0). This matrix has the form [15]

$$Q(E) = \begin{pmatrix}Q_{11}^{(r)}(E) & Q_{12}^{(r)}(E) & 0 & 0\\ Q_{21}^{(r)}(E) & Q_{22}^{(r)}(E) & 0 & 0\\ 0 & 0 & Q_{33}(E) & 0\\ 0 & 0 & 0 & Q_{44}(E)\end{pmatrix}. \quad (4)$$

Here, $Q_{ij}^{(r)}(E) = G_r(\phi_i, \phi_j; E)$ and $Q_{i+2,\,i+2}(E) = G_i(0, 0; E)$ ($i, j = 1, 2$), where G_r and G_j are the Green functions of the Hamiltonians H_r and H_j, respectively. These well known functions are

$$G_r(\phi, \phi'; E) = \frac{m^*}{2\hbar^2 k}\left[\frac{\exp(i(\phi' - \phi \pm \pi)(\eta - k\rho))}{\sin\pi(\eta - k\rho)} - \frac{\exp(i(\phi' - \phi \pm \pi)(\eta + k\rho))}{\sin\pi(\eta + k\rho)}\right] \quad (5)$$

(the plus sign takes place when $\phi \geq \phi'$ and the minus sign, in the opposite case)

$$G_j(x, x'; E) = \frac{im^*}{\hbar^2 k}[\exp(ik(x + x')) + \exp(ik|x - x'|)], \quad (6)$$

where $k = \sqrt{2m^*E}/\hbar$.

As was shown in [15], the transmission coefficient of a nanodevice is expressed via the matrix B ($B = (B_{ij})_{1 \leq i, j \leq 4}$) adjoint to $Q + 2m^*P/\hbar^2$ or, more precisely,

via the matrix element B_{43}

$$B_{43} = -4(m^*)^2\alpha_1\bar{\alpha}_2 Q_{21}^{(r)}/\hbar^4. \quad (7)$$

TRANSMISSION COEFFICIENT OF A NANODEVICE

Let a wave $\exp(ikx)$ incident onto the ring propagate along the first channel $\mathbb{R}_1^+$. Then, the wave function in the channel $\mathbb{R}_1^+$ of the device has the form [15, 27]

$$\psi_1(x) = \exp(ikx) + \left(1 - \frac{4im^*}{\hbar^2 k}\frac{B_{33}}{\det(Q + 2m^*P/\hbar^2)}\right)\exp(-ikx) \quad (8)$$

in the channel $\mathbb{R}_1^+$ and

$$\psi_2(x) = -\frac{4im^*}{\hbar^2 k}\frac{B_{43}}{\det(Q + 2m^*P/\hbar^2)}\exp(ikx) \quad (9)$$

in the channel $\mathbb{R}_2^+$.

Here, $Q(E)$ and B_{43} are given by formulas (4) and (7), respectively. Comparing expressions (8) and (9), we find the transmission amplitude of a nanodevice:

$$t(E) = -\frac{4im^*B_{43}}{\hbar^2 k\det(Q + 2m^*P/\hbar^2)}. \quad (10)$$

Then, the transmission coefficient is $T(E) = |t(E)|^2$. It follows from Eq. (3) that, in the framework of our model, the transmission coefficient is a function of eight independent real parameters. Let us introduce the scattering length λ for a charged particle scattered by a point potential and assume for simplicity that all the parameters of the boundary conditions are equal to $\lambda/2$. This means that we consider equally perturbed channels and a ring and that the wave functions in both the channels and the ring are matched in the same way. It can easily be seen that, in this case, $Q_{33} = Q_{44} = 2im^*/\hbar^2 k$ and the quantities $Q_{11}^{(r)}$ and $Q_{22}^{(r)}$ are equal to

$$Q_{11}^{(r)} = Q_{22}^{(r)} = \frac{m^*}{\hbar^2 k}\frac{\sin kL\cos kL}{\sin^2\pi\eta - \sin^2 kL} \quad (11)$$

for a real k. Here, $L = \pi\rho$ is the length of the half-ring. For contacts symmetrically attached to diametrically opposite points of the ring ($\phi_1 = 0$ and $\phi_2 = \pi$), the off-diagonal elements $Q^{(r)}$ have the form

$$Q_{12}^{(r)} = Q_{21}^{(r)} = \frac{m^*}{\hbar^2 k}\frac{\sin kL\cos\pi\eta}{\sin^2\pi\eta - \sin^2 kL}, \quad (12)$$

and the matrix determinant is

$$\det Q^{(r)} = \left(\frac{m^*}{\hbar^2 k}\right)^2\frac{\sin^2 kL}{\sin^2\pi\eta - \sin^2 kL}. \quad (13)$$

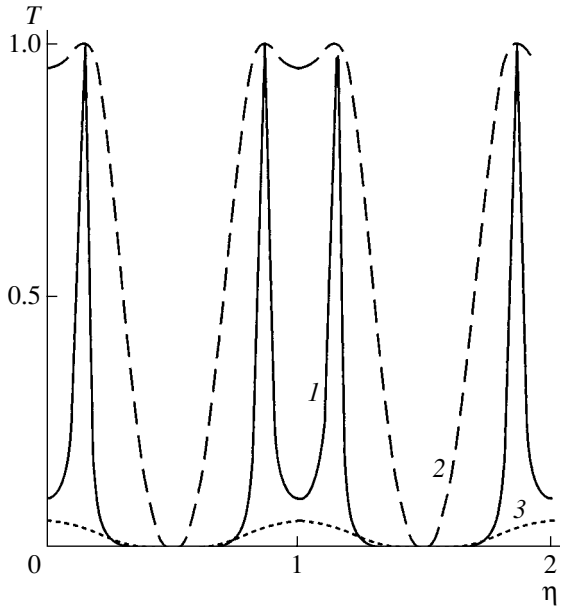

Fig. 2. Transmission coefficient T of the nanodevice vs. magnetic field flux at fixed values of the electron energy: kL = (*1*) 2.9π, (*2*) 2.7π, and (*3*) 2.4π. The scattering length for electrons in the region of contacts is λ = 75 nm. Here and in the subsequent figures, the radius of the ring is ρ = 350 nm.

After simple algebra, expressions (10)–(13) result in

$$T(k,\eta) = \left\{1 + \left[\left(\frac{3}{4} - \frac{1}{\lambda^2 k^2}\right)\sin kL - \frac{2}{\lambda k}\cos kL - \frac{\sin^2\pi\eta}{\sin kL}\right]^2\frac{1}{\cos^2\pi\eta}\right\}^{-1}. \quad (14)$$

It follows from Eq. (14) that $T(k, \eta)$ is a periodic function of the flux η with a period equal to the flux quantum (Fig. 2). At $kL = \pi n$, where $n = 1, 2, \ldots$ (i.e., when the electron energies coincide with the energy levels of electrons in an isolated ring), the transmission coefficient $T(k, \eta)$ turns to zero.

Let us first consider the dependence of T on k in the case of a zero magnetic field (the same result takes place in the case of an integer magnetic flux through the ring). The corresponding plots are shown in Fig. 3. It can be seen that the oscillations of the curve $T(k)$ are not equidistantly distributed along the kL axis. It follows from Eq. (14) that the peaks of $T(k)$ are located at the points corresponding to the roots of the transcendental equation

$$2\cot kL = \frac{3}{4}\lambda k - \frac{1}{\lambda k}. \quad (15)$$

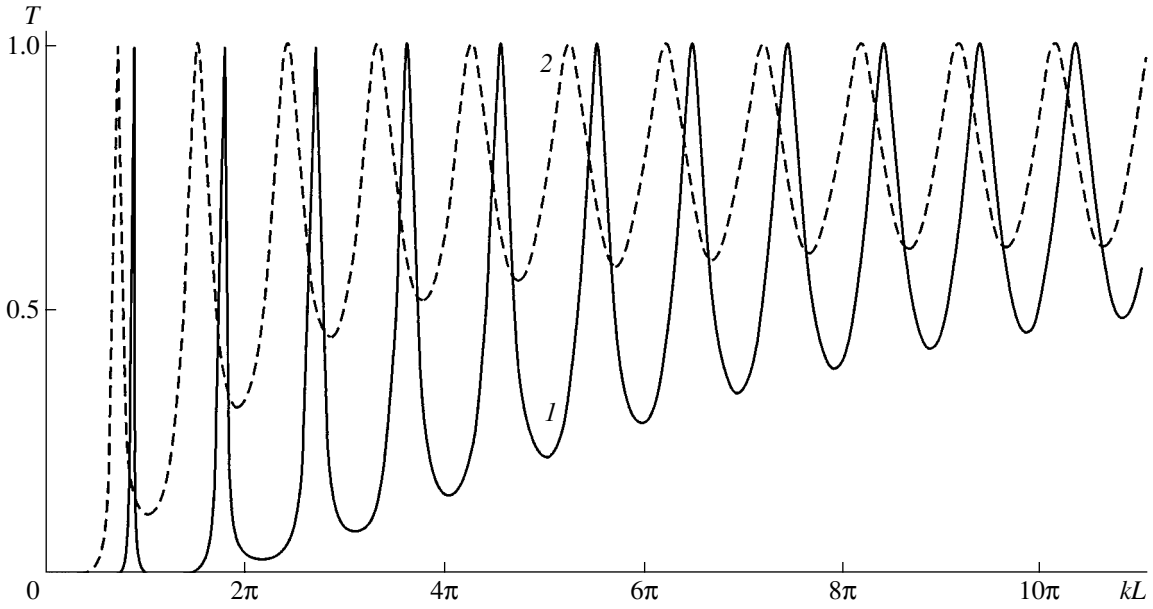

Fig. 3. Transmission coefficient T of the nanodevice vs. electron wavenumber at zero (or integer-valued) magnetic field flux through the ring for two values of the scattering length: λ = (*1*) 75 and (*2*) 250 nm.

These roots have the following asymptotics at $k \longrightarrow \infty$:

$$k = \frac{\pi n}{L} + \frac{8L}{3\pi\lambda n} + O\left(\frac{1}{n^2}\right). \quad (16)$$

The peak values of the amplitudes of all the oscillations are equal to unity. The minima of the oscillations are located at the envelope

$$f(k) = \left(\frac{5}{4} + \frac{1}{\lambda^2 k^2}\right)^{-2}. \quad (17)$$

It follows from boundary conditions (2) that, in the case of a single contact, the expressions for the transmission (t) and reflection (r) amplitudes have the form

$$t = \frac{2\lambda k}{3\lambda k + 2i}, \quad r = -\frac{\lambda k + 2i}{3\lambda k + 2i}. \quad (18)$$

Under the assumptions adopted, these expressions are the same for the scattering at any contact and for any direction of electron motion (from the ring to the channel and vice versa). It follows from Eq. (18) that the phase of the wave function changes by

$$\delta(k) = \arctan\frac{2}{3\lambda k} \quad (19)$$

when an electron passes through the contact. That is why the peaks of $T(k)$ are not spaced equidistantly, especially at relatively low k values. The change in the sign of λ yields the corresponding change in the sign of $\delta(k)$.

Now, we consider a nanodevice subject to a magnetic field. The corresponding plots are shown in Fig. 4. The zeros of $T(k, \eta)$ appear at the points $kL = \pi n$ (n = 1, 2, …) only in the case of noninteger flux η (the zeros of $T(k, \eta)$ coincide with those of $Q_{21}^{(r)}$, whereas the element $Q_{21}^{(r)}$ has no zeros for an integer flux). In the case of a half-integer flux η, $T(k, \eta)$ is zero at any k. These interference effects can be explained as follows. In a magnetic field, the phase shifts of the wave functions of the electrons propagating along the different half-rings have opposite signs $k\rho + \eta$ and $k\rho - \eta$. At odd values of $k\rho$, the imaginary parts of the wave functions have opposite signs and are canceled (no matter how low the magnetic field B and, hence, the phase shift $\pi\eta$). The real parts of the wave functions of the electrons that have passed through the entire ring in the opposite directions are canceled at the entry point. The zero value of $T(k, \eta)$ for a half-integer flux η is explained in a similar way with the only difference that, in this case, the phase difference of the wave functions of the electrons that have passed through the different half-rings is a multiple of π. Therefore, the wave functions are canceled at the exit point at any k.

Additional peaks in the plot of $T = T(k, \eta)$ of unit height (only in the case of a noninteger magnetic field flux) are caused by the term $\sin^2\pi\eta/\sin kL$ in Eq. (10). This term results in additional roots of the expression in brackets. Any peak of the function $T(k, \eta)$ is caused by the interference of the wave functions of the electrons that experience multiple reflections at the points of contacts between the conductors and the ring. An asymmetric distribution of the peaks is explained by the phase shift of the wave functions at the points of contacts.

Let us consider a ring with asymmetrically attached channels ($\phi_1 = 0$ and $\phi_2 \neq \pi$). In this case, the diagonal

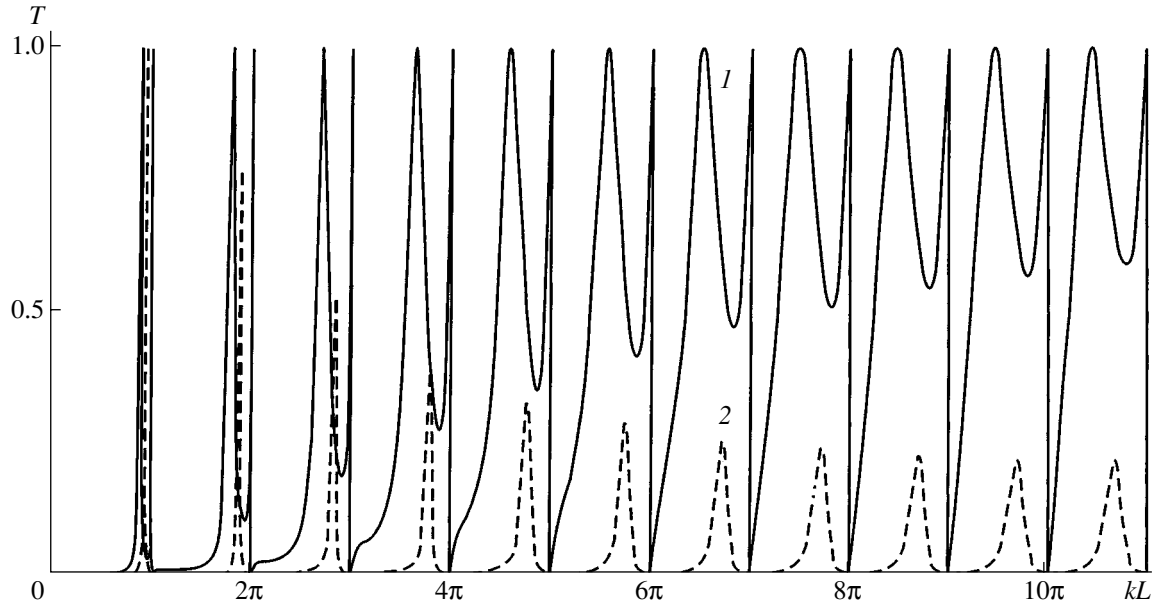

Fig. 4. Transmission coefficient T of the nanodevice vs. electron wavenumber in the presence of a magnetic field for η = (*1*) 0.14 and (*2*) 0.47. Here and in the subsequent figure, the scattering length is λ = 75 nm.

elements $Q^{(r)}$ are again given by formulas (11) and the off-diagonal elements have the form

$$Q_{12}^{(r)} = \overline{Q_{21}^{(r)}} = \frac{2m^*}{\hbar^2 k}\exp(-i(\pi-\phi_2)\eta)\frac{\sin kL\cos\pi\eta\cos(\pi-\phi_2)k\rho + i\cos kL\sin\pi\eta\sin(\pi-\phi_2)k\rho}{\sin^2\pi\eta - \sin^2 kL}. \quad (20)$$

The determinant $Q^{(r)}$ is

$$\det Q^{(r)} = 2\left(\frac{m^*}{\hbar^2 k}\right)^2\frac{\sin^2 kL - \sin^2(\pi-\phi_2)k\rho}{\sin^2\pi\eta - \sin^2 kL}. \quad (21)$$

Then, the transmission coefficient is written as

$$T(k,\eta) = [1 + F_1^2(k,\eta)/F_2(k,\eta)]^{-1}, \quad (22)$$

where

$$F_1(k,\eta) = \left(\frac{3}{4} - \frac{1}{\lambda^2 k^2}\right)\sin^2 kL - \frac{1}{\lambda k}\sin 2kL + \left(\frac{1}{4} + \frac{1}{\lambda^2 k^2}\right)\sin^2(\pi-\phi_2)k\rho - \sin^2\pi\eta, \quad (23)$$

$$F_2(k,\eta) = \sin^2 kL\cos^2\pi\eta\cos^2(\pi-\phi_2)k\rho + \cos^2 kL\sin^2\pi\eta\sin^2(\pi-\phi_2)k\rho. \quad (24)$$

Let us first analyze the expression for the transmission coefficient in the absence of a magnetic field. In this case, expressions (23) and (24) are reduced to

$$F_1(k,0) = \left(\frac{3}{4} - \frac{1}{\lambda^2 k^2}\right)\sin^2 kL - \frac{1}{\lambda k}\sin 2kL + \left(\frac{1}{4} + \frac{1}{\lambda^2 k^2}\right)\sin^2(\pi-\phi_2)k\rho, \quad (25)$$

$$F_2(k,0) = \sin^2 kL\cos^2(\pi-\phi_2)k\rho. \quad (26)$$

Formulas (22), (25), and (26) result in the peaks of $T(k)$ appearing in the zeros of the function $F_1(k)$. In the case of arbitrary positions of contacts (in contrast to the case of symmetric contacts), the peaks of T have a doublet structure for low-energy electrons. The reason is that expression (25) has additional roots at $k < 2/\lambda\sqrt{3}$ due to the presence of a factor by the last term [in the expression in brackets in Eq. (14), this factor is constant]. The zeros of the function $T(k)$ coincide with those of F_2: they appear at the points $kL = \pi n$ ($n \in \mathbb{N}$), except for both the points $q\pi n$, at which the function F_1 turns to zero, and the points $kL = q\pi(1/2 + n)$, at which q is the denominator of the fraction $\phi_2 = p\pi/q$. Figure 5a shows the plot of the function $T(k)$ in the absence of a magnetic filed. Because of the factor $\cos^2(\pi - \phi_2)k\rho$ in the function F_2, the curve $T(k)$ consists of wave packets of width $\pi^2/|\pi - \phi_2|$ (except for the first packet, whose width is two times smaller). We note that the shapes of the additional oscillatory peaks that appear in the case of asymmetric contacts are similar to those in the case of symmetric contacts in a magnetic field. This is related to the fact that the magnetic field shifts the phase of the wave function, and, hence, the effective path of the electrons that move along the different half-rings either increases or decreases.

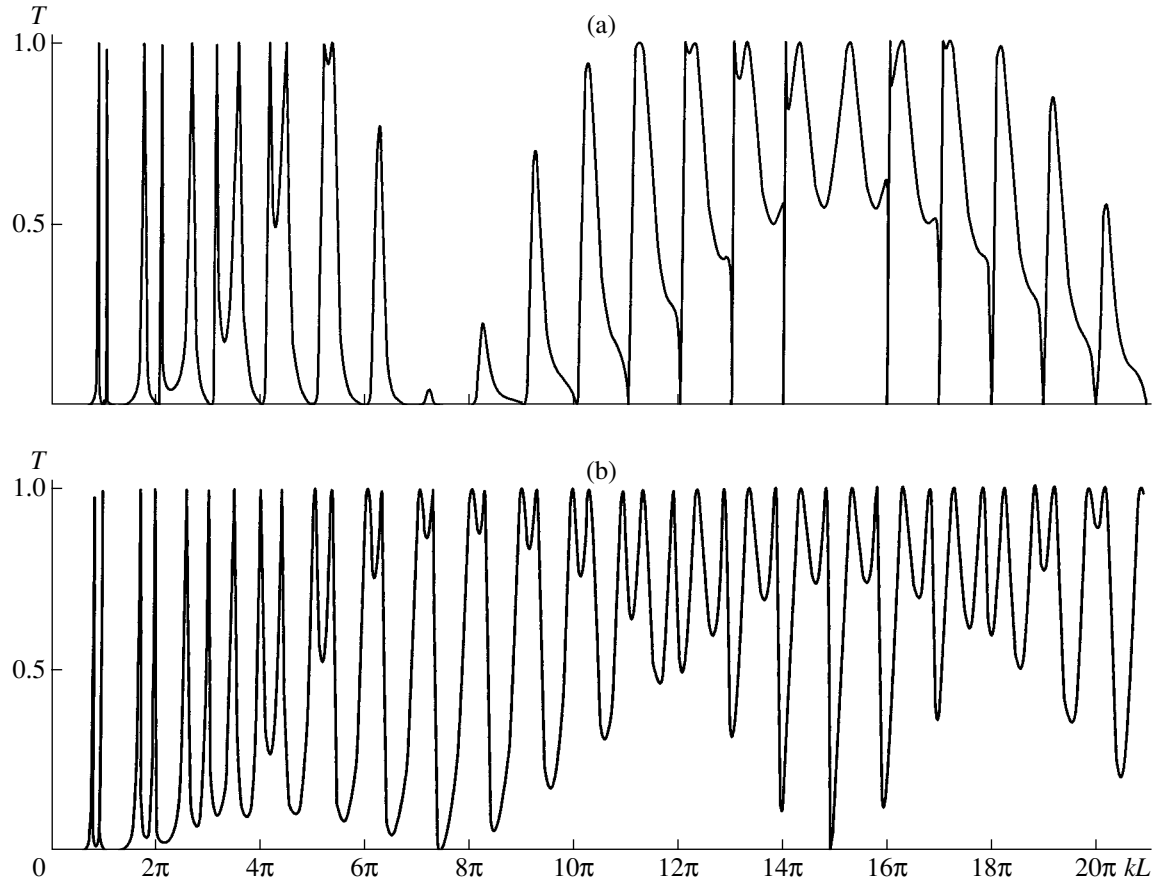

Fig. 5. Transmission coefficient T of the nanodevice vs. electron wavenumber in the case of an asymmetric location of the contacts on the ring ($\phi_1 = 0$, $\phi_2 = 14\pi/15$) in the absence (a) and in the presence (b) of a magnetic field ($\eta = 0.07$).

In the case of a nanodevice subject to a magnetic field (with a flux that is neither integer nor half-integer), the peaks of $T(k)$ are smoothed and the oscillatory peaks are similar to beatings (Fig. 5b). In this case, the zeros at the points $kL = \pi n$ (where $n \in \mathbb{N}$) are lifted and turn into minima, whereas the zeros are retained at the points $kL = q\pi n$ and $kL = q\pi(1/2 + n)$. The last circumstance is related to the specific behavior of the functions F_1 and F_2, which are determined by expressions (23) and (26).

CALCULATION OF THE TRANSMISSION COEFFICIENT OF A RING WITH THE HELP OF THE FEYNMAN SUM RULE FOR THE PROBABILITY AMPLITUDES

In this section, we will show that the expression of the transmission amplitude $t^{(r)}$ for a one-dimensional Aharonov–Bohm ring with two channels attached can be found using the Feynman sum rule for the amplitudes of the wave functions of electrons describing all the topologically different trajectories in a ring. For simplicity, we assume that identical contacts ($\alpha_\pm = \beta_\pm = \gamma_\pm = \lambda/2$) are symmetrically attached to the ring ($\phi_1 = 0$ and $\phi_2 = \pi$) and the magnetic field flux is equal to zero. In this case, expression (10) for the transmission amplitude of a nanodevice can be rewritten as

$$t^{(r)}(k) = \frac{\lambda^2 k^2}{(\lambda^2 k^2 + 2i\lambda k)\cos kL - i(5\lambda^2 k^2/4 + i\lambda k - 1)\sin kL}. \quad (27)$$

Let the incident wave $\exp(ikx)$ propagate along the left channel. We consider wave functions of electrons describing all the possible trajectories in a ring and entering the right channel. The coefficients of transmission and reflection from a channel into any half-ring and vice versa are determined by formulas (18). There are only two waves passing the half-ring path: the wave moving along the upper half-ring and that moving along the lower half-ring (at the exit they have the form $t^2\exp(ikL)$). Each of these waves has a component passing into the ring again. Then, in view of the device symmetry, it is sufficient to consider only the waves moving along the upper half-ring and double the result, because the waves moving along the lower half-ring make the same contribution. There are four waves reflected back into a ring and passing the path $3L$: $t^4\exp(ikL)$, $t^2rt\exp(ikL)$, $trt^2\exp(ikL)$, and $t^2r^2\exp(ikL)$. Each of

these four waves has a component passing into a ring again. Summing up all the waves entering the right channel, we will find the sum in the form of the geometric progression

$$\begin{aligned} t^{(r)} &= 2[t^2\exp(ikL) + t^2(t^2 + 2tr + r^2)\exp(3ikL) \\ &\quad + t^2(t^2 + 2tr + r^2)^2\exp(5ikL) + \ldots] \\ &= 2t^2\exp(ikL)[1 + (t + r)^2\exp(2ikL) \\ &\quad + (t + r)^4\exp(4ikL) + \ldots] \\ &= \frac{2t^2}{\exp(-ikL) - (t + r)^2\exp(ikL)}. \end{aligned}$$

Then, in view of formulas (18), we find

$$\begin{aligned} t^{(r)} &= \frac{8\lambda^2k^2}{(3\lambda k + 2i)^2\exp(-ikL) - (\lambda k - 2i)^2\exp(ikL)} \\ &= \frac{\lambda^2k^2}{(\lambda^2k^2 + 2i\lambda k)\cos kL - i(5\lambda^2k^2/4 + i\lambda k - 1)\sin kL}. \end{aligned} \quad (28)$$

It can be seen that expressions (27) and (28) coincide. Similar, but rather cumbersome, calculations can also be performed in the general case.

CONCLUSIONS

As follows from the foregoing, the dependence of the transmission coefficient of a nanodevice on the electron energy is of an oscillatory nature. The oscillations result from the interference of the electron wave functions in a ring. In the general case of contacts asymmetrically attached to a ring ($\phi_1 = 0$ and $\phi_2 \neq \pi$) in the absence of a magnetic field, the peaks make up packets (Fig. 4). At comparatively low energies, these peaks have a doublet structure. One of the peaks in the doublet structure decreases with increasing energy and evanesces at sufficiently high energies. The width of each packet is $\pi^2/|\pi - \phi_2|$; the only exception is the first packet, whose width is two times smaller. As the magnetic field increases, the packets vanish, the oscillatory peaks are smoothed out, and the peaks acquire the form of beating.

In a special case of contacts symmetrically attached to the ring ($\phi_1 = 0$, $\phi_2 = \pi$) in the absence of a magnetic filed, the minima of $T(k)$ are located at envelope (17) and the maxima of T are equal to unity. At positive values of the scattering length λ, the maxima of $T(k)$ are located to the right of the points πn ($n = 0, 1, \ldots$) and asymptotically approach them with increasing k. At negative λ, the maxima of $T(k)$ are located to the left of the points $n = 1, 2, \ldots$ and asymptotically approach them with increasing k. Similar behavior of $T(k)$ takes place in the model considered in [12].

In the presence of a magnetic field with a flux that is neither integer nor half-integer, the curve $T(k)$ has deep (down to zero) minima at the points corresponding to the energy levels of an isolated ring. At $\lambda > 0$, sharp peaks of unit height appear to the left of the minima. When the magnetic field flux approaches a half-integer value, the minima widen, the peaks are smoothed out, and the transmission coefficient vanishes for all k. In contrast to the curves $T(k)$ found in [6], in our case, $T(k)$ is not a strictly periodic function of the wavenumber k. At large k, the curve $T(k)$ shown in Fig. 4 qualitatively coincides with the curves plotted in [6]; however, it is strongly asymmetric at low energies. The strict periodicity in k derived in [6] is an artifact of the model, which is close in this sense to strong coupling models, in which all zones have the same structure.

ACKNOWLEDGMENTS

This work was supported by the Russian Foundation for Basic Research (project no. 01-02-16564 a), DFG (grant no. 436 RUS 113/572/1), and INTAS (grant no. 00-257).

REFERENCES

1. J. Liu, K. Ismail, K. Y. Loe, *et al.*, Phys. Rev. B **47**, 13039 (1993).
2. A. F. Morpurgo, J. P. Heida, B. J. van Wees, *et al.*, Physica B **251**, 509 (1998).
3. H. R. Shea, R. Martel, and Ph. Avouris, Phys. Rev. Lett. **84**, 4441 (2000).
4. S. Pedersen, A. E. Hansen, A. Kristensen, *et al.*, Phys. Rev. B **61**, 5457 (2000).
5. M. Büttiker, Y. Imry, and R. Landauer, Phys. Lett. A **96A**, 365 (1983).
6. M. Büttiker, Y. Imry, and M. Ya. Azbel, Phys. Rev. A **30**, 1982 (1984).
7. Y. Gefen, Y. Imry, and M. Ya. Azbel, Phys. Rev. Lett. **52**, 129 (1984).
8. D. Takai and K. Ohta, Phys. Rev. B **48**, 14318 (1993).
9. I. A. Ryzhkin, Fiz. Tverd. Tela (St. Petersburg) **41**, 2070 (1999) [Phys. Solid State **41**, 1901 (1999)].
10. J.-B. Xia, Phys. Rev. B **45**, 3593 (1992).
11. J. Li, Z.-Q. Zhang, and Y. Liu, Phys. Rev. B **55**, 5337 (1997).
12. N. T. Bagraev, A. D. Buravlev, V. K. Ivanov, *et al.*, Fiz. Tekh. Poluprovodn. (St. Petersburg) **34**, 846 (2000) [Semiconductors **34**, 817 (2000)].
13. O. A. Tkachenko, V. A. Tkachenko, D. G. Baksheev, *et al.*, Pis'ma Zh. Éksp. Teor. Fiz. **71**, 366 (2000) [JETP Lett. **71**, 255 (2000)].
14. P. Exner and P. Šeba, in *Ideas and Methods in Quantum Physics*, Ed. by S. Albeverio *et al.* (Cambridge Univ. Press, Cambridge, 1992), Vol. 2, p. 227.
15. V. A. Geyler and I. Yu. Popov, Teor. Mat. Fiz. **107** (1), 12 (1996).
16. B. S. Pavlov, I. Yu. Popov, V. A. Geyler, *et al.*, Europhys. Lett. **52**, 196 (2000).
17. P. Kurasov and F. Stenberg, Preprint No. 7 (Stockholm University, Stockholm, 2001).

18. V. A. Geyler and A. V. Popov, Rep. Math. Phys. **42**, 347 (1998).
19. V. B. Bogevolnov, A. B. Mikhailova, and B. S. Pavlov, Preprint No. 413 (University of Auckland, Auckland, 1999).
20. P. Exner, M. Tater, and D. Vanek, J. Math. Phys. **42**, 4050 (2001).
21. C. Texier and G. Montambaux, J. Phys. A **34**, 10307 (2001).
22. A. B. Mikhailova, B. S. Pavlov, I. Yu. Popov, *et al.*, Preprint No. 420 (University of Auckland, Auckland, 2000).
23. A. B. Mikhailova and B. S. Pavlov, Preprint No. 474 (University of Auckland, Auckland, 2001).
24. B. S. Pavlov, Usp. Mat. Nauk **42** (6), 99 (1987).
25. S. Albeverio, F. Gesztesy, R. Hoegh Krohn, and H. Holden, *Solvable Models in Quantum Mechanics* (Springer-Verlag, New York, 1988; Mir, Moscow, 1991).
26. Yu. G. Shondin, Teor. Mat. Fiz. **106** (2), 179 (1996).
27. J. Brüning and V. A. Geyler, *Scattering on Compact Manifolds with Infinitely Thin Horns* (Humboldt University, Berlin, 2001).

Translated by M. Fofanov

Technical Physics, Vol. 48, No. 6, 2003, pp. 669–676. Translated from Zhurnal Tekhnicheskoĭ Fiziki, Vol. 73, No. 6, 2003, pp. 9–16.

THEORETICAL AND MATHEMATICAL PHYSICS

Charge and Force Acting on a Spherical Body near a Planar Electrode in a Polarizable Conducting Medium

A. A. Shutov

Karpov Research Institute of Physical Chemistry, Russian State Scientific Center, Obninsk branch, Obninsk, Moscow oblast, 249030 Russia

e-mail: fci@meteo.ru

Received August 12, 2002; in final form, December 3, 2002

Abstract—A method for calculating electric fields in conducting polarizable media with interface is suggested. An integral equation for the density of surface charge induced at the interface is derived. The value of this density is used to find the field in the volume. The total charge induced at the interface and the force acting on a spherical body touching a planar electrode are calculated. It is found that the total charge and the force are alternating functions of the relative conductivity of the media; that is, both repulsion from and attraction to the electrode are possible depending on the conductivity. The near-electrode force acting on solid particles, bubbles, and drops in an immiscible liquid is studied experimentally. *© 2003 MAIK "Nauka/Interperiodica".*

INTRODUCTION

Interaction with electrodes and the electrification of bodies are basic processes in filtration, sedimentation, and contact charging of particles [1–3]. Filtration technologies involve sedimentation processes, which can be described by various physicochemical mechanisms [1], and the problem of holding particles on filtering elements is very acute. It is generally believed that, when reaching the surface, particles stay on it. Empirical data, however, indicate that fine particles of diameter <0.1 μm, for which the effect of van der Waals forces is significant, adhere to the surface most strongly. Coarser particles, after being charged, may leave the electrode surface. Here, the conductivity of participating bodies and environmental parameters control the situation. In particular, particles are repelled from electrodes if the conductivity of the bodies do not exceed $\sim 10^{-2}$ $(\Omega\ \mathrm{m})^{-1}$ [2, 3]. Otherwise, electrodes hold particles. However, quantitative relationships that specify repulsion or attraction conditions are still lacking.

The electrostatic approach, which considers perfect conductors with induced surface charge or perfect insulators with induced polarization charge, fails in describing particle–electrode interaction and particle charging in a conducting medium. In this work, an electric field in conducting polarizable media will be determined by including both free and polarization charges of density controlled by charge carrier transfer.

STATEMENT OF THE PROBLEM

Consider the stationary problem of charge transfer through the surface of a bounded body placed in an electric field. We assume that the body and the environment have different conductivities and permittivities. Charge transfer is described by the equations [4]

$$\nabla \cdot \mathbf{D} = \rho, \tag{1}$$

$$\nabla \cdot \mathbf{j} = 0 \tag{2}$$

with boundary conditions

$$\varphi_e = \varphi_i, \tag{3}$$

$$D_{ne} - D_{ni} = \sigma, \tag{4}$$

$$\lambda_e E_{ne} = \lambda_i E_{ni}. \tag{5}$$

Here, φ is the electric field potential, $\mathbf{E} = -\nabla\varphi$, $\mathbf{D}_{i,e} = \varepsilon_0\varepsilon_{i,e}\mathbf{E}_{i,e}$, $\mathbf{j}_{i,e} = \lambda_{i,e}\mathbf{E}_{i,e}$, ρ is the volume charge density, σ is the surface charge density at the interface, λ is the conductivity, ε is the permittivity, and ε_0 is the dielectric constant. The subscript n by the field and induction means the projection onto the normal that is outer to the body, and the subscripts i and e refer to internal and external parameters relative to the body, respectively. As follows from conditions (3) and (4), the problem is considered for a charge monolayer.

In the general case of inhomogeneous media, the conductivity and permittivity are coordinate dependent. Then, the compatibility condition for Eqs. (1) and (2) requires the volume charge to be introduced into the problem. In homogeneous media, on the contrary, electric charge transfer proceeds in the absence of the volume charge and relationships (1) and (2) are replaced by the Laplace equation

$$\Delta\varphi = 0, \tag{6}$$

where Δ is the Laplacian.

However, in the presence of an interface, a surface charge is certain to be induced at it. The amount of this

1063-7842/03/4806-0669$24.00 © 2003 MAIK "Nauka/Interperiodica"

surface charge is given by Eq. (4). If the subject of investigation is the amount of transferred electric charge, it will suffice to consider the problem stated with (3), (5), and (6) [4].

A solution to the problem stated with (3)–(6) will be sought using the Grinberg method. According to this method, a solution to the electrostatic problem stated with (3), (4), and (6) at a point **r** can be represented as [5]

$$\varphi(\mathbf{r}) = \varphi_0(\mathbf{r}) + \frac{1}{4\pi\varepsilon_0}\oint_S \frac{\sigma_{tot}(\mathbf{r}')dS}{|\mathbf{r}-\mathbf{r}'|}, \quad (7)$$

where φ_0 is the external field potential away from the body.

The total surface charge density $\sigma_{tot} = \sigma_p + \sigma$ is the sum of the polarization (bound) charge density, σ_p, and physical (free) charge density, σ. Integration in (7) is over points **r**' on the body surface. The Grinberg method uses duality in the description of an electric field in polarizable media [5, 6]. The field can be described by introducing permittivity and postulating some relationship between the electric strength, **E**, and electric induction, **D**, vectors or by introducing an additional electric field through the polarization vector (polarization charges). Both approaches are equivalent and differ only in the way ε is taken into account (explicitly or implicitly). For example, potential (7) does not involve ε, since the medium in this case is characterized by using polarization charges. The normal components of the electric field on a surface element are given by

$$E_{ne} = \frac{\sigma_{tot}}{2\varepsilon_0} + E_n, \quad E_{ni} = -\frac{\sigma_{tot}}{2\varepsilon_0} + E_n, \quad (8)$$

where E_n is the normal component of the total field, which is the superposition of the external field and the field due to interface charges.

Substituting (8) into (4) yields

$$\sigma_{tot} = \frac{2\sigma}{\varepsilon_i + \varepsilon_e} + \frac{2\varepsilon_0(\varepsilon_i - \varepsilon_e)}{\varepsilon_i + \varepsilon_e}E_n. \quad (9)$$

Calculating the gradient of φ (see (7)) and substituting $E_n = -(\mathbf{n}\cdot\nabla\varphi)$ into (9), we arrive at the integral equation for the total charge density

$$\sigma_{tot} = \frac{2\sigma}{\varepsilon_i + \varepsilon_e} + \frac{\varepsilon_i - \varepsilon_e}{\varepsilon_i + \varepsilon_e}\left(2\varepsilon_0 E_n^0 + \frac{1}{2\pi}\oint_S \frac{\sigma_{tot}(\mathbf{n}\cdot\nabla r_{NM})}{r_{NM}^2}dS\right), \quad (10)$$

where $r_{NM} = |\mathbf{r}_N - \mathbf{r}_M|$, $\mathbf{r}_N$ and $\mathbf{r}_M$ are the point of observation and a running point, respectively, on the surface, and $E_n^0 = -(\mathbf{n}\cdot\nabla\varphi_0)$.

In electrostatic problems for insulators, the free charge density σ is considered given. In this case, expression (5) gives an additional relationship between the charge densities. Substituting E_n from (9) into (8), we get

$$E_{ne} = \frac{\varepsilon_i\sigma_{tot} - \sigma}{\varepsilon_0(\varepsilon_i - \varepsilon_e)}, \quad E_{ni} = \frac{\varepsilon_e\sigma_{tot} - \sigma}{\varepsilon_0(\varepsilon_i - \varepsilon_e)}. \quad (11)$$

The substitution of (11) into (5) yields a relation between the total density and the free charge density:

$$\sigma_{tot} = \sigma\frac{\lambda_e - \lambda_i}{\lambda_e\varepsilon_i - \lambda_i\varepsilon_e}. \quad (12)$$

In view of (12), the normal components of the strength and the physical charge density are related as

$$E_{ne} = \frac{\lambda_i\sigma}{\varepsilon_0(\lambda_i\varepsilon_e - \lambda_e\varepsilon_i)}, \quad E_{ni} = \frac{\lambda_e\sigma}{\varepsilon_0(\lambda_i\varepsilon_e - \lambda_e\varepsilon_i)}. \quad (13)$$

Substituting (12) into (10) yields the integral equation for the free charge density

$$\sigma(\mathbf{r}_N) = \frac{\lambda_i\varepsilon_e - \lambda_e\varepsilon_i}{\lambda_i + \lambda_e}2\varepsilon_0 E_n^0(\mathbf{r}_N) + \frac{\lambda_i - \lambda_e}{\lambda_i + \lambda_e}\frac{1}{2\pi}\oint_S \frac{\sigma(\mathbf{r}_M)(\mathbf{n}\cdot\nabla r_{NM})}{r_{NM}^2}dS. \quad (14)$$

From (12) and (14), we find that, at $\lambda_i = \lambda_e$, the total density equals zero, with $\sigma = -\sigma_p = (1 - \varepsilon)\varepsilon_0 E_n^0$. If $\lambda_e\varepsilon_i = \lambda_i\varepsilon_e$, the physical density vanishes and the charge distribution becomes of purely polarization character. Such a construction of the solution meets conditions (3)–(5). Thus, solution (14) in view of relation (12) specifies the field potential in the problem stated. Physically, this solution means that charge transfer in the presence of an interface induces both free and polarization charges at it.

The total charge Q on the body is given by the integral

$$Q = \oint_S \sigma dS. \quad (15)$$

A force acting on a rigid body is defined by the electric tension tensor on the outer side of the surface of the body [6, 7]. Then, the force per unit surface area is [7]

$$\mathcal{F} = D_{ne}\mathbf{E}_e - \frac{D_e E_e \mathbf{n}}{2}. \quad (16)$$

The resultant electric force acting on the body in the direction d is given by the integral

$$F_d = \oint_S \mathcal{F}\mathbf{d}dS. \quad (17)$$

Consider the problem of charge transfer for a sphere of radius r_0 that is in contact with the electrode surface in the cylindrical coordinates r, θ, z (Fig. 1a). The elec-

tric field $\mathbf{E}_0$ is aligned with the z axis: $\varphi_0 = -E_0 z$. The plane of the electrode, $z = 0$, is assumed to be equipotential:

$$z = 0 : \varphi = 0.$$

To meet this condition, it is necessary to make a substitution in the integral in expression (7):

$$\frac{1}{|\mathbf{r} - \mathbf{r}'|} \longrightarrow \frac{1}{|\mathbf{r} - \mathbf{r}'|} - \frac{1}{|\mathbf{r} - \mathbf{r}''|}$$

$$= \frac{1}{\sqrt{(x-x')^2 + (y-y')^2 + (z-z')^2}}$$

$$- \frac{1}{\sqrt{(x-x')^2 + (y-y')^2 + (z+z')^2}},$$

where $\mathbf{r}''$ is the mirror image of the point $\mathbf{r}'$ on the sphere relative to the plane $z = 0$.

Accordingly, the kernel of integral equations (10) and (14) is modified as follows:

$$\frac{\mathbf{n} \cdot \nabla \tau_{NM}}{r_{NM}^2} \longrightarrow G(\mathbf{r}, \mathbf{r}') = \frac{\mathbf{n} \cdot \nabla r_{NM}}{r_{NM}^2} - \frac{\mathbf{n} \cdot \nabla r_{NM'}}{r_{NM'}^2}. \quad (18)$$

Next, consider the problem in the dimensionless variables

$$\bar{r} = \frac{r}{r_0}, \quad \bar{z} = \frac{z}{r_0}, \quad \bar{\varphi} = \frac{\varphi}{E_0 r_0}, \quad \bar{E} = \frac{E}{E_0},$$

$$\bar{\sigma} = \frac{\sigma}{\varepsilon_e \varepsilon_0 E_0}, \quad \bar{Q} = \frac{Q}{\varepsilon_e \varepsilon_0 E_0 r_0^2},$$

$$\bar{\mathscr{F}} = \frac{\mathscr{F}}{\varepsilon_e \varepsilon_0 E_0^2}, \quad \bar{F} = \frac{F}{\varepsilon_e \varepsilon_0 E_0^2 r_0^2}, \quad \lambda = \frac{\lambda_e}{\lambda_i}, \quad \varepsilon = \frac{\varepsilon_i}{\varepsilon_e}.$$

Since the equation for surface, $r = f(z)$, does not involve Θ, the problem is axisymmetric and one can perform integration over the azimuth angle in (14). Eventually, Eq. (14) will take the dimensional form

$$\sigma(z) = \frac{1 - \lambda\varepsilon}{\lambda + 1} 2n_z + \frac{1 - \lambda}{2\pi(\lambda + 1)} \int_0^{f_m} \sigma(z') g(z, z') dz', \quad (19)$$

where

$$g(z, z') = \int_0^{2\pi} G(\mathbf{r}, \mathbf{r}') d\Theta'.$$

Hereafter, the bar over the dimensionless quantities is omitted, n_z is the z component of the vector $\mathbf{n} = (n_r, n_z) = (1, -f')/\sqrt{1 + f'^2}$, $f'(z) = df/dz$, and f_m is the maximal dimension of the body along the z axis. Tension

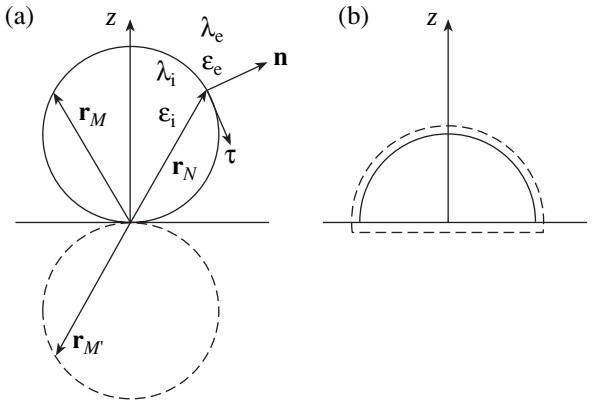

Fig. 1. (a) Unit-radius sphere $f(z) = \sqrt{1 - (z-1)^2}$, $0 \le z \le 2$. The point $\mathbf{r}_{M'}$ of the sphere is the mirror reflection of the point $\mathbf{r}_M$ relative to the plane $z = 0$. (b) Surface $f(z) = \sqrt{1 - z^2}$, $0 \le z \le 1$. The dashed line depicts the contour of integration on the hemisphere. On the base of the contour, the field vanishes.

(16) is given by

$$\mathscr{F}_z = D_{ne} E_{ze} - \frac{D_{ne} E_{ne} + D_\tau E_\tau}{2} n_z$$

$$= \left(\frac{\sigma^2}{2(1 - \lambda\varepsilon)^2} - \frac{E_\tau^2}{2} \right) n_z + \frac{\sigma E_\tau}{1 - \lambda\varepsilon} \tau_z. \quad (20)$$

Here, we used expression (13) and the relationship $E_z = E_n n_z + E_\tau \tau_z$, where the magnitude of the vector tangent to the surface is $\tau = (\tau_r, \tau_z) = (f', 1)/\sqrt{1 + f'^2}$. According to (7) and (13), the field tangent to the sphere is given by

$$E_\tau = \tau_z + \frac{\lambda - 1}{4\pi(\lambda\varepsilon - 1)} \oint_S \sigma \left(\frac{(\tau \nabla r_{NM})}{r_{NM}^2} - \frac{(\tau \nabla r_{NM'})}{r_{NM'}^2} \right) dS.$$

For the spherical surface, expressions (15) and (17) for the total free charge and the z component of the resultant force, respectively, take the form

$$Q = 2\pi \int_0^{f_m} \sigma(z) dz, \quad (21)$$

$$F_z = 2\pi \int_0^{f_m} \mathscr{F}_z dz. \quad (22)$$

SOLUTION OF THE PROBLEM

The known solution [8] for a perfectly conducting bead near a plane gives the following expressions for

the charge density, total charge, and resultant force:

$$\sigma(z) = \frac{1}{(2z)^{3/2}}\int_0^\infty \frac{\nu^2 J_0(\nu\sqrt{(2-z)/4z})}{\sinh(\nu/2)}d\nu, \qquad (23)$$

$$Q = 4\pi\zeta(2) \simeq 20.68,$$
$$F_z = 4\pi\left(\zeta(3) + \frac{1}{6}\right) \simeq 17.20, \qquad (24)$$

where J_0 is the Bessel function and $\zeta(t)$ is the Riemann zeta function.

The positiveness of Q and F_z implies that the charges of the bead and electrode have like signs and the force is aligned with the positive z direction (repulsion from the electrode).

Consider this problem for a hemisphere of radius $r = 1$ resting on a planar electrode in a polarizable medium (Fig. 1b). The solution to problem (3)–(6) has the form

$$\varphi_i = -\frac{3\lambda}{2\lambda+1}z, \quad \varphi_e = -z + \frac{1-\lambda}{2\lambda+1}\frac{z}{(r^2+z^2)^{3/2}},$$
$$\sigma = 3z\frac{1-\lambda\varepsilon}{2\lambda+1}. \qquad (25)$$

This solution coincides with that for a closed sphere placed in a uniform field and is of interest for two reasons. First, the potentials and fields (and, hence, force (20)) are independent of the permittivity. This is no surprise, since relationship (4) allows the determination of σ alone, as follows from the statement of the problem. Second, with $\lambda\varepsilon > 1$, the inverse charge distribution over the spherical surface is possible in a polarizable medium, unlike the similar electrostatic problem for a perfect conductor or insulator. This means that, when the body is detached from the electrode, the signs of the objects may differ. The force and charge on the semisphere are

$$F_z = \frac{9\pi(1+2\lambda-\lambda^2)}{4(2\lambda+1)^2}, \qquad (26)$$

$$Q = \pi\frac{1-\lambda\varepsilon}{2\lambda+1}. \qquad (27)$$

When calculating the force, we took into account that the integral over the base of the hemisphere equals zero, since the field vanishes outside this domain. From (26), it follows that, with $\lambda > 1 + \sqrt{2}$, repulsion from the electrode changes to attraction. With $\lambda \longrightarrow 0$, $F_z = 9\pi/4$; with $\lambda \longrightarrow \infty$, $F_z = -9\pi/16$. The charge changes sign at $\lambda\varepsilon = 1$. Comparing the positions of the points of sign reversal in (26) and (27), we find that the semisphere may separate out of the oppositely charged electrode if $1/\varepsilon < \lambda < 1 + \sqrt{2}$. Dependence (26) is plotted in Fig. 2.

Integral equation (19) for the sphere (Fig. 1a) was solved numerically by reducing it to a set of linear algebraic equations. Since the initial equation has a weak singularity at the point $z = z'$ [9, 10], we numerically solved its equivalent form

$$\sigma(z) = \frac{1-\lambda\varepsilon}{\lambda+1}2n_z(z) + \sigma(z)\frac{1-\lambda}{2\pi(\lambda+1)}\int_0^{f_m} g(z,z')t\,dz'$$
$$+ \frac{1-\lambda}{2\pi(\lambda+1)}\int_0^{f_m} g(z,z')(\sigma(z')-\sigma(z))dz'. \qquad (28)$$

It is assumed that, after such a modification, the integrand in the second integral becomes continuous and can be interpolated. In (28), the first integral of the discontinuous function is taken by the formula of numerical integration on a nonuniform grid whose nodes are defined by the zeros of the Chebyshev polynomial [11]. The second integral is approximated by the Euler quadrature formula on a uniform grid

$$\int_a^b y\,dz = h\left(\frac{y_0}{2} + y_z + \ldots + y_{m-1} + \frac{y_m}{2}\right)$$
$$+ \frac{h^2}{12}(y_0' - y_m') - \frac{h^4}{720}(y_0''' - y_m'''),$$
$$h = \frac{b-a}{m}, \quad y_0 = y(a), \quad y_m = y(b)$$

with a residual $\sim h^6$ [11]. Such a computing scheme yields at least four valid decimal signs in analytical results obtained with (23)–(27) on a grid with $m = 100$ nodes.

Figure 2 plots the force F_z vs. relative conductivity λ (see (22)). As for the case of the hemisphere, the force

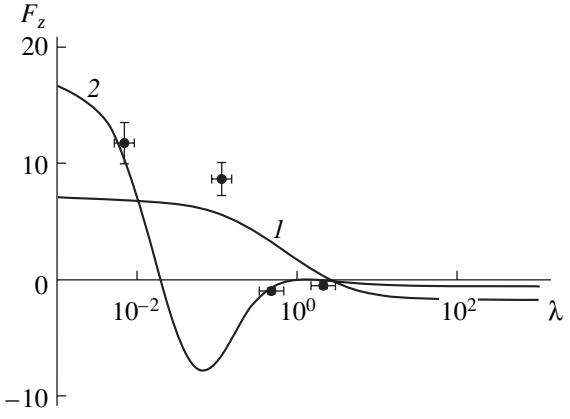

Fig. 2. Electric force vs. relative conductivity: (*1*) hemisphere resting on the electrode and (*2*) sphere.

is independent of ε, and there is a point (λ = 0.0196 for the sphere) where F_z changes sign. The λ dependence of the total charge on the sphere (relationship (21)) is given in Fig. 3. For ε lying between 0.05 and 100, the associated curves (see, e.g., the curve for ε = 20) fall into the domain bounded by the curves for ε = 0.05 and 100. As ε increases, so does the range of λ where the sphere and the electrode have opposite signs. The absolute value of the negative charge also grows. The curves $F_z(\lambda)$ and $Q(\lambda)$ touch the horizontal axis at the point λ = 1. The curve $Q(\lambda)$ crosses the horizontal axis at the points λ = 1/ε.

In the case of a perfectly conducting sphere, λ = 0 and the results of numerical solution coincide with (23) and (24). The other electrostatic limit, a perfect uncharged insulator, is not a specific case of the problem stated. Indeed, the limit $\lambda \longrightarrow \infty$ corresponds to the situation when an insulating sphere is placed in a conducting environment; however, a free charge is induced at the interface in this case. Formally, in a body–medium system, a polarization charge alone is induced only if λε = 1 with $\sigma_{tot} = \sigma_p$. For an insulating sphere at a planar electrode, relationships (10) and (16) in the dimensional quantities have the form

$$\sigma_p(z) = \frac{\varepsilon - 1}{\varepsilon + 1}\left(2n_z(z) + \int_0^{f_m} dz' g(z, z')\sigma_p(z')\right),$$

$$\mathcal{F}_z = \left(\frac{\varepsilon^2\sigma_p^2}{(\varepsilon - 1)^2} - \frac{E_\tau^2}{2}\right)n_z + \frac{\varepsilon\sigma_p E_\tau}{\varepsilon - 1}\tau_z.$$

Here, unlike the problem of charge transfer, ε_e does not enter into the normalization of polarization density: $\bar{\sigma}_p = \sigma_p/\varepsilon_0 E_0$. The ε dependence of the force acting on an insulating sphere is demonstrated in Fig. 4. Since the electrostatic problem for uncharged insulators and the problem of charge transfer in conducting media are stated identically, the solution to the former must coincide with that to the latter if the formal substitution $\varepsilon \longrightarrow 1/\lambda$ is made [4]. This equivalence refers to only the potential and field distributions. Since the force is calculated by using the electric tension tensor outside the body, the data for the force in Figs. 3 (curve *2*) and 4 also coincide upon the substitution $\varepsilon \longrightarrow 1/\lambda$. At the same time, the analogy breaks down for charge densities and forces acting on a deformable body, since the related tension tensors experience a jump at the interface (e.g., in liquids).

EXPERIMENTAL RESULTS AND DISCUSSION

The experiments were performed with solid particles in air, as well as with bubbles and drops in a liquid. The field in the gap between two plane-parallel electrodes spaced at 1 or 2 cm was produced by a dc voltage source. The particle diameter was no more than 1 mm.

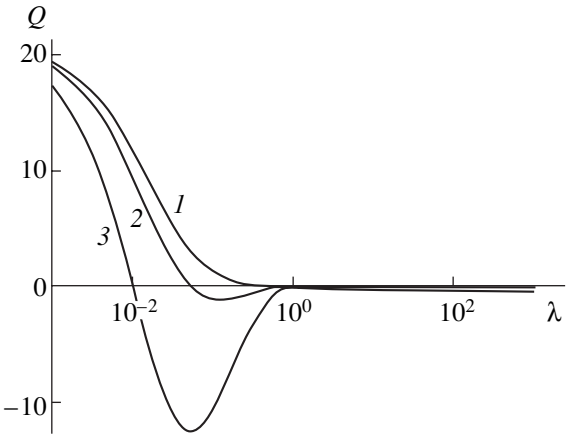

Fig. 3. Total charge on the sphere vs. relative conductivity. ε = (*1*) 0.05, (*2*) 20, and (*3*) 100.

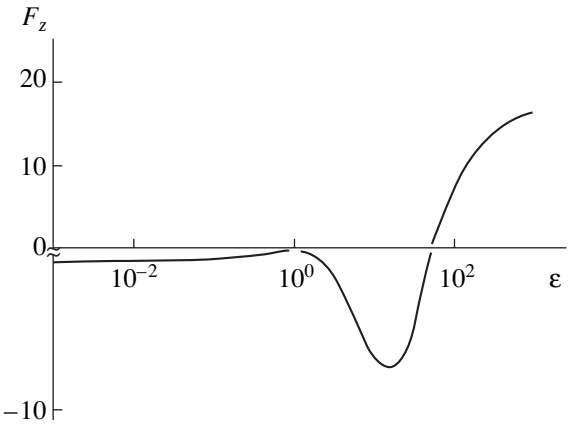

Fig. 4. Electric force vs. relative permittivity. The body and environment are perfect insulators. Negative forces are scaled up twofold. The curve crosses and touches the horizontal axis at ε ≈ 51 and 1, respectively.

In the experiments in air, the test particles were placed on the lower electrode. As the voltage increased, they acquired a charge. When separated out from the electrode, the particle reversed the charge on the upper electrode and returned to the lower electrode. Thus, the particles executed oscillatory motion between the electrodes.

A particle leaves the lower electrode when the gravity force equals the resultant electric force:

$$\varepsilon_0 E_0^2 r_0^2 F_z = \frac{4}{3}\pi r_0^3 \gamma g,$$

where E_0 is the separation field, γ is the density of the particle material, and g is the free-fall acceleration.

Having measured E_0, r_0, and γ, one can find from this equality the experimental value of the dimensional

Table 1. Properties of solids used as spherical particles

Substance	Density, g/cm^3	Conductivity, $(\Omega\ \mathrm{m})^{-1}$	Radius, mm	F_z^{exp}	F_z
Polystyrene	1.05	10^{-16}–10^{-15}	0.5	9.93	–6.95–16.6
Molybdenum glass	2.5	10^{-14}	0.5	18.9	16.6–17.1
Graphite	2.25	2.5×10^4	0.25	17.5	17.2
NaCl	2.19	10^{-11}–10^{-10}	0.5	16.2	17.2
Sugar	1.59	–	0.25	11.8	–

electric force:

$$F_z^{\mathrm{exp}} = \frac{4\pi r_0 \gamma g}{3\varepsilon_0 E_0^2}. \qquad (29)$$

The properties of the particles used in the experiments are listed in Table 1. The density and the conductivity were taken from references books [12, 13]. Separation from the planar electrode was observed at $E_0 = 5 \times 10^5$, 5.6×10^5, and 3.5×10^5 V/m for polystyrene granules, glass beads, and graphite spheres, respectively. For NaCl and sugar granules, $E_0 = 4 \times 10^5$ V/m. As E_0, we took the field strength averaged over ten particles of each of the substances. The maximal discrepancy between the measured and calculated values of E_0 was 5%. The total error of F_z^{exp} calculated by (29) was 10%.

According to [12, 13], the conductivity of air under normal conditions is on the order of 10^{-18}–10^{-17} $(\Omega\ \mathrm{m})^{-1}$. Then, for polystyrene, $10^{-3} < \lambda < 10^{-1}$. Hence, the experimental value of the force F_z^{exp} evaluated by (22) varies between –6.95 and 16.6. The analytical values of F_z are listed in the last column of Table 1. The experimental and analytical values coincide up to the experimental error.

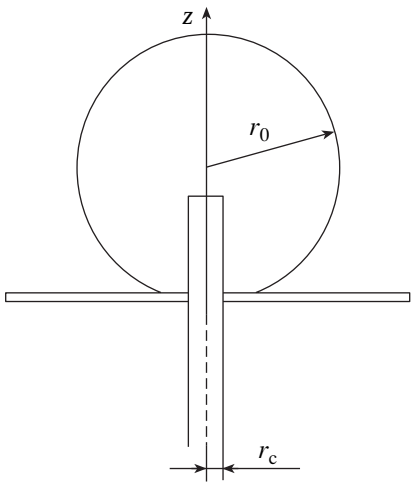

Fig. 5. z axis is directed toward the target electrode. At the instant of separation, the drop touches the electrode.

The separation of the particles was observed on polydisperse powders of polystyrene, polyvinyl chloride, corundum, and coal. The mean size of the particles was 10–50 μm, and their conductivity covered the entire range of interest: 10^{-16}–10^4 $(\Omega\ \mathrm{m})^{-1}$. For each of the materials, single-particle separation was observed, with the particles having a near-spherical shape. However, in each of the powders applied on the lower electrode, there were particles that remained on the plane up to the breakdown field strength.

The electric force acting on a drop placed in an immiscible liquid was determined in a rectangular cell. A constant voltage was applied to the 4×4-cm metallic bases. The side glass walls were 1.2 cm high. The electric field causes convective motion in the closed space of the cell, and the drop can be transferred by its environment. The higher the field, the more intense the motion. Therefore, the electric force was calculated for the case when the drop was separated from the capillary under a low (no more than 1 kV) potential difference. For a sinking (floating-up) drop, the capillary was placed at the center of the upper (lower) electrode. The length of the part of the capillary advanced in the electrode gap was such that, at the instant of separation, the drop just touched the electrode rather than being spread over it (Fig. 5). At the instant of separation, the gravity force, buoyant force, electric force, and surface tension force must be counterbalanced:

$$\frac{4\pi}{3}(\gamma_i - \gamma_e) g r_0^3 = \varepsilon_0 \varepsilon_e E_0^2 r_0^2 F_z + 2\pi r_c T, \qquad (30)$$

where r_c is the radius of the capillary and T is the interfacial tension.

It is assumed here that the drop separates out along a neck of radius equal to the outer radius of the capillary. The surface tension was calculated by (30) for $E_0 = 0$ using the measured radius of the drop being detached. If the radius of the drop separated increases in the field, the force attracts to the electrode: $F_z < 0$. Otherwise, we are dealing with the repulsive force; $F_z > 0$. The field E_0 was varied from 7×10^4 to 10×10^4 V/m. The outer radius of the capillary was 0.11 and 0.23 mm. The radius of separating drops was found to be 0.6–0.7 mm.

Table 2. Properties of drop–environment liquid systems

System	γ_i, g/cm^3	λ_i, $(\Omega\ \mathrm{m})^{-1}$	λ	ε	T, dyn/cm	E_z^{exp}
SO + SR	1.1	1.43×10^{-12}	0.47	1.95	1.54	−0.95
SpO + SR	0.87	6×10^{-14}	0.11	1.95	1.07	8.65
CO + SO	0.97	2.1×10^{-10}	6.8×10^{-3}	2.69	1.21	11.8
SR + SO	0.98	6.7×10^{-13}	2.13	2.69	1.54	−0.49

Drop–liquid systems were made up of four mutually immiscible liquids: silicone oil (SO), silicone rubber (SR), spindle oil (SpO), and castor oil (CO). The parameters of several of the systems are listed in Table 2 (the drop is indicated first). The values of F_z were calculated by formula (30) with an accuracy of 15%. The conductivity was measured with the same accuracy (15%); therefore, the error in determining λ was 30%. Data points are shown in Fig. 2. The significant deviation from analytical curve *2* may be due to convection in the drop and environment and also to the field-induced deformation of the drop shape. Another possibility of separation is the emission of fine droplets from the top of the electric-field-deformed initial drop. For example, an SR drop in SO flattens, spreading over the electrode, when the field grows. At a certain field strength, a sharp droplet-emitting protrusion forms on the drop surface. The same mechanism may work in the case of highly conducting coarse drops, which transform into a spheroid extended along the field.

In bubbles ($\lambda \longrightarrow \infty$), both the surface density and the tangent field are always other than zero. From Eq. (19), one can see that $\sigma(z)$ is proportional to ε, while the tangent field is ε independent. Figure 6 shows the curves $\sigma(z)$ (for $\varepsilon = 0.5$) and $E_\tau(z)$. For comparison, charge distribution (23) on a perfectly conducting sphere (curve *3*) is demonstrated. On the bubble surface,

$$E_{\mathrm{ne}} = 0, \quad E_{\mathrm{ni}} = -\frac{\sigma}{\varepsilon}, \quad E_\tau \neq 0.$$

This means that the external lines of force cross the surface along the tangent and are refracted inside according to the relationship between E_{ni} and E_τ. The qualitative picture of lines of force for a bubble is depicted in Fig. 7. With tangent tensions (i.e., liquid circulation inside the bubble) neglected, the resultant electric force ($F_z = -0.655$) attracts to the electrode. In the experiments with the liquids listed in Table 2, bubbles of diameter 0.1–1.0 mm did not separate out even in the case of fully developed convection. Under the same conditions, the emission of fine bubbles also did not occur. It was shown [14] that bubbles can also be held in place at the lower electrode if they are subjected to the nonuniform field of a hole in the electrode. The electric field may both detach and hold the particles at the electrodes. Here, the ratio between the conductivities of the environment and body and the shape of the latter are governing parameters. Our calculation shows that the separation of a bubble in a uniform field may take place only if the conductivity ratio is less than 0.0196. For a hemisphere on the electrode, this value is $1 + \sqrt{2}$. The associated experiments were carried out with organic and inorganic solids and drops, as well as with bubbles. Sometimes, the environments were anisotropic. We tested both poor (polystyrene and polyvinyl chloride) and good conductors.

The experiments with disperse particles of irregular shape demonstrated that, in this case, fine particles can be held irrespective of the conductivity ratio (although coarse particles are detached from the electrode). For particles of regular spherical shape (for the materials

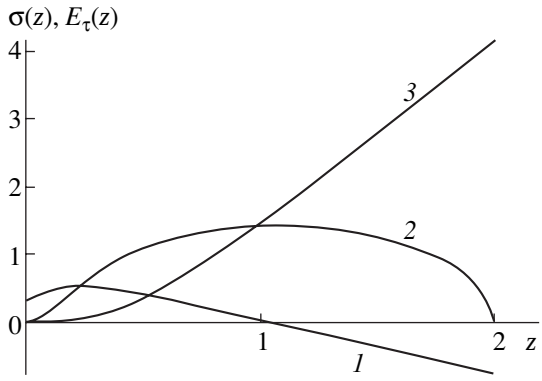

Fig. 6. Charge density and tangent field on the bubble vs. distance for $\varepsilon = 0.5$. (*1*) $\sigma(z)$, (*2*) $E_\tau(z)$, and (*3*) $\sigma(z)$ for a perfectly conducting sphere.

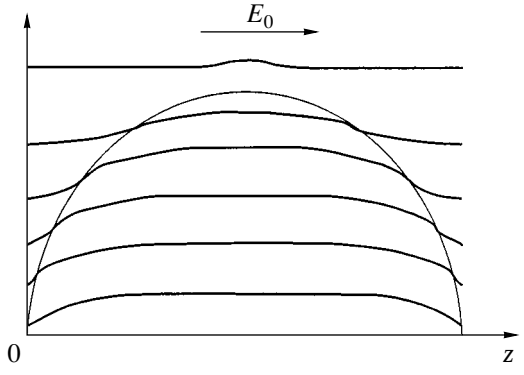

Fig. 7. Distribution of lines of force for the bubble.

cited above), this is impossible, according to the calculation.

Drops placed in a poorer conducting medium tend to separate from the electrode. Conducting bubbles are extremely hard to hold by an electric field. In the intermediate conductivity range, the influence of tangent interfacial tensions becomes significant. In this case, a part of the energy is spent on convection inside the drop and environment and the problem of interaction with the electrode should be stated in terms of electrohydrodynamics.

Current passing in homogeneous media may charge only interfaces. The amount and sign of the free charge are governed by the self-consistent process of charge transfer in the effective field, including the initial field and the induced fields of physical and polarization charges. The signs of the charge on the sphere and electrode coincide if the product of the relative conductivity and permittivity $\lambda\varepsilon < 1$. The reverse inequality reverses the sign on the sphere.

ACKNOWLEDGMENTS

This work was financially supported by the Russian Foundation for Basic Research (project no. 02-01-96035) and the administration of Kaluga oblast.

REFERENCES

1. S. S. Dukhin, V. R. Éstrela-L'opes, and É. K. Zholkovskiĭ, *Electrical Surface Phenomena and Electrical Filtration* (Naukova Dumka, Kiev, 1985).
2. M. N. Livshits and V. M. Moiseev, *Electrical Phenomena in Aerosols and Their Applications* (Énergiya, Moscow, 1965).
3. P. A. F. White and S. E. Smith, *High Efficiency Air Filtration* (Butterworth, London, 1964; Atomizdat, Moscow, 1967).
4. L. D. Landau and E. M. Lifshitz, *Course of Theoretical Physics*, Vol. 8: *Electrodynamics of Continuous Media* (Nauka, Moscow, 1982; Pergamon, New York, 1984).
5. G. A. Grinberg, *Selected Issues of the Mathematical Theory of Electric and Magnetic Phenomena* (USSR Academy of Sciences, Moscow, 1948).
6. N. N. Mirolyubov, M. V. Kostenko, M. L. Levinshteĭn, and N. N. Tikhodeev, *Calculation of Electrostatic Fields* (Vysshaya Shkola, Moscow, 1963).
7. I. E. Tamm, *The Principles of Electricity Theory* (Nauka, Moscow, 1976).
8. N. N. Lebedev and I. P. Skal'skaya, Zh. Tekh. Fiz. **32**, 375 (1962) [Sov. Phys. Tech. Phys. **7**, 268 (1962)].
9. P. P. Zabreyko, A. I. Koshelev, M. A. Krasnosel'skiĭ, *et al.*, *Integral Equations: A Reference Text* (Nauka, Moscow, 1968; Noordhoff, Leyden, 1975).
10. L. V. Kantorovich and V. I. Krylov, *Approximate Methods of Higher Analysis* (Fizmatgiz, Moscow, 1962; Wiley, New York, 1964).
11. I. S. Berezin and N. P. Zhidkov, *Computing Methods* (Nauka, Moscow, 1966; Pergamon, Oxford, 1965), Vol. 1.
12. *Handbook of Physical Quantities*, Ed. by I. S. Grigoriev and E. Z. Meilikhov (Énergoatomizdat, Moscow, 1991; CRC, Boca Raton, 1997).
13. *Tables of Physical Quantities*, Ed. by I. K. Kikoin (Atomizdat, Moscow, 1976).
14. Yu. K. Bratukhin and V. A. Semenov, Zh. Éksp. Teor. Fiz. **83**, 2170 (1982) [Sov. Phys. JETP **56**, 1257 (1982)].

Translated by V. Isaakyan

Technical Physics, Vol. 48, No. 6, 2003, pp. 677–681. Translated from Zhurnal Tekhnicheskoĭ Fiziki, Vol. 73, No. 6, 2003, pp. 17–20.

ATOMS, SPECTRA, RADIATION

Electron Transitions and the Spectrum of Radiation Emitted by an Atom Interacting with an Ultrashort Electromagnetic Pulse

V. I. Matveev

Lomonosov Pomor State University, Arkhangel'sk, 163006 Russia
e-mail: matveev.victor@pomorsu.ru

Received May 20, 2002

Abstract—Electron transitions and the spectrum of radiation emitted by an atom interacting with an ultrashort electromagnetic pulse are studied based on the sudden perturbation approximation. The excitation and ionization probabilities, the spectra of pulse reradiation by the atom, and the reradiation cross sections are calculated. It is concluded that the reradiation of ultrashort electromagnetic pulses by multielectron atoms is of a coherent nature. © *2003 MAIK "Nauka/Interperiodica".*

INTRODUCTION

The recent work [1], in which the possibility of generating electromagnetic pulses of duration $\tau = 10^{-22}$–10^{-21} s was reported, may be a great step forward in investigating the interaction of ultrashort electromagnetic pulses with a substance. In particular, there appears the feasibility of studying processes attendant to the interaction of atoms with high-intensity ultrashort electromagnetic pulses. It should be noted that a nonperturbative approach to the interaction of atoms with intense electromagnetic pulses of duration exceeding the characteristic atomic time scale is very tedious and requires the use of numerical solution methods. Work [2], in which the excitation and ionization of helium atoms by intense electromagnetic pulses as short as 3.8–15.2 fs were considered, is an example of this (see also [3–6] and the references cited there). In our case, the characteristic atomic time scale is $\tau_a \sim 10^{-17}$ s; i.e., it far exceeds the width of ultrashort pulses. Therefore, the sudden perturbation approximation [7], which poses no limitation on the perturbation amplitude and only requires the inequality $\tau/\tau_a \ll 1$ to be met, seems to be a natural basis for the solution. One can cite many examples where a random excitation causes the excitation or ionization of an atom. This is first of all true for nuclear reactions [8, 9], e.g., for the β decay of a nucleus, where atom electrons perceive the escape of a fast β electron as a sudden change in the charge state of the nucleus. Another example is neutron–nucleus collision, where the nucleus suddenly gains the momentum, etc. The sudden perturbation approximation is applied in the consideration of multielectron transitions in complex atoms, in which transitions between inner shells are perceived as instantaneous events by rather slow outer-shell electrons (see, e.g., [10]). Sudden perturbations result in inelastic processes upon collisions between fast multiply charged ions and atoms [11–13] and between charged particles and highly excited atoms [14]. In many cases of practical importance, perturbations are not small and the perturbation theory fails. However, situations often happen where the perturbation duration is much less than the characteristic atomic time scale [13, 15–17]. In this case, one can solve the problem without limiting the perturbation amplitude and perform analytical calculations.

In this work, based on the sudden perturbation approximation, we analyze the excitation and ionization of an atom when it interacts with an ultrashort electromagnetic pulse. In particular, the excitation and ionization probabilities are determined, and the spectra and cross section of the reradiation of the pulse by the atom are calculated.

EXCITATION AND IONIZATION OF AN ATOM

The interaction of atom electrons with a Gaussian electromagnetic pulse

$$\mathbf{E}(t) = \mathbf{E}_0 \exp(-\alpha^2 t^2)\cos(\omega_0 t) \tag{1}$$

can be expressed in the form

$$V(t) \equiv V(\mathbf{r}_a, t) = \mathbf{E}(t)\sum_{a=1}^{N} \mathbf{r}_a, \tag{2}$$

where $\mathbf{r}_a$ and N are the coordinates and number of atom electrons, respectively ($a = 1, \ldots, N$). The potential $V(t)$ is other than zero only within a time interval $\tau \sim \alpha^{-1}$ that is much shorter than the characteristic period of an undisturbed atom described by the Hamiltonian H_0.

1063-7842/03/4806-0677$24.00

Then, when solving the exact Schrödinger equation (from here on atomic units are used)

$$i\frac{\partial\psi}{\partial t} = (H_0 + V(t))\psi, \tag{3}$$

one can ignore the evolution of the ψ function subjected to the undisturbed Hamiltonian H_0 for the time period τ. Therefore, the amplitude of atom transition from the initial state φ_0 to any final state φ_n under the action of a sudden perturbation $V(t)$ will have the form

$$a_{0n} = \left\langle \varphi_n \middle| \exp\left(-i\int_{-\infty}^{+\infty} V(t)dt\right) \middle| \varphi_0 \right\rangle, \tag{4}$$

where φ_0 and φ_n belong to the complete orthonormal set of eigenfunctions of the undisturbed Hamiltonian H_0; that is, $H_{0\varphi_n} = \epsilon_n\varphi_n$.

From the above formulas, one can calculate the probabilities $w_{0n} = |a_{0n}|^2$ of atom excitation and ionization. With a perturbation taken in form (2), w_{0n} can be expressed through the well-known [9, 18] inelastic atomic form factors:

$$w_{0,n} = \left| \left\langle \varphi_n \middle| \exp\left(-\left(i\mathbf{q}\cdot\sum_a \mathbf{r}_a\right)\right) \middle| \varphi_0 \right\rangle \right|^2, \tag{5}$$

where

$$\mathbf{q} = \int_{-\infty}^{+\infty} dt\mathbf{E}(t) = \frac{\sqrt{\pi}}{\alpha}\mathbf{E}_0\exp\left(-\frac{\omega_0^2}{4\alpha^2}\right). \tag{6}$$

For example, for a hydrogen atom, the probability of transitions from the 1*s* state to all states with the principal quantum number *n* takes the form

$$w_{0,n} = 2^8 q^2 n^7\left[\frac{n^2-1}{3} + (qn)^2\right] \times \frac{[(n-1)^2 + (qn)^2]^{n-3}}{[(n+1)^2 + (qn)^2]^{n+3}}. \tag{7}$$

For ionization from the 1*s* state,

$$w_{0,k} = \frac{2^5}{\pi k} \times \frac{\left[\left(\mathbf{q}\cdot\frac{\mathbf{k}}{k}\right)^2 + (q^2 - \mathbf{q}\cdot\mathbf{k})^2\right]\exp\left\{-\frac{2}{k}\arctan\frac{2k}{q^2+1-k^2}\right\}}{[(\mathbf{q}-\mathbf{k})^2+1]^4[(q^2+1-k^2)^2 + (2k)^2]}, \tag{8}$$

where $\mathbf{k}$ is the electron momentum in the continuum of hydrogen atoms.

For transitions in a hydrogen-like atom with a nuclear charge Z_a, one has to make substitutions $\mathbf{r} \longrightarrow \mathbf{r}Z_a$, $\mathbf{q} \longrightarrow \mathbf{q}/Z_a$, and $\mathbf{k} \longrightarrow \mathbf{k}/Z_a$. Figure 1 shows the results of calculation for several lower levels of the hydrogen atom and also the probability of complete ionization. Figures 1 and 2 illustrate the unitarity of the approach used ($\Sigma_n w_{0,n} = 1$, where Σ_n means summation over all possible final states of the atom).

Consider transitions in complex atoms by an example of a helium atom. According to [19, 20], arbitrary two-electron states $|n_1, n_2\rangle$ of a helium-like atom are designated by two sets of quantum numbers, n_1 and n_2. Then, in view of (5), the probability of transition from the ground state $|0, 0\rangle$ to a state $|n_1, n_2\rangle$ has the form

$$w_{0,0;\,n_1,n_2} = \left|\langle n_1, n_2|\exp\{-iq(r_1 + r_2)\}|0, 0\rangle\right|^2. \tag{9}$$

Thus, a multielectron transition is the result of the direct action [21, 22] of a high external field. Consider the single and double ionizations of a helium atom. The total probability of double ionization is found by summing (9) over all n_1 and n_2 from a two-electron continuum. The states in this continuum will be denoted by $|\mathbf{k}_1, \mathbf{k}_2\rangle$, where $\mathbf{k}_1$ and $\mathbf{k}_2$ are the electron momenta in the continuum. Then, the total probability is given by

$$w^{2+} = \int d^3\mathbf{k}_1 \cdot d^3\mathbf{k}_2 \times \left|\langle \mathbf{k}_1, \mathbf{k}_2|\exp\{-i\mathbf{q}\cdot(\mathbf{r}_1 + \mathbf{r}_2)\}|0, 0\rangle\right|^2. \tag{10}$$

In a helium atom, a singly ionized state is caused by those transitions where one of the electrons falls into the continuum and the other, into any state from the discrete spectrum or in any state from the total (discrete–continuous) spectrum. In the latter case, the contribution due to the presence of both electrons in states of the two-electron continuum, i.e., in doubly ionized states, must be subtracted. Therefore, the probability of single ionization is

$$w^{1+} = \int d^3\mathbf{k}_1 \cdot \sum_{n_2} |\langle \mathbf{k}_1, n_2| \times \exp\{-i\mathbf{q}\cdot(\mathbf{r}_1 + \mathbf{r}_2)\}|0, 0\rangle|^2 - w^{2+}, \tag{11}$$

where Σ_{n_2} means summation over all possible values of n_2 for the total (discrete–continuous) spectrum.

Figure 2 shows the probabilities of single and double ionizations for a helium atom. In the calculations, we, following [19, 20], described two-electron states of a helium-like atom in the form of the symmetrized products of hydrogen-like single-electron wave functions. To avoid the orthogonalization procedure (which is usually ill-defined, since, strictly speaking, all states

of the continuous and discrete spectra should be orthogonalized), we chose single-electron hydrogen-like wave functions in the field of a nucleus with the same effective charge: $Z_1 = 1.37$ for single-electron transitions and $Z_2 = 1.97$ for two-electron transitions.

RERADIATION OF A SHORT PULSE BY AN ATOM

To calculate the reradiation cross section for a suddenly arisen electromagnetic pulse, one should proceed as follows. In the sudden perturbation approximation, the evolution of the initial state has the form

$$\Psi_0(t) = \exp\left(-i\int_{-\infty}^{t} V(t')dt'\right)\varphi_0, \tag{12}$$

where $\Psi_0(t)$ satisfies the equation

$$i\frac{\partial\Psi_0(t)}{\partial t} = V(t)\Psi_0(t) \tag{13}$$

with $\Psi_0(t) \longrightarrow \varphi_0$ at $t \longrightarrow -\infty$.

Let us introduce a complete and orthonormal set of functions

$$\Phi_n(t) = \exp\left(i\int_{t}^{+\infty} V(t')dt'\right)\varphi_n, \tag{14}$$

which satisfy Eq. (13) with $\Phi_n(t) \longrightarrow \varphi_n$ for $t \longrightarrow +\infty$. Clearly, amplitude (4) can be recast as

$$a_{0n} = \langle\Phi_n(t)|\Psi_0(t)\rangle. \tag{15}$$

Therefore, the amplitude of photon emission will be calculated in the first order of the perturbation theory as a correction to states (12) and (14) due to the interaction of atom electrons with the electromagnetic field [23]:[1]

$$U = -\sum_{a,k,\sigma}\left(\frac{2\pi}{\omega}\right)^{\frac{1}{2}}\mathbf{u}_{\mathbf{k}\sigma}\cdot(a^{+}_{k\sigma}e^{-i\mathbf{k}\cdot\mathbf{r}_a} + a_{k\sigma}e^{-i\mathbf{k}\cdot\mathbf{r}_a})\hat{\mathbf{p}}_a, \tag{16}$$

where $a^{+}_{\mathbf{k}\sigma}$ and $a_{\mathbf{k}\sigma}$ are the production and loss operators for a photon with a frequency ω, momentum $\mathbf{k}$, and polarization σ ($\sigma = 1$ or 2); $\mathbf{u}_{k\sigma}$ are the unit vectors of polarization; $\mathbf{r}_a$ are the coordinates of atom electrons ($a = 1, \ldots, N$); and $\hat{\mathbf{p}}_a$ are the momentum operators of atom electrons.

In the dipole approximation, the amplitude of photon emission when the atom undergoes the transition from the state φ_0 to a state φ_n then has the form

$$b_{0n}(\omega) = i\left(\frac{2\pi}{\omega}\right)^{\frac{1}{2}}\mathbf{u}_{k\sigma}\cdot\int_{-\infty}^{+\infty}dt e^{i\omega t}\left\langle\Phi_n(t)\left|\sum_a\hat{\mathbf{p}}_a\right|\Psi_0(t)\right\rangle. \tag{17}$$

[1] The sudden unlimited perturbation $V(t)$ is included in the functions $\Phi_n(t)$ and $\Psi_0(t)$.

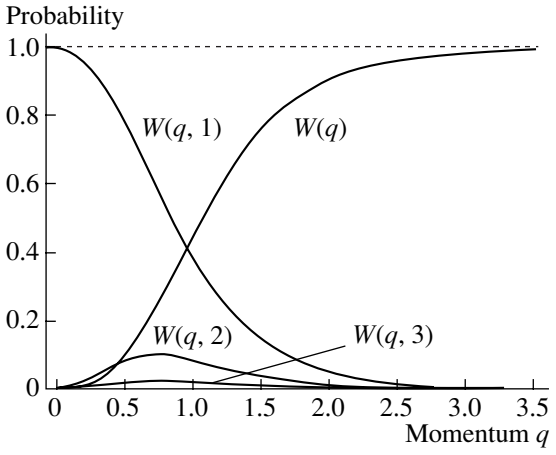

Fig. 1. Results of calculation. $W(q, 1)$ is the probability that an atom will remain in the ground state, $W(q, 2)$ is the excitation probability for states with the principal quantum number $n = 2$, $W(q, 3)$ is the excitation probability for states with the principal quantum number $n = 3$, and $W(q)$ is the probability of complete ionization of an atom.

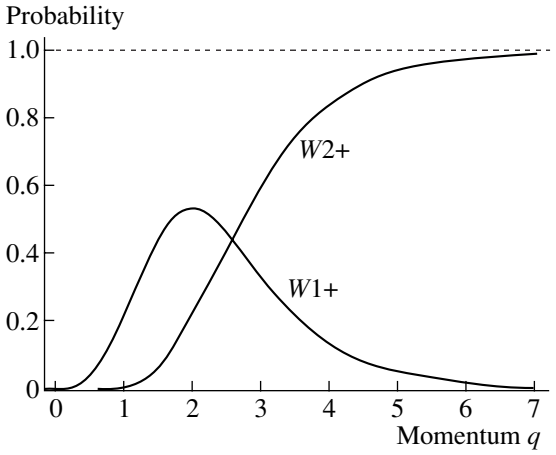

Fig. 2. Probabilities of single ($W1+$) and double ($W2+$) ionization for a helium atom.

Integrating by parts with respect to time and omitting terms that vanish when the interaction with the electromagnetic field ceases (at $t \longrightarrow \pm\infty$) yields

$$b_{0n}(\omega) = -\left(\frac{2\pi}{\omega}\right)^{\frac{1}{2}}\mathbf{u}_{k\sigma}\cdot\int_{-\infty}^{+\infty}dt\frac{e^{i\omega t}}{i\omega} \times\left\langle\varphi_n\left|\sum_a\frac{\partial V(t)}{\partial\mathbf{r}_a}\exp\left(-i\int_{-\infty}^{+\infty}V(t')dt'\right)\right|\varphi_0\right\rangle. \tag{18}$$

The summation of $|b_{0n}(\omega)|^2$ over polarizations, the integration over exit angles, and the summation over all final states φ_n of the atom yield the total radiation spec-

trum

$$\frac{dW}{d\omega} = \frac{2}{3\pi}\frac{1}{c^3\omega}\left\langle \varphi_0 \left| \sum_a \frac{\partial \hat{V}^*(\omega)}{\partial \mathbf{r}_a} \cdot \sum_{a'} \frac{\partial \hat{V}(\omega)}{\partial \mathbf{r}_{a'}} \right| \varphi_0 \right\rangle, \quad (19)$$

where $c = 137$ is the velocity of light and

$$\hat{V}(\omega) = \int_{-\infty}^{+\infty} V(t) e^{i\omega t} dt. \quad (20)$$

Thus, we obtained the complete spectrum of radiation emitted by an atom for a time of action of the sudden perturbation $V(t)$. If $V(t)$ is given by (2), complete spectrum (19) reduces to

$$\frac{dW}{d\omega} = \frac{2}{3\pi}\frac{1}{c^3\omega}\left|\hat{\mathbf{E}}(\omega)\right|^2 N^2, \quad (21)$$

where N is the number of atom electrons and $\hat{\mathbf{E}}(\omega)$ is the Fourier transform of the function $\mathbf{E}(t)$. In view of (20),

$$\hat{\mathbf{E}}(\omega) = \frac{\sqrt{\pi}}{2\alpha} \times \mathbf{E}_0 \left\{ \exp\left[-\frac{(\omega-\omega_0)^2}{4\alpha^2} \right] + \exp\left[-\frac{(\omega+\omega_0)^2}{4\alpha^2} \right] \right\}. \quad (22)$$

Since $\mathbf{E}(t) \neq 0$ only for the time τ and spectrum (21) is proportional to $|\hat{\mathbf{E}}(\omega)|^2$, an atom emits photons that have a continuous spectrum with characteristic frequencies $|\omega - \omega_0| \leq 1/\tau$. Formula (18) also allows one to obtain a photon spectrum when the atom undergoes a transition from the state φ_0 to any state φ_n due to perturbation (2):

$$\frac{dW_{0n}}{d\omega} = \frac{2}{3\pi}\frac{1}{c^3\omega}\left|\hat{\mathbf{E}}(\omega)\right|^2 N^2 \times \left| \left\langle \varphi_n \left| \exp\left(-i\mathbf{q} \cdot \sum_a \mathbf{r}_a \right) \right| \varphi_0 \right\rangle \right|^2. \quad (23)$$

In this case, total spectrum (21) becomes $dW/d\omega = \Sigma_n dW_{0n}/d\omega$, where Σ_n means summation over the complete set of atomic states. As a result, the relative contribution (to total spectrum (21)) of transitions by which the atom is excited into an arbitrary state φ_n takes the form

$$\frac{dW_{0n}/d\omega}{dW/d\omega} = \left| \left\langle \varphi_n \left| \exp\left(-i\mathbf{q} \cdot \sum_a \mathbf{r}_a \right) \right| \varphi_0 \right\rangle \right|^2. \quad (24)$$

As follows from (5), the right of (24) equals $w_{0,n}$; therefore, Figs. 1 and 2 illustrate the respective relative contributions to the total pulse reradiation spectra from those transitions causing the excitation or ionization of hydrogen and helium atoms.

Obviously, to find the pulse reradiation cross sections, it is necessary to multiply spectra (19), (21), and (23) by ω and divide the product by the energy flux I [24]. The energy flux is expressed through the time integral of the Poynting's vector magnitude $S(t) = c(4\pi)^{-1}\mathbf{E}^2$:

$$I = \int_{-\infty}^{+\infty} dt S(t) = \frac{c}{4\pi}\mathbf{E}_0^2 \frac{\sqrt{\pi}}{2\sqrt{2}\alpha}\left\{ \exp\left[-\frac{\omega_0^2}{2\alpha^2} \right] + 1 \right\}. \quad (25)$$

Note an important feature of sudden-perturbation-induced radiation: for multielectron atoms, the radiation intensity, according to (21) and (23), varies as the number of atom electrons squared. This indicates the coherent character of ultrashort pulse reradiation.

ACKNOWLEDGMENTS

This work was financially supported by the Ministry of Education of the Russian Federation (grant no. E00-3.1-390) and the Russian Foundation for Basic Research (grant no. 01-02-17047).

REFERENCES

1. A. E. Kaplan and P. L. Shkolnikov, Phys. Rev. Lett. **88**, 074801 (2002).
2. A. Scrinzi and B. Piraux, Phys. Rev. A **56**, R13 (1997).
3. A. D. Kondorskiy and L. P. Presnyakov, J. Phys. B **34**, L663 (2001).
4. J. B. West, J. Phys. B **34**, R45 (2001).
5. J. Bauer, J. Plucinski, B. Piraux, *et al.*, J. Phys. B **34**, 2245 (2001).
6. G. Lagmago Kamta, T. Grosges, B. Piraux, *et al.*, J. Phys. B **34**, 857 (2001).
7. A. M. Dykhne and G. L. Yudin, Usp. Fiz. Nauk **125**, 377 (1978) [Sov. Phys. Usp. **21**, 549 (1978)].
8. A. B. Migdal, *Qualitative Methods in Quantum Theory* (Nauka, Moscow, 1975; Benjamin, Reading, 1977).
9. L. D. Landau and E. M. Lifshitz, *Course of Theoretical Physics*, Vol. 3: *Quantum Mechanics: Non-Relativistic Theory* (Nauka, Moscow, 1989, 4th ed.; Pergamon, New York, 1977, 3rd ed.).
10. V. I. Matveev and É. S. Parilis, Usp. Fiz. Nauk **138**, 583 (1982) [Sov. Phys. Usp. **25**, 881 (1982)].
11. J. Eichler, Phys. Rev. A **15**, 1856 (1977).
12. G. L. Yudin, Zh. Éksp. Teor. Fiz. **80**, 1026 (1981) [Sov. Phys. JETP **53**, 523 (1981)].
13. V. I. Matveev, Fiz. Élem. Chastits At. Yadra **26**, 780 (1995).
14. I. C. Percival, in *Atoms in Astrophysics*, Ed. by P. G. Burke, W. B. Eissner, D. G. Hammer, and I. C. Percival (Plenum, New York, 1983; Mir, Moscow, 1988), pp. 87–113.
15. A. J. Baltz, Phys. Rev. Lett. **78**, 1231 (1997).

16. A. J. Baltz, Phys. Rev. A **64**, 022718 (2001).
17. V. I. Matveev, Zh. Éksp. Teor. Fiz. **121**, 260 (2002) [JETP **94**, 217 (2002)].
18. A. R. Holt, J. Phys. B **2**, 1209 (1969).
19. V. I. Matveev and Kh. Yu. Rakhimov, Zh. Éksp. Teor. Fiz. **114**, 1646 (1998) [JETP **87**, 891 (1998)].
20. V. I. Matveev, Kh. Yu. Rakhimov, and D. U. Matrasulov, J. Phys. B **32**, 3849 (1999).
21. J. H. McGuire, Adv. At., Mol., Opt. Phys. **29**, 217 (1992).
22. J. H. McGuire, A. Mueller, B. Shuch, *et al.*, Phys. Rev. A **35**, 2479 (1987).
23. V. B. Berestetskii, E. M. Lifshitz, and L. P. Pitaevskii, *Course of Theoretical Physics*, Vol. 4: *Quantum Electrodynamics* (Nauka, Moscow, 1989; Pergamon, New York, 1982).
24. L. D. Landau and E. M. Lifshitz, *Course of Theoretical Physics*, Vol. 2: *The Classical Theory of Fields* (Nauka, Moscow, 1988; Pergamon, Oxford, 1975).

Translated by V. Isaakyan

Technical Physics, Vol. 48, No. 6, 2003, pp. 682–684. Translated from Zhurnal Tekhnicheskoĭ Fiziki, Vol. 73, No. 6, 2003, pp. 21–23.

GASES AND LIQUIDS

Drop on a Hot Plate: The Appearance of 1/*f* Noise at the Transition to the Ball-Shaped State

V. P. Skripov, A. V. Vinogradov, V. N. Skokov, A. V. Reshetnikov, and V. P. Koverda

Institute of Thermal Physics, Ural Division, Russian Academy of Sciences, Pervomaĭskaya ul. 91, Yekaterinburg, 620216 Russia

e-mail: vnskokov@itp.uran.ru

Received July 10, 2002; in final form, November 25, 2002

Abstract—The evaporation of pentane, hexane, and water drops from a horizontal heating surface is studied experimentally. The temperature at which the drops take a completely spherical shape correlates with limit superheat temperatures under quasi-static conditions, suggesting dynamic burnout. Fluctuations in transient regimes are measured by optical and resistive methods. It is found that the spectral densities of the fluctuation intensity are of a flicker nature. Such behavior is associated with percolation in a set of nonequilibrium phase transitions. *© 2003 MAIK "Nauka/Interperiodica".*

The behavior of a drop evaporating from a hot plate has been the subject of investigation over many years [1–3]. Simple early observations have given impetus to extensive research concerning interaction between a liquid drop and a heated solid wettable surface. As the surface temperature T_s grows, the drop tends to take a spherical shape, ceasing to contact with the surface, and the time of evaporation increases. In the intermediate regime when the spreading drop tends to take a spherical shape, the intermittent "adhesion" and intense boiling of the liquid take place. Such behavior is akin to burnout in a large liquid volume [2]. The spherical shape becomes stable when the temperature T_s approaches the limit superheat temperature T_{sup} at a given external pressure [3, 4], where the liquid can no longer be in contact with the surface. A more careful consideration of drop evaporation from a plate implies the inclusion of forces acting in a gap between the evaporating drop and the solid surface [5]. The drop executes vibrations and experiences internal motions, which are associated with the dependence of the surface tension coefficient on the local temperature on the deformed drop surface and with the pressure variation in the gap.

The aim of this work is to reveal flicker noise (or $1/f$ noise) in the drop-on-hot-plate system. In other words, we seek fluctuations of physical parameters whose power spectrum depends on frequency f as $1/f$. Flicker noise appears in a system when the energy of high-frequency fluctuations is transferred to low-frequency fluctuations with the possible formation of great low-frequency surges. Flicker noise exists in various systems [6]. While the origin of $1/f$ noise is of general interest, it nevertheless remains poorly understood. A commonly accepted general picture of this phenomenon is still absent, and mechanisms responsible for fluctuation spectra of such a type are sometimes unclear. In [7–13], we studied the dynamics of fluctuations in such exotic (in terms of flicker noise investigation) phenomena as burnout under Joule heating, the explosive boiling-up of superheated liquid jets, oscillatory combustion, and an electric arc. It has been established that, under critical and transient regimes, these processes exhibit low-frequency high-energy pulsations with power spectra inversely proportional to the frequency. We managed for the first time to separate out an elementary source of fluctuations with the $1/f$ power spectrum and controllably vary its kinetics in experiments.

Based on our experimental studies, we constructed the original model of flicker noise [8, 14], according to which $1/f$ fluctuations appear when nonequilibrium phase transitions in a system occur simultaneously and interact with each other in the presence of high-intensity white noise.

A drop evaporating from a hot plate represents a flow system (heat and vapor flows) subject to disturbances, which arise when centers of boiling appear at the points of liquid–surface contact. The transition to the ball-shaped state as a manifestation of burnout may be considered as a nonequilibrium phase transition [10]. The drop-on-hot-plate system may therefore generate flicker noise.

We experimented with drops of organic liquids (*n*-pentane and hexane) and water that evaporated from a thick copper plate. First, the total time of evaporation τ_{ev} as a function of the plate temperature T was measured. The results summarized in Fig. 1 are similar to those obtained in [1–3]. The left-hand descending portions of the curves $\tau_{ev}(T)$ are taken at temperatures where the drop is in contact with the heating surface and experiences nucleate boiling. The increase in the time of evaporation with growing temperature means

1063-7842/03/4806-0682$24.00 © 2003 MAIK "Nauka/Interperiodica"

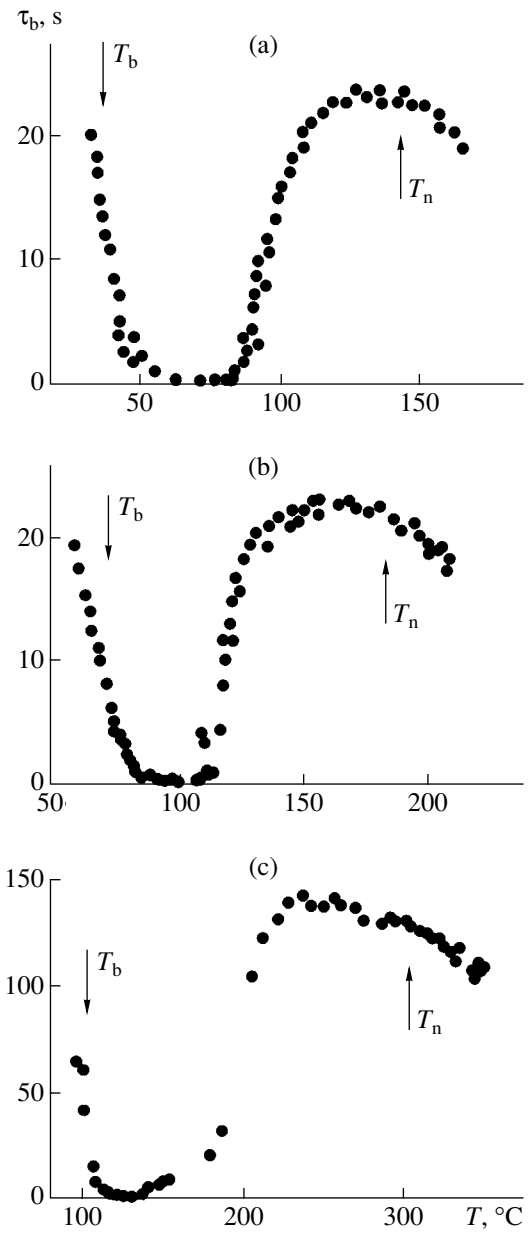

Fig. 1. Total time of drop evaporation from the hot plate vs. plate temperature. (a) *n*-pentane (the boiling temperature under atmospheric pressure T_b = 36°C, T_n = 144°C), (b) hexane (T_b = 69°C, T_n = 182°C), and (c) water (T_b = 100°C, T_n = 304°C). The drop diameter is $d \simeq 4$ mm.

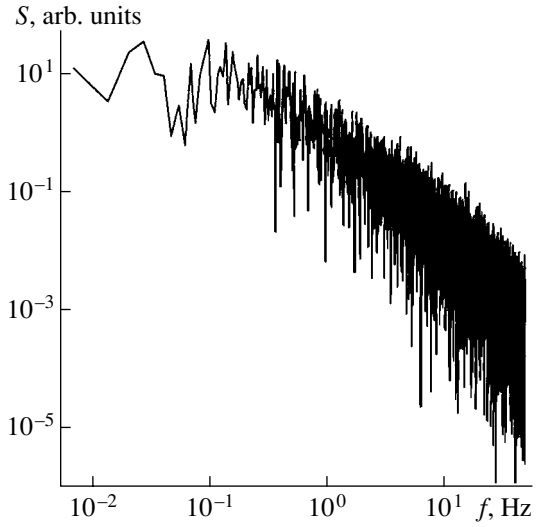

Fig. 2. Fluctuation power spectrum for the developed nucleate boiling of *n*-pentane on the hot plate.

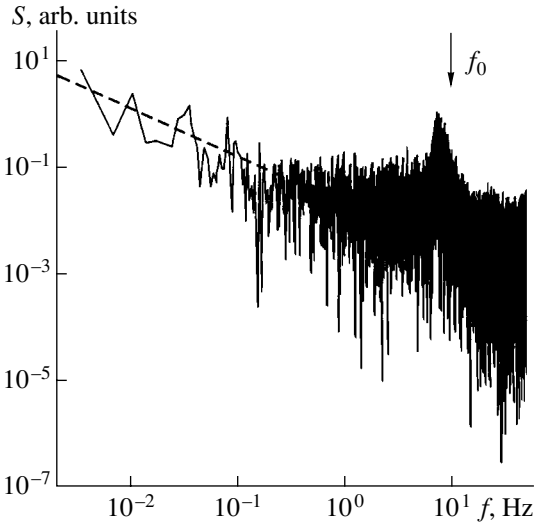

Fig. 3. Fluctuation power spectrum when the drop tends to take a spherical shape. The dashed line shows the dependence $S \sim 1/f$.

that the drop tends to take a spherical shape. When the drop ceases to spread and takes a spherical shape, the heat exchange rate decreases. In this temperature range, the drop–surface contact is intermittent. As the temperature of the plate rises further, the liquid completely separates from the heating surface, the heat exchange rate becomes constant, and the total time of evaporation again smoothly decreases.

Figure 1 shows that the position of the maximal time of evaporation (when the ball-shaped state becomes stable) for pentane and hexane is close to the temperature of homogeneous nucleation: T_n = 144°C (*n*-pentane) or 182°C (hexane) [15]. The water drop acquires a completely spherical shape at temperatures below the temperature of homogeneous nucleation predicted by the theory (T_n = 304°C). However, in this case, too, the position of the maximum correlates with experimental data for limit superheat temperatures found in a clean bubble camera ($T_s \approx 250$°C) [15]. Our findings are consistent with the idea [3] that burnout observed when a

liquid contacts with a hot surface is of a thermodynamic nature.

Fluctuations of drop evaporation from the hot plate were measured by optically recording the intensity of laser radiation passing through the drop edge.

The steadiness of the process was maintained by means of a leak consisting of a controllable-head vessel, a connecting tube, and a glass flask with a capillary. The gap between the capillary and hot surface was varied. The intensity variation of the laser radiation when it passed through the edge of the drop was recorded with an FD256 photodiode with a sensitivity of 0.5 A/W. The signal from the photodetector was applied to an ADC and then memorized.

In the experiments with water, we also used a resistive method to measure electrical fluctuations when the current was passed through a slightly salted drop (a 1% aqueous solution of NaCl). One electrode was the heated copper plate, and the other was placed in the capillary. The fluctuations of the voltage drop across the gap between the plate and capillary were measured. In the resistive measurements, the temperature range was limited by the temperature at which the drop completely separated from the hot surface. The results obtained by both methods coincided.

From the voltage drop fluctuations measured, we found the power spectra by applying the Fourier transformation. In the temperature range corresponding to nucleate boiling and the spherical shape of the drop, the power spectra of the fluctuations had a typical Lorentzian form. Figure 2 shows the fluctuation power spectrum for the developed nucleate boiling of pentane on the hot plate.

In the range where the drop tends to a spherical shape (the ascending portions of the curves τ_{ev}), the spectra change. In particular, a low-frequency portion with the inverse frequency dependence ($1/f$ noise) appears. The fluctuation power spectrum in the transient regime of pentane drop evaporation is depicted in Fig. 3. The fundamental vibrational mode of a drop of radius r has a frequency $f_0 = \sqrt{8\sigma/\rho r^3}/2\pi$ [5], where σ is the surface tension coefficient at the liquid–vapor interface and ρ is the density of the liquid (in Fig. 3, the frequency f_0 is indicated by the arrow). The contribution of this frequency increases as the plate temperature grows and the drop approaches the spherical shape. For hexane and water, the results are similar.

Thus, the temperature at which a drop on a hot plate takes a completely spherical shape correlates with the limit superheat temperature under quasi-stationary conditions. This leads one to argue that burnout observed when a drop is in contact with a hot surface has a thermodynamic origin. In the transient regime when the drop tends to take a spherical shape, the fluctuation power spectrum is inversely proportional to the frequency ($1/f$ noise). The $1/f$ noise discovered is due to nonequilibrium phase transitions occurring in the system in the presence of white noise. White noise is produced by random events of vapor bubble generation.

ACKNOWLEDGMENTS

This work was supported by the Russian Foundation for Basic Research (grant nos. 00-02-16288 and 00-15-96519).

REFERENCES

1. N. A. Pletneva and P. A. Rebinder, Zh. Fiz. Khim. **20**, 961 (1946).
2. S. S. Kutateladze, *Heat Transfer upon Boiling and Condensation* (Mashgiz, Moscow, 1952).
3. V. P. Skripov, Tr. Ural. Politekhn. Inst. (UPI), Ser. Fiz. **123**, 50 (1962).
4. V. P. Skripov, *Metastable Liquid* (Nauka, Moscow, 1972).
5. V. G. Levich, *Physicochemical Hydrodynamics* (Fizmatgiz, Moscow, 1959).
6. Sh. M. Kogan, Usp. Fiz. Nauk **145**, 285 (1985) [Sov. Phys. Usp. **28**, 170 (1985)].
7. V. P. Koverda, V. N. Skokov, and V. P. Skripov, Pis'ma Zh. Éksp. Teor. Fiz. **63**, 739 (1996) [JETP Lett. **63**, 775 (1996)].
8. V. P. Koverda, V. N. Skokov, and V. P. Skripov, Zh. Éksp. Teor. Fiz. **113**, 1748 (1998) [JETP **86**, 953 (1998)].
9. V. N. Skokov, V. P. Koverda, and A. V. Reshetnikov, Pis'ma Zh. Éksp. Teor. Fiz. **69**, 590 (1999) [JETP Lett. **69**, 636 (1999)].
10. V. N. Skokov, A. V. Reshetnikov, and V. P. Koverda, Teplofiz. Vys. Temp. **38**, 786 (2000).
11. A. V. Reshetnikov, N. A. Mazheĭko, V. P. Koverda, *et al.*, Dokl. Akad. Nauk **380**, 176 (2001) [Dokl. Phys. **46**, 612 (2001)].
12. A. V. Reshetnikov, A. V. Vinogradov, V. P. Koverda, *et al.*, Dokl. Akad. Nauk **374**, 481 (2000) [Dokl. Phys. **45**, 515 (2000)].
13. V. N. Skokov, V. P. Koverda, and A. V. Reshetnikov, Zh. Éksp. Teor. Fiz. **119**, 613 (2001) [JETP **92**, 535 (2001)].
14. V. P. Koverda and V. N. Skokov, Dokl. Akad. Nauk **366**, 752 (1999) [Dokl. Phys. **44**, 350 (1999)].
15. *Thermal Physical Properties of Metastable Liquids: A Handbook*, Ed. by V. P. Skripov, E. N. Sinitsyn, P. A. Pavlov, *et al.* (Atomizdat, Moscow, 1980).

Translated by V. Isaakyan

Technical Physics, Vol. 48, No. 6, 2003, pp. 685–692. Translated from Zhurnal Tekhnicheskoĭ Fiziki, Vol. 73, No. 6, 2003, pp. 24–31.

GASES AND LIQUIDS

On Wave Radiation Induced by a Fast Shock Wave in an Oblique Magnetic Field

A. A. Lyubchich* and M. I. Pudovkin**

** Polar Geophysical Institute, Karelian Scientific Center, Russian Academy of Sciences, ul. Fersmana 14, Apatity, Murmansk oblast, 183010 Russia*

e-mail: general@pgi.ru

*** Research Institute of Physics, St. Petersburg State University, Ul'yanovskaya ul. 1, Petrodvorets, St. Petersburg, 198904 Russia*

Received October 21, 2002

Abstract—The corrugation stability of the flat surface of a fast magnetohydrodynamic shock wave in a perfect monoatomic gas with a constant heat capacity is studied with numerical techniques. The magnetic field makes an arbitrary angle with the plane of discontinuity. It is shown that the shock wave remains stable only if it is strictly perpendicular to the magnetic field. At any other angle between the fast shock wave and magnetic field, the former may spontaneously radiate outwardly propagating magnetohydrodynamic waves under certain conditions. Incoming flow characteristics at which these waves are induced are determined. © *2003 MAIK "Nauka/Interperiodica".*

INTRODUCTION

Let us make a brief retrospective journey into the history of studies on hydrodynamic (in the absence of a magnetic field) shock waves. It was shown [1] that, if the conditions

$$j^2\left(\frac{\partial V_2}{\partial P_2}\right)_H < -1 \quad \text{or} \quad j^2\left(\frac{\partial V_2}{\partial P_2}\right)_H > 1 + 2\mathrm{M}_2 \qquad (1)$$

are satisfied, a shock wave is completely unstable. Here, j is the material flux, V is the specific volume, and M is the Mach number. The subscript H means that differentiation is along the Hugoniot adiabat, and subscript 2 refers to quantities in the compressed medium behind the shock front. With the first-order perturbation method, the authors of [1, 2] demonstrated that, if the condition

$$\frac{1 - \mathrm{M}_2^2 - \chi\mathrm{M}_2^2}{1 - \mathrm{M}_2^2 + \chi\mathrm{M}_2^2} < j^2\left(\frac{\partial V_2}{\partial P_2}\right)_H < 1 + 2\mathrm{M}_2 \qquad (2)$$

(where χ is the shock compression) is satisfied, the shock wave will spontaneously generate sound (so-called neutral instability). More detailed analysis (in a second-order approximation) [3] revealed that, under conditions (2), acoustic radiation can be triggered only by an external perturbation that has a high order of smallness and is therefore disregarded in the linear approach. Yet domain (2) does remain physically separated. For example, the amplitude of certain components of noise background behind the front of a shock wave satisfying conditions (2) is expected to be anomalously large [3]. However, it remains unclear whether such a wave can exist for a finite time, since, in the range of the derivative of the Hugoniot adiabat that totally covers domains (1) and (2), the shock wave turned out to be structurally unstable: it may break into a shock wave of greatly differing amplitude and other elements [3]. Hugoniot adiabats satisfying conditions (1) and (2) have a very exotic form and are hard to realize in experiment. In those rare cases when Hugoniot adiabats satisfy the conditions desired, the break of the shock wave, rather than its corrugation instability, is observed [4, p. 536]. Therefore, the issue of whether either corrugation instability or induced acoustic radiation may be observed in conventional hydrodynamics remains open.

In magnetohydrodynamics, stability analysis has been applied largely to parallel shock waves. It was demonstrated [5] that the stability test for these waves is conditions (1). Conditions for spontaneous wave radiation by a parallel shock wave were considered in [6]. It appeared that the upper bound of the neutral stability range is the same as in conventional hydrodynamics (conditions (2)). The lower bound differs from the hydrodynamic value and depends on the magnetic field. Even in a gas with a constant heat capacity, a structurally stable shock wave of sufficiently high intensity must spontaneously emit magnetosonic waves.

However, a parallel shock wave is a mere idealization, since parallelism is violated by any fluctuation of the magnetic field direction. Therefore, in this work the neutral stability of a shock in an oblique magnetic field

1063-7842/03/4806-0685$24.00

is studied under the assumption that the medium behaves as a perfect gas with a constant heat capacity. Our aim is to find magnetic field orientations and discontinuity parameters at which the discontinuity spontaneously generates waves. Results obtained may be useful for the experimental verification of the existing theory of interaction between shock waves and small perturbations.

ANALYTICAL METHOD

We assumed that a magnetized medium behaves as a perfect gas with a constant heat capacity. The polytropic exponent is set equal to $\gamma = 5/3$. Let the gas be divided by an infinitely thin plane shock wave into two half-spaces. The analysis is performed in the laboratory coordinate system related to the undisturbed surface of the shock. The x axis is aligned with the normal to the undisturbed surface and with the flow direction. Then, the undisturbed surface of the shock coincides with the plane yz. The tangential component of the flow velocity before the shock is assumed to be absent. This is provided by appropriately choosing the coordinate system. The problem of shock-induced wave radiation we are interested in is a specific case of the problem of interaction between magnetohydrodynamic (MHD) waves and small perturbations, which was first considered in [7]. This problem is solved in several steps [7].

(i) A complete set of MHD equations includes two Maxwell equations, the continuity equation for material flux, the Euler equation, the entropy conservation equation, and the equation of state:

$$\nabla\cdot\mathbf{B} = 0;\quad \frac{\partial\mathbf{B}}{\partial t} = \nabla\cdot(\mathbf{v}\times\mathbf{B});\quad \frac{\partial\rho}{\partial t} + \nabla\cdot(\rho\cdot\mathbf{v}) = 0;$$
$$\frac{\partial\mathbf{v}}{\partial t} + (\mathbf{v}\cdot\nabla)\mathbf{v} = -\frac{1}{\rho}\nabla P - \frac{1}{4\pi\rho}\mathbf{B}\times(\nabla\times\mathbf{B}); \quad (3)$$
$$\frac{\partial S}{\partial t} + (\mathbf{v}\cdot\nabla)S = 0;\quad P(\rho, S) = \rho^{\gamma}\exp\left(\frac{S}{c_V}\right).$$

The standard procedure of linearization of these equations defines all possible types of plane linear MHD waves and their properties in the laboratory frame of reference.

(ii) A set of boundary conditions that reflect the laws of conservation defines all possible types of plane MHD discontinuities and their properties. The complete set of laws of conservation includes the continuity conditions for the material flux, energy flux, normal and tangential components of the momentum flux, tangential component of the electric field, and normal component of the magnetic field:

$$[\rho\cdot v_n] = 0;$$
$$\left[\rho v_n\left(\frac{v_1^2}{2} + \frac{\gamma}{\gamma-1}\frac{P}{\rho}\right) + \frac{1}{4\pi}(v_n B^2 - B_n(\mathbf{v}\cdot\mathbf{B}))\right] = 0;$$
$$\left[P + \rho v_n^2 + \frac{1}{8\pi}(B_t^2 - B_n^2)\right] = 0; \quad (4)$$
$$\left[\rho v_n\mathbf{v}_t - \frac{1}{4\pi}B_n\mathbf{B}_t\right] = 0;$$
$$[B_n\mathbf{v}_t - \mathbf{B}_t v_n] = 0;\quad [B_n] = 0.$$

Here, $[A] = A_2 - A_1$ means the difference between the values of A on both sides of the discontinuity. The boundary conditions in this form are uniquely derived from the integral MHD equations. The parameters of the shock wave are determined as follows. All necessary parameters of the uncompressed medium (Mach number, ratio between thermal and magnetic pressures β, and magnetic field orientation) are assumed to be known. Solving the cubic equation derived in [8], we find a discontinuity in the tangential component of the magnetic field across the shock. Then, all the necessary flow parameters in the compressed medium are readily found from set (4).

We will consider only those parameter values uniquely defining the characteristics of the medium behind the shock front. Therefore, the case $\mathbf{B}_{t1} = 0$, when two more conditions,

$$v_{x1} > c_{s1}$$
$$\text{and}\quad v_{A1}^2 \le v_{n1}^2 \le \frac{1}{\gamma-1}[(\gamma+1)v_{A1}^2 - 2c_{s1}^2], \quad (5)$$

are satisfied simultaneously, is excluded from consideration.

As is known [9], a solution to set (4) is not only a parallel shock wave but also an inclusion shock wave.

(iii) All the quantities are represented in the form $A = A_0 + \delta A$, where A_0 is the background value of a quantity in the laboratory frame of reference and δA is its small perturbation. Linearizing the boundary conditions yields a relation between perturbed quantities on both sides of the shock wave.

(iv) In the problem stated, the very surface of the shock turns out to be perturbed: it executes linear oscillations about its initial (unperturbed) position. Without loss of generality, we may assume that these oscillations propagate along the y axis. Then, the equation for the perturbed surface will have the form $\eta = \eta_0\exp(i(k_y y - \omega t))$, where η is the displacement of the surface points along the x axis. Such oscillations of the shock front may be either spontaneous or caused by an external reason (an incident linear wave). As a result, the linearized boundary conditions turn out to be written in the local frame of reference related to the oscillating shock front. Our aim is to pass into the laboratory

frame of reference. According to today's concepts, to do this would require taking into account the distortion and an additional velocity of the shock front [7]. With δA represented as $\delta A^W + \delta A^S$, where δA^S describes the transition desired, in view of both effects, we arrive at

$$\begin{aligned}
\delta v_{x1}^S &= \frac{\omega}{k_y} i k_y \eta, \quad \delta v_{x2}^S = -\left(v_{y2} - \frac{\omega}{k_y}\right) i k_y \eta, \\
\delta v_{y1}^S &= v_{x1} i k_y \eta, \quad \delta v_{y2}^S = v_{x2} i k_y \eta, \\
\delta B_{x1}^S &= -B_{y1} i k_y \eta, \quad \delta B_{x2}^S = -B_{y2} i k_y \eta, \\
\delta B_{y1}^S &= B_{x1} i k_y \eta, \quad \delta B_{y2}^S = B_{x2} i k_y \eta.
\end{aligned} \tag{6}$$

One can believe that relationships (6) formally describe the oscillation of the shock front in the laboratory frame of reference.

(v) The disturbed quantities δA^W are expressed through the amplitudes and angles of propagation of the incident wave (if there are any in the problem) and all the theoretically possible outgoing MHD waves. In general, six outgoing waves may exist (one entropy wave, two Alfvén waves, and three magnetosonic waves all propagating behind the fast shock front in the compressed medium). The propagation directions of the outgoing waves are found from the continuity condition at the discontinuity of frequency ω and for the shear component k_y of the wavevector [7]. The continuity of these quantities directly follows from the smallness of the perturbations and does not depend on the specific form of the boundary conditions.

Difficulties associated with determining the angles of propagation of the magnetosonic waves are resolved as follows [10]. Perturbations transferred by magnetosonic waves have the form $\delta A = \delta A_0 \exp(i(k_x x + k_y y - \omega t))$. Here, ω and k_y are given, and k_x (generally complex) is a desired quantity. Linearized set (3) yields the four-degree algebraic equation

$$A_0\left(\frac{k_x}{k_y}\right)^4 + B_0\left(\frac{k_x}{k_y}\right)^3 + C_0\left(\frac{k_x}{k_y}\right)^2 + D_0\left(\frac{k_x}{k_y}\right) + E_0 = 0 \tag{7}$$

with the coefficients

$$A_0 = v_{x2}^4 - v_{x2}^2(c_{s2}^2 + v_{A2}^2) + v_{Ax2}^2 c_{s2}^2,$$

$$B_0 = 2[v_{x2}(c_{ph} - v_{y2})(c_{s2}^2 + v_{A2}^2 - 2v_{x2}^2) + v_{Ax2} v_{Ay2} c_{s2}^2],$$

$$C_0 = -(v_{x2}^2 + (c_{ph} - v_{v2})^2)(c_{s2}^2 + v_{A2}^2) + c_{s2}^2(v_{Ax2}^2 + v_{Ay2}^2) + 6(c_{ph} - v_{y2})^2 v_{x2}^2,$$

$$D_0 = 2[v_{x2}(c_{ph} - v_{y2})(c_{s2}^2 + v_{A2}^2 - 2(c_{ph} - v_{y2})^2) + v_{Ax2} v_{Ay2} c_{s2}^2],$$

$$E_0 = (c_{ph} - v_{y2})^4 - (c_{ph} - v_{y2})^2(c_{s2}^2 + v_{A2}^2) + c_{s2}^2 v_{Ay2}^2,$$

where c_{ph} is the phase velocity common to all the MHD waves and tangent to the boundary and c_{s2} and v_{A2} are the velocity of sound and the Alfvén velocity behind the front. Equation (7), which defines the directions of four (one incident and three outgoing) MHD waves behind the front, has either four real roots or two real and two complex conjugate roots. The latter case will be excluded from consideration for the following reasons. Our goal is to find situations where a shock wave generates a fast outgoing magnetosonic wave. However, a fast magnetosonic wave corresponding to a solution with complex k_x travels (in terms of energy transfer) parallel to the shock front and is not outgoing.

In a number of exotic cases, Eq. (7) behind the shock front may have four real roots corresponding to four slow magnetosonic waves one of which strikes the shock [10]. Such a situation is consistent [7, 10] with the well-known condition $v_{Ax2} < v_{x2} < u_{fast,2}$ for the evolution of a fast shock wave traveling normally to the shock [9]. Such cases will also be eliminated from consideration, since here, too, the shock does not generate a fast outgoing magnetosonic wave.

Thus, below we will consider those conditions when Eq. (7) has four real roots responsible for two fast and two slow magnetosonic waves. The mode with the lowest x component of the wavevector is that falling on the shock front.

Along with the propagation directions, linearized set (3) defines the polarization of magnetosonic waves directly in that frame of reference where the problem is solved (which is very convenient):

$$\delta \mathbf{B} = (\mathbf{B}_2 k^2 - \mathbf{k}(\mathbf{k} \cdot \mathbf{B}_2))(\mathbf{k} \cdot \mathbf{v}_2 - \omega)\delta A,$$

$$\delta \mathbf{v} = [-\mathbf{k}(\mathbf{k} \cdot \mathbf{v}_2 - \omega)^2 + k^2(\mathbf{k} \cdot \mathbf{v}_{A2})]\delta A,$$

$$\delta\rho = -\frac{\mathbf{k} \cdot \delta\mathbf{v}}{\mathbf{k} \cdot \mathbf{v} - \omega}\rho; \quad \delta P = c_s^2 \delta\rho,$$

where δA is the amplitude common to all the perturbations transferred by a given wave.

(vi) The overall perturbations expressed through the known angles of propagation, known amplitude of the incident wave (if any), unknown amplitudes of outgoing waves, and unknown amplitude of front oscillation are substituted into the perturbed set of boundary conditions. Eventually, the eight boundary equations in set (4) transform into a set of seven linearly independent equations in seven unknowns [7].

Analysis of shock wave stability against the generation of outgoing MHD waves can be performed in two equivalent ways [11]. In one, by equating the determinant of the set of linear equations to zero, one arrives at the characteristic equation yielding the phase velocity c_{ph} tangent to the shock and common to all the waves

generated. Then, a search for roots with real c_{ph} will shed light upon the stability of the shock. In the other way, the transformation ratio for the incident wave is analyzed. If it goes to infinity (resonance transmission or reflection), the shock is unstable against the generation of outgoing waves.

We will take advantage of the second approach and study the reflection of a fast magnetosonic wave from a fast shock wave, assuming that the magnetosonic wave is incident from the compressed medium. For a numerical solution, it is necessary to choose a numerical test for resonance reflection. We will assume that a shock is unstable against outgoing wave emission if the pressure oscillation amplitude in a fast magnetosonic wave reflected more than ten times exceeds that in the incident wave. This figure is one order of magnitude higher than the reflection coefficient typical of a stable shock wave in magnetohydrodynamics. As for conventional hydrodynamics, the coefficient of reflection of sound in a perfect gas with a constant heat capacity is always less than unity in magnitude (the reflection of sound from a hydrodynamic shock wave was considered in [12]). Therefore, the test chosen seems to be quite reliable. With a more stringent test (i.e., the reflection coefficient is more than 100), the probability that the range of relative instability will escape notice in numerical calculations increases greatly.

We used a computer program that was written for a more general case of calculating the transformation ratio for MHD waves reflected from shock waves and rotational discontinuities. The program was carefully debugged and tested. For example, we determined that, in those special cases allowing for an analytical solution, it completely coincides with the result of numerical calculation. These special cases are (1) a shock wave in a weakly magnetized plasma ($\beta \gg 1$), (2) a strong shock wave ($c_{s1} \approx v_{A1} \ll v_{x1}$) [10], (3) a shock wave normal to the magnetic field that is aligned with the z axis, i.e., is orthogonal to all the wavevectors [10, 13], (4) normal incidence of MHD waves on a rotational discontinuity [10], and (5) a rotational discontinuity in a highly magnetized plasma ($\beta \ll 1$) [10]. Based on these tests, we can argue that the results obtained by the conventional approach stated above are mathematically valid.

RESULTS OF NUMERICAL CALCULATION

We consider the reflection of a fast linear plane magnetosonic wave propagating in the compressed medium from a fast plane evolving shock wave, seeking discontinuity parameters and the incident wave orientation such that the magnitude of the reflection coefficient in terms of pressure or density oscillation exceeds ten. This situation is treated as a relative instability of a fast shock wave. The angle between the normal to the discontinuity and an external magnetic field is assumed to be arbitrary. In the calculations that follow, emphasis will be on the case $B_x \geq 0$. The medium is viewed as a perfect monoatomic gas ($\gamma = 5/3$).

Let us first consider the case when the velocity of sound c_s before the shock wave is much lower than the flow velocity v_n; that is, $M_1 \longrightarrow \infty$. Figure 1 shows the variation of the relative instability domains with the inclination ψ_1 of the external magnetic field and angle Θ between the tangential components of the wavevectors and magnetic field. The pattern is almost independent of the sign of B_x ("almost" means that, when $B_x \geq 0$, Θ is the angle between the vectors $\mathbf{k}_t$ and $\mathbf{B}_t$, while for $B_x \leq 0$, this angle is made by the vectors $\mathbf{k}_t$ and $-\mathbf{B}_t$). Two domains of relative instability are seen to exist. They occupy mainly sectors bounded by straight lines crossing at an angle of 106°. In our problem, the dependence on the sign of the z component of the magnetic field is absent ($\mathbf{k}_t$ is aligned with the y axis). As a result, the relative instability domains are symmetric about the y axis. Under appropriate conditions (an appropriate value of the magnetic field and an appropriate direction of the wavevector of the incident wave), the fast shock wave will be relatively unstable at any inclination of the magnetic field. The only exception is the case of a strictly perpendicular shock wave ($\psi = 90°$), which remains stable at any Θ. This result is not contradictory to the analytical solution to the problem of MHD wave transformation on a shock wave when the magnetic field is aligned with the z axis [13]. It is of interest that passage to the limit of a perpendicular shock wave cannot be realized. In the presence of a normal component of the magnetic field, however small, the shock wave may become unstable (the numerical calculations were made up to the angle $\psi_1 = 89.999°$). The relative instability domain on the left of Fig. 1 extends up to $\psi_1 = 0.1°$ at $\Theta = 0$, i.e., includes the case of a parallel shock wave. This is in agreement with results in [6].

What is the value of the magnetic field at which the shock wave becomes relatively unstable? The answer to this question is given in Fig. 2. It shows the dependence of the instability domains on the inclination of the magnetic field and on the Alfvén number $M_{A1} = v_{n1}/v_{A1}$ for the case when the magnetic field, all the wavevectors, and the normal to the discontinuity lie in the same plane. It is assumed, as before, that $c_{s1} \ll v_{n1}$. The left-hand side of Fig. 2 corresponds to the situation when $\mathbf{k}_t$ and $\mathbf{B}_t$ are antiparallel; the right-hand side, to the situation when $\mathbf{k}_t$ and $\mathbf{B}_t$ are parallel (we assume that $B_x \geq 0$). The instability domain is rather broad for the nearly perpendicular shock wave (recall that a strictly perpendicular shock wave always remains stable) and decreases with decreasing ψ_1. The lower bound of the instability domain is slightly sensitive to the magnetic field inclination and varies from $M_{A1min} = 3.4$ to 1.8. Conversely, the upper bound is highly sensitive to ψ_1. Since we are dealing with the case $M_1 \longrightarrow \infty$, it can be said that the instability is observed in a high magnetic field ($v_{A1} \gg c_{s1}$).

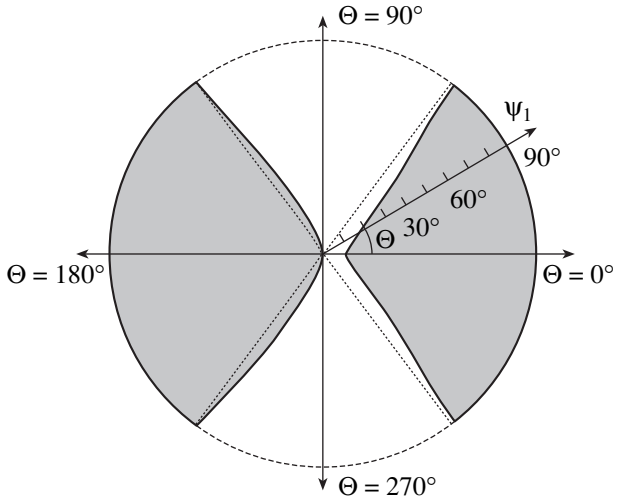

Fig. 1. Variation of the relative instability domains (colored gray) for a fast shock wave with the orientation of an external magnetic field in the uncompressed medium. The magnetic field and the normal to the discontinuity make an angle ψ_1, which is plotted along the radius and varies from 0 (parallel shock wave) to 90° (perpendicular shock wave). The angle Θ between the tangential components of the magnetic field and wavevector of a shock-induced fast magnetosonic wave is plotted as a polar angle. With $\Theta = 0°$ (180°), the vectors $\mathbf{B}_{t1}$ and $\mathbf{k}_t$ are parallel (antiparallel); with $\Theta = 90°$ (or 270°), they are orthogonal to each other. It is assumed that $M_1 \longrightarrow \infty$.

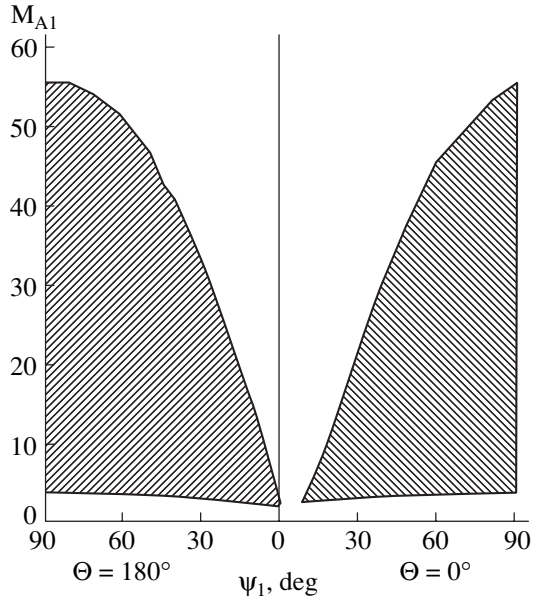

Fig. 2. Variation of the relative instability domains (dashed areas) for a fast shock wave with the inclination ψ_1 of an external magnetic field and Alfvén number M_{A1} in the uncompressed medium. It is assumed that $M_1 \longrightarrow \infty$ and also that $\mathbf{k}_t$, $\mathbf{B}_{t1}$, and the normal to the discontinuity lie in the same plane. The vertical line corresponds to the case of a parallel shock wave ($\psi_1 = 0$). On the left (right) of the vertical line, the vectors $\mathbf{k}_t$ and $\mathbf{B}_{t1}$ are parallel (antiparallel) to each other.

Figure 3 shows angles of incidence at which the resonance reflection of the fast magnetosonic wave is observed vs. the Alfvén number (magnetic field strength). In the calculations, we took $\Theta = 180°$, $\psi_1 = 80°$, and $c_{s1} \ll v_{n1}$. The angle of incidence was counted from the x-axis direction; that is, the angle of normal incidence was 180°. The instability is observed at fairly large angles that are close to the maximum permissible angles of propagation. With such angles of incidence, the angle of reflection is, as a rule, larger than 90°; i.e., the wavevector of the reflected wave is directed toward the discontinuity, but the wave is carried away by the flow. For Alfvén numbers corresponding to the boundary of the instability domain, the resonance angle tends to the limiting angle of incidence. In this case, both the incident wave and the reflected wave propagate (in terms of energy transfer) parallel to the shock front. The strong reflection domain is very narrow: its width, as a rule, does not exceed 1°. When the numerical test for instability is taken to be more stringent, this domain shrinks, tending to a line in the limit. At fixed M_{A1}, resonance reflection is observed at a specific angle of incidence.

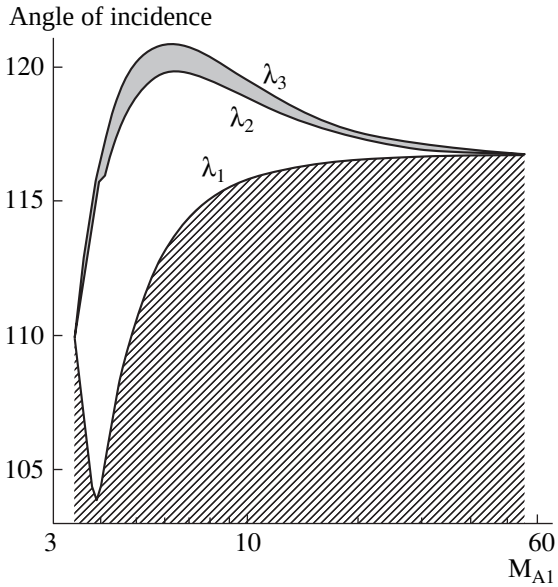

Fig. 3. Angle of incidence of a fast magnetosonic wave at which this wave resonantly reflects (the reflection coefficient for pressure oscillations exceeds ten) vs. Alfvén number in the uncompressed medium. It is assumed that the orientation of the magnetic field is fixed ($\psi_1 = 80°$, $\Theta = 180°$) and Mach number $M_1 \longrightarrow \infty$. The wave is incident from the compressed side. The angle of incidence λ is measured from the x-axis direction ($\lambda = 180°$ corresponds to normal incidence). With the Alfvén number constant, the incident wave can exist only in the range $\lambda_1 < \lambda \le 180°$ (the forbidden range of the angles is dashed). Resonant reflection is observed in the interval $\lambda_2 \le \lambda \le \lambda_3$ (colored gray).

According to our calculations, the fast evolving shock wave subjected to an oblique magnetic field may be unstable against the emission of a fast magnetosonic wave. The question arises as to whether the theoretical conditions for relative instability can be implemented experimentally. To answer this question, let us see how

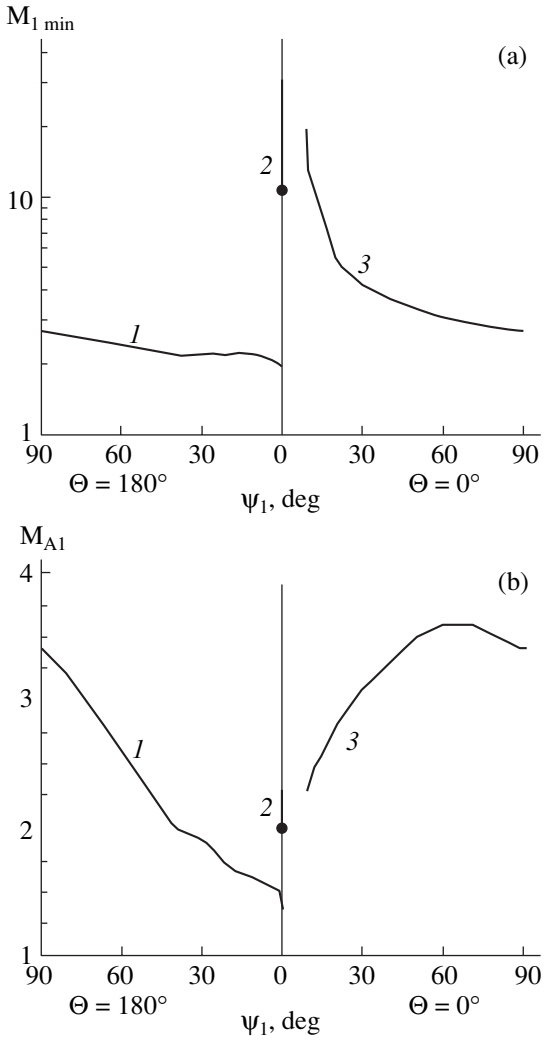

Fig. 4. (a) Minimal Mach number M_{1min} at which a fast shock wave may spontaneously radiate a fast magnetosonic wave vs. inclination of the magnetic field ψ_1 and (b) the Alfvén number M_{A1} at which a shock wave with M_{1min} is relatively unstable. The vertical line corresponds to a parallel shock wave ($\psi_1 = 0$). With $\Theta = 180°$, spontaneous radiation can be observed in the range (*1*) $0° < \psi_1 < 90°$. The limiting case of relative instability of the parallel shock wave is not realized. With $\Theta = 0°$, spontaneous radiation can take place at (*2*) $0° \le \psi_1 \le 0.1°$ and (*3*) $9° \le \psi_1 < 90°$. In these cases, the limiting case of relative instability of the parallel shock wave takes place. A perpendicular wave remains stable at any Θ.

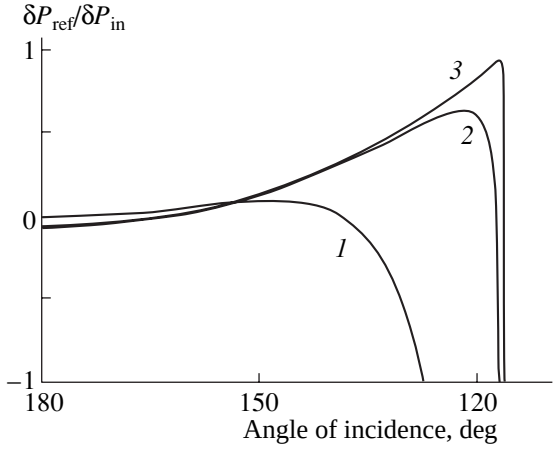

Fig. 5. Coefficient of reflection of an acoustic wave from a hydrodynamic shock wave vs. angle of incident for M_1 = (*1*) 2, (*2*) 10, and (*3*) 100. Normal incidence is observed at 180°.

the instability domains depend on the Mach number. As the Mach number decreases, the magnetic field range favorable to resonance reflection (see Fig. 2) narrows, tending to a certain value in the limit. In Fig. 4, ψ_1 is plotted against the minimal Mach number M_{1min} for which the relative instability and the associated Alfvén number can be observed. The calculation was carried out for the case when the magnetic field, all the wavevectors, and the normal to the discontinuity lie in the same plane. For antiparallel $\mathbf{k}_t$ and $\mathbf{B}_t$, the minimal values of the Mach number fall into the interval 1.9–2.7. A shock wave with such characteristics is easy to generate under laboratory conditions. Hence, one can observe shock-induced wave radiation experimentally and, thus, verify the theory of interaction between shock waves and small perturbations. In Fig. 4, a singularity at small magnetic field inclinations is distinctly seen. Let us comment briefly on this point. In [6], it was concluded that a parallel shock wave may be unstable at $M_1 \ge 7$. Our numerical calculations demonstrated that the parallel evolving shock wave is unstable at $M_1 \ge 10.6$. At smaller Mach numbers, the unstable parallel shock wave no longer meets condition (5) and is disregarded. As ψ_1 tends to zero, the value of M_{1min} varies in relation to the angle between $\mathbf{k}_t$ and $\mathbf{B}_t$. The table lists the values of M_{1min} and their related values of M_{A1} for various angles Θ and for $\psi_1 = 0.00001°$. It is seen that passage to the limit of the instability test for the parallel evolving wave ($M_{1min} = 10.6$ and $M_{A1} = 2.0$) is not possible for all values of Θ.

DISCUSSION

As follows from our numerical calculations, a fast shock wave that is stable against decay may induce MHD waves in an oblique magnetic field. Mathematically, the reason for this effect seems to be the specific features of acoustic reflection from a hydrodynamic shock wave. The reflection coefficient of sound vs. angle of incidence for three Mach numbers is shown in Fig. 5. Here, we used an analytical solution found in [12]. In the case of a strong hydrodynamic strong wave and angles of incidence close to the maximum permissible one, the coefficient of reflection of pressure or density oscillations tends to +1; that is, the reflected

acoustic wave amplifies the incident one. Simultaneously, the coefficients of generation of entropy and vortical waves, as well as of the oscillation of the front surface, reach a maximum:

$$\frac{\delta\rho_{en}}{\delta\rho_s^{in}} \longrightarrow -2; \quad \frac{\delta v_{vor}}{\delta v_s^{in}} \longrightarrow \frac{2}{\gamma-1}; \quad \frac{\delta v_{sur}}{\delta v_s^{in}} \longrightarrow -(\gamma+1)\sqrt{\frac{\gamma}{2(\gamma-1)}}. \tag{8}$$

The hydrodynamic strong wave remains stable under these conditions; however, the situation may change if an external magnetic field is imposed. Probably because of this, the resonant angles of incidence shown in Fig. 1 turn out to be close to the maximum permissible values. When the angle of incidence increases further (for finite-intensity shock waves), the reflection coefficient changes sign and tends to –1, as follows from Fig. 5. In this case, the amplitudes of waves being generated tend to zero; that is, total internal reflection takes place. It should be noted that, for an infinitely strong shock wave, the transition to total internal reflection will not occur. An acoustic wave traveling parallel to the surface of such a wave would cause its oscillation and generate an outgoing entropy–vortical wave with an amplitude given by (8).

When writing this article, we were spurred on by the desire to attract experimentalists' attention to the problem of wave radiation induced by a fast shock wave. In our opinion, experimental investigation in this field would be useful for at least two reasons.

First, experimental observation of acoustic radiation induced by a purely hydrodynamic shock wave is hardly possible, as follows from the Introduction. However, the theoretically predicted conditions for neutral instability of a magnetodynamic shock wave are fairly soft. For shock waves of moderate intensity that are subjected to a moderate magnetic field, this phenomenon may be observed in a medium with normal, rather than exotic, thermodynamic properties. In other words, conditions favorable to the spontaneous radiation of MHD waves can be completely realized in experiment. There is no need to specially generate perturbations falling on a magnetodynamic shock wave: it is sufficient to measure the fluctuation level behind its front. Fluctuations are expected to rise significantly in the domain of instability. Further investigation might elucidate the evolution of a neutrally unstable shock wave.

The second, no less important, reason why experimental studies in this field would be of great value is that the theory of interaction between low-amplitude incident waves and shock waves can be verified experimentally. Unfortunately, the interaction of low-amplitude incident waves with shock waves is poorly understood. We know of only one work [14] where experimental results are numerically compared with theoretical predictions. In [14], the reflection of an acoustic wave from a shock wave in a varying-area channel was studied. It was shown that experimental data agree with the theory if $\omega < \omega^*$, where ω^* is defined by the channel geometry. However, in the "low-frequency" range, a solution may depend not on "wave" but on "steady" processes due to a change in the discontinuity geometry. Therefore, it would be appropriate, in our opinion, to substantiate the existing theory in any other (independent) way.

Table

Θ, deg	$M_{1\,min}$	M_{Al}
1–105	10.6	2.0
106	6.4	1.9
110	5.7	1.9
120	3.9	1.8
130	2.7	1.6
140	2.3	1.5
150	2.2	1.5
160	2.1	1.5
170–180	1.9	1.4

To conclude, the goal of these investigations was the analysis of the conditions for outgoing wave radiation induced by a fast shock wave that are predicted by the existing theory. It should be emphasized that we did not intend to discuss the theory as such. The fact is that the boundary conditions at the disturbed shock front (in the noninertial frame of reference) that are adopted in the conventional approach cast some doubt and poses a number of questions. This point calls for special examination [15].

REFERENCES

1. S. P. D'yakov, Zh. Éksp. Teor. Fiz. **27**, 288 (1954).
2. V. M. Kontorovich, Zh. Éksp. Teor. Fiz. **33**, 1525 (1957) [Sov. Phys. JETP **33**, 1179 (1957)].
3. N. M. Kuznetsov, Usp. Fiz. Nauk **159**, 493 (1989) [Sov. Phys. Usp. **32**, 993 (1989)].
4. Ya. B. Zel'dovich and Yu. P. Raizer, *Physics of Shock Waves and High-Temperature Hydrodynamic Phenomena* (Fizmatgiz, Moscow, 1963; Academic, New York, 1966).
5. C. S. Gardner and M. D. Kruskal, Phys. Fluids **7**, 700 (1964).
6. S. F. Pimenov, Zh. Éksp. Teor. Fiz. **83**, 106 (1982) [Sov. Phys. JETP **56**, 57 (1982)].
7. V. M. Kontorovich, Zh. Éksp. Teor. Fiz. **35**, 1216 (1958) [Sov. Phys. JETP **35**, 851 (1958)].
8. J. C. Whang, J. Geophys. Res. **92**, 4349 (1987).
9. L. D. Landau and E. M. Lifshitz, *Electrodynamics of Continuous Media* (Nauka, Moscow, 1982; Pergamon Press, Oxford, 1984).

10. A. A. Lyubchich, Candidate's (Phys.-Math.) Dissertation (St. Petersburg State Univ., St. Petersburg, 1999).
11. L. D. Landau and E. M. Lifshitz, *Course of Theoretical Physics*, Vol. 6: *Fluid Mechanics* (Nauka, Moscow, 1986; Pergamon, New York, 1987).
12. V. M. Kontorovich, Akust. Zh. **5**, 314 (1959).
13. A. A. Lyubchich and M. I. Pudovkin, *Simulations of Processes in the Upper Polar Atmosphere: Collections of Scientific Works* (Polar Geologic. Inst., Russ. Acad. Sci., Murmansk, 1998), pp. 29–62.
14. V. N. Glaznev, Izv. Sib. Otd. Akad. Nauk SSSR, Ser. Tekh. Nauk, No. 3, 99 (1974).
15. A. A. Lyubchich and M. I. Pudovkin, Interaction of Small Perturbations with Shock Waves, Preprint of Polar Geophysical Inst., Russ. Acad. Sci. (Apatity, Murmansk oblast, 2002).

Translated by V. Isaakyan

Technical Physics, Vol. 48, No. 6, 2003, pp. 693–696. Translated from Zhurnal Tekhnicheskoĭ Fiziki, Vol. 73, No. 6, 2003, pp. 32–35.

GASES AND LIQUIDS

Heat-Transfer Coefficient of a Fine Heated Wire in a Gas Flow

A. A. Pikulev

All-Russia Research Institute of Experimental Physics, Russian Federal Nuclear Center, Sarov, Nizhni Novgorod oblast, 607190 Russia

e-mail: pikulev@expd.vniief.ru

Received November 6, 2002

Abstract—Heat exchange between a linear heat source and a gas flow is studied theoretically. The coefficient of source-to-gas heat transfer is found as a function of the gas velocity in the limit of zero temperature drop. It is shown that the heat-transfer coefficient depends linearly on the gas flow velocity for Peclet numbers 0.1–1.0.

INTRODUCTION

At present, the fine wire method is widely used in experimentally studying the thermodynamics of gas flows. This method is appropriate for (1) measuring the gas flow velocity with single-wire [1] or pulsed two-wire [2, 3] anemometers; (2) determining the thermodynamic parameters of a gas flow, for example, the turbulent thermal conductivity of a gas in a wake [4]; and (3) investigating the amount and distribution of the energy delivered to a gas, for example, the energy applied to the lasing medium of nuclear-pumped lasers [5].

Whether the wire will be heated or cooled by the ambient gas depends on the thermodynamic processes occurring in the cell. The heat transfer coefficient of a fine wire (which defines the characteristic heating time of a wire) depends on the flow type (laminar or turbulent), as well as on the velocity and thermodynamic parameters of the gas. Therefore, to analyze the experimental data for the gas flow velocity [1–3] or the energy distribution in the lasing medium [5], one must know how the coefficient of wire-to-gas heat transfer depends on the gas parameters (first of all, on the velocity).

Note that this problem in its complete statement requires a joint solution of the Navier–Stokes and energy transfer equations. Even for low Reynolds and Peclet numbers, such a procedure is very machine time consuming. In this paper, we suggest a simpler model where a wire of finite thickness is replaced by a linear heat source that does not disturb the gas flow in the limit of zero temperature drop between the wire and gas.

For a low Peclet number, the results of this paper can be extended for the case of a finite-thickness wire. It will be shown that the velocity dependence of the heat transfer coefficient is linear if Pe = 0.1–1.0.

BASIC EQUATIONS

Consider the thermodynamics of a linear heat source placed in a gas uniformly flowing in the x direction (perpendicularly to the source) with a velocity U. The heat source with a linear heat-emission density W is assumed to be aligned with the axis 0z. In this case, all the parameters are z independent. We also assume that the processes involved are steady-state and isobaric and that the presence of the heat source does not disturb the stream lines, which are parallel to axis 0x. From the continuity equation [6], it follows

$$\frac{\partial}{\partial x}\rho U = 0 \longrightarrow \rho U = \rho_0 U_0 = \text{const}, \qquad (1)$$

where ρ_0 and U_0 are the gas density and the gas velocity at infinity, respectively. In view of Eq. (1), the energy equation takes the form (with the temperature dependence of the thermal conductivity μ disregarded) [6, 7]

$$c_p \rho_0 U_0 \frac{\partial T}{\partial x} = \mu \nabla^2 T + W\delta(x)\delta(y), \qquad (2)$$

where c_p is the specific heat at constant pressure, T is the temperature, and δ is the delta function.

As the boundary condition, we assume that the gas temperature is equal to zero at infinity: $T(\infty) = 0$.

The heat-transfer coefficient k can be determined from the relationship [7]

$$W = 2\pi k R T_0, \qquad (3)$$

where R is the wire radius and T_0 is the temperature of the wire surface.

1063-7842/03/4806-0693$24.00

Let us introduce the dimensionless variables

$$\begin{cases} \bar{x} = \dfrac{x}{R}, \quad \bar{y} = \dfrac{y}{R}, \quad \bar{T} = \dfrac{T}{T_0} \\ \overline{W} = \dfrac{W}{\mu T_0}, \quad \bar{k} = k\dfrac{R}{\mu}, \quad \mathrm{Pe} = \dfrac{c_p \rho_0 R U_0}{\mu}, \end{cases} \tag{4}$$

where Pe is the Peclet number [6].

Equation (3) takes the form $\overline{W} = 2\pi\bar{k}$, and Eq. (2) in dimensionless form is recast as

$$\frac{\partial T}{\partial x} = \frac{\nabla^2 T}{\mathrm{Pe}} + \frac{W\delta(x)\delta(y)}{\mathrm{Pe}}. \tag{5}$$

Here, the averaging bar is omitted, and the dimension of the delta function, $\delta(x) = R\delta(x)$, is taken into account. Below, all formulas will be given in the dimensionless form unless otherwise stated.

SOLUTION OF THE ENERGY TRANSFER EQUATION FOR A LINEAR SOURCE

In the Fourier space [8], Eq. (5) is expressed as

$$\begin{cases} i\lambda T(\lambda, \omega) = -\dfrac{\lambda^2 + \omega^2}{\mathrm{Pe}} T(\lambda, \omega) + \dfrac{W}{2\pi\mathrm{Pe}} \\ T(\lambda, \omega) = \dfrac{1}{2\pi}\displaystyle\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \exp(-i\lambda x - i\omega y) T(x, y)\, dx dy. \end{cases} \tag{6}$$

For the temperature in the Fourier space, we obtain the expression

$$T(\lambda, \omega) = \frac{W}{2\pi(\lambda^2 + i\lambda\mathrm{Pe} + \omega^2)}. \tag{7}$$

The Fourier inverse of Eq. (7) is

$$T(x, y) = \frac{W}{4\pi^2}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \frac{\exp(i\lambda x + i\omega y)}{\lambda^2 + i\lambda\mathrm{Pe} + \omega^2} d\lambda d\omega. \tag{8}$$

Equation (8) can be integrated once over the variable λ. Consider the integral

$$I(x, \omega) = \int_{-\infty}^{\infty} \frac{\exp(i\lambda x)}{\lambda^2 + i\lambda\mathrm{Pe} + \omega^2} d\lambda. \tag{9}$$

The integrand satisfies the conditions of the Jordan lemma [9] in the upper half-plane at $x > 0$ and in the lower half-plane at $x < 0$. The roots of the integrand denominator are first-order poles given by

$$\lambda_{1,2} = -i\left\{\frac{\mathrm{Pe}}{2} \mp \sqrt{\frac{\mathrm{Pe}^2}{4} + \omega^2}\right\}. \tag{10}$$

The roots λ_1 and λ_2 are purely imaginary and lie in the lower and upper half-planes, respectively. Integral (9) is calculated by the known residue theorem [9]

$$I(x, \omega) = \begin{cases} 2\pi i\,\mathrm{res}(f(\lambda, \omega, x), \lambda_2), & x > 0 \\ -2\pi i\,\mathrm{res}(f(\lambda, \omega, x), \lambda_1), & x < 0, \end{cases} \tag{11}$$

$$f(\lambda, \omega, x) = \frac{\exp(i\lambda x)}{\lambda^2 + i\lambda\mathrm{Pe} + \omega^2}.$$

The final result is

$$T(x, y) = \frac{W}{2\pi}\int_0^{\infty} d\omega \frac{\cos(\omega y)}{\sqrt{\dfrac{\mathrm{Pe}^2}{4} + \omega^2}} \times \begin{cases} \exp\left(-x\left\{\sqrt{\dfrac{\mathrm{Pe}^2}{4} + \omega^2} - \dfrac{\mathrm{Pe}}{2}\right\}\right), & x > 0 \\ \exp\left(x\left\{\sqrt{\dfrac{\mathrm{Pe}^2}{4} + \omega^2} + \dfrac{\mathrm{Pe}}{2}\right\}\right), & x < 0. \end{cases} \tag{12}$$

It can be shown that integral (12) converges for any $\mathrm{Pe} > 0$.

CALCULATION OF THE HEAT-TRANSFER COEFFICIENT

The lines of equal ratios T/W for the Peclet numbers 0.1 and 0.5 are shown in Figs. 1 and 2. The calculations were carried out by Eq. (12). It is seen that, at $\mathrm{Pe} < 1$ and $r \le 1$, where $r = \sqrt{x^2 + y^2}$, the isotherms are nearly concentric circles centered at the origin. Thus, for low Peclet numbers, heat transfer in the domain $r < 1$ is due to heat conduction alone and becomes convective at $r > 1$. Hence, for $\mathrm{Pe} \ll 1$, the temperature on the surface $r = 1$ is almost constant and can be assigned to the wire temperature.

The distributions of the relative temperature on the surface $r = 1$ are shown in Fig. 3 for different Peclet numbers. At the points $x = 0$ and $y = \pm 1$, the temperature coincides with the mean value. Assuming that the mean wire surface temperature is $\langle T \rangle = 1$, we obtain the following expression for the heat-transfer coefficient:

$$\frac{1}{k} = \int_0^{\infty} \frac{\cos(\omega) d\omega}{\sqrt{\dfrac{\mathrm{Pe}^2}{4} + \omega^2}}. \tag{13}$$

The integrand in (13) is not an elementary function, so that integral (13) can be taken only numerically. For

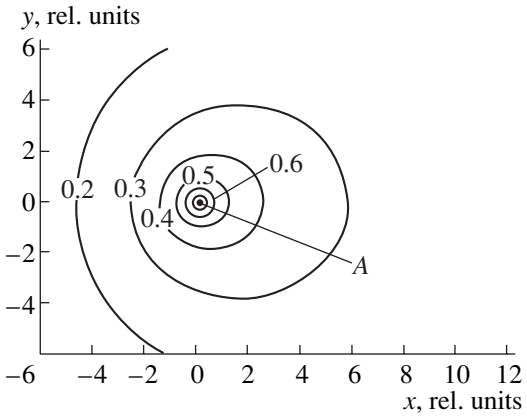

Fig. 1. Lines of equal ratio T/W for Pe = 0.1. A is the position of the heat source.

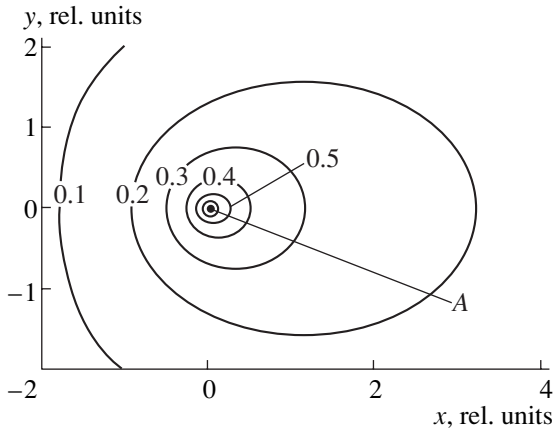

Fig. 2. The same as in Fig. 1 for Pe = 0.5.

Pe $\ll$ 1, Eq. (13) is reduced to

$$\frac{1}{k} \cong -\mathrm{Ei}\left(-\frac{\mathrm{Pe}}{2\sqrt{2}}\right), \quad \mathrm{Ei}(x) = \int_{-\infty}^{x} \frac{\exp(t)}{t} dt, \qquad (14)$$

where Ei is the integral exponent [10].

The calculations show that, in the range 0.1 < Pe < 1, the dependence of the heat-transfer coefficient on the Peclet number, Eq. (13), is well approximated by the linear dependence (in dimension and dimensionless units)

$$\begin{cases} k \cong a + b\mathrm{Pe} \\ k \cong a\mu R^{-1} + bc_p\rho_0 U, \end{cases} \qquad (15)$$

where a = 0.28 and b = 0.93 are the approximation parameters.

It follows from Eq. (15) that the dimensional heat-transfer coefficient consists of two terms, one of which depends on the wire radius and thermal conductivity of the gas and the other is proportional to the gas velocity and independent of the wire geometry. The dimensionless heat-transfer coefficient calculated as a function of the Peclet number by (13)–(15) is shown in Fig. 4.

ERROR ESTIMATE FOR A HEAT-TRANSFER COEFFICIENT

The inaccuracy of the model that is introduced by the replacement of a finite-size heat source by a linear heat source can be estimated by comparing the convective heat transfer through the domain occupied by the wire (r < 1) with the total energy released by the heat source. Obviously, the presence of the convective heat transfer through the domain r < 1 decreases the fraction of the conductive heat transfer in the overall process. In the case of a finite-size heat source, the convective transfer in the domain r < 1 is absent; accordingly, the wire surface temperature grows and the heat-transfer

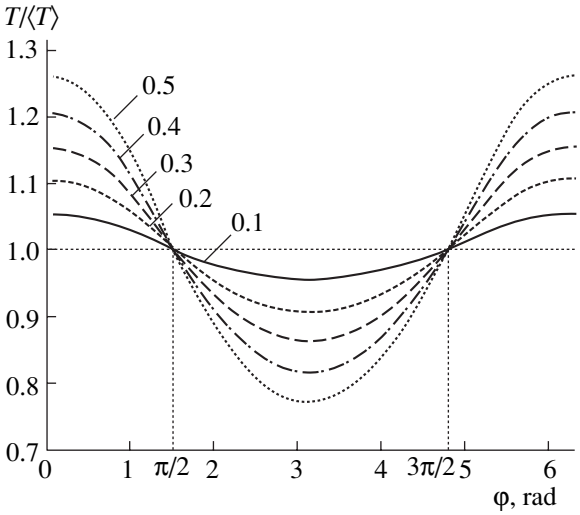

Fig. 3. Relative temperature on the surface r = 1 as a function of the polar angle φ for a linear heat source and different Peclet numbers (the figures by the curves).

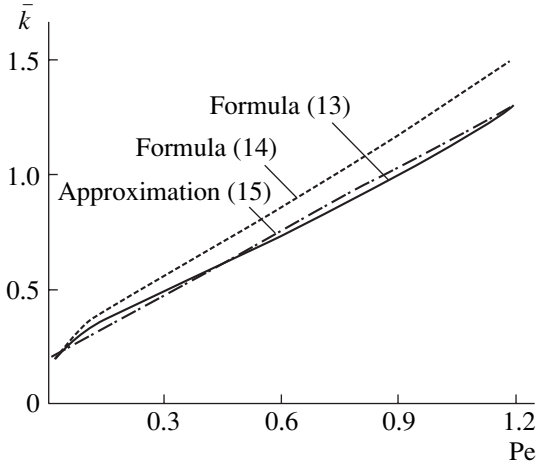

Fig. 4. Calculated dependences of the dimensionless heat-transfer coefficient on the Peclet number.

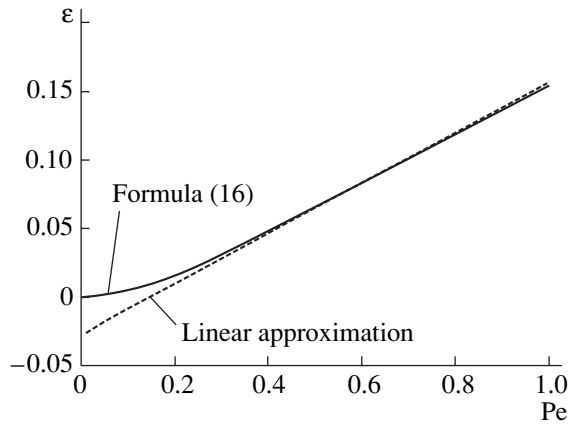

Fig. 5. Fraction of the convective heat flow vs. the Peclet number.

coefficient decreases in comparison with the case of a linear heat source. Thus, the heat-transfer coefficient given by Eq. (13) is overstated. Let us introduce a parameter ε, which is the ratio of the convective heat flow in the domain $r < 1$ to the total power of the heat source. Taking into account the harmonic dependence of the temperature of the surface $r = 1$ on the polar angle φ (Fig. 4), we find from Eq. (12)

$$\varepsilon = \frac{\mathrm{Pe}}{\pi}\sin h\left(\frac{\mathrm{Pe}}{2}\right)\int_0^\infty \frac{\exp\left\{-\sqrt{\frac{\mathrm{Pe}^2}{4}+\omega^2}\right\}}{\sqrt{\frac{\mathrm{Pe}^2}{4}+\omega^2}}d\omega. \qquad (16)$$

The value of ε may serve as a measure of the relative error of Eq. (13).

The results of calculation by Eq. (16) are shown in Fig. 5, where the dashed line is the linear approximation $\varepsilon(\mathrm{Pe}) \approx 0.18\mathrm{Pe} - 0.026$, which is valid for $0.3 < \mathrm{Pe} < 1$. For $\mathrm{Pe} = 1$, ε is about 15%; i.e., for the Peclet numbers less than unity, Eq. (13) for the heat-transfer coefficient is adequate.

CONCLUSIONS

In this paper, we theoretically studied the heat-transfer coefficient of an infinitely fine wire as a function of the gas velocity in the limit of zero temperature drop between the wire and gas. It was shown that a fine heated wire can be replaced with a good accuracy (about 15%) by a linear heat source when Peclet numbers are small ($\mathrm{Pe} < 1$). Such a replacement allows one to perform a comprehensive theoretical investigation of heat exchange between the wire and ambient gas.

For Peclet numbers in the range 0.1–1.0, the velocity dependence of the heat-transfer coefficient is linear. This agrees well with the results of experiments [1, 3], where the linear dependence of the heat-transfer coefficient was observed at much higher Peclet numbers.

REFERENCES

1. A. D. Mamrosov and S. V. Ivashchenko, Prib. Tekh. Éksp., No. 4, 242 (1981).
2. V. M. Kul'gavchuk and V. Yu. Mat'ev, in *Proceedings of the 2nd International Conference on Physics of Nuclear-Excited Plasma and Problems of Nuclear-Pumped Lasers, Arzamas, 1995,* Vol. 2, p. 93.
3. G. V. Vlokh, A. N. Korzenev, A. V. Lisenkov, *et al.*, in *Proceedings of the 5th International Conference "Atomic and Molecular Pulsed Lasers," Tomsk, 2001*, p. 19.
4. V. V. Borovkov, B. V. Lazhintsev, V. A. Nor-Arevyan, *et al.*, in *Proceedings of the 2nd International Conference on Physics of Nuclear-Excited Plasma and Problems of Nuclear-Pumped Lasers, Arzamas, 1995*, Vol. 1, p. 399.
5. G. V. Vlokh, A. I. Konak, V. Yu. Mat'ev, *et al.*, in *Proceedings of the Conference on Physics of Nuclear-Excited Plasma and Problems of Nuclear-Pumped Lasers, Obninsk, 1993*, Vol. 2, p. 55.
6. L. G. Loĭtsyanskiĭ, *Fluid and Gas Mechanics* (Nauka, Moscow, 1973).
7. A. M. Litvin, *Theoretical Principles of Heat Engineering* (Énergiya, Moscow, 1969).
8. A. N. Kolmogorov and S. V. Fomin, *Elements of the Theory of Functions and Functional Analysis* (Nauka, Moscow, 1989, 4th ed.; Graylock, Rochester, 1957).
9. A. M. Markushevich, *Brief Course of the Theory of Analytical Functions* (Nauka, Moscow, 1978).
10. A. P. Prudnikov, Yu. A. Brychkov, and O. I. Marichev, *Integrals and Series* (Nauka, Moscow, 1981; Gordon and Breach, New York, 1986).

Translated by M. Astrov

Technical Physics, Vol. 48, No. 6, 2003, pp. 697–707. Translated from Zhurnal Tekhnicheskoĭ Fiziki, Vol. 73, No. 6, 2003, pp. 36–45.

GASES AND LIQUIDS

On Nonlinear Vibrations of a Charged Drop in the Third-Order Approximation in the Amplitude of Initial Single-Mode Excitation

A. N. Zharov, S. O. Shiryaeva, and A. I. Grigor'ev

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

e-mail: shir@uniyar.ac.ru

Received December 6, 2002

Abstract—Nonlinear axisymmetric motions of the free surface of a charged drop of an ideal liquid under the single-mode initial deformation of its equilibrium shape is investigated in the third-order approximation in the initial perturbation amplitude. An analytical expression for the drop shape generatrix is derived. Nonlinear corrections to the vibration frequencies for the initial perturbation of an arbitrary mode are found for the first time. The effect of vibration nonlinearity on the instability of the drop against its self-charge is studied. © *2003 MAIK "Nauka/Interperiodica".*

(1) Nonlinear vibrations of a charged drop are of considerable interest in various areas of technical physics, instrument making, geophysics, and chemical engineering (see, for example, [1] and the references therein). Over the last two decades, this problem has been solved both theoretically and experimentally with a different degree of accuracy [2–10].

In the pioneering theoretical description of the nonlinear dynamics of the charged drop surface [2], expressions for the drop surface generatrix under the axisymmetric vibration of the drop were derived. Also, corrections to the frequencies of the capillary vibration of the drop surface were found in three particular cases when the initial perturbation of the surface contains only the second, third, or fourth mode.

An experimental investigation of the nonlinear dynamics of the charged surface of a free drop was carried out in [3], where the frequency shift due to nonlinearity was detected under conditions of microgravitation. The authors of [3] noted good agreement between the frequency shift found by them for not too large amplitudes and low viscosities of the drop, on the one hand, and the shift calculated in [2], on the other.

Early attempts to numerically analyze the phenomenon discussed [9, 10] have shown the high nonlinearity of processes accompanying the vibration of a charged drop.

The basic ideas put forward in [2] have been extended in [4, 5], where the surface dynamics of a liquid drop was studied in the second-order approximation in the amplitude of the initial surface perturbation described by an arbitrary mode of an mth order. In [4, 5], it was shown that the drop vibrates not about a sphere but about a body reminiscent of a spheroid of revolution, and that nonlinear interaction between modes in the second-order approximation excites all even modes from the range $[0; 2m]$.

Works [2, 4, 5] have been generalized in [6–8], where the nonlinear dynamics of the surface of a liquid drop is analyzed in the second-order approximation in the initial perturbation for the case when the initial perturbation is represented as the superposition of several modes. It was ascertained that (i) mode interaction causes internal nonlinear resonances and may be responsible for the loss of drop stability [6] and (ii) the initial excitation of two adjacent modes results in the excitation of the first-order (translational) mode [7, 8] with the transformation of a nonlinearly vibrating drop into a source of dipole electromagnetic and acoustic radiation.

Nevertheless, the nonlinear dynamics of the free surface of a charged drop needs further theoretical investigation. Since analytical calculations for an arbitrary form of the initial perturbation were made only in the second order of smallness in its amplitude [4–8], the form of nonlinear corrections to vibration frequencies remained unknown.

Therefore, it is interesting to consider the nonlinear dynamics of the surface of a charged drop in the third order of smallness in the amplitude of an initial perturbation that is proportional to the Legendre polynomial of degree m.

(2) Consider a drop of a perfectly conducting liquid of radius R, density ρ, and surface tension σ. The drop vibrates in a vacuum and carries a charge Q. The liquid motion inside the drop is assumed to be potential with a velocity potential ψ. The electric field potential near the drop is ϕ. We also assume that the drop remains axisymmetric at any instant of time and seek a solution to the problem in the spherical coordinate system (r, ϑ, ϕ)

1063-7842/03/4806-0697$24.00

related to the center of mass of the drop. The equation for the drop surface in dimensionless variables such that $\rho = 1$, $R = 1$, and $\sigma = 1$ is written in the form

$$F(r, \vartheta, t) = r - 1 - \xi(\vartheta, t) = 0. \quad (1)$$

The mathematical statement of the problem of nonlinear capillary vibrations of the charged drop surface [8, 11, 12] contains the Laplace equations for the potentials of the liquid velocity and electric field

$$\Delta\psi = 0, \quad \Delta\phi = 0 \quad (2)$$

(Δ is the Laplacian); the boundedness conditions

$$r \longrightarrow 0: \quad \psi \longrightarrow 0, \quad (3)$$

$$r \longrightarrow +\infty: \quad \Delta\phi \longrightarrow 0; \quad (4)$$

the kinematic and dynamic boundary conditions

$$r = 1 + \xi(\vartheta, t): \quad \frac{dF}{dt} = 0, \quad (5)$$

$$r = 1 + \xi(\vartheta, t):$$
$$\partial_t\psi + \frac{1}{2}(\nabla\psi)^2 = p + p_q - p_{\text{atm}} - p_\sigma; \quad (6)$$

the constancy condition for the drop volume

$$\int_V r^2 \sin\vartheta dr d\vartheta d\varphi = \frac{4\pi}{3}, \quad (7)$$
$$V = \{r, \vartheta, \varphi | 0 \le r \le 1 + \xi; 0 \le \vartheta \le \pi; 0 \le \varphi \le 2\pi\};$$

the immobility condition for the center of mass

$$\int_V \mathbf{r} r^2 \sin\vartheta dr d\vartheta d\varphi = 0; \quad (8)$$

the constancy condition for the total charge

$$r = 1 + \xi(\vartheta, t): \quad \int_V \mathbf{n} \cdot \nabla\phi dS = -4\pi Q, \quad (9)$$
$$S = \{r, \vartheta, \varphi | r = 1 + \xi; 0 \le \vartheta \le \pi; 0 \le \varphi \le 2\pi\};$$

the constancy condition for the electric potential over the drop surface

$$r = 1 + \xi(\vartheta, t): \quad \phi = \phi_S(t); \quad (10)$$

and the initial conditions

$$t = 0: \ \xi = \xi_0 P_0(\cos\vartheta) + \xi_1 P_1(\cos\vartheta) + \varepsilon P_m(\cos\vartheta),$$
$$\partial_t\xi = 0. \quad (11)$$

In expressions (2)–(11), p, p_q, p_{atm}, and p_σ are the hydrodynamic pressure, electric field pressure, atmospheric pressure, and capillary pressure, respectively; **n** is the unit vector normal to the drop surface; ϕ_S is the electric potential of the drop surface; ε is a small parameter characterizing the amount of the initial deformation; $P_m(\cos\vartheta)$ is the Legendre polynomial; and ξ_0 and ξ_1 are quantities defined in such a way that conditions (7) and (8) are fulfilled at zero time. The symbol ∂_t denotes the partial derivative with respect to the variable t.

(3) To solve problem (1)–(11), we apply the method of multiple scales [13] and introduce three time scales $T_k = \varepsilon^k t$ (where $k = 0, 1, 2$). Then, the desired quantities are represented in the form of the expansions

$$\phi = \phi^{(0)} + \varepsilon\phi^{(1)} + \varepsilon^2\phi^{(2)} + \varepsilon^3\phi^{(3)} + O(\varepsilon^4), \quad (12)$$

$$\phi_S = \phi_S^{(0)} + \varepsilon\phi_S^{(1)} + \varepsilon^2\phi_S^{(2)} + \varepsilon^3\phi_S^{(3)} + O(\varepsilon^4), \quad (13)$$

$$\psi = \varepsilon\psi^{(1)} + \varepsilon^2\psi^{(2)} + \varepsilon^3\psi^{(3)} + O(\varepsilon^4), \quad (14)$$

$$\xi = \varepsilon\xi^{(1)} + \varepsilon^2\xi^{(2)} + \varepsilon^3\xi^{(3)} + O(\varepsilon^4), \quad (15)$$

where $\phi^{(0)} = Q/r$ and $\phi_S^{(0)} = Q$ are the zero-order solutions for the equilibrium drop surface.

Substituting (12)–(15) into (1)–(11), we obtain problems of different orders of smallness. The first-order problem has the form

$$\Delta\psi^{(1)} = 0, \quad \Delta\phi^{(1)} = 0; \quad (16)$$

$$r \longrightarrow 0: \quad \psi^{(1)} \longrightarrow 0; \quad (17)$$

$$r \longrightarrow +\infty: \quad \nabla\phi^{(1)} \longrightarrow 0; \quad (18)$$

$$r = 1: \quad \partial_{T_0}\xi^{(1)} = \partial_r\psi^{(1)}; \quad (19)$$

$$r = 1:$$
$$\partial_{T_0}\psi^{(1)} = \frac{1}{4\pi}\partial_r\phi^{(0)}(\partial_r\phi^{(1)} + \xi^{(1)}\partial_{rr}\phi^{(0)}) + 2\xi^{(1)} + \Delta_\Omega\xi^{(1)}; \quad (20)$$

$$\int_{-1}^{1} \xi^{(1)} d(\cos\vartheta) = 0, \quad \int_{-1}^{1} \xi^{(1)} P_1 d(\cos\vartheta) = 0; \quad (21)$$

$$r = 1:$$
$$\int_{-1}^{1} \{\partial_r\phi^{(1)} + \xi^{(1)}(\partial_{rr}\phi^{(0)} + 2\partial_r\phi^{(0)})\} d(\cos\vartheta) = 0; \quad (22)$$

$$r = 1: \quad \phi^{(1)} + \xi^{(1)}\partial_r\phi^{(0)} = \phi_S^{(1)}(t); \quad (23)$$

$$t = 0: \quad \xi^{(1)} = P_m(\cos\vartheta), \quad \partial_{T_0}\xi^{(1)} = 0. \quad (24)$$

In the second order of smallness, we have the boundary-value problem in the form

$$\Delta\psi^{(2)} = 0, \quad \Delta\phi^{(2)} = 0; \quad (25)$$

$$r \longrightarrow 0: \quad \psi^{(2)} \longrightarrow 0; \quad (26)$$

$$r \longrightarrow +\infty: \quad \nabla\phi^{(2)} \longrightarrow 0; \quad (27)$$

$$r = 1:$$
$$\partial_{T_0}\xi^{(2)} + \partial_{T_1}\xi^{(1)} = \partial_r\psi^{(2)} + \xi^{(1)}\partial_{rr}\psi^{(1)} - \partial_\vartheta\xi^{(1)}\partial_\vartheta\psi^{(1)}; \tag{28}$$

$$r = 1:$$
$$\partial_{T_0}\psi^{(2)} + \partial_{T_1}\psi^{(1)} + \xi^{(1)}\partial_{rT_0}\psi^{(1)} + \frac{1}{2}(\partial_r\psi^{(1)})^2 + \frac{1}{2}(\partial_\vartheta\psi^{(1)})^2 = \frac{1}{8\pi}\{2\xi^{(2)}\partial_r\phi^{(0)}\partial_{rr}\phi^{(0)} + (\xi^{(1)})^2((\partial_{rr}\phi^{(0)})^2 + \partial_{rrr}\phi^{(0)}\partial_r\phi^{(0)}) + (\partial_\vartheta\phi^{(1)})^2 + (\partial_r\phi^{(1)})^2 + 2\partial_r\phi^{(2)}\partial_r\phi^{(0)} + 2\xi^{(1)}(\partial_{rr}\phi^{(0)}\partial_r\phi^{(1)} + \partial_{rr}\phi^{(1)}\partial_r\phi^{(0)})\} + 2\xi^{(2)} + \Delta_\Omega\xi^{(2)} - 2(\xi^{(1)})^2 - 2\xi^{(1)}\Delta_\Omega\xi^{(1)}; \tag{29}$$

$$\int_{-1}^{1}(\xi^{(2)} + (\xi^{(1)})^2)d(\cos\vartheta) = 0; \tag{30}$$

$$\int_{-1}^{1}(2\xi^{(2)} + 3(\xi^{(1)})^2)P_1 d(\cos\vartheta) = 0; \tag{31}$$

$$r = 1:$$
$$\int_{-1}^{1}\left\{\partial_r\phi^{(2)} + \xi^{(1)}(\partial_{rr}\phi^{(1)} + 2\partial_r\phi^{(1)}) + \xi^{(2)}(\partial_{rr}\phi^{(0)} + 2\partial_r\phi^{(0)}) + (\xi^{(1)})^2\left(\frac{1}{2}\partial_{rrr}\phi^{(0)} + 2\partial_{rr}\phi^{(0)} + \partial_r\phi^{(0)}\right) - \partial_\vartheta\xi^{(1)}\partial_\vartheta\phi^{(1)}\right\}d(\cos\vartheta) = 0; \tag{32}$$

$$r = 1:$$
$$\phi^{(2)} + \xi^{(1)}\partial_r\phi^{(1)} + \xi^{(2)}\partial_r\phi^{(0)} + \frac{1}{2}(\xi^{(1)})^2\partial_{rr}\phi^{(0)} = \phi_S^{(2)}(t); \tag{33}$$

$$t = 0: \quad \xi^{(2)} = -\frac{1}{2m+1}P_0(\cos\vartheta); \quad \partial_{T_0}\xi^{(2)} + \partial_{T_1}\xi^{(1)} = 0. \tag{34}$$

In the third order of smallness, the problem is stated as follows (see Appendices A and B):

$$\Delta\psi^{(3)} = 0; \quad \Delta\phi^{(3)} = 0; \tag{35}$$

$$r \longrightarrow 0: \quad \psi^{(3)} \longrightarrow 0; \tag{36}$$

$$r \longrightarrow +\infty: \quad \nabla\phi^{(3)} \longrightarrow 0; \tag{37}$$

$$r = 1:$$
$$\partial_{T_0}\xi^{(3)} + \partial_{T_1}\xi^{(2)} + \partial_{T_2}\xi^{(1)} = \partial_r\psi^{(3)} - \partial_\vartheta\xi^{(2)}\partial_\vartheta\psi^{(1)} - \partial_\vartheta\xi^{(1)}\partial_\vartheta\psi^{(2)} + \xi^{(2)}\partial_{rr}\psi^{(1)} + \xi^{(1)}(\partial_\vartheta\xi^{(1)}(2\partial_\vartheta\psi^{(1)} - \partial_{r\vartheta}\psi^{(1)}) + \partial_{rr}\psi^{(2)}) + \frac{1}{2}(\xi^{(1)})^2\partial_{rrr}\psi^{(1)}; \tag{38}$$

$$r = 1:$$
$$\partial_{T_0}\psi^{(3)} + \partial_{T_2}\psi^{(1)} + \partial_{T_1}\psi^{(2)} + \xi^{(1)}\partial_{rT_1}\psi^{(1)} + \partial_\vartheta\psi^{(1)}\partial_\vartheta\psi^{(2)} + \partial_r\psi^{(1)}\partial_r\psi^{(2)} + \xi^{(2)}\partial_{rT_0}\psi^{(1)} + \xi^{(1)}(\partial_{rT_0}\psi^{(2)} + \partial_\vartheta\psi^{(1)}(\partial_{r\vartheta}\psi^{(1)} - \partial_\vartheta\psi^{(1)}) + \partial_r\psi^{(1)}\partial_{rr}\psi^{(1)}) + \frac{1}{2}(\xi^{(1)})^2\partial_{rrT_0}\psi^{(1)}$$
$$= \frac{1}{8\pi}\left\{2\xi^{(3)}\partial_r\phi^{(0)}\partial_{rr}\phi^{(0)} + (\xi^{(1)})^3\left(\partial_{rr}\phi^{(0)}\partial_{rrr}\phi^{(0)} + \frac{1}{3}\partial_r\phi^{(0)}\partial_{rrrr}\phi^{(0)}\right) + 2(\partial_\vartheta\phi^{(1)}\partial_\vartheta\phi^{(2)} + \partial_r\phi^{(1)} \times (\xi^{(2)}\partial_{rr}\phi^{(0)} + \partial_r\phi^{(2)}) + \partial_r\phi^{(0)}\partial_r\phi^{(3)} + \xi^{(2)}\partial_r\phi^{(0)}\partial_{rr}\phi^{(1)}) + 2\xi^{(1)}(\xi^{(2)}((\partial_{rr}\phi^{(0)})^2 + \partial_r\phi^{(0)}\partial_{rrr}\phi^{(0)}) + \partial_{rr}\phi^{(0)}\partial_r\phi^{(2)} + \partial_\vartheta\phi^{(1)}(\partial_{r\vartheta}\phi^{(1)} - \partial_\vartheta\phi^{(1)}) + \partial_r\phi^{(1)}\partial_{rr}\phi^{(1)} + \partial_r\phi^{(0)}\partial_{rr}\phi^{(2)}) + (\xi^{(1)})^2 \times (\partial_{rrr}\phi^{(0)}\partial_r\phi^{(1)} + 2\partial_{rr}\phi^{(0)}\partial_{rr}\phi^{(1)} + \partial_r\phi^{(0)}\partial_{rrr}\phi^{(1)})\right\} + (2 + \Delta_\Omega)\xi^{(3)} + 2\xi^{(1)}((\xi^{(1)})^2 - (2 + \Delta_\Omega)\xi^{(2)}) - 2\xi^{(2)}\Delta_\Omega\xi^{(1)} + 3(\xi^{(1)})^2\Delta_\Omega\xi^{(1)} - (\partial_\vartheta\xi^{(1)})^2\partial_{\vartheta\vartheta}\xi^{(1)} - \frac{1}{2}(\partial_\vartheta\xi^{(1)})^2\Delta_\Omega\xi^{(1)}; \tag{39}$$

$$\int_{-1}^{1}(3\xi^{(3)} + 6\xi^{(1)}\xi^{(2)} + (\xi^{(1)})^3)d(\cos\vartheta) = 0; \tag{40}$$

$$\int_{-1}^{1}(\xi^{(3)} + 3\xi^{(1)}\xi^{(2)} + (\xi^{(1)})^3)P_1(\cos\vartheta)d(\cos\vartheta) = 0; \quad (41)$$

$$r = 1:$$

$$\int_{-1}^{1}\{\partial_r\phi^{(3)} + \xi^{(3)}(\partial_{rr}\phi^{(0)} + 2\partial_r\phi^{(0)}) + \xi^{(2)}(\partial_{rr}\phi^{(1)} + 2\partial_r\phi^{(1)}) + (\xi^{(1)})^3\left(\frac{1}{6}\partial_{rrrr}\phi^{(0)} + \partial_{rrr}\phi^{(0)} + \partial_{rr}\phi^{(0)}\right) + (\xi^{(1)})^2\left(\frac{1}{2}\partial_{rrr}\phi^{(1)} + 2\partial_{rr}\phi^{(1)} + \partial_r\phi^{(1)}\right) + \xi^{(1)}(\xi^{(2)}(\partial_{rrr}\phi^{(0)} + 4\partial_{rr}\phi^{(0)} + 2\partial_r\phi^{(0)}) + 2\partial_r\phi^{(2)} + \partial_{rr}\phi^{(2)} - \partial_\vartheta\xi^{(1)}\partial_{r\vartheta}\phi^{(1)}) - \partial_\vartheta\xi^{(2)}\partial_\vartheta\phi^{(1)} - \partial_\vartheta\xi^{(1)}\partial_\vartheta\phi^{(2)}\}d(\cos\vartheta) = 0; \quad (42)$$

$$r = 1:$$

$$\phi^{(3)} + \xi^{(1)}\partial_r\phi^{(2)} + \xi^{(2)}\partial_r\phi^{(1)} + \xi^{(3)}\partial_r\phi^{(0)} + \frac{1}{2}(\xi^{(1)})^2\partial_{rr}\phi^{(1)} + \xi^{(1)}\xi^{(2)}\partial_{rr}\phi^{(0)} + \frac{1}{6}(\xi^{(1)})^3\partial_{rrr}\phi^{(0)} = \phi_S^{(3)}(t); \quad (43)$$

$$t = 0:$$

$$\xi^{(3)} = -\frac{K_{mmm}}{3(2m+1)}P_0(\cos\vartheta) - \sum_{g=0}^{\infty}K_{mmg}K_{gm1}P_1(\cos\vartheta); \quad (44)$$

$$t = 0: \quad \partial_{T_0}\xi^{(3)} + \partial_{T_1}\xi^{(2)} + \partial_{T_2}\xi^{(1)} = 0.$$

Here, $K_{mln} = (C_{m0l0}^{n0})^2$ and C_{m0l0}^{n0} are the Clebsch–Gordan coefficients [14].

(4) Solutions to the Laplace equations for the potentials of the velocity field and electric field ((16), (25), and (35)) that satisfy boundedness conditions (17), (18), (26), (27), (36), and (37) can be written in the form

$$\psi^{(k)} = \sum_{n=1}^{\infty}r^n D_n^{(k)}(t)P_n(\cos\vartheta); \quad k = 1, 2, 3; \quad (45)$$

$$\phi^{(k)} = \sum_{n=0}^{\infty}\frac{F_n^{(k)}(t)}{r^{n+1}}P_n(\cos\vartheta); \quad k = 1, 2, 3. \quad (46)$$

Note that summation in Eq. (45) starts from $n = 1$, because, as is known, the potential is defined accurate to an arbitrary function of time. Therefore, one may set $D_0^{(k)} = 0$.

A function describing the deviation of the surface from the spherical shape can be represented as the expansion in Legendre polinomials:

$$\xi^{(k)} = \sum_{n=0}^{\infty}M_n^{(k)}(t)P_n(\cos\vartheta); \quad k = 1, 2, 3. \quad (47)$$

Substituting expansions (45)–(47) into sets (19)–(24), (28)–(34), and (38)–(44) gives expressions for the coefficients in expansions (45)–(47).

A solution to first-order problem (16)–(24) allows one to find the dependence of the first-order coefficients $M_n^{(1)}(t)$, $D_n^{(1)}(t)$, and $F_n^{(1)}(t)$ on the fast time scale $T_0 = t$:

$$M_n^{(1)}(t) = \delta_{mn}\cos(\omega_m T_0), \quad (48)$$

$$D_n^{(1)} = \frac{1}{n}\partial_{T_0}M_n^{(1)}; \quad F_n^{(1)} = QM_n^{(1)}. \quad (49)$$

In (48), m is the number of an initially excited mode, $\omega_n = \sqrt{n(n-1)(n+2-W)}$ is the frequency of an nth mode of the capillary vibrations of the drop, and $W = Q^2/(4\pi)$ is the Rayleigh parameter.

Analysis of problem (16)–(34) clarifies the dependences of the coefficients $M_n^{(1)}(t)$, $D_n^{(1)}(t)$, and $F_n^{(1)}(t)$ on the time $T_1 = \varepsilon t$, and of the coefficients $M_n^{(2)}(t)$, $D_n^{(2)}(t)$, and $F_n^{(2)}(t)$ on the time $T_0 = t$.

The functions $M_n^{(1)}(t)$, $D_n^{(1)}(t)$, and $F_n^{(1)}(t)$ turned out to be independent of the time scale T_1, so that expressions (48)–(49) also remain valid in the second order of smallness. Then, the coefficients $M_n^{(2)}(t)$, $D_n^{(2)}(t)$, and $F_n^{(2)}(t)$ have the form

$$M_0^{(2)} = -\frac{1}{2m+1}\cos^2(\omega_m T_0); \quad M_1^{(2)} = 0;$$
$$M_n^{(2)} = \frac{1}{2}[\lambda_{mmn}^{(+)}\cos(2\omega_m T_0) + \lambda_{mmn}^{(-)} - (\lambda_{mmn}^{(+)} + \lambda_{mmn}^{(-)})\cos(\omega_n T_0)]; \quad (50)$$
$$n \geq 2;$$

$$F_0^{(2)} = 0; \quad F_n^{(2)} = QM_n^{(2)} + QmK_{mmn}\cos^2(\omega_m T_0); \quad n \geq 1;$$

$$D_n^{(2)} = \frac{1}{n}\left(\partial_{T_0} M_n^{(2)} + (m(m-1)K_{mmn} - \alpha_{mmn}) \times \frac{\omega_m}{m}\sin(\omega_m T_0)\cos(\omega_m T_0)\right); \quad n \geq 1, \tag{51}$$

where

$$\lambda_{mln}^{(\pm)} = (\gamma_{mln} \pm \omega_m\omega_l\eta_{mln})/(\omega_n^2 - (\omega_m \pm \omega_l)^2),$$

$$\alpha_{mln} = -C_{m0l0}^{n0} C_{m(-1)l1}^{n0}\sqrt{m(m+1)l(l+1)},$$

$$\gamma_{mln} = K_{mln}[\omega_m^2(n-m+1) + 2n(l(l+1)-1) + (l(m+1) - m(2m-2n+7)+3)nW/2] + \alpha_{mln}[\omega_m^2/m + nW/2],$$

$$\eta_{mln} = K_{mln}(n/2 - m + 1) + \alpha_{mln}(1 + n/(2l))/m.$$

Investigating problem (16)–(44) in the third order of smallness makes it possible to find the dependences of the coefficients $M_n^{(1)}(t)$, $D_n^{(1)}(t)$, and $F_n^{(1)}(t)$ on the time $T_2 = \varepsilon^2 t$; the dependences of the coefficients $M_n^{(2)}(t)$, $D_n^{(2)}(t)$, and $F_n^{(2)}(t)$ on the time $T_1 = \varepsilon t$; and those of the coefficients $M_n^{(3)}(t)$, $D_n^{(3)}(t)$, and $F_n^{(3)}(t)$ on the time $T_0 = t$. The coefficients $M_n^{(2)}(t)$, $D_n^{(2)}(t)$, and $F_n^{(2)}(t)$ turned out to be independent of the time scale T_1, so that expressions (50)–(51) hold.

The variation of the coefficients $M_n^{(1)}$ with time is now characterized by two time scales T_0 and T_2:

$$M_n^{(1)} = \delta_{mn}\cos(\omega_n T_0 + b_n T_2), \tag{52}$$

where

$$b_n = \frac{1}{8\omega_n}\left\{\frac{1}{(2n+1)}(2(n-1)\omega_n^2 + 3\Xi_n) - [H_{nnnn}^{1(-)(+)} + H_{nnnn}^{2(-)(-)} + H_{nnnn}^{2(+)(+)}]\right\}. \tag{53}$$

The coefficients $H_{nnnn}^{1(-)(+)}$, $H_{nnnn}^{2(-)(-)}$, $H_{nnnn}^{2(+)(+)}$, and Ξ_n are defined in Appendix C.

The expressions for $D_n^{(1)}$ and $F_n^{(1)}$ can be obtained from formulas (49), and the coefficients $M_n^{(3)}(t)$, $D_n^{(3)}(t)$, and $F_n^{(3)}(t)$ have the form

$$M_0^{(3)} = -\frac{2}{2m+1}M_m^{(2)}\cos(\omega_m T_0) - \frac{K_{mmm}}{3(2m+1)}\cos^3(\omega_m T_0);$$

$$M_1^{(3)} = -3\sum_{k=0}^{\infty} K_{km1}M_k^{(2)}\cos(\omega_m T_0) - \sum_{g=0}^{\infty} K_{mmg}K_{gm1}\cos^3(\omega_m T_0);$$

$$M_n^{(3)} = -\frac{\delta_{nm}}{32\omega_m^2(2m+1)}(2(m-1)\omega_m^2 - \Xi_m) \times (\cos(3\omega_m T_0) - \cos(\omega_m T_0)) - \sum_{g=2}^{\infty}\frac{H_{mgn}^{0(+)}}{4(\omega_n^2 - (\omega_m + \omega_g)^2)} \times (\cos((\omega_m + \omega_g)T_0) - \cos(\omega_n T_0)) - \sum_{g=2}^{\infty}\frac{H_{mgn}^{0(-)}}{4(\omega_n^2 - (\omega_m - \omega_g)^2)} \times (\cos((\omega_m - \omega_g)T_0) - \cos(\omega_n T_0)) + \frac{1}{4}\frac{H_{mmmn}^{1(+)(-)}}{\omega_n^2 - 9\omega_m^2}(\cos(3\omega_m T_0) - \cos(\omega_n T_0)) + \frac{1-\delta_{mn}}{4(\omega_n^2 - \omega_m^2)}(H_{mmmn}^{1(-)(+)} + H_{mmmn}^{2(+)(+)} + H_{mmmn}^{2(-)(-)}) \times (\cos(\omega_m T_0) - \cos(\omega_n T_0)), \quad n \geq 2; \tag{54}$$

$$F_0^{(3)} = Q\frac{m+1}{2m+1} \times \left(\alpha_{mmm} - \frac{m(m+1)}{2}K_{mmm}\right)\cos^3(\omega_m T_0);$$

$$F_n^{(3)} = QM_n^{(3)} + \sum_{k=1}^{\infty}(k+1)K_{kmn}F_k^{(2)}\cos(\omega_m T_0) + Q\sum_{k=0}^{\infty}(m-1)K_{mkn}M_k^{(2)}\cos(\omega_m T_0) - Q\sum_{g=0}^{\infty}\frac{m(m+3)}{2}K_{mmg}K_{gmn}\cos^3(\omega_m T_0),$$

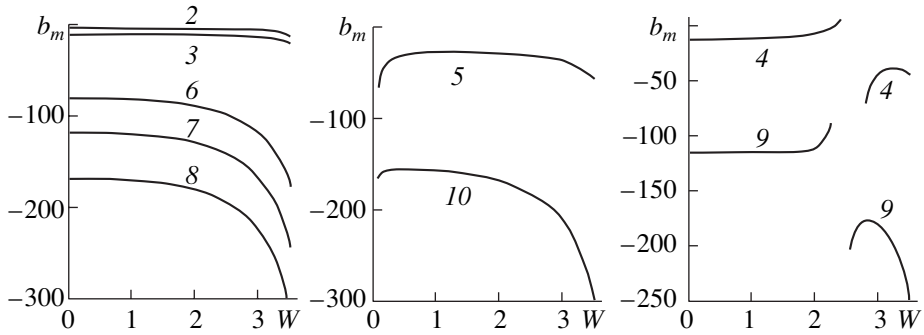

Fig. 1. Coefficient b_m as a function of the Rayleigh parameter $W = Q^2/4\pi$. The curve numbers coincide with the numbers of the initially excited modes.

$$n \geq 1;$$

$$D_n^{(3)} = \frac{1}{n}\partial_{T_0} M_n^{(3)} - \frac{1-\delta_{n1}}{n}\delta_{mn} b_m \sin(\omega_m T_0) \tag{55}$$

$$-\frac{1}{n}\sum_{k=1}^{\infty}(k(k-1)K_{kmn} - \alpha_{kmn})D_k^{(2)}\cos(\omega_m T_0)$$

$$+\frac{\omega_m}{mn}\sum_{k=0}^{\infty}(m(m-1)K_{mkn} - \alpha_{mkn})M_k^{(2)}\sin(\omega_m T_0)$$

$$+\frac{\omega_m}{mn}\sum_{g=0}^{\infty}\left(\frac{m(m-1)(m-2)}{2}K_{mmg} - (m-2)\alpha_{mmg}\right)$$

$$\times K_{gmn}\sin(\omega_m T_0)\cos^2(\omega_m T_0),$$

$$n \geq 1.$$

Substituting expressions (15) and (47) into (1) yields an expression for the generatrix of the shape of a nonlinearly vibrating drop:

$$r(\vartheta, t) = 1 + \varepsilon\sum_{n=2}^{\infty} M_n^{(1)}(t)P_n(\cos\vartheta) \tag{56}$$

$$+\varepsilon^2\sum_{n=0}^{\infty}(M_n^{(2)}(t) + \varepsilon M_n^{(3)}(t))P_n(\cos\vartheta),$$

where the coefficients $M_n^{(k)}$ ($k = 1, 2, 3$) are defined by expressions (50), (52), and (54).

(5) Putting $t = 0$ in (56), we obtain an expression for the drop shape at the initial time instant (see (11)) with the corrections ξ_0 and ξ_1, which result from the constancy of the drop's volume (condition (7)) and the immobility of the drop's center of mass (condition (8)):

$$r(\vartheta) = 1 + \varepsilon P_m(\cos\vartheta) - \frac{\varepsilon^2}{2m+1} \tag{57}$$

$$-\varepsilon^3\left(\frac{K_{mmm}}{3(2m+1)}P_0(\cos\vartheta) + P_1(\cos\vartheta)\sum_{g=m-1}^{m+1}K_{mmg}K_{gm1}\right).$$

From expression (57), it follows that second- and third-order corrections to the zeroth mode arise at zero time, keeping the drop volume constant. Besides, the immobility condition for the center of mass yields a third-order correction to the first-order (translational) mode, whereas in the problem of the second order of smallness, such a correction was absent [5].

For the excitation of the second, third, and fourth modes ($m = 2, 3, 4$) (this situation was analyzed in [2]), expression (57) takes the form

$$r(\vartheta) = 1 + \varepsilon P_2(\cos\vartheta) - \frac{1}{5}\varepsilon^2 - \frac{2}{105}\varepsilon^3 + O(\varepsilon^4)$$

$$\text{at} \quad m = 2;$$

$$r(\vartheta) = 1 + \varepsilon P_3(\cos\vartheta) - \frac{1}{7}\varepsilon^2$$

$$-\frac{36}{385}\varepsilon^3 P_1(\cos\vartheta) + O(\varepsilon^4) \tag{58}$$

$$\text{at} \quad m = 3;$$

$$r(\vartheta) = 1 + \varepsilon P_4(\cos\vartheta) - \frac{1}{9}\varepsilon^2 - \frac{6}{1001}\varepsilon^3 + O(\varepsilon^4)$$

$$\text{at} \quad m = 4.$$

It is easy to see that, according to (57) and (58), the third-order corrections to the initial conditions are different for even and odd modes: for even modes, the third-order correction follows from the volume constancy condition and makes a contribution to ξ_0, while the initial excitation of an odd mode does not affect the amplitude of the mode $n = 0$ in the third order of smallness but gives a third-order correction to the first-order (translational) mode ($n = 1$) because of the immobility condition for the center of mass. Note that in [2], where corrections to the zeroth mode were calculated for the first time, the immobility condition for the center of mass was not considered; therefore, a correction to the translational mode was not found.

From expressions (52) and (53), it is seen that taking into account quantities of the third order of smallness allows one to find the correction $\varepsilon^2 b_m$ to the frequency

of an *m*th mode, this correction depending on the drop charge (Fig. 1). Note that the general expression found for a correction to the frequency coincides with the results in [2] for the particular cases of the second, third, and fourth modes ($m = 2, 3, 4$). The discontinuities of the functions $b_4(W)$ and $b_9(W)$ (where $W = Q^2/4\pi$) are associated with internal nonlinear degenerate resonances, which arise at those values of the parameter W below the Rayleigh stability limit when the relationships $2\omega_4 = \omega_6$ and $2\omega_9 = \omega_{14}$ are fulfilled. With W in the vicinity of the resonances, expressions (53) become invalid. In this situation, the resonance conditions demand special consideration (see, for example, [15]).

The presence of the corrections to the frequencies changes the critical conditions for instability of an *m*th mode against the self-charge of the drop. As the charge builds up, the square of the frequency decreases and vanishes at a certain critical (for a given mode) value. As the charge increases further, the frequency becomes imaginary, which corresponds to the exponential growth of the amplitude of the *m*th mode with time; that is, this mode becomes unstable [16]. The critical condition for the onset of instability of the *m*th mode in view of a nonlinear correction can be written in the form

$$(\omega_m + \varepsilon^2 b_m)^2 \cong \omega_m^2 + 2\varepsilon^2 \omega_m b_m + O(\varepsilon^4) = 0.$$

The stability of the drop as a whole is defined by the critical conditions of instability for the most easily excited mode, which is the fundamental mode ($m = 2$). Figure 2 shows the square of the frequency as a function of the parameter W (characterizing the amount of the drop charge) with (curve *1*) and without (curve *2*) the corrections obtained. The tendency toward lowering the critical value of the parameter W is clearly seen. For example, at the amplitude of the initial perturbation of the principal mode $\varepsilon = 0.1$, numerical estimations give the critical value $W_{cr} = 3.85$ vs. $W_{cr} = 3.54$ at $\varepsilon = 0.3$ (recall that the Rayleigh limit for drop stability corresponds to $W_{cr} = 4$ [17]). However, these numerical estimates are only qualitative, because, with W close to W_{cr}, the small correction $\varepsilon^2 b_m$ becomes comparable to the frequency itself; that is, the expansion used loses its uniformity.

From our analysis, it follows that nonlinear mode interaction in the problem of the third order of smallness in the initial perturbation amplitude results in the broadening of the spectrum of modes responsible for shaping the free drop surface. In particular, when the number m of an initially excited mode is odd, only one mode with the number m is excited in the first order of smallness. In the second order of smallness, the zeroth mode and all even modes with numbers up to $2m$ inclusive are excited. Finally, in the third order of smallness, all odd modes with numbers up to $3m$ inclusive are excited. Thus, the drop surface is shaped by all modes with numbers from the range $0, \ldots, 2m$ and by odd modes with numbers from $2m + 1$ to $3m$. When the

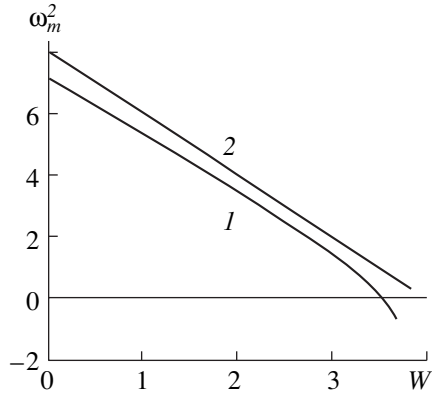

Fig. 2. Square of the frequency of the principal mode ($m = 2$) as a function of the Rayleigh parameter (*1*) with and (*2*) without the correction $\varepsilon^2 b_m$ at $\varepsilon = 0.3$.

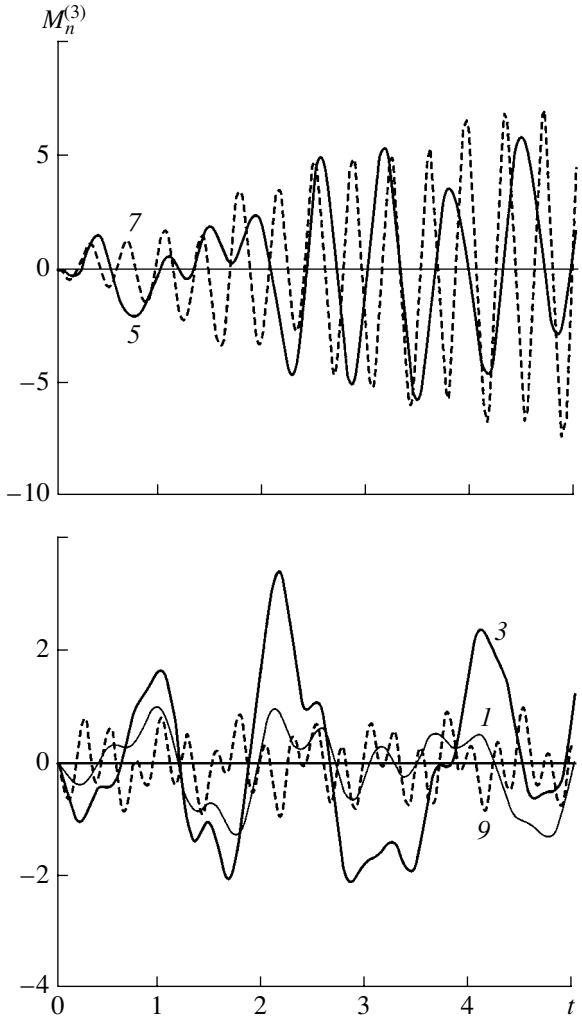

Fig. 3. Time variation of third-order corrections to the mode amplitudes for $W = 2.2$ and the excitation of the third mode. The numbers of the curves coincide with those of initially excited modes.

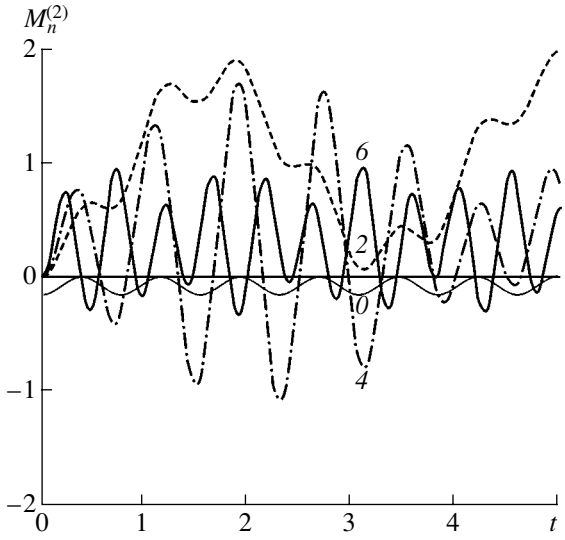

Fig. 4. The same as in Fig. 3 for second-order corrections.

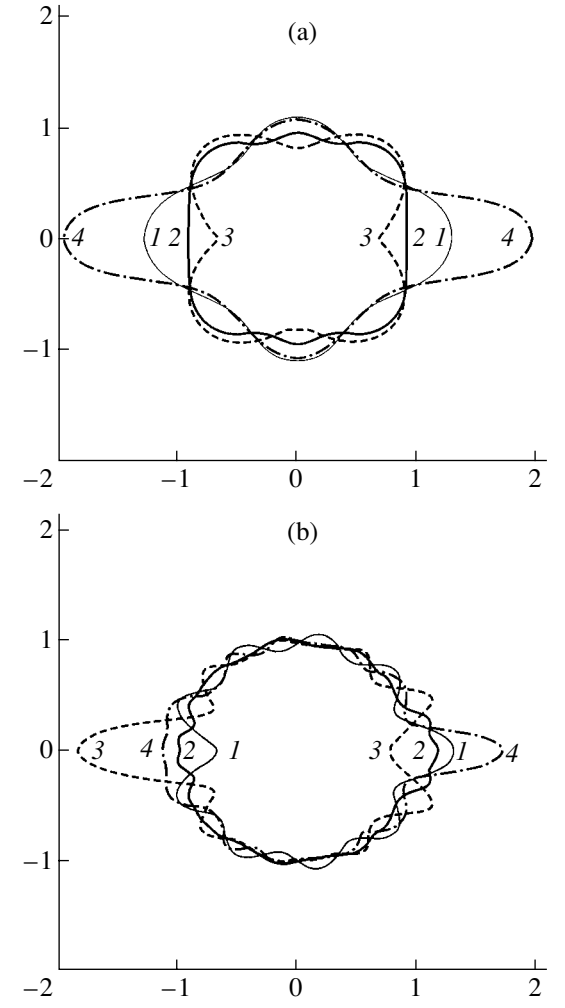

Fig. 5. Generatrix of the drop for the initial excitation of the (a) fourth and (b) ninth modes. For the fourth mode, $\varepsilon = 0.3$, $W = 2.5$, and $t =$ (*1*) 0, (*2*) 0.5, (*3*) 0.9, and (*4*) 1.1. For the ninth mode, $\varepsilon = 0.3$, $W = 2.2$, and $t =$ (*1*) 0, (*2*) 0.1, (*3*) 0.2, and (*4*) 0.3.

number *m* of an initially excited mode is even, the mode with the number *m* is excited in the first order of smallness. In the second order of smallness, the zeroth mode and all even modes with numbers up to 2*m* inclusive are excited, while in the third order of smallness, the zeroth mode and all even modes with numbers up to 3*m* inclusive appear. Thus, the zeroth mode and all even modes with numbers up to 3*m* inclusive shape the drop surface.

Numerical analysis of expressions (50) and (54) for the amplitudes of capillary vibrations showed that, among the third-order corrections $M_n^{(3)}(t)$ to the amplitudes, the largest may be not only the correction to the amplitude of the mode number *m* but also those to the amplitudes of higher modes (Fig. 3). At the same time, among second-order corrections $M_n^{(2)}(t)$ to the amplitudes, the largest is always the correction to the second (fundamental) mode of capillary vibrations (Fig. 4).

Numerical analysis of expression (56) indicates that the areas of the drop surface that are near the polar axis of the spherical coordinate system experience the greatest deviation from the equilibrium state (Fig. 5), because mode superposition occurs only at ϑ close to 0 and π. Away from these values of ϑ, a smoother wavy surface is formed. The larger the number *m* of an initially excited mode, the stronger this tendency (Fig. 5b). For clarity, the drop shapes in Fig. 5 are shown for a large initial perturbation $\varepsilon = 0.3$; therefore, curves *4* reflect the case when the uniformity of the expansion in ε is markedly violated. This shows up in the obvious variation of the initial volumes of drops with the generatrices denoted by figure *4*.

It is interesting that, in the third-order approximation of the single-mode initial deformation due to the broadening of the spectrum of nonlinearly excited modes, internal nonlinear resonances may be found that are absent in the second order of smallness. The four-mode degenerate resonance of the form $3\omega_m = \omega_n$ is an example. In this case, the initial excitation of an *m*th mode causes the buildup of a mode with a higher number *n*.

CONCLUSIONS

Taking into account third-order corrections to the initial excitation amplitude in the problem of the nonlinear vibration of a charged drop considerably broadens the spectrum of modes shaping the drop surface, causes additional nonlinear resonances, and shifts the vibration frequencies toward lower values. Also, there appears a tendency to relieve the critical conditions for the instability of the drop surface against the drop's self-charge with increasing amplitude of the initial perturbation and a tendency toward the formation of emitting asperities through the superposition of higher modes near the poles of the charged drop.

APPENDIX A

Third-Order Approximation of the Capillary Pressure on the Drop Surface

If the drop surface is described by a function of form

$$F(r, \vartheta, t) = r - 1 - \xi(\vartheta, t) = 0,$$

then the normal to the surface is expressed by the formula

$$r = 1 + \xi: \quad \mathbf{n} = \frac{\nabla F}{|\nabla F|} = n_r \mathbf{e}_r + n_\vartheta \mathbf{e}_\vartheta,$$

where

$$n_r = r(r^2 + (\partial_\vartheta \xi)^2)^{-1/2}; \qquad n_\vartheta = -\partial_\vartheta \xi (r^2 + (\partial_\vartheta \xi)^2)^{-1/2}. \tag{1A}$$

Using (1A), we find the Gaussian curvature of the surface:

$$r = 1 + \xi:$$

$$H = \nabla \mathbf{n} = \frac{2}{r} n_r + \partial_r n_r + \frac{1}{r} \partial_\theta n_\vartheta + \frac{n_\vartheta}{r} \cot\vartheta. \tag{2A}$$

The expansions of the expressions entering (2A) have the form

$$r = 1 + \xi:$$

$$\frac{2}{r} n_r = 2 - 2\xi + 2\xi^2 - (\partial_\vartheta \xi)^2 - 2\xi^3 + 3\xi(\partial_\vartheta \xi)^2 + O(\xi^4); \tag{3A}$$

$$r = 1 + \xi: \quad \partial_r n_r = (\partial_\vartheta \xi)^2 - 3\xi(\partial_\vartheta \xi)^2 + O(\xi^4); \tag{4A}$$

$$r = 1 + \xi:$$

$$\frac{1}{r} \partial_\vartheta n_\vartheta = -\partial_{\vartheta\vartheta}\xi + 2\xi \partial_{\vartheta\vartheta}\xi - 3\xi^2 \partial_{\vartheta\vartheta}\xi + \frac{3}{2}(\partial_\vartheta \xi)^2 \partial_{\vartheta\vartheta}\xi + O(\xi^4); \tag{5A}$$

$$r = 1 + \xi:$$

$$\frac{n_\vartheta}{r} = -\partial_\vartheta \xi + 2\xi \partial_\vartheta \xi - 3\xi^2 \partial_\vartheta \xi + \frac{1}{2}(\partial_\vartheta \xi)^3 + O(\xi^4). \tag{6A}$$

Substituting (3A)–(6A) into (2A) yields the Gaussian curvature of the surface in the final form:

$$H = 2 - (2 + \Delta_\Omega)\xi + 2\xi(1 + \Delta_\Omega)\xi - \xi^2(2 + 3\Delta_\Omega)\xi + (\partial_\vartheta \xi)^2 \left(\partial_{\vartheta\vartheta}\xi + \frac{1}{2}\Delta_\Omega \xi\right) + O(\xi^4), \tag{7A}$$

where $\Delta_\Omega \equiv \partial_{\vartheta\vartheta} + \cot\vartheta \partial_\vartheta$.

Represent a perturbation ξ of the drop surface in the form of the series

$$\xi = \varepsilon \xi^{(1)} + \varepsilon^2 \xi^{(2)} + \varepsilon^3 \xi^{(3)} + O(\varepsilon^4) \tag{8A}$$

and, substituting it into (7A), we obtain the expression for the Laplace pressure:

$$P_\sigma = H = 2 - \varepsilon(2 + \Delta_\Omega)\xi^{(1)} - \varepsilon^2(2 + \Delta_\Omega)\xi^{(2)} + 2\varepsilon^2 \xi^{(1)}(1 + \Delta_\Omega)\xi^{(1)} - \varepsilon^3(2 + \Delta_\Omega)\xi^{(3)} - 2\varepsilon^3 \xi^{(1)}((\xi^{(1)})^2 - (2 + \Delta_\Omega)\xi^{(2)}) + 2\varepsilon^3 \xi^{(2)} \Delta_\Omega \xi^{(1)} - 3\varepsilon^3 (\xi^{(1)})^2 \delta_\Omega \xi^{(1)} + \varepsilon^3 (\partial_\vartheta \xi^{(1)})^2 \partial_{\vartheta\vartheta} \xi^{(1)} + \varepsilon^3 \frac{1}{2} (\partial_\vartheta \xi^{(1)})^2 \Delta_\Omega \xi^{(1)} + O(\varepsilon^4).$$

APPENDIX B

Third-Order Approximation for the Electric Field Pressure on the Surface of a Charged Drop

The pressure of an electric field on a liquid surface is given by

$$r = 1 + \xi:$$

$$P_q = \frac{1}{8\pi}(\nabla \phi)^2 = \frac{1}{8\pi}\left[(\partial_r \phi)^2 + \frac{1}{r^2}(\partial_\vartheta \phi)^2\right]. \tag{1B}$$

Expand the derivatives entering (1B) into a series in the vicinity of a sphere in view of the fact that ϕ and $\partial_r \phi$ are of the zeroth order of smallness and $\partial_\vartheta \phi$ is of the first order of smallness (because the potential near a charged spherical drop is defined by expression $\phi^{(0)} = Q/r$):

$$\partial_r \phi|_{r=1+\xi} = \partial_r \phi|_{r=1} + \xi \partial_{rr} \phi|_{r=1} + \frac{\xi^2}{2} \partial_{rrr} \phi|_{r=1} + \frac{\xi^3}{6} \partial_{rrrr} \phi|_{r=1} + O(\xi^4), \tag{2B}$$

$$\partial_\vartheta \phi|_{r=1+\xi} = \partial_\vartheta \phi|_{r=1} + \xi \partial_{r\vartheta} \phi|_{r=1} + \frac{\xi^2}{2} \partial_{rr\vartheta} \phi|_{r=1} + O(\xi^4). \tag{3B}$$

Using (2B) and (3B) we find

$$(\partial_r \phi)^2\big|_{r=1+\xi} = (\partial_r \phi)^2\big|_{r=1} + 2\xi \partial_r \phi|_{r=1} \partial_{rr} \phi|_{r=1} + \xi^2 \partial_r \phi|_{r=1} \partial_{rrr} \phi|_{r=1} + \frac{1}{3}\xi^3 \partial_r \phi|_{r=1} \partial_{rrrr} \phi|_{r=1} + \xi^2 (\partial_{rr} \phi)^2\big|_{r=1} + \xi^3 \partial_{rr} \phi|_{r=1} \partial_{rrr} \phi|_{r=1} + O(\xi^4), \tag{4B}$$

$$\frac{1}{r^2}(\partial_\vartheta \phi)^2\big|_{r=1+\xi} = (\partial_\vartheta \phi)^2\big|_{r=1} + 2\xi \partial_\vartheta \phi|_{r=1} \partial_{r\vartheta} \phi|_{r=1} - 2\xi (\partial_\vartheta \phi)^2\big|_{r=1} + O(\xi^4). \tag{5B}$$

Substituting (4B) and (5B) into (1B) gives

$$r = 1:$$

$$P_q = \frac{1}{8\pi}[(\partial_r\phi)^2 + 2\xi\partial_r\phi\partial_{rr}\phi + \xi^2\partial_r\phi\partial_{rrr}\phi + \frac{1}{3}\xi^3\partial_r\phi\partial_{rrrr}\phi + \xi^2(\partial_{rr}\phi)^2 + \xi^3\partial_{rr}\phi\partial_{rrr}\phi + (\partial_\vartheta\phi)^2 + 2\xi\partial_\vartheta\phi\partial_{r\vartheta}\phi - 2\xi(\partial_\vartheta\phi)^2] + O(\xi^4). \quad (6B)$$

The electric field potential is represented in the form of the expansion

$$\phi = \phi^{(0)} + \varepsilon\phi^{(1)} + \varepsilon^2\phi^{(2)} + \varepsilon^3\phi^{(3)} + O(\varepsilon^4). \quad (7B)$$

Substituting (7B) and (8B) into (6B) and taking into account that $\partial_\vartheta\phi^{(0)} = 0$, we arrive at an expression for the electric field pressure on the drop surface:

$$r = 1:$$

$$P_q = \frac{1}{8\pi}(\partial_r\phi^{(0)})^2 + \frac{\varepsilon}{4\pi}\partial_r\phi^{(0)}(\partial_r\phi^{(1)} + \xi^{(1)}\partial_{rr}\phi^{(0)}) + \frac{\varepsilon^2}{8\pi}\{2\xi^{(2)}\partial_r\phi^{(0)}\partial_{rr}\phi^{(0)} + (\xi^{(1)})^2 \times ((\partial_{rr}\phi^{(0)})^2 + \partial_{rrr}\phi^{(0)}\partial_r\phi^{(0)}) + (\partial_\vartheta\phi^{(1)})^2 + (\partial_r\phi^{(1)}) + 2\partial_r\phi^{(2)}\partial_r\phi^{(0)} + 2\xi^{(1)}(\partial_{rr}\phi^{(0)}\partial_{rr}\phi^{(1)} + \partial_{rr}\phi^{(1)}\partial_r\phi^{(0)})\} + \frac{\varepsilon^3}{8\pi}\Bigg\{2\xi^{(3)}\partial_r\phi^{(0)}\partial_{rr}\phi^{(0)} + (\xi^{(1)})^3 \times\left(\partial_{rr}\phi^{(0)}\partial_{rrr}\phi^{(0)} + \frac{1}{3}\partial_r\phi^{(0)}\partial_{rrrr}\phi^{(0)}\right) + 2(\partial_\vartheta\phi^{(1)}\partial_\vartheta\phi^{(2)} + \partial_r\phi^{(1)}(\xi^{(2)}\partial_{rr}\phi^{(0)} + \partial_r\phi^{(2)}) + \partial_r\phi^{(0)}\partial_r\phi^{(3)} + \xi^{(2)}\partial_r\phi^{(0)}\partial_{rr}\phi^{(1)}) + 2\xi^{(1)}(\xi^{(2)}((\partial_{rr}\phi^{(0)})^2 + \partial_r\phi^{(0)}\partial_{rrr}\phi^{(0)}) + \partial_{rr}\phi^{(0)}\partial_r\phi^{(2)} + \partial_\vartheta\phi^{(1)}(\partial_{r\vartheta}\phi^{(1)} - \partial_\vartheta\phi^{(1)}) + \partial_r\phi^{(1)}\partial_{rr}\phi^{(1)} + \partial_r\phi^{(0)}\partial_{rr}\phi^{(2)}) + (\xi^{(1)})^2 \times(\partial_{rrr}\phi^{(0)}\partial_r\phi^{(1)} + 2\partial_{rr}\phi^{(0)}\partial_{rr}\phi^{(1)} + \partial_r\phi^{(0)}\partial_{rrr}\phi^{(1)})\Bigg\}.$$

APPENDIX C

Expressions for the Coefficients Used to Find the Solution

$$H^{1(+)(-)}_{kmln} = \sum_{g=2}^{\infty}\beta^{1(+)}_{kmgln}\lambda^{(+)}_{lmg} + \sum_{g=1}^{\infty}\mu^{1(-)}_{kmgln} + \sum_{g=0}^{\infty}\mu^{0(-)}_{kmgln},$$

$$H^{1(-)(+)}_{kmln} = \sum_{g=2}^{\infty}\beta^{1(-)}_{kmgln}\lambda^{(-)}_{lmg} + \sum_{g=1}^{\infty}\mu^{1(+)}_{kmgln} + \sum_{g=0}^{\infty}\mu^{0(+)}_{kmgln},$$

$$H^{2(+)(+)}_{kmln} = \sum_{g=2}^{\infty}\beta^{2(+)}_{kmgln}\lambda^{(+)}_{lmg} + \sum_{g=1}^{\infty}\mu^{1(+)}_{kmgln} + \sum_{g=0}^{\infty}\mu^{0(+)}_{kmgln},$$

$$H^{2(-)(-)}_{kmln} = \sum_{g=2}^{\infty}\beta^{2(-)}_{kmgln}\lambda^{(-)}_{lmg} + \sum_{g=1}^{\infty}\mu^{1(-)}_{kmgln} + \sum_{g=0}^{\infty}\mu^{0(-)}_{kmgln},$$

$$H^{0(+)}_{mgn} = (\Pi^0_{mgn} - \Pi^1_{mgn}\omega_m\omega_g - \Pi^2_{mgn}\omega_g^2)(\lambda^{(+)}_{mmg} + \lambda^{(-)}_{mmg}),$$

$$H^{0(-)}_{mgn} = (\Pi^0_{mgn} + \Pi^1_{mgn}\omega_m\omega_g - \Pi^2_{mgn}\omega_g^2)(\lambda^{(+)}_{mmg} + \lambda^{(-)}_{mmg}),$$

$$\beta^{1(+)}_{kmgln} = \Pi^0_{kgn} - \Pi^1_{kgn}\omega_k(\omega_l + \omega_m) - \Pi^2_{kgn}(\omega_l + \omega_m)^2,$$

$$\beta^{1(-)}_{kmgln} = \Pi^0_{kgn} - \Pi^1_{kgn}\omega_k(\omega_l - \omega_m) - \Pi^2_{kgn}(\omega_l - \omega_m)^2,$$

$$\beta^{2(+)}_{kmgln} = \Pi^0_{kgn} + \Pi^1_{kgn}\omega_k(\omega_l + \omega_m) - \Pi^2_{kgn}(\omega_l + \omega_m)^2,$$

$$\beta^{2(-)}_{kmgln} = \Pi^0_{kgn} + \Pi^1_{kgn}\omega_k(\omega_l - \omega_m) - \Pi^2_{kgn}(\omega_l - \omega_m)^2,$$

$$\mu^{1(-)}_{kmgln} = \Lambda^1_{kmgln} - \Gamma^1_{kmgln}\omega_m\omega_k,$$

$$\mu^{1(+)}_{kmgln} = \Lambda^1_{kmgln} + \Gamma^1_{kmgln}\omega_m\omega_k,$$

$$\mu^{0(-)}_{kmgln} = \Lambda^0_{kmgln} - \Gamma^0_{kmgln}\omega_m\omega_k,$$

$$\mu^{0(+)}_{kmgln} = \Lambda^1_{kmgln} + \Gamma^0_{kmgln}\omega_m\omega_k,$$

$$\Lambda^0_{kmgln} = \frac{1}{2k}\Bigg\{K_{gln}(\alpha_{kmg}(kn(l + 3l^2 - 2(k+2)W) + 2(k-2)\omega_k^2) + K_{kmg}(kn(4 - 6k(k+1) + (k^3 - 2(m+1)(m+2) - k^2(n-9) - k(3n + 2m(m+3) - 22))W) - (k-1)k(k-n-2)\omega_k^2)) - 2kn\alpha_{kmg}\sum_{\nu=1}^{[l/2]}(2l - 4\nu + 1)K_{g,\,l-2\nu,\,n}\Bigg\},$$

$$\Lambda^1_{kmgln} = ((g-n-1)K_{gln} - \alpha_{gln}/g)$$
$$\times((m-1)K_{kmg} - \alpha_{kmg}/m)\omega_m^2$$
$$+ Wnk((g+1)(l+n-g-2)K_{gln} + \alpha_{gln})K_{kmg},$$

$$\Gamma^0_{kmgln} = ((k-1)(k-2(n+1))K_{kmg}/2$$
$$-((k-1)(m+n)-m)\alpha_{kmg}/(km))K_{gln}$$
$$+((k-1)(k-2)K_{klg}/2 - (k-2)\alpha_{klg}/k)K_{gmn},$$

$$\Gamma^1_{kmgln} = -((g-n-1)K_{gkn} - (n+k)\alpha_{gkn}/(kg))$$
$$\times((m-1)K_{lmg} - \alpha_{lmg}/m)$$
$$-((g-n-1)K_{gln} - \alpha_{gln}/g)$$
$$\times((m-1)K_{kmg} - \alpha_{kmg}/m),$$

$$\Pi^0_{kmn} = (\omega_k^2(n-k+1) + 2kn(k+1) + 2mn(m+1)$$
$$-4n + nW((n-k-5)(k-1) + (m+1)$$
$$\times(k+n-m-2)))K_{kmn} + (\omega_k^2/k + nW)\alpha_{kmn},$$

$$\Pi^1_{kmn} = (m+k-n-2)K_{kmn} - (n+k+m)\alpha_{kmn}/(mk),$$

$$\Pi^2_{kmn} = (m-n-1)K_{kmn} - \alpha_{kmn}/m,$$

$$\Xi_k = \omega_k^2 + 2k^2(k+1) - 4k - 5k(k-1)W.$$

REFERENCES

1. A. I. Grigor'ev and S. O. Shiryaeva, Izv. Akad. Nauk, Mekh. Zhidk. Gaza, No. 3, 3 (1994).
2. J. A. Tsamopolous and R. A. Brown, J. Fluid Mech. **127**, 519 (1983).
3. T. G. Wang, A. V. Anilkumar, and C. P. Lee, J. Fluid Mech. **308**, 1 (1996).
4. D. F. Belonozhko and A. I. Grigor'ev, Zh. Tekh. Fiz. **70** (8), 45 (2000) [Tech. Phys. **45**, 1001 (2000)].
5. S. O. Shiryaeva, Zh. Tekh. Fiz. **71** (2), 27 (2001) [Tech. Phys. **46**, 158 (2001)].
6. S. O. Shiryaeva, Pis'ma Zh. Tekh. Fiz. **26** (22), 76 (2000) [Tech. Phys. Lett. **26**, 1016 (2000)].
7. S. O. Shiryaeva, Izv. Akad. Nauk, Mekh. Zhidk. Gaza, No. 3, 173 (2001).
8. S. O. Shiryaeva, Zh. Tekh. Fiz. **72** (4), 15 (2002) [Tech. Phys. **47**, 389 (2002)].
9. N. A. Pelekasis, J. A. Tsamopolous, and G. D. Manolis, Phys. Fluids A **2**, 1328 (1990).
10. V. N. Gorshkov and M. G. Chaban, Zh. Tekh. Fiz. **69** (11), 1 (1999) [Tech. Phys. **44**, 1259 (1999)].
11. L. D. Landau and E. M. Lifshitz, *Course of Theoretical Physics*, Vol. 6: *Fluid Mechanics* (Nauka, Moscow, 1986; Pergamon, New York, 1987).
12. L. D. Landau and E. M. Lifshitz, *Course of Theoretical Physics*, Vol. 8: *Electrodynamics of Continuous Media* (Nauka, Moscow, 1982; Pergamon, New York, 1984).
13. A. H. Nayfeh, *Perturbation Methods* (Wiley, New York, 1973; Mir, Moscow, 1976).
14. D. A. Varshalovich, A. N. Moskalev, and V. K. Khersonskii, *Quantum Theory of Angular Momentum* (Nauka, Leningrad, 1975; World Scientific, Singapore, 1988).
15. S. O. Shiryaeva, Zh. Tekh. Fiz. **73** (2), 19 (2003) [Tech. Phys. **48**, 152 (2003)].
16. A. I. Grigor'ev, Zh. Tekh. Fiz. **55**, 1272 (1985) [Sov. Phys. Tech. Phys. **30**, 736 (1985)].
17. Lord Rayleigh, Philos. Mag. **14**, 184 (1882).

Translated by N. Mende

Technical Physics, Vol. 48, No. 6, 2003, pp. 708–713. Translated from Zhurnal Tekhnicheskoĭ Fiziki, Vol. 73, No. 6, 2003, pp. 46–52.

SOLIDS

Heat Evolution Due to Pulsed Thermal Perturbations in Superconducting Composites with a Nonuniformly Distributed Current

V. R. Romanovskii

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

Received May 14, 2002; in final form, December 27, 2002

Abstract—Dissipative processes in superconducting composites under the action of current and pulsed thermal perturbations are studied. Results obtained are compared with the predictions following from the existing theory of thermal stabilization of superconducting composites. It is shown that, if the current is nonuniformly distributed over the composite's cross section and the thermal perturbations are short, intense additional energy dissipation may arise both before and after the occurrence of thermomagnetic instability. Associated Joule heat evolution may substantially exceed the thermal loss predicted by the adopted theory of thermal stabilization throughout the temperature range of the composite. Additional dissipation in composite superconductors narrows the range of stable superconductivity, especially when the current is applied continuously, even if external thermal perturbations are weak. © *2003 MAIK "Nauka/Interperiodica".*

INTRODUCTION

It is known that current applied to a superconducting magnetic system may initiate its premature transition to the normal state. The current characteristics of superconducting magnetic systems vary because of degradation and training processes [1]. In the first case, the operating parameters of magnets are markedly worse than those of short samples made of the same material. Training means that the current characteristics of a superconducting magnet are gradually improved after several superconducting–normal transitions. The basic reason for both phenomena is the instability of the superconducting state of a composite against various thermal perturbations, which is largely associated with high current-induced mechanical stresses. Under these conditions, the turns of a magnet are inevitably displaced and rub against each other, causing the material of the turns to plastically deform and an electrical compound (if any is present) to crack. These perturbations, being of both continuous and intermittent action, result in additional heat evolution. Eventually, the superconductivity of the current-carrying elements may locally break. Depending on the perturbation intensity, the system may either recover the superconductivity or irreversibly pass to the normal state. Therefore, physical processes taking place in superconducting current-carrying elements under the action of current and thermal perturbations should be studied in order to provide the stable operation of superconducting magnets (especially those exhibiting high mechanical stresses).

The theory of thermal stabilization is of great help in finding ways of retaining the superconducting state of current-carrying elements. Its basic concepts are clarified by a model where it is assumed that the current distribution over the composite's cross section is uniform and the Joule heat power is defined by instantaneous changes in the temperature field [1, 2]. It is also assumed that the Joule heat evolution in a superconducting composite starts at the instant the composite temperature exceeds the temperature T_c of current splitting. At T_{cI}, the applied current density becomes equal to the critical current density of the composite. When T_{cI} is exceeded, the current is divided among the superconducting inclusion and the matrix, and when the temperature reaches the superconducting transition critical temperature T_c, the current passes only through the matrix. These assumptions simplify the analysis and, in addition, make it possible to derive analytical tests for superconductivity stability [1, 2]. However, this model applies in the case of stationary currents [3, 4]. It should be mentioned that the existence of the stationary state is in many respects a matter of convention. First, the electromagnetic field inside a superconductor will necessarily respond to any external perturbation, which causes heat release. Second, the applied current distribution over the cross section of the composite is usually nonuniform. Note also that adequate analysis of electrodynamic states must take into consideration the correlation between the superconductor temperature and magnetic flux diffusion as the current passes through the system [5, 6]. For this reason, based on the numerical solution of the Fourier and Maxwell equations, we study dissipation in a composite superconductor due to pulsed thermal perturbations of different duration. Our goal is to substantiate simplified models that are used within the theory of stationary stabilization, which disregard thermal and electromagnetic phenomena concurrently proceeding in superconductors.

1063-7842/03/4806-0708$24.00

MODEL

In terms of a nonisothermal model, we will consider energy dissipation in a multifilament superconducting composite of circular cross section to which a current is applied. Near the outer surface of the composite, an extended thermal perturbation is placed. At zero time, the current in the composite kept at the coolant temperature is absent. It is also assumed that (i) the fibers are uniformly distributed over the cross section, (ii) the *I–V* characteristic of the superconductor includes a magnetic flux creep and is described by the exponential [7], (iii) convective heat exchange with a constant heat-transfer coefficient takes place on the composite surface, (iv) the current varies with a constant rate, (v) the current is applied from an external source after a given external magnetic field has fully occupied the superconducting region, (vi) the power of the external heat source remains constant during its action, and (vii) the heat source is uniformly distributed along the length of the composite and heats a thin surface layer $\Delta r = r_0 - r_q = 0.1r_0$.

According to the statement of the problem, current-induced lines of magnetic force are parallel to the cylinder axis and the current penetrates into the composite from the surface. Therefore, the space–time temperature variation in cylindrical coordinates is described by the one-dimensional non-steady-state equation

$$c\frac{\partial T}{\partial t} = \frac{1\partial}{r\partial r}\left(\lambda r\frac{\partial T}{\partial r}\right) + EJ + q(r, t), \qquad (1)$$

where the current density and the associated electric field strength are related (in terms of the continuum model) as

$$\mu_0\frac{\partial J}{\partial t} = \frac{1\partial}{r\partial r}\left(r\frac{\partial E}{\partial r}\right),$$
$$E = J_s\rho_s\exp\left(\frac{J_s}{J_\delta} + \frac{T - T_c}{T_\delta}\right) = J_m\rho_m, \qquad (2)$$
$$J = \eta J_s + (1 - \eta)J_m.$$

Here, c and λ are the specific heat and thermal conductivity of the composite, respectively; ρ_s and ρ_m are the resistivities of the superconductor and matrix, respectively; J_s and J_m are the currents passing through the superconductor and matrix, respectively; η is the volume fraction of the superconductor in the composite (filling factor); T_c is the critical temperature of the superconductor; J_δ and T_δ are the rise parameters of the *I–V* characteristic of the superconductor; and $q(r, t)$ is a given power of the external heat perturbation:

$$q(r, t) = \begin{cases} q_0 = \text{const}, & r_q \le r \le r_0, \quad t_1 \le t \le t_1 + \Delta t \\ 0, & t > t_1 + \Delta t. \end{cases}$$

In our problem, Eqs. (1) and (2) should be complemented by the initial and boundary conditions

$$T(r, 0) = T_0, \quad E(r, 0) = 0,$$
$$\lambda\frac{\partial T}{\partial r} + h(T - T_0)\Big|_{r = r_0} = 0, \qquad (3)$$
$$\left.\frac{\partial E}{\partial r}\right|_{r = r_0} = \frac{\mu_0}{2\pi r_0}\frac{dI}{dt},$$

where r_0 is the cross-sectional radius of the composite, h is the heat-transfer coefficient, dI/dt = const is the rate of current variation, and T_0 is the coolant temperature.

Problem (1)–(3) is stated in the most general form and allows one to study processes taking place in superconducting composites both before and after the onset of instability without separating out the thermal and electrodynamic subproblems. The results obtained in [5, 6] suggest that precisely such an approach should be used for studying phenomena occurring in superconductors at high rates of current or magnetic field variation and under adiabatic conditions for the insulation of current-carrying elements in superconducting magnetic systems. It should also be noted that, while simple, the form of a thermal perturbation adopted in this work makes it possible to reflect the effect of the most dangerous short perturbations acting on the surface of the composite. In this case, the form of a thermal pulse affects the final thermal state of the composite only slightly [1, 2].

The set of Eqs. (1) and (2) is nonlinear and is difficult to solve analytically. Therefore, we performed numerical simulation for Nb–Ti superconducting filaments embedded in a copper matrix. The averaged values of the starting parameters were as follows: $r_0 = 5 \times 10^{-4}$ m, $dI/dt = 10^4$ A/s, $c = 10^3$ J/(m^3 K), $\lambda = 100$ W/(m K), $h = 10$ W/(m^2 K), $T_0 = 4.2$ K, $T_c = 9$ K, $\eta = 0.5$, $\rho_s = 5 \times 10^{-7}$ Ω m, $\rho_m = 2 \times 10^{-10}$ Ω m, $J_\delta = 4 \times 10^7$ A/m^2, and $T_\delta = 0.048$ K. Note that taking into account the temperature dependences of these parameters introduces only quantitative corrections without affecting qualitative inferences.

RESULTS

Figure 1 shows the temperature dependences of Joule heat evolution per unit length of the composite:

$$G_s = \int_0^{r_0} EJr\,dr.$$

Heat evolution was initiated by perturbations of different duration and intensity when the current applied to the composite was much lower than the so-called quenching current (the current at which the superconducting composite irreversibly passes to the normal state). According to [5], this current is defined as the current at which a superconducting composite sponta-

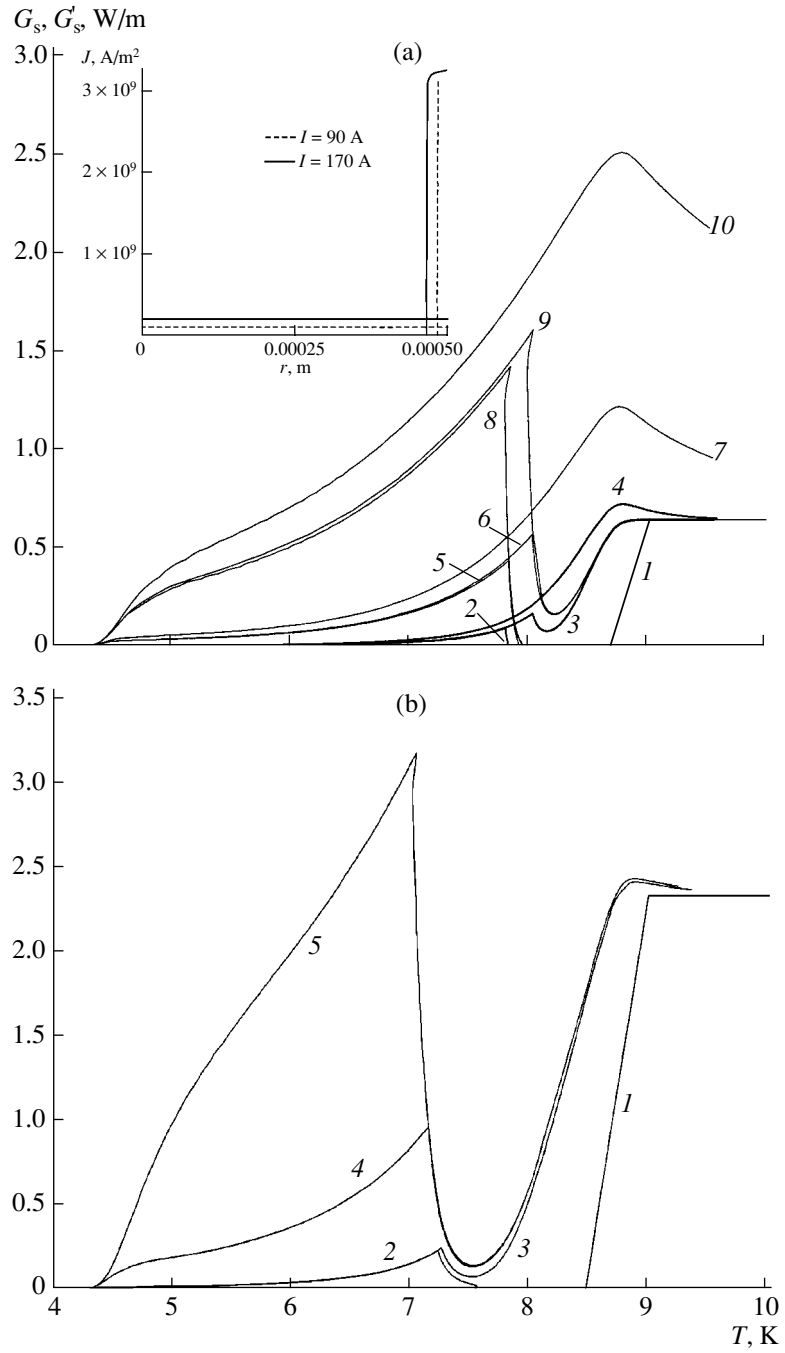

Fig. 1. (a) Temperature dependence of Joule heat evolution in the composite superconductor under the action of external thermal pulses of various duration and power after the current I_0 = 90 A has been applied. (*1*) Analytical calculation by formula (4); (*2*)–(*10*) numerical calculation based on problem (*1*)–(*3*) for various Δt and q_0. Δt = (*2–4*) 10^{-3}, (*5–7*) 10^{-4}, and (*8–10*) 10^{-5} s; q_0 = (*2*) 0.19×10^8, (*3*) 0.2×10^8, (*4*) 0.3×10^8, (*5*) 1.9×10^8, (*6*) 2.0×10^8, (*7*) 3.0×10^8, (*8*) 19×10^8, (*9*) 20×10^8, and (*10*) 30×10^8 W/m^3. (b) The same as in (a) for I_0 = 170 A. (*1*) Calculation by (4); (*2–10*) numerical calculation for various Δt and q_0. Δt = (*2*, *3*) 10^{-3}, (*4*) 10^{-4}, and (*5*) 10^{-5} s; q_0 = (*2*) 0.13×10^8, (*3*) 0.14×10^8, (*4*) 1.4×10^8, and (*5*) 14×10^8 W/m^3.

neously passes to the normal state (with a change in the composite temperature prior to the onset of instability properly taken into consideration) even if the application of the current ceases. Subcritical (without quenching the superconductivity) and above-critical (when superconductivity quenching is observed) perturbations specify the thermal stability range for the superconducting state of the composite for each value of the current applied. The calculation was carried out under the assumption that the current was switched off at $t \geq$

$t_1 + \Delta t$. The inset to Fig. 1a shows the radial current distribution before the action of the perturbation ($t = t_1$). Figure 1 also demonstrates the heat evolution powers calculated by the formulas that are adopted in the theory of thermal stabilization. In our notation, the thermal power is given by

$$G_s' = \frac{S}{2\pi}\frac{J_e^2\rho_m}{1-\eta} \times \begin{cases} 1, & T > T_c \\ \dfrac{(T - T_{cI})}{(T_c - T_{cI})}, & T_{cI} \le T \le T_c \\ 0, & T < T_{cI} = T_c - (T_c - T_0)\dfrac{J_e}{J_{c0}}, \end{cases} \quad (4)$$

where J_{c0} is the critical current density for a superconductor with the ideal I–V characteristic (J_δ, $T_\delta \longrightarrow 0$). This parameter is a conventional quantity [7], appearing when the exponent of the exponential describing the I–V characteristic is set equal to zero: $J_{c0} = J_\delta(T_c - T_0)/T_\delta$. With the above values of the parameters, $J_{c0} = 4 \times 10^9$ A/m^2. The density of the current J_e was calculated under the assumption that the current applied is uniformly distributed over the cross section. The values of J_e are given in the inset to Fig. 1a.

The curves depicted in Fig. 1 clearly demonstrate that perturbation-induced dissipative processes, which take place even when the composite is in the stable superconducting state, are accompanied by additional Joule heat evolution of intensity depending on the thermal pulse power. For near-critical perturbations, the thermal losses increase as the pulse narrows. In this case, the real curves $G_s(T)$ not only deviate markedly from the simplified curves $G_s'(T)$ (formula (4)) but also exhibit nonmonotonically varying (multivalued) portions. As a result, external thermal perturbations may significantly raise the composite's background temperature preceding the onset of instability. Because of this, in the case of short thermal pulses, the temperature range where additional thermal losses can be neglected may noticeably shrink. It is also noteworthy that model (1)–(3) for finding the Joule heat release when thermomagnetic instability develops in a superconducting composite results in the appearance of a peak in the dependence $G_s(T)$ (unlike expression (4)). Moreover, this model overestimates the power of heat evolution, which is observed even after the temperature of the superconductor has exceeded the critical value. This is especially true for above-critical perturbations. The fact is that instability develops in an avalanche-like manner: the electric field strength in the composite first drastically grows, reaches a maximal value, and then drops for a finite time.

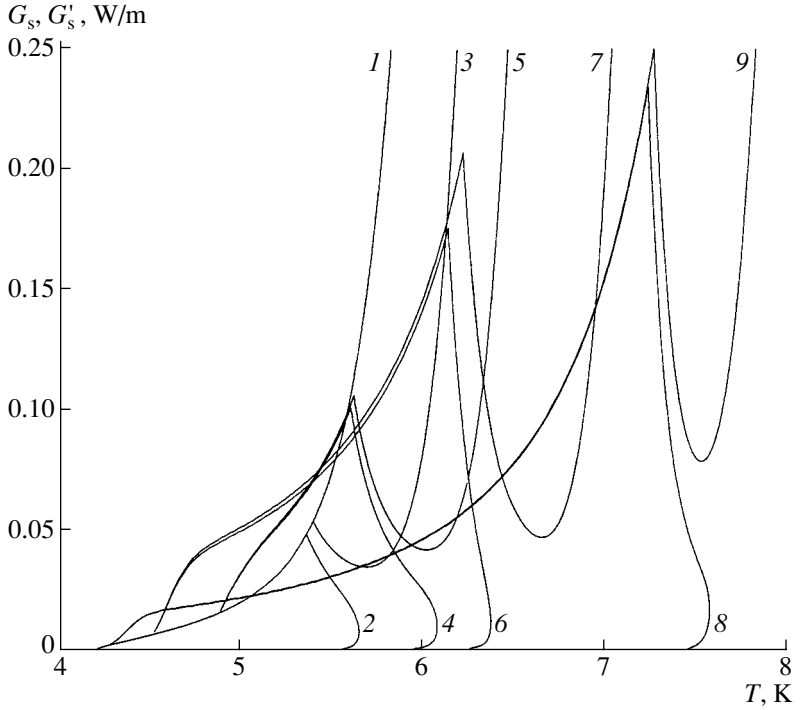

Fig. 2. Temperature variation of Joule heat evolution due to external thermal pulses of duration $\Delta t = 10^{-3}$ s under the action of a near-quenching current I_0. (*1*) Continuous application of the current; q_0 = (*1–3*) 0, (*4*) 0.189×10^7, (*5*) 0.192×10^7, (*6*) 0.6×10^7, (*7*) 0.62×10^7, (*8*) 1.3×10^7, and (*9*) 1.4×10^7 W/m^3; I_0 = (*2*) 359, (*3*) 360, (*4*, *5*) 330, (*6*, *7*) 270, and (*8*, *9*) 170 A.

From Fig. 1, it follows that, in the range of applied currents studied, the temperature at which thermal-pulse-induced instability arises is lower than the current splitting temperature. At the same time, as the current applied to the composite increases and approaches the quenching current, the permissible power of external perturbations must decrease to zero because of the development of thermomagnetic instability [5]. Because of this, additional heat evolution at temperatures between T_0 and T_{cI} and currents close to the quenching current is of great interest. Curves *2* and *3* in Fig. 2 describe the thermoelectrodynamic states of the composite in the absence of perturbations after near-quenching currents have been switched off. Curve *2* corresponds to a current slightly below the quenching value; curve *3*, to a current that somewhat exceeds this value. In the former case, superconductivity quenching is not observed even with the finite value of Joule heat evolution. In the latter case, resulting thermomagnetic instability quenches the superconducting state of the composite even when the current is switched off. To analyze heat evolution near the quenching current, we took external perturbations of powers corresponding to two extreme, superconducting and normal, states of the composite. As was expected, when the current was raised up to the quenching value, the additional heat evolution initiated by subcritical perturbations decreased. It should be taken into account here that instability always arises at temperatures below the current splitting temperature. Therefore, the substantial additional heat evolution, which is observed at temperatures below T_{cI}, is due to instability arising in this temperature range. The existing theory ignores the heat evolution in this temperature range, as follows from (4).

The results shown in Figs. 1 and 2 correspond to the case when the current is switched off during the action of a thermal perturbation or at the instant the perturbation disappears. Actually, the current is applied continuously. Let us therefore consider dissipative processes when the current is continuously applied to the composite in view of the existence of the quenching current. Figure 3 shows the temperature dependences $G_s(T)$ of the Joule heat evolution initiated by perturbations of different intensity and duration $\Delta t = 10^{-3}$ s. For comparison, the associated dependences $G_s'(T)$ obtained by simplified formula (4) for currents after the application of which the perturbation starts acting on the composite are also shown. The inset to Fig. 3 demonstrates the nontrivial behavior of the dependences $G_s(T)$ immediately after the action of the thermal perturbations.

It is seen that, when the current is applied continuously, the Joule heat evolution depends significantly on the external thermal perturbation intensity in the entire temperature range from T_0 to T_c. The actual dependences differ from the simplified results (following from the existing theory of thermal stabilization), as might be expected. Of interest is the unexpected (at first glance) premature onset of instability under the action

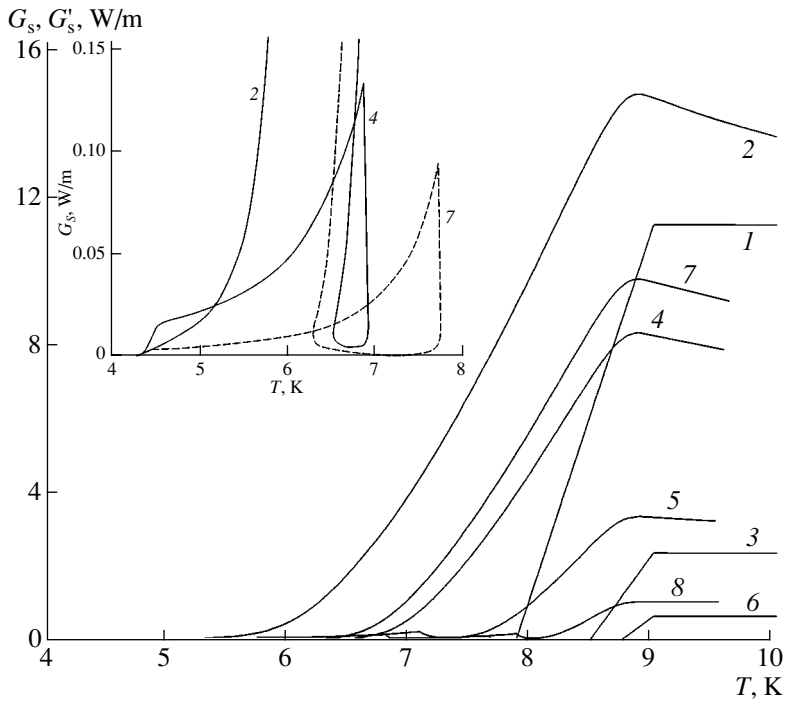

Fig. 3. Temperature variation of Joule heat evolution due to external thermal pulses under the continuous application of the current. (*1*, *3*, *6*) Calculation by formula (4) for I_0 = (*1*) 360, (*3*) 170, and (*6*) 90 A. q_0 = (*2*) 0, (*4*) 0.12×10^8, (*5*) 0.13×10^8, (*7*) 0.18×10^8, and (*8*) 0.19×10^8 W/m^3; I = (*4*, *5*) 170 and (*7*, *8*) 90 A.

of subcritical perturbations. As follows from Figs. 1 and 2, the term "subcritical" or "above-critical," as applied to perturbations, means the retention or quenching of the superconducting state after the current has been switched off. Curves *4* and *7* in Fig. 3 describe the variation of the thermal state of the composite under the continuous application of the current and the action of subcritical (in the above sense) perturbations. Both curves run below curve *2*, which describes the spontaneous quenching of the superconductivity when the current exceeds the quenching value. One can thus infer that the variation of the thermal state of the composite that is shown by curves *4* and *7* (and, hence, *5* and *8*) is eventually due to the development of thermomagnetic instability. Such a singularity in the concurrent occurrence of thermal and electrodynamic processes under the action of a subcritical (for a given current) thermal perturbation on superconducting composites directly follows from the thermal nature of thermomagnetic instability in superconductors. According to [5], the total amount of heat evolved in a superconductor is always limited irrespective of the nature of the perturbation. Therefore, the material passes to the normal state when the heat evolved exceeds a certain critical value for a specific case.

Thus, external thermal perturbations acting on superconducting composites with nonuniformly distributed current may give rise to high thermal losses both when the superconductivity is stable and at the stage of instability development. As a result, the temperature of the composite rises continuously during the redistribution of the thermal-pulse-induced current. Therefore, the simplified formula for finding Joule heat evolution that is derived within the existing theory of thermal stabilization under the assumption of time-invariable current uniformly distributed over the cross section may underestimate the heat release in the entire temperature range if the thermal pulses are short. Moreover, when the current is applied continuously, the additional heat release may initiate the premature quenching of the superconducting state even if the perturbation is weak. If actual dissipative processes in current-carrying elements, which are a direct consequence of the interrelation between the thermal and electromagnetic states, are disregarded, changes in the superconducting state stability are impossible to characterize correctly. Our results demonstrate, first, the physics of dissipative processes taking place in current-carrying elements under the action of current and external thermal sources. Second, they make it necessary to consider the problem of trouble-free operation of superconducting magnets with allowance for the inevitable narrowing of the operating current range under the action of various thermal pulses.

ACKNOWLEDGMENTS

This work was supported by the Russian Foundation for Basic Research (project no. 01-02-16252a).

REFERENCES

1. M. N. Wilson, *Superconducting Magnets* (Oxford Univ. Press, London, 1983; Mir, Moscow, 1985).
2. V. A. Al'tov, V. B. Zenkevich, M. G. Kremlev, and V. V. Sychev, *Stabilization of Superconducting Magnetic Systems* (Énergoatomizdat, Moscow, 1984).
3. A. A. Pukhov and A. L. Rakhmanov, Sverkhprovodimost: Fiz. Khim. Tekh. **5**, 1620 (1992).
4. A. A. Pukhov and A. L. Rakhmanov, Sverkhprovodimost: Fiz. Khim. Tekh. **6** (6), 1165 (1993).
5. V. E. Keilin and V. R. Romanovskii, Cryogenics **33**, 986 (1993).
6. V. R. Romanovskii, Dokl. Akad. Nauk **336**, 341 (1994) [Phys. Dokl. **39**, 363 (1994)].
7. G. L. Dorofejev, A. D. Imenitov, and E. Yu. Klimenko, Cryogenics **20**, 307 (1980).

Translated by V. Isaakyan

Technical Physics, Vol. 48, No. 6, 2003, pp. 714–716. Translated from Zhurnal Tekhnicheskoĭ Fiziki, Vol. 73, No. 6, 2003, pp. 53–55.

SOLIDS

Localization of Plastic Strain under Shock-Wave Loading of Titanium Alloy with a Crack

A. I. Petrov and M. V. Razuvaeva

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

e-mail: an.petrov@mail.ioffe.ru

Received November 20, 2002

Abstract—The effect of shock-wave loading with an amplitude of 4 GPa and a duration of 10 μs on the localization of plastic flow and the dynamic fracture of VT-1 cast titanium alloy containing a disk-shaped crack is considered. It is found that the crack size decreases in the direction of shock wave propagation, the dislocation density grows and adiabatic shear bands appear near the crack, and secondary cracks originate in the areas of localized flow near the crack edges. The strain, strain rate, dislocation density, rate of dislocation generation, and dislocation velocity near and away from the healing crack are estimated. *© 2003 MAIK "Nauka/Interperiodica".*

The healing of pores and cracks in crystalline materials is known to proceed by the dislocation mechanism under static or dynamic compression at medium temperatures [1, 2]. The healing of discontinuities depends on the number of dislocation loops generated by Frank–Read dislocation sources under the action of shear stresses that appear near pores and cracks under pressure. It was shown [3] that the effect of pore healing is much more efficient upon shock-wave loading (pulsed magnetic compression and shock loading) than under static loading of the same amplitude. One reason for this effect may be an increase in the dislocation mobility because of the material heating near a discontinuity, since various internal inhomogeneity regions in the material cause changes in the wave pattern and, hence, an increase in the temperature during dynamic loading [4]. Another reason for the increase in dislocation mobility upon dynamic compression is local heating in slip bands due to adiabatic processes arising when dislocations move with high velocities. In low-diffusivity materials, local heatings in slip bands are known to reach a critical value for adiabatic shear [5]. Adiabatic shear bands would thus be expected to form in such materials near healing-up discontinuities in some modes of dynamic loading. The aim of this work was to study the shear strain localization near discontinuities under explosive dynamic compression.

The material to be investigated was VT-1 cast titanium alloy with a grain size of 900 ± 200 μm, which is prone to the formation of adiabatic shear bands [5, 6]. In a 20-mm-thick plate of the initial material, we detected a 2.5-mm-high disk-shaped crack 15 mm in diameter, using X-ray inspection. The crack plane was almost parallel to the plane of the plate. A 70×70-mm sample with a crack in the middle was cut from the plate and subjected to explosive loading. The sample was loaded by a plane shock wave so that the wave front propagated normally to the crack plane. The shock amplitude and duration were 4 GPa and about 10 μs, respectively. To prevent the transverse spreading of the material, the sample was placed into a yoke. After shock-wave loading, the sample was metallographically examined with a Neophot-30 optical microscope and a Comscan scanning electron microscope on a polished section cut normally to the crack plane.

Let us discuss the results obtained. Metallographic examination of the polished section whose plane coincides with the plane shock wave direction and passes through the center of the crack evidences the following. The general strained structure of the alloy changes significantly after the shock-wave loading. Moreover, the strained structures near the crack and away from it are considerably different. Away from the crack (at distances exceeding several grain sizes), the material deforms by shear, as a rule, in two slip systems. The strained structure of the alloy is characterized by straight rows of dislocations. As the crack is approached, other shear systems come into play: multiple slip along nonparallel planes occurs and apparent slip proceeds over three or four slip systems. In grains that are in immediate contact with the crack, the plastic strain rate increases to the point where individual slip bands become indistinguishable. The activation of dislocation sources near the crack is accompanied by a considerable decrease in the crack size in the plane shock wave direction; that is, the crack heals up. After the loading, the residual strain in the plate at the center of the crack was estimated as 0.12.

We also found that adiabatic heating causes localized shear bands to appear near the crack. Grains adjacent to the crack contain 15 adiabatic shear bands.

1063-7842/03/4806-0714$24.00

Their width varied from 4 to 10 μm, and the bands occupy one or several grains. Figure 1 shows an adiabatic shear band. The bands are similar to those observed in titanium [6, 7] and its alloys [5, 7]. To determine the plastic shear strain γ inside an adiabatic shear band, we used grain boundaries running perpendicularly to the shear band as references. The value of γ was found to be about 500%. The SEM examination of the section showed that the shear bands contain many pores several micrometers in size.

Optical and scanning electron microscopy studies revealed a small amount of deformation twins in a number of grains near the crack.

Metallographic analysis of the polished section also revealed two types of ≈0.2-mm-wide extended secondary cracks with a width-to-length ratio of 1 : 30 appearing at the edges of the macrocrack after the shock loading (Fig. 2). Cracks of the first type propagate in the direction of the shock wave; that is, these are shearing cracks. Cracks of the second type propagate at an angle of 50° to the macrocrack plane. Cracks of both types develop in the adiabatic shear bands at the edges of the initial crack, as indicated by the shifts along the grain boundaries crossed by the secondary cracks (Fig. 3). The shift of the grain boundary due to the shear band in Fig. 3 is 25 μm, which is close to the grain-boundary shift in the adiabatic shear bands considered above. Another indication of secondary cracking in adiabatic shear bands is the fact that the bands of intense localized shear are the extensions of the cracks. Note that the tilted adiabatic shear bands are generated by shear stresses that are close to the maximum shear stresses acting at an angle of 45° to the shock wave front.

Thus, upon shock loading, the plastic strain is localized in adiabatic shear bands at the macrocrack edges. The localized flow causes secondary cracking. According to [7], cracking in adiabatic shear bands is initiated by tensile stresses that appear when the shock wave reflects from the back surface of a sample.

Let us perform some estimations. First, we will estimate the change in the dislocation density owing to the healing up of the macrocrack upon shock loading. Under static loading in the absence of dynamic recovery, the dislocation density ρ is known [8] to depend linearly on the true strain e:

$$\rho = \rho_0 + Ae, \quad (1)$$

where ρ_0 is the dislocation density in the initial material and A is the coefficient characterizing the rate of dislocation multiplication ($A \sim 10^{15}$ m^{-2} for polycrystals).

It was experimentally shown that Eq. (1) is also valid in the case of dynamic loading if plastic deformation is related to dislocation slip. Examples of ρ vs. e dependences for shock-loaded materials are given in [4, 9]. It was shown that the coefficient A for metals and alloys with various crystal lattices is approximately the same and, on average, is equal to $(1.2 \pm 0.2) \times 10^{15}$ m^{-2}. Assuming that the contribution of deformation twins to the plastic strain near the macrocrack at a pressure of 4 GPa is insignificant, we find that the maximum true strain near the crack, $e = \ln(1 + \varepsilon) = 0.11$, corresponds to the dislocation density 1.3×10^{14} m^{-2} (the dislocation density ρ_0 in the cast alloy is lower by several orders of

Fig. 1. Adiabatic shear band near the initial macrocrack in VT-1 alloy after dynamic loading.

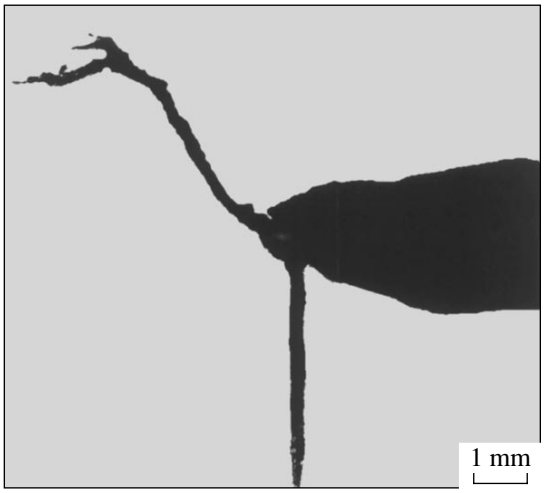

Fig. 2. Shock-wave generated secondary cracks near the edges of the initial macrocrack.

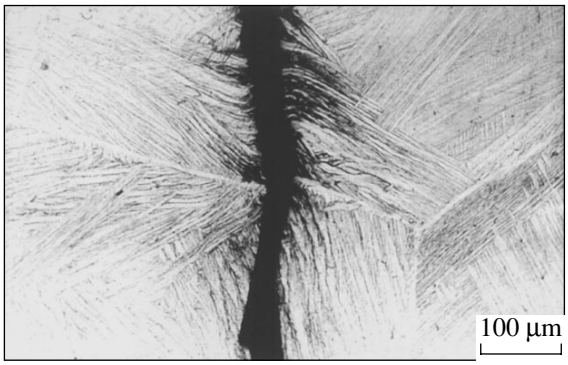

Fig. 3. Intersection of a secondary crack and a grain boundary.

magnitude and can be neglected). At a compression pulse duration $t_p = 10^{-5}$ s, the rate of dislocation generation is $\dot{\rho} = \rho/t_p = 1.3 \times 10^{19}$ m^{-2} s^{-1} at an average strain rate $\dot{\varepsilon} = \varepsilon/t_p = 1.2 \times 10^4$ s^{-1}. As follows from the Orowan equation, the average dislocation velocity

$$\vartheta_d = \dot{\varepsilon}/\rho b, \quad (2)$$

where $b = 2.95 \times 10^{-10}$ m is the Burgers vector, equals 0.3 m s^{-1}.

Let us estimate the same parameters away from the crack. The true plastic strain of a material subjected to dynamic pressure is calculated by the formula

$$e = \frac{4}{3}\ln V/V_0, \quad (3)$$

where V/V_0 is the ratio of the volume during the passage of the shock wave to the initial volume. At a pressure of 4 GPa, $e = 1.86 \times 10^{-2}$ for the material under study. According to Eqs. (1) and (2), this strain at a compression pulse duration of 10 μs yields $\dot{\varepsilon} = 1.86 \times 10^3$ s^{-1}, $\rho = 2.2 \times 10^{13}$ m^{-2}, $\dot{\rho} = 2.2 \times 10^{18}$ m^{-2} s^{-1}, and $\vartheta_d = 0.3$ m s^{-1}.

These estimates show that a disk-shaped crack in the bulk of the shock-loaded material acts as a shear stress concentrator. Therefore, the plastic strain, strain rate, dislocation density, and rate of dislocation generation near the crack are higher than those away from it by almost one order of magnitude. This means that plastic deformation localizes near the crack. Other places of localized plastic flow are secondary cracks appearing near the crack and adiabatic shear bands, in which the shear strain and strain rate are one more order of magnitude higher.

Thus, high shear stresses arising in the shock wave near the crack lead to severe straining of the material via the activation of various shear systems and the formation of adiabatic shear bands. The healing up of the macrocrack in the bulk of the material under shock-wave loading is accompanied by secondary cracking at the edges of the macrocrack at sites of plastic strain localization in adiabatic shear bands.

ACKNOWLEDGMENTS

We thank G.A. Malygin for useful discussions and valuable comments.

REFERENCES

1. Ya. E. Geguzin and V. G. Kononenko, Fiz. Khim. Obrab. Mater., No. 2, 60 (1982).
2. V. I. Betekhtin, A. B. Pakhomov, B. P. Peregud, *et al.*, Zh. Tekh. Fiz. **59** (6), 136 (1989) [Sov. Phys. Tech. Phys. **34**, 668 (1989)].
3. A. I. Petrov, M. V. Razuvaeva, A. B. Sinani, *et al.*, Zh. Tekh. Fiz. **68** (11), 125 (1998) [Tech. Phys. **43**, 1385 (1998)].
4. *Shock Waves and High-Strain-Rate Phenomena in Metals*, Ed. by M. A. Meyers and L. E. Murr (Plenum, New York, 1981; Metallurgiya, Moscow, 1984).
5. S. P. Timothy and L. M. Hutchings, Acta Metall. **33**, 667 (1985).
6. M. A. Mogilevskiĭ, T. O. Sanchaa, and Yu. D. Shishkin, Zh. Prikl. Mekh. Tekh. Fiz., No. 3, 118 (1986).
7. H. A. Grebe, H.-R. Pak, and M. A. Meyers, Metall. Trans. A **16**, 761 (1985).
8. *Strain Hardening and Failure of Polycrystalline Metals*, Ed. by V. I. Trefilov (Naukova Dumka, Kiev, 1987).
9. R. N. Whight and D. E. Mikkola, Metall. Trans. A **16**, 981 (1985).

Translated by K. Shakhlevich

Technical Physics, Vol. 48, No. 6, 2003, pp. 717–721. Translated from Zhurnal Tekhnicheskoĭ Fiziki, Vol. 73, No. 6, 2003, pp. 56–60.

SOLIDS

Use of the Temperature Peak Model for the Description of Track Formation in Semiconductor Crystals Irradiated by Fast Heavy Ions

F. F. Komarov and V. N. Yuvchenko

Research Center, Sevchenko Institute of Applied Physical Problems, Belarussian State University, Minsk, 220064 Belarus

e-mail: paliichuk@pfp.bsu.unibel.by

Received November 25, 2002

Abstract—An attempt to apply the temperature peak model to describe the formation of defects and tracks in semiconductor crystals is made for the first time. The temperature dependences of model parameters, such as specific heat, thermal conductivity, and electron–phonon coupling coefficient are obtained. Agreement between the theoretical results and experimental data for InP and Ge crystals irradiated by ultrafast heavy ions indicates the adequacy of the model, with which one can evaluate the temperature of the local area near the ion trajectory, as well as the diameters of the molten region and experimentally observed track region. The diameter of the cylindrical molten region that forms along the path of 250-MeV Xe^+ ions in InP is predicted to be 20 nm, and the measured cross-sectional diameters of the tracks fall into the range 7–15 nm.

INTRODUCTION

To date, several approaches to describing track formation in solids under high-energy ion implantation have been suggested. The most popular ones are the temperature peak model and the model of Coulomb explosion [1], the model of shock wave generation [2], the model of modified lattice potential [3], and the model of the plasma core of a track [4].

Most experimental data concerning the generation of defects and tracks in metals, semiconductors, and insulators irradiated by ultrafast heavy ions are discussed under the assumption that a highly excited region forms along the ion trajectory, which melts and then is rapidly quenched. The physicomathematical description of these processes (the so-called temperature peak model) was first proposed in [5]. To date, this model has been applied to advantage for the explanation of the track formation effect in a number of metals [6] and insulators [7]. In this work, we for the first time consider track formation in semiconductors on a quantitative basis in terms of this model.

DESCRIPTION OF THE MODEL

The temperature peak model takes into account the fact that a heavy fast ion, when passing through a material, gives up most of the energy to electrons (for ions with a kinetic energy $E \geq 1$ MeV/u, the energy transferred to the electron subsystem exceeds that imparted to the nuclear subsystem by three of four orders of magnitude). Energy transfer from the electron subsystem to the lattice is assumed to proceed largely via electron–phonon interaction. Since the time taken to establish thermodynamic equilibrium in the electron and nuclear subsystems is much smaller than the mutual relaxation time, it is valid to speak of lattice temperature and electron temperature. A track is assumed to form when the temperature of a local target area rises to the melting point.

Thus, for the electron gas and atomic lattice of a crystal, we can write the following set of coupled equations for heat fluxes in radial coordinates:

$$C_e(T_e)\frac{\partial T_e}{\partial t} = \frac{1}{r}\frac{\partial}{\partial r}\left[rK_e(T_e)\frac{\partial T_e}{\partial r}\right] - g(T_e - T_a) + A(r, t),$$
$$C_a(T_a)\frac{\partial T_a}{\partial t} = \frac{1}{r}\frac{\partial}{\partial r}\left[rK_a(T_a)\frac{\partial T_a}{\partial r}\right] + g(T_e - T_a). \quad (1)$$

Here, C_e, K_e and C_a, K_a are the specific heats and thermal conductivities of the electron and atomic subsystems, respectively; T_e and T_a are the respective temperatures of the electrons and lattice; g is the coefficient that characterizes electron-to-lattice energy transfer (the coefficient of electron–phonon coupling); and A is the energy density delivered to the electron subsystem.

The energy density $A(r, t)$ is expressed as [7]

$$A(r, t) = b\frac{dE}{dz}\exp\left(-\frac{(t - t_0)^2}{2t_0^2}\right)F(r), \quad (2)$$

where dE/dz are the linear inelastic energy losses of the ion, t_0 is the time taken for the electrons to reach the equilibrium distribution (for δ electrons of medium energy, the mean free time is $t_0 = (1–5) \times 10^{-15}$ s),

1063-7842/03/4806-0717$24.00

r_0 is the width of the highly excited region, b is a normalizing constant such that the relationship

$$\int_{t=0}^{\infty}\int_{r=0}^{r_m} A(r,t)2\pi r\,dr\,dt = dE/dz \quad (3)$$

holds, r_m is the maximal range of δ ions, and $F(r)$ is the spatial distribution of the δ electron energy in a stopping medium given in [8].

The coefficient of electron–phonon coupling is a free parameter and can be expressed through other quantities. For semiconductors, the value of g can be determined in two ways. One is the use of the formula for metals [9], which applies when the lattice temperature is not much less than the Debye temperature:

$$g = \frac{\pi^2 m_e n_e v_s^2}{6\tau_e(T_e)T_e}, \quad (4)$$

where m_e is the mass of an electron, n_e is the electron concentration, v_s is the velocity of sound in the target material, and τ_e is the mean free time of electrons.

However, the rigorous determination of $\tau_e(T_e)$ still remains a challenge today. For semiconductors, one may set τ_e to be constant, as a first approximation, and equal to 10^{-16}–10^{-15} s [10].

The other way is to express g through the mean free path length λ:

$$\lambda^2 = D_e C_e/g = K_e/g, \quad (5)$$

where D_e is the coefficient of thermal diffusion.

In [7], for various inorganic insulators with the energy gap E_g varying from nearly zero to 12 eV, an empirical relationship $\lambda(E_g)$ was derived. This relationship can probably be applied to finding g in the case of semiconductors.

Experimental temperature dependences of the lattice specific heat, $C_a(T_a)$, and thermal conductivity, $K_a(T_a)$, for Ge and InP are known (see, e.g., [11]). It should be noted, however, that the thermophysical parameters also depend on the dopant concentration, as well as on the type and concentration of defects.

The specific heat and thermal conductivity of the electron subsystem in semiconductor crystals are, strictly speaking, unknown. For metals, the dependences $C_e(T_e)$ and $K_e(T_e)$ can be found from the theory of free electron gas: the specific heat of metal electrons is a linear function of T_e at low T_e and becomes constant,

$$C_e = 3/2 n_e k_B, \quad (6)$$

at temperatures above the Fermi temperature (k_B is the Boltzmann constant).

The specific heat of electron gas is given by $K_e(T_e) = C_e(T_e)D_e(T_e)$. For insulators, the values of $C_e(T_e)$ and $K_e(T_e)$ were obtained [7] under the assumption that hot electrons in the conduction band of an insulator behave like hot electrons in a metal. In this case, C_e is given by (6). If the concentration of electrons involved in the diffusion transfer of energy is approximately 5×10^{22} cm^{-3}, C_e may be set equal to 1 J/(cm^3 K). The value of K_e is determined from the model of free electron gas on the assumption that the mean free path in the presence of electron–electron collisions is minimal and equal to the interatomic spacing. Then, $K_e \approx 2$ J/(cm K). We believe that these values of C_e and K_e can be used to roughly evaluate temperature peaks in semiconductor crystals.

The linear inelastic energy losses of an ion were calculated by the Bethe formula

$$dE/dz = \frac{4\pi N e^4 Z_{eff}^2}{m_e V^2} Z_b \ln\left[\frac{2m_e V^2}{I}\right], \quad (7)$$

where N is the concentration of target atoms; e is the electron charge; Z_b is the number of bound electrons per target atom; Z_{eff} is the effective charge of an incident ion, which depends on the energy of the ion; and I is the mean ionization potential of target atoms.

The effective charge of the ion is calculated by the formula [12]

$$Z_{eff} = Z_1(1 - \exp[-V/(V_0 Z_1^{2/3})]), \quad (8)$$

where Z_1 is the atomic number of the ion and $V_0 = 2\pi e^2/h$.

For two-component targets, the mean ionization potential I was evaluated by using the Bragg rule. It gives a fairly accurate value for the stopping power of a two-component target, expressing it through the partial stopping powers of the components (without allowance for chemical bonding):

$$S_e(A_nB_m) = nS_e(A) + mS_e(B), \quad (9)$$

where A_nB_m is a two-component material (n and m are the number of related atoms) and $S_e = 1/N(dE/dz)$ is the electron stopping cross section. With (7) and (9), we have

$$Z_b \ln\left[\frac{2m_e V^2}{I}\right] = Z_b^{(1)} \ln\left[\frac{2m_e V^2}{I^{(1)}}\right] + Z_b^{(2)} \ln\left[\frac{2m_e V^2}{I^{(2)}}\right], \quad (10)$$

from which dE/dz can be found.

Thus, we determined the physical parameters and coefficients that are necessary to solve the set of Eqs. (1) for semiconductor materials. The numerical solution of this set is described at length elsewhere [13].

RESULTS OF CALCULATION

To compare the model with earlier experimental data for semiconductor crystals [14–16], we calculated time and space temperature profiles for the electron and lattice subsystems near ion trajectories in InP irradiated by 250- and 340-MeV xenon ions and in Ge irradiated

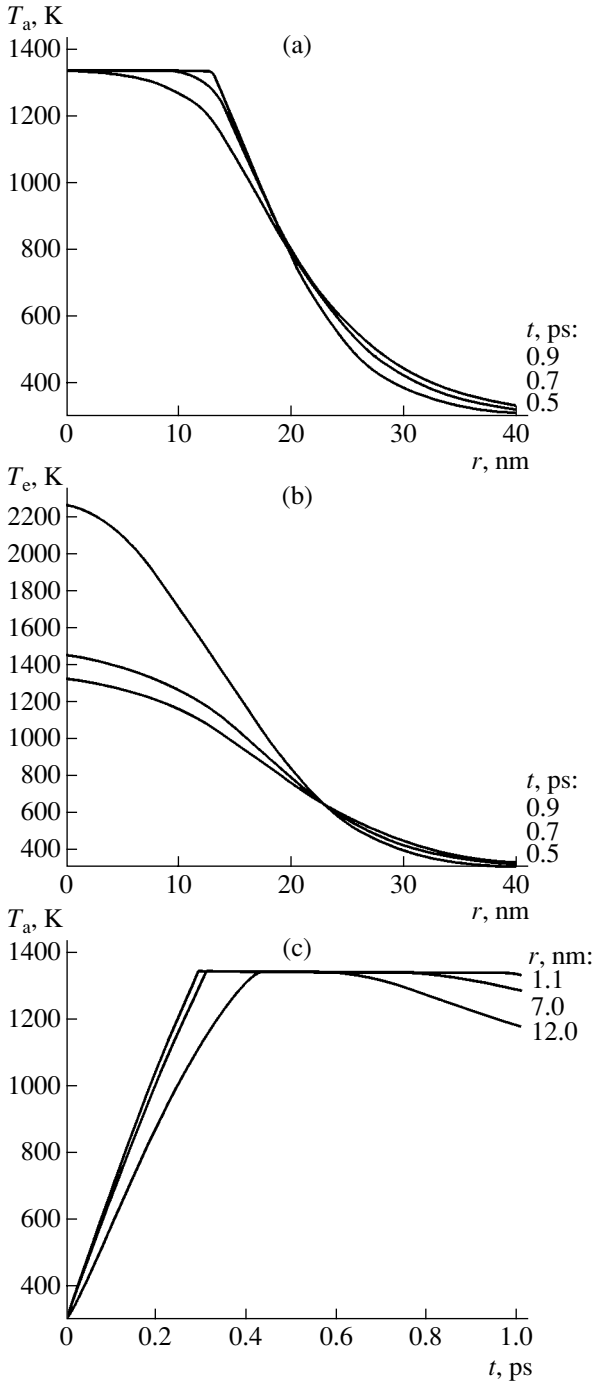

Fig. 1. Space and time dependences of the temperature for the InP crystals irradiated by Xe^+ ions with E = 250 MeV for $(dE/dx)_e$ = 23 keV/nm. The coefficient g is calculated by formula (4). (a) Lattice temperature vs. distance to the ion trajectory, (b) electron temperature vs. distance to the ion trajectory, and (c) lattice temperature vs. time.

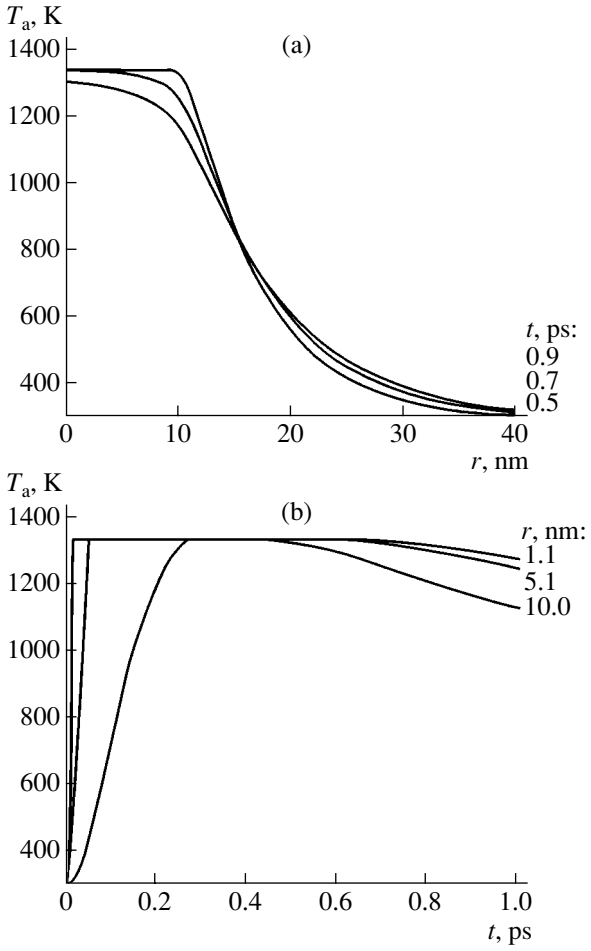

Fig. 2. The same as in Fig. 1 for g calculated by formula (5). (a) Lattice temperature vs. distance to the ion trajectory and (b) lattice temperature vs. time.

by 710-MeV bismuth ions and 1.3-GeV uranium ions. The initial temperature of the crystals was assumed to be equal to room temperature. Inelastic energy losses were calculated for an ion located at a depth of 1 μm. The heat of melting was ignored; that is, we assumed that heating to the melting point will suffice to form a track. The basic results are shown in Figs. 1–4.

The temperature distributions for InP irradiated by 250-MeV Xe^+ ions with the electron–phonon coupling coefficient g calculated by (4) and (5) are shown in Figs. 1 and 2, respectively. With $g \sim I/T_e$, the maximal radius of the molten zone is roughly equal to 12 nm, the melting point is reached within $\approx 3 \times 10^{-13}$ s, and the molten region persists for about 6×10^{-13} s. If g is assumed to be T_e independent, the lattice temperature rises to the melting point much faster (within $\approx 4 \times 10^{-14}$ s at a distance of 5 nm from the ion trajectory) and the molten region radius becomes ≈10 nm. Obviously, for time intervals less than 10^{-13} s, the notion of temperature becomes invalid. In this case, the term "quasi-temperature" seems to be more appropriate. From here on, g will be evaluated by (4), since the molten zone dimensions thus found differ insignificantly. For the InP crystals irradiated by 340-MeV Xe^+ ions, the maximal radius of the molten zone estimated as ≈13 nm (Fig. 3).

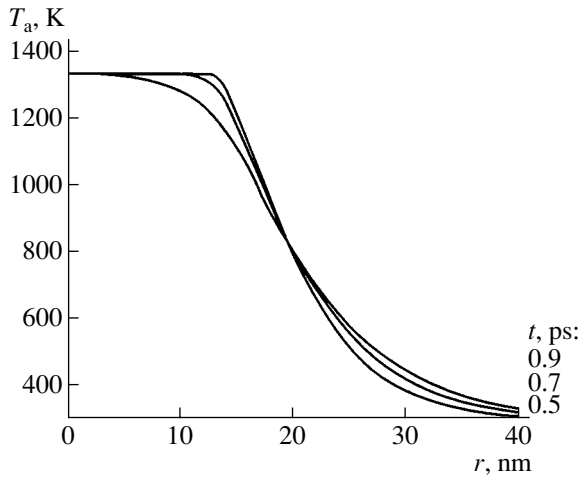

Fig. 3. Lattice temperature vs. distance to the ion trajectory for InP crystals irradiated by Xe^+ ions with E = 340 MeV. $(dE/dx)_e$ = 23.9 keV/nm.

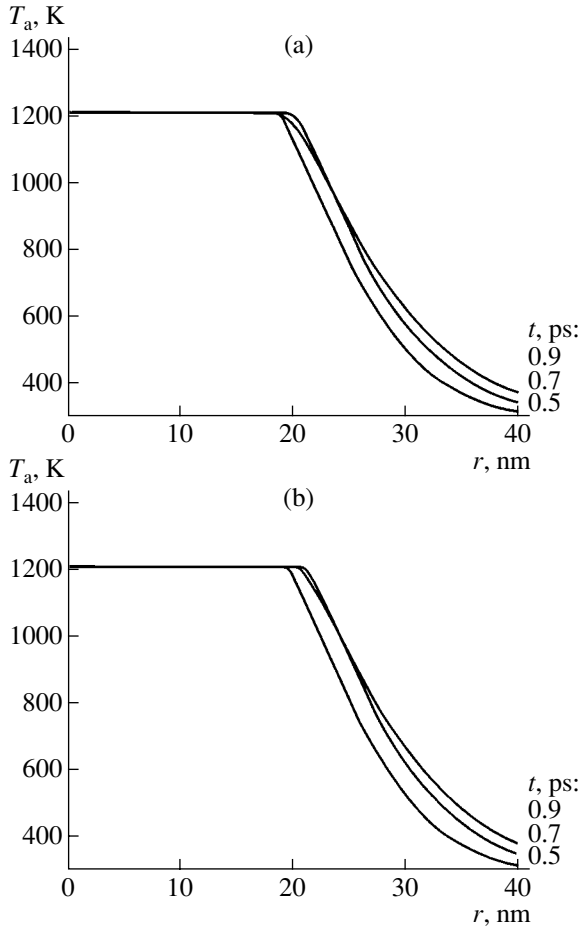

Fig. 4. Lattice temperature vs. distance to the ion trajectory for Ge crystals irradiated by (a) Bi^+ ions with E = 0.710 GeV for $(dE/dx)_e$ = 37 keV/nm and (b) U^+ ions with E = 1.3 GeV for $(dE/dx)_e$ = 42 keV/nm.

The TEM study of InP crystals irradiated by Xe^+ ions with E = 250 MeV [14] revealed a region containing continuous tracks of diameter ranging from 7 to 15 nm at a depth of 0.1–1.0 μm. The calculated and experimentally found dimensions of the molten zone in InP are in satisfactory agreement. The calculated dimensions of the tracks are somewhat higher. This may be associated with the fact that, when calculating the energy density given up to the electron subsystem of the crystal, we assumed that the entire energy imparted to electrons is confined in some local region around the ion trajectory and is spent on heating the atomic subsystem. Actually, however, some of the fast δ electrons that have a considerable transverse velocity component go too far away from the initially excited region to be attracted back by the Coulomb field of the remaining (unneutralized) charge, because the rate of neutralization in semiconductors is very high. These fast electrons carry away no less than one-third of the energy given up to the electron subsystem of the crystal [4].

Figure 4 shows the calculated time profiles of the lattice temperature for the Ge crystals irradiated by Bi^+ ions with E = 710 MeV and U^+ ions with E = 1.3 GeV. The maximal radii of the molten zones are 20 and 21 nm, respectively. At the same time, the TEM examination of Bi^+- and U^+-irradiated Ge crystals, as well as the selective etching of these crystals, revealed discontinuous tracks of diameter from 10 to 20 nm at depths of 12–15 μm from the surface [16]. For the Ge crystals, the discrepancy between the analytical and experimental results for the track radii is somewhat greater than for the InP crystals.

It should be noted that, both in InP and Ge, the maximal track density observed in experiments is much lower (by roughly two orders of magnitude) than the ion fluence. Moreover, at fluences below a threshold value (5×10^{12} cm^{-2} for InP irradiated by 250-MeV Xe^+ ions), continuous tracks do not form at all; instead, small defects and clusters appear. This allows us to suggest that not only the melting of the local zone but also the solidification features of the melt around the ion trajectory is of importance for track formation.

In such semiconductors as Ge, Si, and III–V compounds, the molten zone surrounded by the crystalline matrix solidifies very rapidly. Hence, in the case of low ion fluences, where each molten center is surrounded by the perfect lattice, perfect epitaxial recrystallization may occur. Yet each event of such recrystallization generates a number of lattice defects. With the ion fluence increased to $n \times 10^{11}$–$n \times 10^{12}$ cm^{-2}, when surface layers of the crystal experience multiple recrystallization, the density of defects grows to the point where they may disturb the epitaxial recrystallization of the molten zone. In this case, continuous and discontinuous tracks are generated. When the values of $(dE/dx)_e$ are high (as in the case of fullerene-irradiated Ge and Si [17, 18]), each individual ion forms a track without preirradiation.

CONCLUSIONS

In terms of the temperature peak model, we calculated the time and space profiles of the temperature of the local zones near ion trajectories in InP crystals irradiated by 250- and 340-MeV Xe^+ ions and in Ge crystals irradiated by 0.71-GeV Bi^+ ions and 1.3-GeV U^+ ions. From these profiles, we calculated the transverse dimensions of the molten zones, which are in good agreement with the radii of tracks observed experimentally. This leads us to conclude that the track formation model considered in this paper fits experimental data for semiconductor crystals [14–16] reasonably well.

REFERENCES

1. R. L. Fleischer, MRS Bull. **20** (12), 17 (1995).
2. I. S. Bitensky and E. S. Parilis, Nucl. Instrum. Methods Phys. Res. B **21**, 26 (1987).
3. C. C. Watson and T. A. Tombrello, Radiat. Eff. **89**, 263 (1983).
4. V. A. Belyĭ and F. F. Komarov, Zh. Tekh. Fiz. **68** (9), 42 (1998) [Tech. Phys. **43**, 1048 (1998)].
5. I. M. Lifshits, M. I. Kaganov, and L. V. Tanatarov, At. Energ. **16**, 391 (1959).
6. Z. G. Wang, C. Dufour, E. Paumier, *et al.*, J. Phys.: Condensed Matter **6**, 6733 (1994).
7. M. Toulemonde, C. Dufour, A. Meftah, *et al.*, Nucl. Instrum. Methods Phys. Res. B **166–167**, 903 (2000).
8. M. P. R. Waligorski, R. N. Hamm, and R. Katz, Int. J. Rad. Appl. Instrum. D **11**, 309 (1986).
9. M. I. Kaganov, I. M. Lifshits, and L. V. Tanatarov, Zh. Éksp. Teor. Fiz. **31**, 232 (1956).
10. C. Dufour, E. Paumier, and M. Toulemonde, Radiat. Eff. Defects Solids **126**, 119 (1993).
11. M. Landolt and J. Börnstein, in *Numerical Data and Functional Relationships in Science and Technology: New Series* (Heidelberg, Berlin, 1982), Vol. 17a.
12. W. Booth and I. Grant, Nucl. Phys. **63**, 481 (1965).
13. A. A. Kamarou, D. A. Malafei, and V. S. Shchehlik, in *Proceedings of the 2nd International Conference "Finite-Difference Methods: Theory and Applications" (CFDM-98), Minsk, 1998*, Vol. 2, pp. 56–62.
14. O. Herre, W. Wesch, E. Wendler, *et al.*, Phys. Rev. B **58**, 4832 (1998).
15. P. I. Gaiduk, F. F. Komarov, V. S. Tishkov, *et al.*, Phys. Rev. B **61**, 15785 (2000).
16. L. A. Vlasukova, F. F. Komarov, P. I. Gaiduk, *et al.*, Vacuum (in press).
17. A. Colder, O. Marty, B. Canut, *et al.*, Nucl. Instrum. Methods Phys. Res. B **174**, 491 (2001).
18. B. Canut, N. Bonardi, and S. M. Ramos, Nucl. Instrum. Methods Phys. Res. B **146**, 296 (1998).

Translated by V. Isaakyan

Technical Physics, Vol. 48, No. 6, 2003, pp. 722–726. Translated from Zhurnal Tekhnicheskoĭ Fiziki, Vol. 73, No. 6, 2003, pp. 61–65.

SOLIDS

Dislocation Structure and Transport Properties of Low-Angle Tilt Boundaries in High-Temperature Superconductors

S. V. Bobylev and I. A. Ovid'ko

Institute of Problems of Mechanical Engineering, Russian Academy of Sciences, Vasil'evskiĭ Ostrov, Bol'shoĭ pr. 61, St. Petersburg, 199178 Russia

e-mail: ovidko@def.ipme.ru

Received November 27, 2002

Abstract—A theoretical model that describes split grain-boundary dislocations in low-angle tilt boundaries of high-temperature superconductors is suggested. It is shown that the dissociation of dislocations in low-angle tilt boundaries is usually accompanied by a decrease in their elastic energy and causes an increase in the critical current density across the boundaries in high-temperature superconductors. *© 2003 MAIK "Nauka/Interperiodica".*

INTRODUCTION

A decrease in the critical current density J_{cr} due to the presence of grain boundaries in high-temperature polycrystalline superconductors has been the subject of extensive study [1–18]. The effect of grain boundaries (GBs) on J_{cr} is related to stress fields due to GB dislocations [6–8], as well as crystal order violation [6–8], a change in the chemical composition [9–11], and charge distribution nonuniformity [5, 12–15] near grain boundaries. The symmetry of the order parameter and the faceted structure of grain boundaries may also be factors [16]. In the models [5–15], which theoretically describe GB transport properties, low-angle GBs are considered as periodic walls of perfect lattice dislocations. However, such boundaries in high-temperature superconductors often undergo structural transformations, which affect noticeably their transport properties (see, e.g., experimental data in [19–22] and theoretical models [23–27]). According to the model [26], each dislocation making up [100] low-angle tilt boundaries observed experimentally [19] dissociates into three partial dislocations. Recently, the high-resolution transmission electron microscopy examination of YBaCuO has revealed [001] low-angle tilt grain boundaries in which each perfect dislocation is dissociated into two partial dislocations. Such a dissociation cannot help affecting the superconducting transport properties of grain boundaries. In this paper, we theoretically describe the dissociation of perfect GB dislocations in [001] low-angle tilt boundaries into partial dislocations (observed experimentally in [21]) and estimate the influence of the dissociation on the critical current density across low-angle GBs in high-temperature superconductors.

ENERGY CHARACTERISTICS OF DISLOCATION DISSOCIATION IN LOW-ANGLE TILT GRAIN BOUNDARIES

Let us consider the dissociation of perfect lattice dislocations that are structural elements of a [001] low-angle tilt boundary in the (110) plane in $YBa_2Cu_3O_{7-\delta}$ superconductor. An approach to evaluate the energy gain from the dissociation of perfect edge dislocations is as follows. Consider a periodic wall of perfect edge dislocations that simulates a low-angle tilt grain boundary and the transformation of such a dislocation wall (Fig. 1a) into a new dislocation structure (Fig. 1b) in which one of the perfect dislocations dissociates into two partial dislocations with a stacking fault in between. The new dislocation structure is a usual dislocation wall that contains a structural component of a low-angle tilt boundary (wall) experimentally observed in [21] and consisting of dissociated dislocations (Fig. 1c). For convenience, we represent such a dissociated dislocation configuration as a perfect dislocation and two dipoles made of partial dislocations (Fig. 2). In this representation, the lower dislocation of the upper dipole and the upper dislocation of the lower dipole are located in the same position as the initial perfect dislocation. As a consequence, the superposition of these three dislocations (one perfect and two partial ones) is

1063-7842/03/4806-0722$24.00

equivalent to the absence of a dislocation at the center of the split configuration.

Let us find the difference between the total energies (per unit dislocation length) of the new dislocation structure (Fig. 1b) and a usual wall of dislocations (Fig. 1a). For brevity, a usual dislocation wall (Fig. 1a) and the new dislocation structure (Fig. 1b) will be referred to as structure 1 and structure 2, respectively. The energy of structure 1 has two components:

$$W_1 = RW_1^{\mathrm{el}} + \frac{R}{h_1}W_1^{\mathrm{c}}, \tag{1}$$

where R is the length of a tilt grain boundary, W_1^{el} is the elastic energy density of structure 1 per unit length, and W_1^{c} is the energy of the perfect (nondissociated) dislocation core.

The energy of structure 2 (Fig. 1b) can be represented as follows:

$$W_2 = RW_1^{\mathrm{el}} + \left(\frac{R}{h_1} - 1\right)W_1^{\mathrm{c}} + 2W_{\mathrm{dip}}^{\mathrm{el}} + 2W_2^{\mathrm{c}} + W_{\mathrm{int}}^{\mathrm{dip}} + 2W_{\mathrm{int}}^{\mathrm{dip}-b1} + 2(p - r_{0_2})\gamma. \tag{2}$$

Here, $W_{\mathrm{dip}}^{\mathrm{el}}$ is the intrinsic elastic energy of a dipole of partial dislocations, W_2^{c} is the energy of the partial dislocation core, $W_{\mathrm{int}}^{\mathrm{dip}}$ is the energy of dipole-dipole interaction, $W_{\mathrm{int}}^{\mathrm{dip}-b1}$ is the energy of elastic interaction between a dipole and structure 1, r_{0_2} is the radius of the partial dislocation core, and γ is the energy of the stacking fault between the partial dislocations in the extended dislocation configuration (Fig. 2). Using Eqs. (1) and (2), we find the change in the total energy of the system when structure 1 transforms into structure 2:

$$\Delta W_{1-2} = W_2 - W_1 = -W_1^{\mathrm{c}} + 2W_{\mathrm{dip}}^{\mathrm{el}} + 2W_2^{\mathrm{c}} + W_{\mathrm{int}}^{\mathrm{dip}} + 2W_{\mathrm{int}}^{\mathrm{dip}-b1} + 2(p - r_{0_2})\gamma. \tag{3}$$

The equation $\Delta W_{1-2} = 0$ defines the critical conditions for such a transformation.

Using the equations from [26] and their modifications for the terms appearing on the right of Eq. (3), we find

$$\Delta W_{1-2} = \frac{Gb^2}{4\pi(1-\nu)}\left(2Z_2 - \frac{B^2}{b^2}Z_1 + 2\ln\frac{2(p - r_{0_2})^2}{r_{0_2}(p + r_{0_2})} - 2\frac{B}{b}\ln\frac{1 - \cos\tilde{p}}{1 - \cos\tilde{r}_{0_1}}\right) + 2(p - r_{0_2})\gamma. \tag{4}$$

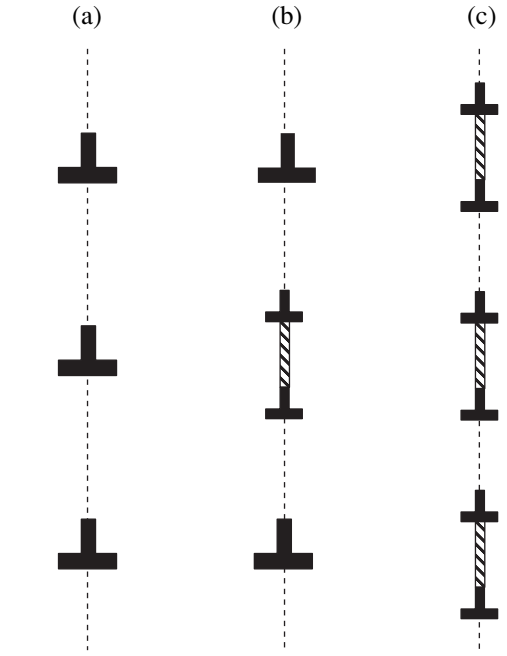

Fig. 1. Dislocation structures of low-angle tilt boundaries: (a) usual, (b) with one dissociated dislocation, and (c) with all dislocations dissociated.

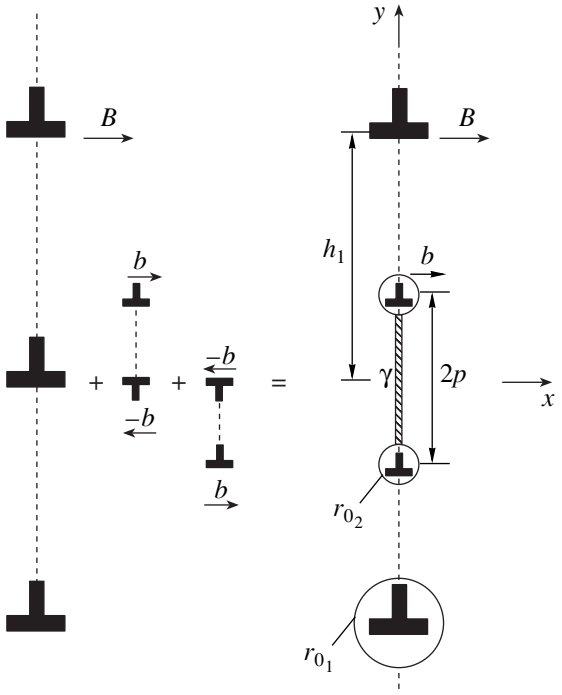

Fig. 2. Dissociated dislocation configuration represented as a perfect dislocation and two dipoles made of partial dislocations.

Here, Z_1 and Z_2 are dimensionless constants on the order of unity, $\tilde{p} = 2\pi p/h_1$, $\tilde{r}_{0_1} = 2\pi r_{0_1}/h_1$, r_{0_1} is the radius of the nondissociated dislocation core, and h_1 is

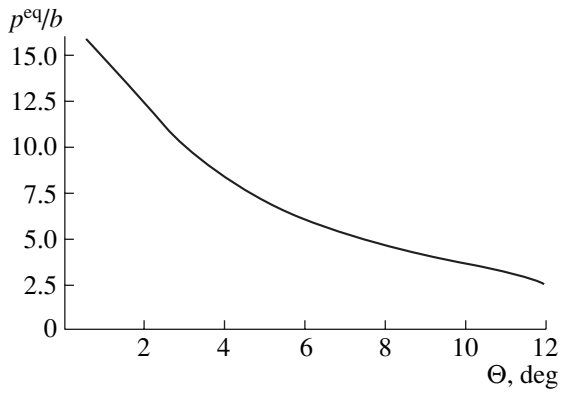

Fig. 3. Dependence of the equilibrium spacing p (in units of b) between neighboring partial dislocations on the GB misorientation angle Θ.

the period of the dislocation wall. The last term of Eq. (4) contains the stacking fault energy γ, which is a basic parameter of the problem. In [26], this energy was estimated as $\gamma = 7GB/[324\pi(1-\nu)]$. With the same estimator and using Eq. (4), we numerically studied the dependence of ΔW_{1-2} on the GB misorientation angle Θ (which is related to the parameters of dislocation structures through the Frank's relationship $B = 2h_1\sin(\Theta/2)$) for characteristic values of the parameters $B = 2b$, $Z_1 \approx Z_2 \approx 1$, $r_{0_1} \approx B$, and $r_{0_2} \approx b$. According to the calculations, the dissociation is energetically favorable (i.e., the energy change ΔW_{1-2} is negative) in a rather wide range of misorientations Θ from 0° to 12°. Also, the calculations indicate that the equilibrium spacing between neighboring partial dislocations (that is, the spacing corresponding to the highest gain $|\Delta W_{1-2}|$ in energy upon the transition from structure 1 to structure 2) decreases with increasing GB misorientation Θ (Fig. 3).

TRANSPORT PROPERTIES OF LOW-ANGLE BOUNDARIES WITH SPLIT DISLOCATION CONFIGURATIONS

Using the results of the previous section, we will evaluate a change in the superconductor critical current density passing across a tilt boundary when all the dislocations are dissociated. Following the approach [5, 15], we assume that the dislocations are surrounded by regions with a nonstoichiometric composition and anomalous charge distribution ("bad" regions in terms of superconductivity). In the framework of such an approach, the critical current density J_{cr} through a tilt boundary with a misorientation Θ can be written in the first approximation as follows [15]:

$$J_{cr}(\Theta)/J_{cr}(0^\circ) \approx \frac{1}{S}\int_S \exp\{-d(y,z)/\xi\}\,dydz. \qquad (5)$$

Here, d is the thickness of a bad region, ξ is the characteristic decay length, and S is the GB surface area. Denoting the critical currents through the tilt boundaries formed by nondissociated and dissociated dislocations (Figs. 1a, 1c) as $J_{cr}^{(1)}$ and $J_{cr}^{(2)}$, respectively, and taking into account that the properties of the dislocation structures under consideration do not depend on the coordinate z, we can express their ratio as

$$J_{cr}^{(2)}/J_{cr}^{(1)} = \frac{\int_S \exp\{-d^{(2)}(y)/\xi\}\,dy}{\int_S \exp\{-d^{(1)}(y)/\xi\}\,dy}, \qquad (6)$$

where $d^{(1)}$ and $d^{(2)}$ are the thicknesses of the bad regions in tilt grain boundaries in which all the dislocations are dissociated and nondissociated, respectively.

Obviously, due to the periodicity of the dislocation wall, it is sufficient to perform integration in (6) over one period.

For the boundary with nondissociated dislocations (Fig. 1a), the y dependence of $d^{(1)}$ (within a period of the wall) is given by the following relationships (in the coordinate system shown in Fig. 2) according to whether the bad regions overlap or not:

$$d^{(1)} = \begin{cases} \begin{cases} 2\sqrt{R^{(1)^2} - y^2} & \text{for} \quad -R^{(1)} \le y \le R^{(1)} \\ 0 \quad \text{for} \quad -R^{(1)} \le y \le h_1 - R^{(1)} \end{cases} \\ \text{if} \quad R^{(1)} \le h_1/2 \quad \text{(no overlap)} \\ \\ 2\sqrt{R^{(1)^2} - y^2} \quad \text{for} \quad -h_1/2 \le y \le h_1/2 \\ \text{if} \quad R^{(1)} > h_1/2 \quad \text{(overlap)}. \end{cases} \qquad (7)$$

Here, $R^{(1)}$ is the radius of the bad region of a nondissociated dislocation.

In the case of the boundary with the dissociated dislocations (Fig. 1c), the situation becomes more complicated, since bad regions of two types may overlap: regions with partial dislocations formed from various perfect dislocations and those near the cores of partial dislocations originating from the same perfect dislocation. The calculations show that, as the misorientation Θ increases, the overlap of the first type (Fig. 4a) initially occurs, while the overlap of the second type (Fig. 4b) takes place at high misorientations. Thus, an

expression for $d^{(2)}$ can be written as follows:

$$d^{(2)} = \begin{cases} \begin{cases} 2\sqrt{R^{(2)^2} - y^2} \\ \text{for} \quad -R^{(2)} \le y \le R^{(2)}, \\ 2\sqrt{R^{(2)^2} - (y-2p)^2} \\ \text{for} \quad -R^{(2)} + 2p \le y \le R^{(2)} + 2p, \\ 0 \quad \text{for} \quad R^{(2)} \le y \le -R^{(2)} + 2p \\ \text{or} \quad R^{(2)} + 2p \le y \le h_1 - R^{(2)} \end{cases} \\ \text{if} \quad R^{(2)} < p \quad \text{and} \quad R^{(2)} < h_1/2 - p \\ \begin{cases} 2\sqrt{R^{(2)^2} - y^2} \\ \text{for} \quad -h_1/2 + p \le y \le R^{(2)}, \\ 2\sqrt{R^{(2)^2} - (y-2p)^2} \\ \text{for} \quad -R^{(2)} + 2p \le y \le h_1/2 + 2p, \\ 0 \quad \text{for} \quad R^{(2)} \le y \le -R^{(2)} + 2p \end{cases} \\ \text{if} \quad R^{(2)} < p \quad \text{and} \quad R^{(2)} > h_1/2 - p \\ \begin{cases} 2\sqrt{R^{(2)^2} - y^2} \\ \text{for} \quad -R^{(2)} \le y \le p, \\ 2\sqrt{R^{(2)^2} - (y-2p)^2} \\ \text{for} \quad p \le y \le h_1/2 + p \end{cases} \\ \text{if} \quad R^{(2)} > p \quad \text{and} \quad R^{(2)} > h_1/2 - p. \end{cases} \tag{8}$$

Using Eqs. (7) and (8), as well as the estimates $R^{(i)} = \lambda r_{0_i}$, where $\lambda \approx 3$, and $\xi \approx 1.5$ nm [15], one can find the ratio of superconductor critical currents through the boundaries consisting of dissociated and nondissociated dislocations from formula (6). The results of calculations are presented in Fig. 5.

According to Fig. 5, the critical current in the dissociated configuration is always higher than in the nondissociated one. Geometrically, this is easy to explain by the smaller surface area of the bad regions in the case of the dissociated configuration (the surface area of two circles of equal radii is smaller than the surface area of one circle with a doubled radius). The slow increase in the ratio $J_{cr}^{(2)}/J_{cr}^{(1)}$ at small misorientations is explained by an increase in the fraction of the bad regions and by the above-mentioned advantage of the dissociated configuration, which becomes more and more significant as the volume occupied by the "good" regions decreases. The sharp increase in the ratio

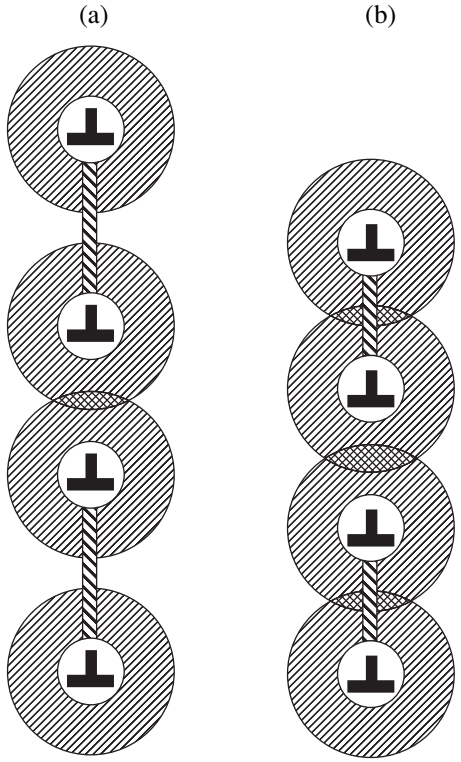

Fig. 4. Overlapping of bad regions in the dissociated configuration: (a) bad regions of partial dislocations formed when neighboring perfect dislocations dissociate at small misorientation angles and (b) bad regions of partial dislocations formed when one perfect dislocation dissociates at large misorientation angles.

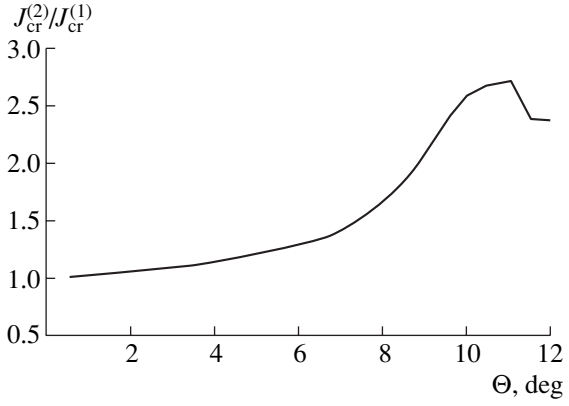

Fig. 5. Ratio of the critical currents across tilt GBs consisting of dissociated dislocation ($J_{cr}^{(2)}$) and nondissociated dislocations ($J_{cr}^{(1)}$) as a function of GB misorientation angle Θ.

$J_{cr}^{(2)}/J_{cr}^{(1)}$ near $\Theta = 7°$ is related to overlapping in the dissociated configuration (Fig. 4a). In this case, the rate of fall of the critical current, which is controlled by the bad regions, slows down with growing misorientation angle in the dissociated configuration. The slower

growth of the current ratio, which starts at $\Theta = 9.5°$, is explained by bad region overlapping in the nondissociated configuration.

CONCLUSIONS

Thus, in a wide range of tilt GB misorientations Θ (from 0 to 12°), the dissociation of perfect dislocations into partial ones is energetically favorable. This is in good agreement with experimental data [19–23]. Generally, the dissociation of dislocations increases the critical current density through a grain boundary. This assertion, which follows from our theoretical consideration, should be taken into account, e.g., in calculations of the critical current in polycrystalline superconductors in terms of the percolation theory [17].

ACKNOWLEDGMENTS

This work was supported in part by the Russian Foundation for Basic Research (project no. 01-02-16853) and the Office of US Naval Research (project no. 00014-99-1-0896).

REFERENCES

1. D. Dimos, P. Chaudhari, J. Mannhart, *et al.*, Phys. Rev. Lett. **61**, 219 (1988).
2. D. Dimos, P. Chaudhari, and J. Mannhart, Phys. Rev. B **41**, 4038 (1990).
3. Z. G. Ivanov, P.-A. Nilsson, D. Winkler, *et al.*, Appl. Phys. Lett. **59**, 3030 (1991).
4. M. Prester, Supercond. Sci. Technol. **11**, 333 (1998).
5. N. D. Browning, E. M. James, K. Kyosuke, *et al.*, Rev. Adv. Mater. Sci. **1**, 1 (2000).
6. M. F. Chisholm and S. J. Pennycook, Nature **351**, 47 (1991).
7. D. Agassi, C. S. Pande, and R. A. Masumura, Phys. Rev. B **52**, 16237 (1995).
8. A. Gurevich and E. A. Pashitskii, Phys. Rev. B **57**, 13878 (1998).
9. D. M. Kroeger, A. Choudhury, J. Brynestad, *et al.*, J. Appl. Phys. **64**, 331 (1988).
10. A. M. Campbell, Supercond. Sci. Technol. **2**, 287 (1989).
11. H. Betouras and R. Joynt, Physica C **250**, 256 (1995).
12. A. Schmehl, B. Goetz, R. R. Schulz, *et al.*, Europhys. Lett. **47**, 110 (1999).
13. H. Hilgenkamp and J. Mannhart, Appl. Phys. Lett. **73**, 265 (1998).
14. J. Mannhart and H. Hilgenkamp, Supercond. Sci. Technol. **10**, 880 (1997).
15. I. A. Ovid'ko, Mater. Sci. Eng. A **313**, 207 (2001).
16. H. Hilgenkamp, J. Mannhart, and B. Mayer, Phys. Rev. B **53**, 14586 (1996).
17. R. Haslinger and R. Joynt, Phys. Rev. B **61**, 4206 (2000).
18. M. Amer, J. Maguire, L. Cai, *et al.*, J. Appl. Phys. **89**, 8030 (2001).
19. M. F. Chisholm and D. A. Smith, Philos. Mag. A **59**, 181 (1989).
20. I.-F. Tsu, J.-L. Wang, D. L. Kaiser, *et al.*, Physica C **306**, 163 (1998).
21. H. Kung, J. P. Hirth, S. R. Foltyn, *et al.*, Philos. Mag. Lett. **81**, 85 (2001).
22. I.-F. Tsu, S. E. Babcock, and D. L. Kaiser, J. Mater. Res. **11**, 1383 (1996).
23. E. Z. Meilikhov, Physica C **271**, 277 (1996).
24. S. A. Kukushkin, I. A. Ovid'ko, and A. V. Osipov, Pis'ma Zh. Tekh. Fiz. **26** (7), 36 (2000) [Tech. Phys. Lett. **26**, 609 (2000)].
25. I. A. Ovid'ko, J. Phys.: Condens. Matter **13**, L97 (2001).
26. M. Yu. Gutkin and I. A. Ovid'ko, Phys. Rev. B **63**, 064515 (2001).
27. S. V. Bobylev, I. A. Ovid'ko, and A. G. Sheinerman, Phys. Rev. B **64**, 224507 (2001).

Translated by M. Lebedev

Technical Physics, Vol. 48, No. 6, 2003, pp. 727–735. Translated from Zhurnal Tekhnicheskoĭ Fiziki, Vol. 73, No. 6, 2003, pp. 66–75.

SOLIDS

Numerical Study of Shock Compression of Graphite and Its Conversion to Diamond in Conical Targets

I. V. Lomonosov*, V. E. Fortov*, A. A. Frolova, K. V. Khishchenko*, A. A. Charakhchyan**, and L. V. Shurshalov****

** Institute for High Energy Densities, Joint Institute for High Temperatures, Russian Academy of Sciences, ul. Izhorskaya 13/19, Moscow, 125412 Russia*

*** Dorodnitsyn Computing Center, Russian Academy of Sciences, Moscow, 119991 Russia*

e-mail: chara@ccas.ru

Received December 10, 2002

Abstract—The shock compression of carbon placed into lead conical targets that are subjected to a blow from aluminum strikers moving at a velocity of 4 km/s is numerically studied. The statement of the problem includes the hydrodynamic equations for compressed media, semi-empirical wide-range equations of state, and a kinetic model of nonequilibrium graphite-to-diamond conversion that is calibrated with known experimental data. All stages of the process are analyzed, starting from the cumulative effect at the symmetry axis of the target to the release of the sample due to the rarefaction wave coming from the rear side of the striker with allowance for spalling. *© 2003 MAIK "Nauka/Interperiodica".*

INTRODUCTION

Graphite-to-diamond conversion under shock compression has been studied over several decades (see, e.g., [1–16]). These and most other studies dealt with plane shock waves. The idea of using a conical solid target for artificial diamond production was first put forward in [17]. An evident disadvantage of such targets in comparison to plane ones is the spatial nonuniformity of compression. However, conical targets offer a number of valuable advantages, which may improve the efficiency of diamond synthesis in graphite areas that are close to the symmetry axis of the target. The process can be controlled by varying the size and shape of the target, the striker and target material, and the velocity and thickness of the striker. Due to the cumulative effect, the pressure near the symmetry axis of a conical target can be increased for a short time to a value that is unattainable with a plane target at the same striker velocity. In addition, one can easily make a specially shaped outlet at the top of the conic recess to extract test samples and provide conditions for their nondestructive unloading, cooling, and deceleration.

The pioneering work on numerical simulation [17] did not take into account graphite-to-diamond conversion and considered only the cumulative effect at the early stage of compression. In this study, new semi-empirical wide-range equations of state for carbon polymorphous modifications are constructed and the statement of the problem includes a kinetic model of graphite-to-diamond phase transformation. For the most part, our simulation is based on a model corresponding to the experiment [10] on the shock compression of highly oriented quasi-single-crystalline graphite. The influence of the graphite sort on the character of its flow is illustrated by separate computations within a model corresponding to the experiment [9] with isotropic artificial graphite. We simulated all stages of the process, including sample release. Two types of conic targets are simulated: with a closed recess and with a long cylindrical channel at the end.

Let us note several points that are disregarded in our consideration. It is well known that the shock compression of graphite may result in the formation of either cubic diamond [1, 2] or its hexagonal modification (lonsdalite) [4]. In our statement, no difference is made between these two phases: both are described by the same equation of state. Another drawback of our physical model is that it ignores the reverse diamond–graphite transformation (graphitization) [6, 11]. One problem in the adequate simulation of the reverse process lies in the much differing kinetic characteristics of graphitization for diamond and lonsdalite [13]. Note that heat conduction is also left beyond the scope of our consideration because of the lack of reliable data on the transport properties of carbon over the required range of thermodynamic parameters. One more "roughness" is the use of the hydrodynamic approximation, which disregards the shear strength of the materials. In general, such an approach is reasonable in view of extremely high pressures involved. At the same time, this factor may somewhat affect the development of the instability of the diamond–target interface. Yet despite the drawbacks indicated above, this model allows one to study the flow characteristics as functions of problem parameters and substantiate the choice of the parameters for a physical experiment.

1063-7842/03/4806-0727$24.00

EQUATIONS OF STATE FOR DIAMOND AND GRAPHITE

Of high importance in the numerical simulation of nonstationary hydrodynamic processes is the adequate specification of the thermodynamic properties of a material. The reliability of numerical results to a great extent depends on the equation of state. In this study, semi-empirical equations of state [16, 18, 19] are used to describe the thermodynamic characteristics of the polymorphous modifications of carbon over a wide range of densities and temperatures. The validity of these equations can be seen by comparing the calculated shock (Hugoniot) adiabats with data points obtained in experiments on the dynamic compression of diamond [20, 21] and graphite [7, 14–16, 21] samples with different initial densities (Fig. 1). It is evident that, for both polymorphous modifications of carbon, the experimental values of the shock compressibility are fitted well by the curves over the entire interval of stress intensities corresponding to the solid state of the material behind the wave front.

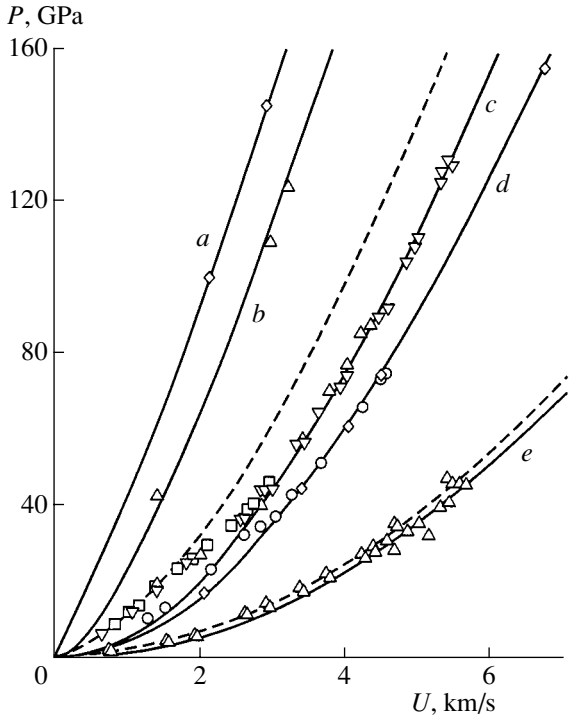

Fig. 1. Shock adiabats of diamond (solid curves) and graphite (dashed curves) with initial densities ρ_{00} = (*a*) 3.51, (*b*) 3.19, (*c*) 2.15, (*d*) 1.9, and (*e*) 1.011 g/cm^3. Experiment: ◇ [20], △ [21], ▽ [7], □ [14], and ○ [15, 16].

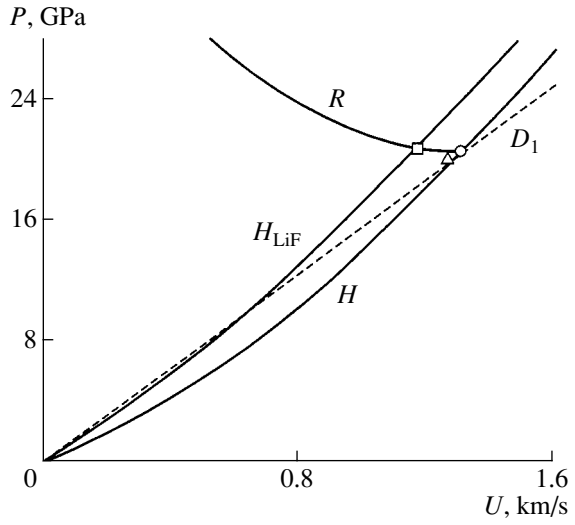

Fig. 2. (*H*) Shock adiabat of graphite with the initial density ρ_{00} = 2.259 g/cm^3 and (*R*) the loading curve of diamond with the initial state $P = P_1$ and $V = V_1$. H_{LiF}, shock adiabat of LiF [21]; D_1, ray $P = \rho_{00} D_1 U$; □, experiment [10]; △, estimate of the state on the shock adiabat of graphite [10]; and ○, calculation made in this study.

Let us consider the results of study [10], where a two-wave structure was observed in quasi-single-crystalline graphite subjected to dynamic loading under a pressure P = 26.5–41.0 GPa. In order to determine the state variables corresponding to the ultimate compression of graphite before its fast transformation into the high-pressure phase, the authors of [10] measured the velocity D_1 of the first shock wave in the sample, behind the front of which the carbon remains in the initial phase modification (with a pressure P_1, specific volume V_1, and mass velocity U_1), and the mass velocity U_R behind the front of the shock wave propagating over a transparent LiF obstacle after the first shock wave reaches the sample–obstacle interface (the velocity U_R also characterizes the motion of the high-pressure carbon phase behind the front of the shock wave reflected from LiF). The pressure P_R behind the reflected shock wave in the carbon can be determined from the measured value of U_R and the LiF shock adiabat given in [21]. Figure 2 shows the shock adiabat of graphite and the curve of repeat loading for forming diamond that were calculated based on our equations of state. It should be noted that the parameters of the initial point of transformation of the shock-compressed graphite into diamond that were calculated from the starting values of D_1 and U_R [10] with our equations of state for the crystalline modifications of carbon, $P_1 \simeq 20.6$ GPa and $U_1 \simeq 1.31$ km/s, somewhat differ from those found in [10] under the simplifying assumption concerning the repeat loading adiabat $P(U)$ for pressures between P_1 and P_R ($P_1 = 19.6 \pm 0.7$ GPa, $U_1 \simeq 1.27$ km/s).

STATEMENT OF THE PROBLEM AND NUMERICAL SIMULATION

The schematic of the process is shown in Fig. 3. A lead target (calculations were also carried out for other materials) occupies a half-space and has an axisymmetric cavity filled with graphite. Cavities of two shapes are considered: an ordinary cone (the dashed lines in Fig. 3) and a cone with an extracting cylindrical channel (solid lines). The radius of the cavity on the target surface is R, the cone angle is ϑ, and the radius of the

cylindrical channel is r_0. An aluminum striker of thickness H hits the target with a velocity $v = 4$ km/s.

We assume here that the target is sufficiently large, so that the influence of its lateral and rear sides can be disregarded, unlike the actual experimental conditions.

The flow is described by the hydrodynamic equations for compressible media

$$\frac{d\rho}{dt} + \rho \operatorname{div} \mathbf{U} = 0, \quad \rho \frac{d\mathbf{U}}{dt} + \nabla P = 0,$$
$$\rho \frac{d(E + \mathbf{U}^2/2)}{dt} + \operatorname{div} P\mathbf{U} = 0, \tag{1}$$

where $d/dt = \partial/\partial t + \mathbf{U}\nabla$ is the Lagrange derivative with respect to time, $\mathbf{U}$ is the velocity vector, ρ is the density, P is the pressure, and E is the specific internal energy.

In the case of aluminum and lead, the system of Eqs. (1) is closed by the wide-range multiphase equations of state for metals [22]. The thermodynamic properties of carbon are described within the model of nonequilibrium phase transformations [23], and each phase is characterized by the equations of state mentioned in the previous section:

$$P_1 = P_1(V_1, T), \quad E_1 = E_1(V_1, T),$$
$$P_2 = P_2(V_2, T), \quad E_2 = E_2(V_2, T), \tag{2}$$

where $V = \rho^{-1}$ is the specific volume and T is the temperature. Subscripts 1 and 2 refer to diamond and graphite, respectively.

Both carbon phases have the same temperature and pressure, while the internal energy and the specific volume of the mixture are additive:

$$P = P_1(V_1, T) = P_2(V_2, T),$$
$$E = \alpha E_1(V_1, T) + (1 - \alpha)E_2(V_2, T), \tag{3}$$
$$V = \alpha V_1 + (1 - \alpha)V_2,$$

where α is the weight fraction of the diamond phase.

Equations (1)–(3) are closed by an equation for α, which defines the phase transformation kinetics.

The main parameters of a phase transition in a shock wave are the transition time τ and the threshold pressure P_0 of the transition. Both parameters are known to depend strongly on the sort of graphite. Our investigation is centered largely on highly oriented graphite used in the experiment [10], for which $P_0 = 21$ GPa (see the previous section) and $\tau < 10$ ns.

The equation for α is taken in the form [9]

$$d\alpha/dt = 0 \quad \text{at} \quad P \le P_0,$$
$$d\alpha/dt = (1 - \alpha)(P/P_0 - 1)/\tau \tag{4}$$
$$\text{at} \quad P > P_0. \tag{5}$$

We assumed that $\tau = 5$ ns. For the targets considered ($R \ge 5$ mm), a further decrease in τ did not change the results, since the thickness of the phase transition

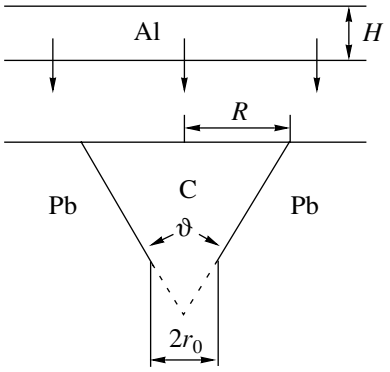

Fig. 3. Schematic of the process.

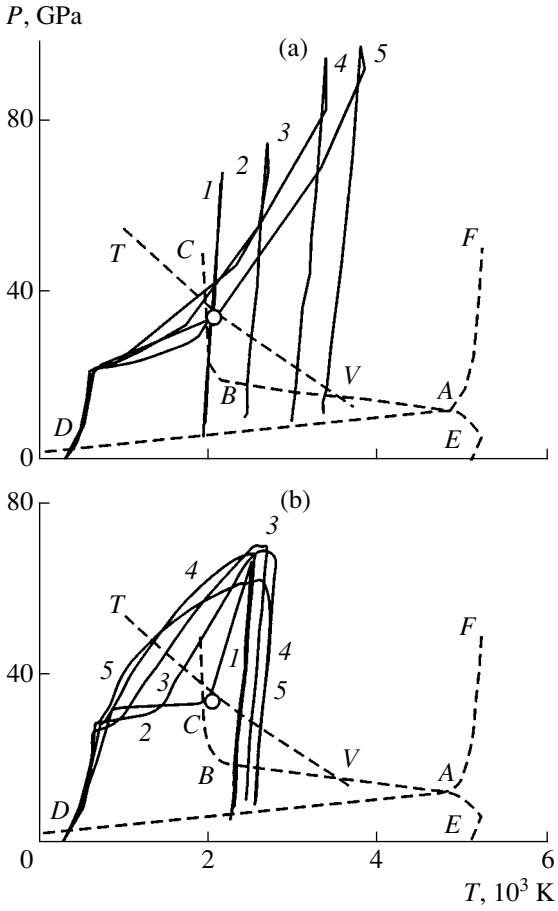

Fig. 4. Phase diagram of carbon and the evolution of five Lagrangean particles lying on the symmetry axis of the target with the channel at (a) $P_0 = 21$ GPa, $\tau = 5$ ns and (b) $P_0 = 24$ GPa, $\tau = 90$ ns. Solid curves *1*–*5*, calculated trajectories of the particles; dashed curves, phase transition boundaries according to [12] (*AD* is the line of graphite–diamond thermodynamic equilibrium; *ABC* is the boundary of rapid graphite-to-diamond conversion region; and *AE* and *AF* are the graphite and diamond melting curves, respectively) and [13] (*VT* is the lower boundary of the region of shock-induced graphite-to-cubic diamond conversion). The open circle is the point (T^*, P^*) from the solution to the one-dimensional problem.

region appeared to be much smaller than R even at τ = 5 ns.

We use the cylindrical coordinates (r, z) with the z axis originating at the initial position of the cone vortex (z = 0) and pointing to its base. Hydrodynamic equations (1) were solved with a program package constructed on mobile regular grids with the explicit separation of interfaces as grid lines and on explicit quasi-monotonic second-order schemes. Equations (2) and (3) were solved jointly with an inexplicit scheme for Eq. (4) at given V and E, which allowed for using grids with a mesh size much larger than the width of the phase transition region. The total number of meshes was varied between 25000 and 35000, of which the carbon accounts for 5000 and 9000–11000 for the usual cone target and for the target with the channel, respectively.

Along with Eq. (4), we considered the kinetic equation [13], where the ultimate value of α, instead of being unity, is presented by a function $\alpha_m(P)$ constructed from experimental data and the factor $(P/P_0 - 1)$ is replaced by a temperature-dependent exponential. Comparative calculations showed that both equations yield close results in our case. This circumstance can be explained by the fact that, for a striker velocity of 4 km/s, the pressure P^* behind the phase transition front is such that $\alpha_m(P^*) \simeq 0.99$ (i.e., close to unity). The discrepancy between the factors defining the front structure also appears to be insignificant because of a small front thickness.

RESULTS OF NUMERICAL SIMULATIONS

Consider first the one-dimensional problem when a shock wave is produced by an aluminum plate striking against a planar graphite sample. It has the approximate solution [9, 10] in the form of a two-wave structure. The first wave propagates in graphite and raises the pressure to the phase transition threshold P_0. The velocity of the first wave, D_1, and the material parameters behind the wave front are determined from the shock adiabat of graphite. The formation of diamond occurs in the second shock wave. The velocity of the second wave, D_2, and the material parameters behind its front are determined from a closed set of algebraic equations including the Hugoniot relations for this wave and for the wave in the striker, as well as the continuity conditions for pressure and velocity at the contact surface. For an aluminum striker moving with the velocity v = 4 km/s, the calculation yields $D_1 \simeq 7.1$ km/s, $D_2 \simeq 5.5$ km/s, $P^* \simeq 33$ GPa, and the temperature of diamond behind the second wave front $T^* \simeq 2010$ K. The only distinction between the numerical and approximate solutions of the one-dimensional problem is the presence of the second wave of width defined by the parameter τ in Eq. (4).

In the two-dimensional case, the process in a conic target may be subdivided into three stages. First, graphite-to-diamond conversion occurs and in the planar two-wave structure both under the action of the shock pulse reflected from the lateral wall of the taper. This stage is characterized by the cumulative effect due to the pulse collapse at the symmetry axis. In the absence of the relieving wave, next goes the stage of diamond pressing in the target at a nearly constant rate. At this stage, the instability of the carbon–target interface plays an important role. If the pressing stage lasts for a sufficiently long time, the instability may result in the complete mixing of the carbon with the target material. The final stage is carbon release in the rarefaction wave coming from the rear side of the striker.

The case of a thin striker, $H \ll R$, is of no interest, since the fast penetration of the rarefaction wave into the carbon precludes the cumulative effect; therefore, we assumed that $H \geq R$ in all the calculations.

THE CUMULATIVE EFFECT

The evolution of five Lagrangean carbon particles located along the symmetry axis of the target with the channel is plotted in the (T, P) plane for different values of the phase transition kinetic parameters (Fig. 4). The target parameters are taken to be $H = R = 5$ mm, $\vartheta = 60°$, and $r_0 = 1$ mm. Initially, the particles were almost regularly arranged between the striker surface and the cone vertex. They are numbered (1 to 5) starting from the surface.

Before proceeding with the analysis of Fig. 4, we recall that, according to dissipation-free hydrodynamic equations (1), the only allowed trajectory for a Lagrangean particle on the phase diagram in the region of one-phase flow is an adiabat (which, for solid particles, is close to an isotherm), unless this particle lies in the way of the shock wave. In the latter case, the entropy of the particle may instantly increase and the particle jumps to another adiabat. In the phase transition region, the motion of the particle depends on the function $\alpha(t)$ and the number of trajectories may greatly increase.

The results presented in Fig. 4a correspond to the parameters of highly oriented graphite studied in [10]. The width of the phase transition region here is much smaller than the size of the target. Until the compression pulse from the lateral wall of the cone reaches the symmetry axis, the flow along the axis retains the two-wave structure. This is demonstrated in Fig. 4a with the first and second particles, which are first compressed to the threshold pressure P_0 = 21 GPa and then pass into the (T^*, P^*) state. The further increase in the pressure is due to the compression pulse coming from the lateral wall of the cone to the symmetry axis, after which adiabatic release begins. It is seen that the growth of the pressure in the first two particles is also almost adiabatic and proceeds with a slight increase in the temperature.

After the collapse, the compression pulse catches up with the second shock wave, as illustrated by the behavior of the last three particles in Fig. 4a. They are also subjected to the action of the two-wave structure, like the first two, but the intensity of the second wave is appreciably higher.

The point (T^*, P^*) in Fig. 4a lies above the region of fast graphite-to-diamond conversion, which was shown [12] to be a possibility in dynamic experiments, but below the region where the shock-compressed graphite transforms into cubic diamond, according to estimates in [13]. The authors of [13] suggested that, at temperatures and pressures within this curve, graphite can transform into lonsdalite alone. This means that, in a plane two-wave structure, the transformation of graphite into lonsdalite, rather than into cubic diamond, is more realistic. However, the compression pulse collapsing at the symmetry axis will be subsequently certain to bring the Lagrangean particles into the region of cubic diamond formation. Note that the rate of graphitization for lonsdalite is appreciably higher than for cubic diamond [13].

Similar calculations were carried out for isotropic artificial graphite with the parameters taken from [9]. For this material, the characteristic transition time τ is nearly 20 times greater than for highly oriented graphite. In this case, the width of the transition region in the plane wave is comparable to the target size and the greater part of graphite is converted under the action of the compression wave reflected from the lateral surface of the cone. In isotropic graphite, one may consider the kinetic parameters of the phase transition to be the same irrespective of whether the transition is induced by the plane wave or by the compression pulse from the conical surface. It should be mentioned, however, that the neglect of porosity somewhat roughens the solution obtained for the graphite material in [9]. The results of our calculation are shown in Fig. 4b. Here, the distribution of the diamond parameters along the symmetry axis is much more uniform and all the Lagrangean particles reliably fall into the region of cubic diamond formation as before. As was expected, with an increase in the target size, the evolution of the Lagrangean particles becomes much the same as that in Fig. 4a.

An increase in the striker thickness H affects the trajectories of the Lagrangean particles on the phase diagram (Fig. 4) only slightly. This is because the basic factor defining such a trajectory is the maximal entropy gained by the particle during shock compression. Even at $H = R$, the time of arrival of the relieving wave from the outer boundary of the striker turns out to be sufficiently large, and this wave begins to have an appreciable influence on the shock compression of carbon. With an increase in H, this time only grows.

In the case of a target without a channel, the shock wave intensity near the cone vertex increases considerably, and the fifth Lagrangean particle even jumps over the melting curve of diamond.

Peak pressures attained in different targets, GPa

Material	ϑ, deg		
	60	75	90
Lead	170	305	405
Aluminum	115	–	–
Diamond	670	–	–

According to [17], the cumulative effect at the symmetry axis depends considerably on the cone angle ϑ and target material. The amount of this effect can be characterized by the peak pressure achieved during the process. The table lists these pressures in channel-free conic targets for different ϑ and different target materials. As ϑ grows from 60° to 90°, the cumulative effect builds up. For the aluminum target, the effect is weak. For diamond, which is chosen as an example of a superhard material, the effect exceeds greatly that for the other materials.

ON THE INFLUENCE OF THE TARGET SHAPE ON INSTABILITY BUILDUP

In the case of a channeled target, the development of instability of the carbon–target interface depends noticeably on the length and shape of the transition region between the cone and cylindrical channel. In cylindrical coordinates (r, z), this region is represented by a curve connecting the generating lines of the cone and the cylinder. The natural choice for such a curve is a circle tangent to both straight lines. This circle is defined by one additional parameter: the size of the transition region h, i.e., the distance along the z axis between the two tangency points, is convenient to take.

At $h = 0$, the interface has a kink. As is known, shock-wave-induced interface instability may start with a rapid twist of the interface in the vicinity of a kink. A similar phenomenon is observed in our case. Figure 5 illustrates the results of calculation for three targets with different h for the same time instant. When h is small ($h = 0.5r_0 = 0.5$ mm), a fold is formed on the interface profile (Fig. 5a). At $h = r_0$, the fold is much smaller (Fig. 5b), and for $h = 2r_0$, the profile remains practically smooth by the time instant shown in the plot (Fig. 5c).

With time, the interface becomes unstable at $h = 2r_0$ as well. We managed to reduce the instability still further by employing a more intricate geometry of the transition region. Instead of being directly adjacent to a cylindrical channel, the conical surface grades into another conical surface with a smaller cone angle, which, in turn, passes into a cylindrical channel. Such a target will be referred to as a biconical channeled target. For a target with $R = 5$ mm, $\vartheta = 60°$, and $r_0 = 0.5$ mm, the straight line in the plane (r, z) that corresponds to the intermediate conical surface had a length

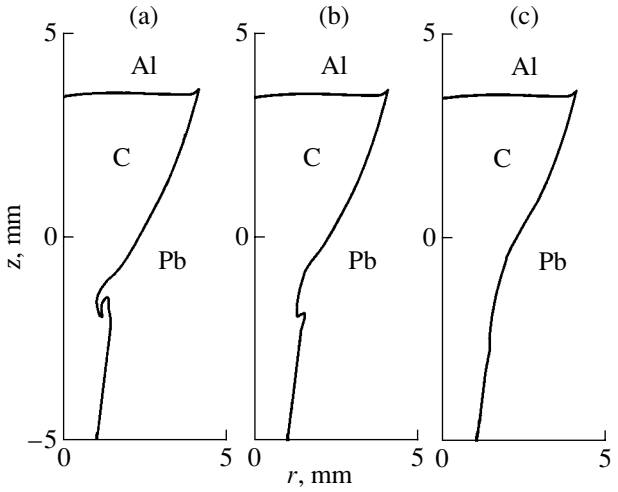

Fig. 5. Carbon boundary at $t = 1.75$ μs. $R = 5$ mm, $\vartheta = 60°$, $r_0 = 1$ mm, and h = (a) 0.5, (b) 1, and (c) 2 mm.

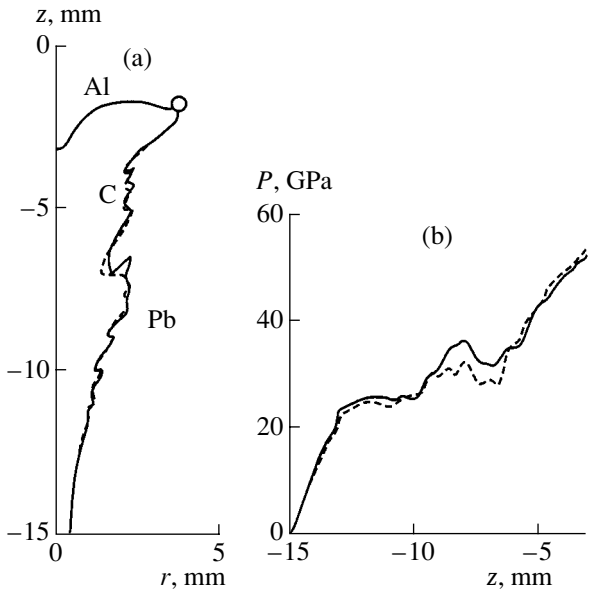

Fig. 6. Comparative calculation of the (a) carbon boundary and (b) pressure distribution along the symmetry axis of the target at $\varphi_{max} = 0$ (dashed curve) and 45° (solid curve).

$l = 8$ mm in the direction of the z axis and the radius along this line decreased from $r_1 \simeq 1$ mm near the main cone to r_0. The straight lines representing the conical surfaces were joined by a circular arc with a length $h = 2$ mm along the z axis.

QUASI-STATIONARY PRESSING

Assume that the striker thickness $H \gg R$. Then, the time it takes for the rarefaction wave to reach the carbon, which is a linear function of H, greatly exceeds the characteristic time of disturbance propagation inside the carbon. In the absence of the rarefaction wave, the rate with which carbon presses in the target is nearly constant. Pressing causes the instability of the carbon–lead interface to build up, leading, in general, to the formation (see, e.g., [24]) of a turbulent mixing region, which extends with time. However, one can suppose that there is a time interval within which the instability of the interface has only a slight influence on the distribution of the thermodynamic functions along the symmetry axis. Our aim was to find this distribution and the rate of the pressing-in process.

In the calculations, we encountered two difficulties, which were overcome at the expense of the accuracy of the method. The first one is related to the interface instability mentioned above. As time passes, the interface exhibits self-intersections, which indicate the onset of turbulent mixing. This stage is evidently impossible to calculate by explicitly separating out the interface as a grid line. In order to get around the problem, we imposed limitations on the extent of instability. Specifically, the angle φ between the grid line that represents the interface and the r axis cannot exceed a certain value φ_{max}. The positive direction along this line is defined so that $\varphi < 0$ at the initial position of the interface. For example, for the boundary of the cylindrical channel in Fig. 3, $\varphi = -90°$.

The influence of such roughening on the accuracy of flow calculation near the symmetry axis was estimated by performing the calculations with various φ_{max}. The results for $\varphi_{max} = 0$ and 45° are shown in Fig. 6. For the time instant shown, the carbon–lead interface contains at least one portion where different limits result in noticeably different profiles of the interface (Fig. 6a). However, the pressure distributions along the symmetry axis closely correlate for these two cases (Fig. 6b).

The second difficulty, because of which we had to roughen the method, is associated with the "triple" point where the boundaries of carbon, aluminum, and lead intersect (the circle in Fig. 6a). The angle of intersection decreases with time, which entails a decrease in the time step. Therefore, this angle was limited from below by a fixed value. As with the first limitation, such a roughening of the method affects the accuracy of computation near the symmetry axis insignificantly.

The pressing-in process with $\varphi_{max} = 0$ was calculated for a biconical channeled target and for a similar target without a channel at $r = 5$ mm and $\vartheta = 60°$ (Figs. 7, 8). In both cases, the interface contains points where the artificial expedient described above substantially limits the instability. It would be natural to assume that these points are centers around which the flux twists and the turbulent mixing region arises. The pressure distributions along the symmetry axis (Fig. 8) weakly vary with time but depend considerably on the type of the target. In the channel-free target, the pressure at the axis varies within the narrow range $P = 55$–69 GPa; in the presence of the channel, the pressure varies over substantially wider limits: from $P = 50$–55 GPa near the striker to $P = 25$–30 GPa behind the

front of the bow wave propagating through the channel. The pressing-in rates were also found to differ appreciably: ≈2 km/s for the target without the channel and ≈4.5 km/s for the channeled target.

RELEASE WAVE

Assume that the striker thickness $H = R$. In this case, carbon release starts nearly at the same moment the action of the cumulative effect comes to an end. The profiles of the carbon at different instants of time are shown in Fig. 9. The effect of the rarefaction wave can be seen by comparing these profiles with the corresponding ones plotted in Fig. 7 in the absence of release. The decrease in the radial dimension of the carbon volume with time indicates that the release is essentially nonuniform. In the channeled target, release may produce a waist in the channel with the resulting formation of two isolated carbon volumes, as follows from our calculations.

The calculations predict the appearance of a region with an extremely high negative pressure ($P < -5$ GPa) in aluminum when the rarefaction wave reflects from the aluminum–lead interface. This points to the possibility of spalling in the actual flow (see, e.g., [25]).

The influence of this region on release was estimated as follows. We specified a certain critical pressure $P_{min} < 0$. Those meshes where the pressure dropped below the critical value, $P < P_{min}$, were rejected from the numerical calculation and the pressure at their boundaries was set equal to atmospheric. The calculation was carried out for different values of P_{min} in the range from –3 to –1 GPa. In spite of the large number of the cells rejected, including the part of the striker over the carbon volume, no changes in the flow parameters inside the carbon were observed.

Figure 10 shows the time dependences of the pressure in a Lagrangean particle for targets with and without the channel, along with the solution to the one-dimensional problem where the striker hits a graphite layer on a semiinfinite lead substrate. The graphite layer thickness is equal to the cone height $R/\tan(\vartheta/2)$. Initially, the particle is located approximately at the half-height of the cone without the channel and in the middle of the graphite layer in the one-dimensional case. In the target with the channel, the particle was initially situated slightly above the cone half-height in order that the pressure change in it be comparable to that experienced by the particle in the channel-free target. For the target with the channel, the pressure in the particle lying strictly at the cone half-height is significantly larger because of the waist in the channel.

The process of relieving in the planar layer is seen to differ qualitatively from that in the conical target. In the latter, the pressure in the particle slowly decreases and remains at a level $P \approx 7$ GPa at $t = 5$ μs. In the planar graphite layer, the pressure drops to zero even at $t \simeq$

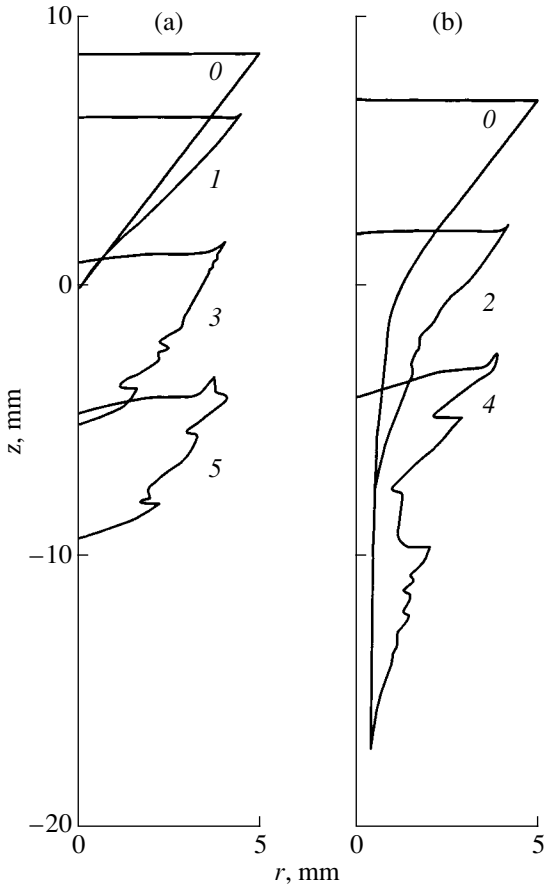

Fig. 7. Carbon boundary at different time instants for (a) channel-free and (b) channeled targets under the action of a striker with $H \gg R$. The figures by the curves are time instants in μs.

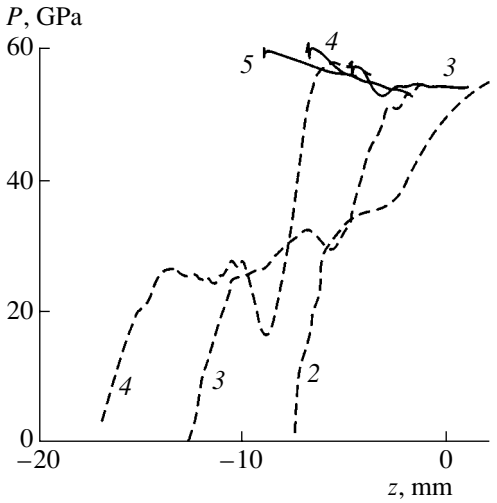

Fig. 8. Pressure distribution along the symmetry axis of channel-free (solid curves) and channeled targets (dashed curves) at different time instants (marked by the figures by the curves in μs) at $H \gg R$.

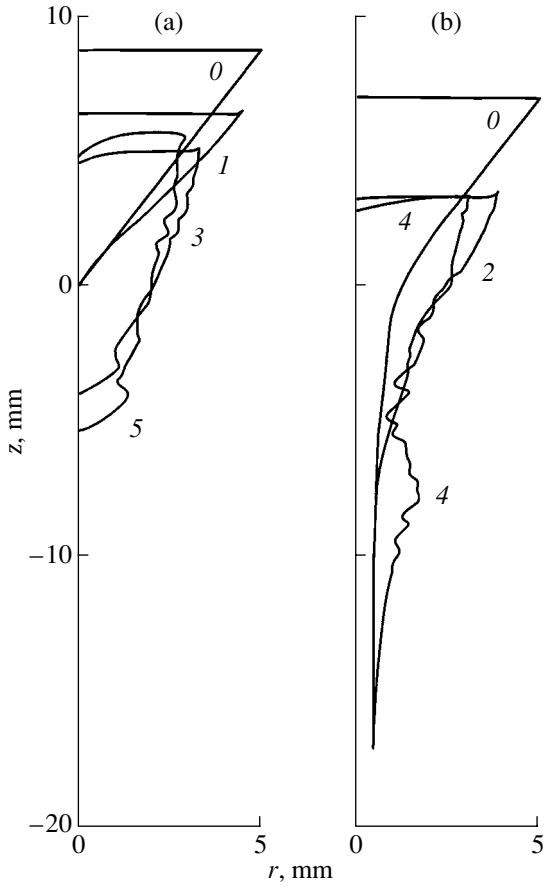

Fig. 9. Carbon boundary in different time instants (marked by the figures, μs) for (a) channel-free and (b) channeled targets for $H = R$.

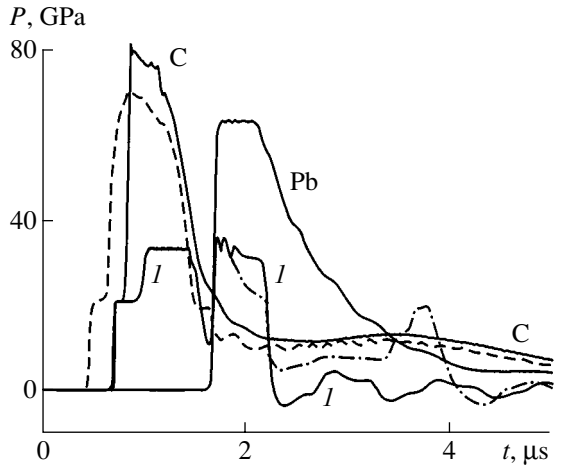

Fig. 10. Time dependence of the pressure in the Lagrangean particle. C, two-dimensional simulation of channel-free (solid curve) and channeled targets (dashed curve); (*1*) one-dimensional simulation with (solid curve) and without (dash-and-dot curve) considering spalling; and Pb, one-dimensional simulation with lead in the place of carbon.

2.2 μs. Note also the strong effect of spalling in the striker on the pressure evolution in the planar layer.

We also solved the one-dimensional problem for a homogeneous lead target without carbon. The associated curve is plotted in Fig. 10 for a Lagrangean particle located at nearly the same depth as the carbon particle in the conical target in view of pressing in the target. The time dependence of the pressure in the lead particle is qualitatively similar to that for the carbon particle in the conical target. One more coincidence should be noted. Relieving in lead, as well as in carbon in the conical target, is nearly independent of whether spalling in the aluminum striker is taken into consideration or not. These data lead us to the important inference that the release of carbon in a conical lead target is governed mainly by the properties of lead rather than of carbon.

CONCLUSIONS

At the initial stage of compression, the cumulative effect at the symmetry axis is essential. Its amount can be varied in a wide range by changing the cone angle and choosing the appropriate target material.

From comparison of our results with the phase diagram for carbon available from the literature, we conclude that, due to the cumulative effect, the carbon at the symmetry axis is certain to fall into the region of cubic diamond formation.

In the presence of an extracting channel in the target, the action of the cumulative effect along the symmetry axis is more uniform. Cumulation is the most uniform when the width of the phase transition region in the bow plane wave is comparable to the conical cavity dimension and the diamond formation is completed in the collapsing pressure pulse. Such a situation arises in a target with a cone base radius $R \approx 5$ mm for isotropic graphite with a relatively large characteristic time of conversion.

At later stages of carbon evolution, the instability of the carbon–target interface plays an important part. In channeled targets, the instability develops rapidly because of the abrupt transition from the cone to the channel. The instability can be controlled to some extent by varying the cone profile and, possibly, the target material.

When thick strikers are used, the carbon is pressed in the target at a nearly constant rate until the rarefaction wave comes. The pressing-in process causes the instability to grow up to the complete mixing of carbon with the target material. Because of this, there is no point in using strikers that are too thick and channels that are too long.

We determined the pressing-in rate for channeled and channel-free lead targets and found the pressure distribution along the symmetry axis, which exists as long as the influence of the interface instability is small.

We also simulated the stage of carbon release induced by the rarefaction wave coming from the external surface of the striker (including the rough simulation of spalling). For a conical lead target, release is shown to depend largely on the properties of the target material rather than those of the sample. In particular, spalling in the striker produces no effect on the unloading.

No consideration was given to the case when a channel-free conical target has a hole at the vertex, i.e., when diamond formed is extracted into an atmospheric-pressure environment. A probable result in this case is the fast destruction of the diamond product into tiny pieces, which rapidly cool down and fly apart. The freezing of the degrees of freedom (i.e., the stabilization of the state reached), which is usually observed under these conditions, is expected to prevent the graphitization of the diamond particles. The ultradisperse diamond powder thus obtained is of independent technical value. Note that there is a basic possibility to decelerate the spreading diamond pieces without repeated substantial heating and, thus, to graphitize them. According to [26], even cosmic dust entering into the Earth's atmosphere at a cosmic velocity may remain fairly cool in a certain range of particle sizes.

The general conclusion to emerge from the results presented above is that the production of artificial diamond in conical targets is quite possible. Therefore, a real physical experiment with such targets, which does not seem to be an extremely difficult task, would be of great interest.

ACKNOWLEDGMENTS

We thank A.Z. Zhuk and V.V. Milyavskiĭ for the valuable discussions and helpful remarks.

This work was supported by the Russian Foundation for Basic Research (project nos. 00-01-00551 and 00-02-16687), the Russian Academy of Sciences (grant no. 31 of the 6th Expertise Competition of Young Scientist, 1999), and the Native Science Foundation of the Russian Federation.

REFERENCES

1. P. S. De Carli and J. C. Jamieson, Science **133**, 1821 (1961).
2. B. J. Alder and R. H. Christian, Phys. Rev. Lett. **7**, 367 (1961).
3. M. N. Pavlovskiĭ and V. P. Drakin, Pis'ma Zh. Éksp. Teor. Fiz. **4**, 169 (1966) [JETP Lett. **4**, 116 (1966)].
4. H. P. Bundy and J. S. Kasper, J. Chem. Phys. **46**, 3437 (1967).
5. G. E. Duvall and R. A. Graham, Rev. Mod. Phys. **49**, 523 (1977).
6. A. V. Kurdyumov and A. N. Pilyankevich, *Phase Transformations in Carbon and Boron Nitride* (Naukova Dumka, Kiev, 1979).
7. W. H. Gust, Phys. Rev. B **22**, 4744 (1980).
8. M. F. Gogulya, Fiz. Goreniya Vzryva **25** (1), 95 (1989).
9. A. Z. Zhuk, A. V. Ivanov, and G. I. Kanel', Teplofiz. Vys. Temp. **29**, 486 (1991).
10. D. J. Erskine and W. J. Nellis, J. Appl. Phys. **71**, 4882 (1992).
11. A. V. Kurdyumov, V. G. Malogolovets, N. V. Novikov, A. N. Pilyankevich, and L. A. Shul'man, *Polymorphous Modifications of Carbon and Boron Nitride* (Metallurgiya, Moscow, 1994).
12. F. P. Bundy, W. A. Bassett, M. S. Weathers, *et al.*, Carbon **34** (2), 141 (1996).
13. E. Vlodarchik and R. Trebinski, Shock Waves **7**, 231 (1997).
14. Yu. N. Zhugin, K. K. Krupnikov, and V. I. Tarzhanov, Khim. Fiz. **18** (5), 96 (1999).
15. M. N. Pavlovskiĭ, K. V. Khishchenko, M. V. Zhernokletov, *et al.*, in *Physics of Extreme State of Matter-2001: Collection of Articles*, Ed. by V. E. Fortov (Inst. Khim. Probl. Khim. Fiz. RAN, Chernogolovka, 2001), pp. 96–97.
16. K. V. Khishchenko, V. E. Fortov, I. V. Lomonosov, *et al.*, in *Shock Compression of Condensed Matter-2001: Collection of Articles*, Ed. by M. D. Furnish, N. N. Thadhani, and Y. Horie (AIP, New York, 2002), pp. 759–762.
17. I. V. Lomonosov, V. E. Fortov, A. A. Frolova, *et al.*, Dokl. Akad. Nauk **360** (2), 199 (1998) [Phys. Dokl. **43**, 306 (1998)].
18. K. V. Khishchenko, I. V. Lomonosov, and V. E. Fortov, in *Physics of Extreme State of Matter-2001: Collection of Articles*, Ed. by V. E. Fortov *et al.* (Inst. Probl. Khim. Fiz. RAN, Chernogolovka, 2001), pp. 94–96.
19. K. V. Khishchenko, V. E. Fortov, and I. V. Lomonosov, Int. J. Thermophys. **24**, 2003 (in press).
20. M. N. Pavlovskiĭ, Fiz. Tverd. Tela (Leningrad) **13**, 893 (1971) [Sov. Phys. Solid State **13**, 741 (1971)].
21. *LASL Shock Hugoniot Data*, Ed. by S. P. Marsh (University of California Press, Berkeley, 1980).
22. A. V. Bushman, G. I. Kanel', A. L. Ni, and V. E. Fortov, *Thermal Physics and Dynamics of Intense Pulsed Effects* (Inst. Khim. Fiz. AN SSSR, Chernogolovka, 1988).
23. R. I. Nigmatulin, *Mechanics of Multiphase Media* (Nauka, Moscow, 1987).
24. V. E. Neuvazhaev, Mat. Model. **3** (7), 10 (1991).
25. G. I. Kanel', S. V. Razorenov, A. V. Utkin, and V. E. Fortov, *Shock Wave Phenomena in Condensed Media* (Yanus-K, Moscow, 1996).
26. *Comets and the Origin of Life*, Ed. by S. Ponnamperuma (Reidel, Dordrecht, 1981; Mir, Moscow, 1984), pp. 69–77.

Translated by A. Sidorova

Technical Physics, Vol. 48, No. 6, 2003, pp. 736–741. Translated from Zhurnal Tekhnicheskoĭ Fiziki, Vol. 73, No. 6, 2003, pp. 76–81.

SOLID-STATE ELECTRONICS

Bending Vibrations of Semiconductor Wafers with Local Heat Sources

A. M. Orlov, A. A. Skvortsov, and O. V. Litvinenko

Ul'yanovsk State University, Ul'yanovsk, 432700 Russia

e-mail: scvor@sv.uven.ru

Received July 22, 2002; in final form, December 15, 2002

Abstract—Experimental and theoretical analysis of temperature fields generated by the pulsed current heating of metallization paths on a semiconductor wafer surface is performed. It is shown that any step change in the heating power causes bending vibrations of the wafer. Within the safety margins for semiconductor device operation, the vibration amplitude is proportional to the step amplitude. The damping factor for the entire wave packet is found ($\Gamma = 1103\ s^{-1}$), and frequency components of acoustic radiation excited in a 300-μm-thick silicon wafer are determined. *© 2003 MAIK "Nauka/Interperiodica".*

In advanced ICs, component interconnections have usually the form of fine-line thin-film metallization paths. When semiconductors are switched on and off, as well as under the pulsed regime of operation, high potential and current jumps acting on the metallization cause local thermal disturbances in the chip. Nonuniform thermal expansion, in turn, generates mechanical vibrations, which, when propagating, not only distort electrical signals [1] but also, along with thermal effects [2], accelerates IC performance degradation. Therefore, combined analysis of thermal and vibratory processes in the chip may shed light upon its thermoacoustic behavior. Such an analysis is the subject of our investigation.

TEMPERATURE FIELDS GENERATED BY A SURFACE HEAT SOURCE

Let us find a temperature field in a semiconductor wafer when the current is passed through a rectangular metallization path on the surface.

To begin with, we will construct the Green function for the problem of a rectangular heat source, that is, solve the problem first for a point heat source of power w placed at the origin of the spherical coordinate system and acting from the zero time $\tau = 0$. The unsteady equation of heat conduction for a spherically symmetric problem has the form [3]

$$\frac{\partial T(r,\tau)}{\partial \tau} = a\left(\frac{\partial^2 T(r,\tau)}{\partial r^2} + \frac{2}{r}\frac{\partial T(r,\tau)}{\partial r}\right), \qquad (1)$$

where T is temperature, K; r is the radial coordinate, m; $a = \lambda/(cd)$ is the thermal diffusivity of the wafer, m^2/s; λ is the thermal conductivity, W/(m K); c is the specific heat, J/(kg K); and d is the density, kg/m^3.

The wafer is assumed to be so thin that bending vibrations dominate over all others. On the other hand, we require the time domain of analysis to satisfy the condition of semiinfinity of the wafer thickness. This condition is readily met if the thermal front does not reach the opposite surface upon heating a metallization layer, e.g., by current. Then, the boundary conditions and the initial condition are written as

$$w = \lim_{r\to 0}\left(-2\pi r^2 \lambda \frac{\partial T(r\tau)}{\partial r}\right),$$
$$T(\infty,\tau) = T_0, \quad T(r,0) = T_0. \qquad (2)$$

Here, w is the power of the point heat source, W; $2\pi r^2$ is the surface area of the semispherical front of heat propagation; and T_0 is the initial temperature, K. Applying the time Laplace transformation to (1) and (2), we can write

$$\frac{\partial^2}{\partial r^2}(rT_L(r,s)) - \frac{s}{a}(rT_L(r,s)) + \frac{rT_0}{a} = 0, \qquad (3)$$

$$\frac{w}{s} = \lim_{r\to 0}\left(-2\pi r^2 \lambda \frac{\partial T_L(r,s)}{\partial r}\right), \quad T(\infty,s) = T_0, \qquad (4)$$

where $T_L(r,s)$ is the Laplace transform of the function $T(r,\tau)$ and s is the operator variable, s^{-1}.

Substituting conditions (4) into the general solution to Eq. (3) [3],

$$rT_L(r,s) - \frac{rT_0}{s} = A\exp\left(\sqrt{\frac{s}{a}}r\right) + B\exp\left(-\sqrt{\frac{s}{a}}r\right),$$

we find the coefficients A and B and the final solution

for the transform $T_L(r,s)$:

$$sT_L(r,s) - T_0 = \frac{w}{2\pi\lambda r}\exp\left(-\sqrt{\frac{s}{a}}r\right). \quad (5)$$

Using the tables of Laplace transforms [3], one can find the original of expression (5):

$$T(r,\tau) - T_0 = \frac{w}{2\lambda\pi r}\left(1 - \Phi\left(\frac{r}{\sqrt{4a\tau}}\right)\right) = wG(r,\tau), \quad (6)$$

where $\Phi(\xi) = 2/\sqrt{\pi}\int_0^\xi \exp(-\xi^2)\,d\xi$ is the probability integral.

Now, knowing the Green function $G(r,\tau)$ for our problem (6), we can obtain a solution for a surface heat source of arbitrary shape:

$$T(r,\tau) - T_0 = q\iint_{\Sigma'} G(r - r',\tau)\,d\Sigma'. \quad (7)$$

Here, q is the density of the heat flux (W/m^2) through the surface Σ' having the coordinate r'. In order to find a temperature field produced by a rectangular heat source of length l (m) and width b (m), we introduce the Cartesian system with the origin placed at the center of the path. The X, Y, and Z axes are directed along the length, along the width, and into the wafer, respectively. If the metallization resistance is $R(\Omega)$, then direct current of strength I (A) passing along the X axis generates a heat flux of density

$$q = I^2R/lb. \quad (8)$$

In the new coordinate system, the substitution of (6) and (8) into (7) yields a solution in quadratures,

$$T(x,y,z,\tau) - T_0 = \frac{I^2R}{4\lambda lb\sqrt{\pi^3 a}}\int_0^\tau \frac{dt}{\sqrt{t^3}}\int_{-l/2}^{l/2} dx'\int_{-b/2}^{b/2} dy' \times \exp\left(-\frac{(x-x')^2 + (y-y')^2 + z^2}{4at}\right), \quad (9)$$

which describes the temperature distribution in a wafer with a rectangular heat source.

In experiments, it is most convenient to keep track of the temperature variation on the chip surface (at $z = 0$), since the surface temperature coincides with that of the metallization layer (this layer is, as a rule, very thin, and the thermal conductivity of metals is much higher than that of semiconductors). Therefore, with (9) one can find not only the average temperature of the path but also the surface temperature distribution, which is responsible for the generation of dislocation loops in the semiconductor during the relaxation of elastic stresses.

To simplify the integration of (9), we assume that $l \gg b$ and the wafer thickness is semiinfinite as before. In this case, one may ignore heat leakage from the end of the conductor and set the limits of integration with respect of x' in (9) infinite. Then, the temperature profile $T(y,\tau)$ can be represented in the form

$$T(y,\tau) - T_0 = \frac{I^2R}{2\pi\lambda lb}\left\{\left(\frac{b}{2} - y\right)E_1\left(\frac{(b/2-y)^2}{4a\tau}\right) + \left(\frac{b}{2} + y\right)E_1\left(\frac{(b/2+y)^2}{4a\tau}\right)\right\} + \frac{I^2R\sqrt{a\tau}}{\sqrt{\pi\lambda lb}}\left\{\Phi\left(\frac{b/2-y}{4a\tau}\right) + \Phi\left(\frac{b/2+y}{4a\tau}\right)\right\}, \quad (10)$$

where

$$E_1(z) = \int_z^\infty \exp(-\xi)\,d\xi/\xi$$

is the integral exponential.

If R, a, and λ strongly depend on temperature, the numerical calculation of $T(y,\tau)$ should be performed by using the integral-averaged (over temperature) values of these parameters.

Note that, if (10) is averaged over y and the path width b tends to infinity, the result will completely coincide with the time dependence of the temperature of an infinite heating surface [2]:

$$T(\tau) - T_0 = \frac{q\sqrt{a\tau}}{\lambda} = \frac{I^2R\sqrt{a\tau}}{\lambda lb}.$$

This solution can be used for short heating times or wide paths.

GENERATION OF BENDING VIBRATIONS IN A WAFER BY THERMAL SHOCK

It is known [4] that the unsteady temperature field of a point heat source generates mechanical stresses in a wafer, which, in turn, cause bending waves:[1]

$$\sigma_{ik} = -K\beta\{T(r,\tau) - T_0\}\delta_{ik}. \quad (11)$$

Here, σ_{ik} is the stress tensor, Pa; K is the bulk modulus, Pa; D is the cylindrical stiffness of the wafer, J; β is the thermal expansion coefficient, K^{-1}; and δ_{ik} is the Kronecker tensor.

Wafer vibrations under the action of external bending forces are described by the equation [4]

$$hd\frac{\partial^2\varepsilon_z}{\partial\tau^2} + D\Delta^2\varepsilon_z = \int_0^h \frac{\partial\sigma_{zk}}{\partial x_k}dz. \quad (12)$$

[1] Heat sources will be considered as point sources, because their dimensions are much smaller than the wafer diameter.

Substituting (6) and (11) into this equation and differentiating with respect to time τ allows one to analyze the vibrational velocities v_z (m/s) of surface points rather than their displacements ε_z (m):

$$hd\frac{\partial^2 v_z}{\partial \tau^2} + D\Delta^2 v_z = \frac{K\beta w}{4\lambda\sqrt{\pi^3\tau^3 a}}\exp\left(-\frac{x^2+y^2}{4a\tau}\right)\left\{1-\exp\left(-\frac{h^2}{4a\tau}\right)\right\}. \quad (13)$$

It is seen that the driving force per unit surface area grows when approaching the heat source (x and $y \longrightarrow 0$) and reaches a maximum at $\tau \longrightarrow 0$. Hence, the most intense vibrations take place at the instant the heat source is switched on and heats up surface layers of the semiconductor.

EXPERIMENTAL RESULTS

We experimented on (111)Si wafers (diameter D_w = 76 mm, thickness h = 350 μm, and resistivity 30 Ω cm) on which test aluminum structures (the metal film thickness 1.5 μm) were patterned by optical lithography. The patterns comprised rectangular metallization paths (l = 2.9 mm, b = 75 μm, $D_w \gg l \gg b$, $l \gg h$) and contact pads to which the current was applied and from which the voltage across various parts of the structure was picked up [2]. Thermoacoustic processes arising in response to single dc pulses of density $j \le 7 \times 10^{10}$ A/m^2 were studied. Note that such current densities preclude the melting of Al/Si contacts.[2] The pulse duration τ_0 was selected such that the temperature front did not reach the opposite surface of the wafer: $\tau = h^2/4a \sim$ 0.8 ms.

The bending vibrations of the wafer were detected by a 300-μm-thick planar piezoelectric transducer of diameter 1 cm made of PZT-19 ceramic and rosin-glued to the wafer over its principal cut. It was shown [6] that the electric signal $U_s(\tau)$ from the transducer thus attached is proportional to the vibrational speed of the surface. Our instruments provided the reliable detection of acoustic responses at current densities from $j \ge 10^8$ A/m^2 up to limiting values at which the metallization layer starts melting.

Concurrently with the acoustic response $U_s(\tau)$, time voltage waveforms $U(\tau)$ across the metallization path, so-called switching waveforms, were recorded. Both signals were memorized in a computer through an S9-8 digital storage two-beam oscilloscope.

Experimental switching voltage waveforms $U(\tau)$ were compared with results of theoretical calculation. To this point, formula (10) was written in terms of voltage,

$$U(y,\tau) = IR_0(1+\alpha\{T(y,\tau)-T_0\}) = jl\rho_0(1+\alpha\{T(y,\tau)-T_0\}), \quad (14)$$

by using the known parameters of the metallization path, R_0 = 0.78 Ω and $\rho_0 = 2.7 \times 10^{-8}$ Ω m for T_0 = 290 K, and the thermal expansion coefficient α (K^{-1}) of aluminum. The temperature dependences of c, λ, and d for silicon and of ρ_0 and α for aluminum were taken from [7]; the approximations of the functions $E_1(\xi)$ and $\Phi(\xi)$, from [8].

Numerical analysis showed that the temperature at the center (y = 0) of the path is always somewhat higher than at the edges ($y = \pm b/2$). Qualitatively, this is supported by the fact that, when the critical current density is applied, the path starts melting at the center in all cases. For below-critical current densities, the analytical temperature difference is no greater than 11°C. Hence, the maximal spread of R across the path must be within 5% and can well be neglected. Further calculations were carried out with (10) for y = 0. Agreement between the theory and experiment is good (Fig. 1); so this approach can be used not only for constructing temperature profiles across a semiconductor wafer but also for the analysis of thermal-shock-induced bending vibrations.

Experimental signals $U_s(\tau)$ from the acoustic transducer were extracted by the differential technique in order to pick out undistorted information. For this purpose, single equishaped pulses of different polarity were applied to the structure and acoustic responses to them (Fig. 2, curves *2*, *4*) were recorded. The desired signal (Fig. 2, curve *3*) was found as the arithmetic mean.

The difference between the resulting curve and initial curves (Fig. 2, curves *2*, *4*) is the greatest at the beginning of pulse application. Among all the curves $U_s(\tau)$, curve *3* is the only one exhibiting an actual time delay of the acoustic response. Using this information, we determined the propagation velocity of the bending vibrations, which was found to be ≈1.5 km/s.

This value must correspond to the highest frequency harmonic, since bending vibrations feature the increasing dispersion law [9]. As follows from the numerical Fourier transformation of the signal, this is the 61-kHz harmonic. The frequency spectrum of thermal-shock-induced bending vibrations was studied at length in [10].

The acoustic response to the total current pulse does not discriminate between three basic stages of vibration evolution. The first one is associated with the thermal shock at the instant of pulse application; the second stage, with the period of long-term heating (when most of the thermal energy is dissipated in the wafer); and the third stage, with instant the pulse is switched off. Therefore, we carried out a series of runs where acous-

[2] The range of j is close to that used in modern semiconductors, $j \le 4 \times 10^{10}$ A/m^2 [5]. In the pulsed operating mode, the limiting value of j may be increased by one order of magnitude.

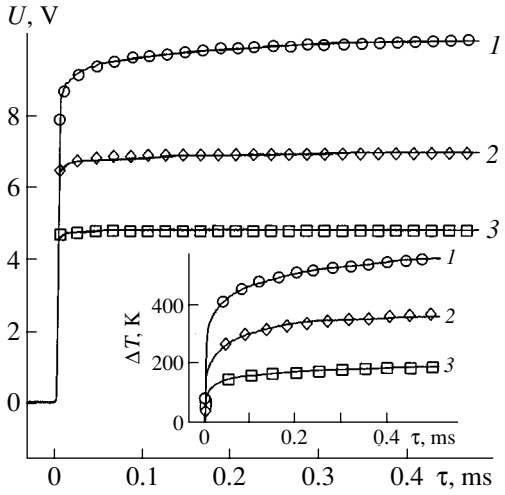

Fig. 1. Experimental (curves) and analytical (symbols) variation of the voltage across and temperature (inset) of the metallization path upon the application of a current pulse of density j = (*1*) 3×10^{10}, (*2*) 2.6×10^{10}, and (*3*) 2.2×10^{10} A/m^2.

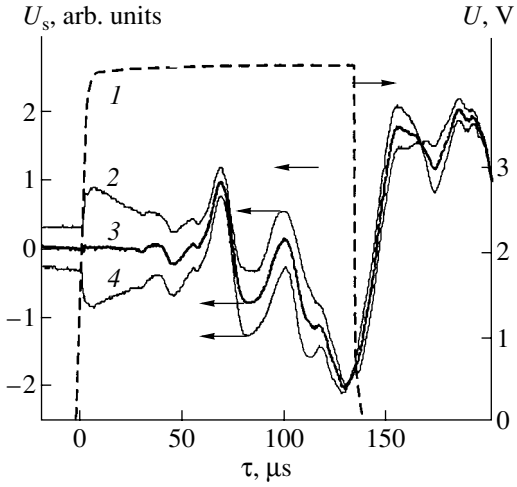

Fig. 2. Signal from the acoustic transducer upon the application of a current pulse of (*2*) direct and (*4*) reverse polarity to the metallization path. (*1*) Switching-on voltage waveform recorded synchronously ($j = 1.7 \times 10^{10}$ A/m^2) and (*3*) arithmetic mean of signals *2* and *4* (curves *2* and *4* are displaced up and down, respectively, along the vertical axis for clarity).

tic signals from each of the stages were separated out to estimate their partial roles.

The response combining the first two stages (Fig. 3, curve *3*) is detected most readily if the acoustic signal recording time (i.e., 0.99 ms) is slightly less than the current pulse duration 1.0 ms (Fig. 3, curve *1*).

To separate out the third stage (pulse switching off), it is necessary to generate a series of independent acoustic responses by heating the same test structure with pulses equal in amplitude but differing in duration (τ_{0i} < 0.99 ms). The difference between curve *3* in Fig. 3a and any of the acoustic signals thus obtained (Fig. 3a, curve *4*) gives the switching-off response (Fig. 3a, curve *5*).

It is noteworthy that, when the beginning of acoustic response *5* is displaced to zero time (Fig. 3b), it becomes the specular reflection of curve *3* up to a sample of the recorder regardless of the pulse duration τ_{0i} (i.e., regardless of the second stage duration). This means that the signal is generated only at the on and off times, which directly follows from (13). It should be mentioned here that the total amplitude of the responses to two current jumps due to j switching on and off depends on τ_{0i}, which specifies the time delay [10].

When the current density j and, accordingly, the current pulse power

$$w = I\frac{1}{\tau_0}\int_0^{\tau_0} U(\tau)d\tau$$

increase, the frequency and phase characteristics of the acoustic response do not change. They depend only on the way the wafer edges are fixed [10]. However, the amplitude of the acoustic response grows.

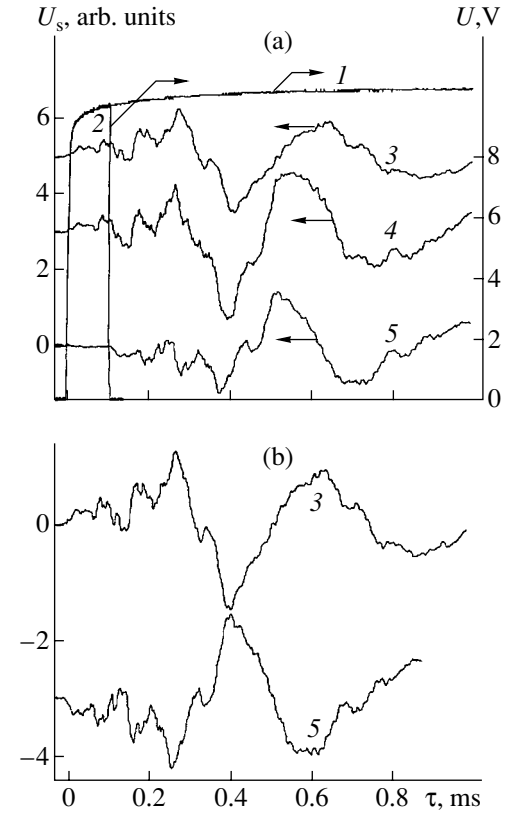

Fig. 3. Voltage across the metallization path upon the application of current pulses with $j = 2.6 \times 10^{10}$ A/m^2 and a duration of (*1*) 1 and (*2*) 0.1 ms. (*3*, *4*) Related acoustic responses and (*5*) the result of subtraction of curve *3* from curve *4*. Curves *3* and *4* in (a) are displaced relative to curve *5* (along the vertical axis) by 0.5 and 0.3 arb. units, respectively, and curve *5* in (b) is displaced by τ_0 = 0.1 ms to the left along the abscissa axis and by 3 arb. units along the vertical axis.

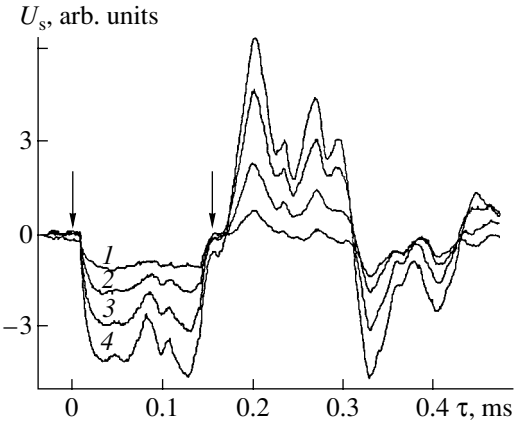

Fig. 4. Growth of the acoustic response amplitude with increasing current density j of the exciting pulse. j = (*1*) 1.2×10^{10}, (*2*) 1.4×10^{10}, (*3*) 1.7×10^{10}, and (*4*) 1.9×10^{10} A/m^2. The arrows indicate the instants the pulses are switched on and off.

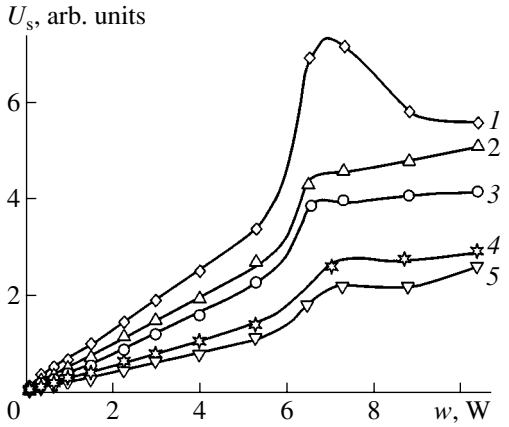

Fig. 5. Local maxima of the acoustic response amplitude $U_s(\tau)$ vs. Joule heat dissipated in the metallization path. Each of the curves corresponds to a certain peak in Fig. 4. τ = (*1*) 0.195, (*2*) 0.270, (*3*) 0.332, (*4*) 0.230, and (*5*) 0.405 ms.

The change in the acoustic signal amplitude was traced for each of the local maxima in the voltage waveforms $U_s(\tau)$ depicted in Fig. 4. The amplitude of each of the peaks is seen to grow in proportion with the exciting pulse power w (Fig. 5), as follows from (13). The extrapolation of $U_s(\tau)$ to the origin indicates that the acoustic response arises, however low the pulsed current action. When $j > 3.4 \times 10^{10}$ A/m^2 ($w > 5.2$ W), the dependence of $U_s(\tau)$ on w becomes nonlinear, suggesting that the energy is spent largely on degradation processes: the breakdown of the metallization and the formation of defects in the single-crystal substrate. This allows designers to find a safety margin in terms of current load on semiconductor structures.

In the final part of our investigation, we will determine the damping parameters of the acoustic signal. The most reliable relevant information can be obtained by studying the attenuation of the acoustic response to the total current pulse or to its last stage, when the heat flux is instantly interrupted. If the time of recording the decaying acoustic signal is sufficiently long,[3] the damping factor can be easily found from the envelope $U_s(\tau - \tau_0)$ of the entire wave packet. The damping factors of all harmonics filtered out by the crystal are of much greater interest. They can be found by successively processing self-sufficient fragments of the acoustic response recorded.[4]

The essence of the procedure is as follows. Fourier transformation is applied to a selected window, and the amplitudes of various harmonics are fixed. The window is then shifted by one sample, and the calculation is repeated. As a result, we find the amplitudes $A(\tau)$ of the harmonics under study vs. time. These dependences are plotted in the semilogarithmic scale (Fig. 6), and the time damping factors

$$\Gamma = \frac{1}{\tau}\ln\frac{A(0)}{A(\tau)} \qquad (15)$$

are determined by linearization.

The corresponding results are listed in the table. They well correlate with the damping factor for the

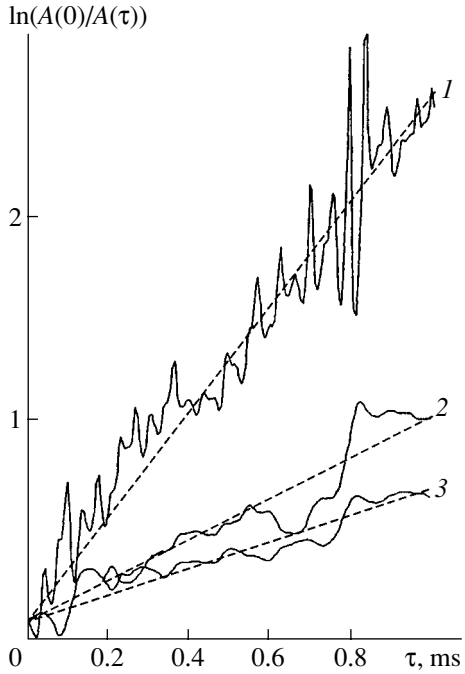

Fig. 6. Linear approximation (dashed lines) of the harmonic amplitude vs. pulse travel time dependences (continuous curves): (*1*) 26.7, (*2*) 9.1, and (*3*) 3.7 kHz.

[3] That is, the envelope amplitude decreases by no less than e times.

[4] By a self-sufficient fragment ("window") of the response recorded, we mean that number of data points from the entire data array that allow for reliable Fourier transformation. In our results, the window covers 50% (2048 data points) of the entire array.

Time damping factor for bending harmonics with various frequencies in a thin silicon wafer

f, kHz	Γ, s^{-1}
3.7	671
6.2	1032
9.1	1164
11.4	2402
26.1	2678

entire wave packet, $\Gamma = 1103\ s^{-1}$, which was determined from the signal envelope.

Thus, we investigated the influence of current pulses on the temperature conditions in metallization layers and the amplitude characteristics of thermally induced bending vibrations in the substrate. It is shown that acoustic waves are generated at any time instant the current density changes stepwise. The intensity of the acoustic waves varies as the power of the pulse. Slow variations of the resistance (hence, electrical power) do not cause intense vibrations.

It is established that safety margins for semiconductor operation can be set by examining the intensity variation of acoustic radiation in the substrate.

REFERENCES

1. V. A. Kalitenko, I. Ya. Kucherov, and V. M. Perga, Ukr. Fiz. Zh. **31**, 2006 (1986).
2. A. M. Orlov, A. V. Pirogov, and T. G. Emel'yanova, Neorg. Mater. **29**, 1559 (1993).
3. A. V. Lykov, *Theory of Heat Conduction* (Vysshaya Shkola, Moscow, 1967).
4. L. D. Landau and E. M. Lifshitz, *Course of Theoretical Physics*, Vol. 7: *Theory of Elasticity* (Nauka, Moscow, 1987; Pergamon, New York, 1986).
5. E. É. Glikman, É. D. Ivanov, and I. A. Osipov, Mikroélektronika **19**, 132 (1990).
6. V. I. Domarkas and R.-I. Yu. Kazhis, *Control and Measuring Piezoelectric Transducers* (Mintis, Vil'nyus, 1975).
7. *Physical Quantities: A Handbook*, Ed. by I. S. Grigor'ev and E. Z. Meilikhov (Énergoatomizdat, Moscow, 1988).
8. G. A. Korn and T. M. Korn, *Mathematical Handbook for Scientists and Engineers* (McGraw-Hill, New York, 1968; Nauka, Moscow, 1984).
9. V. A. Krasil'nikov and V. V. Krylov, *Introduction to Physical Acoustics* (Nauka, Moscow, 1984).
10. A. M. Orlov, A. A. Skvortsov, A. V. Pirogov, *et al.*, Pis'ma Zh. Tekh. Fiz. **25** (5), 57 (1999) [Tech. Phys. Lett. **25**, 191 (1999)].

Translated by V. Isaakyan

Technical Physics, Vol. 48, No. 6, 2003, pp. 742–744. Translated from Zhurnal Tekhnicheskoĭ Fiziki, Vol. 73, No. 6, 2003, pp. 82–85.

SOLID-STATE ELECTRONICS

Formation of Cubic Silicon Carbide Layers on Silicon under the Action of Continuous and Pulsed Carbon Ion Beams

R. M. Bayazitov, I. B. Khaibullin, R. I. Batalov, and R. M. Nurutdinov

Zavoĭskiĭ Physicotechnical Institute, Kazan Scientific Center, Russian Academy of Sciences, Sibirskiĭ Trakt 10/7, Kazan 29, 420029 Tatarstan, Russia

e-mail: bayaz@kfti.knc.ru

Received November 5, 2002

Abstract—The structure and infrared absorption of cubic silicon carbide (β-SiC) layers produced by the continuous high-dose implantation of carbon ions (C^+) into silicon (E = 40 keV and $D = 5 \times 10^{17}$ cm^{-2}), followed by the processing of the implanted layers with a high-power nanosecond pulsed ion beam (C^+, τ = 50 ns, E = 300 keV, and W = 1.0–1.5 J/cm^2), are investigated. Transmission electron microscopy and electron diffraction data indicate the formation of a coarse-grained polycrystalline β-SiC layer with grain sizes of up to 100 nm. A characteristic feature of such a layer is the dendritic surface morphology, which is explained by crystallization from the melt supercooled well below the melting point of β-SiC. © *2003 MAIK "Nauka/Interperiodica".*

INTRODUCTION

Owing to the wide forbidden gap (E_g = 2.3–3.5 eV depending on the polytype), high thermal conductivity, as well as high physical and chemical stability, silicon carbide (SiC) is promising for high-temperature and high-frequency devices [1, 2]. Among the most advanced methods of producing SiC layers on silicon is the method of ion-beam synthesis (the implantation of C^+ ions into silicon), which is well compatible with Si device technology. This method makes it possible to form SiC layers containing a minor amount of foreign impurities at a given depth. Continuous SiC layers are obtained when the ion implantation is performed either at elevated temperatures of the silicon substrate (≥500°C) or at room temperature with subsequent high-temperature (>1100°C) and long-term (>1 h) thermal annealing. Under such conditions, epitaxial or polycrystalline cubic silicon carbide (β-SiC) layers form [3–6]. However, the use of high implantation and annealing temperatures gives rise to additional technological problems and is often undesirable.

The rapid annealing of implanted layers with high-power pulsed laser or corpuscular beams is an alternative to high-temperature implantation/annealing. Pulsed processing does not heat the bulk of the substrate and can be localized within isolated regions on the surface. Earlier, it was shown that the high dose (above 10^{17} cm^{-2}) and low-energy (below 40 keV) implantation of C^+ ions into silicon with subsequent processing by nanosecond (~10^{-8}–10^{-7} s) laser [7] or electron [8] beams leads to the growth of epitaxial or polycrystalline β-SiC layers.

Nanosecond pulsed ion-beam treatment (PIBT) has a number of advantages over laser and electron treatments. Due to the uniform energy deposition (within the ion range) [9], the synthesis of thin-film compounds can be performed under conditions of more uniform heating of near-surface layers without overheating and surface damage [10, 11]. Deep melting (≈1 μm) in combination with a high solidification rate (~10^9 K/s) enables one to form thick heavily doped layers where the concentration of implants exceeds considerably their equilibrium solubility in a crystal [12]. In addition, this method offers a high production rate [13] and makes it possible to perform implantation and annealing simultaneously [14]. In this work, we investigate the feasibility of forming continuous β-SiC layers on silicon by using the high-dose implantation of C^+ ions followed by nanosecond processing with a high-power C^+ ion beam.

EXPERIMENT

C^+ ions with an energy E = 40 keV and a dose $D = 5 \times 10^{17}$ cm^{-2} (ion-current density j < 6 μA/cm^2) were implanted into *n*-Si (100) wafers. After the ion implantation, the samples were subjected to PIBT. A pulsed ion accelerator generated wide-aperture ($d \approx 4$ cm) beams with a duration τ = 50 ns and an ion energy E = 300 keV. The pulse energy density was W = 1.0–1.5 J/cm^2, which corresponds to the ion current density j = 65–100 A/cm^2. The ion beam consisted mainly of C^+ and H^+ ions (≈80 and ≈20%, respectively). The total dose of these ions per pulse did not exceed 10^{14} cm^{-2}. To compare the methods, some of the samples were subjected to laser and thermal annealing. Pulsed laser annealing was performed with a ruby laser (λ = 0.69 μm, pulse duration τ = 50 ns, and energy density W = 0.5–2.0 J/cm^2); thermal annealing (TA), in a nitrogen atmosphere at temperatures of 950–1050°C for 10–15 min. The structure of the layers thus obtained was

examined by transmission electron microscopy (TEM) and electron diffraction. The morphology of the silicon surface and the optical absorption of the β-SiC layers were investigated by optical microscopy and infrared (IR) spectroscopy.

RESULTS AND DISCUSSION

The TEM data (not shown) demonstrate that the implantation of C^+ ions (at a dose of 5×10^{17} cm^{-2}) into the silicon produces a continuous layer of randomly oriented β-SiC grains about 10 nm in size. The microdiffraction pattern is severely blurred, which indicates that the layer implanted is highly disordered. Figure 1a shows the bright-field TEM image of the silicon surface after the implantation and PIBT with a pulse energy density $W = 1.0$ J/cm^2. It is seen that the average size of a β-SiC grain is approximately equal to 100 nm. The microdiffraction pattern contains many sharp rings (up to 12), testifying to the polycrystalline nature of the β-SiC layer obtained. The increase in the pulse energy density to 1.5 J/cm^2 leads to the formation, along with coarse β-SiC grains, of graphite grains about 100 nm in size and also causes visual darkening of the sample (Fig. 1b). The presence of the graphite grains is confirmed by an additional ring appearing in the microdiffraction pattern. Under the action of the laser pulses on the implanted layer, the graphite grains form at energy densities beginning from $W = 0.5$ J/cm^2, which is close to the Si melting threshold. The thermal annealing of the implanted layers at a temperature of 1050°C for 15 min does not result in any noticeable modification of the structure compared to the initial (postimplantation) one.

After the PIBT, the calculated melting depth in the silicon equals ≈0.8 μm at a pulse energy density $W = 1.0$ J/cm^2 [12]. This value exceeds the distribution profile width for implanted carbon ions (≈0.2 μm at 40 keV) that was calculated by using the TRIM program [15]. When formed, the carbon-containing silicon melt is already in the supercooled state with respect to the SiC melting point, since the Si and β-SiC melting points significantly differ (1700 versus 2830 K). Clearly, the radiation energy is insufficient to break Si–C bonds and melt β-SiC grains formed during the ion implantation. In this case, there appears a high probability of unoriented polynuclear crystallization, giving rise to a dendritic surface morphology with typical sizes of dendrites of up to 500 μm (Fig. 2). In [8], a similar surface morphology was observed after the processing of irradiated silicon layers with pulsed electron beams.

The method of IR absorption is one most informative method in the study of SiC formation. After the ion implantation, the wide absorption band with a maximum near ≈720 cm^{-1} is observed. This maximum is shifted toward the long-wave range with respect to the maximum of single-crystal SiC (Fig. 3, curve *1*). The thermal annealing of these samples at 1050°C for 15 min does not noticeably change the spectra. At the same time, the PIBT of the implanted layer causes the absorption band to narrow and its maximum to shift toward ≈770 cm^{-1} (Fig. 3, curve *2*), indicating that the structure of the implanted layer is crystalline. To eliminate residual defects and mechanical stresses as a result of the crystallization from the melt during the PIBT, we

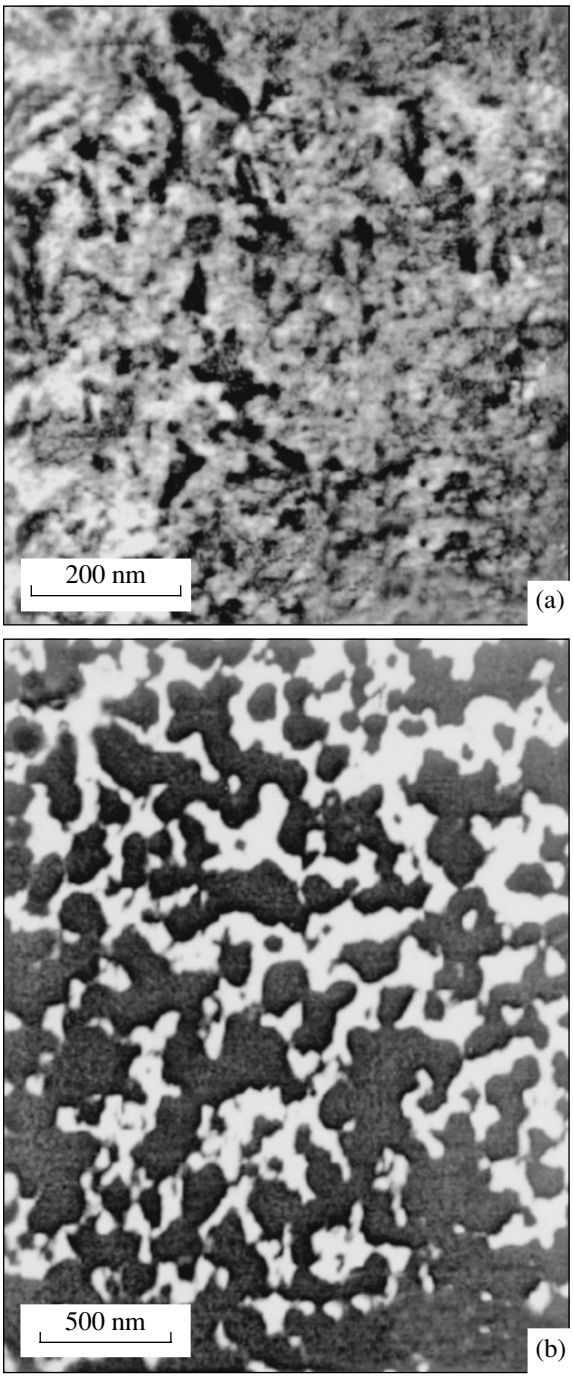

Fig. 1. Bright-field TEM image of the Si surface after implantation (40 keV, 5×10^{17} C^+/cm^2) and PIBT (300 keV, 50 ns) at the energy density W = (a) 1.0 and (b) 1.5 J/cm^2.

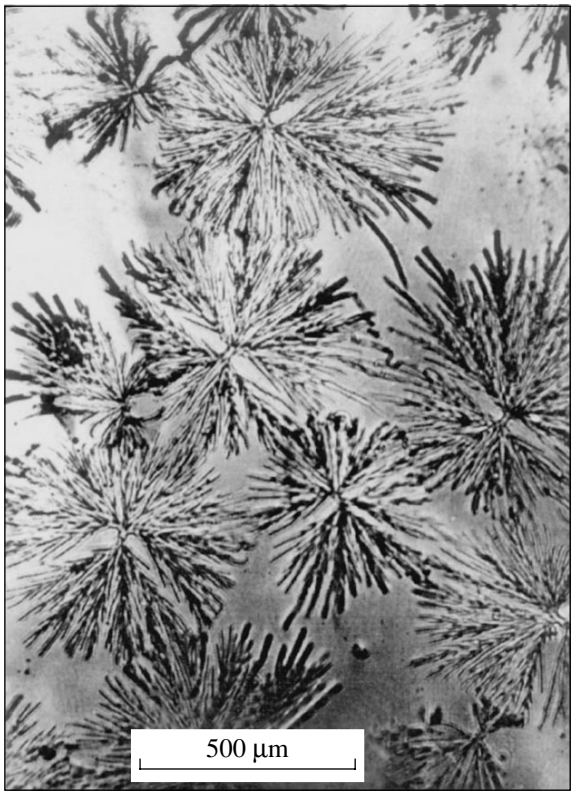

Fig. 2. Optical micrograph of the Si surface after implantation (40 keV, 5×10^{17} C^+/cm^2) and PIBT (300 keV, 50 ns) at the energy density $W = 1.0$ J/cm^2.

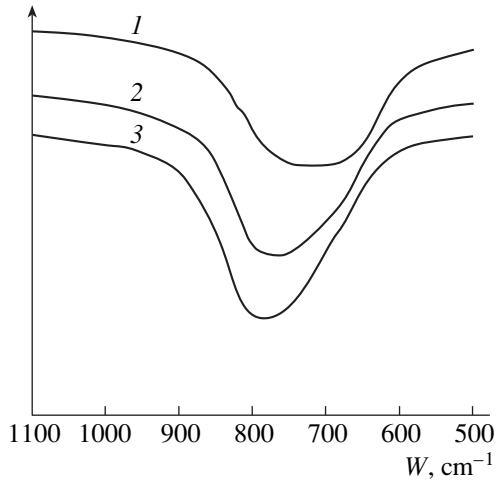

Fig. 3. IR absorption spectra of Si (*1*) after ion implantation (40 keV, 5×10^{17} C^+/cm^2), (*2*) after implantation and PIBT (300 keV, 50 ns, 1.0 J/cm^2), and (*3*) after implantation, PIBT (300 keV, 50 ns, 1.0 J/cm^2), and additional thermal annealing (950°C, 10 min). For clarity, curves *1* and *2* are shifted downward in the vertical direction.

used additional thermal annealing at 950°C for 10 min (Fig. 3, curve *3*). As is seen, the absorption band narrows further and its maximum tends toward ≈790 cm^{-1}, which is close to the value for the transverse optical mode of β-SiC ($\omega_{TO} = 794$ cm^{-1}). Thus, the most appropriate conditions for the formation of the β-SiC phase are PIBT followed by short thermal annealing.

CONCLUSIONS

The action of high-power nanosecond pulsed ion beams on silicon subjected to the high-dose implantation of carbon leads to the formation of a coarse-grained polycrystalline β-SiC layer with grain sizes of about 100 nm. With an increase in the pulse energy density, there appear (along with the polycrystalline β-SiC phase) graphite grains up to 100 nm in size. The basic mechanism responsible for the formation of the β-SiC layers is liquid-phase unoriented crystallization from the melt supercooled with respect to the β-SiC point melting. The feature of this process is the formation of dendritic surface morphology.

ACKNOWLEDGMENTS

This work was financially supported by the Russian Foundation for Basic Research (project nos. 00-15-96615 and 01-02-16649).

REFERENCES

1. P. A. Ivanov and V. E. Chelnokov, Fiz. Tekh. Poluprovodn. (St. Petersburg) **29**, 1921 (1995) [Semiconductors **29**, 1003 (1995)].
2. C. I. Harris, S. Savage, A. Konstantinov, *et al.*, Appl. Surf. Sci. **184**, 393 (2001).
3. E. K. Baranova, K. D. Demakov, K. V. Starinin, *et al.*, Dokl. Akad. Nauk SSSR **200**, 869 (1971).
4. N. N. Gerasimenko, O. N. Kuznetsov, L. V. Lezheĭko, *et al.*, Mikroélektronika **3**, 467 (1974).
5. J. K. N. Lindner, Nucl. Instrum. Methods Phys. Res. B **178**, 44 (2001) and references cited there.
6. Z. Q. Liu, J. F. Liu, J. Y. Feng, *et al.*, Mater. Lett. **50**, 275 (2001).
7. Z. Kantor, E. Foqarassy, A. Grob, *et al.*, Appl. Phys. Lett. **69**, 969 (1996).
8. P. Durupt, D. Barbier, and A. Laugier, Mater. Res. Soc. Symp. Proc. **23**, 747 (1984).
9. R. Hodgson, J. E. E. Baglin, R. Pal, *et al.*, Appl. Phys. Lett. **37**, 187 (1980).
10. L. J. Chen, L. S. Hung, J. W. Mayer, *et al.*, Appl. Phys. Lett. **40**, 595 (1982).
11. R. Bayazitov and R. I. Batalov, J. Phys.: Condens. Matter **13**, L113 (2001).
12. R. Bayazitov, L. Zakirzyanova, I. Khaibullin, *et al.*, Vacuum **43**, 619 (1992).
13. I. Isakov, V. Kolodii, M. Opekunov, *et al.*, Vacuum **42**, 159 (1990).
14. J. Piekoszewski, Z. Werner, and W. Szymczyk, Vacuum **63**, 475 (2001).
15. J. F. Ziegler, J. P. Biersack, and U. Littmark, *The Stopping and Range of Ions in Solids* (Pergamon, New York, 1985).

Translated by Yu. Vishnyakov

Technical Physics, Vol. 48, No. 6, 2003, pp. 745–748. Translated from Zhurnal Tekhnicheskoĭ Fiziki, Vol. 73, No. 6, 2003, pp. 86–89.

SOLID-STATE ELECTRONICS

Preparation of Ultrathin Gold Films by Oxygen-Ion Sputtering and Their Optical Properties

A. I. Stognij, N. N. Novitskii, S. D. Tushina, and S. V. Kalinnikov

Institute of Solid-State and Semiconductor Physics, National Academy of Sciences of Belarus, ul. Brovki 17, Minsk, 220072 Belarus

e-mail: stognij@ifttp.bas-net.by

Received November 10, 2002

Abstract—The optical and electrical properties of gold films of thickness varying from less than 1 to 8 nm are studied. The films are obtained by sputtering with argon ion and oxygen ion beams. It is shown that the properties of the films are independent of the type of ions used for sputtering. The 1- to 5-nm-thick films are continuous and offer a high transparency. Sputtering by oxygen ion beams is used to produce NiO_x/Au ohmic contacts to *p*-GaN. *© 2003 MAIK "Nauka/Interperiodica".*

It is known [1] that native oxide does not form on a gold surface and ultrathin (several nanometers thick) gold films can transmit a major part of incident radiation (the absorption coefficient is 10–20%) [2, 3]. These properties make ultrathin gold films promising for transparent protective coatings in multilayer metallic structures [4] and for transparent ohmic contacts to GaN-based light-emitting semiconductor structures [3, 5].

Problems in the application of ultrathin gold films by evaporation in a vacuum are associated with thickness control because of a high deposition rate: 0.1–1.0 μm/min [6]. Ion-beam sputtering seems to be more promising in this respect [6]. Here, the deposition rate is controlled by varying the parameters of an ion source and can be kept at a level of 1 nm/s. However, ion-beam sputtering is usually carried out in a rough vacuum (~10^{-2} Pa), when the effect of uncontrollable organic impurities, which have an adverse effect on the optical and electrical properties of the films, becomes significant [7]. The use of an oxygen plasma [8] and the oxygen-ion sputtering of targets in the oxygen atmosphere [9] are conventional methods for decreasing the organic impurity concentration in films deposited. In this work, high-conductivity transparent gold films are prepared by sputtering a gold target with oxygen ions.

EXPERIMENTAL EQUIPMENT AND INSTRUMENTS

Our ion-beam sputterer consisted of an ion source (using a self-sustained low-pressure cold-hollow-cathode discharge) for target sputtering and an open-ended modified Kaufmann ion source for substrate surface conditioning before film deposition (for details, see [9, 10]). The target, a gold plate of 4N purity, was sputtered with argon or oxygen ions of energy 0.8 keV and current density 0.3 mA/cm^2. During deposition, the Kaufmann source was switched off, argon was not passed through it, and the thermionic cathode was kept on to maintain the substrate temperature at about 120°C. The ultimate pressure in the sputterer was below 10^{-3} Pa, and the operating pressure during deposition was no more than 10^{-2} Pa.

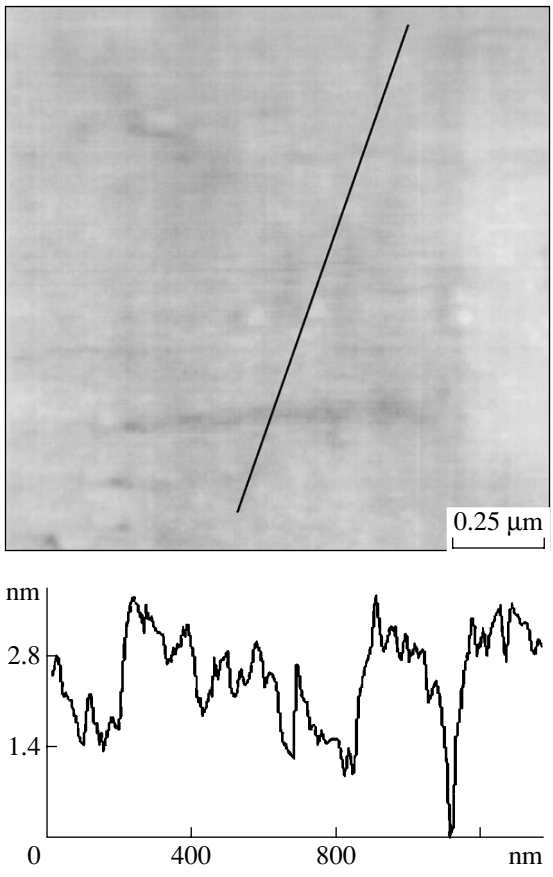

Fig. 1. AFM image of the quartz substrate surface.

1063-7842/03/4806-0745$24.00 © 2003 MAIK "Nauka/Interperiodica"

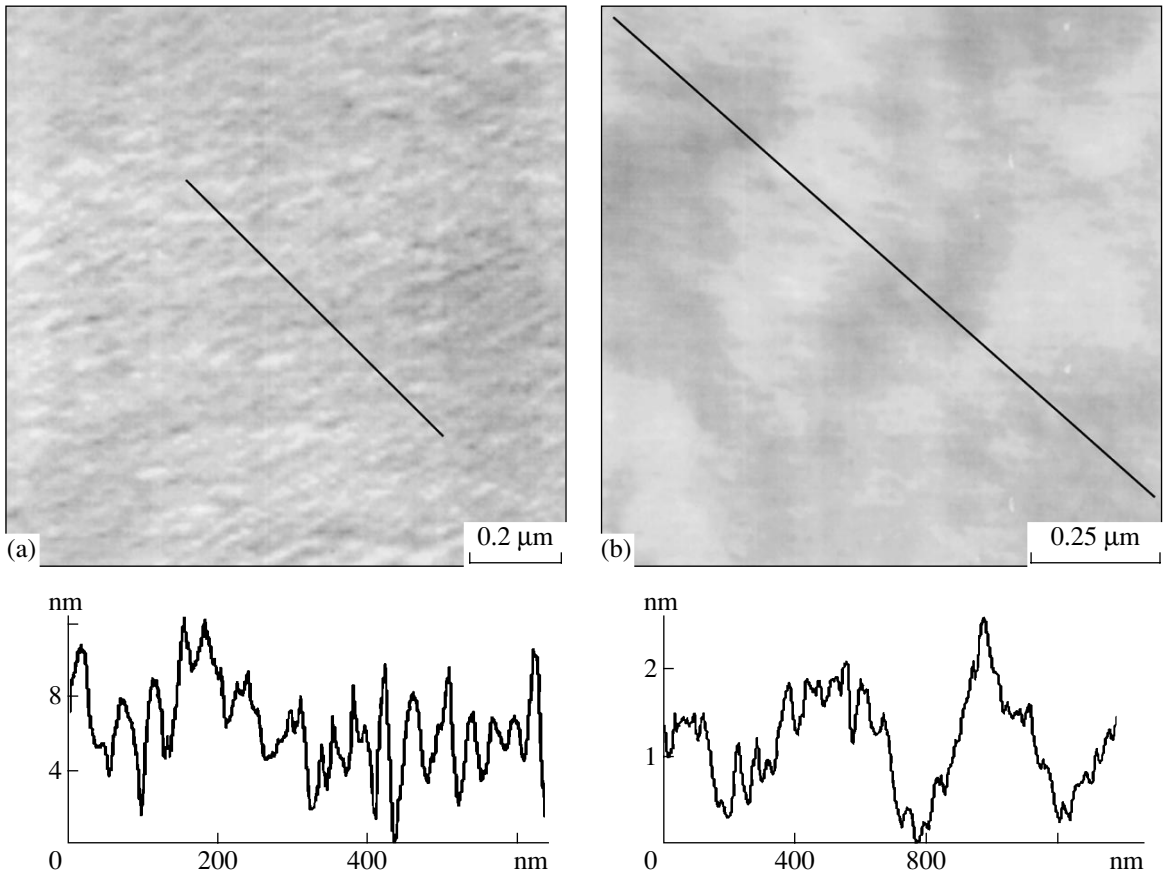

Fig. 2. AFM images of the gold films (a) 0.4 and (b) 1.2 nm thick.

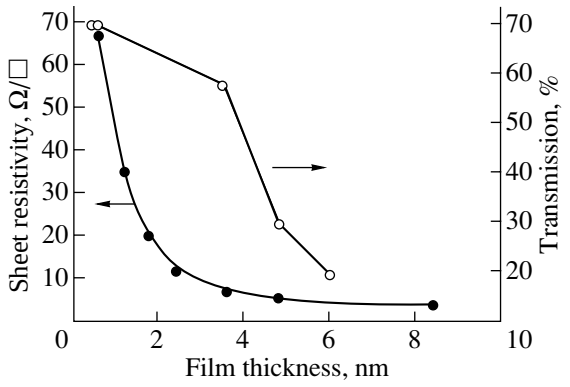

Fig. 3. Sheet resistivity and transmission (λ = 400 nm) of the ultrathin gold films deposited by means of argon ions vs. their thickness.

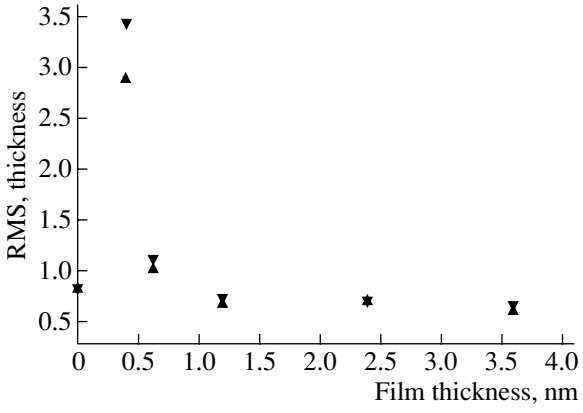

Fig. 4. Root-mean-square roughness of the gold film surface vs. thickness.

The morphology of the substrate and film surfaces was examined by a Femtoskan-001 atomic force microscope (AFM) (Center for Promising Technologies, Moscow State University) with a 5×5-μm scan field. The instrument operated in the contact mode using MicroMach (www.spmtips.com) 50-μm-long cantilevers with a tip radius of 10 nm and an apex angle of cone of 20°. The sheet resistivity of the films was measured by the four-point probe method. The transmission spectra of the films deposited were taken in the range 280–550 nm to within 0.3% with a HITACHI-340 spectrophotometer. The transmission losses

in the films were measured on a bench consisting of a He–Ne laser (λ = 632.8 nm), a rotary table with samples, and a photometry system.

RESULTS AND DISCUSSION

Figure 1 shows the AFM image of the initial surface of the quartz substrate. The surface is smooth: the height difference over a characteristic surface area of linear dimension 4 μm (along the line depicted in Fig. 1) does not exceed 2 nm and the rms roughness is less than 1 nm. Typical AFM images of gold films of thickness (a) 0.4 and (b) 1.2 nm obtained by argon-ion sputtering are shown in Fig. 2. As follows from the images in Fig. 2 and also from the measured values of the sheet resistivity illustrated in Fig. 3, the properties of the films less than 1 nm thick differ greatly from those of the thicker ones. The ultrathin films (Fig. 3) exhibit a very high sheet resistivity, which drops sharply with increasing thickness, and a high surface roughness: the rms roughness is more than 3 nm and the height difference over a characteristic surface area is several times that on the substrate surface. This leads us to suggest that these films are discontinuous and consist of grains up to 6 nm in height and 50 nm across randomly distributed over the substrate (Fig. 2a). The surface of the thicker films (Fig. 2b) is smoother: the rms roughness is less than 1 nm and varies with thickness only slightly (Fig. 4). For the 4-nm-thick film, the sheet resistivity decreases to 5 Ω/□ and the optical transmission amounts to more than 40% (the transmission and absorption are, respectively, 46 and 21%) (Fig. 3). The transmission drops as the film thickens largely because of an increased absorption. Such behavior differs from that of silver films with a comparable thickness, where the transmission changes with thickness because of a more intense reflection with the absorption remaining nearly constant [2]. The weak thickness dependence of the rms roughness (Fig. 4) suggests that, once a continuous primary gold layer has formed on the substrate, the film grows uniform in thickness and the film/substrate interface copies the relief of the substrate surface (cf. Figs. 1 and 2b).

The sheet resistivity and the transmission spectra of the ultrathin films obtained by oxygen-ion sputtering are close to those of the films obtained by sputtering with argon atoms. However, in the case of oxygen, the deposition rate is nearly twice as low as that in the case of argon as a working gas under the same experimental conditions (pressure as well as ion beam density and energy). The morphologies of the films were also similar to each other; however, in the case of oxygen, the surface of the films is smoother and the rms roughness is somewhat smaller (Fig. 4). The fact that the deposition rate in the case of oxygen sputtering is two times lower than in the case of argon can be explained in terms of interaction between molecular oxygen ions O_2^+ and the target surface [6]. It is believed that, when striking a target, an oxygen molecule decomposes into two atoms, with the energy of the molecule being equally shared by the atoms. The oxygen ion source used in this work generated more than 60% O_2^+ and more than 25% O^+ [11]. Therefore, under the same experimental conditions and in view of the fact that the mass of an oxygen atom is 2.5 times smaller than that of an argon atom, the energy contribution of oxygen

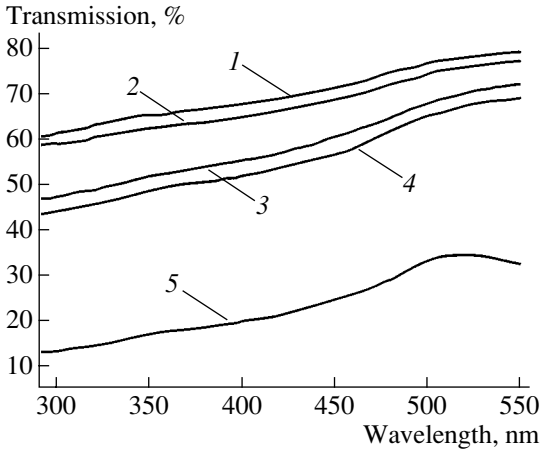

Fig. 5. Transmission spectra of the gold films deposited by means of (*1*, *4*) oxygen and (*2*, *3*, *5*) argon ions. The film thickness is (*1*) 3.0, (*2*) 0.4, (*3*) 3.6, (*4*) 4.0, and (*5*) 4.8 nm.

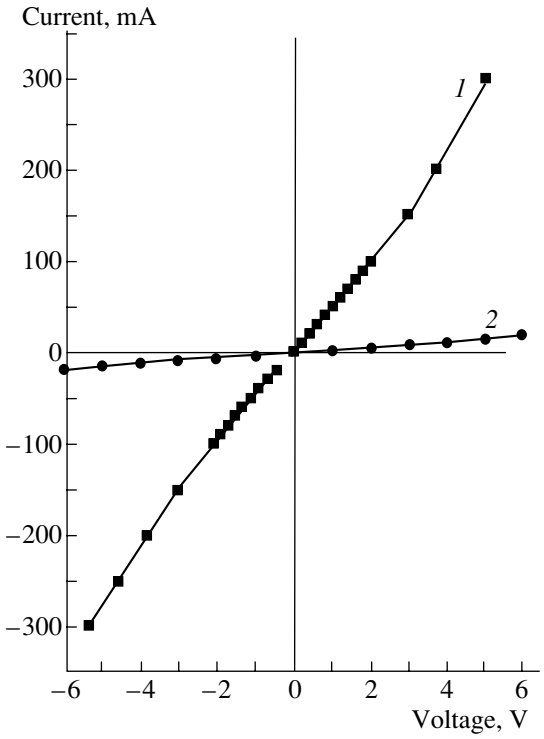

Fig. 6. *I–V* characteristics of the contacts to *p*-GaN. (*1*) *p*-GaN/NiO_x(2 nm)/Au (4 nm)/SiO_2 and (*2*) *p*-GaN/NiO_x (4 nm)/Au (4 nm)/SiO_2.

ions to the sputtering process is more than 30% lower than the energy contribution of argon ions.

For the gold films of thickness less than 5 nm obtained both by argon- and oxygen-ion sputtering, the transmission spectra in the range 290–550 nm were taken (Fig. 5). The transmission is seen to depend not on the type of sputtering ions but on the film thickness. The films with thicknesses from 1 to 4 nm have a transmission coefficient of more than 40%. Moreover, they are continuous with a sheet resistivity ranging from 25 to 5 Ω/□.

The correspondence between the properties of the films deposited by means of argon and oxygen ions indicates that oxygen is promising as a working gas for gold film deposition, especially when the films are used in two-layer ohmic contacts to *p*-type III–V semiconductors. These contacts consist of a lower nanometer-thick *p*-type metal oxide film and an upper gold layer of thickness less than 10 nm [5]. In this case, the surface of *p*-type GaN can be covered first by a nickel oxide layer (by sputtering a nickel target with oxygen ions) and then by a gold film less than 6 nm thick in a single vacuum cycle. In this way, we managed to produce *p*-GaN/NiO_x(2 nm)/Au (4 nm)/SiO_2 and *p*-GaN/NiO_x (4 nm)/Au (4 nm)/SiO_2 film structures with a transmission coefficient of more than 50% for wavelengths ranging from 290 to 550 nm. These structures have transmission spectra similar to those shown in Fig. 5, and their *I–V* characteristic is ohmic (Fig. 6).

CONCLUSIONS

We studied the thickness dependences of the optical and electrical properties of gold films with thicknesses from less than 1 to 8 nm. The films less than 1 nm thick are discontinuous: they consist of grains randomly distributed over the substrate surface and poorly connected electrically. The films with a thickness of 5 nm are continuous, highly transparent, and promising for ohmic contacts to GaN light-emitting structures. The films can also be used as protective films for multilayer nanostructures that degrade when contacting with the substrate surface and environment.

ACKNOWLEDGMENTS

We thank M. Heuken, vice-president of AIXTRON AG, for placing the gallium nitride samples at our disposal; O.M. Stukalov for the assistance in the AFM studies; and G.P. Yablonskiĭ and E.V. Lutsenko for valuable discussions and criticism.

This work was partly supported by the ISTC (project no. B-176).

REFERENCES

1. *Thin Films: Interdiffusion and Reactions*, Ed. by J. M. Poate, K. Tu, and J. Meier (Wiley, New York, 1978; Mir, Moscow, 1982).
2. G. S. Landsberg, *Optics* (Nauka, Moscow, 1976).
3. J. K. Sheu, Y. K. Su, G. C. Chi, *et al.*, Appl. Phys. Lett. **74**, 2340 (1999).
4. A. I. Stognij, N. N. Novitskii, and O. M. Stukalov, in *Proceedings of the 18th International School–Seminar "New Magnetic Materials in Microelectronics"* (Moscow, 2002), pp. 303–305.
5. H. Jin-Kuo, J. Chang-Shyang, C. C. Chien, *et al.*, Appl. Phys. Lett. **74**, 1275 (1999).
6. *Sputtering by Particle Bombardment III: Sputtering of Alloys and Compounds, Electron and Neutron Sputtering, and Surface Topography*, Ed. by R. Behrisch and K. Wittmaak (Springer, Berlin, 1983; Mir, Moscow, 1986).
7. P. Duval, *High Vacuum Production in the Microelectronic Industry* (Elsevier, 1988).
8. S. Tada, Y. Sakamoto, T. Suzuki, *et al.*, Vacuum **53**, 321 (1999).
9. A. I. Stognij, V. T. Svirin, S. D. Tushina, *et al.*, Prib. Tekh. Éksp., No. 3, 151 (2001).
10. A. I. Stognij, N. N. Novitskii, and O. M. Stukalov, Pis'ma Zh. Tekh. Fiz. **28** (1), 39 (2002) [Tech. Phys. Lett. **28**, 17 (2002)].
11. A. I. Stognij and V. V. Tokarev, Prib. Tekh. Éksp., No. 3, 142 (1990).

Translated by V. Isaakyan

Technical Physics, Vol. 48, No. 6, 2003, pp. 749–756. Translated from Zhurnal Tekhnicheskoĭ Fiziki, Vol. 73, No. 6, 2003, pp. 90–98.

OPTICS, QUANTUM ELECTRONICS

Photochromic Reactions in Silver Nanocomposites with a Fractal Structure and Their Comparative Characteristics

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

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

e-mail: karpov@iph.krasn.ru

*** Krasnoyarsk State Technical University, Krasnoyarsk, 660074 Russia*

Received October 2, 2002

Abstract—Conditions for a change in the polarization selectivity of dips in the plasmon absorption spectra of fractal silver nanocomposites irradiated by pulsed laser radiation are studied. The energy thresholds of the polarization selectivity are evaluated, and the polarization and spectral threshold characteristics are compared. Mechanisms behind the correlation between the fractal structure of the nanocomposites, on the one hand, and their optical and photochromic properties, on the other hand, are discussed. *© 2003 MAIK "Nauka/Interperiodica".*

INTRODUCTION

Photochromic processes in silver nanocomposites containing fractal clusters, or fractal aggregates (FAs), were discovered and characterized for the first time in [1]. These processes in fractal media are of interest for many reasons. First, they may be related to the unique nonlinear optical properties of FAs [2]. The giant (10^6 times) enhancement of degenerate four-photon parametric scattering (DFPS) upon silver particle aggregation, the effect observed as early as in pioneering experiments [3, 4], gives special significance to the study of such media, including the need to estimate the role of attendant photochromic processes. Second, such media appear to be promising for polychromic data recording [5].

The effect discovered in [1] was as follows. The pulsed laser irradiation of colloidal silver aggregates embedded in gelatin matrices (the volumetric aggregate concentration was on the order of 10^{-5}) "burned out" a relatively narrow (80–100 nm) long-lived dip near the irradiation wavelength in the spectrum of absorption of light with the same linear polarization as that of the irradiating laser pulse. Furthermore, the width of the dip was close to the plasmon absorption bandwidth for an individual particle.

In transmitted light, one can see a spot near the irradiation area that has the same color as incident radiation. This phenomenon is associated with the so-called photomodification of FAs (the photoinduced modification of their structure), which occurs in the case of the local (spatially selective) absorption of radiation with a given frequency and polarization by resonant particles of FAs.

This phenomenon has been clearly explained by invoking the theory of optical properties of fractal clusters (the OPFC theory) [6–8], which establishes a one-to-one correspondence between the structural and optical properties of FAs.

The aim of this work is (i) the determination of energy thresholds for the polarization selectivity of photoinduced dips in the absorption spectra taken from association silver colloids and (ii) the interpretation of related data in terms of the fractal structure of the aggregates, which is responsible for a number of unique properties of these objects.

OPTICAL PROPERTIES OF FRACTAL SILVER NANOCOMPOSITES

The optical properties of silver nanoparticles are specified by surface plasmon resonance with a near-Lorentzian profile of the absorption band with a uniform width of about 80 nm and a peak at $\lambda \simeq 420$ nm in hydrosols.

It is known that, during the evolution of a sol, the majority of isolated particles aggregate to form fractal structures (see references in [6]). Such structures are observed, in particular, when clusters are deposited on a substrate during the laser evaporation of a metallic target [9].

When combined into FAs, the particles cannot be considered as isolated; therefore, the interaction of radiation-induced oscillating dipole moments should be taken into account. Let us consider the effect of such interaction on the absorption spectra in greater detail.

Works [6, 7, 9, 10] have evidenced the strong influence of nearby particles on the spectrum of any particle incorporated into an FA. This causes a substantial shift in the frequency ω_r of natural optical resonances of the particles. The reason for the shift is electrodynamic par-

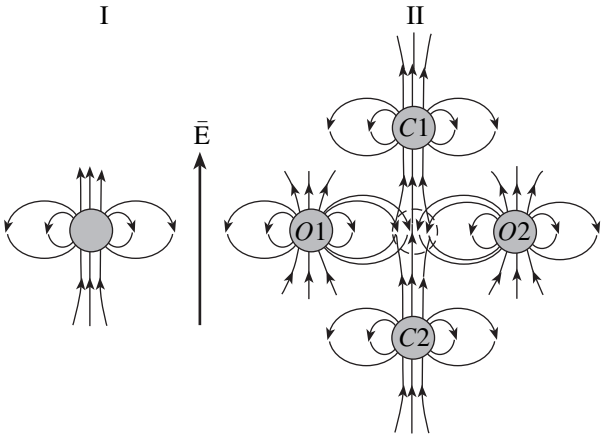

Fig. 1. Formation of the self-consistent field around an *i*th particle (marked by the dashed circle at the center) upon the dipole–dipole interaction of the particle in a varying electric field **E** (for the symmetric environment of the *i*th particle). I, configuration of the lines of electric force of the self-induced dipole at the *i*th particle; II, configuration of the lines of electric force of the dipoles induced by pairs of particles that are collinear with (*C*1, *C*2) and orthogonal to (*O*1, *O*2) the field at the site of the *i*th particle.

ticle interaction (with the prevailing effect of the nearest particle). It was shown (e.g., in [6]) that the effect of nearby particles shows up as a split of the intrinsic resonance (with a frequency ω_r) of a single initially isolated particle into two, lower frequency (ω_l) and higher frequency (ω_h), peaks. It is noteworthy that the shift of the lower frequency resonance is nearly twice as large as that of the higher frequency one [6, 10]. Because of this, the plasmon absorption spectrum of FAs broadens significantly but asymmetrically, covering the entire visible range and a part of the IR range (see, e.g., [10]).

Due to the interaction of a particle with its nearest neighbor, the resonance frequency shifts by a value of $\Delta\omega \approx \Omega_f$, where Ω_f is the characteristic (mean) value of the frequency shift. This parameter defines the characteristic scale of the broadening and of the low-frequency shift (and, accordingly, the extension of the long-wave wing) of the FA spectrum relative to the spectrum of an isolated particle [8]:

$$\Omega_f = \omega_m(R_n/R_0)^3, \quad (1)$$

where R_0, R_n, and ω_m are, respectively, the characteristic spacing between neighboring particles in an FA, particle size, and resonance frequency of the particles. In other words, the frequency shift is inversely proportional to the particle spacing cubed. In the presence of an adsorption layer around the particles, this spacing may exceed R_n only slightly; hence, $\Omega_f \simeq \omega_m$. That is to say, the effect of electrodynamic interaction between nearby particles in a fractal is so high that their resonances at frequencies comparable to the natural resonance frequency shift. An increase in the particle spacing, e.g., by thickening the adsorption layer or increasing the particle size, would decrease the resonance shifts and, accordingly, shrink the long-wave wing of the fractal spectrum.

If the frequency of external radiation falls into the fractal's absorption band, there will always be found resonant particles (or, more specifically, pairs of resonant particles) that are far apart within an FA. The fraction of such particles is $\Gamma/\Omega_f \ll 1$, where Γ is the plasmon absorption bandwidth. Therefore, different parts of the fractal will absorb monochromatic radiation independently because of the heavily asymmetric broadening (due to the dispersion of pair interactions) of the absorption spectrum of the fractal.

In hydrosols, the long-wave wing of the spectrum expands when FAs form. This is because of the great diversity of possible environments near a particular particle involved in an FA. In this case, the relative fraction of closely spaced pairs of particles grows. In addition, as the FA circumference elongates, the local anisotropy of the environment around the particles is enhanced. The appearance of the short-wave branch of the FA spectrum, which is twice as short as the long-wave one (on the frequency scale), is associated with a high-frequency peak of interacting pairs that appears in the absorption spectrum.

Electrodynamic interactions in a set of closely spaced metal particles are a ready illustration of how the fractal geometry and local anisotropy of association colloids influence their optical properties. Qualitatively, this effect is described in Fig. 1.

Consider the isotropic case. A feature of dipole–dipole interaction is that the self-consistent field produced near an *i*th particle by all other fractal particles vanishes when averaged over the spherically symmetric particle distribution. From Fig. 1 it follows that lines of electric force of the dipoles induced on particles that are oriented orthogonally to and collinearly with the applied electric field enter the *i*th particle in opposite directions and almost completely cancel each other (provided that the contribution from all its neighbors is strictly averaged). Obviously, a net addition to the local field induced by all the particles evenly distributed around the *i*th particle tends to zero as the environment of the *i*th particle becomes progressively more isotropic. Note that such a statement holds only in the 3D case, when averaging must take into account not only collinear pairs oriented along the *Z* axis but also orthogonal pairs oriented along the *X* and *Y* axes (the number of the latter pairs turn out to be twice that of collinear pairs). The need for considering collinear pairs is dictated by the fact that, in dipole–dipole interaction, the interaction energy of collinear pairs is twice as high as that of orthogonal ones [11].

The situation changes drastically when the particles are integrated into FAs, i.e., when most particles around a given one are surrounded by either orthogonally oriented or collinear pairs (in terms of their projections to the field direction). In other words, the environment of

an ith particle is severely anisotropic under these conditions. Note that colloidal FAs typically produce branched chain structures. Certainly, one may find a number of less anisotropic local configurations of the particles. In Fig. 1, they would appear if one of the particles in collinear or orthogonal pairs were absent, i.e., if the self-consistent field were generated by three (or five in the 3D case) surrounding particles.

Let us consider the effect of local anisotropy for the highest anisotropy case when only orthogonal ($O1$, $O2$) or only collinear ($C1$, $C2$) pairs are left in Fig. 1. In terms of this simplified model, the addition to the local field around an ith particle is negative if the orthogonal pairs are left. It will also be recalled that, for this configuration, the frequency of the applied field falls into the short-wave part of the fractal spectrum.

For a constant initial strength of the applied field and the same damping constants, this is equivalent to a decrease in the polarizability χ of the ith particle in an aggregate under the effect of the orthogonal pairs. In turn, a decrease in the polarizability is equivalent to an increase in the detuning $\Omega = |\omega_r - \omega|$ ($\chi \sim \Omega^{-1}$) because of the shift of ω_r toward higher frequencies. That is the reason why the high-frequency branch of the FA absorption spectrum somewhat broadens.

Using a similar line of reasoning for the collinear ($C1$, $C2$) pairs and taking into account that the addition to the local field near the ith particle is now positive, we arrive at the conclusion that this situation is equivalent to the case when the polarizability of the particle grows. Here, an increase in the polarizability of the ith particle within a fractal due to dipole–dipole interaction corresponds to a decrease in the detuning $\Omega = |\omega_r - \omega|$ because of the shift of the resonance frequency ω_r toward lower frequencies. This explains the appearance of the long-wave wing in the FA absorption spectrum.

Quantitatively, corrections to the local fields and to the positions of resonant absorption bands were considered in detail by Aver'yanov [12–14]. His model has been used to advantage in liquid-crystal optics.

Let us analyze several expressions that shed light upon the problem solved in [12–14]. To a linear approximation, with the Lorentz tensor L introduced in the expression $\mathbf{P} = N\chi_0 f\mathbf{E}$, a shift of the absorption band maximum in response to a change in the local field is given by formula (2). In the above expression, $\mathbf{P}$ is the polarization of a medium that is related to a macroscopic field $\mathbf{E}$ (including a correction for depolarization), $f = 1 + L(\varepsilon - 1)$ is the local-field tensor, χ_0 is the tensor of the dipole polarizability of particles, and N is the particle concentration in the medium. Then [12],

$$\omega_r' = \omega_r - \frac{\omega_p^2 F_{12} L_j (f_j)_b}{2\omega_r}\left(1 - \frac{(f_j)_b}{6L_j(\varepsilon_j)_b}\right). \qquad (2)$$

Here, ω_r is the initial resonance position with regard for the environmental influence, ω_r' is the resonance position after the local field has changed, ω_p is the plasma frequency, F_{12} is the transition oscillator strength (in our case, it corresponds to the excitation of a particle's surface plasmon), L_j are the diagonal components of the Lorentz tensor, $j = \{xx, yy, zz\}$ (in general, $L_j = 1/3 - B_j/4\pi N$), B_j is the sum over dipoles within the Lorentz sphere (in particular, $B_{xx} = \Sigma(1/r_i^3 - 3X_i^2/r_i^5)$, where $\mathbf{r}_i$ is the radius vector of an ith particle in the Lorentz sphere with the origin placed at the center of the sphere and X_i is the coordinate of the vector), $(f_j)_b = 1 + L_j[(\varepsilon_j)_b - 1]$ are the diagonal components of the local-field tensor that take into account an increase in the effective local field strength with respect to the strength of the macroscopic field of a light wave in the medium ($\mathbf{E}_{\text{eff}}^{\text{loc}} = f\mathbf{E}$), $(\varepsilon_j)_b = (\varepsilon)_b\delta_{ij}$ is the background permittivity imposed by all resonances in the medium except for that under consideration, and δ_{ij} is the Kronecker symbol. The product $L_j(f_j)_b$ is responsible for the configuration and interaction of the particles in the system.

For the isotropic arrangement of the particles, the equilibrium value of the diagonal components is $L_j = 1/3$. For pairs of particles that are collinear with the field (when the field is directed along the Z axis), the maximal value of L_j in expression (2) tends to unity (the components L_j and ε_j with $j = \{zz\}$ are used). For orthogonal pairs, L_j tends to zero (L_j and ε_j with $j = \{xx, yy\}$ are used).

With such values of L_j, the parenthesis in (2), which specifies the sign of the resonance frequency shift $(\omega_r' - \omega_r)$, can be either positive or negative depending on the particle configuration. Note that expression (2) is accurate if the inequality $F_{12}/3 \le \omega_r\Gamma(\varepsilon_j)_b/\omega_p^2(f_j)_b^2$ holds (medium- and low-absorption bands). However, it is helpful in qualitatively treating the effect.

For strong-absorption bands, we have expression

$$\omega_r' = \omega_r - \frac{\omega_p^2 F_{12} L_j (f_j)_b}{2\omega_r} + \Gamma/2\sqrt{3}, \qquad (3)$$

which yields the same law for frequency shift sign under similar conditions.

Thus, since the environments of most of the particles in an FA are locally anisotropic, additions to the local field do not cancel each other (Fig. 1). For this reason, the effect of significant enhancement of local electromagnetic fields [15] is inherent mostly in fractal structures. To check this assumption, the absorption spectra of Ag sols with a stochastically uniform particle distribution were calculated in terms of the OPFC theory [7]. For this case, the broadening of the absorption bands appeared to be the same as for media with ordered particle distributions or for media with isolated particles.

Note also that the FA spectra will be broader for metal particles, which have larger matrix elements of transition electric dipole moments that characterize

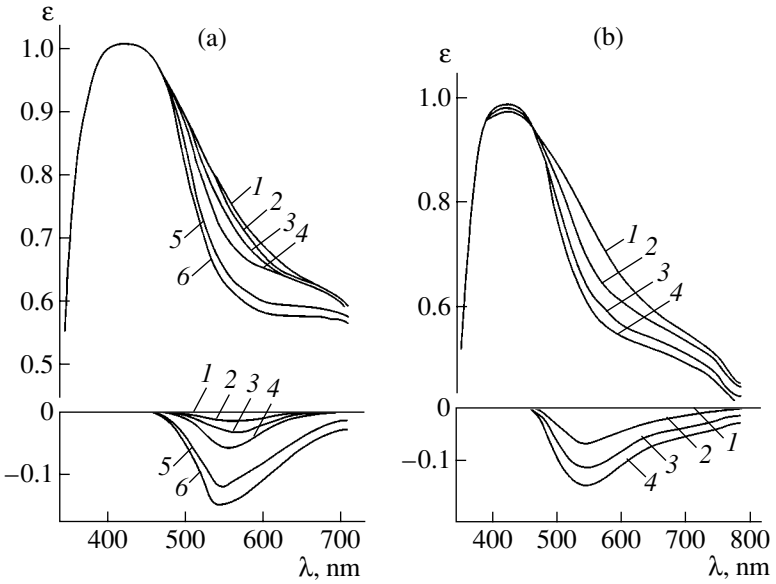

Fig. 2. Absorption spectrum (at the top) and differential spectrum (at the bottom) of irradiated and (*1*) nonirradiated silver aggregates subjected to picosecond and nanosecond pulses with λ = 540 nm. (a) The number of pulses with W = 2 mJ/cm^2 and τ = 30 ps is (*2*) 1, (*3*) 10, (*4*) 20, (*5*) 80, and (*6*) 230. (b) Single-pulse irradiation with W = (*2*) 8, (*3*) 10, and (*4*) 12.4 mJ/cm^2; τ = 15 ns.

optical resonance. For particles that have optical polarizabilities lower than those of metal particles, dipole–dipole interaction causes a smaller shift of resonances.

Analytical expressions that relate quantitatively the fractal dimension of aggregates to the broadening of their absorption spectra can be derived only in the binary approximation to the OPFC theory (in view of the well-known limitations involved in this approximation) [8].

A description of the FA absorption spectra that is based on the rigorous OPFC model is given elsewhere [6, 7, 10]. In this case, however, an analytical relationship between the broadening and fractal dimension is impossible to find, because (i) the model is additive and (ii) the number of equations for FA polarizability by which the properties of the model are described exceeds N_0^3 (N_0 is the number of particles). Yet, a correlation between the absorption spectrum and the fractal dimension persists, since fractal geometry is featured by aggregates over which the summation is carried out.

EXPERIMENTAL RESULTS AND DISCUSSION

FA photomodification is most pronounced in media with a near-limiting degree of association. Association silver sol consists of aggregates 0.1–1.0 μm or more in size. Aggregates, in turn, integrate particles of radius 5–10 nm. Comparing the number of particles and the sizes of aggregates in micrographs yields the fractal dimension $D \simeq 1.7$, which is typical of such objects.

The plasmon absorption spectrum of a hydrosol with unassociated Ag particles has a sharp peak at $\lambda_{pl} \simeq$ 420 nm and Γ = 80–100 nm. Of association media, a broad absorption band in the visible and near-IR (λ = 1.0–1.5 μm) ranges is characteristic (see, e.g., [10]).

Aggregated Ag sols (or Ag aggregates in gelatin) were irradiated by second-harmonic pulses from $Y_3Al_5O_{12}$: Nd and $YAlO_3$: Nd lasers (λ = 532 and 540 nm, respectively) and also by radiation generated by stimulated Raman scattering in acetone (λ = 641 nm). Giant (τ ≃ 10 ns) and ultrashort (τ ≃ 30 ps) pulses were applied.

Figures 2 and 3 show irradiation-induced changes in the absorption spectra of the sols, i.e., dips near the lasing wavelengths. It is seen that the depth and width of the gaps grow with both the number and energy of irradiation pulses. This is because the burning out of the dips is a process of a threshold nature (or highly nonuniform in energy). In the hydrosol (Figs. 2a, 2b) at τ = 15 ns, the spectral dip near λ = 532 nm is observed at a pulse energy density W = 4–10 mJ/cm^2. For picosecond pulses, the threshold energy can be several times lower than for nanosecond radiation.

The characteristic width of the dips (about 3500 cm^{-1}) is close to the symmetric width of the absorption band for individual particles in an unassociated hydrosol (3500–5000 cm^{-1}) and is much smaller than the width of the FA absorption band (more than 20 000 cm^{-1}). The narrowest dips (1800 cm^{-1}) were observed in the spectra of aggregates recorded in holographic emulsion [1]. It should be noted that an asymmetric broadening (within Δλ = 15 nm) of the absorption band is observed even in the unassociated hydrosol due to its polydisper-

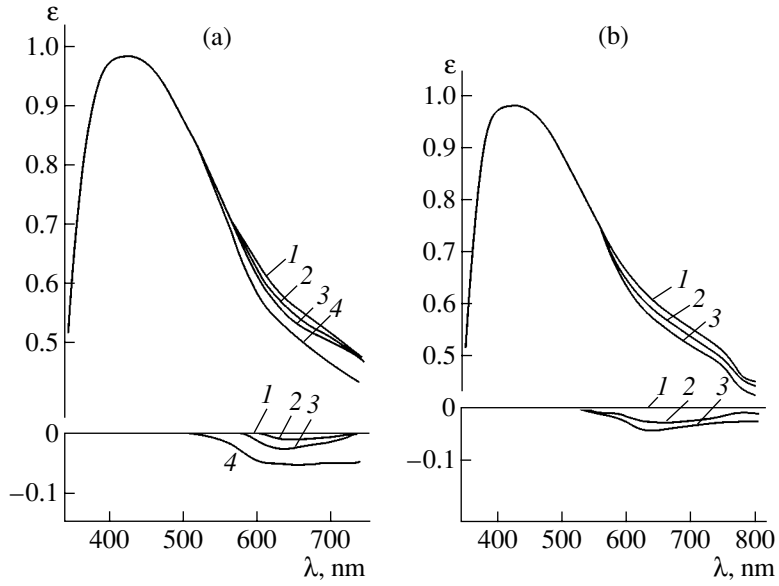

Fig. 3. The same as in Fig. 2 for λ = 641 nm. (a) (*2*) 20 pulses with W = 2 mJ/cm^2, (*3*) 120 pulses with W = 2 mJ/cm^2, and (*4*) one 8-mJ/cm^2 pulse; τ = 30 ps. (b) (*2*) 20 and (*3*) 40 pulses with W = 4 mJ/cm^2; τ = 15 ns.

sity and a spread in the resonance frequency for particles from 5 to 30 nm in size. Also, even microaggregates present in the sol cause an additional broadening of the plasmon absorption band.

According to today's concepts, dips related to photomodification are explained by burning out particles for which the shift of the resonance frequency (because of dipole–dipole interaction) equals the incident radiation frequency. It is this burning out (breakage) that is responsible for the threshold character of the photomodification process.

Estimates show that each resonant particle of a fractal for W = 1.5 mJ/cm^2 and a particle concentration of 10^{12} cm^{-3} absorbs about 3×10^5 photons per pulse. This value equals or somewhat exceeds the energy needed to sublimate a particle with a radius of 10 nm. With τ = 30 ps, the area of absorbed energy localization, $d \propto \sqrt{\tau}$, is on the order of the particle size; with τ = 15 ns, this area covers several particles. This is the reason for the high value of the threshold energy and the partial loss in frequency and polarization selectivity in the case of long pulses, which is observed in experiments.

Among the specific features of photomodification, of interest is the fact that, roughly within an hour after the irradiation of the entire cuvette with the hydrosol, a spectral dip may broaden and shift slightly toward longer waves. In this case, the peak due to plasmon absorption somewhat grows. This finding may be related to photostimulated aggregation [16–19] when microaggregates or a small fraction of free particles are present in the medium.

Stable (at least for a month) dips in the absorption spectrum were observed for silver aggregates in gelatin (Figs. 2, 3). Still more stable dips were found upon using polyacrylamide matrices in our experiments. For example, a number of the samples irradiated have retained spectral selectivity since 1987. Such a slow relaxation can be attributed to the fact that here particles have a lower mobility than in the liquid.

Note that the threshold energy of photomodification may somewhat increase if sol particles are surrounded by a polymeric adsorption layer.

Along with the dip near the laser wavelength in the FA absorption spectrum, there appears polarization selectivity in the irradiation zone, which depends on the incident radiation polarization. This effect is explained by the fact that resonance in the collinear configurations of pairs, shifts, as was already noted, toward lower frequencies relative to the resonance frequency of an isolated particle ($\omega_r' < \omega_r$). For orthogonal pairs, a shift toward higher frequencies ($\omega_r' > \omega_r$) is observed. Hence, an important conclusion follows: when interacting with the aggregate, linearly polarized monochromatic radiation must impart nearby particles a certain arrangement. It will be recalled that a radiation with a wavelength from the long-wave wing interacts only with collinear pairs.

Figure 4 demonstrates the variation of the dip near the laser wavelength λ = 540 nm (τ = 15 ns) in the differential absorption spectrum of aggregated silver (embedded in the gelatin matrix). Bearing in mind that polarization selectivity arises when the radiation interacts with collinear or orthogonal pairs of particles, one can infer that, in an FA, the radiation interacts with both simultaneously. Accordingly, a radiation with a wavelength from only the long-wave or short-wave branch of

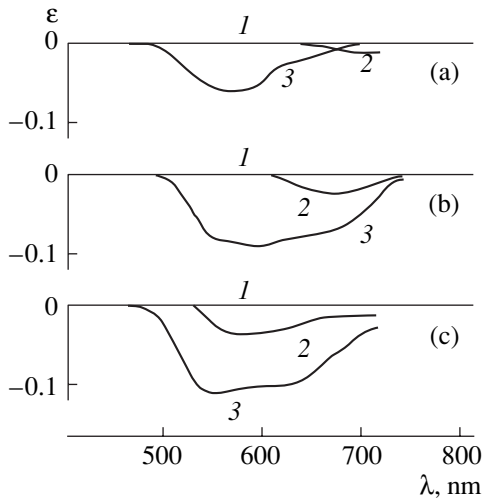

Fig. 4. Variation of the polarization of the spectral dip (in the differential absorption spectra) upon irradiating by pulses with $\lambda = 540$ nm and $\tau = 15$ ns. W = (a) 7, (b) 30, and (c) 60 mJ/cm^2. (*1*) Nonirradiated medium, (*2*) irradiated medium with linear polarization orthogonal to incident light, and (*3*) polarization corresponding to that of incident light.

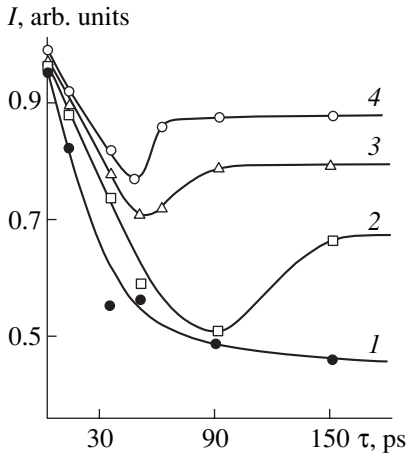

Fig. 5. Amplitude of the optical response to DFPS in the aggregated silver hydrosol vs. elapsed time from the application of the pulses ($\lambda = 540$ nm, $\tau = 30$ ps). W = (*1*) 8.25, (*2*) 20.76, (*3*) 24.75, and (*4*) 30 mJ/cm^2.

the absorption spectrum causes two dips to appear simultaneously: one at $\lambda > \lambda_{pl}$ and the other at $\lambda < \lambda_{pl}$ with their polarizations being mutually orthogonal.

As follows from Fig. 4, the increase in the pulse energy density from 7 to 60 mJ/cm^2 leads to a broadening of the low-frequency dip primarily to the long-wave wing. This is related to the fact that nonresonant pairs of particles start being involved in the process with the resulting degradation of the spectral selectivity. Also, the dip in the orthogonal polarization becomes deeper possibly because of the pulsed-induced break of local anisotropy near resonant particles.

The aforesaid is in complete agreement with current concepts [20] of processes disturbing the locality of interactions above the energy threshold.

To observe the polarization effect upon burning out the spectral dip, the FAs irradiated should be spatially fixed with the use of nondestructive matrices. The relaxation times of photochromic processes far exceed the lifetime of the polarization effect. For example, the visually estimated time of tint diffusion blurring in the irradiation zone of hydrosols is not more than a few seconds. As for polarization selectivity, its lifetime τ_l in the irradiation zone is several microseconds or less, depending on the rotational diffusion of the aggregates upon Brownian motion. As is known, the angle of rotation Θ_0 of a Brownian particle and the time of rotation are related by the expression

$$\sin^2\Theta_0 = 2/3[1 - \exp(-6D_{rot}t)],$$

where $D_{rot} = kT/8\pi\eta(2R_a)^3$ is the rotational diffusion coefficient, k is the Boltzmann constant, T is the ambient temperature, $2R_a$ is the dimension of an aggregate, and η is the dynamic viscosity.

With $\eta = 0.001$ Pa s [16], $a = 10^{-7}$ m, and $\Theta_0 = 45°$, the rotational diffusion time corresponding to this angle is on the order of 10^{-4}–10^{-3} s. From this value, the lifetime τ_l can be found.

To estimate the evolution time of the photomodification process, we experimentally studied the kinetics of degenerate four-photon parametric interaction (under conditions of phase conjugation) with FAs in Ag hydrosols at different energy densities of picosecond second-harmonic ($\lambda = 540$ nm) pulses from a $YAlO_3$: Nd laser [21], including those exceeding the energy threshold of photomodification.

A diffraction grating in the medium under study was formed by applying 30-ps pulses, with the probe pulse being applied with a controllable delay. Figure 5 shows the amplitude of the optical response to DFPS in the associated Ag hydrosol versus elapsed time from irradiating pulses ($\tau = 30$ ps) for different energy densities of the pulses. The inertial part of the amplitude is related to the variation of the sol absorption coefficient due to the photomodification of silver aggregates in the sol irradiated.

It was found that, when the pulse energy density exceeds the threshold, along with the fast response (observed at $t \le 50$ ps), the slow response, which correlates with the onset of FA photomodification [4], appears (at $t > 50$ ps). The time of its appearance depends on the pulse energy density and varies from 150 to 50 ps as the total energy density of pumping pulses (without considering Fresnel losses) increases from $W = 20.76$ to 24.75 and then to 30 mJ/cm^2. With $W = 8.25$ mJ/cm^2 or less, the inertial response was absent (Fig. 5). To be precise, we note that the energy of a laser pulse was distributed among two reference pulses, which formed the refractive index grating, and one probe pulse in roughly equal proportions.

The kinetics of the photomodification process, which starts when the absorbed energy per pulse is sufficient to sublimate a particle of radius 10 nm, is defined by the following expansion of sublimation products or by the modification of the FA structure until the depth of optical modulation (relative to the equilibrium value) due to these processes becomes below the sensitivity limit of a measuring method used.

Another view of the photomodification process (and, hence, its kinetics) should also be taken into consideration. In [22], the possibility of photoinduced modifications in the FA structure that arise in the light wave field under the action of resulting dipole–dipole interactions was analyzed. The authors of [22] believe that this process may both induce dichroism and favor the coalescence of resonant particles heated up (even if insignificantly) by the radiation absorbed.

The threshold (in energy) character of photomodification was analyzed in [20, 23, 24], where the presence of two thresholds was reported. The first one is related to the dip burned out on the short-wave side of the FA absorption peak. Physically, this threshold is due to a change in the polycrystalline structure of individual particles. The second threshold is associated with two dips appearing simultaneously. One arises in the long-wave wing near the frequency of a modifying radiation; the other, in the short-wave branch. In this case, as was noted above, the frequency of the low-frequency dip differs from the maximal absorption frequency of the unassociated hydrosol by a value that is twice as large as the shift of the high-frequency dip. Measurements [20, 23] showed that the low-frequency dip is deeper and its polarization is orthogonal to that of the high-frequency dip, which is fully consistent with the existing theoretical concepts.

In the case of short pulses, the energy threshold of photomodification depends largely not on the pulse intensity but on the energy density or, eventually, on the number of photons absorbed by a resonant particle. However, the radiation intensity, which specifies the strength of the optical wave field, should also be taken into account, since this parameter is responsible for optical damage to the crystal lattice of a particle during energy exchange with free electrons generated by the electromagnetic wave. In the case of long pulses (in particular, free-running pulses with $\tau = 0.1$ ms and $W = 0.1$ J/cm^2) or continuous radiation, photomodification ceases [1]. An increase in the pulse duration is also accompanied by more intense heat exchange.

The duration of irradiating pulses also has an effect on the spectral characteristics of photomodification. In going from nanosecond to picosecond pulses, the area of absorbed energy localization shrinks and the spectral dips become narrower. With nanosecond pulses, the collapse (reconstruction) of resonant domains in FAs also covers nearby particles whose resonance frequencies may differ from those of the resonant domains. This, in particular, explains why the spectral width of the dips somewhat increases, as demonstrated in Figs. 2 and 3.

One more feature of photomodification is that the dip becomes wider as the wavelength approaches that of absorption maximum (410–420 nm) [5, 20, 23, 24]. This is because the closer the wavelength to the initial plasmon maximum, the greater the number of resonant particles. On the contrary, the number of closely spaced pairs of particles with a large shift of resonances decreases away from the plasmon absorption maximum. This feature is substantiated by the fact that the threshold energy drops with distance to this maximum, since the radiation is absorbed by a lesser number of resonant particles. This is one more confirmation that the energy absorbed is distributed over the whole aggregate but localizes in resonant domains in the first case and in domains fairly distant from each other in the second case. Since, as was mentioned above, the collapse of resonant domains may partially cover adjacent FA areas, such a delocalization may involve a great amount of nonresonant particles in the process.

In [5, 23], structural changes in FAs induced by pulsed laser irradiation were examined. It appeared that FAs irradiated contain large spherical and ellipsoidal forms instead of initial particles. The fraction of the new forms depended on the incident radiation wavelength. For $\lambda = 450$, 540, and 1079 nm, this fraction equaled 0.7, 0.3, and 0.1, respectively.

The results of numerical simulation of photomodification in terms of the OPFC theory were reported in [20, 23, 24]. It was shown that a dip similar to that observed in our experiments appears in the FA absorption spectrum after the removal of resonant particles from FAs.

In spite of major achievements in studying the photomodification of noble-metal-based FAs, a physical model of this phenomenon has not yet been elaborated and a mechanism that can explain the entire body of experimental data and include the effect of attendant processes is still lacking. Research in this field continues.

CONCLUSIONS

(1) Photomodification of colloid fractals aggregates of metallic sols experimentally confirms the fact that the plasmon absorption band in their spectra broadens asymmetrically, since the width of the dip burned out is close to that of the symmetric absorption band for isolated particles.

(2) The polarization effect also unambiguously indicates the asymmetric broadening of the FA spectrum and, hence, the locality of photomodification. Furthermore, this effect follows from the local anisotropy of the fractals and also from the interaction of radiation with pairs of particles configured collinearly with and orthogonally to the field.

(3) The transition from nanosecond to picosecond irradiation pulses makes it possible to narrow the area of absorbed energy localization and, accordingly, produce narrower spectral dips.

(4) The energy threshold of photomodification largely depends not on the pulse intensity but on the energy density or, eventually, on the number of photons absorbed by each resonant particle in an FA. The energy threshold drops with increasing pulse wavelength.

REFERENCES

1. S. V. Karpov, A. K. Popov, S. G. Rautian, *et al.*, Pis'ma Zh. Éksp. Teor. Fiz. **48**, 528 (1988) [JETP Lett. **48**, 571 (1988)].
2. Yu. E. Danilova, N. N. Lepeshkin, S. G. Rautian, *et al.*, Physica A **241**, 231 (1997).
3. S. G. Rautian, V. P. Safonov, P. A. Chubakov, *et al.*, Pis'ma Zh. Éksp. Teor. Fiz. **47**, 200 (1988) [JETP Lett. **47**, 243 (1988)].
4. A. V. Butenko, Yu. E. Danilova, P. A. Chubakov, *et al.*, Z. Phys. **17**, 283 (1990).
5. V. P. Safonov, V. M. Shalaev, V. M. Markel, *et al.*, Phys. Rev. Lett. **80**, 1102 (1998).
6. V. A. Markel', L. S. Muratov, and M. I. Shtokman, Zh. Éksp. Teor. Fiz. **98**, 819 (1990) [Sov. Phys. JETP **71**, 455 (1990)].
7. V. M. Markel, V. M. Shalaev, E. V. Stechel, *et al.*, Phys. Rev. B **53**, 2425 (1996).
8. V. M. Shalaev and M. I. Shtokman, Zh. Éksp. Teor. Fiz. **92**, 509 (1987) [Sov. Phys. JETP **65**, 287 (1987)].
9. A. I. Plekhanov, G. L. Plotnikov, and V. P. Safonov, Opt. Spektrosk. **71**, 775 (1991) [Opt. Spectrosc. **71**, 451 (1991)].
10. S. V. Karpov, A. L. Bas'ko, A. K. Popov, *et al.*, Kolloidn. Zh. **62**, 773 (2000).
11. L. D. Landau and E. M. Lifshitz, *Course of Theoretical Physics*, Vol. 2: *The Classical Theory of Fields* (Nauka, Moscow, 1973; Pergamon, Oxford, 1975).
12. E. M. Aver'yanov, *Local Field Effects in Liquid Crystal Optics* (Nauka, Novosibirsk, 1999).
13. E. M. Aver'yanov, Pis'ma Zh. Éksp. Teor. Fiz. **66**, 805 (1997) [JETP Lett. **66**, 847 (1997)].
14. E. M. Aver'yanov, Opt. Zh. **65** (7), 5 (1998).
15. V. M. Shalaev, E. Y. Poliakov, and V. A. Markel, Phys. Rev. B **53**, 2437 (1996).
16. *Tables of Physical Quantities*, Ed. by I. K. Kikoin (Nauka, Moscow, 1970).
17. S. V. Karpov, A. L. Bas'ko, S. V. Koshelev, *et al.*, Kolloidn. Zh. **59**, 765 (1997).
18. S. V. Karpov, A. K. Popov, and V. V. Slabko, Pis'ma Zh. Éksp. Teor. Fiz. **66**, 97 (1997) [JETP Lett. **66**, 106 (1997)].
19. S. V. Karpov, V. V. Slabko, and G. A. Chiganova, Kolloidn. Zh. **64**, 474 (2002).
20. Yu. E. Danilova, S. G. Rautian, and V. P. Safonov, Izv. Ross. Akad. Nauk, Ser. Fiz. **60** (3), 56 (1996).
21. S. V. Karpov, A. K. Popov, and V. V. Slabko, Izv. Ross. Akad. Nauk, Ser. Fiz. **60** (6), 43 (1996).
22. V. P. Drachev, S. V. Perminov, S. G. Rautian, *et al.*, Zh. Éksp. Teor. Fiz. **121**, 1051 (2002) [JETP **94**, 901 (2002)].
23. Yu. E. Danilova and V. P. Safonov, in *Proceedings of the 3rd Conference on Fractal in the Natural and Applied Sciences (IFIP), Marseille, 1995* (Chapman and Hall, London, 1995), pp. 102–112.
24. Yu. E. Danilova, V. A. Markel', and V. P. Safonov, Opt. Atmos. Okeana **6**, 1436 (1993).

Translated by V. Isaakyan

Technical Physics, Vol. 48, No. 6, 2003, pp. 757–760. Translated from Zhurnal Tekhnicheskoĭ Fiziki, Vol. 73, No. 6, 2003, pp. 99–102.

OPTICS, QUANTUM ELECTRONICS

Formation of Two-Dimensional Periodic Nanostructures on Fused Quartz, Polyimide, and Polycrystalline Diamond by Pulsed Four-Wave Interference Laser Modification

Yu. K. Verevkin, N. G. Bronnikova, V. V. Korolikhin, Yu. Yu. Gushchina, V. N. Petryakov, D. O. Filatov, N. M. Bityurin, A. V. Kruglov, and V. V. Levichev

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

e-mail: verevkin@appl.sci-nnov.ru

Received October 23, 2002

Abstract—Atomic force microscopy is used to examine the topography of submicron periodic structures formed on the surfaces of synthetic polycrystalline diamond and polyimide films. The films are deposited on fused quartz substrates by four-wave interference modification using a pulsed 308-nm UV XeCl excimer laser. It is demonstrated that a two-dimensional periodic relief with a submicron period can be formed on the diamond surface directly by laser evaporation in the absence of a photoresist. Depending on the exposure, two mechanisms of polyimide film modification are observed. At exposures less than 100 mJ/cm^2, the relief is formed due to swelling at the positions of interference maxima. At exposures greater than 100 mJ/cm^2, holes are formed in the films. A periodic relief on the fused quartz surface is formed by using a UV photoresist exposed to pulsed interference laser radiation and subsequent Ar ion etching. *© 2003 MAIK "Nauka/Interperiodica".*

INTRODUCTION

The modification of materials under the action of several coherent UV laser beams interfering at the site of incidence enables one to pattern sub-100-nm periodic structures in solids. Such structures are of both fundamental and applied interest [1–4]. The coming generation of semiconductor structures will be based on sub-50-nm design rules. This goal can be achieved by nonlinear modification with a 308-nm XeCl laser. Of applied interest is creating phase masks in transparent insulators and Bragg fiber-optic and plane reflectors with various spectral characteristics.

Interference laser modification may take two paths. In the first case, the surface is covered by a photoresist with subsequent development and etching by means of lithography [1, 2, 5]. The alternative is direct (one-step) modification by laser irradiation [6–9]. The first variant employs cw and pulsed laser sources, whereas the second one needs high-power pulsed lasers. In this work, we apply both. The high sensitivity of photoresists makes it possible to form regular structures over an area of more than 10 cm^2 [2, 5]. This approach is promising for the creation of extended subwavelength hole arrays in metal films and for patterning phase masks on fused quartz in fiber-optics and planar technologies. However, the corresponding process is complex and involves many steps. The second variant makes it possible to apply rapid (up to 10^{11}°C/s) annealing to the samples in local nanodimensional areas, causing material evaporation, the recrystallization of films of various chemical composition [10–12], the formation of ferromagnetic structures with a high coercive force [4, 6], laser oxidation, etc.

In this work, we study the formation of nanodimensional periodic structures in synthetic polycrystalline diamond films and polyimide films deposited on fused quartz. The films are exposed to UV radiation of a 308-nm nanosecond XeCl laser. The morphology of the structures obtained is examined by atomic force microscopy (AFM). The films under study play an important role in microelectronics and fiber optics. In particular, synthetic polycrystalline diamonds offer a unique thermal conductivity, as well as a high transparency and radiation hardness in various spectral ranges. Polyimide films are used as high-contrast photoresist layers in the production of conducting polymers by laser irradiation, ion implantation, etc. An important applied problem is the patterning of the fused quartz surface. Here, we suggest that the surface of synthetic diamonds be modified by interference laser irradiation. It is demonstrated that the pattern on the polyimide film surface depends on the incident beam power. A periodic relief on the fused quartz surface is obtained with a UV photoresist exposed to interference pulsed laser radiation and Ar ion etching.

EXPERIMENTAL

The objects to be modified by pulsed interference laser irradiation were synthetic polycrystalline diamonds and 1-μm-thick polyimide films deposited on fused quartz substrates. A three-cascade XeCl excimer laser generated pulsed UV radiation with the following

1063-7842/03/4806-0757$24.00

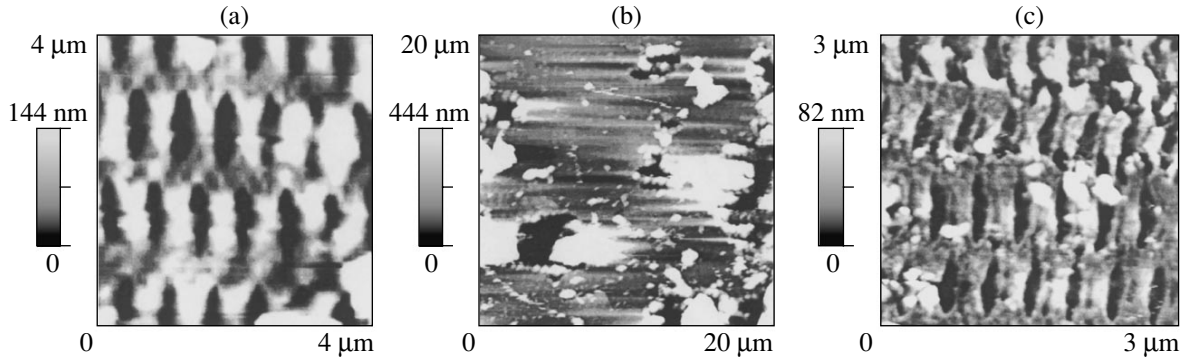

Fig. 1. AFM images of the modified surface of synthetic polycrystalline diamond (at an exposure of 10 mJ/cm^2): (a) the modified surface of a single-crystal grain, (b) severe surface damage in the area of opaque impurities, and (c) the boundary between single-crystal grains.

parameters: wavelength, 308 nm; coherence length, greater than 30 cm; pulse duration, 3–45 ns; and pulse energy, up to 0.1 J. The laser beam exhibited a divergence close to the diffraction limit.

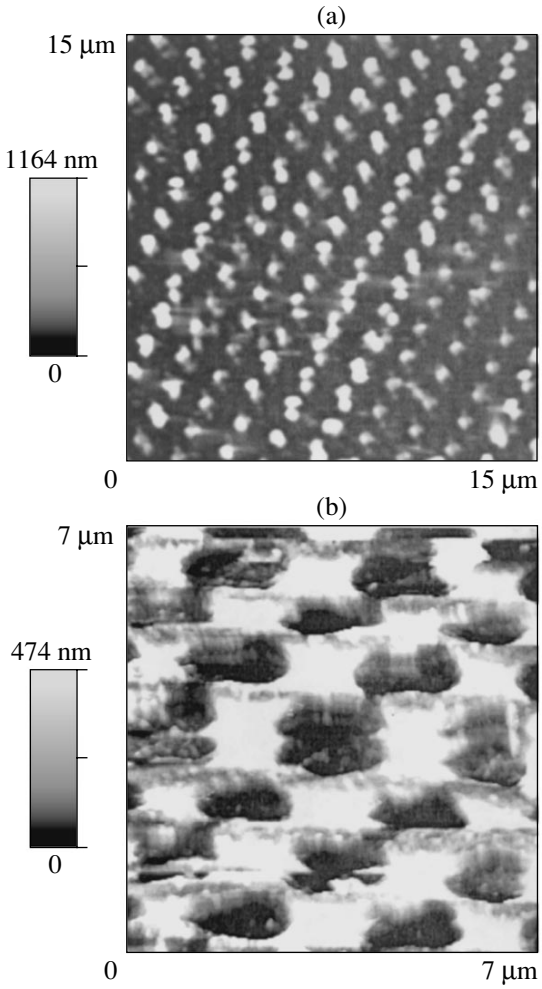

Fig. 2. AFM images of the modified polyimide film at exposures of (a) 60 and (b) 120 mJ/cm^2.

The setup consisted of a master oscillator, a thermally stabilized optical scheme for spectral selection, a spatial filter, a four-way amplifier with an SBS (stimulated Brillouin scattering) mirror, a final two-way amplifier, an optical system for splitting the beams and bringing them together on the object, a high-resolution spectrograph, a unit for taking energy and time characteristics, a station for evacuating the optical amplifiers and preparing gas mixtures, and power supply and synchronization units.

The application of the thermally stabilized optics for spectral selection, the spatial filters, and the SBS mirrors makes it possible to maintain a high quality of the beam during amplification. An SBS mirror was also used to shorten the laser pulses and to filter out the amplified spontaneous radiation. To split the emergent beam into four beams and to bring them together on the sample, we employed an optical scheme based on dielectric mirrors with reflectances of 50 and 99%. For the most part, the duration of the laser pulses was about 7 ns.

To form a periodic pattern on the fused quartz surface, we used an FP-4-04-V positive UV photoresist film with a thickness of 800–1000 nm. The photoresist surface was examined after various exposures and after wet development in a KOH solution. After the development, periodic structures on the fused quartz surface were patterned by liquid etching in HF and argon-ion etching (at a dose of $4.4–5.0 \times 10^{-4}$ cm^{-2} and an ion energy of 2.5–3.0 kV) for 12 h in a URM 3.279.029 etcher.

The surface pattern obtained by the interference modification was examined in a TopoMetrix® TMX-2100 Accurex™ atomic force microscope operating in the tapping mode with a silicon cantilever produced by the μmach® Co. (the stiffness is 0.35 N/m and the tip radius, 40 nm).

RESULTS AND DISCUSSION

Figure 1 shows the AFM image of the surface of the synthetic polycrystalline diamond exposed to the laser

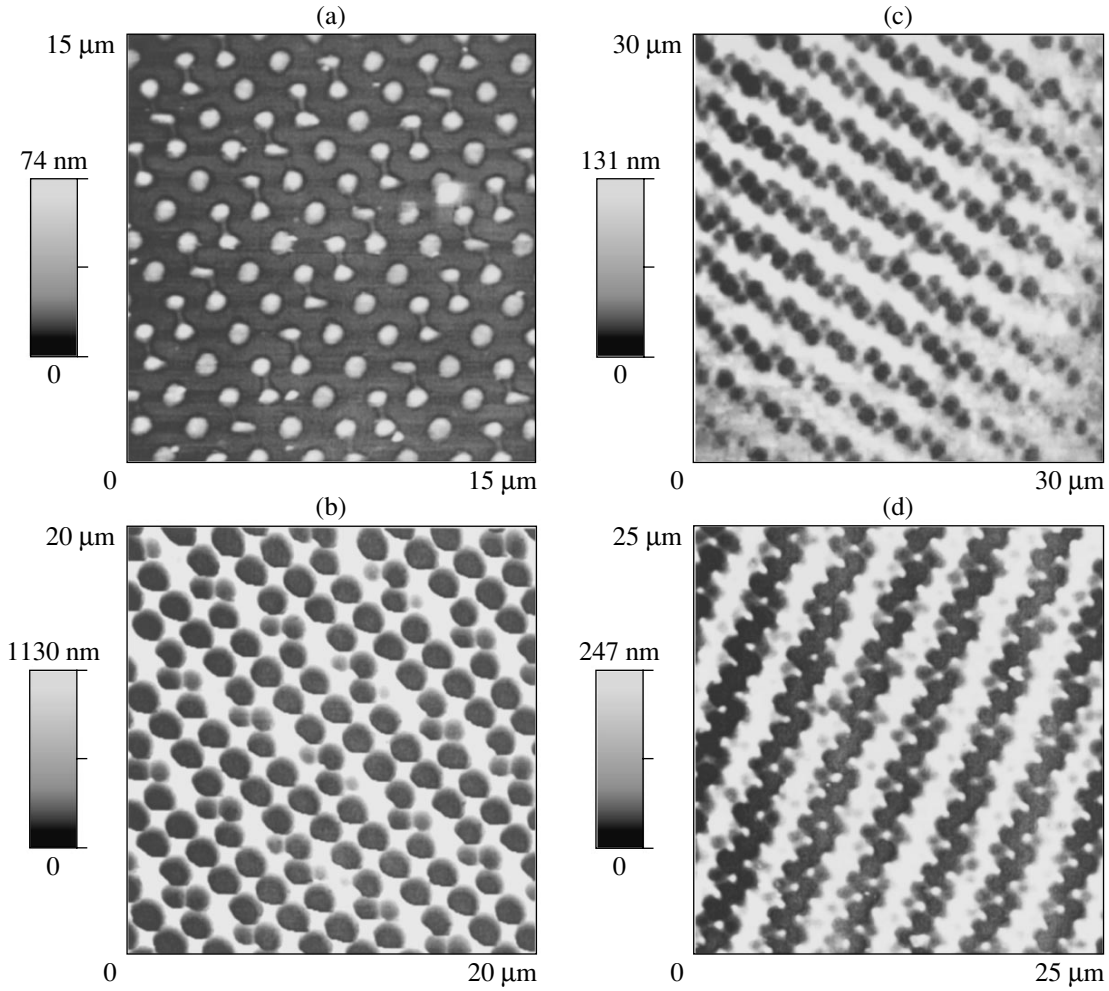

Fig. 3. AFM images of the UV photoresist on the fused quartz substrate: (a) after an exposure of about 20 mJ/cm^2, (b) after development, (c) fused quartz after etching in hydrofluoric acid, and (d) fused quartz after Ar ion etching.

radiation with an energy density of about 10 J/cm^2. At such an exposure, different areas exhibit different types of damage. There are areas where the damage pattern correlates with the interference pattern of the incident radiation (Fig. 1a). These are the surfaces of single-crystal grains. In places, severe damage independent of the incident radiation structure is observed (Fig. 1b). Diamond crystals are transparent to 308-nm radiation, and the exposure 10 J/cm^2 corresponds to the radiation-induced damage threshold for diamond single crystals. The severely damaged areas are highly absorbing intergranular regions.

Figure 1c demonstrates a modified polycrystalline diamond fragment lying between crystallites where the absorbing layer and, hence, severe damage are absent. We are planning to study the chemical composition in the modified zones of polycrystalline diamond. Thus, with coherent UV radiation, one can pattern polycrystalline diamond at a submicron level (including for marking purposes).

The laser modification of the polyimide films indicates that there exist two exposure ranges where the modification proceeds in a different way. At exposures of less than 100 mJ/cm^2, the film surface swells. Figure 2a shows the AFM image of such a fragment. The height of the swellings depends on the exposure and ranges from 300 to 1000 nm. At exposures greater than 100 mJ/cm^2, holes appear in the film. Figure 2b shows the AFM image of the modified film with holes. Such an object can be used as a mask in submicron lithography.

The patterning of the fused quartz surface is an important applied problem. The breakdown threshold for fused quartz exposed to four UV beams with a wavelength of 308 nm is about 15 J/cm^2. Such a high value does not allow the formation of a periodic coherent relief over an area of about 1 cm^2 by direct writing. Therefore, we used a UV photoresist. A 800- to 1000-nm-thick photoresist layer was applied on the quartz substrate by spinning. At an exposure of about 20 mJ/cm^2, we observed ~50-nm-high swellings on the photoresist

surface (see Fig. 3a). Wet development in the KOH solution produces 1000-nm-deep through holes (Fig. 3b). Note also the presence of shallow pits by the beams reflected from the back surface of the substrate. The fused quartz was etched through the developed photoresist. Figure 3c shows the AFM image of the quartz surface etched in hydrofluoric acid; Fig. 3d, after argon ion etching. In the first case, the maximum depth of the pattern is about 100 nm. In the second case, it is greater than 200 nm. It turned out that, when radiation with a wavelength of 308 nm is diffracted by the ion-etched relief, less than 50% of the energy is concentrated in the zeroth order. Similar measurements at wavelengths of 248 and 193 nm will be included in the subsequent study.

CONCLUSIONS

The results of this work demonstrate the feasibility of a periodic relief on a diamond surface by interference pulsed laser evaporation. In particular, one can use this method to mark jewelry diamonds. Two exposure-dependent mechanisms of polyimide film modification are discovered. At exposures of less than 100 mJ/cm^2, the relief is patterned through swelling in the positions of the interference maxima. At exposures greater than 100 mJ/cm^2, holes appear in the film. Such films with periodic arrays of holes obtained at high exposures may serve as effective masks for submicron lithography. A periodic relief on a fused quartz surface is obtained with a UV photoresist exposed to interference pulsed laser radiation and subsequent argon ion etching.

ACKNOWLEDGMENTS

This work was supported by the Russian Foundation for Basic Research (grant nos. 98-02-1636 and 02-02-17139) and the joint Russian–American program "Fundamental Research and Higher Education" financed by the Civilian Research and Development Foundation (CRDF) and the Ministry of Education of the Russian Federation (grant no. REC-001-1).

REFERENCES

1. A. Fernandez, H. T. Nguyen, J. A. Britten, *et al.*, J. Vac. Sci. Technol. B **15**, 729 (1997).
2. A. Fernandez, P. J. Bedrossian, S. L. Baker, *et al.*, IEEE Trans. Magn. **32**, 4472 (1996).
3. H. I. Smith, Physica E (Amsterdam) **11**, 104 (2001).
4. A. M. Alekseev, Yu. K. Verevkin, N. V. Vostokov, *et al.*, Pis'ma Zh. Éksp. Teor. Fiz. **73**, 214 (2001) [JETP Lett. **73**, 192 (2001)].
5. T. A. Savas, M. Farhoud, H. I. Smith, *et al.*, J. Appl. Phys. **85**, 6160 (1999).
6. M. Zheng, M. Yu, Y. Liu, *et al.*, Appl. Phys. Lett. **79**, 2606 (2001).
7. H. M. Phillips, D. L. Callaham, and R. Sauerbrey, Appl. Phys. A **54**, 158 (1992).
8. S. Pelissier, D. Biauc, M. P. Andrews, *et al.*, Appl. Opt. **38**, 6744 (1999).
9. M. Ezaki, H. Kumagai, K. Toyoda, *et al.*, Jpn. J. Appl. Phys. **32**, 1308 (1993).
10. Y. Chung and R. Shi-Woo, J. Mater. Sci.: Mater. Electron. **12**, 697 (2001).
11. M. Lee, S. Moou, M. Haatana, *et al.*, Appl. Phys. A **73**, 317 (2001).
12. M. D. Efremov, V. V. Bolotov, V. A. Volodin, *et al.*, Fiz. Tekh. Poluprovodn. (St. Petersburg) **36**, 109 (2002) [Semiconductors **36**, 102 (2002)].

Translated by A. Chikishev

Technical Physics, Vol. 48, No. 6, 2003, pp. 761–767. Translated from Zhurnal Tekhnicheskoĭ Fiziki, Vol. 73, No. 6, 2003, pp. 103–109.

RADIOPHYSICS

Parametric Interaction of Space-Charge Waves in Asymmetric Thin-Film *n*-GaAs Structures

A. A. Barybin* and A. I. Mikhailov**

** St. Petersburg State Electrotechnical University, ul. Prof. Popova 5, St. Petersburg, 197376 Russia*
e-mail: barybin@scientist.com
*** Chernyshevsky State University, ul. Dvadtsatiletiya VLKSM 112a, Saratov, 410071 Russia*

Received July 10, 2002

Abstract—Based on a theory developed previously, parametric interaction between space-charge waves in thin-film semiconductor structures with negative differential conductivity is analyzed. The analysis is carried out in the approximation that the drift flux of charge carriers has a rigid boundary and under the assumption that the frequency of low-frequency pumping equals the cutoff frequency f_c of waves being amplified (f_c roughly equals 30 GHz in our case). For asymmetric structures, a general multimode set of coupled equations is reduced to a pair of differential equations for the excitation amplitudes of the fundamental space-charge mode at the signal frequency ω_s and idler frequency $\omega_i = \omega_s - \omega_p$. The equations are solved numerically for *n*-GaAs-based structures, and the solution obtained is discussed. © *2003 MAIK "Nauka/Interperiodica".*

INTRODUCTION

Among thin-film semiconductor structures, those with negative differential conductivity (NDC), such as those based on *n*-GaAs and *n*-InP, are of special interest because of the remarkable behavior of space-charge waves (SCWs) when electrons are heated by a strong electric field. Structures of this type can be used in integrated millimeter-wave signal-processing devices [1–11]. For example, a traveling-wave amplifier made on a thin *n*-GaAs film can amplify and generate signals, provide a group delay, change the signal phase, switch channels, etc. [1–3]. More complex functions, such as controlled filtering with conversion and frequency spectrum synthesis, can be performed by using parametric regimes (of which low-frequency pumping is the most attractive) [4–11].

In [12–14], we developed the general theory of parametric interaction between SCWs and drifting carriers in thin-film structures with NDC. The theory includes carrier diffusion, frequency dispersion, and differential mobility asymmetry with real conditions set at the boundary of the semiconductor film. It reflects the multifrequency and multimode character of a wave process.

In this work, based on this general theory, we perform a theoretical study of the parametric interaction of the SCW fundamental mode under the assumption that the boundary of the carrier flux is rigid and the low-frequency pumping frequency f_p equals the cutoff frequency f_c of SCWs being amplified (f_c is roughly equal to 30 GHz in our case).

BASIC EQUATIONS AND SOLUTION

In the framework of the electrodynamic theory of waveguide excitation by external currents, we derived a set of equations [12–14] for the excitation amplitudes of parametrically coupled SCW modes. It is assumed that the modes propagate with electrons drifting in a thin-film structure with NDC and have frequencies $\omega_\nu = \omega_s + \nu\omega_p$ and $\omega_{\nu\pm1} = \omega_s + (\nu \pm 1)\omega_p$, where ω_s is the signal frequency and ω_p is the pumping frequency ($\nu = 0, \pm1, \pm2, \ldots$). The applicability of this theory will be demonstrated with a case of practical importance subject to several simplifications and refinements.

A highly asymmetric thin-film structure ($\varepsilon_H = \varepsilon$, $\varepsilon_b \ll \varepsilon_H$) is shown in two sections in Fig. 1. Here, $2a$ and $2b$ are the thickness and width of the semiconductor

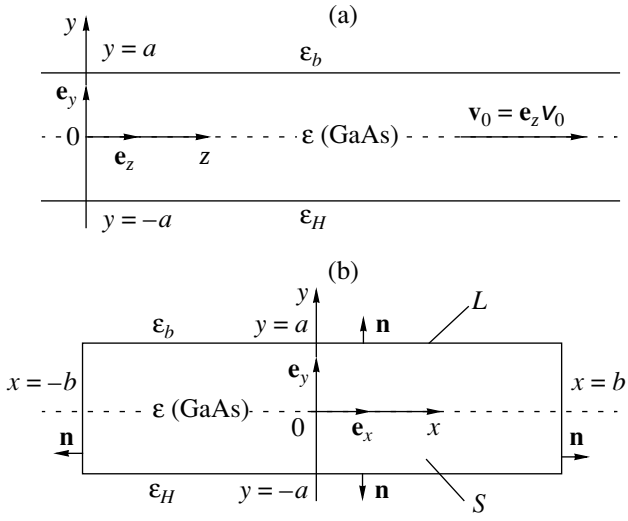

Fig. 1. Highly asymmetric thin-film structure. Shown are the sections through the plane (a) *y*–*z* and (b) *x*–*y*. *L*, boundary contour; *S*, cross section.

1063-7842/03/4806-0761$24.00 © 2003 MAIK "Nauka/Interperiodica"

film ($2a \ll 2b$), respectively; $\mathbf{e}_x$, $\mathbf{e}_y$, and $\mathbf{e}_z$ are the unit vectors in the Cartesian coordinates; S is the cross-sectional area of the semiconductor film ($S = 4ab$); L is the boundary contour of the cross section S; $\mathbf{n}$ is the unit outer normal to the contour L in the cross-sectional plane S; and $\mathbf{v}_0 = \mathbf{e}_z^* v_0$ is the vector of the static drift velocity of electrons within the NDC portion of the velocity–field characteristic of *n*-GaAs.

The effect of the electrophysical parameters of semiconductor films, as well as of differential mobility anisotropy and electron diffusion, on the structure and dispersion of SCW eigenmodes has been comprehensively discussed in monograph [15]. The influence of the frequency dispersion of differential electron mobility in *n*-GaAs and *n*-InP on the propagation of SCW eigenmodes in thin-film structures was investigated in [16, 17]. Thus, the structure and dispersion of the SCW eigenmodes in thin-film structures with NDC may be considered known.

Here, three points need to be made.

(1) The parametric interaction of the SCW fundamental mode is of greatest interest, since higher order modes have lower amplitudes and are hard to generate in real devices [15].

(2) For simplicity, we will consider the case $v = 0$; that is, $\omega_v = \omega_0 = \omega_s$ is the signal frequency, $\omega_{v-1} = \omega_s - \omega_p = \omega_i$ is the idler frequency, and $\omega_{v+1} = \omega_s + \omega_p$ is the image frequency.

(3) We assume that the frequency dispersion of the electron differential mobility and the frequencies ω_p and ω_s are taken such that the SCW amplitude at the frequency $\omega_s + \omega_p$ is negligible compared with the SCW amplitude at the frequencies ω_s and ω_i.

Hereafter, the subscripts s, p, and i are assigned to quantities related to signal, pumping, and idler frequency, respectively. Unlike [12–14], only subscripts are used in this work, since there is no need to number modes.

In view of the aforesaid, an infinite set of coupled equations that is obtained in the general theory [12–14] is reduced to a pair of ordinary first-order differential equations for the excitation amplitudes A_s and A_i of the SCW fundamental mode at the signal and idler frequencies:

$$
\begin{aligned}
N_s \frac{dA_s(z)}{dz} e^{-\gamma_s z} &= -\int_S \hat{\varphi}_s^* \Big\{ (\hat{\rho}_p \hat{v}_{iz} + \hat{\rho}_i \hat{v}_{pz}) \\
&\times \Big[\frac{dA_i(z)}{dz} - (\gamma_i + \gamma_p) A_i(z) \Big] e^{-(\gamma_i + \gamma_p) z} \Big\} dS \\
&- \oint_L \hat{\varphi}_s^* [\mathbf{n} \cdot (\hat{\rho}_p \hat{\mathbf{v}}_i + \hat{\rho}_i \hat{\mathbf{v}}_p) e^{-(\gamma_i + \gamma_p) z} A_i(z)] dl \\
&+ \int_S [(\hat{\rho}_p \hat{\mathbf{v}}_i + \hat{\rho}_i \hat{\mathbf{v}}_p) e^{-(\gamma_i + \gamma_p) z} A_i(z)] \cdot \nabla \hat{\varphi}_s^* dS,
\end{aligned}
\tag{1}
$$

$$
\begin{aligned}
N_i \frac{dA_i(z)}{dz} e^{-\gamma_i z} &= -\int_S \hat{\varphi}_i^* \Big\{ (\hat{\rho}_p^* \hat{v}_{sz} + \hat{\rho}_s \hat{v}_{pz}^*) \\
&\times \Big[\frac{dA_s(z)}{dz} - (\gamma_s + \gamma_p^*) A_s(z) \Big] e^{-(\gamma_s + \gamma_p^*) z} \Big\} dS \\
&- \oint_L \hat{\varphi}_i^* [\mathbf{n} \cdot (\hat{\rho}_p^* \hat{\mathbf{v}}_s + \hat{\rho}_s \hat{\mathbf{v}}_p^*) e^{-(\gamma_s + \gamma_p^*) z} A_s(z)] dl \\
&+ \int_S [(\hat{\rho}_p^* \hat{\mathbf{v}}_s + \hat{\rho}_s \hat{\mathbf{v}}_p^*) e^{-(\gamma_s + \gamma_p^*) z} A_s(z)] \cdot \nabla \hat{\varphi}_i^* dS,
\end{aligned}
\tag{2}
$$

where ρ, φ, and $\mathbf{v}$ are the electron charge density, the potential, and the variable component of the drift velocity, respectively; γ is the SCW propagation constant; N are normalizing factors that specify the power transferred by the SCW. The roof designates the membrane functions of physical quantities that define the dependence of the latter on transverse coordinates (on y in our case), the bar means averaging over the cross section S, and the asterisk denotes complex conjugates.

Eventually, the set of Eqs. (1) and (2) can be recast as

$$
\frac{dA_s(z)}{dz} = C_{11}(z) A_s(z) + C_{12}(z) A_i(z), \tag{3}
$$

$$
\frac{dA_i(z)}{dz} = C_{21}(z) A_s(z) + C_{22}(z) A_i(z), \tag{4}
$$

where

$$
\begin{aligned}
C_{11}(z) &= \hat{C}_{11} M(z) e^{-\alpha_p z}; \\
C_{12}(z) &= \hat{C}_{12} M(z) e^{(\alpha_s - \alpha_i) z};
\end{aligned}
\tag{5}
$$

$$
\begin{aligned}
C_{21}(z) &= \hat{C}_{21} M(z) e^{(\alpha_i - \alpha_s) z}; \\
C_{22}(z) &= \hat{C}_{22} M(z) e^{-\alpha_p z};
\end{aligned}
\tag{6}
$$

$$
\hat{C}_{11} = -\frac{\overline{\hat{G}_{sz}}}{\overline{N}_s \overline{N}_i} \Big[(\gamma_s + \gamma_p^*) \overline{\hat{G}_{iz}} + \overline{\hat{Q}_i} - \frac{1}{2a} \Delta \hat{G}_{iy} \Big]; \tag{7}
$$

$$
\hat{C}_{12} = \frac{1}{\overline{N}_s} \Big[(\gamma_i + \gamma_p) \overline{\hat{G}_{sz}} + \overline{\hat{Q}_s} - \frac{1}{2a} \Delta \hat{G}_{sy} \Big]; \tag{8}
$$

$$
\hat{C}_{21} = \frac{1}{\overline{N}_i} \Big[(\gamma_s + \gamma_p^*) \overline{\hat{G}_{iz}} + \overline{\hat{Q}_i} - \frac{1}{2a} \Delta \hat{G}_{iy} \Big]; \tag{9}
$$

$$\hat{C}_{22} = -\frac{\overline{\hat{G}_{iz}}}{\overline{N}_s\overline{N}_i}\left[(\gamma_i+\gamma_p)\overline{\hat{G}_{sz}}+\overline{\hat{Q}_s}-\frac{1}{2a}\Delta\hat{G}_{sy}\right]; \quad (10)$$

$$M(z) = \frac{e^{-\alpha_p z}}{1-M_p e^{-2\alpha_p z}}, \quad M_p = \frac{\hat{G}_{sz}\hat{G}_{iz}}{\overline{N}_s\overline{N}_i}, \quad \overline{N}_s = \frac{N_s}{S}, \quad \overline{N}_i = \frac{N_i}{S}; \quad (11)$$

and α is the amplitude constant, which is equal to the real part of the propagation constant for an associated mode ($\gamma = \alpha + i\omega/v_0$).

Other parameters are clarified in Appendix 1.

From Eq. (3), we express $A_i(z)$ as

$$A_i(z) = \frac{1}{C_{12}(z)}\frac{dA_s(z)}{dz} - \frac{C_{11}(z)}{C_{12}(z)}A_s(z) \quad (12)$$

and substitute (12) into (4). After rearrangements, this procedure yields a second-order differential equation for the excitation amplitude at the signal frequency:

$$\frac{d^2A_s(z)}{dz^2} - \{(\alpha_s-\alpha_p-\alpha_i) + M(z)e^{-\alpha_p z} \times [(\hat{C}_{11}+\hat{C}_{22}) - 2\alpha_p M_p]\}\frac{dA_s(z)}{dz} + M(z)\{\hat{C}_{11}e^{-\alpha_p z}(\alpha_s+\alpha_p-\alpha_i) + M(z)(\hat{C}_{11}\hat{C}_{22}e^{-2\alpha_p z} - \hat{C}_{12}\hat{C}_{21})\}A_s(z) = 0. \quad (13)$$

The coefficients in Eq. (13) are functions of z and include combinations of exponentials. It seems likely that this equation can be integrated analytically [18], at least for the limiting case $\alpha_p z \ll 1$, by expanding the exponential into series. Such a procedure was carried out in [19] to solve a simpler equation with the zero coefficient by the first derivative. However, even such a solution, if any, is a very awkward expression that is hard to treat physically. Therefore, we will make another simplifying assumption to clarify the result.

Let us take the pumping frequency ω_p such that $\alpha_p = 0$; that is, $\omega_p = \omega_c$, where $\omega_c = 2\pi f_c$ is the cutoff frequency of the SCWs being amplified. This means that, first, the pumping SCW propagates with a constant amplitude and, second, $\omega_s > \omega_c$ (the signal SCW decays in the absence of pumping) and $\omega_i < \omega_c$ (the SCW at the idler frequency builds up in the absence of pumping). With this assumption, Eq. (13) turns into the second-order differential equation with constant coefficients

$$\frac{d^2A_s(z)}{dz^2} - [(\alpha_s-\alpha_i) + M(\hat{C}_{11}+\hat{C}_{22})]\frac{dA_s(z)}{dz} + M[\hat{C}_{11}(\alpha_s-\alpha_i) + M(\hat{C}_{11}\hat{C}_{22}-\hat{C}_{12}\hat{C}_{21})]A_s(z) = 0, \quad (14)$$

where $M = 1/(1-M_p)$ for $\alpha_p = 0$ according to (11).

A solution to Eq. (14) is found in the form $A_s(z) \sim e^{\Gamma_s z}$. Substituting this expression into differential equation (14) yields the second-order characteristic equation for Γ_s

$$\Gamma_s^2 - [(\alpha_s-\alpha_i) + M(\hat{C}_{11}+\hat{C}_{22})]\Gamma_s + M[\hat{C}_{11}(\alpha_s-\alpha_i) + M(\hat{C}_{11}\hat{C}_{22}-\hat{C}_{12}\hat{C}_{21})] = 0. \quad (15)$$

Solutions to Eq. (15) are two roots, Γ_{s1} and Γ_{s2}:

$$\Gamma_{s1,2} = \frac{M(\hat{C}_{11}+\hat{C}_{22}) + (\alpha_s-\alpha_i)}{2} \pm \frac{M(\hat{C}_{11}-\hat{C}_{22})(\alpha_s-\alpha_i)}{2}\sqrt{\frac{1}{1-F_{si}}}, \quad (16)$$

$$F_{si} = \left\{1 + \frac{\left[(\hat{C}_{11}-\hat{C}_{22}) - \frac{(\alpha_s-\alpha_i)}{M}\right]^2}{4\hat{C}_{12}\hat{C}_{21}}\right\}^{-1}. \quad (17)$$

The coefficient F_{si} is very much like the coefficient introduced in [20, p. 28] for dispersionless systems and called the transfer coefficient. Expression (17) for F_{si} differs from the expression for the similar coefficient in [20] in that it has the term $(\alpha_s - \alpha_i)/M$, which appears in dissipative systems, such as a semiconductor with NDC. The coefficient F_{si} shows the strength of parametric coupling between modes at the signal and idler frequencies, which occurs owing to pumping. If the parametric coupling is weak, when $\sqrt{1/(1-F_{si})} \approx 1$, the solutions to the characteristic equation have the simplest form $\Gamma_{s1} = M\hat{C}_{11}$ and $\Gamma_{s2} = M\hat{C}_{22} + (\alpha_s - \alpha_i)$. Otherwise, when $\sqrt{1/(1-F_{si})} \neq 1$, one must use formula (16) to calculate the roots.

Thus,

$$A_s(z) = A_{s01}e^{\Gamma_{s1}z} + A_{s02}e^{\Gamma_{s2}z}, \quad (18)$$

where the subscripts 1 and 2 correspond to the plus and minus signs in (16).

In view of (12) and (18), we have

$$A_i(z) = \frac{1}{\hat{C}_{12}Me^{(\alpha_s-\alpha_i)z}} \times [(\Gamma_{s1}-\hat{C}_{11}M)A_{s01}e^{\Gamma_{s1}z} + (\Gamma_{s2}-\hat{C}_{11}M)A_{s02}e^{\Gamma_{s2}z}]. \quad (19)$$

Setting the boundary conditions at the point $z = 0$ in the form

$$A_s(z=0) = 1, \quad A_i(z=0) = 0, \quad (20)$$

we find

$$A_{s01} = \frac{\Gamma_{s2}-\hat{C}_{11}M}{\Gamma_{s2}-\Gamma_{s1}}, \quad A_{s02} = -\frac{\Gamma_{s1}-\hat{C}_{11}M}{\Gamma_{s2}-\Gamma_{s1}}. \quad (21)$$

Then, expressions (18) and (19) take the form

$$A_s(z) = \frac{1}{(\Gamma_{s2}-\Gamma_{s1})} \times [(\Gamma_{s2}-\hat{C}_{11}M)e^{\Gamma_{s1}z} - (\Gamma_{s1}-\hat{C}_{11}M)e^{\Gamma_{s2}z}], \quad (22)$$

$$A_i(z) = \frac{(\Gamma_{s1}-\hat{C}_{11}M)(\Gamma_{s2}-\hat{C}_{11}M)}{(\Gamma_{s2}-\Gamma_{s1})\hat{C}_{12}Me^{(\alpha_s-\alpha_i)z}}(e^{\Gamma_{s1}z}-e^{\Gamma_{s2}z}). \quad (23)$$

With formula (A.2.24) in Appendix 2, we represent $\hat{C}_{12}$ as

$$\hat{C}_{12} = \hat{C}'_{12}\frac{\hat{\varphi}_p(a)\hat{\varphi}_i(a)}{\hat{\varphi}_s(a)}, \quad (24)$$

where

$$\hat{C}'_{12} = \frac{\varepsilon\mu_e}{4\omega_s\varepsilon_H}\left[\left(\gamma_i + i\frac{\omega_p}{v_0}\right)K_1 - K_3 - K_5\right]. \quad (25)$$

Then, expression (23) can be written in the form

$$A_i(z) = \tilde{A}_i(z)\frac{\hat{\varphi}_s(a)}{\hat{\varphi}_p(a)\hat{\varphi}_i(a)}, \quad (26)$$

where

$$\tilde{A}_i(z) = \frac{(\Gamma_{s1}-\hat{C}_{11}M)(\Gamma_{s2}-\hat{C}_{11}M)}{(\Gamma_{s2}-\Gamma_{s1})\hat{C}'_{12}Me^{(\alpha_s-\alpha_i)z}}(e^{\Gamma_{s1}z}-e^{\Gamma_{s2}z}). \quad (27)$$

Now we take into account that the mean natural power $P_\omega(z)$ (without considering the mutual power of the modes) transferred by the fundamental mode with a frequency ω and a propagation constant γ_ω is given by [21]

$$P_\omega(z) = \frac{1}{4}N_\omega A_\omega(z)A^*_\omega(z)e^{-(\gamma_\omega+\gamma^*_\omega)z} \quad (28)$$

and that N_ω in a highly asymmetric thin-film structure is [21]

$$N_\omega = 2b\omega\varepsilon_H|\hat{\varphi}_\omega(a)|^2. \quad (29)$$

Then, in view of expressions (22) and (26), one can calculate the amplification coefficient of a thin-film structure of length z at a frequency ω_s,

$$K_a = 10\log\left[\frac{P_s(z)}{P_s(z=0)}\right] = 10\log[A_s(z)A^*_s(z)e^{-2\alpha_s z}], \quad (30)$$

and the transfer coefficient that converts the signal frequency ω_s at the input ($z = 0$) to the idler frequency ω_i at the output (conversion factor):

$$K_{con} = 10\log\left[\frac{P_i(z)}{P_s(z=0)}\right] = 10\log\left[\frac{\omega_i}{\omega_s}\frac{\tilde{A}_i(z)\tilde{A}^*_i(z)}{|\hat{\varphi}_p(a)|^2}e^{-2\alpha_i z}\right]. \quad (31)$$

With $\alpha_p = 0$ and $A_p = 1$, the pumping power is $P_p = N_p/4$, where $N_p = 2b\omega_p\varepsilon_H|\hat{\varphi}_p(a)|^2$.

The results of numerical calculation for an *n*-GaAs film by the formulas given above are shown in Figs. 2 and 3. In the calculations, the pumping frequency was constant: $f_p = \omega_p/2\pi = f_c \approx 31.275$ GHz. The parameters used were $\varepsilon = \varepsilon_H = 12.9\varepsilon_0$, $\varepsilon_b = \varepsilon_0$, $2a = 0.25$ μm, $2b =$

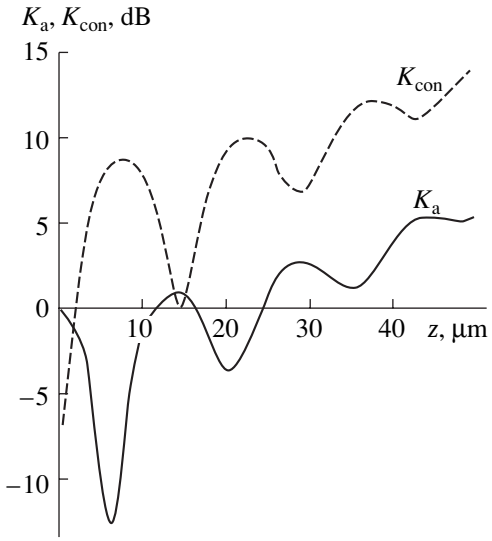

Fig. 2. K_a and K_{con} vs. z at the signal frequency.

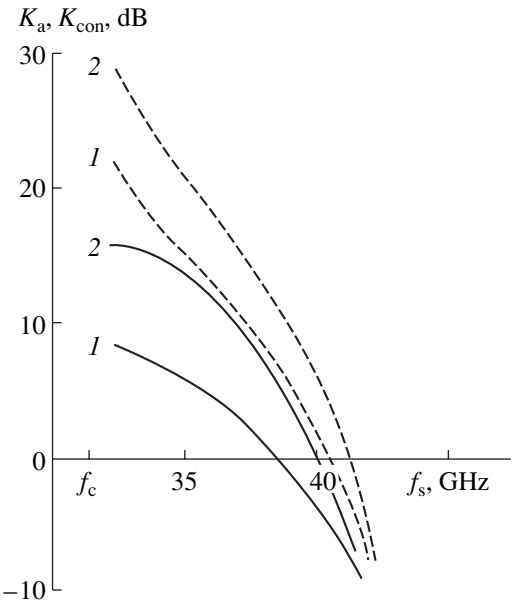

Fig. 3. K_a (solid lines) and K_{con} (dashed lines) vs. signal frequency f_s.

300 μm, $n_0 = 8 \times 10^{14}$ cm^{-3}, $D = 200$ cm^2/s, and $v_0 = 1.7 \times 10^{17}$ cm/s. This corresponds to the electric field strength $E_0 = 5.5$ kV/cm. The frequency dependence of the electron differential mobility was taken from [22], where it was evaluated by the Monte Carlo method for *n*-GaAs with the same value of E_0.

Figure 2 demonstrates the longitudinal (z) dependences of the amplification coefficient K_a and conversion factor K_{con} at the signal frequency $f_s = 35.275$ GHz and the pumping power $P_p = 0.1$ mW. As the signal and idler SCWs interact and propagate with drifting electrons, they periodically exchange power, with their amplitudes retaining the general tendency to grow. At the end of the interaction region (at $z = 50$ μm), both parameters become positive, thereby indicating effective parametric coupling between the SCWs at the signal and idler frequencies.

Figure 3 shows the dependences of K_a and K_{con} on the signal frequency f_s for $z = 50$ μm and $P_p = 0.1$ (curves *1*) and 0.2 mW (curves *2*). For the thin-film structure with the parameters listed above, low-frequency pumping ($f_p < f_s$) is seen to raise markedly (by 22–27%) the upper frequency limit of the SCWs being amplified and provide autodyne conversion at an input signal frequency of up to 41 GHz.

CONCLUSIONS

Thus, we demonstrated the feasibility of numerical calculation for specific cases in the framework of the general theory elaborated in our previous works [12–14]. The calculation performed in this article suggests that parametric interaction between SCWs in thin-film structures with NDC upon low-frequency pumping is a promising way of raising the upper operating frequency bound for functional devices of this type. Moreover, such a mechanism can provide effective frequency conversion in the millimeter range.

The approach used to work out the general theory of SCW parametric coupling can also be used to construct a theory of parametric interaction of waves in thin-film waveguide systems of other types.

APPENDIX 1

In the text,

$$\hat{\mathbf{G}}_s = \hat{\varphi}_s^*(\hat{\rho}_p\hat{\mathbf{v}}_i + \hat{\rho}_i\hat{\mathbf{v}}_p),\quad \hat{\mathbf{G}}_i = \hat{\varphi}_i^*(\hat{\rho}_p^*\hat{\mathbf{v}}_s + \hat{\rho}_s\hat{\mathbf{v}}_p^*) \quad \text{(A.1.1)}$$

are the vector membrane functions, which depend, in general, only on x and y (in our case, only on y);

$$\begin{aligned}\hat{Q}_s &= (\hat{\rho}_p\hat{\mathbf{v}}_i + \hat{\rho}_i\hat{\mathbf{v}}_p)\nabla\hat{\varphi}_s^*,\\ \hat{Q}_i &= (\hat{\rho}_p^*\hat{\mathbf{v}}_s + \hat{\rho}_s\hat{\mathbf{v}}_p^*)\nabla\hat{\varphi}_i^*\end{aligned} \quad \text{(A.1.2)}$$

are the scalar membrane functions, which depend, in general, only on x and y (in our case, only on y);

$$\overline{\hat{G}_{sz}} = \frac{1}{S}\int_S \hat{G}_{sz}dS,\quad \overline{\hat{G}_{iz}} = \frac{1}{S}\int_S \hat{G}_{iz}dS, \quad \text{(A.1.3)}$$

$$\overline{\hat{Q}_s} = \frac{1}{S}\int_S \hat{Q}_s dS,\quad \overline{\hat{Q}_i} = \frac{1}{S}\int_S \hat{Q}_i dS \quad \text{(A.1.4)}$$

are the values averaged over the cross section S; and

$$\Delta\hat{G}_{sy} = \hat{G}_{sy}(y = +a) - \hat{G}_{sy}(y = -a), \quad \text{(A.1.5)}$$

$$\Delta\hat{G}_{iy} = \hat{G}_{iy}(y = +a) - \hat{G}_{iy}(y = -a). \quad \text{(A.1.6)}$$

APPENDIX 2

Calculation of $\hat{C}_{11}$, $\hat{C}_{12}$, $\hat{C}_{21}$, $\hat{C}_{22}$, and M_p for a Highly Asymmetric Thin-Film Structure

The quantities $\hat{C}_{11}$, $\hat{C}_{12}$, $\hat{C}_{21}$, $\hat{C}_{22}$, and M_p are defined by formulas (7)–(11). We suppose that the longitudinal propagation constants γ_s, γ_p, and γ_i and the transverse wavenumbers ζ_s, ζ_p, and ζ_i for SCWs with frequencies ω_s, ω_p, and ω_i are known. With the subscripts omitted, one can write the following expressions for a highly asymmetric thin-film structure:

$$\varphi = \hat{\varphi}(a)\cos[\zeta(y-a)]e^{-\gamma z}, \quad \text{(A.2.1)}$$

$$\rho = \varepsilon(\zeta^2 - \gamma^2)\varphi, \quad \text{(A.2.2)}$$

$$v_y = \mu_e\zeta\hat{\varphi}(a)\sin[\zeta(y-a)]e^{-\gamma z}, \quad \text{(A.2.3)}$$

$$v_z = \mu_e\kappa\gamma\hat{\varphi}(a)\cos[\zeta(y-a)]e^{-\gamma z}. \quad \text{(A.2.4)}$$

In view of expressions (A.2.1)–(A.2.4), we arrive at

$$\hat{\varphi} = \hat{\varphi}(a)\cos[\zeta(y-a)], \quad \text{(A.2.5)}$$

$$\hat{\rho} = \varepsilon(\zeta^2 - \gamma^2)\hat{\varphi}(a)\cos[\zeta(y-a)], \quad \text{(A.2.6)}$$

$$\hat{v}_y = \mu_e\zeta\hat{\varphi}(a)\sin[\zeta(y-a)], \quad \text{(A.2.7)}$$

$$\hat{v}_z = \mu_e\kappa\gamma\hat{\varphi}(a)\cos[\zeta(y-a)]. \quad \text{(A.2.8)}$$

Taking into account (11), one can write [21]

$$\overline{N} = \frac{N}{S} = \frac{1}{2a}\omega\varepsilon_H|\hat{\varphi}(a)|^2. \quad \text{(A.2.9)}$$

Bearing in mind the above formulas and also formulas (A.1.1)–(A.1.6), we obtain after straightforward, though cumbersome, transformations

$$\overline{\hat{G}_{sz}} = \frac{1}{8a}\hat{\varphi}_s^*(a)\hat{\varphi}_p(a)\hat{\varphi}_i(a)\varepsilon\mu_e K_1, \quad \text{(A.2.10)}$$

$$\overline{\hat{G}_{iz}} = \frac{1}{8a}\hat{\varphi}_s(a)\hat{\varphi}_p^*(a)\hat{\varphi}_i^*(a)\varepsilon\mu_e K_2, \quad \text{(A.2.11)}$$

$$\Delta\hat{G}_{sy} = \frac{1}{4}\hat{\varphi}_s^*(a)\hat{\varphi}_p(a)\hat{\varphi}_i(a)\varepsilon\mu_e K_3, \quad \text{(A.2.12)}$$

$$\Delta\hat{G}_{iy} = \frac{1}{4}\hat{\varphi}_s(a)\hat{\varphi}_p^*(a)\hat{\varphi}_i^*(a)\varepsilon\mu_e K_4, \quad \text{(A.2.13)}$$

$$\overline{\hat{Q}_s} = -\frac{1}{8a}\hat{\varphi}_s^*(a)\hat{\varphi}_p(a)\hat{\varphi}_i(a)\varepsilon\mu_e K_5, \quad \text{(A.2.14)}$$

$$\overline{\hat{Q}_i} = -\frac{1}{8a}\hat{\varphi}_s(a)\hat{\varphi}_p^*(a)\hat{\varphi}_i^*(a)\varepsilon\mu_e K_6. \quad \text{(A.2.15)}$$

Here,

$$K_1 = \left[\left(\zeta_p^2 + \frac{\omega_p^2}{v_0^2}\right)\kappa_i\gamma_i + (\zeta_i^2 - \gamma_i^2)\kappa_p i\frac{\omega_p}{v_0}\right] \times (I_1 + I_2 + I_3 + I_4); \quad \text{(A.2.16)}$$

$$K_2 = \left[\left(\zeta_p^{*2} + \frac{\omega_p^2}{v_0^2}\right)\kappa_s\gamma_s - (\zeta_s^2 - \gamma_s^2)\kappa_p i\frac{\omega_p}{v_0}\right] \times (I_5 + I_6 + I_7 + I_8); \quad \text{(A.2.17)}$$

$$K_3 = \left(\zeta_p^2 + \frac{\omega_p^2}{v_0^2}\right)\zeta_i[I_2(\zeta_s^* - \zeta_p + \zeta_i) - I_1(\zeta_s^* - \zeta_p - \zeta_i) + I_4(\zeta_s^* + \zeta_p + \zeta_i) - I_3(\zeta_s^* + \zeta_p - \zeta_i)] \quad \text{(A.2.18)}$$
$$+ (\zeta_i^2 - \gamma_i^2)\zeta_p[I_3(\zeta_s^* + \zeta_p - \zeta_i) - I_1(\zeta_s^* - \zeta_p - \zeta_i) + I_4(\zeta_s^* + \zeta_p + \zeta_i) - I_2(\zeta_s^* - \zeta_p + \zeta_i)];$$

$$K_4 = \left(\zeta_p^{*2} + \frac{\omega_p^2}{v_0^2}\right)\zeta_s[I_6(\zeta_s - \zeta_p^* + \zeta_i^*) + I_5(\zeta_s + \zeta_p^* - \zeta_i^*) + I_8(\zeta_s + \zeta_p^* + \zeta_i^*) + I_7(\zeta_s - \zeta_p^* - \zeta_i^*)] \quad \text{(A.2.19)}$$
$$+ (\zeta_s^2 - \gamma_s^2)\zeta_p^*[I_5(\zeta_s + \zeta_p^* - \zeta_i^*) - I_7(\zeta_s - \zeta_p^* - \zeta_i^*) + I_8(\zeta_s + \zeta_p^* + \zeta_i^*) - I_6(\zeta_s - \zeta_p^* + \zeta_i^*)];$$

$$K_5 = \zeta_s^*\left[\left(\zeta_p^2 + \frac{\omega_p^2}{v_0^2}\right)\zeta_i(I_1 - I_2 + I_3 - I_4) + (\zeta_i^2 - \gamma_i^2)\zeta_p(I_1 + I_2 - I_3 - I_4)\right]; \quad \text{(A.2.20)}$$

$$K_6 = \zeta_i^*\left[\left(\zeta_p^{*2} + \frac{\omega_p^2}{v_0^2}\right)\zeta_s(I_5 - I_6 + I_7 - I_8) + (\zeta_s^2 - \gamma_s^2)\zeta_p^*(I_5 + I_6 - I_7 - I_8)\right],$$

$$I_1 = \frac{\sin[2(\zeta_s^* - \zeta_p - \zeta_i)a]}{(\zeta_s^* - \zeta_p - \zeta_i)},$$

$$I_2 = \frac{\sin[2(\zeta_s^* - \zeta_p + \zeta_i)a]}{(\zeta_s^* - \zeta_p + \zeta_i)},$$

$$I_3 = \frac{\sin[2(\zeta_s^* + \zeta_p - \zeta_i)a]}{(\zeta_s^* + \zeta_p - \zeta_i)},$$

$$I_4 = \frac{\sin[2(\zeta_s^* + \zeta_p + \zeta_i)a]}{(\zeta_s^* + \zeta_p + \zeta_i)}, \quad \text{(A.2.21)}$$

$$I_5 = \frac{\sin[2(\zeta_s + \zeta_p^* - \zeta_i^*)a]}{(\zeta_s + \zeta_p^* - \zeta_i^*)},$$

$$I_6 = \frac{\sin[2(\zeta_s - \zeta_p^* + \zeta_i^*)a]}{(\zeta_s - \zeta_p^* + \zeta_i^*)},$$

$$I_7 = \frac{\sin[2(\zeta_s - \zeta_p^* - \zeta_i^*)a]}{(\zeta_s - \zeta_p^* - \zeta_i^*)},$$

$$I_8 = \frac{\sin[2(\zeta_s + \zeta_p^* + \zeta_i^*)a]}{(\zeta_s + \zeta_p^* + \zeta_i^*)}.$$

As is known [15], at frequencies for which $\mu_d < 0$, trigonometric waves alone propagate in a thin-film structure; in other words, the transverse wavenumbers are real, so that $\zeta_s^* = \zeta_s$, $\zeta_p^* = \zeta_p$, and $\zeta_i^* = \zeta_i$. In this case, the formulas derived are somewhat simplified.

In view of the aforesaid, the desired parameters are given by

$$\hat{C}_{11} = -\frac{|\hat{\varphi}_p(a)|^2\varepsilon^2\mu_e^2}{16\omega_s\omega_i\varepsilon_H^2}K_1 \times \left[\left(\gamma_s - i\frac{\omega_p}{v_0}\right)K_2 - K_4 - K_6\right], \quad \text{(A.2.22)}$$

$$\hat{C}_{22} = -\frac{|\hat{\varphi}_p(a)|^2\varepsilon^2\mu_e^2}{16\omega_s\omega_i\varepsilon_H^2}K_2 \times \left[\left(\gamma_i + i\frac{\omega_p}{v_0}\right)K_1 - K_3 - K_5\right], \quad \text{(A.2.23)}$$

$$\hat{C}_{12} = \frac{\hat{\varphi}_s^*(a)\hat{\varphi}_p(a)\hat{\varphi}_i(a)\varepsilon\mu_e}{4\omega_s\varepsilon_H|\hat{\varphi}_s(a)|^2} \times \left[\left(\gamma_i + i\frac{\omega_p}{v_0}\right)K_1 - K_3 - K_5\right], \quad \text{(A.2.24)}$$

$$\hat{C}_{21} = \frac{\hat{\varphi}_s(a)\hat{\varphi}_p^*(a)\hat{\varphi}_i^*(a)\varepsilon\mu_e}{4\omega_i\varepsilon_H|\hat{\varphi}_i(a)|^2} \times \left[\left(\gamma_s + i\frac{\omega_p}{v_0}\right)K_2 - K_4 - K_6\right], \quad \text{(A.2.25)}$$

$$M_{\mathrm{p}} = \frac{\varepsilon^2 \mu_e^2 K_1 K_2}{16\omega_s \omega_i \varepsilon_H^2} |\hat{\varphi}_{\mathrm{p}}(a)|^2. \qquad \text{(A.2.26)}$$

REFERENCES

1. A. A. Barybin *et al.*, Mikroélektronika **8** (1), 3 (1979).
2. R. Din and R. Matareze, Tr. Inst. Inzh. Élektrotekh. Radioélektron. **60** (12), 23 (1972).
3. K. Kumabe and H. Kanbe, Int. J. Electron. **58** (4), 587 (1985).
4. A. I. Mikhaĭlov and S. A. Sergeev, Izv. Vyssh. Uchebn. Zaved., Radioélektron. **38** (10), 43 (1995).
5. G. L. Gurevich, M. A. Kitaev, A. L. Kogan, *et al.*, Radiotekh. Élektron. (Moscow) **33**, 1272 (1988).
6. M. A. Kitaev, A. L. Kogan, M. L. Pikovskaya, and E. I. Ryzhova, Élektron. Tekh., Ser. Élektron. SVCh, No. 2, 14 (1990).
7. A. I. Mikhaĭlov and S. A. Sergeev, Pis'ma Zh. Tekh. Fiz. **22** (24), 75 (1996) [Tech. Phys. Lett. **23**, 1 (1997)].
8. A. I. Mikhaĭlov, S. A. Sergeev, and Yu. M. Ignat'ev, RF Inventor's Certificate No. 9351, Byull. Izobret., No. 2 (1999).
9. A. I. Mikhaĭlov, S. A. Sergeev, and Yu. M. Ignat'ev, RF Patent No. 2,138,116, Byull. Izobret., No. 26 (1999).
10. A. I. Mikhaĭlov, S. A. Sergeev, and A. A. Goryachev, Izv. Vyssh. Uchebn. Zaved., Radioélektron. **43** (2), 16 (2000).
11. A. I. Mikhaĭlov, Pis'ma Zh. Tekh. Fiz. **26** (5), 80 (2000) [Tech. Phys. Lett. **26**, 217 (2000)].
12. A. A. Barybin and A. I. Mikhaĭlov, in *Materials of All-Russia Interuniversity Conference "Modern Problems of Electronics and Microwave Radiophysics," Saratov* (Gos. UNTs "Kolledzh," Saratov, 1997), pp. 109–110.
13. A. A. Barybin and A. I. Mikhaĭlov, Zh. Tekh. Fiz. **70** (2), 48 (2000) [Tech. Phys. **45** (2), 189 (2000)].
14. A. A. Barybin, A. I. Mikhaĭlov, and A. A. Kletsov, Élektrodin. Tekh. SVCh KVCh **7** (2), 88 (1999).
15. A. A. Barybin, *Waves in Thin-Film Hot-Electron Semiconductor Structures* (Nauka, Moscow, 1986).
16. A. I. Mikhaĭlov, Pis'ma Zh. Tekh. Fiz. **21** (21), 89 (1995) [Tech. Phys. Lett. **21**, 902 (1995)].
17. A. I. Mikhaĭlov and S. A. Sergeev, Pis'ma Zh. Tekh. Fiz. **25** (4), 85 (1999) [Tech. Phys. Lett. **25**, 162 (1999)].
18. E. Kamke, *Gewöhnliche Differentialgleichungen* (Academic, Leipzig, 1959; Nauka, Moscow, 1976).
19. G. L. Gurevich, A. L. Kogan, and G. M. Korobkov, Radiotekh. Élektron. (Moscow) **29** (2), 333 (1984).
20. W. H. Louisell, *Coupled Mode and Paramagnetic Electronics* (Wiley, New York, 1960; Inostrannaya Literatura, Moscow, 1963).
21. A. A. Barybin and M. G. Stepanova, Izv. Leningr. Élektrotekh. Inst. im. V. I. Ul'yanova **437**, 61 (1991).
22. H. D. Rees, Solid State Commun. **7** (2), 267 (1969).

Translated by V. Isaakyan

Technical Physics, Vol. 48, No. 6, 2003, pp. 768–775. Translated from Zhurnal Tekhnicheskoĭ Fiziki, Vol. 73, No. 6, 2003, pp. 110–117.

RADIOPHYSICS

Self-Oscillations in a Gyro-Backward-Wave Tube with Coupled Electrodynamic Structures

A. A. Koronovskii, D. I. Trubetskov, and A. E. Khramov

Colledge State Research and Educational Center, Chernyshevsky State University, ul. Dvadtsatiletiya VLKSM 112a, Saratov, 410026 Russia

e-mail: aeh@cas.ssu.runnet.ru

Received August 28, 2002; in final form, November 10, 2002

Abstract—Self-oscillations in a gyro-backward-wave tube with coupled guiding structures are studied. It is shown analytically and numerically that the critical current for the onset of self-modulation in such a structure can be significantly increased and a high power can be achieved in the single-frequency oscillation mode. Numerical simulations predict a significant increase in the gyrotron efficiency at the optimal coupling between the guiding structures. © *2003 MAIK "Nauka/Interperiodica".*

INTRODUCTION

In recent years, the generation and amplification of microwave radiation due to the interaction between helical electron beams and traveling waves have received much attention [1–6]. The generation of microwave radiation in a regular waveguide takes place when the synchronism condition between the electron beam and backward electromagnetic wave,

$$\omega \approx \hat{\omega}, \quad \hat{\omega} + \beta_0(\hat{\omega})v_\parallel - \omega_c = 0, \qquad (1)$$

is satisfied [7–12]. Here, $\hat{\omega}$ is the synchronism frequency, ω_c is the cyclotron frequency, $v_\parallel$ is the longitudinal electron velocity (i.e., the velocity parallel to the applied magnetic field),

$$\beta_0(\hat{\omega}) = -\frac{\omega_{cr}}{c}\sqrt{(\omega/\omega_{cr})^2 - 1}$$

is the propagation constant for a waveguide without an electron beam, and ω_{cr} is the critical frequency of the mode under consideration. This system, which is called the gyro-backward-wave tube (gyro-BWT), allows one to vary the radiation frequency by varying the longitudinal electron velocity or the external magnetic field [1, 13]. However, the efficiency of gyro-BWTs is relatively low (no higher than 20%) [14]. This is related to the specific features of the interaction between the beam electrons and the microwave field along the interaction space: the microwave radiation is directed oppositely to the beam, and the oscillating electrons bunch under the action of the strong output microwave field. Having bunched rapidly, the beam also rapidly becomes unbunched and does not have enough time to give up its energy to the field. Further, a secondary bunch is formed as the beam travels along the system. As a consequence, the current and field distributions in gyro-BWTs are characterized by the presence of many maxima due to the multiple rebunching of the beam. The increase in the beam current or the length of the gyro-BWT results in the self-modulation of the output signal, i.e., the generation of a broadband spectrum [5, 14, 15]. To increase the efficiency and the output power (which is determined by the beam current) in the single-frequency operating mode, the field distribution along the structure should be changed so that the field is small near the exit and has one maximum in the collector region, where the bunched electrons will give up a significant portion of their energy to the field. One of the ways of solving this problem is to employ the distributed extraction of the microwave power in devices with long-term interaction by using identical coupled guiding structures, along one of which a helical electron beam propagates [16–18].

In this paper, we study self-oscillations in a gyro-BWT with coupled guiding structures (coupled-structure gyro-BWT). In particular, we analyze the possibility of extending the parameter range corresponding to the single-frequency operating mode by using the distributed extraction of the microwave power.

THEORETICAL MODEL OF A COUPLED-STRUCTURE GYRO-BWT

Let us derive equations that describe the amplification and generation of a microwave signal in a coupled-structure gyro-BWT. A cylindrical weakly relativistic helical electron beam travels through one of the coupled structures. We consider the interaction between the beam and the TE mode of the waveguide under the assumption that the field is uniform over the beam cross section. We will neglect the interaction between the electrons and the high-frequency components of the magnetic field, which means that the longitudinal velocity of the beam is constant: $v_\parallel \approx$ const. We will

also assume that, throughout the operating frequency band, the helical beam interacts only with the backward wave. This means that, near the synchronism frequency $\hat{\omega}$, the dispersion relation can be linearized and represented in the form

$$\omega \approx \hat{\omega} + \left(\frac{d\omega}{d\beta_0}\right)_{\omega = \hat{\omega}} [\beta_0(\omega) - \beta_0(\hat{\omega})], \quad (2)$$

where $(d\omega/d\beta_0)_{\omega = \hat{\omega}} = v_g(\hat{\omega})$ is the group velocity of the wave at the frequency $\hat{\omega}$ (see condition (1)).

Then, the steady-state amplitudes $E_1(x)$ and $E_2(x)$ of the electric field in two weakly coupled guiding structures, along one of which (for the sake of definiteness, along the first one) the helical electron beam propagates, can be written as [19, 20]

$$\frac{dE_1}{dx} + j\beta_0 E_1 + \alpha_1 E_2 = i_1, \quad (3)$$

$$\frac{dE_2}{dx} + j\beta_0 E_2 + \alpha_2 E_1 = 0, \quad (4)$$

where i_1 is proportional to the amplitude of the high-frequency component of the current of the helical electron beam and $\alpha_{1,2}$ are the coupling parameters.

Since the guiding structures are identical, we have $|\alpha_1| = |\alpha_2| = \hat{\alpha}$. Then, from the power flow conservation law at $i_1 = 0$, we obtain $\alpha_1 = -\alpha_2^* = j\hat{\alpha}$.

Let us rewrite Eqs. (3) and (4) in terms of coupled waves:

$$\frac{dE_s}{dx} + j(\beta_0 + \hat{\alpha})E_s = i_1, \quad (5)$$

$$\frac{dE_f}{dx} + j(\beta_0 - \hat{\alpha})E_f = i_1, \quad (6)$$

where $E_s = E_1 + E_2$ is the amplitude of the slow normal wave (its propagation constant is $\beta_s = \beta_0 + \hat{\alpha}$, where β_0 is the propagation constant of the backward wave at the cold synchronism frequency) and $E_f = E_2 - E_1$ is the amplitude of the fast ($\beta_f = \beta_0 - \hat{\alpha}$) normal wave of the coupled structure.

Let us switch to the new variables $\hat{E}_f = E_f \times \exp[-j\beta_f x]$ and $\hat{E}_s = E_s \exp[-j\beta_s x]$ in Eqs. (5) and (6) and use the following normalized variables: the dimensionless coordinate $\xi = \beta_0(\hat{\omega})\varepsilon x$, the dimensionless time $\tau = \hat{\omega}\varepsilon\,(t - x/v_\parallel)(1 + v_\parallel/|v_g|)^{-1}$, $F_{f,s} = \hat{E}_{f,s}/2\varepsilon^2 V_0\beta_0$, $I_1 = (2i_1/\beta_0^2 KI_0)\exp[-j\beta_0 x]$, and the interaction parameter $\varepsilon = \sqrt{I_0 K/4V_0(1 + v_{\perp 0}^2/v_\parallel^2)}$ (which is physically equivalent to the Pierce amplification parameter C in O-type devices with long-term interaction [21]; $\varepsilon \ll 1$). Here, K is the coupling impedance, $v_{\perp 0}$ is the initial transverse velocity, and I_0 and V_0 are the constant components of the electron beam current and voltage. Then, we change from the time-independent equations to time-dependent ones [22] to obtain the final set of equations in terms of the dimensionless variables,

$$\frac{\partial F_s}{\partial \tau} - \frac{\partial F_s}{\partial \xi} = -I_1 \exp[j\alpha\xi], \quad (7)$$

$$\frac{\partial F_f}{\partial \tau} - \frac{\partial F_f}{\partial \xi} = -I_1 \exp[-j\alpha\xi], \quad (8)$$

which describe the field part of the problem of signal amplification and generation in coupled-structure gyro-BWTs. The coupling coefficient can be written in new variables as $\alpha = \hat{\alpha}/\beta_0\varepsilon$. The normal waves $F_{s,f}$ of the structure are related to the slowly varying complex amplitudes $F_{1,2}$ ($E_{1,2}(x, t) = \mathrm{Re}\{F_{1,2}(x, t)E_0 \exp[j(\hat{\omega}t + \beta_0 x)]\}$) of the microwave field of each of the guiding structures as

$$\begin{aligned} F_s &= (F_1 + F_2)\exp[j\alpha\xi], \\ F_f &= (F_2 - F_1)\exp[-j\alpha\xi]. \end{aligned} \quad (9)$$

The set of Eqs. (7) and (8) should be solved under the following initial and boundary conditions:

$$F_{1,2}(\xi, \tau = 0) = f_{1,2}^0(\xi), \quad F_{1,2}(\xi = A, \tau) = 0, \quad (10)$$

where A is the dimensionless length of the structure.

Let us now consider the electron part of the problem. The equation of motion of a weakly relativistic helical electron beam was derived in [10, 11]. In our notations, this equation can be written as

$$\begin{aligned} &\frac{d\beta}{d\xi} - j\mu(1 - |\beta|^2)\beta \\ &= \frac{1}{2}(F_s \exp[-j\alpha\xi] - F_f \exp[j\alpha\xi]), \end{aligned} \quad (11)$$

where $\beta = r\exp[j\Theta]$ is the complex-valued radius of the beam electrons that were initially uniformly distributed in phase with respect to the microwave field $B(\xi = 0) = \exp[j\Theta_0]$ ($\Theta_0 \in [0, 2\pi]$) and μ is the nonisochronism parameter, which is associated with the relativistic dependence of the electron mass on energy and describes the degree to which the system is inertial.

The amplitude I_1 of the high-frequency component of the beam current, which is determined by the phase bunching of the oscillating electrons in the helical beam propagating along the first guiding structure, is given by

$$I_1 = \frac{1}{2\pi}\int_0^{2\pi} \beta\, d\Theta_0. \quad (12)$$

Thus, the control parameters of the model are the dimensionless structure length A, the nonisochronism parameter μ, and the coupling coefficient α. If we study

a structure with a constant beam current I_0, then changing A at a constant μ is equivalent to changing the structure length l. At a constant length l and variable A, the parameter μA remains constant due to the variation in the current I_0.

LINEAR THEORY OF A COUPLED-STRUCTURE GYRO-BWT

Let us consider the conditions for the self-excitation of a coupled-structure gyro-BWT. To this end, we linearize the nonlinear equations that describe the dynamics of such a gyro-BWT. Under the assumption that $|F_{s,f}| \ll 1$, the following set of linear equations can easily be obtained from Eqs. (7)–(12):

$$\frac{d\bar{F}_s}{d\xi} - j(b+\mu)\bar{F}_s = -\bar{I}_1 \exp[j\alpha\xi], \quad (13)$$

$$\frac{d\bar{F}_f}{d\xi} - j(b+\mu)\bar{F}_f = -\bar{I}_1 \exp[-j\alpha\xi], \quad (14)$$

$$\frac{d^2\bar{I}_1}{d\xi^2} + \bar{I} = -j\frac{b}{2}(\bar{F}_s \exp[-j\alpha\xi] + \bar{F}_f \exp[j\alpha\xi]), \quad (15)$$

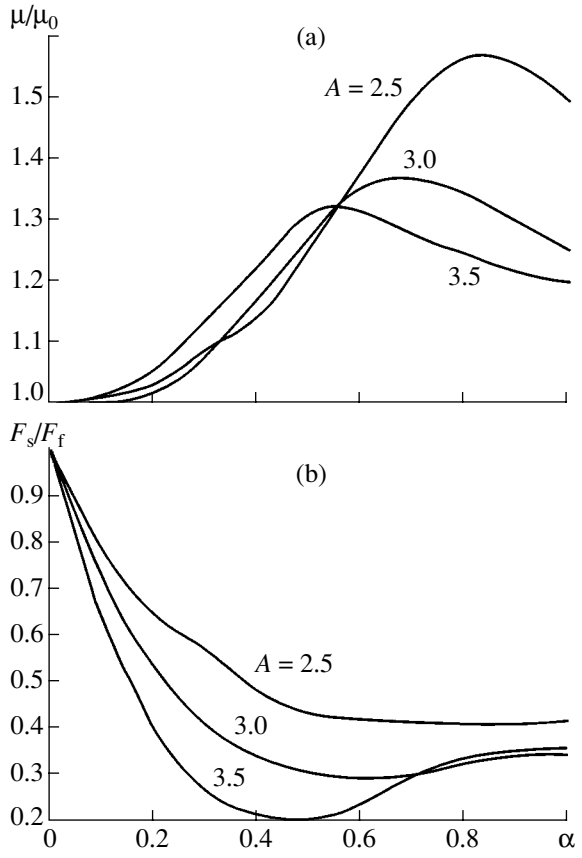

Fig. 1.

where $b = (\omega + \beta_0 v_\parallel - \omega_c)/k\varepsilon v_\parallel$ is the off-synchronism parameter, $\bar{I}_1 = I_1 \exp[j\mu\xi]$, $\bar{F} = F\exp[j\mu\xi]$, and $k = \omega/c$.

The numerical solution of the set of ordinary linear differential equations (13)–(15) shows that the threshold for the self-excitation of a gyro-BWT increases with increasing coupling between the guiding structures.

Figure 1a presents the threshold value of the nonisochronism parameter μ at which the tube is self-excited and a steady-state generation mode is established versus the coupling coefficient α for three lengths A. The nonisochronism parameter μ is normalized to the quantity μ_0, which refers to the excitation of the tube at the coupling coefficient $\alpha = 0$. It can be seen from the figure that, when the coupling between the guiding structures is nonzero, the threshold value of the nonisochronism parameter μ at which self-excitation occurs is larger than that in the case of $\alpha = 0$. It can also be seen that the curve $\mu(\alpha)$ has one maximum. As the structure length A increases, the relative increase in the threshold value of the nonisochronism parameter $\mu(\alpha)/\mu_0$ becomes smaller: at high coupling coefficients, the greater the structure length A, the lower the curve in Fig. 1a.

Let us find out what the reason is for such a dependence of the threshold value of the parameter μ on the coupling between the guiding structures. To this end, we will study the behavior of the normal modes excited in the coupled-structure gyro-BWT.

Figure 1b plots the ratio F_s/F_f between the amplitudes of the slow and fast normal waves at the output of the structure $\xi = 0$ near the threshold for self-excitation versus the coupling coefficient α for different values of the interaction length A. A comparison of Figs. 1a and 1b allows us to conclude that the maximum increase in the threshold value of the nonisochronism parameter is associated with the minimum values of the ratio F_s/F_f. This means that an increase in the coupling parameter leads to the predominant excitation of the fast wave in the coupled guiding structures; i.e., the beam interacts with the normal waves in a selective manner. The coupling impedance of each of the normal waves is twice as low as the coupling impedance K of a gyro-BWT with the same guiding structures. When there is no coupling ($\alpha = 0$), the expansion of the field in normal waves means that each of the normal waves has the same amplitude $F_s \equiv F_f$. Since $F_f > F_s$ at $\alpha > 0$, the self-excitation of the coupled-structure gyro-BWT requires a higher beam current than that for a gyro-BWT with $\alpha = 0$. This means that, at a fixed structure length A, the self-excitation of the coupled-structure gyro-BWT occurs at a larger value of the nonisochronism parameter μ of the oscillating electrons.

RESULTS FROM NUMERICAL SIMULATIONS OF A COUPLED-STRUCTURE GYRO-BWT

Now, we consider the results of simulations of a coupled-structure gyro-BWT in terms of nonlinear time-dependent equations (7)–(12). The equations were solved using the method of coarse particles, the two-sweep predictor–corrector scheme for integrating the equations of motion, and the Lax–Wendorff procedure for integrating the partial differential equations for the excitation of the normal waves. The numerical scheme used provides a second-order accuracy [23–25].

Figure 2a shows the critical value of the nonisochronism parameter μ_{self} at which self-modulation appears versus the coupling coefficient α for different interaction lengths A. It can be seen from the figure that μ_{self} significantly increases with increasing coupling between the guiding structures. The maximum increase is observed at $\alpha \approx 1.0$. The shorter the interaction length A, the greater the increase. The maximum relative increase slightly depends on A; for instance, at $A = 2.5$, 3.0, and 3.5, we have $\mu_{self}/\mu_{self}(\alpha = 0) \approx 2.52$, 2.77, and 3.76, respectively.

We note that the onset of self-modulation in the coupled-structure gyro-BWT differs from that in an uncoupled gyro-BWT. At $\alpha = 0$, self-modulation builds up in a hard manner (i.e., at $\mu = \mu_{self}(\alpha = 0)$, the self-modulation amplitude is finite), whereas at $\alpha > 0.2$–0.4, self-modulation builds up in a soft manner (i.e., the self-modulation amplitude is zero at $\mu = \mu_{self}(\alpha)$). Further, the self-modulation amplitude gradually increases with μ.

Let us study what the reason is for such a behavior of the self-modulation boundary in a coupled-structure gyro-BWT. It is well known [15, 26, 27] that self-modulation in microwave devices with long-term interaction is associated with an additional distributed feedback that develops in the structure. When the helical electron beam that was bunched in the strong microwave field arrives at the collector end $\xi = A$ of the structure with the speed $v_\parallel$, it has already been rebunched. The field excited by the bunched current propagates toward the entrance of the structure ($\xi = 0$) with the backward-wave group velocity v_g. Then, the beam bunched in the weak field excites the strong field, in which the helical electron beam is rebunched. This process repeats with the period $T \sim 2A(1/v_\parallel + 1/v_g)$, and the time $T/2$ can be regarded as the characteristic delay time in the additional feedback loop.

When the two guiding structures are coupled to each other, the additional feedback is suppressed because a fraction of the power of the backward wave (with the group velocity v_g), which provides the feedback, is transferred to the structure without a helical beam. As a result, the boundary of the self-modulation region shifts toward the higher values of the nonisochronism parameter μ, i.e., the stronger beam currents.

To better understand the process of transferring the microwave power between the guiding structures, Fig. 2b plots the ratio P_2/P_1 of the power fluxes in the guiding structures with and without a beam versus the coupling coefficient α. It can be seen that, as the coupling coefficient α increases from 0 to 0.7, the power transferred from the first structure to the second one rapidly grows: $\sqrt{P_2/P_1} \sim \exp[\gamma\alpha]$, where the coefficient γ lies within 1.2–1.8, depending on the structure length. At sufficiently large coupling coefficients ($\alpha > 0.7$), the saturation of the power flux P_2 is observed. However, $\mu_{self}(\alpha)$ keeps growing, which is related to the specific distributions of the fields $F_{1,2}$ in the coupled guiding structures at large values of the coupling coefficient. At large α values, a significant fraction of the backward wave power in the structure with a beam is transferred to the second structure as early as near the entrance $\xi = A$ to the tube, which also contributes to the suppression of the feedback that is responsible for self-modulation. This effect is illustrated in Fig. 3, which plots the distributions of fields $F_{1,2}$ in the coupled structures at $A = 3.5$ for different coupling coefficients α in the steady-state oscillation mode at the nonisochronism parameter close to the critical value, $\mu \lesssim \mu_{self}(\alpha)$. At $\alpha = 0.7$, which corresponds to the maximum P_2/P_1 ratio, the field distribution in the structure with a beam (Fig. 3, solid line) exhibits two maxima at the beginning and end of the interaction space. The field distribution $F_2(\xi)$ in the second guiding structure increases monotonically

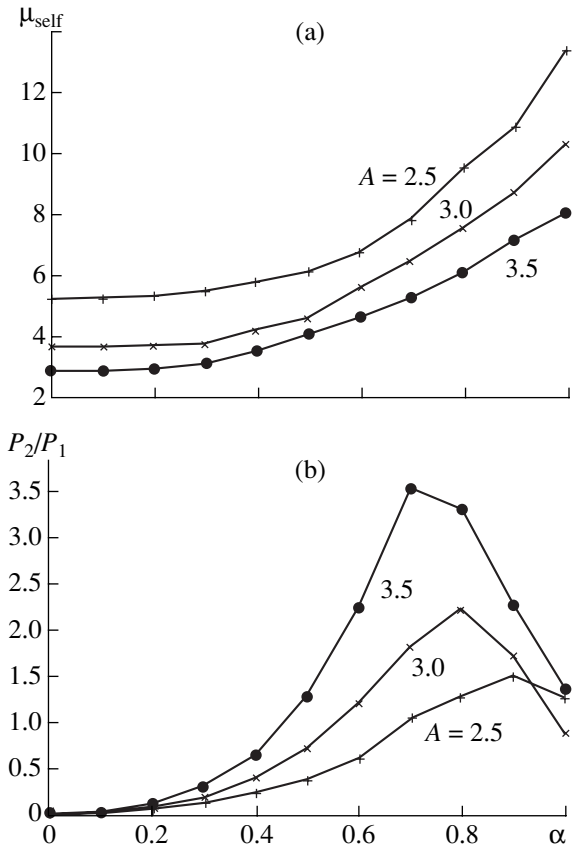

Fig. 2.

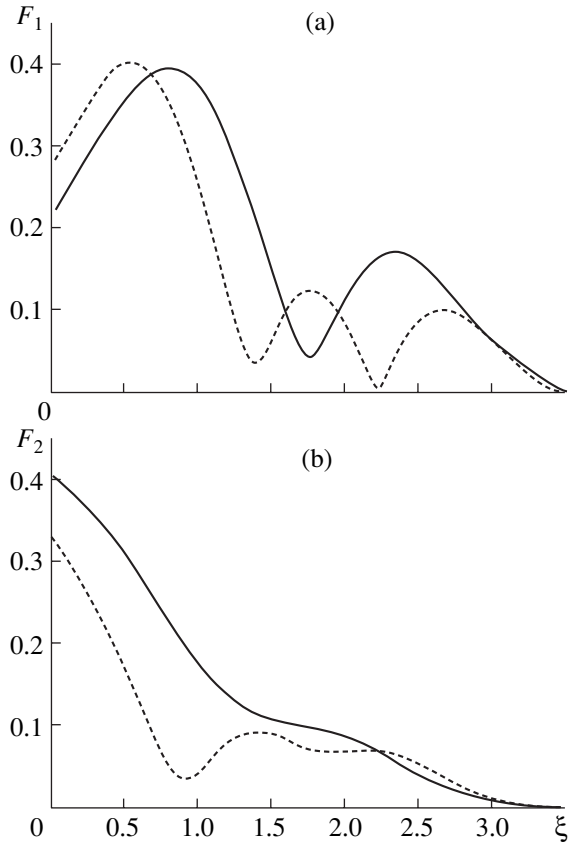

Fig. 3.

from the entrance ($\xi = A$) toward the exit ($\xi = 0$) of the coupled-structure gyro-BWT.

At a stronger coupling ($\alpha = 1.0$), the self-modulation threshold shifts toward the greater values of the nonisochronism parameter μ. As a result, the field distribution $F_1(\xi)$ exhibits three maxima (Fig. 3, dashed line), which is due to the multiple phase rebunching of the beam electrons over the interaction length A. Simultaneously, the amplitude of the field F_1 near the entrance to the tube ($\xi > 0.4A$) becomes significantly lower than that at $\alpha = 0.7$. At $\xi > 0.4A$, the field $F_2(\xi)$ in the second guiding structure shows a more complicated distribution with several maxima, but has the same amplitude as for $\alpha = 0.7$. Since $F_1(\alpha = 0.7\xi) > F_1(\alpha = 1.0\xi)$ at $\xi > 0.4A$, we find that, when the coupling is strong, a fraction of the power that is transferred to the second structure near the collector end is significantly greater than that at $\alpha = 0$. Therefore, although the fraction of the microwave power that is transferred to the second guiding structure decreases, the field distribution remains such that the additional feedback is weak and the self-modulation of the output signal does not occur.

The increase in the coupling between the guiding structures is accompanied by an increase in the efficiency of steady-state generation (by a factor of 1.2 to 1.4). The conversion efficiency of the transverse kinetic energy of the helical beam into the field energy (the transverse efficiency) is defined as [11, 23]

$$\eta_\perp = 1 - \frac{1}{2\pi}\int_0^{2\pi} |\beta|^2 d\Theta_0. \qquad (16)$$

The total efficiency of the gyro-BWT is given by

$$\eta = \eta_\perp \frac{\kappa^2}{1+\kappa^2},$$

where $\kappa = v_{\perp 0}/v_\parallel$.

At a constant κ, the gyro-BWT efficiency is determined by $\eta_\perp$ alone.

Figure 4 plots the transverse efficiency of the gyro-BWT versus the coupling coefficient α for different values of A. It follows from these plots that the maximum efficiency in the single-frequency oscillation mode is attained at $\alpha \approx 0.5$–0.6. When the structure length is small ($A \le 3.2$), the increase in the efficiency is insignificant and, at $\alpha \sim 1.0$, the efficiency is $\eta_\perp < \eta_\perp$ ($\alpha = 0$). For longer structures, the maximum of $\eta_\perp(\alpha)$ shifts to smaller α and $[\eta_\perp(\alpha)]_{max}$ significantly increases over that for the case of $\alpha = 0$. Such behavior is attributed to the fact that, at these coupling coefficients, the field distribution in the first guiding structure is optimal for the efficient energy extraction from the beam electrons.

To analyze this situation, we again refer to Fig. 3, which plots the distributions of the fields $F_{1,2}$ at the coupling coefficient $\alpha = 0.7$, which is close to optimal in terms of the energy conversion efficiency ($\eta_\perp = 0.22$, $\mu = 5.2$), and at $\alpha = 1.0$ ($\eta_\perp = 0.18$, $\mu = 8.0$), at which this efficiency is appreciably lower. The distributions F_1 are such that the field amplitudes near the exit side of the tube ($\xi = 0$) are approximately the same in both cases. However, at the optimal coupling ($\alpha = 0.7$), the field amplitude near the collector side ($\xi = A$) is significantly higher than that at $\alpha = 1.0$. This means that, in the former case, due to the higher amplitude F_1 in the region $\xi > 0.6A$, the electron bunch that was formed near the exit side of the tube ($\xi = 0$) under approximately the same conditions gives up a greater fraction of its energy than in the latter case, in which $F_1(\alpha = 1.0) < F_1(\alpha = 0.7)$ near the collector side of the tube. As a result, the efficiency of the gyro-BWT at a larger coupling coefficient α is lower than that observed under the optimal coupling.

Let us now consider the oscillation regimes observed in the coupled-structure gyro-BWT for different control parameters μ and A at a constant coupling coefficient $\alpha = 0.6$, which is optimal in terms of the generation efficiency.

Figure 5 shows the plane (A, μ) (dimensionless length vs. nonisochronism parameter), which is divided into domains corresponding to the typical modes of spatiotemporal oscillations in the coupled-structure

gyro-BWT. The following domains are indicated: the domain with no oscillations (*S*) and the domains corresponding to the single-frequency mode (*G*) and the periodic and chaotic self-modulation of the output signal (*T* and *C*, respectively). Figure 5 also shows the boundaries of the steady-state generation (dashed curve *1*) and the periodic and chaotic self-modulation (dashed-and-dotted curves *2* and *3*, respectively) in an uncoupled gyro-BWT ($\alpha = 0$).

An analysis of the map of the oscillation modes presented in Fig. 5 allows us to draw the following conclusions:

(i) The steady-state generation domain in the coupled-structure gyro-BWT is significantly greater than that at $\alpha = 0$. Simultaneously, as was discussed above, the energy conversion efficiency in the single-frequency mode increases.

(ii) The boundary of the self-modulation domain has a complicated irregular shape. This is related to the presence of several stable maxima in the electromagnetic-field distributions (Fig. 3), which is a characteristic feature of coupled-structure gyro-BWTs (as well as uncoupled gyro-BWTs [15]). As was noted in [15], this is due to the competition between two nonlinear mechanisms—the phase and inertial ones. The former mechanism, which is related to the phase nonlinearity of the beam electrons, disturbs the phase synchronism between the electromagnetic and electron waves. The latter mechanism, which is related to the inertial nature of the interaction between the electromagnetic and electron waves, causes the phase bunches to decay and the oscillating electrons to rebunch. The complicated shape of the self-modulation boundary is caused by the fact that the system can change to the periodic self-modulation regime from different steady-state generation regimes characterized by different spatial distributions. A similar effect was observed in a relativistic O-type BWT [28, 29]. We note that, at $A = 4.5$, below and above the steady-state generation domain, which has a complicated shape, there are domains of self-modulation with different frequencies and amplitudes. This also corroborates the statement that the complicated shape of the self-modulation domain is related to the competition between fundamentally different self-modulation modes of the coupled-structure gyro-BWT.

(iii) The boundary of the chaotic self-modulation domain in the coupled-structure gyro-BWT, as well as the boundary at which single-frequency generation becomes unstable, shifts toward the higher values of the parameter μ (cf. curve *3* for an uncoupled gyro-BWT and the boundary of the chaotic self-modulation domain *C* of the coupled-structure gyro-BWT in Fig. 5). At large *A*, the boundary of the chaotic self-modulation domain also shifts toward higher μ, but to a lesser extent.

The coupled-structure gyro-BWT changes to chaotic generation in a manner different from that of a gyro-BWT without distributed power extraction (an uncoupled gyro-BWT). As the gyro-BWT changes to a chaotic mode, it typically passes through periodic self-modulation, which gradually becomes more complicated (for instance, as *A* increases, the onset of chaos takes place after the self-modulation period doubles once or twice). The situation changes when the coupling (with $\alpha > 0.3$–0.4) is introduced into the structure. Numerical simulations show that, unlike the gyro-BWT or O-type BWT, self-modulation modes with irregular periods are not typical of the coupled-structure gyro-BWT. In the range of the control parameters under study, the following two main types of transition from the regular to chaotic regime were observed.

At large nonisochronism parameters μ and small structure lengths *A*, the transition to a chaotic regime proceeds through the intermittence regime. The amplitude $|F_{1,2}(\tau)|$ of the output signal behaves such that the system alternates between periodic oscillations (which are close to oscillations in the periodic self-modulation mode) and short chaotic bursts, which significantly differ in amplitude and phase from the regular oscillations. As the nonisochronism parameter μ increases over its critical value μ_{self}, at which single-frequency genera-

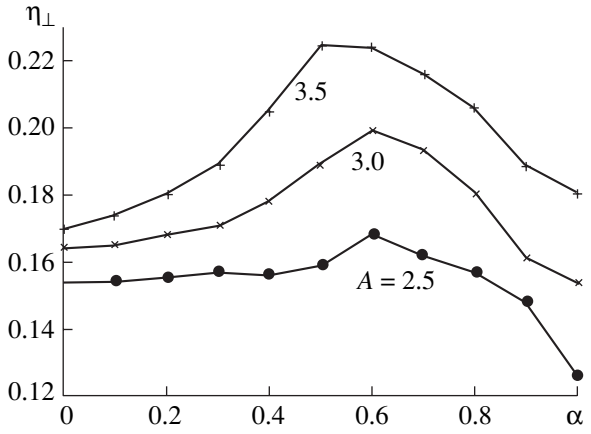

Fig. 4.

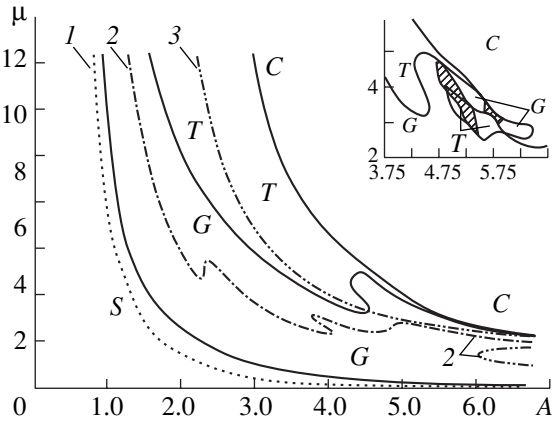

Fig. 5.

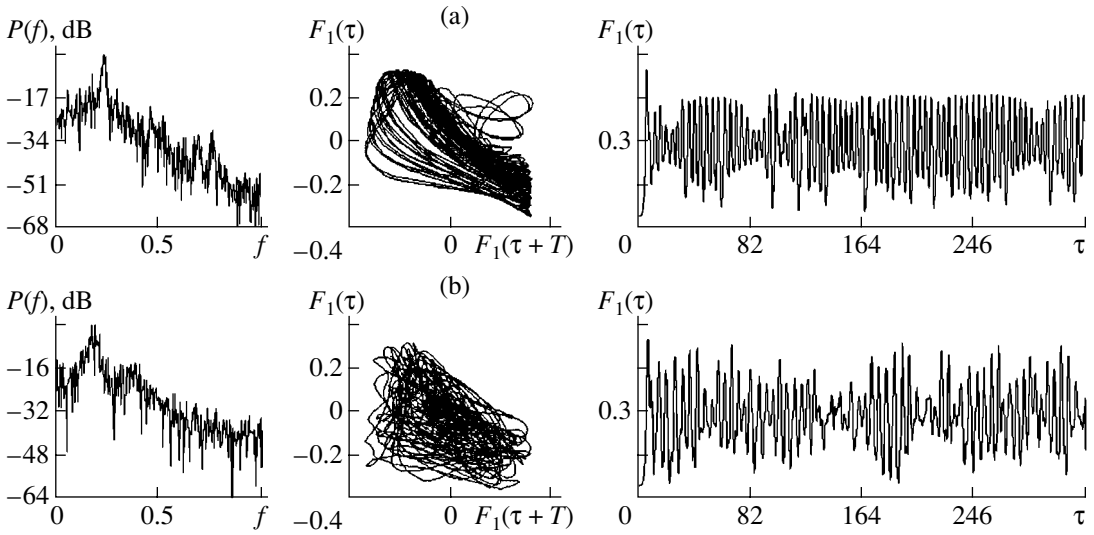

Fig. 6.

tion becomes unstable, the duration of the intervals with regular oscillations decreases until they disappear almost completely. The characteristics of this chaotic self-modulation mode are given in Fig. 6a, which plots the power spectra and phase portraits obtained by the delay embedding method (Takens' method [30])[1] and the realizations of the amplitude of the output field $|F_1(\tau)|$ in a guiding structure along which a helical beam propagates. In the realization, one can still distinguish the time behavior close to regular oscillations, which alternate with amplitude bursts in a random manner.

At large A and small μ, the transition to chaos proceeds through a gradual complication of the dynamics of the output signal as μ increases over its critical value μ_{self}. The characteristics of developed chaotic oscillations for this case are given in Fig. 6b. A comparison of Figs. 6a and 6b allows us to conclude that, in the latter case (at greater A), the chaotic oscillations are more complicated than at smaller A. In particular, at greater A, the noise pedestal in the power spectrum falls off slower with increasing frequency and the phase portrait is structureless and uniformly fills a certain region of phase space. In contrast, at smaller A (Fig. 6a), a chaotic attractor appears from a limiting set of the periodic slowly varying self-modulation of the output signal $|F|$ (limiting cycle).

At $A > 5$–6, the coupled-structure gyro-BWT changes to a chaotic generation mode directly from steady-state generation without passing through periodic self-modulation.

The inset in Fig. 5 shows the structure of the parameter space near the line at which the single-frequency generation becomes unstable at a large dimensionless length A. This fine structure is not shown on the main map in order not to overload the figure. The magnified fragment of the map reveals that, in this case, the mode pattern of the coupled-structure gyro-BWT is fairly complicated. We note several narrow domains corresponding to the periodic, quasi-periodic, and chaotic modes, which alternate with changing parameters. It may be expected that these regions, as well as the complicated shape of the boundary of the periodic self-modulation domain T, are associated with the competition between different regimes of spatiotemporal dynamics in the system.

CONCLUSION

We have shown that, at a certain value of the coupling coefficient, employing coupled guiding structures in a gyro-BWT allows one to shift the self-excitation boundary toward higher currents. In this case, the values of the control parameters A and μ corresponding to the onset of self-modulation increase, which provides a higher output power in the single-frequency mode. Simultaneously, the use of the guiding structures with an optimum coupling coefficient allows one to increase the efficiency with which the energy of the helical beam electrons in the single-frequency mode is converted into the microwave field energy. This is related to the establishment of the optimum current and field distributions over the interaction space in terms of the efficiency of energy extraction from the helical beam electrons by the backward wave. Chaotic oscillation modes have also been studied and three types of transition from regular oscillations to chaos have been shown to exist in coupled-structure gyro-BWTs.

[1] The function $x(t)$ is used to construct an n-dimensional vector $\{x(t), x(t + T), x(t + 2T), \ldots, x(t + nT)\}$, which, as was shown by Takens, can be treated as a vector in phase space. Here T is an arbitrary delay time [30].

ACKNOWLEDGMENTS

This work was supported by the Russian Foundation for Basic Research (project nos. 01-02-17392 and 02-02-16351) and by the Nonlinear Dynamics and Biophysics Research and Educational Center at Chernyshevsky State University (grant no. REC-006 of the US Civilian Research and Development Foundation for the Independent States of the Former Soviet Union (CRDF)).

REFERENCES

1. K. L. Felch, B. G. Danly, H. R. Jory, *et al.*, Proc. IEEE **87**, 752 (1999).
2. J. Rodgers, H. Guo, G. S. Nusinovich, *et al.*, IEEE Trans. Plasma Sci. **48**, 2434 (2001).
3. G. S. Nusinovich, W. Chen, and V. L. Granatstein, Phys. Plasmas **8**, 631 (2001).
4. G. S. Nusinovich, O. V. Sinitsyn, and A. Kesar, Phys. Plasmas **8**, 3427 (2001).
5. G. S. Nusinovich, A. N. Vlasov, and T. M. Antonsen, Phys. Rev. Lett. **87**, 218301 (2001).
6. D. I. Trubetskov and A. E. Khramov, Pis'ma Zh. Tekh. Fiz. **28** (18), 34 (2002) [Tech. Phys. Lett. **28**, 767 (2002)].
7. A. V. Gaponov, Izv. Vyssh. Uchebn. Zaved. Radiofiz. **2**, 443 (1959).
8. A. V. Gaponov, Izv. Vyssh. Uchebn. Zaved. Radiofiz. **2**, 450 (1959).
9. A. V. Gaponov, Izv. Vyssh. Uchebn. Zaved. Radiofiz. **4**, 547 (1961).
10. V. K. Yulpatov, Vopr. Radioélektron., Ser. Élektronika, No. 12, 15 (1965).
11. V. K. Yulpatov, Vopr. Radioélektron., Ser. Élektronika, No. 12, 24 (1965).
12. *Gyrotron* (Inst. Prikl. Fiz. Akad. Nauk SSSR, Gorki, 1981).
13. Z. G. Chen and H. Döring, Int. J. Infrared Millim. Waves **5**, 691 (1984).
14. A. Yu. Dmitriev, D. I. Trubetskov, and A. P. Chetverikov, Izv. Vyssh. Uchebn. Zaved. Radiofiz. **34**, 595 (1991).
15. D. I. Trubetskov and A. P. Chetverikov, Izv. Vyssh. Uchebn. Zaved. Prikl. Nelineinaya Din. **2** (5), 3 (1994).
16. D. I. Trubetskov, in *Proceedings of the 3rd Winter School on Microwave Electronics and Radiophysics, Saratov, 1978* (Saratov Gos. Univ., Saratov, 1978), Vol. 5, p. 89.
17. V. A. Isaev, V. L. Fisher, and A. P. Chetverikov, in *Proceedings of the 7th Winter School on Microwave Electronics and Radiophysics, Saratov, 1986* (Saratov Gos. Univ., Saratov, 1986), Vol. 2, p. 3.
18. V. A. Balakirev, A. O. Ostrovskiĭ, and Yu. V. Tkach, Zh. Tekh. Fiz. **61** (9), 94 (1991) [Sov. Phys. Tech. Phys. **36**, 1021 (1991)].
19. W. H. Louisell, *Coupled Mode and Paramagnetic Electronics* (Wiley, New York, 1960; Inostrannaya Literatura, Moscow, 1963).
20. D. I. Trubetskov and V. P. Shakhin, *Problems on Microwave Electronics*, Ed. by V. N. Shevchik (Saratov Gos. Univ., Saratov, 1973), Vol. 7, p. 44.
21. V. N. Shevchik and D. I. Trubetskov, *Analytical Methods in Microwave Electronics* (Sov. Radio, Moscow, 1970).
22. S. P. Kuznetsov and D. I. Trubetskov, *Backward-Wave Tube* (Saratov Gos. Univ., Saratov, 1975), p. 135.
23. A. Yu. Dmitriev, A. E. Konevets, L. A. Pishchik, *et al.*, in *Proceedings of the 7th Winter School on Microwave Electronics and Radiophysics, Saratov, 1986* (Saratov Gos. Univ., Saratov, 1986), Vol. 3, p. 61.
24. R. Hockney and J. Eastwood, *Computer Simulation Using Particles* (McGraw-Hill, New York, 1981; Mir, Moscow, 1987).
25. P. J. Roach, *Computational Fluid Dynamics* (Hermosa, Albuquerque, 1976; Mir, Moscow, 1980).
26. N. S. Ginzburg, S. P. Kuznetsov, and T. N. Fedoseeva, Izv. Vyssh. Uchebn. Zaved. Radiofiz. **21**, 1037 (1978).
27. B. P. Bezruchko, S. P. Kuznetsov, and D. I. Trubetskov, *Nonlinear Waves: Stochasticity and Turbulence* (Inst. Prikl. Fiz. Akad. Nauk SSSR, Gorki, 1980).
28. D. I. Trubetskov, V. G. Anfinogentov, N. M. Ryskin, *et al.*, Radiotekhnika (Moscow) **63** (4), 61 (1999).
29. N. M. Ryskin and V. N. Titov, Izv. Vyssh. Uchebn. Zaved. Radiofiz. **42**, 566 (1999).
30. F. Takens, in *Lectures Notes in Mathematics, Warwick, 1980*, Ed. by D. Rand and L.-S. Young (Springer-Verlag, New York, 1981), p. 366.

Translated by A. Khzmalyan

Technical Physics, Vol. 48, No. 6, 2003, pp. 776–779. Translated from Zhurnal Tekhnicheskoĭ Fiziki, Vol. 73, No. 6, 2003, pp. 118–122.

**SURFACES,
ELECTRON AND ION EMISSION**

Heat–Field Treatment of Tips Made of Tungsten–Hafnium Alloy

O. L. Golubev and V. N. Shrednik

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

e-mail: V.Shrednik@pop.ioffe.rssi.ru

Received December 10, 2002

Abstract—The combined effect of a high electric field and a high temperature on the morphology of W–Hf field-emission tips is studied. It is shown that, while the tip reconfiguration stages for this alloy are basically the same as for pure tungsten, they have a number of specific features due to the surface segregation of hafnium. The heat–field treatment of these emitters narrows the solid angle of emission significantly, improves the emissivity of the tips, and is accompanied by the high-temperature evaporation primarily of Hf atoms. © *2003 MAIK "Nauka/Interperiodica".*

INTRODUCTION

The simultaneous action of a high electric field (comparable to that causing field emission) and a temperature that may activate surface diffusion leads to a change in the sample shape. For example, a tip (tipped samples are considered in this paper), initially rounded, first becomes faceted (so-called reconstructed shapes) as the electric field F and temperature T grow. Then, microprotrusions appear on the edges and vertices of the faceted crystal. These microprotrusions favor the formation of larger hillocks, outgrowths, mainly on close-packed faces. The edges and vertices remain covered by microprotrusions. As a rule, these changes in the tip shape are observed when the field direction is such that the tip is positively charged. However, this kind of treatment can also be applied to tips charged negatively if the growth of the field and temperature does not result in catastrophic emission (eruption) in this case. Thermal–field action on pure refractory metals has been explored fairly comprehensively [1–3]. In alloys, such effects are poorly understood.

In this work, we study the combined effect of electric field and temperature on W–Hf alloy with (low) hafnium concentrations corresponding to the solid solution. In crystallographic and emission properties, W–Hf is similar to Mo–Hf and W–Zr alloys, which were subjected to heat–field treatment in several works. Field reconstruction in W–Zr was described in [4], where it was shown that this process differs markedly from the reconstruction in pure tungsten. In [5, 6], field emission from the microprotrusions on the surface of the solid solution of Hf in Mo was studied. Finally, in [7] the appreciable surface segregation of Hf was found by using the atomic probe method. However, the works listed did not touch upon problems of morphology and also the problem of producing point electron emitters generating narrow electron beams. As for W–Hf, the brief discussion of the spectrum of field emission from the microprotrusions formed on the surface of an alloyed emitter was given in [5], where the enrichment of the ion flux by hafnium was mentioned.

Our aim was to carefully study the temperature- and field-induced morphological modifications and emission properties of single-crystal W–Hf tips by field emission microscopy. The study of combined thermal–field effect on emitters made from alloys like W–Hf is of great interest in relation to the problem of producing efficient point sources of electrons and ions for various microelectronics and nanoelectronics applications [8]. A standard emitter, a tip several fractions of micrometer in size, is an essentially point electron (ion) source. However, conventional emitters may turn out to be unsuitable for microelectronics and nanoelectronics purposes, because they emit highly divergent beams (the solid angle α of emission may reach 1 to 2 sr).

An emitter that has a small effective size (down to one atom on its surface) and provides a small solid angle of emission will be called a point source. There exist at least three ways of reducing the emitting surface and, thus, emission confinement within a small solid angle [9]. The first is to decrease the work function φ within a small area on the emitter surface. To do this would require the presence of a higher emissivity adsorbate island at a given site on the surface. The second implies a local increase in the field F at a given site by creating a pyramidal protrusion. The third way integrates the first and second ones: a decrease in φ is combined with an increase in F [9, p. 221].

Adsorbates, such as Zr and Hf, decrease greatly the work function φ on the tungsten surface. Moreover, Zr and Hf layers on W confine electron emission from W{001} planes within a narrow solid angle, producing high-emissivity islands upon heating even without an electric field [10, 11]. The heat–field treatment of metallic (unalloyed) emitters also makes it possible to localize emission in a narrow solid angle by forming

1063-7842/03/4806-0776$24.00

sharp microprotrusions [12]. Once could therefore expect that the application of such treatment to W–Hf field emitters will combine the advantages of both methods of emission localization mentioned above. Note also that the study of morphology modifications during heat–field treatment is of fundamental interest in terms of crystal growth in electric field [2].

EXPERIMENTAL

The study was performed with the conventional FIM technique. The objects were alloyed tips prepared by the deposition of several tens of monoatomic Hf layers on the tungsten surface, followed by heating at $T \geq 1600$ K to provide the inward diffusion of Hf atoms. The emission pattern was typical of a homogeneous material, suggesting that we dealt with the solid solution of Hf (with a concentration of no more than 3 at.%) in W, as follows from the W–Hf phase diagram [13] and also from the measured values of φ, which was found to vary between 3.4 and 3.8 eV, depending on the treatment conditions. The reference book [14] gives φ = 3.51 eV for W–3 at.% Hf alloy in the range T = 1300–1900 K.

Upon heat–field treatment, one should discriminate between the initial field of treatment, F_{in}, which is always specified by the initial shape of the tip, and the final field F_{fin}, which is the surface field resulting after the tip shape has been changed. The values of F and φ were determined in the usual way, i.e., from the slope of the Fowler–Nordheim curve under the assumption that the total slope corresponds to the variation of F with φ (remaining constant) known and vice versa.

EXPERIMENTAL RESULTS

First, we will list our findings for the heat–field treatment of tips made from pure tungsten. The initial stage of tip reconfiguration is observed at relatively low T and F: $F_{in} \approx 0.4$ V/Å and $T \approx 1600$ K. An increase in F_{in} and T to 0.45–0.50 V/Å and 1800 K, respectively, leads to a substantial reconfiguration: {011}, {001}, and {211} close-packed faces expand. With T = 1800 K and F_{in} = 0.55 V/Å, small outgrowths appear around {001} faces. Finally, when T = 1850 K and F_{in} = 0.6 V/Å, the field-assisted crystal growth is the most pronounced: large outgrowths appear primarily on the close-packed faces of the tip, and microprotrusions form at the vertices and on the edges of the well-developed faces. At these stage, the field grows from initial F_{in} to final F_{fin} = 2.0–2.3 V/Å.

For W–Hf tips, the sequence of stages is basically the same but the stages have a number of specific features and take place at other T and F. After the condensation of 10–20 monoatomic Hf layers on the tungsten surface with the subsequent heating of the coating at T = 1500–1600 K without an electric field, intensely emitting 2D Hf spots or rings are observed on {001}W faces (Fig. 1a). Simultaneously, the work function decreases from φ = 4.5 eV for pure tungsten to 3.4–3.8 eV, depending on the amount of the adsorbate and heating temperature. The effective angle α of electron emission declines to 0.01–0.03 sr (versus the initial value ≈2 sr). The emissivity of field emitters is conveniently characterized, along with the quantities mentioned above, by the value of the applied voltage needed to achieve some fixed value of the electron current. We used the parameter U_{10}, which is the voltage at which the emission current i = 10 nA. For the W–Hf emitter shown in Fig. 1a, U_{10} = 5700 V, while for the initial (all-tungsten) tip, U_{10} = 6850 V. The treatment of such emitters modifies the structure of the emitting rings on {001} planes and diminishes α. At the same time, heating at T = 1600 K and F_{in} = 0.4 V/Å does not lead to the reconfiguration of the tip except that the initially round features (rings) around {001} faces gradually turn into squares. As F_{in} rises to 0.45–0.55 V/Å with the same T (1600 K), the rings turn into squares completely, reflecting the fourfold symmetry of {001}W faces (Fig. 1b). Simultaneously, the reconfiguration of {111} faces is evident, which has never been observed in pure tungsten, and U_{10} lowers to 4070 V. A further increase in F_{in} to 0.60 V/Å (with the same temperature 1600 K) causes the high-emissivity squares on {001} faces to break into much smaller emitting spots. These spots are arranged into chains propagating from {001} faces to the central (011) face. In Fig. 1c, corresponding to this state, these chains are not seen, because they emit much more weakly than the strongly emitting areas around {001} faces. In this case, U_{10} drops to 3450 V.

With F_{in} = 0.6 V/Å and T increased to 1700 K, several sharp microprotrusions grow on {001} faces. Four

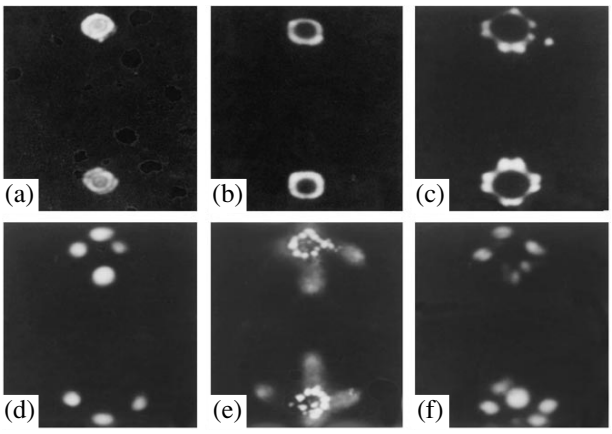

Fig. 1. Field emission microscopy images of the W–Hf emitter surface subjected to combined thermal–field action at the early stages of emitter reconfiguration. (a) After heating at T = 1600 K without an applied field, (b) after heating at T = 1600 K with F_{in} = 0.53 V/Å, (c) after heating at T = 1600 K with F_{in} = 0.6 V/Å, (d) after heating at T = 1700 K with F_{in} = 0.6 V/Å, (e) after repeatedly heating the emitter in the state (d) at 1600 K without the field, and (f) after heating at T = 1800 K with F_{in} = 0.6 V/Å.

of them around the upper (001) face and three around the lower (001) face are seen in Fig. 1d. In this case, electron emission is strongly localized (α = 0.002 sr) and U_{10} drops to 2300 V. Such a low angle of emission for a W–Hf emitter is hard to reach without using heat–field treatment. With the microprotrusions smoothed out with care by heating at 1600 K without an electric field, one can see (Fig. 1e) that the rest of the tip is left almost intact, i.e., that reconfiguring proceeds only on {001} faces, where the Hf concentration is likely to be the highest. The fact is that hafnium, like Zr, may produce a close-packed square lattice on these faces [15], because the atomic diameter of Hf, 3.18 Å [16], is close to the atomic spacing on {001}W planes, 3.16 Å. This circumstance favors the formation of Hf islands on these faces [10].

For the microprotrusions shown in Fig. 1d, F_{fin} usually equals 2.3–2.4 V/Å. With such values of F and T, high-temperature field sublimation takes place. The emission of Hf ions may be noticeable from the microprotrusions in this case. If so, we are dealing with a point source of Hf ions with α the same as for electrons. However, the emission current density will certainly be much lower, since the ion current itself is several orders of magnitude smaller than the electron current. Usually, the ion current varies between 10^{-11} and 10^{-9} A, while the electron current ranges from 10^{-7} to 10^{-5} A.

If T is increased to 1800 K at the same F_{in} = 0.6 V/Å, the amount of microprotrusions on the cube faces grows (Fig. 1f). The emission confinement in this case is the highest (α = 0.001–0.002 sr), and U_{10} takes the least value of 2040 V.

The stage of intense field-assisted growth, which is marked by the appearance of large outgrowths, starts at still higher temperatures. With T = 1850 K and the relatively low F_{in} = 0.35 V/Å, individual outgrowths, though not very large, appear on {001} faces and on the central (011) face. These outgrowths smoothed out by heating at 1600 K without an electric field are demonstrated in Fig. 2a. The surface of the outgrowths on {001} faces is seen to be covered by hafnium islands.

If T is raised to 1850 K and F_{in} to 0.7 V/Å, microprotrusions cover the tip almost entirely (Fig. 2b). After they have been smoothed out by heating, outgrowths of various size can be seen on many faces of the tip (Fig. 2c). The final value of the electric field is the highest in this case, F_{fin} = 2.45 V/Å, and so is the ion current, $i \sim 10^{-9}$ A. Under these conditions, the evaporation of microprotrusions goes in parallel with their growth. In field-desorption images, this dynamic process shows up as the collapse of rings [17]. It should be noted here that the value of F_{fin} is related only to the areas on the surface where the field strength is maximal: to the vertices of microprotrusions or to those sites on {001} faces where heating causes local field-enhancing steps [5]. In the remaining part of the surface, the field strength is certainly lower. In addition, these steps may exhibit a lower work function, down to $\varphi \approx 2.5$ eV, as in the case of Zr on the W surface [15]. With such values of φ, F_{fin} found from the slope of the Fowler–Nordheim curve may be somewhat underestimated.

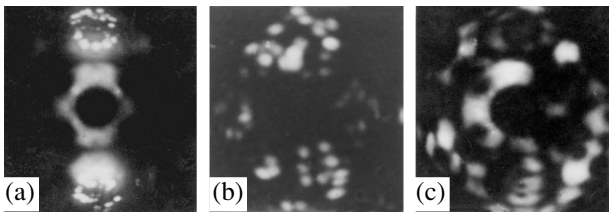

Fig. 2. The same as in Fig. 1 for the stage of intense field-assisted crystal growth. (a) After heating at T = 1850 K with F_{in} = 0.35 V/Å followed by heating at 1600 K without the field, (b) after heating at T = 1800 K with F_{in} = 0.7 V/Å, and (c) after repeatedly heating the emitter in the state (d) (Fig. 1) at T = 1600 K in the absence of the field.

DISCUSSION

First of all, it should be noted that the treatment of the alloy causes considerable segregation: the surface is enriched with Hf relative to its concentration in the volume, of which the value of φ = 3.53 eV (approaching that for pure Hf) [14] is an indication. In [5, 6], where the high-temperature evaporation of W–3 at.% Hf and Mo–7 at.% Hf alloys was studied, the same effect was found. Surface segregation of Hf explains why tip reconfiguration in the case of the alloy differs from that in the case of pure tungsten. Treatment-related microprotrusions grow on {001} faces, where the Hf concentration is the highest, at temperatures and fields F_{in} such that the remaining surface of the emitter remains nearly intact. This suggests that the material for the growth of microprotrusions is taken from the monoatomic surface layer, whereas underlying W and Hf atoms still migrate slowly. Microprotrusions grow when the pressure of ponderomotive forces $P_{pf} = F^2/8\pi$ exceeds the pressure of surface tension forces $P_\gamma = 2\gamma/r$ [2], where γ is the surface tension coefficient and r is the tip radius. Hafnium segregated at the surface decreases γ and P_γ, allowing for the formation of microprotrusions at relatively low temperatures and fields F_{in} when the tip reconfiguration is in its incipient stage. The values of T and F corresponding to various reconfiguration stages are listed in the table.

As follows from this work, the combined action of temperature and field, which results in the formation of microprotrusions, makes it possible to highly localize emission from the alloyed emitter: the angle α diminishes by one order of magnitude compared with that for the untreated alloy. Since φ of the alloy is significantly less than for W, the emissivity of microprotrusions on the W–Hf alloy will be much higher; that is, for the same i and current density j, microprotrusions on the W–Hf alloy will emit at lower F and U.

As was already noted, when subjected to heat–field treatment, the W–Hf emitter can generate not only electron but also ion currents. The latter, varying from sev-

Values of T and F at which W and W–Hf emitters undergo various stages of reconfiguration

Material	Reconstruction	Profound reconstruction	Microprotrusions	Macroprotrusions and outgrowths
W	T = 1600 K F = 0.4 V/Å	T = 1800 K F = 0.45 V/Å	T = 1800 K F = 0.55 V/Å	T = 1850 K F = 0.6 V/Å
W–Hf	Not observed	Not observed	T = 1700 K F = 0.6 V/Å	T = 1850 K F = 0.35 V/Å

eral fractions of a nanoampere to several nanoamperes, are certainly lower than the former at the same degree of emission confinement. The ion flux consists largely of variously charged Hf ions. As was shown [5, 6], mass spectra recorded during the high-temperature field evaporation of a Mo–7 at.% Hf emitter often contain only Hf^{3+} and Hf^{2+} ions. Consequently, using the heat–field treatment of alloyed emitters, one can generate currents of ions of various elements, including those for which the fabrication of emitting tips is extremely difficult (e.g., Hf and Zr) or altogether impossible.

Thus, the combined action of high temperatures and fields on emitters made of emission-active alloys, such as W–Hf, greatly improves their emissivity and makes it possible to create point electron and ion sources with unique properties. Also, the reconfiguration of alloyed emitters due to heat–field treatment has a number of intriguing features, which are of independent interest.

CONCLUSIONS

We studied the behavior of field emitters made of W–Hf alloy that were subjected to the combined thermal–field action over wide ranges of temperatures and fields. It was shown that, while the process of tip reconfiguration in this case has basically the same stages as for pure metals, it demonstrates a number of specific features and takes place at other values of T and F due to the surface segregation of hafnium.

The combined action of high field and temperature on these emitters increases the concentration of hafnium at the surface relative to its volume concentration and, hence, markedly improves the emission of electrons and ions.

Heat–field treatment substantially improves the confinement of electron emission. The angle α decreases from 0.01–0.02 sr, which is typical of the untreated emitter, to 0.001–0.002 sr after the treatment.

Heat–field treatment is accompanied by high-temperature field evaporation, which causes the emission of ions (mainly Hf ions). The values of α in this case are the same as for electron emission, and the emission current varies from 10^{-11} to 10^{-9} A.

Heat–field treatment of emission-active alloys like W–Hf is an efficient means of creating point emitters of electrons and ions.

ACKNOWLEDGMENTS

This work was supported by the Russian Foundation for Basic Research (project no. 01-02-17803) and the program of the Ministry of Science and Technology of the Russian Federation (contract no. 40.012.1.1.1152).

REFERENCES

1. V. N. Shrednik, *Growth of Crystals* (Nauka, Moscow, 1980), Vol. 13, pp. 68–79.
2. Yu. A. Vlasov, O. L. Golubev, and V. N. Shrednik, *Growth of Crystals* (Nauka, Moscow, 1991), Vol. 19, pp. 5–21.
3. V. G. Pavlov, A. A. Rabinovich, and V. N. Shrednik, Fiz. Tverd. Tela (Leningrad) **17**, 2045 (1975) [Sov. Phys. Solid State **17**, 1335 (1975)].
4. V. N. Shrednik, Fiz. Tverd. Tela (Leningrad) **1**, 1134 (1959) [Sov. Phys. Solid State **1**, 1037 (1959)].
5. E. L. Kontorovich, M. V. Loginov, and V. N. Shrednik, J. Vac. Sci. Technol. B **15**, 495 (1997).
6. M. V. Loginov and V. N. Shrednik, Zh. Tekh. Fiz. **68**, 69 (1998) [Tech. Phys. **43**, 327 (1998)].
7. M. K. Miller and G. D. W. Smith, *Atom Probe Microanalysis: Principles and Applications to Material Problems* (Materials Research Society, Pittsburgh, 1989; Mir, Moscow, 1993).
8. V. N. Shrednik, Poverkhnost, No. 2, 102 (1998).
9. V. N. Shrednik, *Nonincandescent Cathodes*, Ed. by M. I. Elinson (Sov. Radio, Moscow, 1974), Chap. 6, pp. 165–169.
10. V. N. Shrednik and G. A. Odishariya, Izv. Akad. Nauk SSSR, Ser. Fiz. **33**, 536 (1969).
11. L. W. Swanson and L. C. Crouser, J. Appl. Phys. **40**, 4741 (1969).
12. V. N. Shrednik, V. G. Pavlov, A. A. Rabinovich, *et al.*, Phys. Status Solidi A **23**, 373 (1974).
13. C. Bill, Trans. Metall. Soc. AIME **224**, 61 (1962).
14. V. S. Fomenko, *Emission Properties of Materials: A Handbook* (Naukova Dumka, Kiev, 1981).
15. V. N. Shrednik, Fiz. Tverd. Tela (Leningrad) **3**, 1750 (1961) [Sov. Phys. Solid State **3**, 1268 (1961)].
16. *Hafnium: Collection of Articles*, Ed. by L. N. Komissarova (Inostrannaya Literatura, Moscow, 1962).
17. V. G. Butenko, O. L. Golubev, V. N. Shrednik, *et al.*, Pis'ma Zh. Tekh. Fiz. **18** (8), 86 (1992) [Sov. Tech. Phys. Lett. **18**, 506 (1992)].

Translated by V. Isaakyan

Technical Physics, Vol. 48, No. 6, 2003, pp. 780–786. Translated from Zhurnal Tekhnicheskoĭ Fiziki, Vol. 73, No. 6, 2003, pp. 123–129.

EXPERIMENTAL INSTRUMENTS AND TECHNIQUES

The Action of Pulsed Electron Beams on Carbon Bisulfide Impurity in Air

Yu. N. Novoselov, A. I. Suslov, and D. L. Kuznetsov

Institute of Electrophysics, Ural Division, Russian Academy of Sciences, Komsomol'skaya ul. 34, Yekaterinburg, 620016 Russia

e-mail: nov@lep.uran.ru

Received November 21, 2002

Abstract—The results are presented of experimental and theoretical studies of the action of nano- and microsecond electron beams on an atmospheric-pressure nitrogen–oxygen mixture with a CS_2 impurity. It is demonstrated that there are two modes of impurity conversion that depend on the beam parameters and differ in both the mechanism involved and the final products. The effect of the external electric field on the impurity conversion is studied experimentally. A computer model of the processes is developed, and an analysis of the experimentally observed effects is performed. © *2003 MAIK "Nauka/Interperiodica".*

1. INTRODUCTION

Gaseous wastes of some chemical plants contain carbon bisulfide (CS_2) impurities. The concentration of carbon bisulfide is especially high in the waste gases of viscose production. One of the methods to clean the air is to create a weakly ionized plasma in a gas flow and activate plasmochemical processes leading to the elimination of the toxic component. To generate plasma, one can use electron beams or electric discharges. Recently, such an approach has found wide application [1, 2].

The results of experiments with sulfur dioxide [3] have shown the advantages of using pulsed electron beams for the ionization of air with the aim of eliminating toxic molecules as compared to continuous beams. The application of pulsed beams initiating non-self-sustained volume discharges allows one to optimize the conversion of the toxic component in terms of energy losses by varying the parameters of the beam and discharge. In particular, the application of pulsed electron beams makes it possible to realize a chain mechanism of SO_2 oxidation in plasma [3]. The efficiency of this process is several times higher than the efficiency provided by continuous electron beams. In this paper, we present the results of experimental studies and numerical simulations of the physicochemical processes induced by pulsed electron beams in a waste gas, which is modeled by an atmospheric-pressure nitrogen–oxygen mixture with a carbon bisulfide impurity.

2. EXPERIMENTAL

In this study, all the experiments were carried out in model gas mixtures containing nitrogen and oxygen. The relative contents of molecular nitrogen and molecular oxygen varied from 89 to 99 and from 0.1 to 10%, respectively. The concentration of CS_2 varied from 50 ppm to 1%. The mixtures were prepared in a special mixer equipped with an impeller for the forced mixing of gases. All the experiments were carried out at atmospheric pressure. The composition of the mixture and the reaction products was analyzed by the method of gas chromatography with the use of a heat conduction detector and a 3-m column filled with a Silochrome-120 sorbent. A gas sample with a volume of 2 cm^3 was taken from a constantly pumped (at a rate of 100 cm^3/min) loop of the sampler valve. The typical error of the measurements of the impurity concentration was no greater than 10% in the range 10–100 ppm and 3% for the impurity concentration exceeding 100 ppm.

Model gas mixtures were irradiated by two electron accelerators with different parameters. In nanosecond experiments, we employed a small-size Radan accelerator with an explosive emission cathode [4]. The parameters of its electron beam are as follows: the electron energy is 180 keV, the full width at half maximum is $\tau \sim 3$ ns, the current density in a beam with a cross section of 1 cm^2 is $j \sim 800$ A/cm^2, and the pulse repetition rate is up to 10 s^{-1}. We used this accelerator to irradiate a small gas volume (about 10 cm^3) in a chamber with a total volume of 3 l. In some experiments, we observed a non-self-sustained discharge initiated by a nanosecond electron beam. The discharge was ignited in a 1-cm-long gap connected to a 12-nF capacitor charged to a voltage of 1.0–25.0 kV.

In microsecond experiments, we used a single-pulse accelerator with a plasma cathode similar to that described in [5]. This accelerator produces a radially divergent electron beam with a cross section of about 1.5 m^2, electron energy of 280–300 keV, current density of $j \sim 10^{-4}$–10^{-2} A/cm^2, and pulse duration of $\tau \sim 48$ μs.

1063-7842/03/4806-0780$24.00

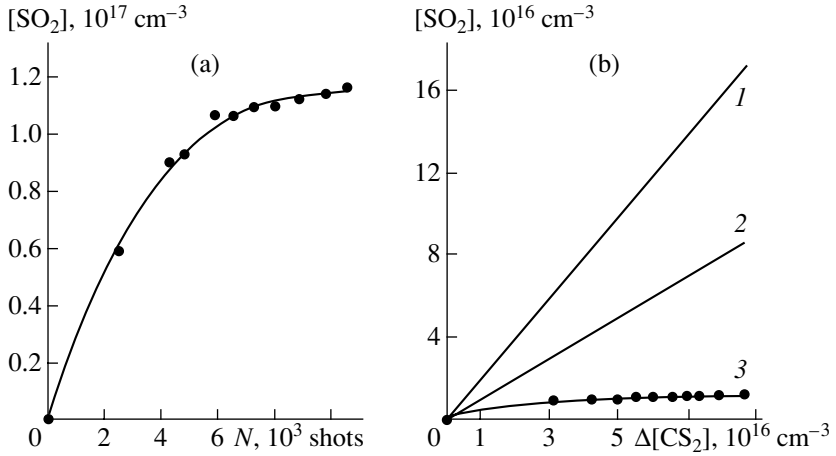

Fig. 1. Density of SO_2 molecules produced in the $N_2 : O_2 : CS_2 = 89 : 10 : 1$ mixture exposed to nanosecond electron beams vs. (a) the number N of pulses and (b) the amount $\Delta[CS_2]$ of the impurity molecules eliminated: (*1*) complete oxidation of CS_2, (*2*) incomplete oxidation of CS_2, and (*3*) experimental data.

The irradiated volume was 170 l. The design of the setup allows the ignition of a non-self-sustained discharge maintained by an electron beam. The discharge was ignited between two coaxial electrodes. The first electrode was the grounded exit foil of the accelerator, and the second one was a metal cylinder connected to a 9-nF capacitor charged to a voltage of 0–20 kV. The interelectrode distance was 10 cm.

The gas under study was irradiated by a series of pulses. The energy deposited in the gas by one nanosecond pulse was about 3×10^{-4} J/cm^3. For a microsecond pulse, this energy was about 10^{-4} J/cm^3, which means that the energies deposited were of the same order of magnitude. In experiments, we determined the absolute decrease ΔC in the impurity concentration, the degree of conversion, and the energy needed for the elimination of one impurity molecule. The degree of conversion (the degree of gas purification) η and the energy loss ε per one molecule (in eV/molecule) were calculated as

$$\eta = \Delta C/C_0, \quad \varepsilon = (W_b + W_E)N/e\Delta C.$$

Here $\Delta C = C_0 - C$ is the decrease in the concentration of carbon bisulfide after a series of pulses (in cm^{-3}); C_0 and C are the initial and current concentrations of the impurity in the mixture under study (in cm^{-3}); W_b and W_E are the energies deposited per one pulse by the electron beam and the external electric field, respectively (in J/cm^3); N is the number of pulses in a series; and e is the electron charge (in C). The total error of the measurements of ε, with allowance for the error in determining the specific absorbed energy $W = W_b + W_E$, is no greater than 0.3.

3. MODES OF CARBON BISULFIDE CONVERSION

3.1. Nanosecond Mode

In experiments on the irradiation of a mixture by a nanosecond electron beam, we obtained the dependences of the carbon bisulfide concentration on the number of pulses. The results obtained show a linear decrease in the CS_2 concentration with increasing number of pulses (energy deposition). Note that the initial concentration of CS_2 weakly affects the slope of the curves obtained. In these experiments, the minimum energy loss (ε = 0.8 eV/molecule) was significantly lower than the dissociation energy of carbon bisulfide (4.5 eV). Therefore, the mechanism for CS_2 elimination cannot be related to the dissociation resulting from direct electron impact. Recently, we observed a similar effect of pulsed irradiation in experiments with an N_2–O_2–SO_2 mixture exposed to pulsed electron beams [3]. The reason was the chain mechanism for SO_2 oxidation. A possible reason for the low energy spent on the elimination of carbon bisulfide may also be the chain mechanism for CS_2 oxidation [6, 7]. A piece of evidence in support of the oxidation of the impurity is the presence of sulfur dioxide as the final product of CS_2 conversion in experiments with a nanosecond electron beam. In addition, we detected small amounts of carbon (CO and CO_2) and nitrogen (NO and NO_2) oxides. Figures 1a and 1b show the resulting SO_2 concentration versus the number of pulses and the amount $\Delta[CS_2]$ of the converted carbon bisulfide molecules, respectively, for the $N_2 : O_2 : CS_2 = 89 : 10 : 1$ mixture.

With a sufficiently high concentration of oxygen, the complete oxidation of CS_2 proceeds via the reaction

$$CS_2 + 3O_2 \longrightarrow CO_2 + 2SO_2. \quad (1)$$

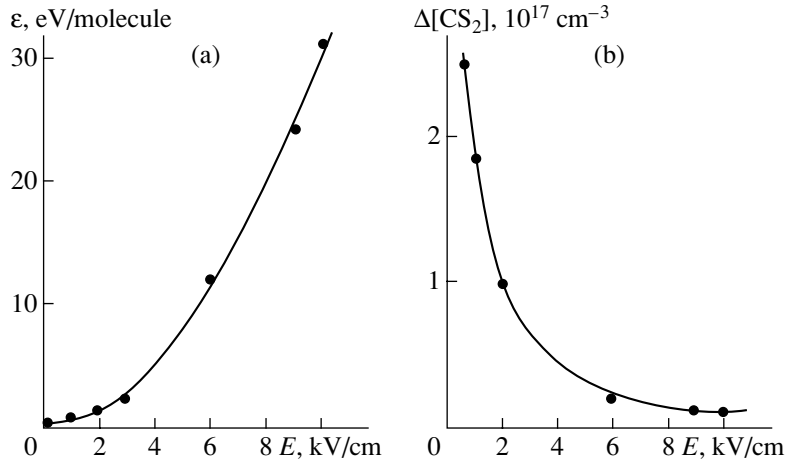

Fig. 2. (a) Energy ε spent on the elimination of one impurity molecule and (b) the amount $\Delta[CS_2]$ of the molecules eliminated vs. electric field strength E in a non-self-sustained discharge initiated by a nanosecond electron beam: (circles) experimental data and (solid lines) the results of calculation.

At an insufficient concentration of oxygen, an incomplete oxidation resulting in the production of CS radicals takes place:

$$CS_2 + O_2 \longrightarrow CS + SO_2. \quad (2)$$

The stoichiometric ratio of carbon bisulfide and sulfur dioxide (the oxidation products of the above reactions) shows that in the case of oxidation of equal amounts of CS_2, the yield of SO_2 in reaction (1) is twice as large as that in reaction (2).

Curve *3* in Fig. 1b shows the sulfur dioxide concentration as a function of the amount $\Delta[CS_2]$ of carbon bisulfide molecules eliminated in the model mixture exposed to a nanosecond beam. It is seen that the concentration of sulfur dioxide first increases with $\Delta[CS_2]$ and then stabilizes at a certain level. The experimental dependence of the SO_2 concentration on the number of pulses initiating CS_2 decomposition (curve *3*) is not described by reactions (1) and (2) represented in Fig. 1b by curves *1* and *2*, respectively.

In non-self-sustained volume discharges, the discharge is maintained regardless of the voltage applied to the plasma. In view of this fact, one can optimize the energy transfer to certain degrees of freedom of the molecule by varying the electric field strength [8]. For example, in [3] it was experimentally demonstrated that such an approach allows one to significantly decrease the energy ε spent on the elimination of one SO_2 molecule [3]. In particular, the chain processes of SO_2 oxidation are accelerated in weak fields with a strength of about 150–200 V/cm due to the increase in the concentration of the vibrationally excited oxygen molecules O_2^* [9].

In this context, it is expedient to analyze the effect of the external electric field on the conversion of carbon bisulfide molecules. Figure 2 shows the amount $\Delta[CS_2]$ of the molecules eliminated and the energy loss ε versus the electric field strength in the volume discharge. It can be seen (Fig. 2a) that at the electric field strength in the range 0–1.5 kV/cm, the value of ε varies insignificantly in the range 0.6–0.9 eV/molecule. This level is substantially lower than the dissociation energies of both carbon bisulfide and oxygen. The minimum value ε = 0.6 eV/molecule is obtained at the electric field strength E = 0.5 kV/cm, which corresponds to the maximum value of $\Delta[CS_2]$ in Fig. 2b. The increase in the electric field results in a substantial increase in the energy loss ε: for E > 10 kV/cm, ε is greater than 35 eV/molecule. Such an increase is related not only to an increase in the energy deposited in the gas but also to a decrease in the amount $\Delta[CS_2]$ of the molecules converted (Fig. 2b).

Thus, the results of experiments with the gas mixture exposed to nanosecond electron beams show that (i) the energy needed for the conversion of one CS_2 molecule is significantly lower than the binding energy, (ii) the probable mechanism for conversion involves the oxidation of CS_2 yielding SO_2, and (iii) the conversion efficiency decreases in the presence of an external electric field.

3.2. Microsecond Mode

When the gas mixture was exposed to low-current microsecond beams, we also observed a linear decrease in the CS_2 concentration with increasing number of pulses. Figure 3 shows the amount $\Delta[CS_2]$ of the molecules eliminated and the energy loss ε versus the initial concentration of CS_2. In the interval of the initial concentrations $[CS_2]_0$ = 0.3–1.0%, the amount of the molecules eliminated and the energy loss ε range from 3.7×10^{16} to 5.25×10^{16} cm^{-3} and from 1.2 to 1.8 eV/molecule, respectively.

In a gas mixture irradiated by a low-current microsecond electron beam, the chain mechanism for the oxi-

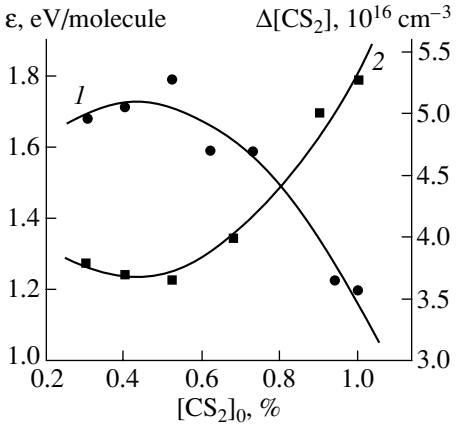

Fig. 3. (*1*) Energy loss ε and (*2*) the amount $\Delta[CS_2]$ of the molecules eliminated vs. initial relative concentration $[CS_2]_0$ of carbon bisulfide in the mixture exposed to a low-current microsecond electron beam.

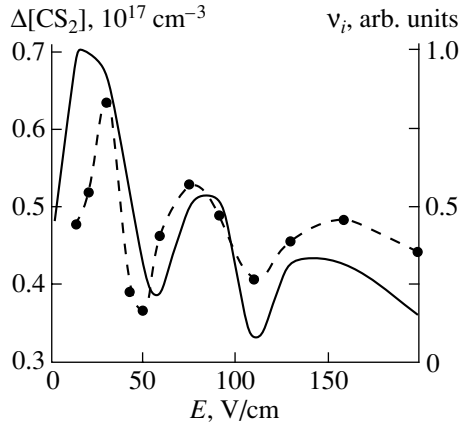

Fig. 4. The amount $\Delta[CS_2]$ of carbon bisulfide molecules eliminated by a series of 30 microsecond electron beams (circles) and the calculated rate ν_i of the generation of O_2^- negative ions in reaction (12) (solid line) vs. electric field strength E in a non-self-sustained discharge.

dation of CS_2 cannot be the reason for low values of ε because, in this mode, we have detected a very low (almost trace) concentration of sulfur dioxide, which is the final product of the chain mechanism.

After the irradiation of the model gas mixture by a microsecond electron beam, the side surfaces and the bottom of the plasmochemical reactor were covered with a yellow solid substance, whereas the top horizontal surface was covered with a viscous brown liquid. An elemental analysis showed that the composition of the final products was represented by $(–CS–)_n$ compounds. The phase state of a $(–CS–)_n$ compound depends on its molecular mass (i.e., on the number n). Low-molecular compounds are sublimated and condensed on the upper part of the reactor as a liquid. The compounds with large n precipitate as a solid powder. However, we did not determine the molecular masses of the condensed phases. Thus, for the given mode, the probable mechanism for carbon bisulfide conversion involves the polymerization of CS radicals.

We note that the dependences of the amount $\Delta[CS_2]$ of the molecules eliminated and the energy loss ε on the electric field strength in the ionized gas mixture are very different. In experiments with a volume discharge maintained by a microsecond electron beam at an electric field strength in the interelectrode gap of $E \sim 10$–500 V/cm, we observed a pronounced nonmonotonic dependence of $\Delta[CS_2]$ on E (Fig. 4).

4. SIMULATION OF CARBON BISULFIDE CONVERSION

4.1. Numerical Model

To analyze mechanisms for CS_2 conversion, we developed an original model of the plasmochemical processes initiated by pulsed electron beams in nitrogen–oxygen mixtures with a carbon bisulfide impurity. The model includes the calculation of the beam energy deposition in the gas mixture and (in the case of the volume discharge ignition) the energy supply from a capacitive storage to the discharge gap, as well as the calculation of the plasma parameters, the rate constants of the elementary processes involving electrons, and the kinetics of the charged particles and free radicals. We start by calculating the electron energy distribution function using the data on the spectra of the plasma electrons (the rates of the electron production in a unit energy interval in 1 cm^3 per one second) obtained by the Monte Carlo method. The rate constants for the production of electrons, ions, and radicals under the action of an electron beam are determined by calculating the contributions from the secondary electrons of the degradation cascade to the ionization and dissociation of O_2 and N_2. In these calculations, we determine the distribution of the electron beam energy over the electronic and vibrational states of the gas mixture components, the initial concentrations of these components, and the concentrations of the charged particles. Based on the data obtained, we calculate the kinetics of the charged particles and free radicals. Besides the reactions involving electrons, the model takes into account the processes involving positive and negative ions, molecules, and radicals in the ground and excited states. In total, the model includes more than 300 processes. Thus, we solve a self-consistent problem: on the one hand, we take into account the effect of all types of energy losses on the electron energy distribution function and, on the other, we calculate the rate constants and electron energy losses for various excitation channels using this distribution function. Based on the simultaneous numerical solution of the kinetic and Boltzmann equations, we determine the plasma param-

eters, such as the electron mobility and plasma conductivity.

4.2. Oxidative Conversion of CS_2 Exposed to High-Current Nanosecond Beams

The results of experiments show that the final products of CS_2 decomposition contain gaseous sulfur and carbon oxides. This allows us to assume that the conversion of carbon bisulfide involves oxidation mechanisms. The low energy spent on the conversion of one molecule points to the possibility of a chain process.

It follows from the calculations by the above model that the main part of the beam energy is spent on the ionization and dissociation of N_2 and O_2. In particular, it is found that, in the case of excitation by high-current nanosecond electron beams, the electron density is about 10^{14} cm^{-3}. An analysis of the kinetics of plasmochemical processes shows that the characteristic times of the main impurity conversion reactions are significantly longer than the duration of the pulsed electron beam. Figure 5 demonstrates the time dependences of the concentrations of the main components of a plasma created by (a) electron beams and (b) non-self-sustained volume discharge at an electric field strength of 10 kV/cm. Here, we do not present the dependences for charged particles, because the major part of electrons and ions is lost in electron–ion recombination reactions in a time of $t < 1$ μs.

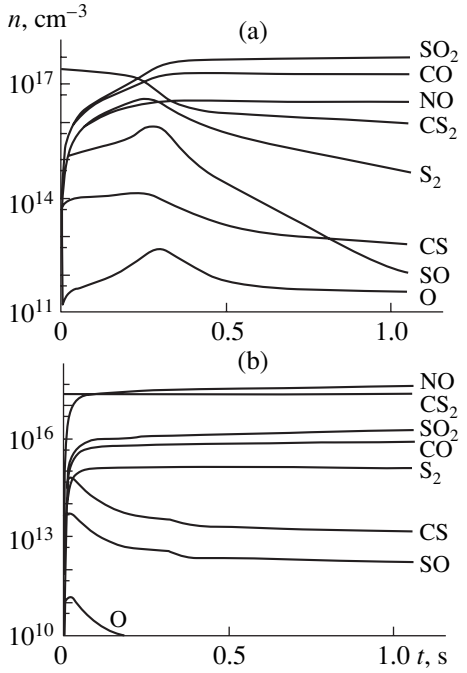

Fig. 5. Time dependences of the concentrations of the main plasma components produced in the $N_2 : O_2 : CS_2 = 89 : 10 : 1$ mixture by (a) a nanosecond electron beam and (b) a non-self-sustained volume discharge with an electric field of 10 kV/cm.

It can be seen from Fig. 5a that the decrease in the CS_2 concentration takes place after the end of the beam current pulse. In this case, the main final products are the gaseous oxides of sulfur, nitrogen, and carbon. Our analysis shows that the ionization by a high-current nanosecond beam gives rise to the chain oxidation of CS_2 with atomic oxygen as an initiating reagent:

$$CS_2 + O \longrightarrow CS + SO, \quad (3)$$

$$SO + CS \longrightarrow CO + S_2, \quad (4)$$

$$O + S_2 \longrightarrow SO + S, \quad (5)$$

$$SO + O_2 \longrightarrow SO_2 + O, \quad (6)$$

$$S + O_2 \longrightarrow SO + O. \quad (7)$$

In the atmospheric-pressure N_2–O_2–CS_2 mixture exposed to an electron beam, atomic oxygen is generated via O_2 electronic states and reactions involving electronically excited N_2^* molecules:

$$e + O_2 \longrightarrow 2O, \quad (8)$$

$$N_2^* + O_2 \longrightarrow N_2 + 2O. \quad (9)$$

In the case at hand, the contribution of dissociative attachment into the generation of O radicals is significantly lower than that of the reactions proceeding via electronically excited states of oxygen and nitrogen. We note that nearly half of the electron beam energy is spent on the formation of such excited states. If the beam intensity is high enough, the amount of O radicals generated in reactions (8) and (9) is sufficient for the initiation of chain mechanism (3)–(7). When the CS_2 concentration is lower than a certain critical level, the chain mechanism is terminated, atomic oxygen is not produced in reactions (6) and (7), and the oxidation of carbon bisulfide comes to a stop. The results of calculations show that the main products of CS_2 decomposition are sulfur and carbon oxides (SO_2 and CO). Small amounts of NO, SO, and S_2 are also produced (Fig. 5).

The processes of CS_2 oxidation in a N_2–O_2–CS_2 mixture are similar to the processes that take place in a CO chemical laser [10, 11]. However, the mechanism for the conversion of carbon bisulfide is somewhat modified due to the presence of molecular nitrogen. The primary reason for this is the reaction of the vibrationally excited nitrogen molecules with atomic oxygen [12]:

$$N_2(v) + O \longrightarrow NO + N, \quad (10)$$

which leads to a decrease in the amount of free oxygen radicals due to the synthesis of nitrogen oxides and a reduction in the efficiency of CS_2 conversion as a whole. Another process increasing the rate of NO production in an ionized mixture is the reaction

$$N + O_2 \longrightarrow NO + O. \quad (11)$$

These reactions provide an effective chain mechanism for the synthesis of nitrogen monoxide [13].

When the mixture is excited only by an electron beam, the concentration of the vibrationally excited nitrogen molecules is low. Therefore, the mechanism involving reactions (10) and (11) plays a significant role only with a sufficiently high electric field corresponding to the ignition of a non-self-sustained volume discharge initiated by an electron beam. In this case, the concentration of the vibrationally excited nitrogen molecules $N_2(v)$ becomes fairly high and the efficiency of reactions (10) and (11) increases. At the electric field strength $E \sim 10$ kV/cm, the loss of O radicals is so intense that this mechanism almost completely inhibits the chain process of CS_2 oxidation.

It can be seen from Fig. 5b that the ignition of a volume discharge results in a decrease in the concentrations of atomic oxygen and sulfur and carbon oxides and a substantial (by more than one order of magnitude) increase in the NO concentration. It is the production of nitrogen oxide and the corresponding decrease in the amount of oxygen radicals that lead to the experimentally observed (Fig. 2) decrease in the amount $\Delta[CS_2]$ of the molecules converted and an increase in the energy loss ε.

Our model makes it possible to calculate the degree of conversion and the energy spent on the elimination of one carbon bisulfide molecule.

4.3. Ion–Molecular Processes of CS_2 Conversion in a Gas Exposed to a Low-Current Microsecond Beam

Another mechanism for carbon bisulfide conversion is realized in the case of the excitation of an N_2–O_2–CS_2 mixture by a low-current microsecond beam. The intensity of such a beam is fairly low and the density of the plasma electrons is lower than that in the case of a high-current beam by two orders of magnitude. The density of charged particles generated by a low-current beam is low and comparable to the density of free oxygen radicals (Fig. 6). In this case, the chain mechanism is not realized, although CS_2 oxidation takes place. A comparison of Figs. 5a and 6 shows that the concentration of SO_2 (the main gaseous product of CS_2 oxidation) in experiments with a microsecond beam is lower than that in the case of a nanosecond beam by three orders of magnitude. Based on the results of calculations, we can conclude that the loss of ions due to electron–ion recombination sharply decreases; as a result, favorable conditions for CS_2 conversion via the formation of cluster ions are produced. Under our experimental conditions, the density of such ions is $A^{\pm} \sim 10^{13}$ cm^{-3} (Fig. 6). The CS_2 concentration in the ionized mixture decreases in fast reactions of cluster formation. The rate of these processes is substantially higher than the rate of reactions involving free oxygen radicals. The presence of $A^{\pm}$ clusters was demonstrated in [14] using mass-spectrometry technique; in that study, the $A^{\pm}$ ions were also produced by an electron beam. In clusters consisting mainly of CS radicals, there are favorable conditions for effective polymerization yielding condensed $(–CS–)_n$ compounds.

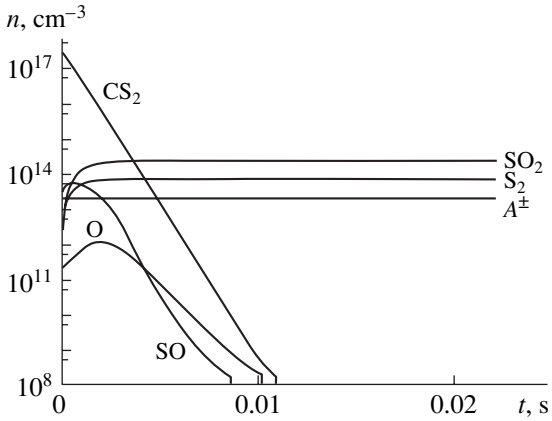

Fig. 6. Same as in Fig. 5 for a plasma created by a low-current microsecond electron beam.

The results of calculations show that the amount of electrons decreases mainly due to the three-body attachment to oxygen:

$$e + O_2 + M \longrightarrow O_2^- + M. \qquad (12)$$

Here, the third particle M is either an oxygen or a nitrogen molecule.

It is known [15] that the rate of such a process significantly depends on the electron energy (i.e., on the external electric field). Therefore, the electric field in the non-self-sustained discharge maintained by an electron beam significantly affects the conversion process (Fig. 4). In the case under consideration, the dependence of the rate ν_i of process (12) on the external field strength is nonmonotonic (solid line in Fig. 4). It can be seen that the model proposed qualitatively describes the experimental data. However, due to the absence of the necessary data on the rate constants of ion–molecular reactions resulting in the formation of clusters with the participation of O_2^- ions, we failed to perform a quantitative comparison of the conversion efficiency.

5. CONCLUSION

The results of a comprehensive study of the action of pulsed electron beams on the carbon bisulfide impurity in an atmospheric-pressure nitrogen–oxygen mixture are presented. The mixture is irradiated by either a high-current nanosecond electron beam or a low-current microsecond electron beam.

It is demonstrated that the energy spent on the conversion of one CS_2 molecule is significantly lower than its dissociation energy. It is experimentally demonstrated that there are two mechanisms for CS_2 conversion that differ in the final products. The final products of the impurity conversion proceeding by the first mechanism are the oxides of the gas mixture compo-

nents (SO_2, CO, CO_2, NO, and NO_2), and those of the conversion proceeding by the second mechanism are liquid or solid $(–CS–)_n$ compounds.

A theoretical model for calculating CS_2 conversion in an ionized nitrogen–oxygen mixture is developed. Two modes of irradiation are investigated at nearly the same specific energies deposited in the gas mixture. It is shown that, in the case of a high-current nanosecond beam, the production rate of free oxygen radicals is high and the main mechanism for impurity conversion involves chain oxidation. In the case of a low-current microsecond beam, the concentration of radicals is low and the CS_2 conversion proceeds via the ion channel with the formation of clusters and impurity polymerization. At intermediate pulse durations and beam current densities, these mechanisms operate simultaneously.

The effect of the electric field on the CS_2 conversion is investigated both experimentally and numerically for two regimes of the interaction of an electron beam with a gas mixture that are characterized by the predominant generation of either free radicals or ions. It is demonstrated that, in a plasma created by a high-density electron beam, the presence of an external electric field leads to the production of nitrogen oxides (NO and NO_2) and a corresponding decrease in the rate of the main conversion mechanism—the chain oxidation of CS_2. Another mechanism is realized when the mixture is exposed to a low-current beam. The decrease in the intensity of electron–ion recombination results in an increase in the contribution of ion–molecular processes, the formation of clusters, and carbon bisulfide polymerization. In a weak electric field, the efficiency of conversion of CS_2 molecules contained in ionized air becomes very sensitive to the electric field strength. The reason is that the production rate of ions forming clusters strongly depends on the electric field strength.

REFERENCES

1. H. Matzing, *Non-Thermal Plasma Techniques for Pollution Control*, Ed. by B. M. Penetrante and S. E. Schultheis (Springer-Verlag, Berlin, 1993), NATO ASI Series, Vol. G34, Part A, p. 59.
2. A. K. Pikaev, Khim. Vys. Énerg. **34** (3), 163 (2000).
3. Yu. N. Novoselov, G. A. Mesyats, and D. L. Kuznetsov, J. Phys. D **34**, 1248 (2001).
4. A. S. El'chaninov, A. S. Kotov, V. G. Shpak, *et al.*, Élektron. Tekh., Ser. 4, No. 2, 33 (1987).
5. A. M. Efremov, B. M. Koval'chuk, Yu. E. Kreĭndel', *et al.*, Prib. Tekh. Éksp., No. 1, 167 (1987).
6. G. V. Denisov, Yu. N. Novoselov, and I. E. Filatov, Pis'ma Zh. Tekh. Fiz. **21** (18), 89 (1995) [Tech. Phys. Lett. **21**, 765 (1995)].
7. G. V. Denisov, D. L. Kuznetsov, and Yu. N. Novoselov, Khim. Vys. Énerg. **31**, 471 (1997).
8. Yu. I. Bychkov, Yu. D. Korolev, G. A. Mesyats, *et al.*, *Injection Gas Electronics* (Nauka, Novosibirsk, 1982).
9. Yu. N. Novoselov, Pis'ma Zh. Tekh. Fiz. **19** (23), 58 (1993) [Tech. Phys. Lett. **19**, 760 (1993)].
10. B. N. Bronfin and W. Q. Jeffers, *Handbook of Chemical Lasers*, Ed. by R. W. Gross and J. F. Bott (Wiley, New York, 1976; Mir, Moscow, 1980).
11. A. V. Eletskiĭ, in *Plasma Chemistry*, Ed. by B. M. Smirnov (Atomizdat, Moscow, 1974), p. 67.
12. Yu. N. Novoselov and A. I. Suslov, Dokl. Akad. Nauk **384** (4), 1 (2002).
13. V. D. Rusanov and A. A. Fridman, *Physics of Chemically Active Plasma* (Nauka, Moscow, 1984).
14. K. Hiraoka, S. Fujimaki, K. Aruga, and S. Yamabe, J. Phys. Chem. **98**, 1802 (1994).
15. H. S. W. Massey, *Negative Ions* (Cambridge Univ. Press, 1976; Mir, Moscow, 1979).

Translated by A. Chikishev

Technical Physics, Vol. 48, No. 6, 2003, pp. 787–792. Translated from Zhurnal Tekhnicheskoĭ Fiziki, Vol. 73, No. 6, 2003, pp. 130–135.

EXPERIMENTAL INSTRUMENTS AND TECHNIQUES

Effect of Intense Energy Fluxes on Vacuum-Tight Rubber

B. A. Demidov, V. P. Efremov, M. V. Ivkin, A. N. Meshcheryakov, and V. A. Petrov

Russian Research Centre Kurchatov Institute, pl. Kurchatova 1, Moscow, 123128 Russia

e-mail: kingsep@dap.kiae.ru

Received June 11, 2002; in final form, December 15, 2002

Abstract—Vacuum-tight rubber is widely used in various structures. Composite shields protecting space vehicles against micrometeorites include natural-rubber-like materials, which are similar to vacuum-tight rubber in properties. It is, therefore, of interest to study the behavior of vacuum-tight rubber subjected to an intense pulsed energy flux. From the response of this material to a high-energy action, one can construct a behavioral model for natural-rubber-like materials used in the protective shields. The behavior of vacuum-tight rubber exposed to a high-energy pulsed electron beam (with an energy density of up to 600 J/cm^2 and a half-width of 10^{-7} s) is investigated. The rate of material expansion toward the beam, as well as the velocities of propagation of strong and weak perturbations over the material, is measured. Spalls on the front and back surfaces of the sample are detected. The Grüneisen constant (a fundamental thermodynamic parameter that characterizes the pressure in a medium upon isochoric heating) is found. © *2003 MAIK "Nauka/Interperiodica".*

INTRODUCTION

From the behavior of condensed materials under fast energy release conditions, one can extract information on the equation of state in parameter ranges that are beyond the reach of shock-wave experiments. Under the conditions of pulsed volumetric energy release, the behavior of complex-structure materials, such as rubber, which contains long molecular chains, may be essentially nonequilibrium, since the energy distribution over degrees of freedom takes a certain time. Therefore, if an external action is short, the pressure developed in the material may differ substantially from that predicted from the equation of state based on shock adiabats.

When the action of corpuscular radiation is short, energy release may have no effect on the material density and be accompanied by a pressure increase in the absorption zone and by the excitation of compression–rarefaction waves in nearby layers. Available information on equations of state for rubberlike materials is today contradictory. For example, there is a spread in the speed of sound C_s in vacuum-tight rubber: from 1500 m/s [1] to 1040 m/s [2]. The longitudinal speed of sound in natural rubber was reported to be C_s = 1480 m/s [3]. Data for the Grüneisen constant are lacking. That is why new experimental information regarding the properties of rubberlike materials under the action of pulsed corpuscular radiation is of great value. These data would be helpful in constructing behavioral models for impulse-loaded polymer materials used in space technology.

In this work, we study the effect of a high-current pulsed electron beam ($J \le 20$ kA, E = 300 keV, half-width $\tau_0 \le 10^{-7}$ s) on white vacuum-tight rubber. The rate of expansion of irradiated rubber toward the electron beam, as well as the propagation velocity of large and small perturbations over the material, is measured. Conditions for spalling the back surface of rubber targets are investigated.

The Grüneisen constant Γ_0 of white vacuum-tight rubber is determined from the thermoelastic response of the material to the pulse electron beam. Upon rapid heating, when the time τ_r of acoustic relaxation of the energy release zone far exceeds the electron pulse duration τ_0, the material heats up with its volume remaining almost constant. In this case, the pressure of the thermoelastic response is proportional to the energy absorbed and the Grüneisen constant. Detecting the thermoelastic response with a differential laser interferometer and knowing a fraction of the electron beam energy absorbed in the target, one can find the Grüneisen constant. The accuracy of determining the Grüneisen constant rises with decreasing electron pulse duration. Such an approach to measuring the Grüneisen constant (for details, see [4]) was used to find this parameter in PMMA [5].

EXPERIMENTAL

The experimental scheme for studying the interaction of a pulsed high-current electron beam (electron energy E = 300 keV, current J = 20 kA, pulse half-width τ_0 = 100 ns) with a rubber target is depicted in Fig. 1. The exit part of a Kal'mar accelerator [6] (which incorporates casing *1*, output transformer *2*, insulating separating diaphragm *3*, and cylindrical cathode *4* with a diameter of 20 mm) communicates with vacuum chamber *5*. Electron beam *6* falls on rubber sample *7*, which is pressed against anode *8* by a 5-mm-thick brass plate.

1063-7842/03/4806-0787$24.00

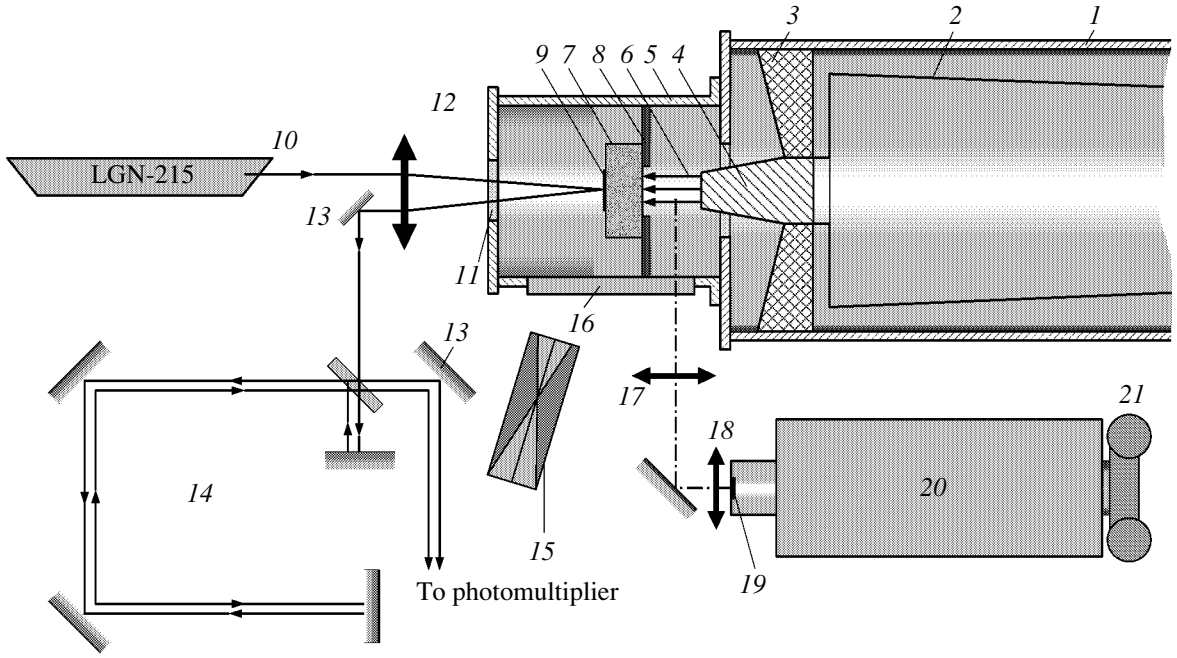

Fig. 1. Experimental scheme.

The plate has a hole of diameter 9 mm, which corresponds to the size of the central part of the target's rare surface. Aluminum foil *9* 10 μm thick is glued to the central part of the target's rare surface with a vacuum lubricant. Beam *10* (λ = 0.63 μm) from an LGN-215 laser is passed through transparent vacuum-tight window *11* and focused on the central part of foil *9* by lens *12*. The Doppler-modified beam reflects from mirror *13* toward differential laser interferometer *14*. Having passed through the interferometer, the beam strikes the photocathode of a photomultiplier, whose signal is detected by an S1-75 oscilloscope. The time synchronization of the interferometer signal is accomplished by using the X-ray label of a stilbene sensor, which records the bremsstrahlung of the electron beam. The sensor is placed inside vacuum chamber *5*. Scintillations of the X-ray sensor are applied to the input of the photomultiplier through an optical fiber.

Scintillations due to beam–rubber interaction pass through transparent lateral window *16* made of organic glass and are focused on entrance slit *19* of FÉR-7 streak camera *20* by means of lenses *17* and *18*. The streak camera provides the linear transfer of scintillations from the beam–rubber interaction area. In the experiments, we used a sweep rate of 0.25 μs/cm. The width of the entrance slit was 0.1 mm. The image produced on the exit screen of the streak camera was displayed in photographic attachment *21*. The samples of white vacuum-tight rubber (type 78-89) were 32–50 mm in diameter and 1.75–6.5 mm in thickness.

The diameter of the interaction area was checked from the bremsstrahlung intensity with camera-obscura *15* placed at an angle of 30° to the axis of the setup. In other experiments where the dimension of the energy release area was studied more carefully, the camera-obscura was placed behind rubber target *7* on the axis of the setup.

The total current of the accelerator and the cathode voltage were measured with a shunt and an electronic meter both mounted in the diode chamber of the accelerator.

Experiments in which the Grüneisen constant for the rubber was determined employed another scheme. To sharpen the current pulse, the high-current electron beam was directed to the accelerator's metallic anode, which had a central hole of diameter 3 mm covered by a 10-μm-thick aluminum foil. In this way, the half-width τ_b of the pulsed beam leaving the hole was made several times smaller than the half-width τ_0 of the total current of the accelerator. Then, a set of diaphragms cut a beam of diameter 20 mm from the general divergent beam and the former was directed to the target. A Rogowski loop measured the current of the beam falling on the rubber sample [5].

EXPERIMENTAL RESULTS

Figure 2 shows typical voltage, high-current, and low-current waveforms in the diode gap of the Kal'mar accelerator that were obtained in the sharpening mode. It is seen that the half-width τ_b of the low current, which was measured by the Rogowski loop, is much shorter than that of the accelerator's total current, τ_0, and equals 30–50 ns.

It was found that the action of the high-current (20 kA) electron beam on the rubber samples with

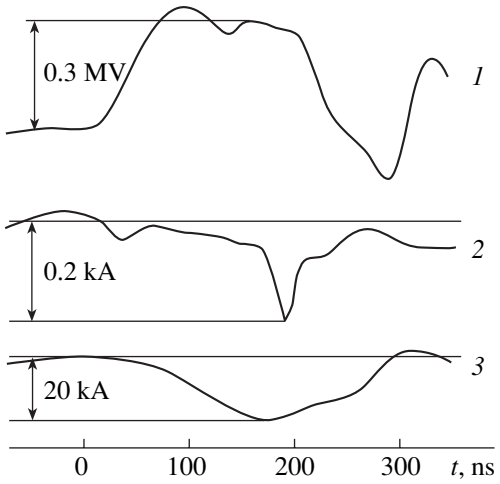

Fig. 2. Waveforms of the (*1*) voltage across the diode gap, (*2*) current measured by the Rogowski loop, and (*3*) accelerator's total current.

thicknesses of up to 4 mm produces a through hole. In thicker samples, a standard platelike ≈2-mm-thick back-surface spall was observed. Simultaneously, a distinct spall 0.20–0.25 mm in depth was observed on the front surface (see Fig. 3). On samples with thicknesses of more than 6 mm, the back spall was absent.

From the photometric measurements of the obscuragrams, it follows that the area of beam–rubber interaction measures 1 cm^2. From this value, one can estimate the energy contribution of the electron beam to the focal spot. With the beam current J = 20 kA, electron energy E = 300 keV, and beam half-width τ_0 = 100 ns, the energy flux density is 600 J/cm^2 and the power flux density is 6×10^9 W/cm^2.

The time instant the shock wave appears on the free surface and the mass velocity profiles were detected with the laser interferometer (Fig. 4). With free-surface velocities exceeding 100 m/s, the resolution of the system was insufficient to record the velocity profile; therefore, the signal from the interferometer was used only to note the instant of shock wave appearance on the free surface. On the left of the interferograms, the narrow pulse of negative polarity (the signal from the X-ray sensor) is seen. This signal is a reference point indicating the beginning of beam–target interaction. From Fig. 4, it follows that the rear surface of the target starts moving with a time delay relative to the reference point and the delay increases with target thickness.

Using targets of varying thickness, one can find the mean velocity of the shock wave and the depth of the energy release zone from the thickness dependence of the time the shock wave appears on the free surface. This dependence is shown in Fig. 5. Curve *1* shows that the mean velocity of the shock is D = 2.0 km/s and the depth of the energy release zone, 0.7 mm.

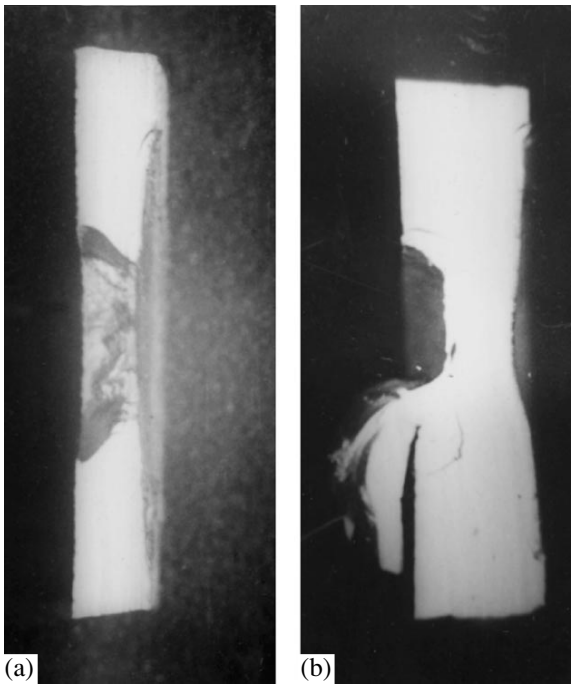

Fig. 3. (a) 3.5- and (b) 5.5-mm-thick rubber targets after the action of the electron beam.

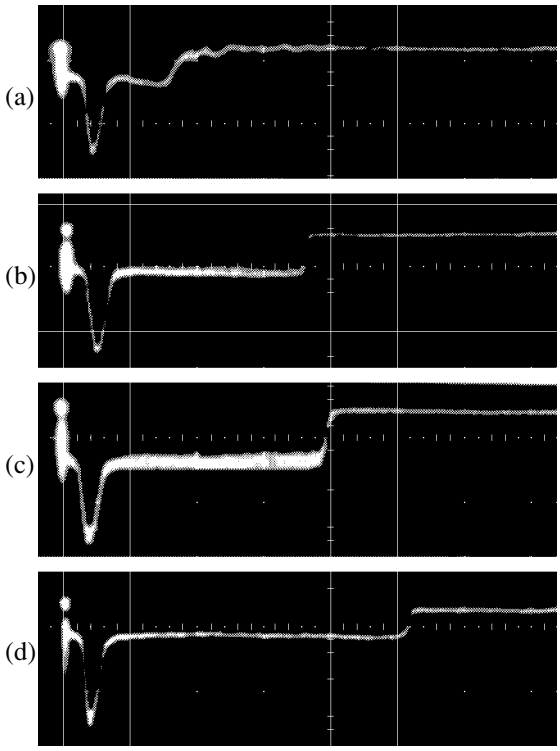

Fig. 4. Interferograms showing the time the shock wave (strong action) reaches the rear surface of the target. The target thickness is (a) 2.0, (b) 4.0, (c) 4.5, and (d) 5.5 mm. The scale division is 0.5 μs.

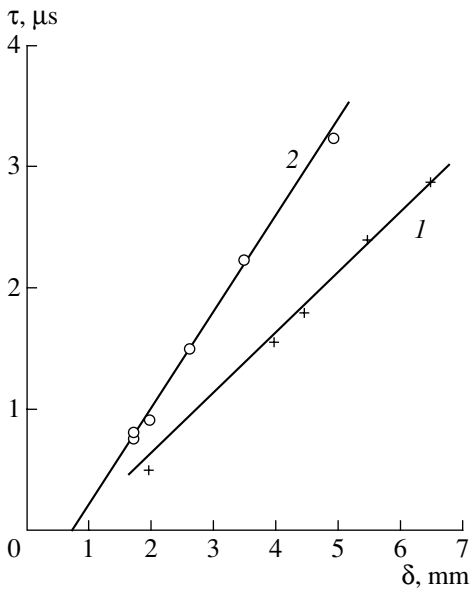

Fig. 5. Time of arrival of (*1*) strong and (*2*) weak perturbations at the free surface vs. target thickness.

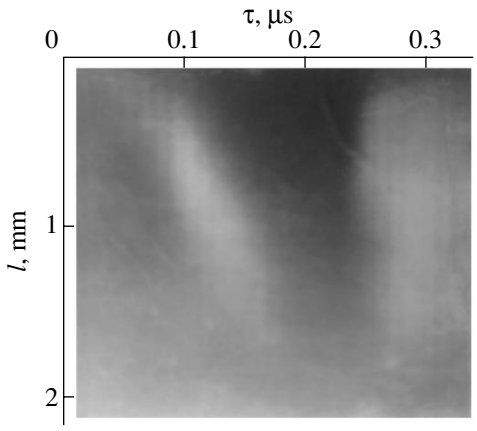

Fig. 6. Streak photograph illustrating the expansion of the rubber toward the electron beam.

When a high-current electron beam interacts with rubber, the latter thermally decomposes and decomposition products from the focal spot propagate toward the beam. The diagram in Fig. 6 demonstrates the dynamics of such interaction. The expansion of the material toward the beam (from top to bottom) is displayed as the inclined bright band with fairly distinct boundaries. Also, the appearance of the cathode plasma (the bright area on the right), whose velocity remained unresolved with the sweep rate used in the experiments, is seen. From Fig. 6, it follows that the expansion velocity of decomposition products reaches 12–15 km/s.

As was noted, the Grüneisen constant was determined under other experimental conditions. It will be recalled that, as the accelerator current was decreased to 8–10 kA (by expanding the cathode–anode spacing), the short divergent electron beam (with a current of 50–400 A) passed through the 3-mm-diameter hole in the anode (Fig. 2). The target was irradiated by an electron beam of diameter 2 cm. A personnel colorimetric dosimeter placed near the target indicated that the cross section of the beam is fairly uniform in intensity: the nonuniformity of energy release zone irradiation was within 10%. In this regime, the interferometer could determine the velocity of the back surface and the time instant the weak perturbation reaches this surface. A set of interferograms in Fig. 7 demonstrates the motion of the free surface of targets with a varying thickness under the action of the electron beam.

Comparing interferograms taken from targets of variable thickness allows one to determine the speed of sound C_s and the depth of the energy release zone from the thermoelastic response. This simplifies experimental data processing.

The time the thermoelastic response reaches the back surface of the target vs. target thickness is plotted in Fig. 5 (curve *2*). It is seen that the velocity of the response, i.e., the longitudinal speed of sound, is C_s = 1.3 ± 0.1 km/s and the energy release depth is h = 0.7 mm. This means that the relaxation time of the energy release zone, $\tau_r = h/C_s$, is 10–12 times longer than the duration τ_b of the electron beam and that the velocity of the back surface must change sign 0.5–0.6 μs after the pulse has been arrived at the back surface. The arrow in Fig. 7c shows the appearance of the thermoelastic response that passed through the target twice. The time it takes for the response to doubly pass the target (the overall thickness 3.5 mm) equals 2.7 μm, which corresponds to C_s = 1.3 km/s (in full accordance with curve *2* in Fig. 5). The interferograms in Figs. 7a and 7b were obtained for the beam energy density 1.0 J/cm^2 and beam durations τ_b = 80 and 40 ns, respectively. From them, the maximal velocity of the back surface is estimated at 11–12 m/s and the maximal travel, at L = 5 and 5.4 μm, respectively. In Fig. 7d, the maximal velocity of the back surface is 13 m/s for the energy density 1.3 J/cm^2 and the maximal travel L = 5.8 μm.

DISCUSSION

The depth of the energy release zone h = 0.7 mm found experimentally correlates well with the value found by numerically simulating electron absorption in the rubber. The extrapolated depth of the energy release zone found by the technique described in [7] equals 0.75 mm for an electron energy of 0.3 MeV.

The depth of the front-surface spall, 0.2–0.3 mm, equals one-third of the energy release zone, which is consistent with [8]. Of interest is the observation of the back-surface spall under conditions that are sometimes hard to realize in explosion experiments. According to [2, 8], the irreversible rupture of rubber under static

stressing takes place when the elongation per unit length reaches 600%. The depth of the back-surface spall lies between 1.5 and 2.0 mm, which is twice that of the energy release zone. This implies a strong spread of the shock wave inside the rubber (the thickness of the piece split off is proportional to the compressing pulse duration). The absence of the spall in 6.0-mm-thick samples and its presence in those with a thickness of 5.5 mm indicates a severe decay in the medium and correlates with the data of shock-wave experiments [8].

From the experimental data for the maximal velocity of the back surface under the action of the high-current electron beam, one can estimate the Grüneisen constant Γ_0 for white vacuum-tight rubber. According to [9],

$$\Gamma_0 = \frac{\rho_0 C_s^2 L \alpha}{Q - \Delta Q},$$

where ρ_0 is the initial density of the rubber, C_s is the speed of sound, L is the maximal travel of the target's back surface, Q is the electron-beam energy density, α is a correction for the finiteness (noninstantaneousness) of the energy release time, and ΔQ is a correction for pressure pulse attenuation with travel in the target.

For thin targets (2.0 to 2.5 times thicker than the energy release zone), the correction for pressure pulse decay can be neglected. The correction α for energy release time finiteness depends on the ratio τ_b/τ_r and the electron beam absorption profile. The effect of finiteness of the energy release time is different for different profiles [5], but this correction can also be neglected, since the ratio $\tau_b/\tau_r \le 0.1$ is small and the contribution from this correction is less than 5%. Assuming that $\rho_0 = 1.2 \times 10^3$ kg/m^3 and $C_s = 1.3$ km/s, we find from the interferograms on Figs. 7a and 7b the effective Grüneisen constant $\Gamma_0 = 1.1 \pm 0.1$.

It should be noted that, while the values of τ_b in Figs. 7a and 7b differ twofold, the Grüneisen constants differ insignificantly, suggesting that the effect of the energy release time finiteness is weak in this work.

From the interferogram in Fig. 7d (thick target), one obtains the somewhat underestimated value of Γ_0, $\Gamma_0 = 0.9$. This means that the correction for pressure pulse decay should be included. In Fig. 7c, the shape of the pulse that doubly passed through the target is changed. This also indicates that the acoustic approximation fails in this case.

Knowing the Grüneisen constant, one can estimate the maximum possible pressure in the focal spot on the rubber for high current densities in the Kal'mar accelerator, assuming that the Grüneisen constant may only decline under these conditions.

With the current density $J = 20$ kA/cm^2, the electron energy $E = 300$ keV, the current pulse duration $\tau_0 = 10^{-7}$ s, and the depth of the energy release zone $h = 0.7$ mm, the volumetric energy density is $w = 8.5 \times 10^9$ J/m^3. Hence, for $\Gamma_0 = 1.1$, the pressure p in the energy release zone is $p \le 55$ kbar.

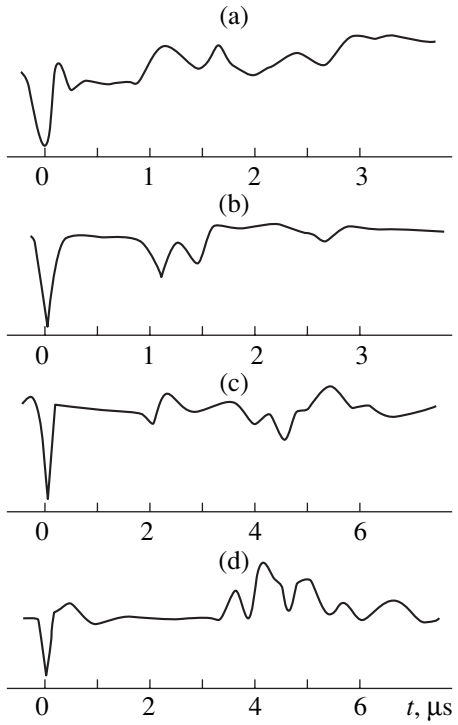

Fig. 7. Interferograms illustrating the motion of the free surface of targets with a thickness of (a) 2.0, (b) 1.75, (c) 1.75, and (d) 5.0 mm.

In order to evaluate the Grüneisen constant for white vacuum-tight rubber over a wide range of energy densities, the numerical simulation of pressure wave generation and propagation is needed. In this work, the Grüneisen constant was measured under the same experimental conditions, which allows one to compare this parameter for various materials and choose the one with a maximal or minimal Grüneisen constant.

CONCLUSIONS

(1) The behavior of the natural-rubber-like material under the action of a short pressure pulse generated by a high-current electron beam was studied.

(2) The thermodynamic Grüneisen constant for white vacuum-tight rubber was measured ($\Gamma_0 = 1.1$), and the pressure wave dynamics under isochoric heating was determined.

(3) Front and rear splits in vacuum-tight rubber samples subjected to a high-current pulsed electron beam were found.

(4) The results obtained could have considerable utility in those areas of technology where short high-current electron and ion beams are used.

REFERENCES

1. G. I. Kanel', S. V. Razorenov, A. V. Utkin, and K. Baumang, Preprint No. 1-394 (Inst. of High Temperatures, Russian Academy of Sciences).
2. Yu. B. Kalmykov, G. I. Kanel', I. P. Parkhomenko, *et al.*, Prikl. Mekh. Tekh. Fiz., No. 1, 126 (1990).
3. *Physical Quantities: A Handbook*, Ed. by I. S. Grigor'ev and E. Z. Melikhov (Énergoatomizdat, Moscow, 1991).
4. F. C. Perry, J. Appl. Phys. **41**, 5017 (1970).
5. B. A. Demidov, M. V. Ivkin, I. A. Ivonin, *et al.*, Zh. Tekh. Fiz. **65** (9), 56 (1995) [Tech. Phys. **40**, 895 (1995)].
6. B. A. Demidov, M. V. Ivkin, V. A. Petrov, *et al.*, At. Energ. **46**, 100 (1979).
7. T. Tabata and R. Ito, Nucl. Sci. Eng. **53**, 1974 (1974).
8. G. I. Kanel', S. V. Razorenov, A. V. Utkin, and V. E. Fortov, *Shock Wave Phenomena in Condensed Media* (Yanus-K, Moscow, 1996).
9. R. B. Oswald, F. B. McLean, D. R. Schallhorn, *et al.*, J. Appl. Phys. **42**, 3463 (1971).

Translated by V. Isaakyan

Technical Physics, Vol. 48, No. 6, 2003, pp. 793–794. Translated from Zhurnal Tekhnicheskoĭ Fiziki, Vol. 73, No. 6, 2003, pp. 136–137.

BRIEF COMMUNICATIONS

Simulation of Ion-Optical Systems for Mass Spectrometers with a Quadrupole Mass Filter

A. A. Trubitsyn

Ryazan State Radio Engineering Academy, ul. Gagarina, Ryazan, 390005 Russia

e-mail: oief@rgrta.ryazan.ru

Received February 20, 2002; in final form, November 25, 2002

Abstract—Ion-optical systems with intermediate focusing and parallel extraction of the ion beam are simulated. The systems may be integrated with a quadrupole mass filter in order to improve the analytical properties of the latter. © *2003 MAIK "Nauka/Interperiodica".*

The numerical simulation of electron (ion)-optical systems conventionally includes the simulation of electrostatic fields and the calculation of charged particle paths in these fields.

A basic advantage of numerical simulation over an analytical approach is the much lower degree of idealization of real structures. This benefit is most significant for techniques based on the calculation of the potential distribution in an unbounded domain, i.e., techniques for solving the outer Dirichlet problem. Among advanced methods appearing in the computer era, the boundary-element approach [1] seems to be the most promising for solving problems of mathematical physics. However, this method suffers from a serious disadvantage, namely, the low accuracy of potential calculation (the error may exceed 100%) near the boundary of a domain of simulation because of the integrand singularity.

Earlier [2], a method for calculating singular integrals appearing in the solution of the inner Dirichlet problem was suggested. Here, this method is extended to estimate singular integrals in the solution of the outer problem. The formulas derived are essentially the same (with minor corrections) as those for the inner problem.

The method suggested for solving the outer Dirichlet problem was verified by calculating the potential (u) distribution around a sphere charged to the unity potential. The lower limit of the calculation accuracy is basically specified by rounding accuracy. In practice, however, the calculation accuracy depends on the number of elementary partitions of the domain of simulation. Specifically, the accuracy of estimating the function $u = 1/R$, where R is the radial coordinate, turned out to be within 10^{-6}–10^{-2}% when the boundary of the upper part of the meridional section of a sphere was split into 50–1000 rectilinear boundary elements.

Based on this method for calculating the potential distribution, as well as on the gradientless method for calculating charged particle paths in an electrostatic field [3] and the correlation method for searching for high-order angular focusing conditions [4], we devised a Fokus application program to simulate axisymmetric particle optics with an arbitrary electrode geometry. The program allows for design entry and modification in the graphical mode, calculates the potential distribution over a given domain, and performs the trajectory analysis of a system. Below, the potentialities of the program are demonstrated with the simulation of ion-optical systems designed for mass spectrometry with a quadrupole mass filter (QMF).

A quadrupole filter has enjoyed the widest application in mass spectrometry [5]. Its most important parameters are transmission and resolution. To simultaneously provide a high resolution and a high transmission is a challenge. In practice, this problem is overcome by using entrance ion-optical systems, which match the ion beam to the phase characteristics (acceptance ellipses) of a QMF.

If ions enter into an analyzer with a coordinate $x_0 = y_0 = 0$, the transverse ion velocity for a 100% transmission is limited by the condition [6]

$$\dot{x}_0, \dot{y}_0 < 0.16 r_0 w \left(\frac{\Delta M}{M}\right)^{1/2},$$

where r_0 is the "field radius," w is the rf voltage frequency, and $\Delta M/M$ is the resolution.

From this inequality, it follows that the ratio of the resolutions of two analyzers with equal resolutions can be approximated as $(\sin\alpha_1/\sin\alpha_2)^2$, where α_1 and α_2 are the entrance apertures for the ion trajectories for the first and second analyzers. Therefore, the use of systems that parallelize divergent beams would greatly improve QMF performance.

As a basic electron-optical system, we took a single lens, for which the simulation in the paraxial approximation was carried in [7]. With the Fokus program, the lens was designed for operation in a wide range of

1063-7842/03/4806-0793$24.00

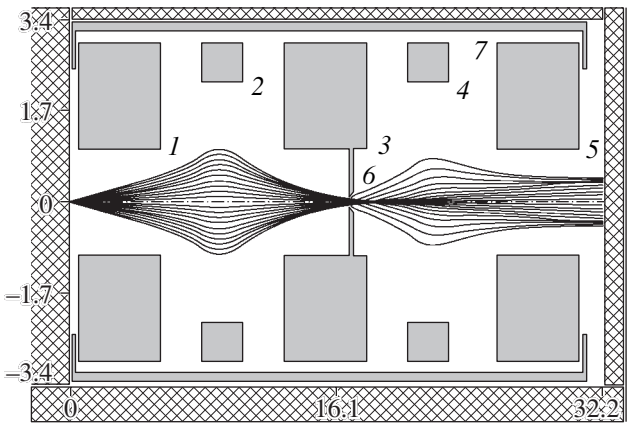

Fig. 1. Ion paths in the lens system with intermediate focusing. (*1–5*) Electrodes, (*6*) diaphragm, and (*7*) electrostatic shield. The potentials (in V) are as follows: $V_1 = 0$, $V_2 = 1$, $V_3 = 0$, $V_4 = 1$, $V_5 = 0$, $V_6 = V_3$, and $V_7 = 0$. The inner diameters of the electrodes (in cm) are: $d_1 = 2$, $d_2 = 4.5$, $d_3 = 2$, $d_4 = 4.58$, and $d_5 = 2$. The ion energy is 1 eV.

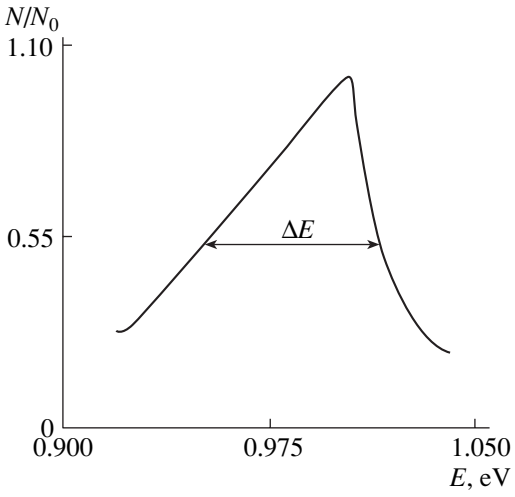

Fig. 2. Transmission curve of the lens system.

entrance apertures. To this end, its geometry was appropriately modified. The analysis showed that the divergence of ion paths may be decreased tenfold with such a lens.

Next, we simulated an ion-optical system that used a combination of the electrodes of the basic lens and provided the intermediate focusing of the ion beam. Ion trajectories in this system are depicted in Fig. 1.

The system makes it possible to narrow the range of entrance apertures from initial ±5° to ±1°. Its other (and major) advantage is the dependence of the relative transmission on the initial ion-beam energy (Fig. 2), which leads us to conclude that the lens system suggested may be used as a device for the preliminary separation of ions by energy. The preseparation of ion beams has been successfully applied to improve the resolution in mass spectrometry applications where the initial energy spread of ions reaches several hundreds or even several thousands of electron volts (e.g., in laser mass spectrometry, spark mass spectrometry, and others). Also, the solution of a number of contemporary physicochemical problems needs energy analysis and mass analysis of ion beams to be performed simultaneously (see, e.g., [8]). The energy resolution $\Delta E/E_0$ of our system (where ΔE is the FWHM of the transmission curve and E_0 is the energy in the transmission maximum) is ≈5%.

REFERENCES

1. C. A. Brebbia, J. C. F. Telles, and L. C. Wrobel, *Boundary Element Techniques* (Springer, Heidelberg, 1984; Mir, Moscow, 1987).
2. A. A. Trubitsyn, Zh. Vychisl. Mat. Mat. Fiz. **35**, 532 (1995).
3. A. A. Trubitsyn, Zh. Vychisl. Mat. Mat. Fiz. **30**, 1113 (1990).
4. A. A. Trubitsyn, Zh. Tekh. Fiz. **71** (5), 126 (2001) [Tech. Phys. **46**, 630 (2001)].
5. A. A. Sysoev and M. S. Chupakhin, *Introduction to Mass Spectrometry* (Atomizdat, Moscow, 1977).
6. P. H. Dawson, Int. J. Mass Spectrom. Ion Phys. **17**, 423 (1975).
7. W. Glaser, *Grundlagen der Elektronenoptik* (Springer, Vienna, 1952; Gostekhizdat, Moscow, 1957).
8. A. Tolstogouzov, S. Daolio, C. Pagura, *et al.*, Surf. Sci. **466**, 127 (2000).

Translated by V. Isaakyan

Technical Physics, Vol. 48, No. 6, 2003, pp. 795–796. Translated from Zhurnal Tekhnicheskoĭ Fiziki, Vol. 73, No. 6, 2003, pp. 138–139.

BRIEF COMMUNICATIONS

On the Feasibility of a Plasma-Chemical Reactor for Hydrogen Production

E. T. Protasevich

Tomsk Polytechnical University, pr. Lenina 30, Tomsk, 634034 Russia

Received August 6, 2002; in final form, November 22, 2002

Abstract—It is proposed that a high-frequency discharge in humid air be used as a model of a plasma-chemical reactor for hydrogen production. © *2003 MAIK "Nauka/Interperiodica".*

In a number of today's applications of a high-frequency discharge, the energy being released upon the combustion of hydrogen produced as a result of the dissociation of water molecules may be comparable to, or even greater than, the HF energy applied to the discharge. Consider this situation in more detail, since it may be of great practical importance.

It is known [1] that the energy barrier for the vibrational excitation of molecules usually coincides with the activation threshold of a reaction and is noticeably lower than the threshold of an electronically excited reaction.

Calculations [1] showed that the effective decomposition of H_2O vapor in a plasma (hence, the production of hydrogen with an efficiency of 50–70%) may proceed both via the dissociative attachment of electrons to water molecules and via the vibrationally excited state of the reagents. In the first case (the dissociative attachment of electrons), it is necessary that the electron temperature be $T_e \geq 1.5$ eV. Our experiments with an HF discharge in humid air [2] showed that the dissociation of water molecules through dissociative attachment due to the cooling of electrons (down to a temperature of about 0.2 eV or lower) is unrealistic.

As to the other way of water dissociation, it can be easily implemented with an HF discharge circuit (Fig. 1). At a pressure of approximately 2 torr (267 Pa) in an unionized air + H_2O mixture and a concentration of water molecules $[H_2O]$ of 4.7×10^{22} m^{-3}, the translational temperature T of hydrogen atoms and the electron concentration n_e that were measured with a Fabry–Perot interferometer were found to be ≈350 K and ~10^{20} m^{-3}. The electron temperature T_e (more precisely, the excitation temperature T_1) calculated from the relative intensities of hydrogen lines in the Balmer series is $T_e \approx 2500$ K. A typical line H_β from which T was determined is shown in Fig. 2.

It was noted in [1] that the vibrational excitation is provided if the electron concentration n_e obeys the inequality

$$n_e/[H_2O] > 3 \times 10^{-4}. \quad (1)$$

Under the above experimental conditions, inequality (1) is reliably satisfied because of the high degree of ionization of the mixture (air + H_2O) and the dissociation of H_2O molecules via vibrational excitation. The most intense *VT* relaxation involves the mode v_2 [1], which is consistent with our results of numerical simulation [3].

It should to be noted that, in the above discharge conditions, for one air molecule there is roughly one water molecule. According to [4], under such conditions for water dissociation, the energy contribution to the process is the highest (6.5 eV per molecule instead of 6.0 eV; i.e., the excess is ≈15%).

In our experiment, along with the lowering of the temperature of electrons and hydrogen atoms, we observed an increase (by more than one order of magnitude) in the effective frequency of collisions of electrons with heavy particles. This also favors the dissociation of water molecules via the vibrationally excited state of the components.

The yield of hydrogen is judged from the discharge emission spectrum. Under optimum water dissociation conditions, the spectrum contained only the hydrogen lines from the Balmer series. The technique of deter-

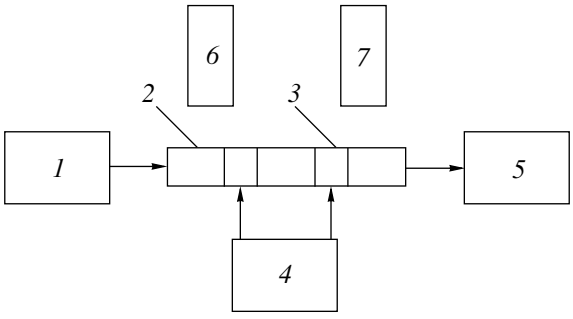

Fig. 1. Block diagram of a plasma-chemical flow-type reactor [2]: (*1*) vapor source, (*2*) discharge tube ($d \approx 2 \times 10^{-2}$ m, $l = 0.3$ m), (*3*) external HF electrodes, (*4*) HF generator ($f =$ 27 MHz, $P = 150$ W), (*5*) vacuum pump, (*6*) spectrograph, and (*7*) Fabry–Perot interferometer.

1063-7842/03/4806-0795$24.00

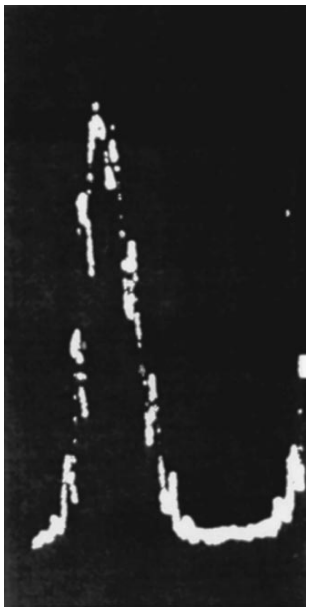

Fig. 2. Typical line H_β recorded by the Fabry–Perot interferometer under the experimental conditions described in the text. $p = 267$ Pa, $N_{H_2O} = 4.7 \times 10^{22}$ m^{-3}.

mining the hydrogen concentration from these spectra is described in detail in [5, 6].

It can thus be argued that an HF capacitive discharge with circular external electrodes is promising for use in an efficient plasma-chemical reactor.

ACKNOWLEDGMENTS

The author is grateful to A. Brabletz and V. Kapicka from Masarik University (Czech Republic) for assistance in the experiments.

REFERENCES

1. V. D. Rusanov, A. A. Fridman, and G. V. Sholin, Usp. Fiz. Nauk **134**, 185 (1981) [Sov. Phys. Usp. **24**, 447 (1981)].
2. A. Brabletz, V. Kapicka, and E. T. Protasevich, Khim. Vys. Énerg. **17**, 258 (1983).
3. V. P. Grigor'ev, E. T. Protasevich, and A. G. Potashev, Sib. Fiz.-Tekh. Zh., No. 3, 161 (1991).
4. E. T. Protasevich, *Cold Non-Equilibrium Plasma: Generation, Properties, and Applications* (CISP, Cambridge, 1999).
5. H. R. Griem, *Spectral Line Broadening by Plasmas* (Academic, New York, 1974; Mir, Moscow, 1979).
6. E. T. Protasevich and V. P. Grygoryev, *Application of Electromagnetic Radiation Plasma for Solving Ecological Problems* (CISP, Cambridge, 2002).

Translated by B. Kalinin

Technical Physics, Vol. 48, No. 6, 2003, pp. 797–799. Translated from Zhurnal Tekhnicheskoĭ Fiziki, Vol. 73, No. 6, 2003, pp. 140–142.

BRIEF COMMUNICATIONS

Effect of Residual Stresses on the Inelastic Properties of Metallic Composites

V. M. Arzhavitin

Kharkov Institute of Physics and Technology, Ukrainian Scientific Center, Akademicheskaya ul. 1, Kharkov, 61108 Ukraine

e-mail: finkel@kipt.kharkov.ua (TO: AR)

Received November 22, 2002

Abstract—The stressed state of in metallic composite systems is analyzed. It is analytically shown that the temperature dependences of the internal friction in the composites may have extrema that are caused by internal stresses. The analytical results are supported by the experimental temperature dependences of the internal friction in (Cu–13% Sn)–30% Nb and Cu–1.3% Cr composites. © *2003 MAIK "Nauka/Interperiodica".*

After mechanical and thermal relieving in composite materials, the composites experience significant internal stresses, which affect their mechanical properties. In this work, we study the effect of such stresses on the inelastic properties (internal friction) of metallic composites. Internal friction in heat-treated steels with a high level of residual microstresses was studied in [1, 2]. The internal friction was found to exhibit a maximum near 200°C that is caused by microstresses in the martensitic phase.

We studied (Cu–13% Sn)–30% Nb and Cu–1.3% Cr composite materials. The (Cu–13% Sn)–30% Nb composites were made in the form of wires 1.5 mm in diameter, and the Cu–1.3% Cr composites were prepared by the reduction of as-cast ingots down to a diameter of 3 mm. The temperature-dependent internal friction $Q^{-1}(T)$ of the (Cu–13% Sn)–30% Nb wires was taken under conditions of free torsional vibrations at a frequency of ~ 1 Hz. The internal friction of the Cu–1.3% Cr samples was measured by the method of resonance bending vibrations at a frequency of ≈50 Hz. In both cases, the measurements were performed at a pressure of 10^{-1} Pa with a strain amplitude $\varepsilon_0 \sim 10^{-4}$. The preparation of the samples and their microstructures are described in detail elsewhere [3].

At $\varepsilon_0 \le 10^{-4}$, the temperature dependence of the internal friction usually represents an exponential background $Q_b^{-1} \sim \exp(-U_b/kT)$ (where U_b is the activation energy of the background), on which peaks corresponding to various physical processes are superposed. However, Fig. 1 shows that the temperature dependences of the internal friction in the (Cu–13% Sn)–30% Nb composite first have a minimum at 50–100°C and only then change to the typical exponential background. To our knowledge, temperature dependences $Q^{-1}(T)$ of this form have not yet been studied, at least in composite materials.

The microplastic internal friction in polycrystalline materials is known to increase as a power function of the applied stress [4]. For composites with a plastic matrix and elastic fibers, the author derived the expression [5]

$$Q^{-1} \sim \frac{\sigma_{av}^{n-2}}{\sigma_{0.2}^{n-1}} \exp(-U/kT), \qquad (1)$$

where $\sigma_{0.2}$ is the static yield strength of the plastic matrix, σ_{av} is the averaged stress in the matrix, n is a constant depending on the variance of the grain stress

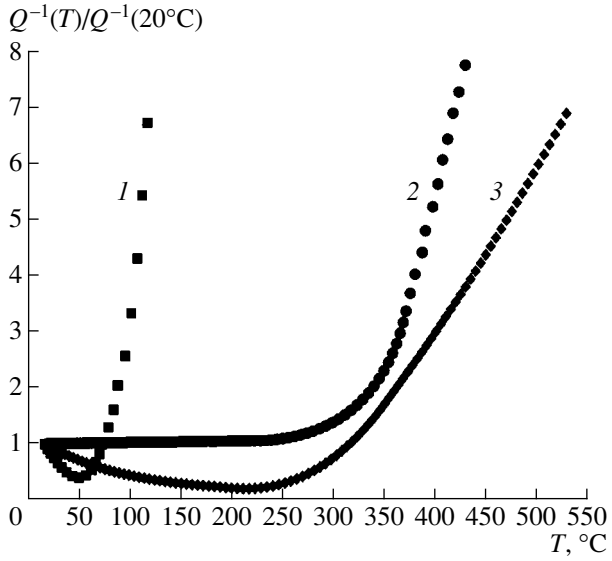

Fig. 1. Variation of the internal friction in the (Cu–13% Sn)–30% Nb wire during sequential heatings (*1–3* are the test numbers).

1063-7842/03/4806-0797$24.00 © 2003 MAIK "Nauka/Interperiodica"

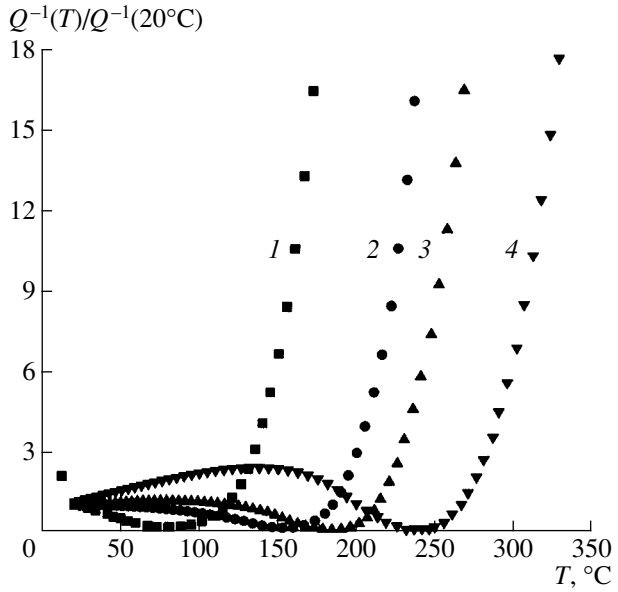

Fig. 2. Calculated dependences of the internal friction in the Cu–1.3% Cr composite for processing-related residual stresses σ^* = (*1*) 0, (*2*) 0.2, (*3*) 0.3, and (*4*) 0.5 kg/mm^2.

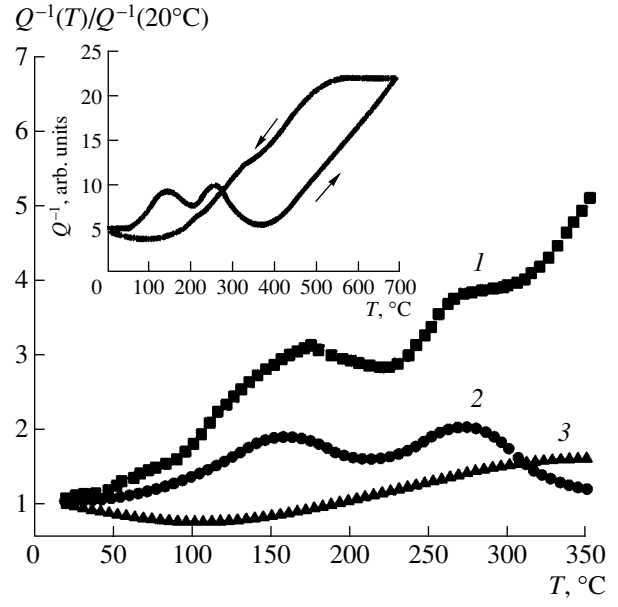

Fig. 3. Variation of the internal friction in the Cu–1.3% Cr sequentially heated to (*1*) 525, (*2*) 600, and (*3*) 500°C. The inset shows the internal friction curve for Cu–1.3% Cr subjected to the heating (600°C)–cooling cycle.

(σ) distribution, and U is the energy required for dislocations to overcome barriers to the plastic flow of the matrix.

Equation (1) was derived under two simplifying assumptions: (1) all matrix grains have the same yield strength $\sigma_{0.2}$ in the direction of an applied force, but their stressed states are different and (2) the grains do not harden during deformation. The latter assumption requires low activation energies ($U \approx 0.1$ eV) of the plastic flow so that barriers to dislocation glide are overcome only by thermal excitations. An example is the thermally activated depinning of dislocations from pinning point defects [6]. The thermal energy is, however, insufficient for dislocations to overcome high-energy barriers ($U \approx 1$ eV) during the flow; here, the applied stress is of significance because of the possible strain hardening of the material.

Average stresses in heated (cooled) composite materials can be calculated by the Garmong equation [7]. When deriving the computational formula, Garmong used the additivity law (for "mixtures") and the continuity condition at the matrix–fiber interface. In the absence of an applied load, the Garmong equation has the form

$$\int_{T_0}^{T} (\alpha_m - \alpha_f)\,dT = (\varepsilon_m)_p - \sigma\left(\frac{V_f E_f + V_m E_m}{V_f E_f E_m}\right),$$

where α, E, V, and σ are the thermal expansion coefficient, the modulus of elasticity, the volume fraction of the matrix (fibers), and the average thermal stress in the matrix; T and T_0 are the actual and initial temperatures of the thermal cycle, respectively; and $(\varepsilon_m)_p$ is the plastic strain in the matrix.

The last-named parameter is the sum of the strains due to the current and diffusion creep. The diffusion component can be neglected at relatively low temperatures (≈200°C). Then, the actual plastic strain of the matrix without allowance for strain hardening is $(\varepsilon_m)_p \approx \sigma_{0.2}/E_m$.

Along with stresses σ induced by temperature variations, the matrix may experience processing-related residual stresses σ^*, such as those introduced by wire-drawing from as-cast ingots. In this case, the total stress in the matrix is $\sigma_m = \sigma + \sigma^*$ and the Garmong equation can be recast as

$$\sigma_m = \left(\frac{\sigma_{0.2}}{E_m} - \int_{T_0}^{T} (\alpha_m - \alpha_f)\,dT\right)\frac{V_f E_f E_m}{V_f E_f + V_m E_m} + \sigma^*. \quad (2)$$

To calculate the temperature dependences of the internal friction in the composites, we will use Eqs. (1) and (2). To a first approximation, let us assume that the average, σ_{av}, and internal, σ_m, stresses in the matrix are equal in order of magnitude: $\sigma_{av} \sim \sigma_m$. This means that we consider the part of the internal friction that is associated solely with internal stresses in the material rather than with stresses induced by an applied load or with a combination of both. Figure 2 shows the internal friction curves (*1–3*) for the Cu–1.3% Cr model composite that were calculated at $n = 4$; $U = 0.15$ eV; and $\sigma^* = 0$, 0.2, 0.3, and 0.5 kg/mm^2. The data on α, E, and $\sigma_{0.2}$ for the components of the composite were taken from reference books [8, 9]. The minimum of the internal friction is the most pronounced at $T \sim 100$°C in the absence

($\sigma^* = 0$) of processing-related residual stresses. When such stresses arise ($\sigma^* \neq 0$), maxima near 150°C appear instead of the minima in the curves $Q^{-1}(T)$.

For comparison, Fig. 3 gives the experimental temperature dependences of the internal friction in the Cu–1.3% Cr composite that were obtained when the same sample was sequentially heated to 525, 600, and 550°C at a rate of 2°C/min. The calculated and experimental dependences qualitatively coincide. The disappearance of the maxima and the appearance of a minimum in the third thermal cycle may be explained by the relaxation of the residual stresses σ^* via the microplastic deformation of the matrix in the two previous cycles. The inset in Fig. 3 shows the complete curve $Q^{-1}(T)$ for the copper–chromium composite in the second thermal cycle. It is seen that the minimum appears during the cooling of the composite. The temperature hysteresis of the internal friction (Fig. 3, inset) indicates the microplastic deformation of the copper matrix [3, 5] at cooling temperatures preceding the minimum of $Q^{-1}(T)$.

To conclude, we note that the temperature positions of the internal friction extrema coincide with the temperatures of so-called ductility dips in copper and copper-based alloys [10, 11]. This fact seems to be not accidental. It was noted [10] that plasticity dips are observed in thermodynamically metastable materials. Stressed materials are in thermodynamically nonequilibrium states, irrespective of the origin of the stresses.

REFERENCES

1. V. I. Sarrak and S. O. Suvorova, *Internal Friction in Metals, Semiconductors, Insulators, and Ferromagnets* (Nauka, Moscow, 1978), pp. 75–80.
2. I. S. Mikhaĭlov, N. V. Gavrilov, and S. B. Mikhaĭlov, Fiz. Met. Metalloved. **93** (4), 107 (2002).
3. V. M. Arzhavitin and V. Ya. Sverdlov, Zh. Tekh. Fiz. **68** (11), 114 (1998) [Tech. Phys. **43**, 1376 (1998)].
4. V. T. Troshchenko, Fiz. Tverd. Tela (Leningrad) **2** (6), 1060 (1960) [Sov. Phys. Solid State **2**, 958 (1960)].
5. V. M. Arzhavitin, Zh. Tekh. Fiz. **72** (10), 128 (2002) [Tech. Phys. **47**, 1330 (2002)].
6. H. G. van Bueren, *Imperfections in Crystals* (North-Holland, Amsterdam, 1961; Inostrannaya Literatura, Moscow, 1962).
7. G. Garmong, Metall. Trans. **8**, 2183 (1974).
8. A. V. Bobylev, *Mechanical Properties and Workability of Metals: A Handbook* (Metallurgiya, Moscow, 1980).
9. L. V. Tikhonov, V. A. Kononenko, G. I. Prokopenko, and V. A. Rafalovskiĭ, *Structure and Properties of Metals and Alloys (Mechanical Properties of Metals and Alloys)* (Naukova Dumka, Kiev, 1986).
10. A. A. Presnyakov and V. V. Chervyakova, *The Nature of Ductility Dips in Metal Alloys* (Nauka, Alma-Ata, 1970).
11. A. K. Mindyuk, Metally, No. 3, 144 (1990).

Translated by K. Shakhlevich